Departament de Física de la Terra y Termodinàmica

Calibración de nuevos sensores térmicos a bordo de satélite y validación de productos de temperatura superficial terrestre

Doctorado en Teledetección

Jesús Jose Puchades Yago

Mayo, 2025

Directores: Raquel Niclòs Corts

César Coll Company

VNIVERSITAT DE VALÈNCIA **Departament de Física de la Terra i Termodinàmica**

Dña. Raquel Niclòs Corts, Profesora titular de Física de la Tierra del Departamento de Física de la Tierra y Termodinámica de la Facultad de Física, Universitat de València, y

D. César Coll Company, Catedrático de Física de la Tierra del Departamento de Física de la Tierra y Termodinámica de la Facultad de Física, Universitat de València.

CERTIFICA/N:

Que la presente memoria, titulada "*Calibración de nuevos sensores térmicos a bordo de satélite y validación de productos de temperatura superficial terrestre*", corresponde al trabajo realizado bajo su dirección por D. Jesús Jose Puchades Yago, para optar al grado de doctor por la Universitat de València dentro del programa de Doctorado en Teledetección.

Y para que conste firman el presente certificado en Valencia, a 22 de Mayo de 2025.

Firmado por RAQUEL NICLOS CORTS - NIF:***4468** el día 22/05/2025 con un certificado emitido por ACCVCA-120

CESAR COLL COMPAÑY - NIF:52670828T

Firmado digitalmente por CESAR COLL COMPAÑY - NIF:52670828T
Fecha: 2025.05.22 12:08:48 +02'00'

Fdo.: Raquel Niclòs Corts

Fdo.: César Coll Company

Este trabajo ha sido posible gracias a la financiación recibida por parte de los siguientes proyectos:

- Nuevas técnicas de teledetección y modelización para la mejora de la predicción de fenómenos extremos asociados al cambio climático (Tool4Extreme) (PID2020-118797RB-I00). Ministerio de Ciencia e Investigación, Agencia Estatal de Investigación (MCIN/AEI/10.13039/501100011033).
- Análisis y modelización de tendencias en la Comunidad Valenciana en el contexto del cambio climático (Teleclima) (PROMETEO/2021/016), Generalitat Valenciana, Conselleria d'Educació, Universitats i Ocupació en el marco del Programa Prometeo (I+D+i) para grupos de investigación de excelencia.
- T-based and r-based validation of the Sentinel-3 SLSTR LST product and an alternative emissivity-dependent Split-Window algorithm dentro del Sentinel-3 Validation team (S3VT) (PP0097749).
- Impacto del cambio climático sobre ecosistemas: patrones a fina escala, índices de vigilancia de temperaturas extremas mediante nuevas técnicas de teledetección (CLIMTEX) (CGL2015-64268-R), Ministerio de Economía y Competitividad y los fondos de desarrollo regional (MINECO/FEDER, UE).
- Estimación y seguimiento de la temperatura y la emisividad de la superficie terrestre mediante los últimos avances en teledetección (CGL2013-46862-C2-1-P), Ministerio de Economía y Competitividad (MINECO).
- Impacto térmico en fenómenos meteorológicos de ecosistemas mediterráneos: desarrollo de técnicas de teledetección y una red automática de radiometría angular de campo (CGL2011-30433-C02-02), Ministerio de Ciencia e Innovación (MICINN).

- Y gracias a la contratación como Técnico Superior de Investigación dentro del proyecto (Tool4Extreme) (PID2020-118797RB-I00).
- Contratación de Técnico Superior de Investigación con la ayuda: Optimización de la precisión de datos térmicos adquiridos con drones (RPAs). Subvenciones para la contratación de Personal de Apoyo vinculado a un Proyecto de Transferencia Tecnológica (APOTI/2018/014), financiado por la Generalitat Valenciana.
- Contratación de Técnico Superior de Investigación con la ayuda: Apoyo técnico para el laboratorio de calibración y validación de sensores térmicos. Ayudas para la promoción de empleo joven e implantación de la garantía juvenil I+D+I (PEJ-2014-A-18514). Financiado por el Ministerio de Economía y Competitividad (MINECO).

AGRAÏMENTS

En primer lloc, m'agradaria donar les gràcies als meus directors de Tesi Doctoral, a la Dra. Raquel Niclòs i al Dr. César Coll sense ells este treball no seria possible. A més, d'estar sempre disponibles per a resoldre qualsevol tipus de dubte i fer-me sentir un més, malgrat no ser físic. Per molt complexa que siga la ciència amb vosaltres és més fàcil d'entendre. Raquel, gràcies per comptar amb mi des del primer moment i ajudar-me en tot este procés. César saps que en teledetecció menges en la taula dels més grans, però de futbol ni idea oblida el RM i passa't al costat bo (VCF).

Al Dr. Enric Valor (ara Degà de la facultat) per resoldre els dubtes, compartir els seus coneixements davant les meues inquietuds que han fet millorar la meua formació científica. I al Dr. Vicent García Santos per esta última temporada en el grup i en la memòria queda la campanya de mesures en el NPL de Teddington.

Als companys antics del grup, Xusa, Carol, Xumi, Lluís gràcies per acollir-me en el grup i pels riures en les hores del dinar. Xumi moltes gràcies també pel bon acolliment en la UCLM. Lluís (Luisito) són molts els dies de mesures en els arrossars, però han valgut la pena, Amunt sempre. I a Juanma pel tracte i el bon període en la UCLM.

I als companys més recents, Jordi, Joan, Óscar, Martí, Gala, Jandro i Sara hem gaudit de les campanyes de mesures en els arrossars, però sobretot de lo surrealistes que eren les converses que teníem en el laboratori, encara que el més important ha sigut el bon ambient que hem creat en el treball.

Menció a part són els cafés, en els quals tractem els temes de conspiració, cuarto mileno i el futbol. A més de totes les boludeses que fan que siga un moment únic, si es gravaren tindríem clips virals no tinc cap mena de dubte.

Al personal de secretaria, Mónica i Quique per ajudar-me amb tot el tema de papers que tant odie. Les xarrades rient-nos del catedràtic madridista són per a recordar.

A la família, als meus pares i al meu germà que m'han donat suport i ajudat no sols en esta etapa sinó, al llarg de la vida. També a la xicoteta de la casa que m'ompli d'energia amb les seues visites.

I a eixe grup d'amics que fa que desconnecte de tot i només riguem i gaudim del temps que passem junts.

ÍNDICE

Resumen……...1

Summary……..7

Resum……..…13

Capítulo 1 Introducción……19

a. Temperatura de la superficie terrestre ……21
b. Métodos para la determinación de la LST……...21
c. Validación LST y calibración de bandas IRT……..23
c.1. Importancia de la validación y la calibración ……23
c.2. Validación t-based ……23
c.3. Validación r-based ……25
d. Sensores……...26
d.1. Sensor SLSTR……26
d.2. Sensor TIRS y TIRS-2……...27
d.3. Sensor MODIS ……..28
d.4. Sensor AVHRR/3 ……..29
d.5. Otros sensores y misiones futuras ……30
e. Objetivo de la tesis ……..30
f. Estructura de la memoria ……31

Capítulo 2. Metodología……33

a. Ecuación de transferencia radiativa (ETR) y simulación de datos de satélite….35
b. Método monocanal (SC) ……..35
c. Medidas in situ de LST, para el método t-based……..36
d. Método r-based……36
e. Algoritmos LST……38

e.1. Algoritmos para SLSTR....38
e.1.1 Producto operativo LST SLSTR....38
e.1.2. Algoritmo SW alternativo para SLSTR....39
e.1.3. Otros algoritmos SW propuestos....40
e.2. Algoritmos para TIRS y TIRS-2....41
e.2.1. Producto operativo LST TIRS y TIRS-2....41
e.2.2. Algoritmos propuestos para TIRS y TIRS-2....41
e.3. Algoritmos para MODIS....43
e.3.1. Producto M*D11....43
e.3.2. Producto M*D21....44
e.4. Algoritmos para AVHRR/3....44
e.4.1. Producto operativo LST de LSA-SAF ELST....44
e.4.2. Algoritmo SW propuesto....44
e.4.3. Otros algoritmos SW....45
f. Métodos para obtener la emisividad de la superficie....45
f.1. Método de la caja....46
f.2. Método TES para medidas de campo....47
f.3. Método de cobertura vegetal....49
f.4. Ecuación paramétrica de la emisividad del agua....49
g. Análisis de sensibilidad....50

Capítulo 3. Áreas de validación, instrumental, datos de satélite y perfiles atmosféricos....53

a. Áreas de validación para el método t-based....55
a.1. Arrozales de Valencia....55
a.2. Matorral de Cortes de Pallás....56
a.3. Desierto de Gobabeb....57
a.4. Lago Constanza....57
a.5. Estaciones de la red LAW....57
a.5.1. Svartberget (SVA)....57
a.5.2. Hyytiälä (HYY)....58
a.5.3. Robson Creek (ROB)....58
a.5.4. KIT forest (KIT)....58
a.5.5. Puéchabon (PUE)....58
b. Áreas de validación para el método r-based....58
b.1. Bosque de Soria....58
b.2. Lago de la Albufera....58
b.3. Viñedo....58
b.4. Salinas....58
b.5. Olivar....58
c. Instrumental....59
c.1. Fuente de calibrado Landcal P80P....59
c.2. Radiómetro térmico multibanda CIMEL CE312-2....60
c.3. Apogee SI-121....62
c.4. Heitronics KT15.85 IIP....63
d. Imágenes Satelitales....64
e. Perfiles atmosféricos....64
f. Datos de emisividad....65

Capítulo 4. Resultados....67

a. Resultados para el sensor SLSTR....69
a.1. Calibración de las bandas TIR del sensor SLSTR....71

a.2. Validación con el método t-based……..73
a.3. Validación con el método r-based ……77
b. Resultados para el sensor TIRS……...80
b.1. Calibración de las bandas TIR del sensor TIRS ……80
b.2. Validación con el método t-based ……..88
c. Resultados para el sensor TIRS-2……94
c.1. Calibración de las bandas TIR del sensor TIRS-2……..94
c.2. Validación con el método t-based……..96
d. Resultados para el sensor MODIS……100
d.1. Validación del método r-based……101
d.2. Validación r-based de los productos M*D11 y M*D21……..102
e. Resultados para el sensor AVHRR/3 ……..…….…114
e.1. Validación t-based para MetOp-A ……..…114
e.2. Validación t-based para MetOp-B ……..115
e.3. Validación t-based para MetOp-C……...118

Capítulo 5. Conclusiones……121

Bibliografía……..127

Anexos ……..i

ANEXO A. Laboratory calibration and field measurement of land surface temperature and emissivity using thermal infrared multiband radiometers.……iii
ANEXO B. Validation of Sentinel-3 SLSTR Land Surface Temperature Retrieved by the Operational Product and Comparison with Explicitly Emissivity-Dependent Algorithms.……xix
ANEXO C. Evaluation of Landsat-8 TIRS data recalibrations and land surface temperature split-window algorithms over a homogeneous crop area with different phenological land covers.……xlv
ANEXO D. Evaluating Landsat-9 TIRS-2 calibrations and land surface temperature retrievals against ground measurements using multi-instruments spatial and temporal sampling along transects……..lxv
ANEXO E. Demonstrating the suitability of the Radiance-based Method for Assessing Accuracy of MODIS Land Surface Temperature Products.……...lxxxi
ANEXO F. Ground Calibration of Sentinel-3 SLSTR Thermal Infrared Bands and Validation of LST Algorithms.……xcix

RESUMEN

Para conocer la precisión y exactitud de la temperatura de la superficie terrestre (LST) obtenida a partir de datos de sensores satelitales es necesario validar los algoritmos con los que se deriva la LST frente a medidas independientes de dicha magnitud. Por otro lado, es también importante comprobar el calibrado de las bandas térmicas de los sensores con referencia a medidas in situ de LST (calibración vicaria). Los datos necesarios para la calibración vicaria y la validación de las LST de satélite se obtienen generalmente a partir de medidas radiométricas in situ de LST en zonas experimentales (método basado en temperatura o t-based). Para la medida in situ de la LST es necesario el uso de radiómetros de campo que normalmente operan en el intervalo espectral de 8 a 14 μm, bien con canales estrechos como el radiómetro multibanda CIMEL312-2, o de un canal ancho como los radiómetros Apogee SI-121 y Heitronics KT15.85IIP, todos ellos utilizados en este trabajo. Estos instrumentos pueden estar instalados de manera fija en una estación de medidas o tomar medidas a lo largo de transectos. Para que las medidas in situ de las LST sean comparables a datos de satélite, las zonas de validación deben ser térmicamente homogéneas y de una sola componente (agua, vegetación completa, suelo desnudo) lo que limita la aplicación del método t-based de validación a unas pocas cubiertas. Para la validación de LST de satélite también puede utilizarse el método basado en radiancia (r-based; Wan y Li, 2008) en el que la LST de referencia se calcula a partir de las radiancias en el techo de la atmósfera, valores de emisividad en las bandas térmicas, perfiles atmosféricos y un modelo de transferencia radiativa (MODTRAN 5.2, Berk et al., 2008). El método r-based puede utilizarse en zonas heterogéneas donde las medidas in situ no son factibles o no están disponibles, siempre que se conozca la emisividad de la superficie y los perfiles atmosféricos.

En cualquier caso, el conocimiento de la emisividad de las superficies de validación en las bandas espectrales utilizadas. La emisividad es necesaria, para: (i) obtener la LST in situ a partir de medidas radiométricas, (ii) la simulación de temperaturas de brillo en el techo de la atmósfera a partir de las medidas de LST in situ, (iii) la aplicación de los distintos algoritmos de LST y (iv) la aplicación del método r-based. Para ello hemos realizado medidas experimentales de emisividad utilizando el método de la caja (Rubio et al., 1997), el método TES (Gillespie et al., 1998), y aplicado el modelo de Valor y Caselles (1996) en superficies terrestres heterogéneas para calcular la emisividad a la escala del píxel satelital y el modelo de Niclòs et al. (2009) para la emisividad de la superficie del agua.

Los algoritmos monocanales (SC) (Chedin et al., 1982) calculan la LST a partir de la temperatura de brillo en una banda espectral y requieren, además de la emisividad, los perfiles atmosféricos en el momento de paso del sensor satelital y un modelo de transferencia radiativa con los que obtener la transmisividad y las radiancias atmosféricas ascendente y descendente en el canal utilizado. Otro tipo de algoritmos son los métodos de absorción diferencial, o de ventana dividida (SW) (McMillin, 1975) que se basan en utilizar dos temperaturas de brillo dentro de la misma ventana atmosférica de tal manera que no es necesario el conocimiento de los perfiles atmosféricos. También existe el método multicanal de separación de temperatura y emisividad (TES; Gillespie et al., 1998), que requiere de al menos tres bandas en el infrarrojo térmico y la corrección monocanal de cada banda de forma independiente. Con las radiancias a nivel de superficie se consigue desacoplar la temperatura de la emisividad y se puede obtener de manera simultánea la LST y la emisividad para cada banda.

Al realizar la calibración del sensor SLSTR frente a medidas de 9 estaciones situadas sobre distintas coberturas (zonas boscosas, desérticas, cultivos de regadío, lagos) los resultados obtenidos muestran una R-RMSD de 0.8 K y 0.7 K para las bandas 8 y 9, respectivamente en ambos satélites (Sentinel-3A y -3B), con medianas cercanas a cero en todos los casos. La validación t-based del producto operativo de LST de SLSTR presentó diferencias medias superiores a 1 K en todas las cubiertas, excepto para las zonas boscosas que estaban dentro de ±1 K. Sin embargo, para la zona de los arrozales de Valencia mostró una diferencia media de 1.6 K y 2.0 K para Sentinel-3A y -3B respectivamente, lo que puede ser debido a la asignación de un coeficiente fijo para el bioma cuando realmente se trata de una cubierta cambiante a lo largo del año. Por otro lado, el SW propuesto por Pérez-Planells et al. (2021) solo muestra diferencias medias excesivas (-1.7 K) en las zonas desérticas con datos diurnos, estando en el intervalo de ±1 K para el resto de cubiertas. Los resultados con el método r-based para el producto de LST de SLSTR presentaron diferencias medias superiores al umbral GCOS en las zonas de cultivos de regadío y desérticas, y el SW propuesto por Pérez-Planells et al. (2021) solo en las zonas desérticas con datos diurnos, análogamente a la validación t-based.

El sensor TIRS a bordo de Landsat-8 sufrió diferentes reprocesados que afectaron a la calibración de las bandas térmicas (bandas 10 y 11). Las temperaturas de brillo con la calibración original mostraron una sobreestimación de 0.7 K en la banda 10 y 2.7 K en la banda 11 respecto a las temperaturas de brillo simuladas a partir de datos in situ medidos en la zona de Arrozales de Valencia. Tras el reprocesado de 2014 obtuvimos subestimaciones de -1.4 K y -1.6 K, respectivamente. Sin embargo, tras la aplicación de la corrección TIRS-on-TIRS (Montanaro et al., 2015) las desviaciones medias fueron muy próximas a cero en ambas bandas. En 2019 se realizó un nuevo reprocesado con el que las diferencias medias pasaron a ser de -0.1 K y 0.5 K para las bandas 10 y 11, respectivamente, introduciendo una desviación sistemática significativa entre las dos bandas. Los resultados de validación t-based de la LST obtenida con el método SC muestran que no existe diferencia entre el procesado de 2017 y el de 2019 en términos de R-RMSD, siendo este 0.7 K en ambos casos. El algoritmo SW que mejor funciona con datos TIRS en el procesado de 2017 es el propuesto por Jiménez-Muñoz et al. (2014) con una R-RMSD de 0.6 K, mientras que en el procesado de 2019 las diferencias medias de todos los SW estudiados (un total de 17) aumentaron entre 1.3 K y 1.9 K dependiendo del algoritmo.

La calibración de las bandas térmicas del sensor TIRS-2 a bordo de Landsat-9 fue evaluada frente a datos in situ en la zona de arrozales de Valencia calibradas para los datos originales y tras el reprocesado de 2023. Los resultados con los datos originales ofrecieron unas diferencias medias de -0.3 K y -0.1 K para las bandas 10 y 11, respectivamente. Tras el reprocesado, la banda 10 mostró una diferencia media despreciable, mientras que fue de 0.2 K en la banda 11, lo que hace que la desviación entre las bandas aumente y afecte al funcionamiento de los SW. La validación del producto operativo de LST y de los SW estudiados con el método t-based presentan diferencias medias dentro del intervalo ±1 K, tanto en los datos originales como en los reprocesados. Aunque se observó una diferencia sistemática de 0.6 K entre los datos originales y los reprocesados debido a los cambios en la calibración de las bandas térmicas.

En el caso del sensor MODIS, inicialmente comparamos las LST de referencia del método r-based frente a un conjunto de medidas in situ de LST en la zona de arrozales de Valencia para estimar la incertidumbre de las primeras en relación con respecto a las segundas. A continuación, hemos realizado la validación r-based en zonas con diversas coberturas de los productos de LST de M*D11 y M*D21 (MOD para el satélite Terra y MYD para Aqua). De acuerdo con estos resultados, el producto M*D11 presentó una subestimación de -1.0 K (RMSD de 1.2 K) para los datos diurnos y de 0.0 K (RMSD de 0.5 K) para los datos nocturnos, teniendo en cuenta todas las zonas de validación conjuntamente. Por otro lado, la desviación media (RMSD) mostrada por el M*D21 fue de 0.1 K (0.7 K) y 0.3 K (0.5 K) para los datos diurnos y nocturnos, respectivamente, para ambos satélites y todo el conjunto de datos. El método de validación r-based proporciona una alternativa valiosa al método t-based para la mayoría de las cubiertas terrestres donde las medidas in-situ de LST no son posibles.

También realizamos la validación t-based al producto operativo de LST de LSA-SAF y algoritmos SW del sensor AVHRR/3 a bordo de MetOp-A, -B y -C. El producto mostró resultados para la versión operativa en R-RMSD de 3.4 K y 2.0 K para los datos diurnos y nocturnos, respectivamente. Se propone un algoritmo SW que presenta medianas cercanas a cero en todas las cubiertas y los tres satélites. El algoritmo propuesto por Jiménez-Muñoz y Sobrino (2008) presenta similares resultados a excepción del caso de la vegetación donde sobreestima las LST en torno a 1 K y para MetOp-C, donde la desviación media para todas las cubiertas es de 1.4 K.

SUMMARY

In order to know the precision and accuracy of land surface temperature (LST) retrieved from satellite sensor data, it is necessary to validate the algorithms with which the LST is derived against independent measurements of this magnitude. On the other hand, it is also important to check the calibration of the thermal bands of satellite sensors with reference to ground LST measurements (vicarious calibration). Data required for the vicarious calibration and validation of satellite LSTs are generally obtained from ground radiometric LST measurements in experimental areas (temperature-based or t-based method). Ground LST measurement require the use of field radiometers that typically operate in 8-14 μm spectral range, either with narrow channels those of the CIMEL312-2 multiband radiometer, or a wide channel those of the Apogee SI-121 and Heitronics KT15.85IIP radiometers, all of which are used in this work. These instruments can be fixedly installed at a measurement station or can be handled to take measurements along transects. For ground LST measurements to be comparable to satellite data, validation zones must be thermally homogeneous and single-component (water, full vegetation, bare soil), which limits the application of the t-based validation method to a few land covers. For satellite LST validation, radiance-based method (r-based; Wan and Li, 2008) can also be used in which reference, LST is calculated from radiances at the top of atmosphere, emissivity values in thermal bands, atmospheric profiles, and a radiative transfer model (MODTRAN 5.2, Berk et al., 2008). R-based method can be used in heterogeneous areas where ground measurements are not feasible or not available, provided that surface emissivity and atmospheric profiles are known.

In any case, it is necessary to know the emissivity for the different validation surfaces and spectral bands used. Emissivity is required in order to: (i) obtain ground LST from radiometric measurements, (ii) simulate brightness temperatures in top of atmosphere from ground LST measurements, (iii) apply different LST algorithms and (iv) apply the r-based method. For this purpose, we have performed experimental emissivity measurements using the box method (Rubio et al., 1997), TES method (Gillespie et al., 1998) and applying the model of Valor and Caselles (1996) on heterogeneous land surfaces to calculate emissivity at the satellite pixel scale, and the model of Niclòs et al. (2009) for water surface emissivity.

Single-channel (SC) algorithms (Chedin et al., 1982) calculate LST from brightness temperature in a spectral band and require, in addition to the emissivity, the atmospheric profiles at the time of satellite overpass and a radiative transfer model, with which to obtain the transmissivity and the upward and downward atmospheric radiances in the channel used. Another type of algorithms are the differential absorption or split-window (SW) methods (McMillin, 1975), which are based on using two brightness temperatures within the same atmospheric window so that knowledge of atmospheric profiles is not necessary. There is also the multichannel method of temperature and emissivity separation (TES; Gillespie et al., 1998), which requires at least three bands in thermal infrared and single-channel correction of each band independently. With surface-level radiances, temperature is decoupled from emissivity and LST, and emissivity can be obtained simultaneously for each band.

The results obtained when calibrating the SLSTR sensor against measurements from 9 stations located over different land covers (forest areas, desert, irrigated croplands, lakes), show an R-RMSD of 0.8 K and 0.7 K for

bands 8 and 9, respectively for both satellites (Sentinel-3A and -3B), with medians close to zero in all cases. T-based validation of the SLSTR operational LST product showed mean differences greater than 1 K for all land covers, except for forest areas, in which mean differences were within ±1 K. However, for Valencia ricepaddy area, mean difference of 1.6 K and 2.0 K were obtained for Sentinel-3A and -3B respectively, which may be due to the assignment of a fixed coefficient for the biome when it is really a changing cover throughout the year. On the other hand, the SW proposed by Pérez-Planells et al. (2021) only shows excessive mean differences (-1.7 K) in desert areas with diurnal data, being in the interval of ±1 K for the rest of the land covers. Results with the r-based method the SLSTR LST product showed mean differences above the GCOS threshold in irrigated croplands and desert areas, and the SW proposed by Pérez-Planells et al. (2021) only in desert areas with diurnal data, analogous to the t-based validation.

Data of the TIRS sensor on board Landsat-8 underwent different reprocessing that affected the calibration of thermal bands (bands 10 and 11). The brightness temperatures with the original calibration showed an overestimation of 0.7 K in band 10 and 2.7 K in band 11 with respect to the simulated brightness temperatures from in situ data measured in the Valencia ricepaddy area. After 2014 reprocessing, we obtained underestimates of -1.4 K and -1.6 K, respectively. However, after application of TIRS-on-TIRS correction (Montanaro et al., 2015) mean deviations were very close to zero in both bands. In 2019 a new reprocessing was performed, with which mean differences became -0.1 K and 0.5 K for bands 10 and 11, respectively, introducing a significant systematic deviation between bands 10 and 11. Results of the t-based validation of the LST obtained with the SC method show that there is no difference between 2017 and 2019 data processing in terms of R-RMSD, this being 0.7 K in both cases. The best performing SW algorithm with TIRS data in the 2017 processing is proposed by Jiménez-Muñoz et al. (2014) with an R-RMSD of 0.6 K, while in 2019 processing mean differences of all studied SWs (a total of 17) increased between 1.3 K and 1.9 K depending on the algorithm.

The calibration of thermal bands of TIRS-2 sensor on board Landsat-9 was evaluated against ground data in the ricepaddy area of Valencia. They were calibrated for the original data and after reprocessing in 2023. Results with original data gave mean differences of -0.3 K and -0.1 K for bands 10 and 11, respectively. After reprocessing, band 10 showed a negligible mean difference, while it was 0.2 K for band 11, which makes the deviation between bands increase and affect performance of SW algorithms. The validation of the LST operational products and SW algorithms studied with t-based method show mean differences within the ±1 K range, both in the original and reprocessed data. However, a systematic difference of 0.6 K was observed between the original and reprocessed data due to changes in thermal band calibration.

In the case of MODIS sensor, we initially compared the reference LSTs of r-based method against a set of ground LST measurements in the rice paddy area of Valencia to estimate the uncertainty of the former with respect to the latter. Then, we have performed r-based validation of LST products of M*D11 and M*D21 (MOD for Terra satellite and MYD for Aqua) in areas with different land covers. According to these results, M*D11 product presented an underestimation of -1.0 K (RMSD of 1.2 K) for daytime data and 0.0 K (RMSD of 0.5 K) for nighttime data, taking into account all validation areas together. On the other hand, the mean deviation

(RMSD) shown by M*D21 was 0.1 K (0.7 K) and 0.3 K (0.5 K) for daytime and nighttime data, respectively, for both satellites and the whole data set. R-based validation method provides a valuable alternative to t-based method for most land covers where ground measurements of LST are not possible.

We also performed t-based validation to the operational LST product of LSA-SAF and SW algorithms of the AVHRR/3 sensor onboard MetOp-A, -B and -C. The results showed for the operational product are R-RMSD of 3.4 K and 2.0 K for daytime and nighttime data, respectively. We proposed an alternative SW algorithm, which showed median values closed to zero for all land covers and the three satellites. The algorithm of Jiménez-Muñoz and Sobrino (2008) presents similar results except for the case of vegetation in which this method overestimates LSTs by around 1 K, and for MetOp-C, where the average deviation for all land covers is 1.4 K.

RESUM

Per a conéixer la precisió i exactitud de la temperatura de la superfície terrestre (LST) obtinguda a partir de dades de sensors satel·litaris és necessari validar els algorismes amb els quals es deriva la LST front a mesures independents d'esta magnitud. D'altra banda, és també important comprovar el calibrat de les bandes tèrmiques dels sensors amb referència a mesures in situ de la LST (calibratge vicari). Les dades necessàries per al calibratge vicari i la validació de les LST de satèl·lit s'obtenen generalment a partir de mesures radiomètriques in situ de LST en zones experimentals (mètode basat en temperatura o t-based). Per a la mesura in situ de la LST és necessari l'ús de radiòmetres de camp que normalment operen en l'interval espectral de 8 a 14 µm, bé amb canals estrets com el radiòmetre multibanda CIMEL312-2, o d'un canal ample com els radiòmetres Apogee SI-121 i Heitronics KT15.85IIP, tots ells utilitzats en este treball. Estos instruments poden estar instal·lats de manera fixa en una estació de mesures o prendre mesures al llarg de transectes. Perquè les mesures in situ de les LST siguen comparables a dades de satèl·lit, les zones de validació han de ser tèrmicament homogènies i d'una sola component (aigua, vegetació completa, sòl nu) el que limita l'aplicació del mètode t-based de validació a unes poques cobertes. Per a la validació de LST de satèl·lit també pot utilitzar-se el mètode basat en radiància (r-based; Wan i Li, 2008) en el qual la LST de referència es calcula a partir de les radiàncies en el sostre de l'atmosfera, valors d'emissivitat en les bandes tèrmiques, perfils atmosfèrics i un model de transferència radiativa (MODTRAN 5.2, Berk et al., 2008). El mètode r-based pot utilitzar-se en zones heterogènies on les mesures in situ no són factibles o no estan disponibles, sempre que es conega l'emissivitat de la superfície i els perfils atmosfèrics.

En qualsevol cas, és necessari el coneixement de l'emissivitat de les diferents superfícies de validació i les bandes espectrals utilitzades. L'emissivitat es necessaria per a: (i) obtindre la LST in situ a partir de mesures radiomètriques, (ii) la simulació de temperatures radiatives al sostre de l'atmosfera a partir de les mesures de LST in situ, (iii) l'aplicació dels diferents algorismes LST i (iv) l'aplicació del mètode r-based. Per a això hem realitzat mesures experimentals d'emissivitat utilitzant el mètode de la caixa (Rubio et al., 1997), el mètode TES (Gillespie et al., 1998) i utilitzant el model de Valor i Caselles (1996) en superfícies terrestres heterogènies per a calcular l'emissivitat a l'escala del píxel satel·litari i el model de Niclòs et al. (2009) per a l'emissivitat de la superfície de l'aigua.

Els algorismes monocanals (SC) (Chedin et al., 1982) calculen la LST a partir de la temperatura radiativa en una banda espectral i requerixen, a més de l'emissivitat, els perfils atmosfèrics en el moment de pas del sensor satel·litari i un model de transferència radiativa, amb els quals obtindre la transmissivitat i les radiàncies atmosfèriques ascendent i descendent de la banda utilitzada. Un altre tipus d'algorismes són els mètodes d'absorció diferencial, o de finestra dividida (SW; McMillin, 1975) que es basen en utilitzar dos temperatures radiatives dins de la mateixa finestra atmosfèrica de tal manera que no és necessari el coneixement dels perfils atmosfèrics. També existix el mètode multicanal de separació de temperatura i emissivitat (TES; Gillespie et al., 1998), que requerix d'almenys tres bandes en l'infraroig tèrmic i la correcció monocanal de cada banda de manera independent. Amb les radiàncies a nivell de superfície s'aconseguix desacoblar la temperatura de l'emissivitat i es pot obtindre de manera simultània la LST i l'emissivitat per a cada banda.

En realitzar el calibrat del sensor SLSTR front a mesures de 9 estacions situades sobre diferents cobertures (zones boscoses, desèrtiques, cultius de regadiu, llacs) els resultats obtinguts mostren una R-RMSD de 0.8 K i 0.7 K per a les bandes 8 i 9, respectivament en els dos satèl·lits (Sentinel-3A i -3B), amb mitjanes pròximes a zero en tots el casos. La validació t-based del producte operatiu de LST de SLSTR va presentar diferències mitjanes superiors a 1 K en totes les cobertes, excepte per a les zones boscoses que estaven dins de ±1 K. No obstant això, per a la zona dels arrossars de València va mostrar una diferència mitjana d'1.6 K i 2.0 K per a Sentinel-3A i -3B respectivament, el que pot ser degut a l'assignació d'un coeficient fix per al bioma quan realment es tracta d'una coberta canviant al llarg de l'any. D'altra banda, el SW proposat per Pérez-Planells et al. (2021) només mostra diferències mitjanes excessives (-1.7 K) en les zones desèrtiques amb dades diürnes, estant en l'interval de ±1 K per a la resta de cobertes. Els resultats amb el mètode r-based per al producte de LST de SLSTR van presentar diferències mitjanes superiors al llindar GCOS en les zones de cultius de regadiu i desèrtiques, i el SW proposat per Pérez-Planells et al. (2021) només en les zones desèrtiques amb dades diürnes, anàlogament a la validació t-based.

El sensor TIRS a bord de Landsat-8 va patir diferents reprocessaments que van afectar el calibrat de les bandes tèrmiques (bandes 10 i 11). Les temperatures radiatives amb el calibratge original van mostrar una sobreestimació de 0.7 K en la banda 10 i 2.7 K en la banda 11 respecte a les temperatures radiatives simulades a partir de dades in situ mesurades en la zona d'arrossars de València Després del reprocessament de 2014 vam obtindre subestimacions de -1.4 K i -1.6 K, respectivament. No obstant això, després de l'aplicació de la correcció TIRS-on-TIRS (Montanaro et al., 2015) les desviacions mitjanes van ser molt pròximes a zero en les dos bandes. En 2019 es va realitzar un nou reprocessament amb el qual les diferències mitjanes van passar a ser de -0.1 K i 0.5 K per a les bandes 10 i 11, respectivament, introduint una desviació sistemàtica significativa entre les dos bandes. . Els resultats de validació t-based de la LST obtinguda amb el mètode SC mostren que no existix diferència entre el processament de 2017 i el de 2019 en termes de R-RMSD, sent este 0.7 K en els dos casos. L'algorisme SW que millor funciona amb dades TIRS en el processament de 2017 és el proposat per Jiménez-Muñoz et al. (2014) amb una R-RMSD de 0.6 K, mentres que en el processament de 2019 les diferències mitjanes de tots els SW estudiats (un total de 17) van augmentar entre 1.3 K i 1.9 K depenent de l'algorisme.

El calibrat de les bandes tèrmiques del sensor TIRS-2 a bord de Landsat-9 va ser avaluada enfront de dades in situ en la zona d'arrossars de València calibrades per a les dades originals i després del reprocessament de 2023. Els resultats amb les dades originals van oferir unes diferències mitjanes de -0.3 K i -0.1 K per a les bandes 10 i 11, respectivament. Després del reprocessament, la banda 10 va mostrar una diferència mitjana menyspreable, mentres que va ser de 0.2 K en la banda 11, la qual cosa fa que la desviació entre les bandes augmente i afecte el funcionament dels SW. La validació del producte operatiu de LST i dels SW estudiats amb el mètode t-based presenten diferències mitjanes dins de l'interval ±1 K, tant en les dades originals com en els reprocessaments. Encara que es va observar una diferència sistemàtica de 0.6 K entre les dades originals i els reprocessaments a causa dels canvis en el calibratge de les bandes tèrmiques.

En el cas del sensor MODIS, inicialment comparem les LST de referència del mètode r-based front a un conjunt de mesures in situ de LST en la zona d'arrossars de València per a estimar la incertesa de les primeres en relació respecte a les segones. A continuació, hem realitzat la validació r-based en zones amb diverses cobertures dels productes de LST de M*D11 i M*D21 (MOD per al satèl·lit Terra i MYD per a Aqua). D'acord amb estos resultats, el producte M*D11 va presentar una subestimació de -1.0 K (RMSD d'1.2 K) per a les dades diürnes i de 0.0 K (RMSD de 0.5 K) per a les dades nocturnes, tenint en compte totes les zones de validació conjuntament. D'altra banda, la desviació mitjana (RMSD) mostrada pel M*D21 va ser de 0.1 K (0.7 K) i 0.3 K (0.5 K) per a les dades diürnes i nocturnes, respectivament, per als dos satèl·lits i tot el conjunt de dades. El mètode de validació r-based proporciona una alternativa valuosa al mètode t-based per a la majoria de les cobertes terrestres on les mesures in-situ de LST no són possibles.

També realitzem la validació t-based al producte operatiu de LST de LSA-SAF i algorismes SW del sensor AVHRR/3 a bord de MetOp-A, -B i -C. El producte va mostrar resultats per a la versió operativa en R-RMSD de 3.4 K i 2.0 K per a les dades diürnes i nocturnes, respectivament. Es proposa un algorisme SW que presenta mitjanes pròximes a zero en totes les cobertes i als tres satèl·lits. L'algorisme proposat per Jiménez-Muñoz i Sobrino (2008) presenta similars resultats a excepció del cas de la vegetació on sobreestima les LST entorn d'1 K i per a MetOp-C, on la desviació mitjana per a totes les cobertes és d'1.4 K.

Capítulo 1

INTRODUCCIÓN

En esta introducción indicamos los conceptos y términos necesarios para terminar definiendo los conceptos de la tesis.

a. Temperatura de la superficie terrestre

A lo largo de los años, una de las variables que más importancia ha adquirido en Teledetección es la Temperatura de la Superficie Terrestre (LST por *Land Surface Temperature* en inglés), que ha sido designada como una *Essential Climate Variable* (Bojinski et al., 2014; Guillevic et al., 2018; Hollmann et al., 2013) por el *Global Climate Observing System* (GCOS). Por tanto, se entiende la importancia de obtener un valor preciso de dicha magnitud para multitud de estudios medioambientales, ya que se trata de una variable principal en muchos de ellos. La LST es necesaria para el estudio del balance energético entre la masa atmosférica y el suelo, del estrés hídrico que puede sufrir la vegetación (Anderson, 1997; Anderson et al., 2012; Boronina and Ramillien, 2008; Er-Raki et al., 2008; Galleguillos et al., 2011; Hall et al., 1992; Jacob et al., 2008; Mokhtari et al., 2019; Sánchez et al., 2008; Sellers et al., 1995), como dato de entrada para modelos de predicción meteorológica (Barton et al., 1989; Ghent et al., 2010; Goutorbe et al., 1994; Trigo et al., 2015; Weng, 2009), y para la detección de incendios (Calle et al., 2006; García-Llamas et al., 2019; Roy et al., 2015). La variable LST se considera como un indicador del cambio climático (Alien et al., 1994) y de la desertificación de una zona (Lambin and Ehrlich, 1997), del efecto isla de calor (Peres et al., 2018), el riesgo de catástrofe natural (Darge et al., 2019) y otras muchas aplicaciones. Utilizar la teledetección y los sensores en el infrarrojo térmico (IRT) a bordo de satélite para obtener la LST es la única posibilidad de obtener dicha variable de manera diaria y a escala tanto regional como global.

b. Métodos para la determinación de la LST

Los sensores a bordo de satélite miden la temperatura de brillo (*Brigthness Temperature,* BT) en el techo de la atmósfera. Sin embargo, lo que nos interesa como usuarios es la temperatura de la superficie terrestre. Para obtener la LST a partir de la BT que proporciona el satélite en el techo de la atmósfera se debe corregir los efectos atmosféricos y de la emisividad de la superficie.

La absorción y la emisión atmosférica en la región del IRT son dominadas por el vapor de agua y por la distribución de la humedad y temperatura atmosférica (Bignell, 1970; Maul and Sidran, 1973), las cuales tienen una gran variación tanto espacial como temporal debido a que depende de la temperatura y presión atmosféricas y la presión parcial del vapor de agua. Por otro lado, el resto de gases atmosféricos absorbentes como el ozono, dióxido de carbono, metano tienen una contribución pequeña y uniforme. Además, la atmósfera afecta de manera diferente en función de la longitud de onda de los canales térmicos. Para realizar la corrección atmosférica es necesario conocer la transmisividad entre la superficie observada y el sensor y la radiancia emitida por la atmósfera tanto en la dirección ascendente como descendente. Estas magnitudes atmosféricas pueden calcularse a partir de los perfiles verticales de presión, temperatura y razón de mezcla o humedad relaiva, y un modelo de transferencia radiativa, como por ejemplo MODTRAN 5.2 (Berk et al., 2008).

La emisividad de la superficie se define como el cociente entre la radiación que emite una superficie y la radiación que emitiría un cuerpo negro a la misma temperatura. La emisividad varía en función de la longitud de onda, del ángulo de observación, del tipo de cobertura y de la humedad del suelo (García-Santos et al., 2012; Olioso et al., 2007; Salisbury and D'Aria, 1992), la dependencia angular puede observarse de manera muy marcada en el caso del agua (Niclos et al., 2005; Niclòs et al., 2014). Usualmente la emisividad de la superficie debe ser determinada a priori a través de medidas experimentales o modelos (Gillespie, 1998; Rubio et al., 1997; Valor y Caselles, 1996).

Para realizar la corrección atmosférica y de emisividad en el IRT y así obtener la LST se puede utilizar distintos métodos.

b.1. La aplicación de la ecuación de transferencia radiativa (ETR) en el método monocanal (*Single Channel*, SC) (Chedin et al., 1982; Price, 1983). El método SC se basa en utilizar un modelo de transferencia radiativa con el que a través de los perfiles atmosféricos de vapor de agua y temperatura calcular la transmisividad y la radiancia emitida por la atmósfera en una banda espectral y por tanto obtener las radiancias a nivel de superficie resolviendo la ETR a partir de las radiancias a nivel del sensor satelital. Es necesario también conocer la emisividad de la superficie en la banda utilizada para obtener la LST. El método SC está limitado a la disponibilidad perfiles atmosféricos coincidentes con la hora de paso del sensor satelital y cercanos a la zona de interés y la aplicación de un modelo de transferencia radiativa.

b.2. El método bicanal de ventana dividida (*Split-Window*, SW). Para evitar tener que utilizar los perfiles atmosféricos, radiancias y transmisividades atmosféricos en métodos SC existen métodos de absorción diferencial (McMillin, 1975), que se desarrollaron originalmente para obtener la temperatura de la superficie del mar (SST). Estos métodos, utilizan medidas satelitales sobre una misma superficie, pero en condiciones de observación diferentes, como el método SW que utiliza dos canales dentro de la ventana de 10-12.5 µm sometidos a diferente absorción atmosférica. El método SW se adaptó para la determinación de la LST teniendo en cuenta la emisividad de la superficie en las dos bandas utilizadas (Becker and Li, 1990; Coll et al., 1994; Coll and Caselles, 1997; Dash et al., 2002; Deschamps and Phulpin, 1980; Jiang and Li, 2008; Niclòs et al., 2011; Pérez-Planells et al., 2021; Prabhakara et al., 1974; Prata, 1994; Sobrino and Romaguera, 2004; Sun and Pinker, 2003; Wan and Dozier, 1996). Este tipo de algoritmos viene caracterizado por su facilidad de aplicación y eficiencia computacional.

Los algoritmos SW se obtienen generalmente mediante regresiones semiempíricas entre la LST y las BTs en las dos bandas utilizadas a partir de simulación de las BTs para un conjunto de escenarios que cubren una variedad amplia de condiciones geofísicas (tipo de superficie, temperatura de la misma, vapor de agua atmosférico y ángulo de observación). Normalmente los algoritmos SW expresan la LST como combinaciones lineales o cuadráticas de las BTs. Los coeficientes de los algoritmos pueden ser fijos o pueden variar en función del tipo de superficie, rangos de vapor de agua o rangos de temperaturas. Independientemente de esto, la necesidad de conocer los valores de la emisividad de la superficie para poder obtener la LST es el principal problema que presentan esto tipo de métodos, ya que tienen una fuerte dependencia con la emisividad. Esto

hace que los algoritmos SW funcionen mejor sobre superficies cuya emisividad es similar a la de un cuerpo gris como agua o vegetación densa, donde la variación espectral de la emisividad es pequeña. Otro dato necesario para obtener la LST a través de los algoritmos SW es el contenido de agua precipitable en la atmosfera, W, que representa la masa de vapor de agua en una columna atmosférica por unidad de área.

b.3. El método multicanal de separación de temperatura y emisividad (*Temperature and Emissivity Separation*, TES) (Gillespie et al., 1998). El método TES requiere de las radiancias a nivel de superficie en tres o más bandas dentro del intervalo 8-12 μm. Para ello se realiza la corrección atmosférica de las radiancias a nivel de satélite en cada banda mediante perfiles de temperatura y humedad y un modelo de transferencia radiativa, para obtener así las radiancias a nivel de superficie para cada banda (Palluconi et al., 1999). En dichas radiancias, la emisividad y la temperatura están acopladas, con lo que es necesario una hipótesis adicional sobre la variación espectral de la emisividad basada en medidas experimentales de esta magnitud para un amplio tipo de superficies naturales (Gillespie et al., 1998). Con esta hipótesis, se consigue desacoplar temperatura y emisividad, obteniéndose simultáneamente la LST y la emisividad de cada banda.

Originalmente, los sensores satelitales tan solo poseían una banda en el espectro térmico por lo que era necesario aplicar algoritmos SC. Más adelante aparecieron los sensores con al menos dos bandas en la región IRT y esto hizo que se utilizaran los algoritmos SW y TES para obtener la LST. Conocer la emisividad de la superficie para cada banda en necesario para poder obtener la LST en el caso de los métodos SC y SW, mientras el método TES proporciona la emisividad en cada banda y la LST conjuntamente.

c. Validación LST y calibración de bandas IRT

c.1. Importancia de la validación y la calibración

Para cualquier producto operativo de LST que esté disponible para los usuarios (ya sea en tiempo real o en tiempo no crítico), realizar una validación de estos productos es una rutina esencial. De esta forma, podemos establecer de forma absoluta la precisión y exactitud de los valores de LST satelitales. Por tanto, es necesario validar los diferentes algoritmos sobre diferentes superficies terrestres en términos de LST. En cuanto la incertidumbre en la LST obtenida desde satélite, GCOS recomienda unos umbrales en términos de la diferencia media (±1.0 K) y la desviación estándar (SD) de las diferencias (1.0 K). La diferencia media y la SD se corresponden con la incertidumbre sistemática y la precisión, respectivamente, de acuerdo con el *Joint Committee for Guides in Metrology* (JCGM, 2008). Por otra parte, es necesario también comprobar la calibración a bordo de los sensores satelitales comparando la BT medida por las bandas térmicas con una referencia independiente (calibración vicaria).

c.2. Validación t-based

El método de referencia para la validación de la LST satelital es la comparación con medidas in situ de LST tomadas simultáneamente al paso del satélite, lo que se conoce como el método t-based. De esta forma se puede

establecer el error de los productos LST de una forma independiente en relación a una "verdad terreno". El problema reside en la dificultad de obtener datos in situ de LST, medidos con radiómetros de campo, cuyo campo de visión es del orden de 1 m^2, que sean representativos de la LST correspondiente a un área del tamaño de un píxel satelital (hasta 1 km^2). Las zonas de validación deben ser extensas y térmicamente homogéneas para que puedan representar varios píxeles de satélite (Slater et al., 1996; Wan et al., 2002). Esta homogeneidad se debe dar tanto en términos de LST como en términos de emisividad, para escalas desde varios metros (el caso de la visión desde las estaciones radiométricas), por debajo del metro (el caso de los radiómetros de mano) y llegar a 1 km (el caso de la mayoría de los sensores térmicos satelitales). Además, las medidas in situ deben ser simultáneas al paso del satélite dada la variabilidad temporal de la LST.

Los motivos anteriormente citados hacen que las zonas de validación sean escasas y normalmente limitadas a superficies con una sola componente, como cuerpos de agua, superficies con una cobertura vegetal completa o casi nula (desiertos, zonas áridas). Son numerosos los trabajos que han utilizado estas zonas para realizar validaciones. El lago Tahoe (EE. UU.) dispone de boyas monitorizadas para medir la temperatura del agua y es utilizado en diversos trabajos para validar diferentes sensores como EOS/MODIS y EOS/ASTER (Hook et al., 2001; Tonooka et al., 2005), de ENVISAT/AATSR (Coll et al., 2009; Prata, 2003) y Landsat-7 (ETM), Landsat-8 (TIRS) y Landsat-9 (TIRS-2) (Barsi et al., 2006, 2014, 2022; Schott et al., 2012). Además de sobre superficies de agua, también se ha realizado con los mismos sensores sobre superficies de cultivos de arroz con distintos tipos de cobertura a lo largo del año (Coll et al., 2005; Niclòs et al., 2018, 2021, 2023; Pérez-Planells et al., 2021). O también sobre superficies desérticas (Ghent et al., 2017; Göttsche et al., 2013, 2016; Zhou et al., 2017). Por otro lado, las medidas in situ de LST pueden utilizarse, junto con medidas de emisividad y perfiles atmosféricos, para simular la BT medida por el sensor satelital y compararla con el valor real.

En ocasiones se ha utilizado como LST in situ medidas realizadas en estaciones meteorológicas y climatológicas mediante sensores de radiación de onda larga o pirgeómetros (Göttsche et al., 2016; Guo et al., 2020; Li et al., 2021, 2023). Sin embargo, los pirgeómetros infrarrojos miden flujos de banda ancha (4-50 μm), integrados hemisféricamente, que no se corresponden necesariamente con las medidas de radiancia direccionales de banda estrecha de los satélites, especialmente para datos diurnos (Wang et al., 2008). Lo ideal sería que las mediciones de LST en tierra se realizaran con radiómetros IRT que tomaran mediciones direccionales cercanas al ángulo de visión del satélite en bandas estrechas similares a las de los instrumentos a bordo de satélite. En esta Tesis, todas las LST in situ han sido medidas mediante radiómetros IRT direccionales de banda estrecha.

Las medidas in situ de temperatura se obtienen generalmente con radiómetros instalados en estaciones fijas o con radiómetros de mano realizando transectos a lo largo de la zona de validación. Los radiómetros de estaciones fijas realizan medidas sobre la misma zona y de manera continua y frecuente (desde 4 s a 1 min dependiendo de la estación), y se promedia las medidas tomadas durante el periodo entre 3 min antes y después del paso del sensor. En los transectos, la manera de proceder consiste en distribuir distintos radiómetros (hasta 4 en nuestro caso) de mano en diferentes puntos cubriendo la zona de interés. Cada uno toma medidas cada 4~5

m, de manera que se toma unas 25~30 medidas durante los 3 min anteriores y posteriores al paso del satélite. Así se logra cubrir unos 150 m con un solo radiómetro, cubriendo en total una parte significativa de un píxel de los sensores de baja resolución espacial, y decenas de píxeles en el caso de los sensores de alta resolución espacial.

La validación t-based es la más utilizada para comprobar el correcto funcionamiento de los producto y algoritmos que proponen los autores de los SW (Coll et al., 2006, 2010, 2012; Galve et al., 2008; Hook et al., 2001; Niclòs et al., 2011, 2018, 2021, 2023; Pérez-Planells et al., 2021; Prata, 2003; Tonooka et al., 2005).

El inconveniente principal que presenta el método t-based es el coste tanto material como personal de mantener las instalaciones de los radiómetros en condiciones adecuadas en el caso de las estaciones meteorológicas o en el caso de la realización de transectos con radiómetros de mano. Otra desventaja es la variabilidad intrínseca de la LST a escalas espaciales y temporales muy pequeñas debido a turbulencias atmosféricas de alta frecuencia que pueden hacer variar la temperatura en varios K, en breves intervalos de tiempo (<2 min) y en áreas no muy extensas (decenas de m) (Lagouarde et al., 2015). Este efecto no puede apreciarse a nivel de satélite, pero si a nivel de campo, con lo cual se puede inducir una desviación entre las medidas de campo y de satélite. Este problema se minimiza en el caso de los transectos al utilizar varios instrumentos abarcando una mayor área de la zona de interés.

c.3. Validación r-based

Debido a las dificultades del método t-based, se ha propuesto un método de validación alternativo denominado r-based, que puede aplicarse potencialmente a un mayor número de tipos de superficie y no requiere medidas in situ de LST (Wan and Li, 2008). El método r-based fue inicialmente propuesto para la validación de productos de LST del sensor EOS/MODIS a escala global. Este método utiliza el algoritmo SC para calcular las LST de referencia a partir de las temperaturas de brillo y las radiancias atmosféricas ascendentes, descendentes y transmisividades en los dos canales de la ventana 10-12 µm. La LST así calculada se puede utilizar como referencia para la validación tanto de los productos operativos que ofrecen las agencias de cada uno de los sensores como de los algoritmos SW propuestos por los diferentes autores para los diferentes sensores. El método se ha utilizado en esta tesis para realizar la validación de los sensores MODIS, SLSTR y Landsat-8.

El método r-based permite realizar la validación de las LST satelitales en zonas donde las mediciones in situ no están disponibles o no son factibles (Coll et al., 2010, 2024; Hulley and Hook, 2012; Li et al., 2023; Wan and Li, 2008). Proporciona información valiosa para ampliar la validación de los productos LST sobre una gama más amplia de cubiertas terrestres y, por lo tanto, consolidar la etapa de validación de los diferentes productos LST (Guillevic et al., 2018). En la sección 2.f.1 se detalla el procedimiento seguido en el método r-based.

d. Sensores

En esta Tesis, hemos utilizado los datos de los siguientes sensores satelitales con bandas en el IRT.

d.1. Sensor SLSTR

El sensor Sea and Land Surface Temperature Radiometer (SLSTR) de la misión Copernicus Sentinel-3 responde a los requerimientos de tener monitorizada en tiempo real la superficie de la tierra a lo largo del tiempo. El instrumento mide las radiancias IRT de la superficie del mar y de la tierra con gran precisión y fiabilidad (Donlon et al., 2012) como ayuda a la monitorización del clima y el medio ambiente bajo el programa de la Unión Europea Copernicus. El satélite Sentinel-3A, que fue lanzado el 16 de febrero de 2016, y el Sentinel-3B, que fue lanzado el 25 de abril de 2018, ofrecerán servicios de LST con cobertura diaria casi global para los próximos años y tendrán continuidad con los satélites 3C y 3D. El SLSTR posee 9 canales espectrales en el visible e infrarrojo de onda corta, uno que está centrado en 3.7 µm y dos en el IRT (banda 8 en 10.8 µm y banda 9 en 12 µm, Figura 1) que son los utilizados en esta tesis.

SLSTR es la continuación del Along Track Scanning Radiometer (ATSR) (Llewellyn-Jones and Remedios, 2012) y su sucesor Advanced Along Track Scanning Radiometer (AATSR) que dejo de estar operativo en abril de 2012 por un fallo en la plataforma Envisat. Dichos sensores (SLSTR y AATSR) son de barrido cónico y tienen la capacidad de obtener datos con la visión doble, una cercana a nadir y otra oblicua. En el SLSTR, la visión a nadir abarca una anchura de barrido de 1420 km, mientras que la visión oblicua tiene 750 km de anchura como se puede observar en la Figura 2. Los sensores tienen una calibración durante el vuelo basada en la observación de un cuerpo negro a alta temperatura a bordo del satélite y el vacío como cuerpo negro a baja temperatura (Smith et al., 2021).

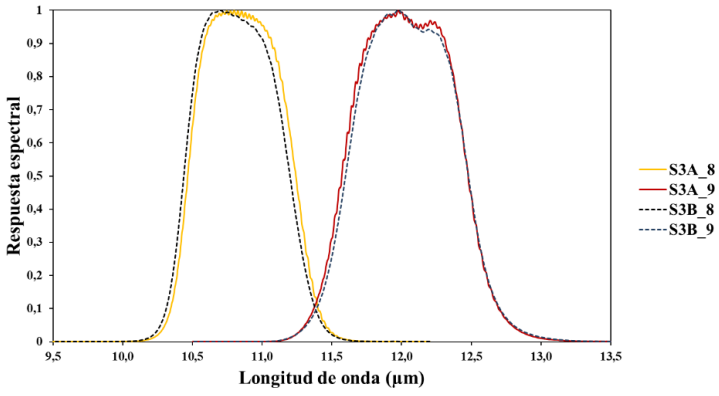

Figura 1. *Respuesta espectral de las bandas IRT de los sensores SLSTR.*

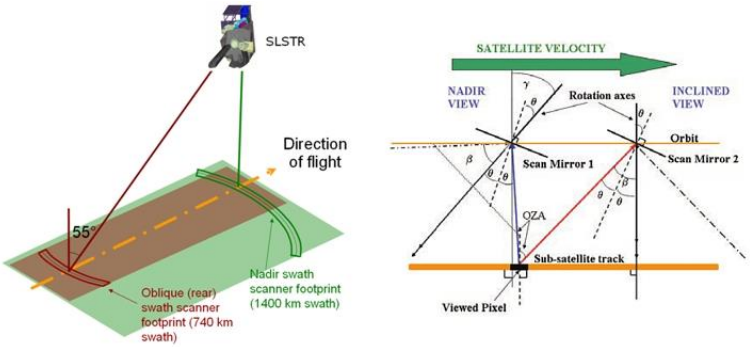

***Figura 2**. Visión del sensor SLSTR en nadir y oblicua. Imagen extraída de la ESA: https://sentiwiki.copernicus.eu/web/s3-mission.*

El sensor SLSTR ofrece un producto operativo de LST de nivel 2 basado en un algoritmo SW utilizando las bandas 8 y 9, con el objetivo de obtener la LST de manera global con una precisión por debajo de 1 K.

d.2. Sensor TIRS y TIRS-2

La Agencia Espacial Norteamericana (NASA) y el Servicio Geológico de los Estados Unidos (USGS) lanzaron el 11 de febrero de 2013 la plataforma Landsat-8 (L8) y en septiembre de 2021 la Landsat-9 (L9). Son la continuación de las misiones previas de la serie Landsat que se remonta a 1972. El sensor Operational Land Imager (OLI, en el espectro solar) y el Thermal Infrared Sensor (TIRS, en el IRT) de L8 empezaron a adquirir imágenes en abril de 2013, y el OLI-2 y TIRS-2 de L9 en febrero de 2022. El L8 fue el primero de la serie Landsat que incorporaba dos bandas en el TIR, la banda 10 (10.60-11.19 µm) y la banda 11 (11.50-12.51 µm), al igual que el L9 como se muestra en la Figura 3. Ambos poseen una gran resolución espacial: 100 m que se reescalan a 30 m, que es la resolución de las bandas de OLI. La resolución temporal (periodo de revisita) que poseen los sensores TIRS y TIRS-2 es de 8 días.

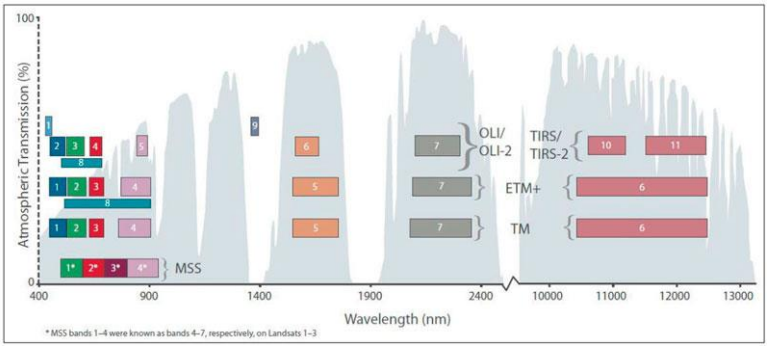

***Figura 3.** Respuesta espectral de las bandas de los sensores OLCI, TIRS y TIRS-2. Imagen extraida de https://landsat.gsfc.nasa.gov/satellites/landsat-8/*

Las bandas 10 y 11 hacen que se pueda aplicar algoritmos SW para obtener la LST evitando el método SC que requiere información detallada sobre los perfiles atmosféricos.

En noviembre de 2013 el equipo de calibración de L8 avisó de la necesidad de reprocesar los datos TIRS debido a unas discrepancias que se observaron al comparar estos con datos obtenidos in-situ sobre superficies de agua de referencia. Se propuso desde el USGS Earth Resource Observation Science (EROS) restar unos valores fijos de radiancia TOA a las bandas 10 y 11 de TIRS, equivalentes en términos de temperatura a 2.1 K y 4.4 K, respectivamente (reprocesado de 2014), pero aun así las radiancias que se obtuvieron eran superiores a las esperadas para las bandas TIRS (Barsi et al., 2014). Además, los usuarios no podían aplicar los SW debido a las grandes incertidumbres que presentaban los datos de la banda 11. Montanaro et al. (2014a, 2014b, 2015) propusieron un algoritmo para reducir la radiancia que se introducía en el sensor de fuera de su campo de visión (*stray-light*). Gerace y Montanaro (2017) mostraron un buen ajuste cuando se comparaban los datos con las escenas EOS-Terra Moderate Resolution Image Spectrometrer (MODIS) como referencia, lo que condujo a un nuevo reprocesado en 2017. Finalmente, en diciembre de 2019 se anunció que se realizaría un recalibrado radiométrico a toda la serie en la llamada Colección 2, en el cual se solucionan los problemas del *stray-light* pero a la vez se realizaban calibraciones auxiliares tomando como referencia medidas in situ en superficies de agua.

Para L9 el equipo de Calibración/Validación de USGS/NASA realizó una calibración radiométrica posterior al lanzamiento sobre el lago Tahoe utilizando boyas para medir la temperatura del agua (Barsi et al., 2022; Montanaro et al., 2022). Los primeros datos se publicaron con la calibración realizada previa al lanzamiento mientras que en marzo de 2023 se realizó una recalibración con los resultados obtenidos por Barsi et al. (2022). Tras el lanzamiento de la colección 2 se ofreció al usuario un producto de LST basado en un algoritmo SC con los datos de la banda 10 tanto para L8 como para L9.

d.3. Sensor MODIS

El sensor MODIS a bordo de las plataformas EOS/Terra y Aqua de la NASA fue lanzado en diciembre de 1999 y en mayo de 2002, respectivamente, los cuales proveen de manera global 4 productos diarios de LST y de emisividad (2 diurnos y 2 nocturnos). Este instrumento ha sido utilizado en multitud de ocasiones como sensor de referencia para la calibración de otros satélites. MODIS tiene unas características en el diseño tanto espectral como en cobertura que mejoraba a sus predecesores AVHRR a bordo de NOAA-14 y NOAA-15 (Barnes et al., 1998). Posee 36 bandas espectrales que se sitúan en la región del visible, el infrarrojo cercano, el infrarrojo medio y en el IRT (bandas 29, 31 y 32 entre 8.0 y 12.5 µm, Figura 4), que son las que se utilizan para el cálculo de la LST. Las bandas IRT tienen una resolución espacial de 1 km y un ángulo de visión de ±65º con respecto a nadir y un ancho de visión de 2330 km. El tamaño del píxel aumenta según se acerca los bordes de la imagen llegando a los ~3 km (Masuoka et al., 1998). Para la calibración del sensor se utiliza un cuerpo negro que lleva incorporado a bordo del satélite y se determinan unos coeficientes de calibración para cada detector, aunque al usuario le llegan las radiancias ya calibradas.

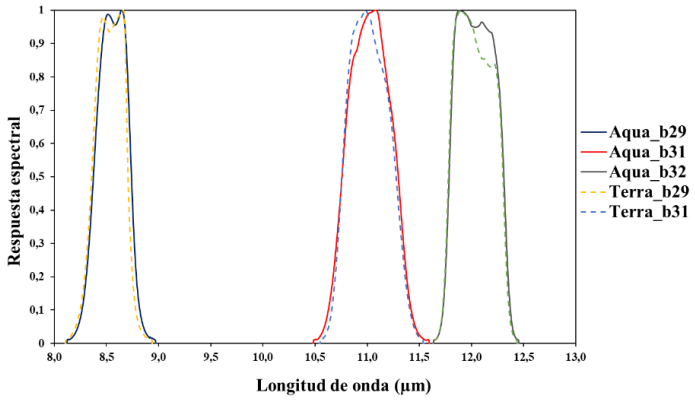

Figura 4. *Respuesta espectral de las diferentes bandas espectrales del infrarrojo térmico del sensor MODIS.*

MODIS ofrece los productos de LST M*D11 y M*D21 (MOD para el satélite Terra y MYD para el satélite Aqua) que se obtienen de dos formas distintas. En el producto M*D11, la LST se obtiene a través de la aplicación de un algoritmo SW con las bandas 31 (11 µm) y 32 (12 µm), mientras que M*D21 proporciona la LST con la aplicación del método TES utilizando también la banda 29 (situada en 8.45 µm).

d.4. Sensor AVHRR/3

El sensor Advanced Very High Resolution Radiometer/3 (AVHRR) que va a bordo de las plataformas MetOp-A, -B y -C posee 6 bandas espectrales en el rango de 0.58 a 12.5 µm, cubriendo el visible, infrarrojo cercano, el medio y el térmico. En este último posee dos canales en 10.8 y 12 µm (canales 4 y 5, respectivamente), como se muestra en la Figura 5, que son las que se utilizan para el cálculo de la LST mediante un algoritmo SW.

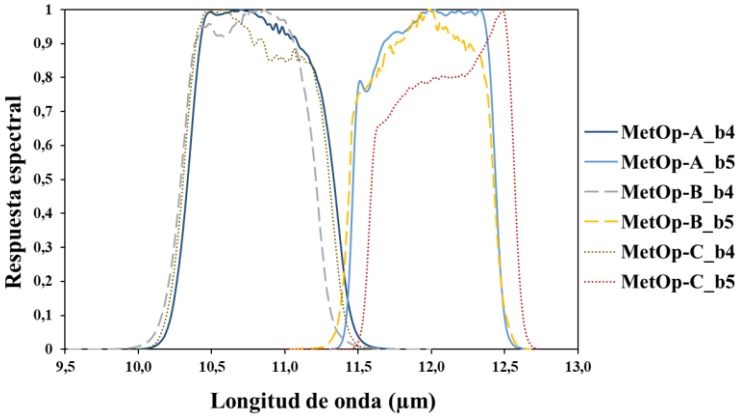

Figura 5. *Respuesta espectral de las bandas del infrarrojo térmico del sensor AVHRR/3 a bordo de MetOp.*

MetOp-A -B y -C fueron lanzados en octubre de 2006, en septiembre de 2012 y en noviembre de 2018 respectivamente. La plataforma MetOp-A fue retirada en noviembre de 2021 tras 15 años de adquisición de

datos dejando paso a sus sucesores MetOp-B y -C. Este sensor ofrece imágenes diurnas y nocturnas de tierra, mar y nubes de manera diaria. La resolución espacial del sensor es de 1 km en nadir y llega a medir a ángulos de observación de ±65º por lo que la imagen tiene una anchura de ±1447 km. Lo que se ofrece al usuario son las radiancias en el techo de la atmosfera (TOA), aunque también se ofrece el producto LST basado en un algoritmo SW por parte de Satellite Application Facilities Land Surface Analysis (LSA SAF) que pertenece a European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT).

d.5. Otros sensores y misiones futuras

Además de los que se estudian en este trabajo, existen en la actualidad otros satélites cuyos sensores poseen bandas en el infrarrojo térmico como Viirs a bordo de Suaomi/NPP, la serie NOAA con el sensor AVHRR o los Meteosat. Esto hace que de manera diaria se tenga datos de distintos sensores de manera diurna y nocturna con una cobertura global (Barton et al., 1989; Krishnan et al., 2020; Niclòs et al., 2011, 2018). Algunos proyectos como el Climate Change Iniciative (CCI) tienen como finalidad ofrecer un valor diario de LST (además de otras tantas variables) tanto diurno como nocturno de manera global a una resolución de 1 km (Martin, 2020; Plummer et al., 2017). CCI unifica y armoniza distintos sensores para poder tener valores de los últimos 20 años y de esta forma observar las tendencias en cuanto a esta variable a lo largo del tiempo (Ghent et al., 2024; Good et al., 2022; Sismanidis et al., 2022).

En la actualidad existen misiones de prueba para llegar a tener una alta resolución espacial como es el ECOsystem Spaceborne Thermal Experiment on Space Station (ECOSTRESS) que contiene diferentes bandas espectrales en la región TIR, y va a bordo de la Estación Espacial Internacional (SSI). Es el precursor del próximo sensor Surface Biology and Geology (SBG) que tiene 5 canales en el IRT, con una alta resolución espacial de 60 m aproximadamente (Fisher et al., 2020; Gerhards et al., 2019; Hulley et al., 2021; Meerdink et al., 2019; Turpie et al., 2023). Entre los productos que ofrece este sensor está la LST, el estrés hídrico, la evapotranspiración y el uso eficiente del agua. Otro precursor es el Hyperespectral Thermal Emission Spectrometrer (HyTES) que incluye 256 bandas en el rango de 7.5–12 µm y en la actualidad vuela a bordo de una avioneta para realizar las pruebas necesarias antes de ser montado en un satélite (Gerhards et al., 2019; Johnson et al., 2009, 2011). Con esta misión se pasará de sensores térmicos multiespectrales a hiperespectrales con una gran cantidad de bandas.

También existen otras futuras misiones propuestas para los próximos años que llevarán a bordo sensores como son el Thermal infraRed Imaging Satellite for High-resolution Natural resource Assessment (TRISHNA) que tendrá 5 canales en el IRT (Lagouarde et al., 2018, 2019; Vidal et al., 2022) y el Land Surface Temperature Monitoring (LSTM) que tendrá 4 canales en el IRT (Alfieri et al., 2017; Koetz et al., 2018).

e. Objetivo de la tesis

El objetivo de la presente tesis es, por una parte, realizar una evaluación de la calibración vicaria de nuevos sensores térmicos a bordo de satélite, como el sensor SLSTR, el TIRS y el TIRS-2, basándonos en medidas in situ. Esto es especialmente importante en el caso de los sensores TIRS y TIRS-2, debido a los distintos

recalibrados y reprocesados que se han ido aplicando a los datos originales. Por otra parte, tenemos como objetivo validar los productos LST que ofrecen las agencias de los diferentes satélites mencionados anteriormente, así como otros algoritmos LST propuestos por diferentes autores para dichos sensores, incluidos los algoritmos SW para SLSTR y AVHRR/3 propuestos en esta tesis

Todos los algoritmos de LST, tanto los productos operativos como los propuestos, son validados frente a medidas in situ de temperatura en diferentes áreas equipadas con estaciones de medida fijas y en otras donde las medidas de campo se adquieren mediante transectos de varios instrumentos a lo largo de la zona de interés (validación t-based). Uno de los objetivos principales es tener unos datos de referencia confiables para que puedan ser comparados con los que se obtienen desde satélite. Además, se ha aplicado el método r-based para la validación en las zonas t-based como comparación y en otras donde las medidas in situ no son posibles, aumentando así el número de tipos de superficie estudiados.

f. Estructura de la memoria

A continuación, se expone la estructura que se ha seguido en la realización de la presente memoria.

En el segundo apartado se describe la metodología de los diferentes procedimientos que se han aplicado en esta memoria. En primer lugar, planteamos la ETR con la que se simula las BTs de satélite y de la que se deriva el método SC. Seguidamente exponemos la metodología de las medidas in situ de LST y del método de validación r-based, y describimos los distintos algoritmos de LST que han sido evaluados. A continuación, abordamos los distintos métodos para obtener la emisividad de las cubiertas de las zonas de validación. Por último, exponemos el análisis de sensibilidad de las temperaturas (BTs en la calibración y LSTs en la validación) obtenidas como referencia es este estudio.

En el apartado 3 se presenta las regiones de interés que servirán como zonas de validación de los productos operativos y de los diferentes algoritmos de LST que se estudian en la memoria. Se describirá los radiómetros utilizados en cada zona y los cuerpos negros para la calibración de dichos instrumentos. También se describirá los datos satelitales que se ha utilizado de cada sensor y la resolución que ofrecen tanto espectral, espacial como temporal. Así mismo, se presentan los perfiles atmosféricos para poder realizar la corrección atmosférica a los datos de calibración y validación.

En el cuarto apartado se muestra los resultados que se ha obtenido para cada uno de los sensores estudiados. En primer lugar, la evaluación de la calibración del sensor SLSTR y la validación con el método t-based y r-based de sus diferentes algoritmos, tanto el operativo como los propuestos, en las diferentes zonas de validación. Siguiendo con el sensor TIRS y TIRS-2, del que se comprueba su calibración antes y después del reprocesado además de validar una gran variedad de algoritmos de LST que han propuesto diversos autores. Para el sensor MODIS de los satélites Terra y Aqua, validamos los productos M*D11 y M*D21 con el método r-based en diferentes superficies. Y por último para el sensor AVHRR/3 a bordo de MetOp-A, -B y -C se valida el producto operativo y otros algoritmos con el método t-based.

En el quinto apartado se expone las conclusiones más relevantes a las que se llegan tras realizar esta Tesis.

Por último, en los Anexos presentamos los artículos indexados incluidos en el Journal Citation Index fruto de la realización de esta investigación.

Capítulo 2

METODOLOGÍA

En esta sección se explican los procedimientos y metodologías que se han empleado para calibrar y validar los sensores satelitales y los algoritmos estudiados.

a. Ecuación de transferencia radiativa (ETR) y simulación de datos de satélite

La radiancia medida por la banda *i* de un sensor a bordo de un satélite (L_i^{sen}) está relacionada con la LST (T_g) y la emisividad de la superficie (ε_i) a través de la ecuación de transferencia radiativa (ETR) según:

$$L_i^{sen} = \left(\varepsilon_i \cdot B_i(T_g) + (1-\varepsilon_i) \cdot L_i^{\downarrow}\right) \cdot \tau_i(\theta) + L_i^{\uparrow}(\theta) \tag{2.1}$$

donde $L_i^{\downarrow}$ es la irradiancia atmosférica descendente dividida por π (o radiancia atmosférica descendente), τ_i es la transmisividad atmosférica y $L_i^{\uparrow}$ es la radiancia atmosférica ascendente, que deben ser calculadas a partir de perfiles atmosféricos con un modelo de transferencia radiativa, y θ es el ángulo cenital de observación. En la ecuación (2.1) asumimos que la reflectividad es $(1-\varepsilon_i)$ y que la reflexión es Lambertiana. B_i es la función de Planck integrada para cada una de las bandas de los sensores (en $W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}$), que viene dada por

$$B_i(T_g) = \frac{c_2}{\lambda_i^5\left[exp\left(\frac{c_1}{\lambda_i \cdot T_g}\right)-1\right]} \tag{2.2}$$

donde $c_1 = 1.1911 \cdot 10^8$ ($W \cdot m^{-2} \cdot \mu m^4 \cdot sr^{-1}$), $c_2 = 1.4388 \cdot 10^4$ ($\mu m \cdot K$), y λ_i es la longitud de onda efectiva para cada canal *i* del sensor.

La ecuación (2.1) puede ser utilizada para simular la BT del sensor a partir de valores de LST medidos en el campo y la emisividad de la superficie, la transmisividad y radiancias atmosféricas. La BT de la banda *i* (T_i) se define a partir de la radiancia en el sensor como

$$B_i(T_i) = L_i^{sen} \tag{2.3}$$

Con lo que podremos comparar las BT originales proporcionadas por el sensor satelital con las simuladas a partir de datos in situ.

b. Método monocanal (SC)

Por otro lado, también podemos utilizar la ecuación (2.1) para obtener la LST mediante el denominado método monocanal (SC) a partir de

$$B_i\left(T_{gi}\right) = \frac{L_i^{sen} - L_i^{\uparrow}(\theta)}{\tau_i(\theta)\cdot\varepsilon_i} - \frac{(1-\varepsilon_i)\cdot L_i^{\downarrow}}{\varepsilon_i} \tag{2.4}$$

y obtener la LST invirtiendo la función de Planck según la banda y el sensor utilizados. En el método SC la LST es denotada como T_{gi} ya que en principio puede depender de la banda *i* utilizada, aunque no deberían si ε_i y los parámetros atmosféricos son los adecuados y la L_i^{sen} se ha medido correctamente.

c. Medidas in situ de LST, para el método t-based

Independientemente de que se utilice radiómetros instalados en estaciones fijas o a lo largo de transectos, para obtener la LST a partir de las medidas de radiancia tomadas, se debe corregir el efecto de la emisividad y la radiancia atmosférica descendente. La LST in situ (T_g) se calcula a partir de la ecuación (2.1) pero teniendo en cuenta que el radiómetro está instalado cerca de la superficie por lo que $\tau_i(\theta) = 1$ y $L_i^{\uparrow}(\theta) = 0$ en (W/m^{-2} sr μm). Por tanto, tenemos

$$B_i(T_g) = \frac{L_i^{sup} - (1-\varepsilon_i)L_i^{\downarrow}}{\varepsilon_i} \tag{2.5}$$

donde L_i^{sup} es la radiancia medida en la banda i por los radiómetros en campo. $L_i^{\downarrow}$ es medida con un radiómetro apuntando al cielo a un ángulo de 53º en el caso de las estaciones fijas o con un panel de oro altamente reflectante en el caso de los radiómetros utilizados para los transectos (Guillevic et al., 2018). Las emisividades se determinan según el tipo de superficie con los métodos explicados en la sección posterior (2.f). En la función de Planck (ecuación (2.2)) se utiliza la longitud de onda efectiva de las bandas de los radiómetros de campo.

Los trabajos de validación t-based realizados durante el desarrollo de esta tesis han dado como fruto diversas publicaciones (Pérez-Planells et al., 2021; Niclòs et al., 2021, 2023; Coll et al., 2024; Puchades et al., 2025) (anexo B, C, D, E y F).

d. Método r-based

El método de validación r-based (Wan and Li, 2008) se basa en obtener la LST de referencia o in situ mediante la aplicación del método SC a las temperaturas de brillo del satélite (ecuación (2.4)) e invirtiendo la función de Planck con la ecuación (2.2) según la banda y el sensor utilizados. De la misma forma que el método t-based, el r-based también requiere del conocimiento de la emisividad de la superficie en las bandas del sensor satelital, y perfiles atmosféricos para obtener la transmisividad y las radiancias atmosféricas con un modelo de transferencia radiativa. Para la aplicación del método r-based es necesario que el sensor disponga de dos bandas en la ventana atmosférica 10-12.5 μm, ya que la ecuación (2.4) debe aplicarse a ambas bandas, que están afectadas por una absorción diferencial. En lo que sigue, denominaremos banda 1 a la situada en longitudes de onda más cortas y banda 2 a la situada en longitudes de onda más largas dentro de la ventana.

En un caso ideal, T_{gi} debería ser idéntico en ambas bandas, pero se ven afectadas por incertidumbres en la emisividad y los perfiles atmosféricos. Para la mayoría de las superficies naturales, como el agua, la vegetación y los suelos, la emisividad en la ventana de 10-12.5 μm es cercana a la unidad y muestra una pequeña variabilidad entre los tipos de superficie (Snyder et al., 1998; Pinheiro et al., 2004; Pérez-Planells et al., 2022). Por lo tanto, las incertidumbres de emisividad pueden mantenerse dentro de límites razonables. Más crítica es la necesidad de perfiles atmosféricos concurrentes precisos. La precisión de los cálculos de transferencia radiativa puede evaluarse a través de la diferencia en la LST obtenida de las dos bandas, $\delta = T_{g1} - T_{g2}$. Los

efectos atmosféricos en esta ventana se deben principalmente a la absorción del vapor de agua, por lo que están muy correlacionados entre bandas y son mayores en banda 1 que en banda 2. Por lo tanto, cuando el perfil atmosférico utilizado está sobrecorrigiendo el efecto atmosférico real en la banda 1, es decir, la diferencia entre la LST real y la derivada, $T_g - T_{g1}$, es negativa, la sobrecorrección en T_{g2} será mayor y entonces $\delta < 0$. Lo contrario ocurre en el caso de una subcorrección en la banda 1 ($T_g - T_{g1} > 0$), lo que dará lugar a una mayor subestimación en la banda 2 y, por tanto, $\delta > 0$. Por lo tanto, la diferencia $\delta = T_{g1} - T_{g2}$ debe estar dentro de un pequeño rango alrededor de cero cuando los perfiles atmosféricos son precisos y las incertidumbres en las emisividades superficiales son pequeñas.

El umbral δ que determina los casos válidos puede obtenerse mediante un análisis de incertidumbre de los diversos factores que intervienen en la derivación de T_{gi}, de modo que puede definirse un rango de valores de δ para mantener la incertidumbre en T_{g1} dentro de ciertos límites (Wan and Li, 2008; Coll et al., 2009). El valor utilizado en caso del sensor SLSTR es de $\delta = \pm 0.7\ K$ (Puchades et al., 2025). En el caso del sensor TIRS y MODIS el valor es de $\delta = \pm 0.5\ K$ (Coll et al., 2024; Niclòs et al., 2021).

El método r-based puede resumirse en los siguientes tres pasos.

1) Utilizando las radiancias del sensor en las bandas 1 y 2, los parámetros atmosféricos y los datos de emisividad, se obtiene la LST para ambas bandas, T_{g1} y T_{g2} con la ecuación (2.4).

2) Se calcula la diferencia $\delta = T_{g1} - T_{g2}$ y sólo se seleccionan como casos válidos aquellos para los que δ se encuentra dentro de un pequeño intervalo alrededor de cero.

3) Para los casos válidos, la temperatura calculada en la banda 1, T_{g1}, se considera la LST de referencia para la validación de la LST por satélite. Se selecciona la banda 1 porque los errores de corrección atmosférica son menores que en la banda 2.

En la formulación original (Wan and Li, 2008), δ se definía como la diferencia de temperatura de brillo ($T_1 - T_2$) obtenida a partir de los datos reales del satélite menos el valor calculado a partir de la simulación directa (ascendente) utilizando T_{g1} como LST de entrada en la ecuación (2.1). La nueva definición de δ (Coll et al., 2024) es más sencilla y se basa en la misma lógica que la definición original. Además, utilizamos una aproximación para la función de Planck promediada en banda, B_i, como una función analítica de la temperatura (ecuación 2.2) de modo que T_{gi} se puede invertir fácilmente a partir de la ecuación (2.4).

Las ventajas que presenta el método r-based es la posibilidad de aplicación en más zonas que el método t-based, siempre que sean homogéneas en emisividad y que al usar datos del mismo satélite la resolución tanto temporal como espacial serán iguales. Tampoco requiere del uso de instrumental en campo ya sea bien para realizar transectos o instalados en estaciones de toma de datos, para la adquisición de datos de referencia. Las desventajas que presenta son la dificultad de conseguir una correcta emisividad de la superficie y el conocimiento de los perfiles atmosféricos cercanos en tiempo y espacio a los datos satelitales.

e. Algoritmos LST

En esta sección presentamos los algoritmos LST de los distintos sensores que hemos estudiado en esta tesis. En general, denominamos banda 1 a la situada en longitudes de onda más cortas (alrededor de 11 µm) y banda 2 a la situada en longitudes de onda más largas (alrededor de 12 µm) dentro de la ventana atmosférica.

e.1. Algoritmos para SLSTR

e.1.1. Producto operativo LST SLSTR

En esta tesis se ha analizado el algoritmo operativo de LST del sensor SLSTR a bordo de Sentinel-3A y -3B (Prata, 2002), que proviene del planteado por (Prata, 1994), que se expresa mediante una ecuación que combina las temperaturas de brillo, según

$$T = a_{f,i,wcv} + b_{f,i}(T_1 - T_2)^{\sec(\frac{\theta}{m})} + (b_{f,i} + c_{f,i})T_2 \tag{2.6}$$

$$a_{f,i,w} = d[\sec(\theta) - 1]W + f a_{v,i} + (1 - f)a_{s,i} \tag{2.7}$$

$$b_{f,i} = f b_{v,i} + (1 - f)b_{s,i} \tag{2.8}$$

$$c_{f,i} = f c_{v,i} + (1 - f)c_{s,i} \tag{2.9}$$

donde T es la LST, T_1 y T_2 son las temperaturas de brillo de los canales 11 µm y 12 µm (bandas 8 y 9 de SLSTR), respectivamente, θ es el ángulo de observación cenital del satélite, m es un parámetro relacionado con el ángulo de observación y $a_{f,i,wcv}$, $b_{f,i}$ y $c_{f,i}$ son los coeficientes del algoritmo. Estos están calculados para 27 tipos de cobertura o biomas (subíndice i) y están definidos como una combinación lineal de dos componentes, una para la vegetación (subíndice v) y otra para el suelo desnudo (s) ponderados por la fracción de vegetación f (ecuaciones 2.7 a 2.9) que pude variar entre 0 y 1 y se obtiene de datos en el visible e infrarrojo cercano del sensor SLSTR. En el coeficiente $a_{f,i,w}$ se incluye un término que depende de W y de θ, a través de un coeficiente d.

Cada píxel se clasifica como uno de los 27 tipos diferentes de cobertura de suelo del producto de biomas Globcover con una resolución de 300 metros (Bicheron, 2008). Para cada caso, los coeficientes $a_{f,i,wcv}$, $b_{f,i}$ y $c_{f,i}$,i se determinan según su valor de f. Sólo en el caso de las aguas permanentes, se dan coeficientes diferentes para el día y la noche, independientes de f.

Nótese que el producto SLSTR es el único algoritmo de SW que no depende explícitamente de la emisividad de la superficie, sino que lo hace a través de coeficientes distintos para cada tipo de cobertura considerada, que son asignados en base a una clasificación previa de biomas. Por ello en este caso no se requiere del conocimiento de la emisividad de la superficie.

e.1.2. Algoritmo SW alternativo para SLSTR

En Pérez-Planells et al., (2021) propusimos un algoritmo alternativo dependiente explícitamente de la emisividad y basado en el propuesto por Niclòs et al., (2011) para el sensor SEVIRI a bordo del METEOSAT de segunda generación (MSG-2). Viene dado por

$$T = T_1 + a_0 + a_1(\sec(\theta) - 1) + \big(a_2 + a_3(\sec(\theta) - 1)\big)(T_1 - T_2) + \big(a_4 + a_5(\sec(\theta) - 1)\big)(T_1 - T_2)^2 + \alpha(1 - \varepsilon) - \beta\Delta\varepsilon \quad (2.10)$$

donde ε es el valor de la media de las emisividades de las bandas 1 y 2, $\varepsilon = 0.5(\varepsilon_1 + \varepsilon_2)$ y $\Delta\varepsilon$ es la diferencia entre las emisividades de las dos bandas, $\Delta\varepsilon = \varepsilon_1 - \varepsilon_2$, θ es el ángulo de observación del satélite, $\alpha = (a_6 + a_7 W^* + a_8 W^{*2})$ y $\beta = (a_9 + a_{10} W^*)$ que son los términos de corrección de la emisividad y $W^* = W/\cos(\theta)$ es el contenido de vapor de agua en la dirección de observación. Los valores de los coeficientes a_0 a a_{10} se dan en la Tabla 1 y el valor de W para aplicar el algoritmo se toma de los datos auxiliares de nivel 1 del SLSTR.

Tabla 1. Coeficientes para el algoritmo propuesto del sensor SLSTR (Pérez-Planells et al., 2021).

Coeficiente	Split-Window
a_0 (K)	0.052 ± 0.013
a_1 (K)	0.15 ± 0.02
a_2	0.95 ± 0.02
a_3	-0.30 ± 0.03
a_4 (K^{-1})	0.305 ± 0.004
a_5 (K^{-1})	0.202 ± 0.007
a_6 (K)	52.51 ± 0.18
a_7 (K cm^{-1})	-0.11 ± 0.12
a_8 (K cm^{-2})	-1.004 ± 0.018
a_9 (K)	75.7 ± 0.2
a_{10} (K cm^{-1})	-11.21 ± 0.06

Los coeficientes a_0 a a_5 se obtienen al realizar un análisis de regresión entre LST-T_1 y $T_1 - T_2$ sobre una base de datos de simulación de medidas de satélite tomando el caso del cuerpo negro ($\varepsilon = 1$ y $\Delta\varepsilon = 0$), ya que dichos coeficientes no dependen de la emisividad de la superficie (Coll et al., 1994). Los términos de corrección de la emisividad son α y β, los cuales dependen del contenido en vapor de agua de la atmósfera (W) (Coll and Caselles, 1997).

La base de datos de simulación utilizada para obtener los coeficientes a_0-a_5 fue generada a partir de la base de datos CLAR (Cloudless Land Atmosphere Radiosounding), que se compone de 383 perfiles atmosféricos

con cielo despejado a partir de radiosondeos recopilados por la Universidad de Wyoming (Galve et al., 2008). Estos perfiles están distribuidos de manera uniforme en las diferentes latitudes por lo que son aptos para generar algoritmos globales como se muestra en la Figura 6. El 40% de los radiosondeos pertenecen a latitudes entre 0 y 30, otro 40 % entre latitudes de 30 y 60 y el 20 % restante a latitudes superiores a 60. Los valores de W que se tomaron de estos perfiles se distribuyen entre 0 y 7 cm, y las temperaturas más bajas de la capa de la atmósfera (T_0) oscilan entre 253 y 313 K. Para generar los datos de entrenamiento del algoritmo, utilizamos como ángulos cenitales de observación los ángulos gaussianos (0º, 11.7º, 26.1º, 40.3º, 53.7º y 65º) y como valores de entrada de LST: $T_0 - 6$ K, $T_0 - 2$ K, $T_0 + 1$ K, $T_0 + 3$ K, $T_0 + 5$ K, $T_0 + 8$ K y $T_0 + 12$ K siguiendo el análisis global realizado en (Galve et al., 2008). En total el conjunto de datos contenía 16044 casos diferentes para obtener los coeficientes del algoritmo.

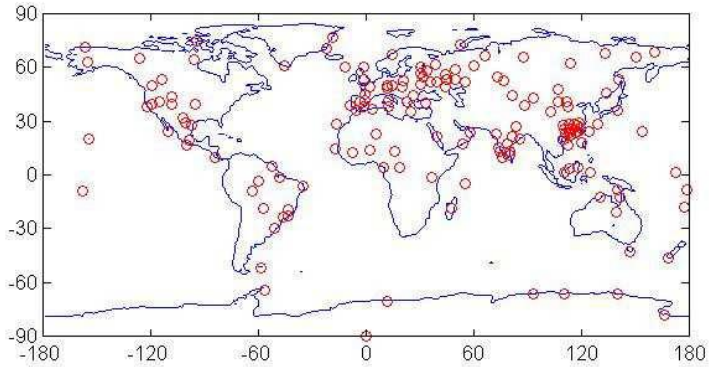

Figura 6. *Representación de las localizaciones de toma de radiosondeos de la base CLAR. Imagen extraída de Galve et al. (2008).*

e.1.3. Otros algoritmos SW propuestos

Otros autores también han propuesto variantes de algoritmos SW dependientes de la emisividad para el sensor SLSTR de Sentinel-3, como son Sobrino et al. (2016):

$$T = T_1 + d_1(T_1 - T_2) + d_2(T_1 - T_2)^2 + d_0 + (d_3 + d_4 W^*)(1 - \varepsilon) + (d_5 + d_6 W^*)\Delta\varepsilon \qquad (2.11)$$

donde los coeficientes d_0 a d_6 se obtienen mediante la simulación de una gran gama de temperaturas T_1 y T_2 con el código de transferencia radiativa MODTRAN 4.0 (Berk et al., 1999) para 61 perfiles atmosféricos seleccionados de la base de datos TIGR-1 (Thermodynamic Initial Guess Retrieval version 1) y 108 espectros de emisividad de la base de datos ASTER Spectral Library (Baldridge et al., 2009). La LST fue variada en 5 valores alrededor de la temperatura de la capa más baja de la atmósfera, y se tomó 5 ángulos de observación distintos.

Otro algoritmo SW es el propuesto por Zhang et al. (2019), que viene dado por

$$T = d_1 T_1 + d_2(T_1 - T_2) + d_3(T_1 - T_2)^2 + d_0(d_4 - d_5 W^*)(1 - \varepsilon) + (d_6 - d_7 W^*)\Delta\varepsilon \qquad (2.12)$$

donde los coeficientes d_0 a d_6 se obtuvieron de la simulación de 60 perfiles atmosféricos con el cielo despejado que se seleccionaron de la base de datos TIGR2000. Estos perfiles se utilizaron como entrada del programa MODTRAN5.2. La temperatura del nivel más bajo de la atmósfera se varia en 5 valores distintos. Los valores de emisividad media ε se varían de 0.9 a 1 en pasos de 0.02 y la diferencia entre bandas $\Delta\varepsilon$ entre -0.02 a 0.02 en pasos de 0.005, con un total de 30456 casos.

Por último, Zheng et al. (2019) propusieron el siguiente algoritmo, basado en el propuesto por Wan (2014) para el sensor MODIS:

$$T = d_0 + \left(d_1 + d_2\frac{1-\varepsilon}{\varepsilon} + d_3\frac{\Delta\varepsilon}{\varepsilon^2}\right)\frac{T_1+T_2}{2} + \left(d_4 + d_5\frac{1-\varepsilon}{\varepsilon} + d_6\frac{\Delta\varepsilon}{\varepsilon^2}\right)\frac{T_1-T_2}{2} + d_7(T_1-T_2)^2 \qquad (2.13)$$

donde los coeficientes d_0 a d_7 se obtienen de simulaciones realizadas con 946 perfiles atmosféricos de la base TIGR-3, utilizando como entrada una variación de rangos de LST, 60 espectros de emisividad de la librería espectral ASTER y la Universidad de California-Santa Bárbara y 5 ángulos de observación distintos que hacen un total de 2550200 casos diferentes.

e.2. Algoritmos para TIRS y TIRS-2

e.2.1. Producto operativo LST TIRS y TIRS-2

El producto operativo de LST de Landsat, tanto para el sensor TIRS como para el sensor TIRS-2, es un algoritmo SC (ecuación (2.4)) aplicado a la radiancia TOA de la banda 10 (alrededor de 11 µm). La emisividad que se utiliza se obtiene de una combinación del NDVI (Normalized Difference Vegetation Index) que se obtiene de las bandas del infrarrojo cercano y el rojo de OLI y la emisividad de ASTER Global Emissivity Dataset (GED), el cual es un archivo de datos estático de emisividades globales del año 2000 a 2008. Para obtener la radiancia ascendente, descendente y transmisividad se ejecuta MODTRAN y como entrada al código se le proporcionan los perfiles atmosféricos de los cuadrantes más cercanos de MERRA-2 (Modern-Era Retrospective analysis for Research and Aplications, Version 2), para posteriormente interpolar en tiempo, posición y en altura los resultados de la salida de MODTRAN con los datos de tiempo, posición y elevación de cada píxel.

e.2.2. Algoritmos propuestos para TIRS y TIRS-2

También se ha propuesto distintos algoritmos SW para los sensores TIRS y TIRS-2, como el de Jiménez-Muñoz et al. (2014) que coincide con la ecuación (2.11) con los coeficientes adaptados al sensor TIRS. Meng et al. (2019), Du et al. (2015) y Gerace et al. (2020) presentaron distintos algoritmos SW basados en el de Wan (2014) (ecuación (2.13)) pero con los coeficientes adaptados al sensor TIRS. Por otro lado, Meng et al. (2019) también propusieron otro algoritmo según

$$T = c_0 + c_1T_1 + c_2(T_1-T_2) + c_3\varepsilon + c_4\varepsilon(T_1-T_2) + c_5\Delta\varepsilon \qquad (2.14)$$

donde los coeficientes c_0 a c_5 vienen dados por Meng et al. (2019).

Yu et al. (2014) propusieron un algoritmo que se basa en evaluar los parámetros p_i mediante regresiones lineales con las T_i y queda de la siguiente forma

$$T = T_1 + B_1(T_1 - T_2) + B_0 \tag{2.15}$$

$$p_i = a_i + b_i T_i \tag{2.16}$$

$$B_0 = \frac{C_2(1-A_1-C_1)p_1 - C_1(1-A_2-C_2)p_2}{C_2A_1 - C_1A_2} \tag{2.17}$$

$$B_1 = \frac{C_1}{C_2A_1 - C_1A_2} \tag{2.18}$$

$$A_{11} = \varepsilon_1\tau_1 \tag{2.19}$$

$$A_{12} = \varepsilon_2\tau_2 \tag{2.20}$$

$$C_1 = (1-\tau_1)(1+(1-\varepsilon_1)\tau_1) \tag{2.21}$$

$$C_2 = (1-\tau_2)(1+(1-\varepsilon_2)\tau_2) \tag{2.22}$$

donde los coeficientes a_i y b_i vienen dados por Yu et al. (2014).

Rozenstein et al. (2014) propusieron un algoritmo similar al propuesto por Yu et al. (2014), pero sin utilizar los parámetros p_i y viene dado según

$$T = A_0 + A_1T_1 - A_2T_2 \tag{2.23}$$

$$A_0 = E_1b_1 - E_2b_2 \tag{2.24}$$

$$A_1 = 1 + A + E_1a_1 \tag{2.25}$$

$$A_2 = A + E_2a_2 \tag{2.26}$$

$$A = D_1/E_0 \tag{2.27}$$

$$E_1 = D_2(1 - C_1 - D_1)/E_0 \tag{2.28}$$

$$E_2 = D_1(1 - C_2 - D_2)/E_0 \tag{2.29}$$

$$E_0 = D_2C_1 - D_1C_2 \tag{2.30}$$

$$C_i = \varepsilon_i\tau_i(\theta) \tag{2.31}$$

$$D_i = [1-\tau_i(\theta)][1+(1-\varepsilon_i)\tau_i(\theta)] \tag{2.32}$$

donde los coeficientes a_i y b_i se dan para el rango de temperaturas de 10-40ºC y pueden obtenerse en Rozenstein et al. (2014).

Wang et al. (2023) propusieron el algoritmo solo para el sensor TIRS-2 y sigue la forma siguiente

$$T = \frac{c_2\lambda_1^{-1}}{LN\left(\frac{c_1\lambda_1^{-5}}{B_1(T)}+1\right)} \tag{2.33}$$

$$B_1(T) = A_0L_1 + A_1L_2 + A_2 \tag{2.34}$$

$$A_0 = \frac{D_2}{C_1 D_2 - C_2 D_1} \tag{2.35}$$

$$A_1 = \frac{-D_1}{k(C_1 D_2 - C_2 D_1)} \tag{2.36}$$

$$A_2 = \frac{bD_1(C_2 + D_2)}{k(C_1 D_2 - C_2 D_1)} \tag{2.37}$$

$$C_{1/2} = \varepsilon_{1/2}\tau_{1/2} \tag{2.38}$$

$$D_{1/2} = \left(1 - \tau_{1/2}\right)\left[\left(1 - \varepsilon_{1/2}\right)\tau_{1/2}\varphi_{1/2} + 1\right] \tag{2.39}$$

$$k = \frac{c_1^2 \lambda_1^{-4} \lambda_2^{-6} \left(c_1 \lambda_1^{-5} L_1^{-1} + 1\right)^{\left(\lambda_1 \lambda_2^{-1} - 1\right)}}{\left[\left(c_1 \lambda_1^{-5} L_1^{-1} + 1\right)^{\left(\lambda_1 \lambda_2^{-1}\right)} - 1\right]^2 L_1^2} \tag{2.40}$$

$$b = \frac{c_1 \lambda_2^{-5}}{\left(c_1 \lambda_1^{-5} L_1^{-1} + 1\right)^{\left(\lambda_1 \lambda_2^{-1}\right)} - 1} - kL_1 \tag{2.41}$$

$$\tau_{1/2} = a_0 W + a_1 \tag{2.42}$$

$$\varphi_{1/2} = a_2 LN(W) + a_3 \tag{2.43}$$

donde LN representa el logaritmo natural, $B_1(T)$ representa la radiancia de un cuerpo negro de la temperatura LST, L_1 y L_2 son las radiancias del sensor TIRS-2; λ_1 y λ_2 son las longitudes de onda efectivas para las bandas 1 y 2 dadas en la Tabla 5 y $\varepsilon_{1/2}$ son las emisividades de los canales 10 y 11 del sensor TIRS-2.

e.3. Algoritmos para MODIS

Para el sensor MODIS a bordo de Aqua y Terra se validan los productos M*D11 (MOD11 para Terra y MYD11 para Aqua) y M*D21 (MOD21 para Terra y MYD21 para Aqua).

e.3.1. Producto M*D11

En el caso del M*D11, el algoritmo SW que se utiliza es una adaptación del propuesto para MODIS por Wan y Dozier (1996) y adaptado por Wan (2014) que se muestra

$$T = A_0\left(A_1 + A_2\frac{1-\varepsilon}{\varepsilon} + A_3\frac{\Delta\varepsilon}{\varepsilon^2}\right)\frac{T_1+T_2}{2} + \left(B_1 + B_2\frac{1-\varepsilon}{\varepsilon} + B_3\frac{\Delta\varepsilon}{\varepsilon^2}\right)(T_1 - T_2) + C_0(T_1 - T_2)^2 \tag{2.44}$$

donde las bandas 1 y 2 se corresponden con la 31 y 32 de MODIS respectivamente, $\varepsilon = 0.5(\varepsilon_1 + \varepsilon_2)$ y $\Delta\varepsilon = \varepsilon_1 - \varepsilon_2$ y los coeficientes A_i se determinan a través de regresiones de perfiles atmosféricos donde se varía el ángulo de observación (con 9 ángulos diferentes), la temperatura de la superficie se varía en dos grupos distintos entre día y noche, el W se varía en saltos de 0.5 cm, la emisividad media entre 0.89 y 1 con pasos de 0.01 y las diferencias de emisividades entre -0.025 y 0.015 con pasos de 0.005 (Wan y Dozier, 1996).

Para el cálculo de las emisividades necesarias se desarrolló modelos matemáticos que toman las medidas de laboratorio de emisividades de diferentes especies, una clasificación de usos de suelo, la fenología de los distintos cultivos y las estaciones del año. A cada uno de los píxeles se le asigna un tipo de superficie que viene

dado en consonancia con la clasificación del producto M*D12 (Strahler et al., 1999) y se calcula la emisividad para cada banda asociada a cada tipo de superficie (Snyder et al., 1998).

e.3.2. Producto M*D21

El M*D21 utiliza el método TES, que se basa en la separación de la temperatura y la emisividad (Gillespie et al., 1998). El método necesita las radiancias de al menos 3 canales en la ventana atmosférica 8 - 12 µm, que en el caso de MODIS son las bandas 29 (8.5 µm), 31 (11 µm) y 32 (12 µm). Se realiza la corrección atmosférica de las mismas y se obtiene las radiancias a nivel de superficie que son las entradas que requiere el método. En el caso del M*D21 la corrección atmosférica se realiza con la corrección estándar propuesta por Palluconi et al. (1999) que requiere los perfiles atmosféricos de temperatura y humedad de la atmósfera que se le dan como entrada al modelo de transferencia radiativa MODTRAN (Berk et al., 1999). Los datos de los perfiles atmosféricos se obtienen de los datos de MERRA-2 que ofrece valores cada 6 horas, con 42 niveles de presión, a una resolución espacial de 0.5º × 0.625º y con una cobertura global. El método TES que sigue el producto M*D21 para desacoplar las emisividades de banda y las LSTs a partir de las radiancias en superficie se detalla en la sección (2.f.2) dedicada a la obtención de la emisividad.

e.4. Algoritmos para AVHRR/3

e.4.1. Producto operativo LST de LSA-SAF ELST

El producto de LST de LSA-SAF EDLST (*EPS EUMETSAT Polar System Daily LST*) utilizan el algoritmo GSW con una formulación similar a la propuesta por (Wan and Dozier, 1996) para MODIS:

$$T = \left(A_1 + A_2\frac{1-\varepsilon}{\varepsilon} + A_3\frac{\Delta\varepsilon}{\varepsilon^2}\right)\frac{T_1+T_2}{2} + \left(B_1 + B_2\frac{1-\varepsilon}{\varepsilon} + B_3\frac{\Delta\varepsilon}{\varepsilon^2}\right)\frac{T_1-T_2}{2} + C + \Delta T \quad (2.45)$$

donde las bandas 1 y 2 corresponden a las 4 y 5 de AVHRR/3, A_i, B_i y C son los coeficientes del algoritmo, y ΔT es el error del modelo. Para cada clase de vapor de agua y ángulos de observación del satélite existen coeficientes A_i, B_i y C específicos que hacen que se minimice el error del modelo.

e.4.2. Algoritmo SW propuesto

En este trabajo hemos propuesto un algoritmo SW propio para el sensor AVHRR/3 a bordo de MetOpA, B y C, que es una adaptación más simple al propuesto en la ecuación (2.10) ya que no tiene dependencia angular y sigue la forma

$$T = T_1 + c_1(T_1 - T_2) + c_2(T_1 - T_2)^2 + c_0 + (c_3 + c_4W + c_5W^2)(1-\varepsilon) + (c_6 + c_7W)\Delta\varepsilon \quad (2.46)$$

donde los coeficientes c_0 a c_7 vienen dados en la Tabla 2 y el valor de W se obtiene de los perfiles atmosféricos NCEP interpolados espacial y temporalmente.

Los coeficientes del algoritmo (Tabla 2) se obtuvieron mediante regresión sobre la base de datos de simulación de la sección 2.e.1.2, pero en este caso aplicando las funciones filtro de las bandas 4 y 5 de cada sensor AVHRR/3 de los satélites MetOp-A, -B y -C.

Tabla 2. Coeficientes para el algoritmo propuesto del sensor AVHRR/3.

Coeficiente	MetOp-A	MetOp-B	MetOp-C
$c_0\ (K)$	0.27 ± 0.025	0.32 ± 0.024	0.17 ± 0.031
c_1	1.25 ± 0.026	1.1 ± 0.024	0.46 ± 0.023
$c_2\ (K^{-1})$	0.45 ± 0.006	0.4 ± 0.005	0.32 ± 0.004
$c_3\ (K)$	49.2 ± 0.18	49.38 ± 0.18	52.9 ± 0.18
$c_4(K/cm)$	1.81 ± 0.12	1.68 ± 0.12	-0.61 ± 0.12
$c_5\ (K/cm^2)$	-1.2 ± 0.02	-1.18 ± 0.02	-0.93 ± 0.02
$c_6\ (K)$	98.27 ± 0.23	89.74 ± 0.21	61.75 ± 0.18
$c_7\ (K/cm)$	-15.21 ± 0.06	-13.71 ± 0.06	-8.91 ± 0.05

e.4.3. Otros algoritmos SW

El algoritmo propuesto por Jiménez-Muñoz y Sobrino (2008) sigue la forma de la ecuación (2.11) pero con los coeficientes particularizados para el sensor AVHRR/3 a bordo de MetOp-A. Para obtener los coeficientes se utilizó un total de 108 espectros de emisividad distintos extraídos de la librería espectral ASTER, con 61 perfiles de la base de datos TIGR como se describe en Sobrino et al. (1993). También utilizaron 5 ángulos de observación y 5 valores de temperatura distintos en las capas más bajas de la atmósfera. Todas estas combinaciones llevaban a un total de 164700 casos para obtener los coeficientes.

Parra et al. (2006) propusieron un algoritmo SW para el sensor AVHRR/3 a bordo de MetOp-A que sigue la forma

$$T = T_1 + (2 + 0.28W)(T_1 - T_2) - (0.4 - 0.48W) + (53 - 4W)(1 - \varepsilon) + (149 - 26W)\Delta\varepsilon \quad (2.47)$$

f. Métodos para obtener la emisividad de la superficie

El conocimiento de la emisividad es imprescindible para la medida de la LST in situ, la simulación de medidas de satélite y la aplicación de algoritmos LST. Por tanto, necesitamos obtener la emisividad de los distintos tipos de superficie en las bandas espectrales de los instrumentos y sensores estudiados, y en el ángulo de observación requerido.

En esta sección describimos las diferentes metodologías que se ha utilizado para obtener los valores de emisividad de las distintas superficies terrestres que hemos estudiado. Hemos utilizado métodos de medida de emisividad en el campo como el método de la caja (Rubio et al., 1997) y el TES (Gillespie et al., 1998). El

método de la caja puede aplicarse también a medidas de laboratorio principalmente para muestras de suelo recogidas en el campo. El método TES solo puede aplicarse en el campo, pero requiere de mucho menos tiempo en la toma de medidas. Por otro lado, ambos métodos solo pueden aplicarse sobre muestras homogéneas.

En el caso de superficies heterogéneas, con varias componentes como suelo y vegetación, la emisividad a escala de satélite la hemos obtenido utilizando el método Valor y Caselles (1996, 2005). En el caso de superficies de agua hemos utilizado el modelo propuesto por Niclòs et al. (2009) que obtiene la emisividad para cada banda en función del ángulo de observación.

f.1. Método de la caja

El método de la caja con dos tapas (Rubio et al., 1997) utiliza una caja con la que se aísla la muestra del entorno y de esta manera, conseguir que no tenga sobre la muestra efecto la radiación ambiental. Dispone de dos tapas intercambiables que se utilizan junto con la caja en las medidas.

La caja utilizada posee un cuerpo central de 80 cm de altura y una base de 30 × 30 cm. Las paredes de la caja están recubiertas en su interior por un aluminio pulido que se trata de un material que es muy reflectante con un valor de emisividad cercano a cero ($\varepsilon_c \sim 0.03$, Rubio et al., 1997). La caja no posee base inferior, que es donde se sitúa la muestra a medir. Todo el cuerpo de la caja está recubierto exteriormente por un material aislante (baja conductividad térmica) para de esta manera conseguir aumentar la estabilidad térmica de las paredes de la caja durante el procedimiento de la toma de medidas. Para cubrir la parte superior, la caja posee dos tapas, una fría y otra caliente. Ambas poseen un agujero central para colocar el radiómetro con el que se realizan las medidas. La tapa caliente tiene una respuesta espectral muy parecida a la del emisor perfecto $\varepsilon_h \sim 0.98$, está pintada de negro Parson y debe estar al menos 20ºC más caliente que la temperatura de la muestra, lo que se consigue mediante un sistema eléctrico. La otra tapa llamada fría está recubierta de aluminio al igual que el interior de la caja y con el mismo valor de emisividad. También se utiliza una tercera tapa fría, pero en este caso se utiliza como base y sin orificio central.

De acuerdo con el método de la caja en su versión de dos tapas, para obtener los valores de emisividad debe realizarse tres medias de radiancia en tres configuraciones distintas de la caja, más una cuarta que sirve para cuantificar y corregir el efecto de la caja no-ideal ($\varepsilon_c \neq 0$ y $\varepsilon_h \neq 1$). El orden de las medidas es el siguiente (Figura 7):

- Medida L^1 (tapa fría – muestra): con esta configuración y entendiendo que la caja es ideal la radiancia medida por el radiómetro es equivalente a la de un cuerpo negro a la temperatura de la muestra, de tal forma que se puede obtener esta temperatura aplicando la Ley de Planck.
- Medida L^2 (tapa caliente – muestra): en esta configuración el radiómetro mide la radiancia emitida por la muestra más la radiancia ambiental (la que aporta la tapa caliente) que se refleja en la muestra hacia el radiómetro.
- Medida L^3 (tapa caliente - base fría): en esta configuración, si se toma la caja como ideal, lo que mide el radiómetro es la radiancia de un cuerpo negro a la temperatura de la tapa caliente.

- Medida L^4 (tapa fría -base fría): en esta configuración se obtiene la contribución de la caja no-ideal, ya que si fuese ideal el valor de esta medida sería nulo.

Este orden no es arbitrario, si no que se realiza de esta manera específica para alcanzar un equilibrio térmico adecuado en cada una de las medidas, dejando pasar un tiempo establecido entre ellas.

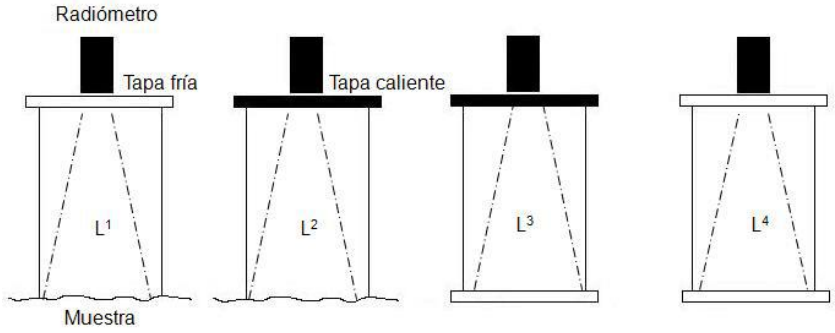

***Figura 7**. Procedimiento a seguir en el método de la caja en su versión de dos tapas. La secuencia que se debe seguir es de izquierda a derecha. Imagen extraída de Rubio et al. (1997).*

De este modo la expresión concreta de la emisividad de la muestra viene dada en Rubio et al. (1997) según la forma

$$\varepsilon = \frac{(L^2-L^1)(1-\varepsilon_c)}{(L^3-L^1)-(L^3-L^2)P+(L^1-L^4)Q} \tag{2.48}$$

donde $P = 0.146$ y $Q = 0.2921$ son factores que depende de la emisividad de la tapa fría y caliente, de la geometría de la caja y corrigen por tanto la no-idealidad de la caja, ya que en caso ideal ($\varepsilon_c = 0$ y $\varepsilon_h = 1$) tendríamos $P = Q = 0$.

f.2. Método TES para medidas de campo

El método TES fue propuesto por Gillespie et al. (1998) para los productos de LST y emisividad del sensor ASTER a bordo de la plataforma EOS Terra de la NASA. El método se aplica a radiancias a nivel de superficie por lo que también puede aplicarse con medidas de campo tomadas con un radiómetro multiespectral (Coll et al., 2019), obteniéndose las emisividades de la superficie y la LST a partir de al menos tres bandas en el IRT.

Para aplicar el método TES es necesario medir las radiancias a nivel de superficie ($L_{sup,i}$) de manera simultaneas en distintos canales. Esta variable está formada por dos contribuciones, la emisión propia de la superficie que se va a estudiar ($\varepsilon_i B_i(T)$) y la reflexión sobre dicha superficie de la radiancia de la atmósfera ($L_i^{\downarrow}$), por tanto, se puede expresar como

$$L_{sup,i} = \varepsilon_i B_i(T) + (1-\varepsilon_i)L_i^{\downarrow} \tag{2.49}$$

donde B_i es la función de Planck para la longitud de onda efectiva de los canales utilizados. La radiancia $L_i^{\downarrow}$ se mide simultáneamente a las radiancias de superficie mediante una placa difusora Infragold Reflectance Target (IRT-94-100) fabricada por Labsphere (Guillevic et al., 2018), la cual refleja en todas las direcciones la radiancia que procede del cielo. Las medidas de radiancia en superficie y radiancia atmosférica deben realizarse de manera casi simultánea (< 1 min).

Inicialmente, el algoritmo TES utiliza un módulo de normalización de la emisividad (NEM) (Gillespie, 1986), donde se escoge un valor de emisividad inicial para todos los canales del radiómetro, que en este caso fue $\varepsilon = 0.98$ (valor típico del cuerpo gris) y una estimación del valor de la radiancia atmosférica descendente. Con esto se obtiene una temperatura de la superficie para cada canal, entre las cuales se elige el valor máximo como una primera aproximación de la LST (T_{NEM}). Con este T_{NEM} se calculan las emisividades de cada canal (ε_i) a partir de la ecuación (2.49). Estos valores se utilizan en el módulo del cociente (ratio module (Schmugge et al., 1998)) para poder calcular los índices independientes de temperatura β_i o espectro β con

$$\beta_i = \frac{L_{sup,i}}{B_i(T_{NEM})}\frac{\bar{B}}{\bar{L}} \tag{2.50}$$

donde $\bar{B}$ y $\bar{L}$ son el valor medio de $B_i(T_{NEM})$ y $L_{sup,i}$ respectivamente para todos los canales (5 en nuestro caso). El siguiente paso es el cálculo de la diferencia entre el máximo y mínimo de los valores β, correspondientes al contraste espectral (MMD) según

$$MMD = max(\beta_i) - \min(\beta_i) \tag{2.51}$$

El valor MMD está relacionado con el valor mínimo de la emisividad (ε_{min}) de acuerdo con una relación empírica, que es la que se encarga de aportar la información que se requiere para romper con la indeterminación en la medida simultanea de LST y emisividad, ya que para N bandas tendríamos N+1 incógnitas (N emisividades más la LST) y tan solo N medidas. La relación empírica $MMD - \varepsilon_{min}$ se obtiene en base a medidas de laboratorio de espectros de emisividad de suelos, vegetación, nieve, rocas y agua de la librería ASTER (Baldridge et al., 2009), y viene expresada para los canales de ASTER por

$$\varepsilon_{min} = 0.994 - 0.687MMD^{0.737} \tag{2.52}$$

En este trabajo, hemos utilizado los radiómetros multibanda CIMEL312-2 descritos en la sección (3.b.2) cuyas bandas son muy similares a las de ASTER. Por tanto, hemos utilizado la ecuación (2.52) sin necesidad de ser modificada para obtener la emisividad en las 5 bandas estrechas de dicho radiómetro. Por otra parte, los canales 2 y 3 del CIMEL 312-2 son muy similares a las bandas SW de los sensores satelitales estudiados en esta tesis.

Con estos resultados de las estimaciones de emisividad, a partir de la ecuación (2.49) e invirtiendo con la ecuación de Planck se calcula la LST para cada uno de los canales. Si las estimaciones de emisividad son correctas, las LST de cada uno de los canales deberían ser iguales. Sin embargo, en la práctica presentan diferencias, por lo que se lleva a cabo un proceso iterativo utilizando las emisividades resultantes en el módulo

del cociente (β_i), para así reducir las diferencias entre las emisividades y mejorar la corrección de la radiancia atmosférica descendente. Cuando se logra que las diferencias entre las LST de cada canal sean inferiores a un cierto umbral (< 0.1K) se da por terminado el proceso y se toman dichas emisividades para cada canal y LST como válidas.

f.3. Método de cobertura vegetal

En el caso de tener superficies heterogéneas que están conformadas por varios componentes, como son las zonas arboladas o con cobertura parcial de la vegetación, la emisividad se obtiene siguiendo el método de Valor y Caselles (1996)

$$\varepsilon_i = \varepsilon_{v,i} f + \varepsilon_{s,i}(1-f) + 4 < d\varepsilon >_i f(1-f) \tag{2.53}$$

donde $\varepsilon_{v,i}$ es la emisividad de la vegetación, $\varepsilon_{s,i}$ es la emisividad del suelo, $< d\varepsilon >_i$ es el término de cavidad y f es la proporción de vegetación o fracción de cobertura vegetal que hay en el píxel.

Posteriormente los mismos autores ofrecieron una versión simplificada de la ecuación (2.53) únicamente dependiente de la emisividad del suelo, la vegetación y la fracción de cobertura vegetal del píxel (Valor y Caselles, 2005), que se muestra de la siguiente forma

$$\varepsilon_i = \varepsilon_{v,i} f + \varepsilon_{s,i}(1-f)(1-1.74f) + 1.7372 f(1-f) \tag{2.54}$$

En esta tesis, la ecuación (2.54) se ha utilizado para el cálculo de las emisividades de los bosques y zonas arboladas para los sensores SLSTR, MODIS y AVHRR/3. La fracción de cobertura vegetal se extrae de los datos que ofrece el producto de SLSTR. De esta manera, la emisividad se calcula de forma diaria para cada sensor.

f.4. Ecuación paramétrica de la emisividad del agua

La emisividad del agua es ampliamente conocida, sin embargo, es importante tener en cuenta su disminución cuando aumenta el ángulo de observación (Niclòs et al., 2005, 2007). En esta tesis, la emisividad de las superficies de agua se ha estimado mediante la ecuación paramétrica que propuso Niclòs et al. (2009). La ecuación proporciona la emisividad en la banda i $\varepsilon_i(\theta)$ en función del ángulo cenital de observación a partir de emisividad en el nadir $\varepsilon_i(0°)$ y la velocidad del viento U, de acuerdo con

$$\varepsilon_i(\theta) = \varepsilon_i(0°)\cos(\theta^{c\cdot U+d})^{b_i} \tag{2.55}$$

donde $c = -0.037$, $d = 2.36$ y b_i es una constante para cada banda espectral. Los valores de $\varepsilon_i(0°)$ y b_i son proporcionados para cada banda IRT de interés por Niclós et al. (2009). En esta tesis, al no disponer de medidas simultaneas de la velocidad del viento se tomó como constante igual a 5 m/s, tal y como propone el propio trabajo.

g. Análisis de sensibilidad

Hemos realizado un análisis de sensibilidad para determinar la incertidumbre de las LSTs medidas en el campo, que se toman como referencia en el método t-based, y de las LSTs calculadas a partir de datos de satélite que se toman como referencia en el método r-based. También hemos realizado el análisis de sensibilidad de las BT simuladas a partir de los datos de campo.

Para proceder con el cálculo de las incertidumbres en las LSTs obtenidas a partir de las medidas radiométricas de campo (ecuación 2.5), tenemos en cuenta tres fuentes de error que se suman cuadráticamente, tal y como se detalla a continuación (Coll et al., 2005, 2019):

1) La incertidumbre de calibración de los radiómetros terrestres, que para el CIMEL 312-2 se estima entre 0.12 K y 0.15 K (Coll et al., 2019). También hemos utilizado radiómetros de campo Apogee SI-121 y Heitronics KT15.85 IIP (ver secciones 3.c.3 y 3.c.4) cuyos errores de calibración según los fabricantes son 0.2 K y 0.3 K respectivamente.

2) La variabilidad espacial y temporal en la zona de medidas. En el caso de medidas tomadas en transectos, se toma el valor de la SD de todas las comprendidas en un intervalo de 6 minutos alrededor del paso del satélite. Para las estaciones fijas, se toma igualmente la SD durante los 6 minutos, aunque para este caso solo tenemos en cuenta la variabilidad temporal.

3) Las incertidumbres resultantes de la propagación de las variables de entrada para el cálculo de las LSTs, es decir, las incertidumbres procedentes de la emisividad y la radiancia atmosférica descendente. Este punto es el que produce mayor incertidumbre en la LST, con valores entre 0.4 K a 0.6 K debido en gran parte al error de emisividad (Coll et al., 2019).

Los resultados que obtuvimos en términos de incertidumbre para las LST de campo varían entre 0.6 y 0.9 K.

Para estimar las distintas componentes de la incertidumbre en la LST obtenida desde satélite en el método SC y el r-based (ecuación 2.4), hemos considerado cuatro fuentes de incertidumbre:

1) Las incertidumbres en las simulaciones de transferencia radiativa incluyendo posibles errores en los perfiles atmosféricos y en el modelo de transferencia radiativa utilizado (Coll et al., 2012; Hook et al., 2004). Estas obtienen valores de 0.6 K y 1.1 K para las bandas 1 (~11 μm) y 2 (~12 μm) respectivamente, para variaciones del 10% de humedad relativa y de 1 K en la temperatura en cada nivel del perfil atmosférico.

2) La incertidumbre en los valores de la emisividad medida que se propagan con la ecuación (2.4).

3) La incertidumbre en las BT que contribuyen al error en las LSTs. Estas incertidumbres se tomaron como el equivalente al ruido (NEΔT) específico para las bandas de cada sensor. Para el sensor SLSTR el NEΔT es de 0.05 K para ambas bandas y una temperatura de 270 K (Donlon et al., 2012). Para las dos bandas del sensor TIRS este se encuentra en torno a los 0.4 K a 300 K (Du et al., 2015; Irons et al., 2012; Meng et al., 2019). En el caso del sensor TIRS-2 se consideró un error < 0.1 K para ambas bandas (Barsi

et al., 2022; Pearlman et al., 2022). Para el sensor MODIS el NEΔT para ambas bandas es de 0.05 K (Wan, 1999). Y para el sensor AVHRR/3 el NEΔT para ambas bandas es de 0.12 K (Chang and Wu, 2012). En el caso de los sensores TIRS y TIRS-2 se consideró el valor medio de LST de una matriz de 17×17 píxeles remuestreados a 30 m (equivalentes a 5×5 píxeles nativos de 100 m) lo que abarca un área de estudio de unos 500×500 m^2. Este tamaño de matriz se seleccionó como compromiso entre disponer de un número suficiente de píxeles dentro de la zona homogénea de validación. Dado que el NEΔT es puramente aleatorio, la contribución en incertidumbre a la LST promediada debe de dividirse por $\sqrt{N}$ (Ghent et al., 2019), siendo N = 17×17 el número de muestras. Las primeras dos fuentes de error no son aleatorias, sino que, están correlacionadas por lo que no se deben de dividir por N.

4) Por último, la variabilidad espacial de las LSTs en las matrices de píxeles seleccionadas para la validación (2×2 en sensores de baja resolución espacial o 17×17 en sensores de alta resolución espacial). Dicha variabilidad se obtiene como la SD de los valores de LST de las matrices consideradas.

Las cuatro contribuciones de error se suman de manera cuadrática para obtener las incertidumbres totales de las LST. Estas varían en el intervalo de 0.6 a 1.5 K dependiendo del algoritmo y el sensor al que se le aplica.

Para el análisis de la sensibilidad de las BT simuladas se consideró el impacto a través de la ecuación (2.1) de las fuentes de error 1, 2 y 4 que se explican en el análisis de incertidumbres anterior, que se suman cuadráticamente para obtener la incertidumbre total en la BT simulada. Los resultados que obtuvimos en términos de incertidumbre para las temperaturas simuladas en la banda de 11 μm varían entre 0.6 a 1.0 K y para la banda de 12 μm entre 0.9 a 1.3 K para todos los sensores.

Capítulo 3

ÁREAS DE VALIDACIÓN, INSTRUMENTAL, DATOS DE SATÉLITE Y PERFILES ATMOSFÉRICOS

En esta sección se detallan las zonas donde se han realizado las validaciones LST, los instrumentos radiométricos utilizados para derivar la LST y los datos satelitales para poder obtener la LST desde satélite.

a. Áreas de validación para el método t-based

En estas áreas las temperaturas de referencia se obtienen mediante radiómetros que o bien están instalados en estaciones fijas o bien se desplazan a lo largo de transectos en la zona. A continuación, describimos las zonas de validación elegidas en esta tesis.

a.1. Arrozales de Valencia

Se trata de una extensa zona de cultivos de arroz en la marjal de la Albufera de Valencia, que abarca alrededor de unos 100 km^2. Debido a las características fenológicas del arroz este cultivo pasa por diversas coberturas. Desde mediados de junio hasta mediados de septiembre se encuentra la cobertura vegetal completa, entre febrero y mayo es la época de suelo desnudo, sin vegetación, que pasa desde tierra húmeda hasta que se seca, y de noviembre a enero tiene una cobertura de suelo inundado. Estas características hacen que en una misma zona existan diferentes coberturas de suelo, por lo que resulta una zona muy interesante para comprobar cómo se comportan los diferentes algoritmos propuestos sobre diferentes coberturas.

Además, la cubierta vegetal completa como el suelo inundado hacen que la homogeneidad térmica sea alta y propicie una buena comparación entre datos de satélite y datos tomados en tierra. Tanto los canales de riego como los caminos de acceso a los campos tienen un tamaño despreciable comparados con el tamaño típico de píxel de los sensores térmicos (1 km), por lo que no se rompe la homogeneidad dentro de la zona. Para los sensores de alta resolución se seleccionan parcelas grandes y no se cubren canales ni caminos. Las características bien conocidas de la emisividad en la cubierta vegetal completa en el rango espectral de 8-14 µm, alta emisividad y baja variación espectral (Rubio et al., 2003; Salisbury and D'Aria, 1992), facilitan las medidas radiométricas de temperatura de la superficie.

En la zona realizamos medidas de LST a lo largo de transectos con varios radiómetros CIMEL 312-2 (sección 3.c.2). La LST de referencia se toma como el valor medio de las medidas de todos los radiómetros disponibles durante 3 minutos antes y después del paso del sensor satelital. Este criterio se ha tomado para todos los casos en los que se utilizan medidas in situ de LST. En la zona también se encuentra instalada una estación fija de medidas continuas que incorpora dos radiómetros Apogee SI-121 (sección 3.c.3), que toman medidas cada 4 segundos. Uno de ellos apunta a la superficie con un ángulo cenital de visión de 10º con respecto al nadir y el otro apuntando al cielo con un ángulo de 53º para medir la radiancia descendente del cielo (Kondratyev, 1969).

Desde el año 2002 se lleva realizando campañas anuales de medida sobre la zona de los arrozales, con campañas intensivas sobre los tres tipos de cubiertas desde los últimos 8 años. Dichas medidas han sido utilizadas para calibrar y validar sensores IRT (TIRS, TIRS-2, SLSTR y MODIS) que han dado como fruto

diversos trabajos a lo largo de esta tesis (Coll et al., 2024; Niclòs et al., 2021, 2023; Pérez-Planells et al., 2021; Puchades et al., 2025).

En la Figura 8 se puede observar la zona de validación de los arrozales de Valencia. La localización de las medidas se encuentra centrada en la zona de los arrozales por lo que garantiza la homogeneidad de la superficie.

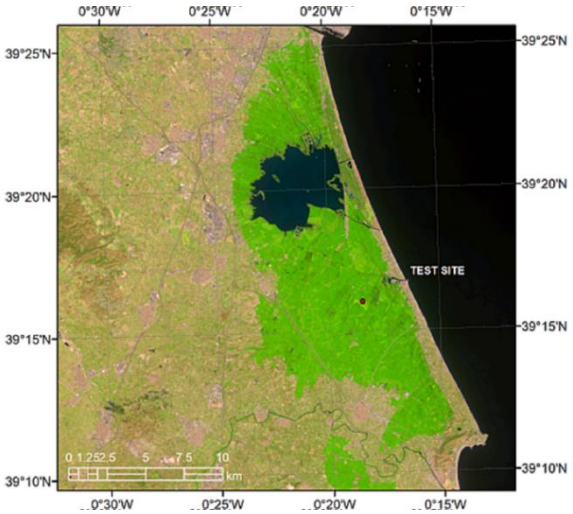

Figura 8. *Zona de los arrozales de Valencia donde se realizan los transectos y se encuentra la estación de medidas LST. Imagen de Niclòs et al. (2023).*

a.2. Matorral de Cortes de Pallás

Esta zona se encuentra en una planicie de matorral mediterráneo con una extensión de unos 200 km^2 al oeste de Valencia. Esta zona está dominada por la vegetación mediterránea que es su gran mayoría es romero y aliaga. La fracción de cobertura vegetal en la zona para un ángulo de visión de unos 20º respecto al nadir es de 0.6 (Niclòs et al., 2011). Se trata de una zona adecuada tanto para los satélites de baja resolución espacial como para los de alta resolución espacial. En la Figura 9 se puede observar una visión de la zona de interés sobre la que se toman las medidas.

Figura 9. *Zona de validación de Cortes de Pallás donde puede verse a la izquierda la zona de toma de medidas y a la derecha un detalle del matorral. (Niclòs et al., 2011).*

En esta zona se posee una estación de medidas similar a la que se encuentra en la zona de los arrozales de Valencia, equipada con radiómetros Apogee SI-121. En el caso de Cortes son tres los radiómetros que miden la temperatura del suelo con diferentes ángulos cenitales de visión (11º, 23º y 46º con respecto al nadir) sobre la misma superficie y otro instalado a un ángulo cenital de 53º (radiancia descendente del cielo). De esta manera se obtiene la temperatura de brillo para diferentes ángulos de visión y se puede comparar con los más próximos a los de los sensores satelitales.

a.3. Desierto de Gobabeb

Se trata de una zona desértica situada en Namibia con una gran homogeneidad térmica. Es un lugar con un clima desértico hiper árido y muy estable en el espacio y el tiempo. Es una estación que forma parte de la red que posee el Karlsruhe Institute of Technology (KIT). La estación de medidas tiene instalados dos radiómetros Heitronics KT15.85 IIP (sección 3.c.4), uno de los radiómetros midiendo al cielo con un ángulo cenital de 53º para medir la radiancia atmosférica descendente y otro mirando a la superficie con un ángulo inferior a los 10º. Ambos radiómetros están configurados para tomar medidas cada minuto. Esta estación también ha sido utilizada en trabajos previos (Göttsche et al., 2016; Martin et al., 2019).

a.4. Lago Constanza

Es otra estación que pertenece a la red del KIT, en el lago Constanza que se sitúa como frontera de los países Alemania, Suiza y Austria. Los radiómetros están instalados sobre un barco que atraviesa el lago 16 veces por día entre Friedrichsafen, Alemania, y Romanshorn, Suiza, y miden la temperatura de la superficie del lago. Los radiómetros instalados son los Heitroniks KT15.85 IIP (los mismo que en Gobabeb), pero el radiómetro que se encarga de obtener la radiancia del cielo se encuentra instalado de manera que mide con un ángulo complementario al de las medidas del agua, debido a la reflexión especular de la radiancia descendente (Niclòs et al., 2007). En esta zona de validación se toman medidas cada 15 segundos y se promedian a lo largo de 1 km cuando el barco se encuentra en el centro del lago.

a.5. Estaciones de la red LAW

Para ampliar el número de zonas de validación con diferentes tipos de superficie se ha utilizado medidas de LST en otras estaciones que pertenecen al proyecto LAW (Copernicus Space Component Validation for Land Surface Temperature, Aerosol Optical Depth and Water Vapor Sentinel-3 Products) (https://law.acri-st.fr/home). Todas ellas utilizan radiómetros Heitronics KT15.85 IIP instalados en torres meteorológicas con la misma configuración experimental que la zona de validación de Gobabeb. Las zonas que son consideradas "supersites" (zonas de validación con instrumentos para obtener todo tipo de magnitudes meteorológicas, climatológicas y atmosféricas) son:

a.5.1. Svartberget (SVA). Es una estación de medidas que se encuentra en Suecia. Actualmente la gestiona la Universidad Sueca de Agricultura. El tipo de bioma del lugar es una mezcla de bosque caducifolio con árboles de hoja perenne en forma de aguja.

a.5.2. Hyytiälä (HYY). Se encuentra instalada en Finlandia, el tipo de bioma que se le atribuye es un tipo de bosque mixto latifoliado cerrado.

a.5.3. Robson Creek (ROB). Está situada en Australia. El bioma que se le asigna es bosque latifoliado perennifolio y semicaducifolio cerrado.

a.5.4. KIT forest (KIT). Está instalada en el campus norte del KIT a unos 10 km al norte de la ciudad de Karlsruhe, Alemania. Los radiómetros están instalados en la parte superior de un mástil meteorológico de 200 m y apuntan directamente a un bioma de bosque caducifolio latifoliado cerrado.

a.5.5. Puéchabon (PUE) es una estación situada a unos 45 km de Montpellier, Francia. Está situada en una zona de bosques, arbustos y praderas y el tipo de bioma que se le asigna es de vegetación escasa. La zona en la que se encuentra la torre de 11 m de altura está rodeada por un dosel de bosque de unos 6 m de altura.

b. Áreas de validación para el método r-based

Para aumentar el número de biomas estudiados, hemos añadido cinco zonas de validación r-based adicionales que se encuentran en España. Estas se suman a las zonas descritas en el apartado anterior, a las que también se les ha aplicado dicho método.

b.1. Bosque de Soria. Se trata de alrededor de 200 km^2 de bosque situado cerca de Soria con un bioma asignado por el producto de Sentinel-3 SLSTR de bosque de coníferas.

b.2. Lago de la Albufera. Es una gran masa de agua dulce situada al sur de Valencia. El lago de la Albufera está clasificado como bioma de agua.

b.3. Viñedo. Se trata de una gran extensión de viñedo cerca de Requena. La cubierta varía a lo largo del año desde el brote de las hojas en abril hasta la caída de estas a mediados de octubre, permaneciendo el resto del año sin hojas. El tipo de bioma para este emplazamiento es cultivo de secano.

b.4. Salinas. Es una zona situada en Torrevieja dedicada a la producción de sal marina mediante la inundación de parcelas con agua salada. El tipo de bioma que se le asigna es agua.

b.5. Olivar. Es una gran extensión de terreno dedicada al olivar en la zona de Jaén. Es un cultivo de secano, formado por hileras de olivos separadas por suelo desnudo. El bioma que se le asigna a este olivar es de vegetación escasa.

En la Tabla 3 se puede observar la localización de cada zona de estudio y si pertenece a las estaciones de validación t-based o r-based. También se muestra la elevación del terreno y la clasificación del bioma según el producto de SLSTR de Sentinel-3.

Tabla 3. Localización de las zonas de validación t-based y r-based con el bioma asignado por el producto LST de SLSTR.

Zona	Latitud	Longitud	Elevación (m)	t-based	r-based	Bioma (SLSTR)
Arrozales de Valencia	39.2729º N	0.3185º O	1	✓	✓	1 (irrigated cropland)
Matorral de Cortes de Pallás	39.2242º N	0.8969º O	800	✓	✓	6/13 (closed broadleaved deciduous forest/ shrubland)
Desierto de Gobabeb	23.5200º S	15.0832º E	426	✓	✓	20 (bare area unknown)
Lake Constance	47.6050º N	9.4440º E	396	✓	✓	26 (water)
Svartberget	64.1708º N	19.7472º E	160	✓	✓	9 (open needleleaved forest)
Hyytiälä	61.8463º N	24.2956º E	170	✓	✓	10 (mixed forest)
Robson Creek	17.1175º S	145.6301º E	710	✓	✓	5 (broadleaved evergreen forest)
KIT forest	49.0760º N	8.4310º E	130	✓	✓	6 (closed broadleaved deciduous forest)
Puéchabon	43.7369º N	3.6008º E	500	✓	✓	15 (sparse vegetation)
Bosque de Soria	41.5912º N	2.5508º O	1120	-	✓	8 (closed needleleaved forest)
Lago de la Albufera	39.3388º N	0.3500º O	1	-	✓	26 (water)
Viñedo	39.5800º N	1.2700º O	770	-	✓	2 (rainfed cropland)
Salinas	37.9966º N	0.7262º O	1	-	✓	26 (water)
Olivar	37.8251º N	3.9680º O	481	-	✓	15 (sparse vegetation)

c. Instrumental

c.1. Fuente de calibrado Landcal P80P

Para asegurar una buena precisión y exactitud de las medidas in situ de LST es necesario calibrar a lo largo del tiempo los radiómetros IRT utilizados frente a una referencia trazable al sistema internacional (SI) con la que establecer la incertidumbre absoluta de los instrumentos de campo. En nuestro caso, realizamos calibraciones anuales de los radiómetros utilizando el cuerpo negro de referencia de temperatura variable Landcal P80P (Figura 10), que está a su vez calibrado frente a un radiómetro de referencia trazable al SI en el National Physical Laboratory (NPL) del Reino Unido. La calibración consiste en la comparación de las medidas de temperatura obtenida mediante los radiómetros con la temperatura del cuerpo negro.

Figura 10. *Fuente de calibración de la casa Landcal modelo P80P.*

Este modelo de cuerpo negro puede variar su temperatura desde -10 ºC hasta 80 ºC. Las especificaciones que ofrece son un tiempo de enfriamiento desde temperatura ambiente (20 ºC) a -10 ºC en un periodo de 90

minutos y de temperatura ambiente a una temperatura de 75 ºC en un periodo de 60 minutos. Con la fuente a una temperatura fija es capaz de mantener la temperatura en ±0.05 ºC durante un periodo de 30 minutos.

Este instrumento fue calibrado en diversas campañas internacionales de intercomparación e intercalibración de radiómetros y fuentes de calibrado en el NPL organizados por ESA y el comité CEOS. La calibración se realizó por primera vez en 2009 y se midió tanto la reproducibilidad de la medida (promedio de la desviación estándar de las medidas) y la repetitividad de la medida (promedio de las diferencias entre las medidas equivalentes de las series) frente al radiómetro de referencia del NPL, que son necesarias para conocer la precisión de los instrumentos (Theocharous et al., 2010). En este caso la fuente de calibrado ofreció un error de ± 0.15 K en el rango de temperaturas de 20-30 ºC (Theocharous et al., 2010). En 2016 volvió a ser calibrada, corrigiendo esta vez el efecto del valor real de la emisividad del cuerpo negro (>0.995, pero no exactamente igual a 1). Se obtuvo una incertidumbre de 0.02 K respecto al radiómetro de referencia del NPL (que tiene una incertidumbre de 53 mK) en el rango de temperaturas de 0 a 50 ºC, con una desviación estándar de 0.04 K (Theocharous et al., 2017). Por último, se realizó la calibración en 2022 y presentó una precisión mejor que 0.1 K en el intervalo de temperaturas de interés (Yamada et al., 2023a, 2023b).

c.2. Radiómetro térmico multibanda CIMEL CE312-2

Para realizar las medidas de LST a lo largo de transectos utilizamos los radiómetros térmicos multibanda CE312-2 de la casa CIMEL Electronique (Brogniez et al., 2003; Legrand et al., 2000) (http://www.cimel.fr). Este modelo en concreto posee seis bandas en el IRT, de las cuales una es ancha (banda 1) y cinco son estrechas (bandas 2 a 6) (véase la Figura 11 donde se muestra sus funciones de respuesta espectral). Las longitudes de onda efectivas (λ_{ef}) y el intervalo de banda a la mitad del máximo ($\Delta\lambda$) se especifican en la Tabla 4. Las bandas espectrales del CIMEL 312-2 son muy similares a las que incorpora el sensor ASTER.

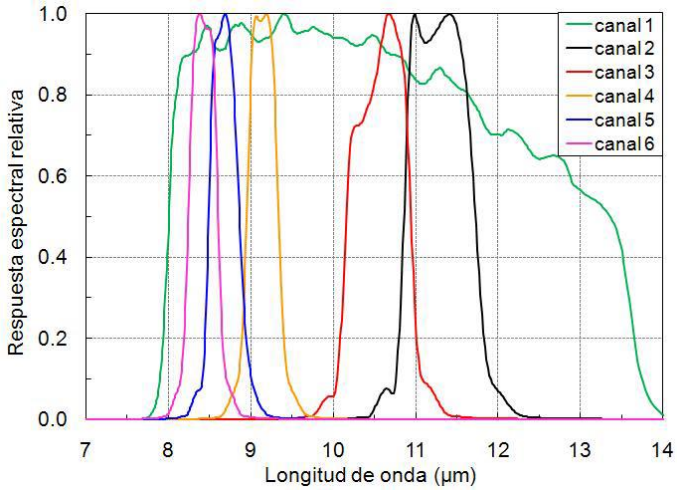

Figura 11. *Respuesta espectral de los canales radiométricos del CIMEL312-2.*

Tabla 4. Longitud de onda efectiva (λ_{ef}) e intervalo de banda a la mitad del máximo ($\Delta\lambda$) del radiómetro multiespectral CIMEL312-2 de la casa CIMEL Electronique. La banda 1 no ha sido utilizada en esta tesis.

CIMEL 312-2 / Canal	λ_{ef} (µm)	$\Delta\lambda$ (µm)
1	-	8.0 – 13.3
2	11.3	10.9 – 11.7
3	10.57	10.2 – 11.0
4	9.15	9.0 – 9.3
5	8.68	8.5 – 8.9
6	8.42	8.3 – 8.6

El instrumento está formado por dos partes, un cabezal óptico y la unidad electrónica. En la Figura 12 se puede ver una imagen de los dos componentes y los cables que los conectan. También es necesario un ordenador para almacenar en tiempo real los datos que se obtienen desde el cabezal óptico y poder configurar la selección de escenarios para la adquisición de los datos.

Figura 12. *Radiómetro térmico CIMEL 312-2 de la casa CIMEL Electronique con su cabezal óptico y su unidad electrónica y los cables para su conexión y carga.*

El cabezal está formado por el detector y la óptica. Tiene como detector una termopila delante de la cual está situada una ventana de Germanio. El cabezal óptico utiliza un arreglo de Kohler para iluminar el detector térmico. Una rueda de filtros con los seis filtros donde se encuentran las lentes y un motor paso a paso se encargan de la selección del filtro. Entre el condensador y el detector hay una ventana de sulfuro de zinc que se encarga de bloquear toda la radiación más allá de los 30 µm. Enfrente del objetivo existe un espejo ocultable recubierto de oro que permite comparar la radiación externa (radiación de la muestra) y la radiación interna (radiación de referencia) dentro del cabezal óptico que se encuentra a temperatura ambiente y por tanto contribuye a la radiancia que llega al detector. El instrumento mide la temperatura de la cavidad óptica mediante

una sonda de platino, con lo que se puede eliminar la radiancia interna del propio cabezal de la radiancia total en el detector. Este paso puede repetirse con el filtro seleccionado por el usuario las veces que se desee, de manera que, existe la posibilidad de realizar diferentes configuraciones de medida con un único canal o alternando los seis canales.

Por otro lado, está la unidad electrónica que es la responsable de recopilar, visualizar y almacenar los datos. El detector da la salida en voltajes en el intervalo de ±3 mV, que pasan por un amplificador con ganancia 1000 por lo que llega a ±3 V. Esta señal es transformada de voltaje a frecuencia (Hz/V), y a través de un microcontrolador se transforma esta frecuencia a cuantas digitales y se muestran por pantalla a la vez que se almacenan en memoria. El rango de cuentas digitales es de 0 a 60000, siendo 0 la radiancia máxima y 60000 la radiancia mínima detectable. La gestión de los filtros y el almacenaje se realiza mediante un software que proporciona la empresa CIMEL Electronique.

El CIMEL312-2 posee un campo de visión de 10º y un tiempo de respuesta de 1 segundo, con una sensibilidad de ±0.008 ºC a los 20 ºC para el canal 1 y de ±0.05 ºC para el resto (Brogniez et al., 2003; Legrand et al., 2000). El rango de temperaturas en que puede medir el radiómetro es desde -80 ºC a 50 ºC.

Estos radiómetros fueron calibrados en campañas internacionales de intercalibración e intercomparación de radiómetros en el experimento FRM4STS (Fiducial Reference Measurements for validation of Surface Temperature from Satellites) en el NPL. Los radiómetros fueron calibrados frente al cuerpo negro estándar del NPL con calibración trazable al SI y una incertidumbre absoluta de 0.04 K. En el año 2016, los dos radiómetros CIMEL 312-2 mostraron una incertidumbre total de 0.15 K teniendo en cuenta todas las fuentes de error (Coll et al., 2019; Theocharous et al., 2019). En el año 2022 volvieron a realizarse estas calibraciones y las incertidumbres obtenidas fueron de 0.2 K en el rango de 0 ºC a 50 ºC para todas las bandas de los dos radiómetros (Yamada et al., 2023a, 2023b).

Por otra parte, los radiómetros son calibrados anualmente en el laboratorio frente a la fuente de calibrado Landcal P80P mencionada anteriormente. Para ello consideramos el intervalo de temperaturas típicas que nos encontraremos en las zonas de validación de 0 ºC hasta 50 ºC con saltos de 5 ºC y en dos series diferentes (subida y bajada de las temperaturas) para comprobar la reproducibilidad de cada radiómetro. Los resultados obtenidos por los radiómetros CIMEL 312-2 frente a la fuente de calibrado Landcal P80P muestran un buen comportamiento a lo largo del tiempo siendo similares a los obtenidos en los experimentos de calibración FRM4STS en el NPL.

c.3. Apogee SI-121

Este instrumento mide la radiancia en la región espectral IRT (8-14 µm), con una longitud de onda efectiva de 11.05 µm, un campo de visión de 36º y una incertidumbre según el fabricante de 0.2 K (https://www.apogeeinstruments.com). En la Figura 13 se muestra la función filtro del Apogee SI-121.

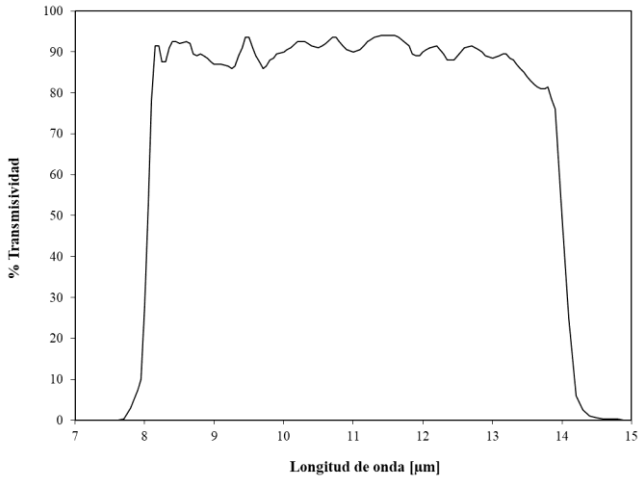

Figura 13. *Función filtro del Apogee SI-121.*

El tiempo de respuesta del instrumento es de 0.6 segundos, aunque en el caso de las estaciones que se utilizan en esta tesis están configurados para adquirir datos cada 4 segundos. Posee un blindaje resistente para reducir las fluctuaciones térmicas rápidas y los componentes electrónicos están completamente encapsulados. Además, trae un soporte que está diseñado para orientar los radiómetros en diferentes ángulos en función de las distintas aplicaciones, de tal forma que pueden montarse en un mástil de manera sencilla. El cable que lleva incorporado es altamente resistente al agua, estabilidad de rayos ultravioleta y posee flexibilidad en condiciones de frío, de esta manera la señal no se ve alterada por ningún agente externo.

Este instrumento fue calibrado en el año 2009 en la campaña de intercalibración e intercomparación internacional frente a fuentes de calibrado de referencia y frente a diferentes tipos de superficies (Theocharous et al., 2010). Además, se calibra de manera anual en el laboratorio frente a la fuente de calibrado Landcal P80P. Estos radiómetros muestran unas incertidumbres en las calibraciones de 0.1 K (Niclòs et al., 2015; Theocharous et al., 2010). Sin embargo, para ser más precavidos se decidió considerar el valor de incertidumbre que estima el fabricante de 0.2 K.

c.4. Heitronics KT15.85 IIP

Se trata de un tipo de instrumento que está diseñado para funcionar correctamente en ambientes muy diversos, desde desiertos y regiones polares hasta embarcaciones marinas. Posee carcasas de protección contra la intemperie para que no afecte la climatología exterior, ya que el rango de medidas en el que puede trabajar es de -20ºC a 60ºC, con una protección a agua y polvo de IP65. Este instrumento mide la radiancia en la región espectral IRT de 9.6-11.5 µm, con una longitud de onda efectiva de 10.55 µm, un campo de visión de 8.5º y una incertidumbre según el fabricante de 0.3 K (https://www.heitronics.com/). Las estaciones que utilizan estos radiómetros en esta tesis están configuradas para adquirir datos cada minuto y el modo de instalación es similar al de los radiómetros Apogee SI-121.

Este instrumento fue calibrado en los años 2009, 2016 y 2022 en las diferentes campañas de intercalibración e intercomparación internacionales frente a fuentes de calibrado de referencia (Barker-Snook et al., 2016, 2017; Theocharous et al., 2010, 2017, 2019; Yamada et al., 2023a, 2023b). En estas campañas de calibración la incertidumbre resultante fue menor que 0.3 K que indica el fabricante, aunque hemos tomado este valor como error de calibración del instrumento en el análisis de sensibilidad.

d. Imágenes Satelitales

Las características de los sensores y las imágenes satelitales utilizadas para realizar esta tesis se muestran en la Tabla 5. En dicha tabla podemos ver las resoluciones tanto espaciales como temporales, las bandas que posee en el IRT, sus longitudes de onda efectivas, la fecha de lanzamiento y el producto LST operativo de cada sensor.

La mayoría de los sensores estudiados tienen una resolución espacial de 1 km (SLSTR, MODIS y AVHRR/3) y una resolución temporal diaria, excepto los sensores de la seria Landsat (TIRS y TIRS-2) que poseen una resolución de 100 m reescalados a 30 m y una resolución temporal de 16 días. En las zonas de los arrozales de Valencia y Cortes los sensores TIRS y TIRS-2 poseen dos "path" diferentes sobre la zona de interés (path 198 y 199), por tanto, en realidad se tienen imágenes cada 8 días.

Tabla 5. Resumen de los diferentes sensores que se han estudiado en esta tesis, su resolución espacial y temporal, el número de bandas en la región 10.0 - 12.5 µm, sus longitudes de onda efectivas, su fecha de lanzamiento y su producto LST operativo.

Satélite	Sensor	Resolución espacial nominal	Resolución temporal	Nº bandas IRT	Banda 1 (µm)	Banda 2 (µm)	Fecha lanzamiento	Producto
Sentinel 3A	SLSTR	1 km	Diaria	2	10.85	12.00	Febrero 2016	LST SLSTR
Sentinel 3B	SLSTR	1 km	Diaria	2	10.81	12.02	Abril 2018	LST SLSTR
Landsat 8	TIRS	100 m	16 días	2	10.89	11.98	Febrero 2013	LST product
Landsat 9	TIRS-2	100 m	16 días	2	10.82	12..00	Septiembre 2021	LST product
Terra	MODIS	1 km	Diaria	3*	11.01	12.03	Diciembre 1999	MOD 11 y 21
Aqua	MODIS	1 km	Diaria	3*	11.02	12.04	Mayo 2002	MYD 11 y 21
MetOp A	AVHRR/3	1 km	Diaria	2	10.80	11.95	Octubre 2006	LSA SAF
MetOp B	AVHRR/3	1 km	Diaria	2	10.74	11.94	Septiembre 2012	LSA SAF
MetOp C	AVHRR/3	1 km	Diaria	2	10.75	12.07	Noviembre 2018	LSA SAF

*MODIS dispone también de una banda IRT en 8.52 µm.

e. Perfiles atmosféricos

Los perfiles atmosféricos de presión, temperatura y humedad atmosférica pueden obtenerse idealmente realizando radiosondeos locales con globos sonda equipados con el instrumental meteorológico necesario para obtener las medidas atmosféricas. El globo sonda envía la información de GPS, presión atmosférica, temperatura y humedad, entre otros, a diferentes alturas pudiendo llegar a 10 km. Los radiosondeos realizados

con globos sonda muestran una ventaja que es la gran coincidencia en espacio y tiempo al paso de los sensores satelitales, pero como contrapartida son costosos económicamente. Con los radiosondeos se obtiene un perfil casi continuo en altura, aunque para la utilización del modelo de transferencia radiativa MODTRAN 5.2 (Berk et al., 2008) se debe seleccionar algunos niveles para poder implementarlo.

Como alternativa a este tipo de radiosondeos existen los perfiles que se pueden obtener de imágenes de satélite de sensores sondeadores. Como ejemplo son los productos atmosféricos MOD07 y MYD07 de MODIS con el que se obtienen los perfiles de temperatura y humedad a 5 km de resolución espacial (Borbas et al., 2011; Seemann et al., 2006). Estos productos ofrecen 20 niveles de presión estándares y con una altura máxima de 35 km. Debido a que la disponibilidad y la periodicidad del paso de los satélites no es continua, es necesario utilizar otros métodos para obtener los perfiles.

Otra posibilidad para obtener los perfiles atmosféricos de una forma más continua tanto de manera espacial como temporal es utilizar los productos de análisis troposférico a escala global que proporciona el National Centers of Environmental Prediction (NCEP, Kalnay et al., 1996). NCEP ofrece perfiles atmosféricos con una resolución temporal de 6 horas y una resolución espacial de 1º en latitud y longitud. Los datos que proporciona son la temperatura del aire, la humedad relativa, y la presión en 26 niveles que llegan a una altura máxima de 31 km. Existen datos desde el año 1999, en dos versiones diferentes, pero para las variables que son necesarias en esta tesis no existe ninguna diferencia entre ellas. Los datos son de acceso libre y pueden ser descargados en el siguiente enlace: https://rda.ucar.edu/datasets/d083002/dataaccess/#.

f. Datos de emisividad

Los valores de emisividad necesarios para el cálculo de la LST de referencia en las zonas de validación t-based se muestran en la Tabla 6 para los radiómetros Apogee SI-121 (8-14 µm) y Heitronics KT15.85IIP (9.6-11.5 µm) en las zonas de validación t-based. En ella se muestra también las referencias en las que se han obtenido las medidas de dichos valores.

Tabla 6. Valores de emisividad en las estaciones de medida de LST in situ y su rango espectral.

Zona de estudio	Referencia	Rango Espectral	Emisividad
Arrozales (Tierra seca)	(García-Santos et al., 2014)	8 – 14 µm	0.967
Arrozales (Tierra húmeda)	(García-Santos et al., 2014)	8 – 14 µm	0.973
Arrozales (Vegetación)	(Niclòs et al., 2011)	8 – 14 µm	0.983
Arrozales (Agua)	(Niclòs et al., 2011)	8 – 14 µm	0.986
Matorral de Cortes de Pallás	(Niclòs et al., 2018)	8 – 14 µm	0.986
Lake Constance	(Thome et al., 1998)	9.6 – 11.5 µm	0.973
Gobabeb	(Göttsche and Hulley, 2012)	9.6 – 11.5 µm	0.940
KIT forest	(Hu et al., 2022)	9.6 – 11.5 µm	0.988
Hyytiälä	Copernicus LAW LST dataset	9.6 – 11.5 µm	0.991
Puéchabon	Copernicus LAW LST dataset	9.6 – 11.5 µm	0.989
Svartberget	Copernicus LAW LST dataset	9.6 – 11.5 µm	0.991
Robson Creek	Copernicus LAW LST dataset	9.6 – 11.5 µm	0.991

Por otra parte, en la Tabla 7 mostramos las emisividades de las superficies estudiadas en las bandas de SW de los sensores satelitales (banda 1 cercana a 11 µm y banda 2 cercana a 12 µm), tanto para las zonas de

validación t-based como para las r-based. Estos valores son necesarios para la simulación de BTs a partir de medidas in situ, la aplicación de los distintos algoritmos SW y el cálculo de la LST de referencia en el método r-based. En ella se muestran las referencias en las que se ha utilizado dichas emisividades y los métodos empleados para obtenerlas. En algunos casos se da un intervalo de valores debido a la variación de la vegetación a lo largo del año, y al ángulo de observación en las superficies agua. En las medidas in situ de LST realizadas en transectos en la zona de los arrozales de Valencia, utilizamos las bandas 2 y 3 de los radiómetros CIMEL 312-2 tomando las emisividades de la Tabla 7 para la corrección de las medidas radiométricas, dada la similitud con las bandas de los sensores satelitales.

Tabla 7. Emisividades para los canales térmicos de los sensores satelitales en las zonas de estudio, y con los métodos y referencias que se han utilizado para obtenerlas.

Zona de estudio	**Método/Referencia**	**Banda 1 (11 µm)**	**Banda 2 (12 µm)**
Arrozales (Tierra seca)	Método TES (Niclòs et al., 2023a)	0.970	0.972
Arrozales (Tierra húmeda)	Método TES (Niclòs et al., 2023a)	0.972	0.977
Arrozales (Vegetación)	Método de la caja (Niclòs et al., 2023a)	0.985	0.980
Arrozales (Agua)	(Niclòs et al., 2009)	0.988-0.992	0.985-0.988
Matorral de Cortes de Pallás	Método VCM (Valor and Caselles, 1996)	0.980-0.985	0.983-0.987
Lake Constance	(Niclòs et al., 2009)	0.991-0.992	0.986-0.988
Gobabeb	Espectroradiómetro (Göttsche, F.-M. et al., 2018)	0.946	0.972
KIT forest	Método VCM (Valor and Caselles, 1996)	0.980-0.986	0.982-0.986
Hyytiälä	Método VCM (Valor and Caselles, 1996)	0.980-0.986	0.980-0.988
Puéchabon	Método VCM (Valor and Caselles, 1996)	0.980-0.986	0.980-0.988
Svartberget	Método VCM (Valor and Caselles, 1996)	0.980-0.986	0.980-0.988
Robson Creek	Método VCM (Valor and Caselles, 1996)	0.980-0.986	0.980-0.988
Bosque de Soria	Método VCM (Valor and Caselles, 1996)	0.982-0.986	0.984-0.986
Albufera	(Niclòs et al., 2009)	0.990-0.992	0.986-0.988
Viñedo	Método VCM (Valor and Caselles, 1996)	0.969-0.976	0.961-0.972
Salinas	Método de la caja (Rubio et al., 1997)	0.991	0.990
Olivar	Método VCM (Valor and Caselles, 1996)	0.972-0.976	0.966-0.973

Capítulo 4

RESULTADOS

En esta sección presentamos los resultados que hemos obtenido para los distintos sensores estudiados. En el caso del sensor SLSTR, la calibración de las bandas térmicas, validación sobre diferentes superficies del producto operativo de LST además de otros SW propuestos, usando tanto el método t-based como el r-based. Se sigue con la calibración de los sensores TIRS y TIRS-2 y la validación de los productos operativos y otros algoritmos propuestos por diversos autores. Por otro lado, validamos los productos de MODIS (M*D11 y M*D21) mediante el método r-based, además de mostrar una comparación de los métodos t-based y r-based en la zona de los arrozales de Valencia. Por último, presentamos los resultados que se obtienen para el sensor AVHRR/3 mediante el método t-based con el producto operativo de LSA-SAF, el algoritmo SW propuesto en esta tesis y algoritmos de otros autores.

Para el caso de los sensores SLSTR, MODIS y AVHRR/3 (baja resolución espacial, 1 km), hemos utilizado una ventana de los 2×2 píxeles más cercanos a la zona de interés para comparar los datos de satélite con los datos tomados en campo. De esta manera se ponderan asignando un peso a cada uno de los píxeles igual a la inversa de la distancia a la estación de toma de datos o de los transectos realizados. De ahí la importancia de que la zona sea lo suficientemente homogénea, para que en esos 4 píxeles los valores de emisividad y temperatura sean similares. En el caso de los sensores TIRS y TIRS-2 (alta resolución espacial, 100 m reescalados a 30 m), hemos considerado la comparación tanto con un solo píxel de 30 m, como con una ventana de 17×17 píxeles (equivalente a 500 m×500 m) centrados en la zona de las medidas in situ. El motivo es realizar una comparación y observar cuál de las dos opciones ofrece mejores resultados.

En todos los casos, hemos calculado las diferencias entre las temperaturas derivadas de satélite (BTs y LSTs para los algoritmos) y las correspondientes temperaturas de referencia (BTs simuladas y LSTs medidas en el campo y obtenidas mediante el método r-based). Para cada caso, calculamos los promedios de las diferencias (bias) y las desviaciones estándar (SD) de las diferencias. La RMSD (root mean square difference) se calcula como la raíz de la suma cuadrática del bias y el SD. Por otro lado, hemos utilizado también estadística robusta (Wilrich, 2007) en la cual en vez del bias se utiliza la mediana y en vez de la SD, la desviación estándar robusta (RSD, robust standard deviation) que viene dada por

$$RSD = mediana\left|\left(T_{sat} - T_g\right)_k - mediana\left(T_{sat} - T_g\right)_k\right| 1.4826 \quad (4.1)$$

donde T_{sat} es la temperatura satelital, T_g es la temperatura de referencia y k es cada par de datos coincidentes. La R-RMSD (robust root mean square difference) se calcula como la raíz de la suma cuadrática de la mediana de las diferencias y la RSD.

a. Resultados para el sensor SLSTR

El producto operativo de LST de SLSTR fue inicialmente validado en Pérez-Planells et al. (2021), donde utilizamos datos de la estación de medida de los arrozales de Valencia tomados con un radiómetro fijo Apogee SI-121. Los resultados estadísticos obtenidos en dicho estudio se presentan en la Tabla 8 para el satélite Sentinel-3A y en la Tabla 9 para el Sentinel-3B, diferenciando los datos diurnos y nocturnos y el tipo de

superficie (agua, tierra y vegetación), además de los resultados globales (datos diurnos y nocturnos, y todas las superficies). En el momento del estudio no se tenía tantos datos de Sentinel-3B como del -3A, ya que fue lanzado dos años más tarde.

Para el Sentinel-3A, los resultados globales que obtuvimos muestran una mediana de 1.3 K, un RSD de 1.3 K y una R-RMSD de 1.8 K. Los resultados que obtuvimos para el satélite Sentinel-3B fueron una mediana de 1.5 K, un RSD de 1.2 K y una R-RMSD de 1.9 K para la combinación de todas las horas de paso y superficies. En general, observamos un mejor comportamiento para los datos nocturnos y que los peores resultados se dan en la superficie de suelo inundado con R-RMSD por encima de 2 K.

Tabla 8. Resultados estadísticos para el producto LST operativo de SLSTR de Sentinel-3A frente a las LST in situ en los tres tipos de superficies de los arrozales de Valencia. Todos los valores son en K y N es el número de puntos (Pérez-Planells et al., 2021).

	Todos los datos				**Día**				**Noche**			
	Mediana	RSD	R-RMSD	N	Mediana	RSD	R-RMSD	N	Mediana	RSD	R-RMSD	N
Todas las superficies	1.3	1.3	1.8	194	1.8	1.2	2.2	98	1.0	1.0	1.4	96
Suelo inundado	1.8	1.1	2.2	44	2.2	0.7	2.3	19	1.8	1.3	2.2	25
Tierra	1.1	0.7	1.3	37	1.3	0.9	1.6	16	0.8	0.6	1.0	21
Vegetación	1.3	1.4	1.9	113	1.7	1.5	2.2	63	1.0	0.9	1.3	50

Tabla 9. Resultados estadísticos para el producto LST operativo de SLSTR de Sentinel-3B frente a las LST in situ en los tres tipos de superficies, en los arrozales de Valencia. Todos los valores son en K y N es el número de puntos (Pérez-Planells et al., 2021).

	Todos los datos				**Día**				**Noche**			
	Mediana	RSD	R-RMSD	N	Mediana	RSD	R-RMSD	N	Mediana	RSD	R-RMSD	N
Todas las superficies	1.5	1.2	1.9	107	1.6	1.3	2.0	41	1.3	1.0	1.7	66
Suelo inundado	2.1	0.6	2.2	48	2.5	1.1	2.8	15	1.9	0.7	2.0	33
Tierra	0.8	1.1	1.3	31	0.8	1.7	1.8	13	0.8	1.0	1.3	18
Vegetación	1.0	1.3	1.6	28	1.4	0.9	1.7	13	0.5	1.2	1.3	15

Las mismas LSTs de referencia han sido utilizadas para validar distintos algoritmos SW (descritos en las secciones 2.e.1.2 y 2.e.1.3). Los resultados se muestran en la Tabla 10 únicamente para datos Sentinel-3A en términos de mediana, RSD y R-RMSD análogamente a las Tablas 8 y 9. Teniendo en cuenta todos los datos, los valores globales de las medianas oscilan entre -0.8 K para el algoritmo propuesto por Sobrino et al. (2016) y 0.4 K el propuesto por Zheng et al. (2019), mientras que las R-RMSD varían entre 1.1 K el de Pérez-Planells et al. (2021) y 1.3 K el de Zhang et al. (2019). Además, observamos que la comparativa entre los algoritmos muestra resultados similares para cualquier tipo de superficie en los cuatro algoritmos estudiados.

Tabla 10. Resultados estadísticos para los diferentes algoritmos SW. Los resultados se muestran para todas las superficies y separando los días de las noches, así como por superficies (Pérez-Planells et al., 2021). N es el número de puntos.

		Todos los datos				**Día**			**Noche**		
		Mediana	RSD	R-RMSD	N	Mediana	RSD	R-RMSD	Mediana	RSD	R-RMSD
Todas las superficies	Sobrino	-0.8	0.9	1.2	198	-0.8	1.2	1.5	-0.9	0.8	1.2
	Zhang	-0.7	1.1	1.3	198	-0.5	1.3	1.4	-0.9	0.8	1.2
	Zheng	0.4	1.1	1.2	198	0.5	1.4	1.5	0.2	1.1	1.1
	Propuesto	-0.4	1.1	1.1	198	-0.3	1.5	1.5	-0.5	0.8	0.9
Suelo inundado	Sobrino	-0.4	0.6	0.7	32	-0.3	0.7	0.7	-0.5	0.6	0.8
	Zhang	-0.5	1.0	1.1	32	-0.4	0.8	0.9	-0.9	0.9	1.2
	Zheng	1.0	0.7	1.2	32	1.1	0.6	1.3	0.9	0.9	1.3
	Propuesto	0.0	0.6	0.6	32	0.2	0.7	0.8	-0.1	0.7	0.7
Tierra	Sobrino	-0.4	0.9	0.9	38	-0.4	1.5	1.5	-0.4	0.8	0.9
	Zhang	-0.5	0.6	0.8	38	0.6	1.5	1.6	-0.6	0.7	1.0
	Zheng	0.9	0.7	1.2	38	0.8	1.4	1.6	1.0	0.5	1.1
	Propuesto	-0.2	0.9	0.9	38	0.2	1.5	1.5	-0.2	0.6	0.6
Vegetación	Sobrino	-1.0	1.0	1.4	128	-1.3	1.5	1.9	-1.0	0.7	1.2
	Zhang	-0.9	1.2	1.5	128	-0.7	1.5	1.7	-0.9	0.8	1.2
	Zheng	-0.1	1.3	1.3	128	0.0	1.4	1.4	-0.1	0.9	0.9
	Propuesto	-0.7	1.2	1.4	128	-0.8	1.6	1.8	-0.7	0.8	1.0

Posteriormente a la publicación del artículo de Pérez-Planells et al. (2021) incrementamos el número de datos en la misma zona y añadimos nuevas zonas de validación con distintos tipos de cobertura. Con estos datos, primeramente, evaluamos la calibración de las bandas 8 y 9 (alrededor de 11 µm y 12 µm, respectivamente) de los sensores SLSTR a partir de las medidas de campo y seguidamente realizamos una nueva validación (t-based y r-based) del producto LST operativo de SLSTR y el algoritmo SW propuesto por Pérez-Planells et al. (2021) (Puchades et al., 2025). A continuación, exponemos los resultados obtenidos en dicho artículo.

a.1. Calibración de las bandas TIR del sensor SLSTR

La evaluación de la calibración se ha realizado sobre las zonas de estudio de la Tabla 6, y tomando como referencia las temperaturas de brillo simuladas con la ecuación (2.1) a partir de las medidas de LST in situ, los valores de emisividad de la Tabla 7 y las magnitudes atmosféricas calculadas mediante los perfiles atmosféricos y MODTRAN 5.2. Las estaciones se han agrupado en cuatro clases distintas de acuerdo con el bioma de Sentinel-3 para cada zona de estudio: bosque (Cortes, KIT, HYY, PUE, SVA y ROB), cultivos de regadío (arrozales de Valencia), agua (Lago Constanza) y desierto (Gobabeb).

La Figura 14 muestra cómo se ajustan las temperaturas de brillo del satélite frente a las temperaturas de brillo simuladas en las bandas 8 y 9 de Sentinel-3A y -3B para los datos tanto diurnos como nocturnos. Se ha separado los datos en los cuatro biomas anteriormente citados. Se observa que la concordancia entre ambas temperaturas es buena para ambas bandas y ambos satélites y para todos los biomas con un amplio rango de temperaturas de estudio (250-310 K). El coeficiente de correlación entre las temperaturas simuladas y las

obtenidas desde satélite es alto y oscila entre 0.98 para el bosque y 0.999 para los cultivos de regadío en ambas bandas.

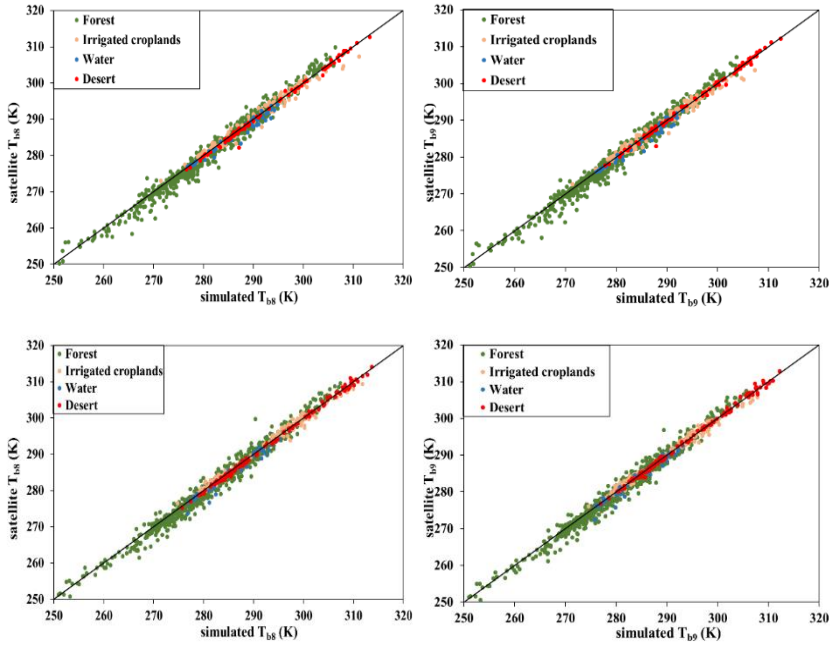

Figura 14. *Temperatura de brillo (T_{bi}) de satélite frente a las temperaturas de brillo simuladas. La parte superior corresponde al satélite Sentinel-3A y la parte inferior al Senintel-3B, mientras que a la izquierda se representa la banda 8 (10.85 µm) y a la derecha la banda 9 (12.05 µm). La línea 1:1 se muestra en negro. (Puchades et al., 2025).*

En la Tabla 11 se muestra la mediana y la R-RMSD de las diferencias para ambos satélites y ambas bandas. Las medianas y R-RMSD para ambos sensores difieren entre 0.2 y 0.3 respectivamente, lo cual muestra un ajuste consistente para ambas bandas y ambos satélites. Según zonas, las medianas varían entre -1.0 K para HYY (bosque) en la banda 8 y -0.7 K para la banda 9, alrededor de 0.3 K para el cultivo de arroz y dando valores cercanos a cero en las zonas de estudio de KIT, Gobabeb y Lago Constanza. Los biomas de bosques muestran cierta variabilidad en los valores de la mediana, siendo prácticamente insignificantes en ambas bandas en PUE y ROB. Por tanto, los valores de R-RMSD en los bosques oscilan entre 1.0 y 1.3 K en HYY y 0.5 y 0.8 K en PUE y ROB, como se puede observar en la mayor dispersión de los datos del bioma bosque en la Figura 14. Las zonas boscosas son complejas y variadas, por lo que obtener la LST y la emisividad en la superficie es una tarea compleja, lo que puede explicar dicha variabilidad. Los valores de R-RMSD más bajos se obtuvieron en Gobabeb (0.4 K y 0.3 K para las bandas 8 y 9 de ambos sensores respetivamente). Tomando todas las zonas juntas, la Tabla 11 muestra pequeñas medianas (entre -0.1 y -0.3 K) y R-RMSD de 0.8 K para la banda 8 y 0.7 K para la banda 9. Los valores de R-RMSD de la Tabla 11 son compatibles con el análisis de incertidumbre de la sección (2.g) para la mayoría de los emplazamientos en las dos bandas y sensores, lo que sugiere un buen ajuste de la calibración de las bandas IRT del SLSTR.

Los elevados valores de R-RMSD en la zona de Cortes se deben a la distinta incertidumbre aleatoria entre los valores diurnos y nocturnos de ambos satélites. Esto puede deberse a la dificultad de obtener valores representativos de los datos satelitales a partir de medidas in situ. Este inconveniente puede darse en zonas de bosques donde la visión de los radiómetros que están instalados en mástiles puede que no sea similar a la vista del satélite.

Tabla 11. Mediana y valores R-RMSD de las diferencias entre las BT de satélite y las BT simuladas con los datos in situ como referencia para las bandas 8 y 9 en cada estación de medidas in situ y el total. Se muestra los valores para Sentinel-3A y -3B.

		Sentinel-3A			Sentinel-3B		
		Mediana	**R-RMSD**	**Nº de datos**	**Mediana**	**R-RMSD**	**Nº de datos**
Cortes	**Band 8**	0.3	1.4	86	0.6	1.5	85
	Band 9	0.3	1.4		0.5	1.4	
Arrozales	**Band 8**	0.3	0.7	136	0.3	0.7	124
	Band 9	0.2	0.7		0.3	0.7	
KIT	**Band 8**	-0.3	0.6	130	-0.3	0.5	158
	Band 9	-0.1	0.5		-0.0	0.5	
Gobabeb	**Band 8**	-0.3	0.4	87	-0.3	0.4	136
	Band 9	0.0	0.3		-0.2	0.3	
Lago Constanza	**Band 8**	-0.4	0.6	54	-0.4	0.6	60
	Band 9	-0.4	0.6		-0.3	0.4	
HYY	**Band 8**	-0.9	1.2	112	-1.0	1.3	113
	Band 9	-0.7	1.0		-0.7	1.0	
PUE	**Band 8**	-0.2	0.7	152	-0.2	0.7	166
	Band 9	0.0	0.7		0.1	0.7	
ROB	**Band 8**	-0.2	0.7	49	-0.2	0.8	65
	Band 9	-0.2	0.5		-0.2	0.8	
SVA	**Band 8**	-0.6	1.0	118	-0.7	1.1	116
	Band 9	-0.5	0.9		-0.5	1.0	
Total	**Band 8**	-0.3	0.8	924	-0.3	0.8	1023
	Band 9	-0.2	0.7		-0.1	0.7	

a.2. Validación con el método t-based

Una vez evaluada la calibración de las bandas de los satélites, realizamos la validación t-based del producto operativo de SLSTR y del algoritmo SW propuesto por Pérez-Planells et al. (2021) utilizando los mismos datos de LST in situ. Las figuras 15 y 16 muestran las medianas y R-RMSDs, respectivamente, para los biomas mostrados en la Figura 14. Los resultados se han separado entre día, noche y todos conjuntamente. Para el caso especial del agua solo existen datos diurnos debido a que el barco que lleva instalado el sensor solo navega de día.

Observando la Figura 15 se puede apreciar que las medianas son similares para Sentinel-3A y -3B independientemente del bioma o del algoritmo LST, mostrando una buena consistencia entre ambos sensores. El producto operativo de SLSTR muestra unas diferencias que superan el umbral de ±1 K para el desierto y los cultivos de regadío tanto de día como de noche con diferencias negativas para el desierto (-2.1 K para Sentinel-3A y -2.2 K para -3B) y positivas para los cultivos de regadío (1.6 K y 2.0 K). Para los biomas de bosque y agua el producto muestra diferencias mucho menores (alrededor de ±0.3 K) en todos los casos. El SW de Perez-

Planells et al. (2021) solo muestra una diferencia excesiva para el bioma de desierto y durante el día (-1.8 K para Sentinel-3A y -1.7 para Sentinel -3B), estando todos los demás biomas dentro de ±0.5 K en la mayoría de los casos.

Las R-RMSD de la Figura 16 muestran valores muy similares para los dos algoritmos en el agua (alrededor de 0.7 K para Sentinel-3A y 0.9 K para Sentinel-3B) y los bosques (inferiores a 1.5 K). Sin embargo, el producto SLSTR muestra mayores R-RMSD para el desierto (2.2 K para Sentinel-3A y 2.3 K para 3B) y los cultivos de regadío (1.8 K y 2.2 K), combinando datos diurnos y nocturnos. Comparativamente, el SW alternativo produce menores R-RMSD en el desierto (1.6 K y 1.3 K para Sentinel-3A y 3B, respectivamente) y en cultivos de regadío (0.9 K y 0.8 K).

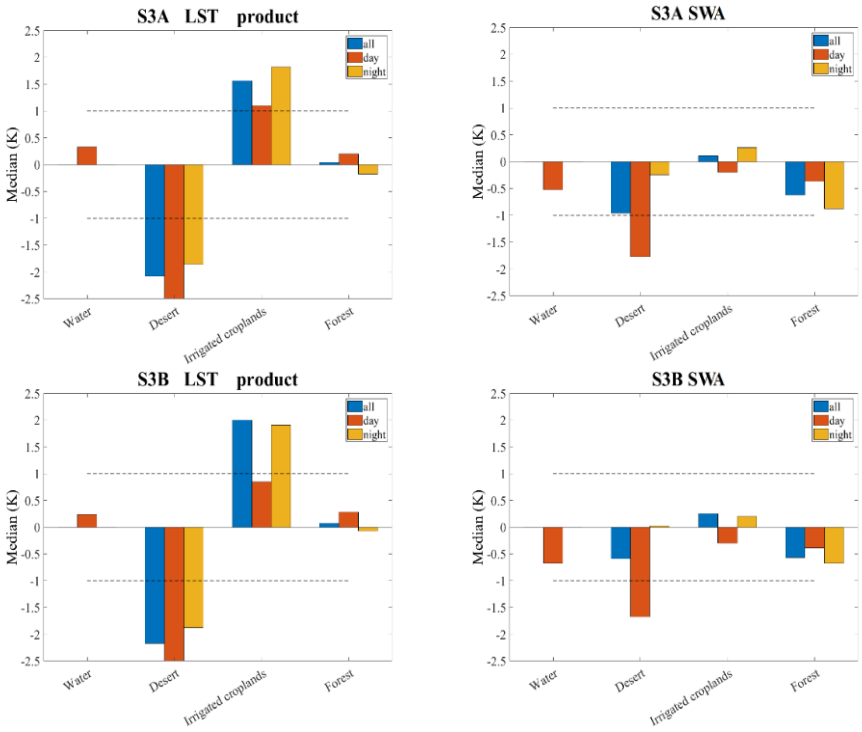

Figura 15. *Medianas de las diferencias entre la LST de satélite y la LST in situ para cada bioma, con el producto operativo de SLSTR a la izquierda y el SW de Pérez-Planells et al. (2021) a la derecha. En la parte superior se muestran los resultados para Senitnel-3A y la parte inferior los resultados para Senitnel-3B. Las líneas discontinuas muestran el umbral de ±1 K establecido por GCOS. (Puchades et al., 2025).*

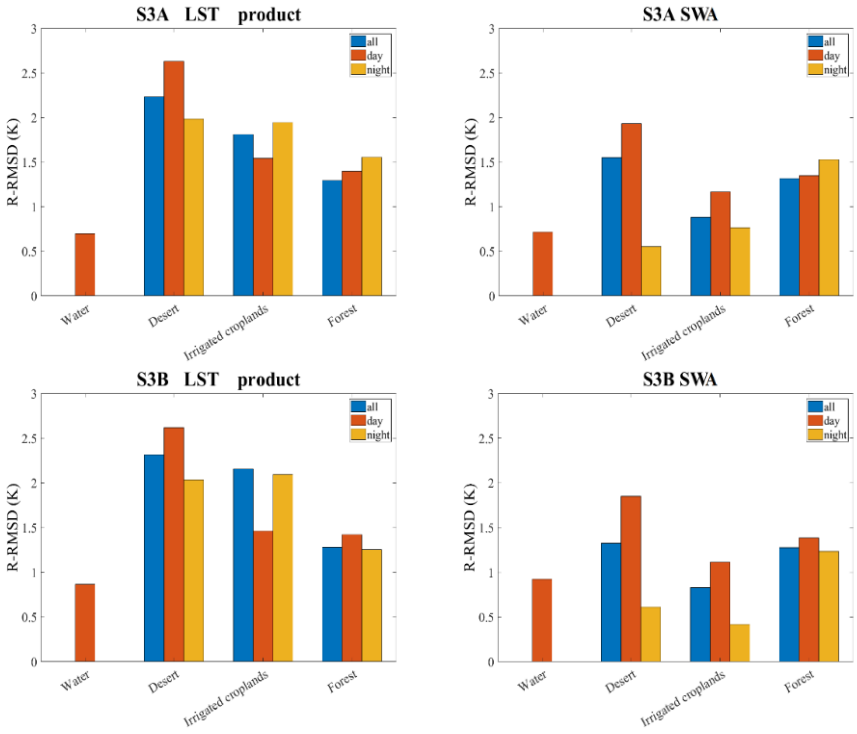

Figura 16. *Valores R-RMSD de las diferencias entre la LST de satélite y la LST in situ para cada bioma. El producto operativo de LST se muestra en la parte izquierda y el SW de Pérez-Planells et al. (2021) en la parte derecha. En la parte superior se muestran los datos para Sentinel-3A y en la parte inferior para Sentinel-3B. (Puchades et al., 2025).*

Para los dos biomas con mayores diferencias (desierto y cultivos de regadío), se muestra en la Figura 17 la dependencia de las diferencias de la LST del satélite menos la in situ con el contenido en vapor de agua (W) y la LST in situ para el producto operativo SLSTR y el SW de Pérez-Planells et al. (2021). El producto sobreestima las LST in situ independientemente de W para los cultivos de regadío, con grandes subestimaciones en el desierto para los valores de W más bajos. También podemos ver que las diferencias son cercanas a cero para altas temperaturas en los cultivos de regadío, mientras que es lo contrario para el desierto. Las diferencias del SW alternativo se aproximan a cero independientemente del W para el desierto y los cultivos de regadío, pero muestran diferencias negativas para LST elevadas, especialmente en el desierto. Esto explica la mediana negativa del algoritmo SW para los datos diurnos en el desierto.

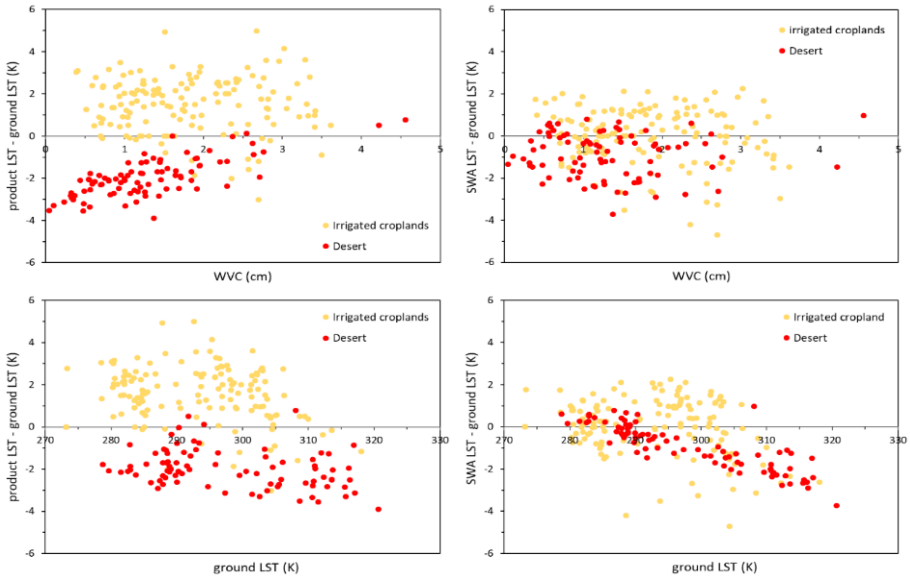

Figura 17. *LST satelitales menos LST in situ para los biomas de cultivos de regadío y desierto frente al contenido de vapor de agua atmosférico (W) (parte superior) y las temperaturas LST in situ (parte inferior), para el producto operativo SLSTR (parte izquierda) y para el SW de Pérez-Planells et al. (2021) (parte derecha). (Puchades et al., 2025).*

En el caso de los cultivos de regadío, los resultados pueden estudiarse más a fondo teniendo en cuenta los cambios de la cubierta terrestre en la zona debido a la fenología del arroz, con tres tipos de superficie diferentes a lo largo del año (suelo desnudo húmedo y seco, vegetación y suelo inundado). La Tabla 12 muestra los resultados específicos separados por los diferentes tipos de cubiertas, tanto para el producto operativo como el SW de Pérez-Planells et al. (2021) para Sentinel-3A y Sentinel-3B. Los cambios de cobertura no se tienen en cuenta en la asignación fija de bioma en el producto operativo LST y ésta puede ser una de las razones por las que el SW alternativo, en el que se utiliza emisividades dinámicas en función del tipo de cobertura, obtiene mejores resultados para este bioma concreto en los tres tipos de coberturas. Puede observarse que ambos satélites muestran resultados semejantes.

Tabla 12. Mediana y R-RMSD (K) de la validación t-based para cada tipo de cobertura en el bioma de cultivos de regadío, para el producto LST y el SW alternativo, para Sentinel-3A y Sentinel-3B.

		Sentinel-3A		**Sentinel-3B**	
Cobertura	**Estadístico**	**Producto SLSTR**	**SW**	**Producto SLSTR**	**SW**
Vegetación	Mediana	1.6	0.3	1.8	0.4
	R-RMSD	1.8	1.0	2.0	1.1
Suelo desnudo	Mediana	1.3	0.8	0.7	0.2
	R-RMSD	1.5	1.0	1.0	0.7
Suelo inundado	Mediana	2.0	-0.2	2.4	0.0
	R-RMSD	2.3	0.6	2.4	0.3

a.3. Validación con el método r-based

Además de la validación utilizando el método t-based, también hemos realizado la validación con el método r-based en todas las zonas de validación de la Tabla 3, que además de las anteriores incluye otras zonas de estudio que cubren nuevos biomas donde no se dispone de medidas in situ de LST. Las zonas r-based permiten aumentar el número de datos en los biomas definidos previamente en la validación t-based (agua, con las zonas de validación Salinas y Albufera, y bosque con la zona Soria) y añadir dos nuevos biomas según la clasificación de SLSTR: vegetación escasa (olivar) y cultivos de secano (viñedo).

Las figuras 18 y 19 muestran, respectivamente, los valores de la mediana y de la R-RMSD de la validación r-based para los seis biomas utilizados. El número total de casos es de 1978 para Sentinel-3A y 2190 para Sentinel-3B, prácticamente el doble que en la validación t-based (Tabla 11). Los resultados fueron similares para los datos de Sentinel-3A y 3B, excepto para los cultivos de secano durante el día en el caso del SW de Pérez-Planells et al. (2021) (0.6 K para 3A y 1.0 K para 3B). En general, los resultados fueron comparables a los de la validación t-based, mostrando el producto SLSTR diferencias medias en torno a -1.5 K para el desierto y 1.5 K para los cultivos de regadío, así como el SW de Pérez-Planells et al. (2021) para el desierto durante el día (-1.5 K). Las diferencias de este fueron próximas a cero en todos los demás casos. Las pequeñas diferencias y R-RMSD mostradas por el SW alternativo pueden deberse al hecho de que utiliza la misma base de datos de emisividad (Tabla 7) que el método r-based. Para el producto operativo, las R-RMSD se situaron en torno a 2.0 K para los cultivos de regadío y el desierto y por encima de 1.0 K para los cultivos de secano. Para el SW alternativo, sólo el desierto de día (R-RMSD>1.5 K) superó el límite de ±1 K.

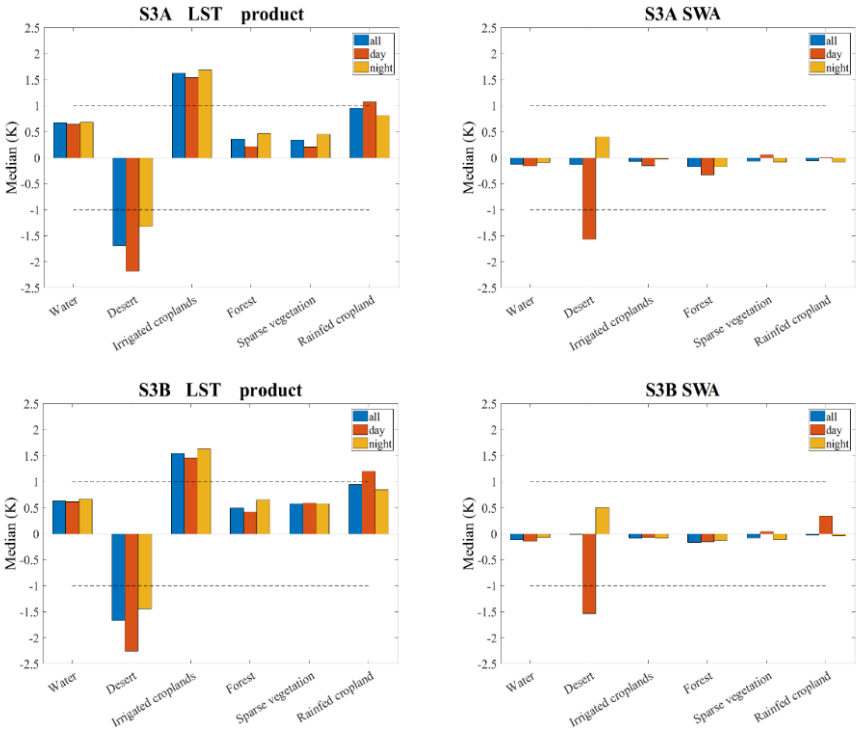

Figura 18. *Mediana de las diferencias entre las LST de satélite y las LST del método de validación r-based para cada bioma. El producto operativo de SLSTR a la izquierda y a la derecha el SW de Pérez-Planells et al. (2021), en la parte superior Sentinel-3A y en la parte inferior Sentinel-3B. Las líneas discontinuas representan el umbral ±1 K establecido por GCOS. (Puchades et al., 2025).*

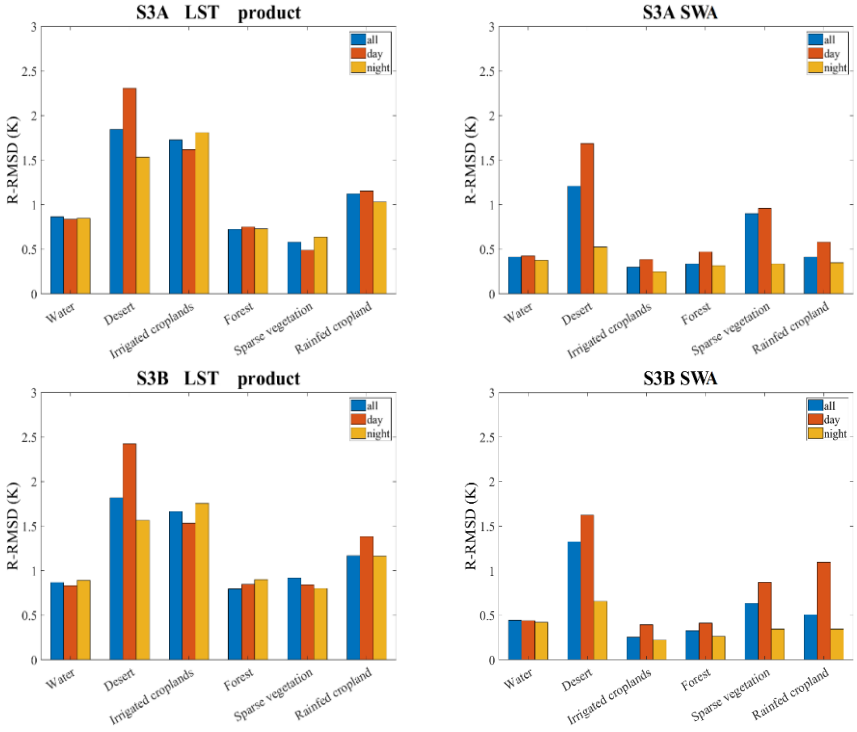

Figura 19. *Valores de R-RMSD de las diferencias entre las LST de satélite y las LST del método r-based para cada bioma. El producto operativo de SLSTR a la izquierda y a la derecha el SW de Pérez-Planells et al. (2021), en la parte superior Sentinel-3A y en la parte inferior Sentinel-3B. (Puchades et al., 2025).*

En la Figura 20 se muestran las diferencias entre la LST de satélite y la de referencia r-based frente a W y la LST de referencia para desierto, cultivos de regadío y cultivo de secano. En el caso del desierto y de cultivos de regadío, los resultados coinciden con los de la Figura 17. El producto operativo muestra diferencias negativas para bajos valores de W y LSTs altas para el bioma desierto, y diferencias positivas independientemente de W y LST para el bioma de cultivo de regadío. Esto también se observa para los cultivos de secano. Las diferencias del SW de Pérez-Planells et al. (2021) sólo muestran una correlación significativa con LST en el desierto, con grandes valores negativos para temperaturas altas, como en la Figura 17. Para los cultivos de regadío y de secano, los sesgos son pequeños independientemente de W y LST.

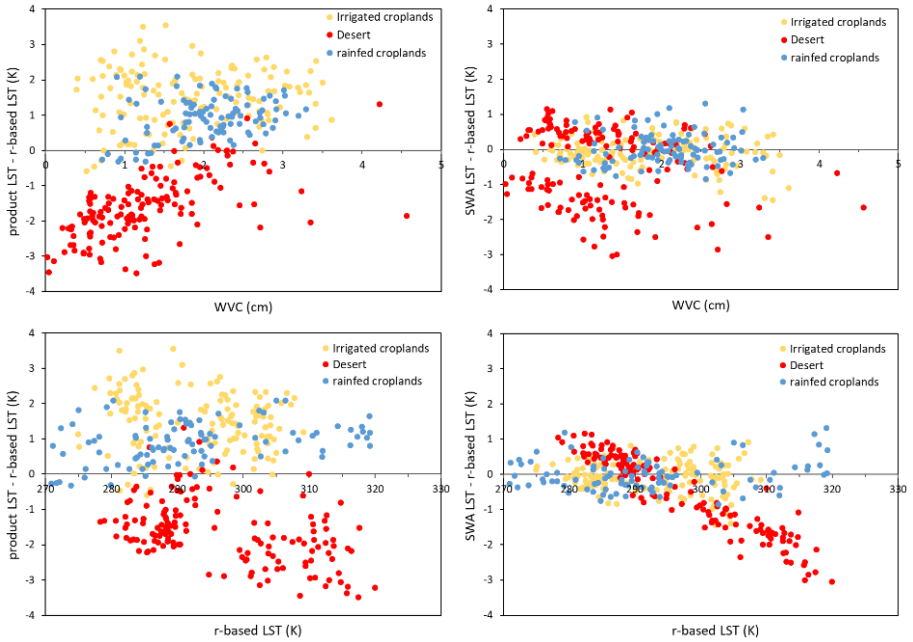

Figura 20. *LST satelitales menos LST r-based para los biomas de cultivos de regadío, desierto y cultivos de secano frente al contenido de vapor de agua atmosférico (W) (parte superior) y las temperaturas LST de referencia r-based (parte inferior), para el producto operativo SLSTR (parte izquierda) y para el SW (parte derecha). (Puchades et al., 2025).*

b. Resultados para el sensor TIRS

En Niclòs et al., (2021) realizamos la evaluación de la calibración del sensor TIRS y la validación de diferentes algoritmos SW propuestos por diversos autores (detallados en la sección 2.e.2.2) sobre la superficie de los arrozales de Valencia. Utilizamos datos TIRS anteriores y posteriores al reprocesado de 2014 y posteriores al reprocesado de 2017 (sección 1.d.2). Todas las medidas in situ que se utilizaron como referencia fueron tomadas a lo largo de transectos con varios radiómetros CIMEL312-2 entre los años 2012 y 2019. Al presentar distintos tipos de cobertura durante el año, esta zona de validación proporciona un amplio intervalo de LSTs (valores diurnos entre 280 K y 315 K en el estudio) para la calibración y la validación de sensores en el IRT.

b.1. Calibración de las bandas TIR del sensor TIRS

Para la evaluación de la calibración, simulamos las radiancias TOA en las bandas 10 y 11 a partir de las medidas in situ mediante la ecuación (2.1) y calculamos la BT equivalente (ecuación 2.3). La comparación con los datos de satélite se hizo tanto en términos de radiancias TOA como de BTs. En la Tabla 13 mostramos la media, SD y RMSD de las diferencias entre los valores de satélite menos los simulados en las dos bandas en cada una de las versiones de reprocesado del TIRS. La comparación se realizó tomando únicamente el píxel

más cercano a la zona de medidas y una matriz de 17×17 píxeles centrada sobre ella. Los resultados que se obtienen en ambos casos son muy similares debido a la homogeneidad térmica de la zona.

Tabla 13. Valores medios de las diferencias ± SD (RMSD) entre los datos TIRS y simulados para cada banda en términos de radiancia y de temperatura, para las tres versiones de reprocesado. Los resultados se muestran para la matriz (17×17) y para un píxel único (1×1). (Niclòs et al., 2021).

		$L_{10}(Wm^{-2}sr^{-1}\mu m^{-1})$	$T_{10}(K)$	$L_{11}(Wm^{-2}sr^{-1}\mu m^{-1})$	$T_{11}(K)$
Datos originales	17×17	0.09 ± 0.10 (0.13)	0.7 ± 0.7 (1.0)	0.33 ± 0.13 (0.35)	2.7 ± 1.0 (2.9)
	1×1	0.08 ± 0.10 (0.13)	0.6 ± 0.7 (0.9)	0.32 ± 0.12 (0.34)	2.7 ± 1.0 (2.8)
Reprocesado 2014	17×17	-0.20 ± 0.10 (0.22)	-1.4 ± 0.7 (1.6)	-0.18 ± 0.13 (0.22)	-1.6 ± 1.1 (2.0)
	1×1	-0.21 ± 0.10 (0.24)	-1.5 ± 0.7 (1.7)	-0.19 ± 0.12 (0.23)	-1.7 ± 1.1 (2.0)
Reprocesado 2017	17×17	0.02 ± 0.17 (0.17)	0.2 ± 1.2 (1.2)	0.01 ± 0.18 (0.18)	0.1 ± 1.6 (1.6)
	1×1	0.00 ± 0.17 (0.17)	0.1 ± 1.2 (1.2)	0.01 ± 0.18 (0.18)	0.1 ± 1.5 (1.5)

La Figura 21 muestra las T_i simuladas frente a las T_i del sensor (valores medios para 17×17 píxeles) para las bandas 10 y 11 con (a) la calibración original, (b) el reprocesado de 2014 y (c) el de 2017. Las incertidumbres en las BTs simuladas, representadas por las barras de error de la figura, fueron estimadas tal como se detalla en la sección (2.e), oscilando entre 0.7 y 0.9 K para la banda 10 y entre 1.0 y 1.2 K para la banda 11 para todos los datos. De acuerdo con la Tabla 13 y la Figura 21, el reprocesado de 2014 no determinó resultados adecuados para la zona de los arrozales de Valencia, mostrando subestimaciones similares para ambas bandas (de alrededor de -1.5 K). Nótese que los datos originales mostraban una sobreestimación cercana a 1 K en la banda 10 y a 3 K en la banda 11. Los resultados de la banda 10, utilizada por los usuarios para obtener la LST con el SC, fueron peores que antes de este reprocesado. Para analizar el reprocesado de 2017, se dispuso de más datos de validación adquiridos en 2016-2019, ampliando así la evaluación de los datos TIRS disponibles hasta 25 casos. Las barras de error en las figuras son las incertidumbres de los datos descritas en la sección 2.g teniendo en cuenta la propagación de errores de las variables requeridas.

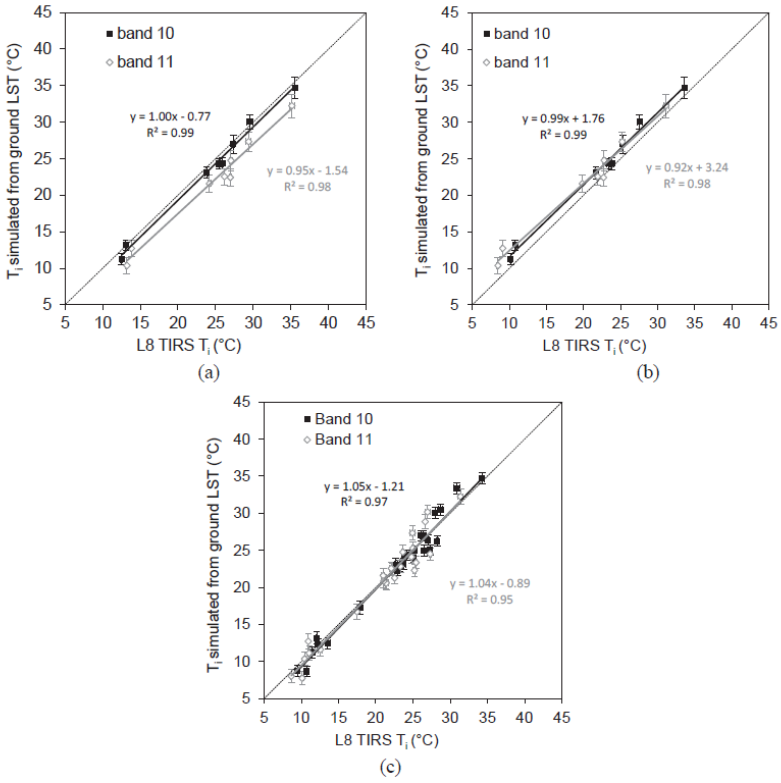

Figura 21. *Comparación de las temperaturas simuladas a partir de las temperaturas in situ y las adquiridas por el sensor TIRS de L8. a) Antes del reprocesado de 2014 (datos originales). b) Después del reprocesado de 2014. c) Después del reprocesado de 2017 (datos de la Colección 1). Las líneas punteadas muestran la regresión lineal de los datos con sus ecuaciones de ajuste y el coeficiente de determinación R^2. (Niclòs et al., 2021).*

A diferencia del reprocesado de 2014, en el que solo se aplicó la sustracción de valores constantes en cada banda, en el reprocesado de 2017 se utilizó el algoritmo "TIRS-on-TIRS" (Montanaro et al., 2015) que permitió modelar y eliminar la radiancia procedente de fuera del campo de visión del sensor, utilizando así una corrección específica para cada escena y píxel. Los resultados de la Figura 21 (c) y la Tabla 13 muestran que esta corrección mejoró los resultados en los datos TIRS disponibles en la Colección 1. Se obtuvo diferencias medias cercanas a cero y RMSDs de 1.2 K y 1.6 K para las bandas 10 y 11, respectivamente.

La Figura 22 (a-b) muestra los resultados de la calibración de 2017 para ambas bandas TIRS por tipo de cubertura sobre la zona de los arrozales de Valencia con el fin de analizar las posibles dependencias de la cubierta terrestre. Las mayores diferencias se observaron principalmente en la condición de suelo desnudo seco. Este hecho podría deberse a la menor homogeneidad observada en estudios anteriores para esta cubierta superficial en la zona (Niclòs et al., 2018). Sin embargo, estas mayores diferencias corresponden a las LST más altas (>307 K) y podrían ser también consecuencia de este hecho.

Los resultados también se analizaron según las dos trayectorias en las que L8 cubre la zona de validación (*paths* 198 y 199), ya que el algoritmo TIRS-on-TIRS utilizado en el reprocesado de 2017 debería de ser igualmente efectivo para cualquier escena, aunque la trayectoria 198 cubre una mayor proporción de mar que la 199. La Figura 22 (c-d) muestra dichos resultados, con diferencias entre los paths 198 y 199 de alrededor de 0.8 K en términos de SD, pero diferencias medias insignificantes. Sin embargo, dado que el número de datos es bastante diferente para ambos casos, con sólo 8 datos coincidentes para la 198 y 17 datos coincidentes para la 199, las diferencias de SD observadas pueden ser sólo consecuencia de este hecho.

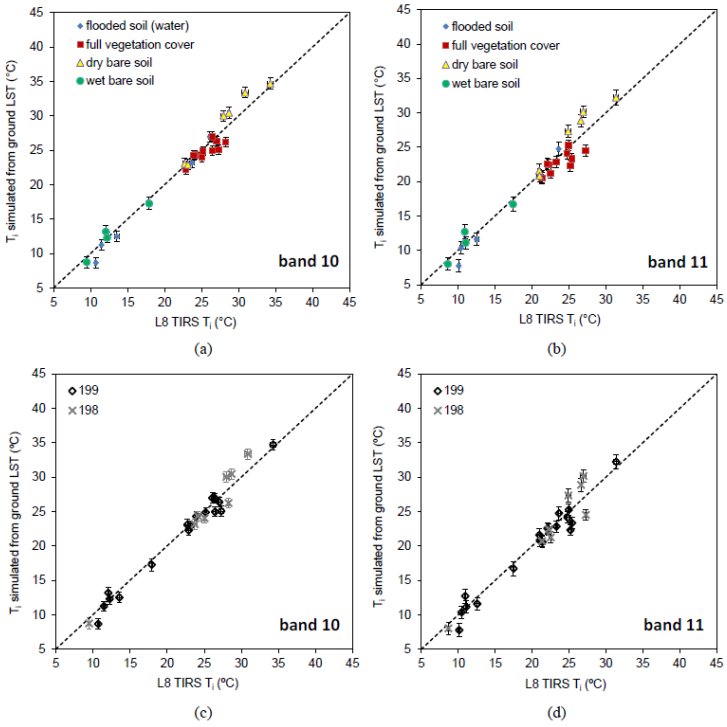

Figura 22. *Comparación de las temperaturas de brillo para las bandas 10 y 11 simuladas a partir de la LST in situ frente a las adquiridas con el sensor TIRS tras el reprocesado de 2017, separadas por: (a)-(b) cobertura del suelo, y (c)-(d) trayectoria de adquisición. (Niclòs et al., 2021).*

En algunos casos de las Figuras 21 y 22, la comparación entre las BT simuladas y las de satélite muestran diferencias mayores que la incertidumbre combinada de ambas mediciones. La mayor fuente de error en las BT simuladas son los cálculos de transferencia radiativa. En algunos casos, los perfiles atmosféricos utilizados pueden diferir del perfil atmosférico real sobre el lugar en el momento del paso del satélite más de lo que supusimos en el análisis de sensibilidad.

Podemos comprobar la idoneidad de los perfiles atmosféricos comparando las diferencias entre las BTs de las dos bandas observada desde satélite, $(T_{10}\text{-}T_{11})_{obs}$ con la misma diferencia de BTs obtenida de la simulación a partir de datos de tierra, $(T_{10}\text{-}T_{11})_{sim}$. Entonces se puede analizar la diferencia $\delta(T_{10}\text{-}T_{11}) = (T_{10}\text{-}T_{11})_{obs} - (T_{10}\text{-}$

$T_{11})_{sim}$; un concepto que está relacionado con el método de validación r-based (Coll et al., 2012; Wan and Li, 2008). Nótese que, en las dos bandas, la absorción atmosférica se debe principalmente al mismo mecanismo (vapor de agua atmosférico), que es más fuerte en la banda 11, por lo que T_{10}-T_{11} suele ser positivo y aumenta con W. Entonces, si el perfil atmosférico se aproxima al perfil real, la diferencia $\delta(T_{10}$-$T_{11})$ debería ser cercana a cero. Sin embargo, si el perfil utilizado sobreestima (subestima) el efecto atmosférico, el valor de $(T_{10}$-$T_{11})_{sim}$ será mayor (menor) que $(T_{10}$-$T_{11})_{obs}$ y, por tanto, $\delta(T_{10}$-$T_{11}) < 0$ (>0).

Esto puede observarse en la Figura 23, donde las diferencias entre T_{10} simulada a partir de datos in situ y T_{10} observado por TIRS se representan gráficamente frente a $\delta(T_{10}$-$T_{11})$ para todos los datos. Los casos en los que $\delta(T_{10}$-$T_{11})$ es pequeño (en valor absoluto) están asociados a valores absolutos pequeños de las diferencias de temperatura, siendo también cierto lo contrario. Así pues, podemos utilizar la diferencia $\delta(T_{10}$-$T_{11})$ como comprobación de la calidad del perfil atmosférico y, en consecuencia, de la precisión de la simulación de la BTs. Para ello, podemos establecer un umbral en $\delta(T_{10}$-$T_{11})$ basado en un análisis de incertidumbre aplicado a las diferencias de temperatura representadas en la Figura 23, utilizando los resultados de los análisis de incertidumbre precedentes para las temperaturas individuales implicadas.

Dado que ambas temperaturas son independientes entre sí, la incertidumbre combinada en la diferencia entre T_{10} simulada y observada son las incertidumbres individuales sumadas cuadráticamente, lo que ofrece un valor medio de 0.9 K para todos los casos. Para la diferencia $\delta(T_{10}$-$T_{11})$, tenemos que considerar tanto $(T_{10}$-$T_{11})_{obs}$ como $(T_{10}$-$T_{11})_{sim}$. Para la primera, la incertidumbre es la suma en cuadratura de las incertidumbres de las BTs de TIRS. Para la segunda, se suman cuadráticamente las incertidumbres de las T_{10} y T_{11} simuladas debidas a la emisividad, NEΔT y variabilidad espacial (sección 2.g). Sin embargo, el error en las simulaciones de transferencia radiativa tiene el mismo signo en ambas bandas, independientemente de que estemos sobreestimando o subestimando el efecto atmosférico real, por lo que se cancelan en parte al calcular la diferencia $\delta(T_{10}$-$T_{11})$. Sumando las cinco fuentes de error cuadráticamente, obtuvimos una incertidumbre media de 0.5 K para $\delta(T_{10}$-$T_{11})$ para todos los datos.

Por lo tanto, según el análisis de incertidumbres anterior y la Figura 23, podemos ver que cuando $\delta(T_{10}$-$T_{11})$ está dentro de ±0.5 K la incertidumbre en las diferencias de BTs estará dentro de ±0.9 K (SD de las diferencias), lo que sugiere tomar la condición $-0.5\ K < \delta(T_{10}$-$T_{11}) < 0.5\ K$ como umbral para seleccionar los casos en los que los datos del perfil atmosférico son fiables. Para todos los datos estudiados, 5 de 25 casos no cumplían dicha condición (véase la Figura 23). Eliminando los 5 casos de la comparación, la media de las diferencias de T_{10} fue finalmente de 0.08 K (en lugar de 0.15 K obtenido anteriormente), y la SD y RMSD fueron de 1.0 K (en lugar de 1.2 K).

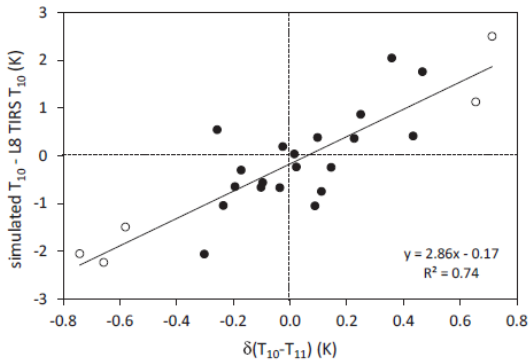

Figura 23. *Diferencias entre BTs simuladas a partir de datos in situ y BTs obtenidas con el sensor TIRS en la banda 10 frente a las diferencias entre las BTs observadas y simuladas para ambas bandas $\delta(T_{10}-T_{11})$. La línea recta es la regresión lineal de los datos con su ecuación de ajuste y el coeficiente de determinación (R^2). Los cirulos abiertos muestran los casos que no cumplen la condición $-0.5\ K < \delta(T_{10}-T_{11}) < 0.5\ K$. (Niclòs et al., 2021).*

Finalmente, se evaluó la actualización de calibración propuesta en diciembre de 2019 en preparación para la Colección 2, que es la única disponible para el usuario en la actualidad. La Figura 24 muestra los resultados después de corregir los datos TIRS reprocesados en 2017 con la actualización propuesta por el Equipo de Calibración L8 en 2019. Los datos de la Figura 24 corresponden a la configuración electrónica principal (*side* A) y a la secundaria (*side* B), que entró en funcionamiento debido a los problemas presentados por un exceso de consumo en corriente eléctrica del *side* A. En la actualidad la electrónica que sigue operativa es la *side* B. Los resultados mostraron que las diferencias medias de BTs para las bandas 10 y 11 fueron -0.08 K y 0.5 K, respectivamente (en lugar de 0.15 K y 0.14 K obtenidos para los datos reprocesados en 2017), sin cambios significativos en términos de SDs y RMSDs.

Por lo tanto, para nuestra zona, la actualización de la calibración mejora ligeramente los resultados para la banda 10, en la que se aplica un algoritmo SC para proporcionar un producto LST en la actual Colección 2 (Gerace et al., 2020; Malakar et al., 2018). Como observamos, la variación en la diferencia media introducida por la actualización de la calibración fue de -0.2 K en la banda 10 para ambas configuraciones electrónicas. Las diferencias medias introducidas en la banda 11 fueron de 0.6 K para el periodo del lado A y de 0.3 K para el periodo del lado B. Así pues, el cambio de respuesta debido a la actualización de la calibración fue mayor para la banda 11 que para la banda 10 en la zona de los arrozales, aumentando las diferencias medias entre bandas. De nuevo, las RMSD se redujeron ligeramente (a 0.9 K para la banda 10) al eliminar los cinco casos mencionados en el párrafo anterior.

Tras la comprobación de las distintas recalibraciones del sensor TIRS, pasamos a la validación t-based de la LST obtenida a partir de aplicación de algoritmos SC (ecuación 2.4) aplicados a las radiancias TOA de las bandas 10 y 11 por separado. Utilizamos los valores de emisividad de la Tabla 7 y los parámetros atmosféricos simulados para las bandas correspondientes.

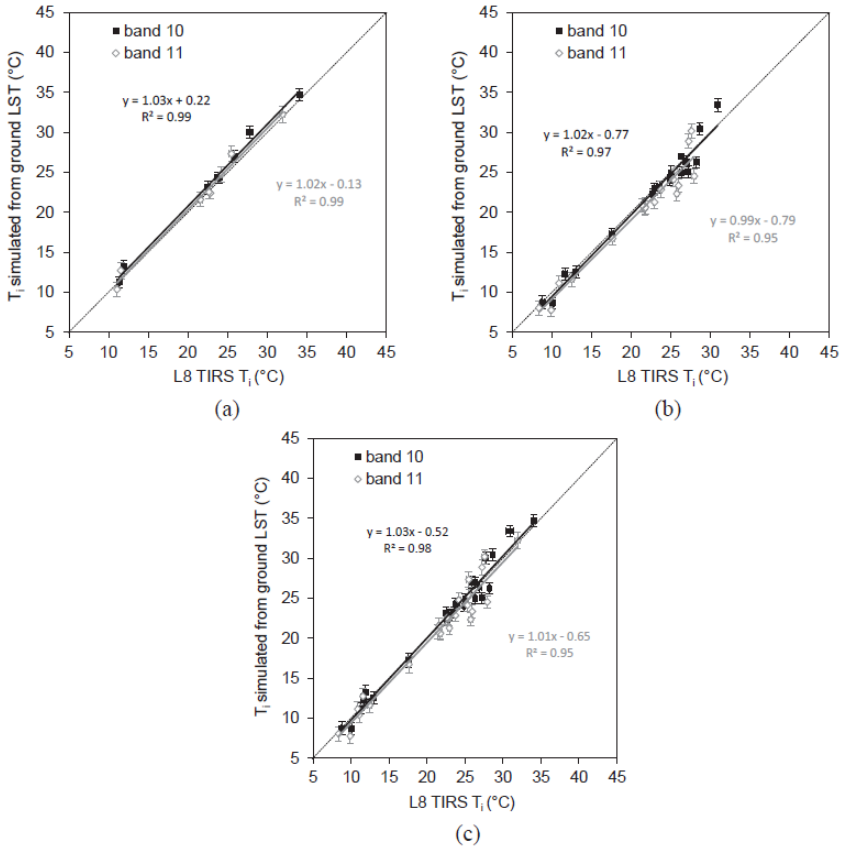

Figura 24. *Comparación de BTs simuladas a partir de LSTs in situ con las BT de satélite con la calibración propuesta de diciembre de 2019, para las bandas 10 y 11: a) para el periodo con el instrumento TIRS operando en el side A; b) para el periodo con el instrumento TIRS operando en el side B; y c) ambos juntos. (Niclòs et al., 2021).*

En la Tabla 14 se puede observar los resultados estadísticos de la validación de las LSTs. Estos muestran que el reprocesado de 2017 proporciona menores diferencias medias, e iguales para ambas bandas, en relación a las obtenidas con los reprocesados anteriores. Para los datos reprocesados en 2017, las SD son las más altas probablemente debido al mayor número de coincidencias utilizadas en este caso (25 casos en el periodo 2014-2019 en lugar de 8 casos en 2014). Los resultados para la banda 10 muestran RMSD más bajos que los de la banda 11, lo que sugiere un mejor funcionamiento de la corrección SC para la banda 10 con los datos después del reprocesado de 2017.

Tabla 14. Diferencias medias ± SD (RMSD) entre LSTs obtenidas con un SC para cada banda y LSTs in situ (17 × 17 píxeles), con los datos originales y tras los reprocesados de 2014 y 2017.

	LST_{10} (K)	**LST_{11} (K)**
Datos originales	0.9 ± 1.0 (1.3)	4.0 ± 2.2 (4.6)
Datos reprocesados 2014	-1.7 ± 0.8 (1.9)	-2.1 ± 1.4 (2.5)
Datos reprocesados 2017	0.2 ± 1.5 (1.6)	0.3 ± 2.3 (2.3)

Guo et al. (2020) obtuvieron la misma conclusión usando el algoritmo SC propuesto por Jiménez-Muñoz et al. (2014) y datos de cinco emplazamientos SURFRAD como referencia, pero con incertidumbres más altas. La red SURFRAD incluye distintas estaciones meteorológicas y climatológicas donde la LST in situ se mide con pirgeómetros infrarrojos de banda ancha (4-50 µm), integrados hemisféricamente. Dichas medidas no se corresponden necesariamente con las medidas direccionales de banda estrecha de los satélites, especialmente para datos diurnos (Wang et al., 2008), a diferencia de los radiómetros IRT utilizados en esta Tesis.

Utilizando datos TIRS tras el reprocesado de 2017, Guo et al. (2020) obtuvieron diferencias medias (RMSDs) de 0.5 K (2.5 K) y 1.6 K (3.6 K) para las bandas 10 y 11, respectivamente. Los resultados fueron diferentes cuando excluyeron un emplazamiento SURFRAD debido a problemas de heterogeneidad y altas temperaturas, con diferencias medias (RMSDs) de 0.10 K (2.1 K) y 1.0 K (2.8 K) para las bandas 10 y 11, respectivamente. Sin embargo, Guo et al. (2020) obtuvieron diferencias medias más bajas (pero RMSD más grandes) para los datos de reprocesado de 2014 al utilizar los cinco sitios SURFRAD: -0.17 K (2.7 K) y 1.0 K (4.3 K) para las bandas 10 y 11, respectivamente. Tras excluir el mismo emplazamiento del análisis de los datos de reprocesado de 2014, se obtuvieron diferencias medias (RMSD) de -0.8 K (2.4 K) y -0.17 K (3.2 K) para las bandas 10 y 11, respectivamente.

El aumento de las diferencias medias observadas por Guo et al. (2020) con la corrección de reprocesado de 2017 al utilizar los datos de los cinco emplazamientos SURFRAD (y también para la banda 11 al excluir los datos de una estación) no concuerda con los resultados SC de la Tabla 14 ni con los resultados de calibración mostrados en la Figura 21 y la Tabla 13.

Como se observa en la Figura 25, la diferencia entre las LST in situ y por satélite es mayor que la incertidumbre combinada de ambas mediciones en algunos casos. Esto puede deberse a la calidad de los perfiles atmosféricos utilizados para calcular las radiancias y transmisividades atmosféricas. De forma análoga al anterior análisis de calibración, podemos utilizar como prueba de calidad de los perfiles la diferencia entre las LST derivadas en las bandas 10 y 11 mediante el método SC, $\delta = T_{SC}(10)-T_{SC}(11)$, de igual forma que en el método r-based (sección 2.d). La diferencia δ debería ser cercana a cero si los perfiles atmosféricos utilizados en la simulación se aproximan a las condiciones reales. Por el contrario, si $\delta > 0$ (< 0) el efecto atmosférico calculado subestima (sobreestima) el efecto atmosférico real y, por tanto, las LST derivadas de SC deberían subestimar (sobreestimar) las LST reales.

La Figura 25 muestra las diferencias entre las LSTs in situ y SC obtenidas de la banda 10, $T_g-T_{SC}(10)$, frente a δ para todo el conjunto de datos. Podemos ver que los casos en los que el valor absoluto de δ es pequeño (grande) se corresponden con valores absolutos pequeños (grandes) de $T_g-T_{SC}(10)$. Se puede establecer un umbral en δ utilizando los resultados de los análisis de incertidumbre precedentes (sección 2.g) para las temperaturas individuales de la Figura 25. La incertidumbre combinada en $T_g-T_{SC}(10)$ es simplemente la suma en cuadratura de las incertidumbres de T_g y $T_{SC}(10)$, dando un valor medio de 1.0 K para todos los casos. La incertidumbre para δ se calculó de manera análoga a la de $(T_{10}-T_{11})_{sim}$, lo que arrojó una incertidumbre

promedio de 0.6 K para δ en todo el conjunto de datos. Por lo tanto, cuando δ está dentro de ± 0.6 K, la incertidumbre en T_g-T_{SC}(10) estará dentro de ± 1.0 K, lo que sugiere tomar la condición -0.6 K < δ < 0.6 K como umbral para seleccionar casos en los que el método SC es confiablemente preciso. Nueve de los 25 casos no cumplieron la condición (ver figura 25), y al eliminarlos de la comparación, las diferencias promedio permanecen igual que en la Tabla 14 (0.2 K) y las SD y RMSD fueron de 0.7 K (en lugar de 1.5 K y 1.6 K).

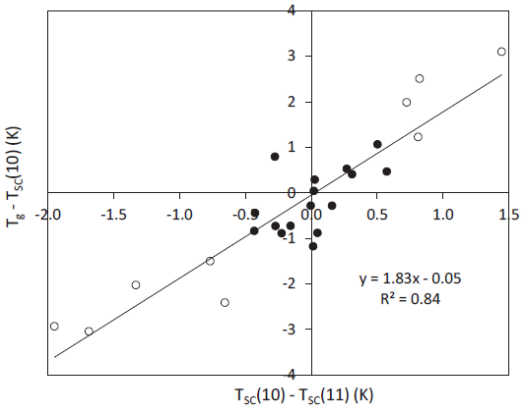

Figura 25. *Diferencias entre la LST in situ y la LST de SC de la banda 10 frente a las diferencias entre el LST de SC de la banda 10 y 11. La línea recta es la regresión lineal de los datos, con la ecuación de ajuste lineal y el coeficiente de determinación (R^2). Los símbolos abiertos representan los casos que no satisfacen la condición -0,6 K < T_{SC}(10)-T_{SC}(11) < 0,6 K. (Niclòs et al., 2021).*

Finalmente, también se evaluó la actualización de la calibración de 2019 en términos de la LST obtenida con el método SC. El valor medio de las diferencias respecto a las LST in situ fue de 0.0 K (en lugar de los 0.2 K que se muestran en la Tabla 14), sin cambios significativos en términos de SD y RMSD. Este resultado confirma, para la zona de los arrozales y su rango de temperatura, la ligera mejora introducida con la actualización de calibración para corregir los datos de la banda 10 con un algoritmo SC, como se utiliza en la Colección 2. Nuevamente, la RMSD se redujo a 0.7 K para la banda 10 al eliminar los nueve casos que no cumplen la condición de tener un perfil atmosférico óptimo.

b.2. Validación con el método t-based

A continuación, mostramos los resultados de la validación de LST (t-based) para los diferentes algoritmos SW propuestos para TIRS, descritos en la sección 2.e.2.2. En este caso, utilizamos los 25 casos (promedio de 17 × 17 píxeles) después del reprocesado de 2017.

La Figura 26 y la Tabla 15 (primera a tercera fila) muestran los resultados de validación para el algoritmo SW descrito en Jiménez-Muñoz et al. (2014) al utilizar los coeficientes propuestos en ese trabajo (JM), pero también con los coeficientes obtenidos para el mismo algoritmo por Meng et al. (2019). Para este algoritmo, los resultados se muestran en la Tabla 15 tanto al utilizar coeficientes por rangos de W (JM_M_W) como al utilizar coeficientes generales para cualquier W entre 0.0 cm y 7.0 cm (JM_M). Las Figuras 26a y 26b muestran

los resultados para JM y JM_M_W, respectivamente. Obtuvimos resultados similares al utilizar los coeficientes generales (JM_M).

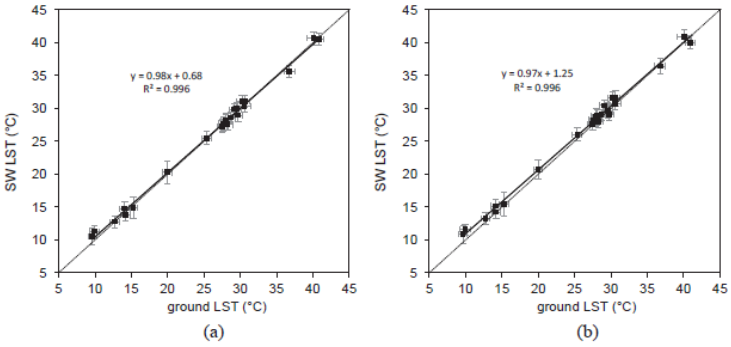

Figura 26. *Comparación de las LSTs obtenidas con el algoritmo Jiménez-Muñoz et al. (2014) SW (después del reprocesado de 2017) frente a datos LST in situ: (a) utilizando los coeficientes originales (JM), (b) utilizando los coeficientes ofrecidos por Meng et al. (2019) por rangos de W (JM_M_W). (Niclòs et al., 2021).*

Los resultados de la Figura 26 y la Tabla 15 muestran un funcionamiento preciso para el algoritmo SW de Jiménez-Muñoz et al. (2014) (JM), con una incertidumbre sistemática despreciable y una RMSD de 0.6 K. Los mismos valores de SD, pero con diferencias medias ligeramente superiores (0.5 K y 0. 6 K), resultan al utilizar los coeficientes proporcionados por Meng et al. (2019) para el algoritmo (JM_M_W y JM_M), con RMSD de 0.8 K. Por lo tanto, la modificación de los coeficientes de Meng et al. (2019) no introdujo mejoras para este algoritmo de SW en la zona de estudio.

Tabla 15. Resultados estadísticos (diferencias medias ± SD (RMSD)), de las diferencias entre LSTs obtenidas con los algoritmos SW (con datos de después del reprocesado de 2017) y LSTs in situ. (Niclòs et al., 2021).

Algoritmo SW	Versión	LST SW – LST in situ (K)
(Jiménez-Muñoz et al., 2014)	JM	0.1 ± 0.6 (0.6)
	JM_M_W	0.5 ± 0.6 (0.8)
	JM_M	0.6 ± 0.6 (0.8)
(Wan and Dozier, 1996) con el refinamiento de (Wan, 2014)	W&D_M_W	0.4 ± 0.7 (0.8)
	W&D_M	0.7 ± 0.7 (1.0)
	W&D_D_W	1.6 ± 0.6 (1.8)
	W&D_D	1.7 ± 0.6 (1.8)
	W&D_D_W_T	0.6 ± 0.9 (1.1)
	W&D_D_G	0.4 ± 0.6 (0.8)
NOAA JPSS algoritmo Enterprise (Meng et al., 2019)	Ent_M_W	0.5 ± 0.7 (0.8)
	Ent_M	0.8 ± 0.6 (1.0)
(Yu et al., 2014)	Yu_SIM	1.1 ± 0.7 (1.3)
	Yu_US	0.8 ± 0.7 (1.1)
	Yu_MLS	2.1 ± 1.2 (2.4)
(Rozenstein et al., 2014)	Ro_SIM	1.1 ± 0.7 (1.3)
	Ro_US	0.6 ± 0.8 (1.0)
	Ro_MLS	0.4 ± 0.8 (0.9)

Meng et al. (2019) validaron los resultados de JM_M_W utilizando datos de varios sitios SURFRAD como referencia, y obtuvieron unas diferencias medias de 2 K y una RMSD de 3 K. Guo et al. (2020) validaron el

algoritmo JM utilizando cinco sitios SURFRAD y obtuvieron unas diferencias medias de 1.4 K y una RMSD de 2.5 K. También proporcionaron resultados después de eliminar los datos de una estación del análisis y mostraron unas diferencias de 1.2 K y una RMSD de 2.3 K. Meng et al. (2019) también utilizaron los datos adquiridos por un radiómetro IRT en 8-14 µm (IR120, Campbell Scientific) instalado en una torre fija en una zona de hierba silvestre (García-Santos et al., 2018), obteniendo diferencias (RMSDs) de -0.4 K (1.6 K) para el algoritmo JM y de -0.2 K (1.7 K) para el JM_M_W. Nuestros resultados en la zona experimental de los arrozales de Valencia (Tabla 15), utilizando las mediciones precisas del radiómetro CE-312-2 a lo largo de transectos sobre diferentes cubiertas terrestres estacionales, muestran resultados mucho mejores para estos algoritmos SW, con los mejores resultados para la versión JM.

La Tabla 15 también muestra los resultados para el algoritmo SW generalizado propuesto por Wan y Dozier (1996), y refinado por Wan (2014), con los coeficientes para el TIRS estimados por Meng et al. (2019) por rangos de W (W&D_M_W) y para cualquier W entre 0 y 7 cm (W&D_M). La Figura 27a muestra gráficamente los resultados de W&D_M_W. La Tabla 15 muestra también los resultados para el mismo algoritmo, pero con los coeficientes estimados previamente en Du et al. (2015), tanto por rangos de W (W&D_D_W) como para cualquier W entre 0.0 y 6.3 cm (W&D_D), utilizando una colección de perfiles atmosféricos distinta a la utilizada por Meng et al. (2019). La Figura 27b muestra el funcionamiento de W&D_D_W. Los resultados del mismo algoritmo, pero con los coeficientes no dependientes de W propuestos por Gerace et al. (2020) (W&D_G) y modificados por Guo et al. (2020) por subrangos de W y temperatura (W&D_D_W_T) también se incluyen en la Tabla 15 y en la Figura 27 c-d. Las incertidumbres estimadas para este algoritmo SW fueron en promedio de 1.0 ± 0.3 K (0.7-1.6 K) y 1.02 ± 0.06 K (0.9-1.2 K) para las versiones W&D_M_W y W&D_M, respectivamente, y de 1.1 ± 0.3 K (0.8-1.8 K), 1.17 ± 0.11 K (1.0-1.5 K), 1.1 ± 0.3 K (0.8-1.7 K) y 1.07 ± 0.13 K (0.9-1.5 K) para las versiones W&D_D_W, W&D_D, W&D_D_W_T y W&D_G, respectivamente.

Los resultados de la Tabla 15 muestran que los algoritmos basados en el de Wan y Dozier (1996) con coeficientes constantes (independientes de W y de la temperatura) proporcionaron los mejores resultados para nuestra zona de validación. Este es el caso del algoritmo propuesto por Gerace et al. (2020) para el futuro producto LST de SW de la Colección 3.

Meng et al. (2019) mostraron resultados de validación para el W&D_M_W utilizando datos SURFRAD como referencia. Se obtuvo una diferencia media de 1.9 K y una RMSD de 3 K. Sin embargo, al utilizar las LST in situ medidas con radiómetro IRT de García-Santos et al. (2018), se obtuvieron las siguientes diferencias medias (RMSDs): 1.4 K (2 K) para el W&D_D_W y -0.2 K (1.7 K) para el W&D_M_W. Así pues, la sobreestimación del algoritmo W&D_D_W observada en la Tabla 15 concuerda con los resultados observados anteriormente. Sin embargo, nuestros datos de validación muestran mejores resultados para el W&D_M_W que en los trabajos anteriores, siendo cercanos a los obtenidos con el algoritmo JM.

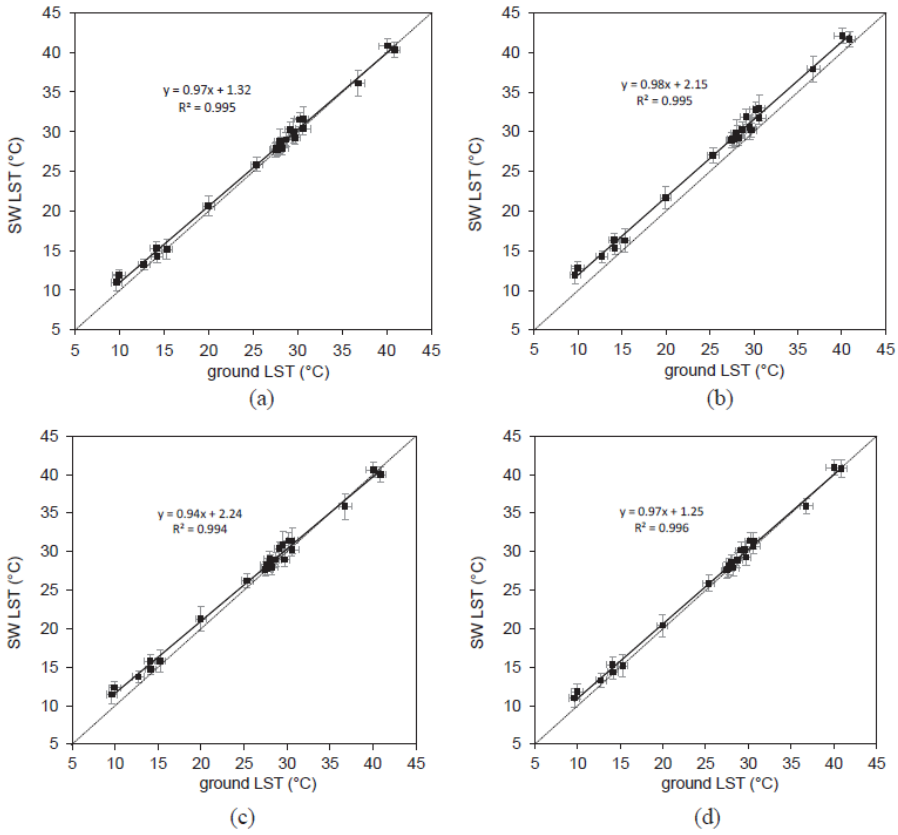

Figura 27. *LST del algoritmo SW (Wan and Dozier, 1996) (refinado por (Wan, 2014)) frente a LST in situ cuando se utilizan coeficientes estimados para TIRS por: (a) Meng et al. (2019) por rangos de W (W&D_M_W); (b) Du et al. (2015) por rangos de W (W&D_D_W); (c) Guo et al. (2020) por rango de W y rangos de temperaturas (W&D_D_W_T); (d) (Gerace et al., 2020) (W&D_G). (Niclòs et al., 2021).*

Guo et al. (2020) evaluaron el algoritmo de Du et al. (2015) con coeficientes por subrangos de W y por subrangos de temperatura (W&D_D_W_T) utilizando datos de cinco emplazamientos SURFRAD como referencia, mostrando unas diferencias medias de 2 K y una RMSD cercano a 3 K. Excluyendo los datos de un emplazamiento, tanto las diferencias medias como el RMSD se redujeron en 0.1 K. Nuestros resultados sobre la zona de los arrozales fueron mucho mejores para la versión W&D_D_W_T, probablemente debido a la mayor precisión de nuestros datos in situ, al ser medidos por radiómetros IRT.

Gerace et al. (2020) evaluaron los algoritmos W&D_D y W&D_G sobre tres emplazamientos SURFRAD y un emplazamiento FLUXNET (que también usa pirgeómetros para la medida de la LST in situ). Se obtuvo una diferencia media (SD) de -1.3 K (2.4 K) para el algoritmo W&D_D, y de -0.4 (2.1 K) para la versión W&D_G cuando la distancia a la nube más cercana era >10 km. Dicha condición se cumple en el presente estudio para todos los casos en los que se ha tomado medidas in situ a lo largo de transectos, puesto que solo se midió en condiciones de cielos despejados.

Gerace et al. (2020) también utilizaron medidas de la temperatura de la superficie del agua de treinta y seis boyas de la NOAA en las proximidades de la costa de EE.UU. y las boyas del Jet Propulsion Laboratory (NASA) en el lago Tahoe y el mar de Salton como sitios de referencia. En este caso, se obtuvo una diferencia media (SD) de 0.0 K (1.5 K) para el algoritmo W&D_D, y de 0.2 (0.8 K) para la versión W&D_G. A la vista de estos resultados, Gerace et al. (2020) indicaron las limitaciones mencionadas de utilizar datos de pirgeómetros como referencia para las validaciones. Los resultados de la Tabla 15 para el algoritmo W&D_G concuerdan con los mencionados anteriormente, lo que demuestra el buen funcionamiento de este algoritmo también en el caso de superficies terrestres.

Además, la Tabla 15 muestra los resultados para el algoritmo SW Enterprise con coeficientes derivados para TIRS por Meng et al. (2019), tanto por rangos de W (Ent_M_W) como para cualquier valor de W de 0 a 7 cm (Ent_M). La Figura 28a muestra los resultados de Ent_M_W frente a los datos in situ. Las incertidumbres del algoritmo Enterprise SW estimadas mediante propagación de errores fueron 1.1 ± 0.3 K (0.8-1.6 K) y 1.3 ± 0.09 K (1.2-1.6 K) para las versiones Ent_M_W y Ent_M, respectivamente.

Meng et al. (2019) también mostraron resultados de validación para el algoritmo Ent_M_W al utilizar datos SURFRAD como referencia. Se obtuvieron unas diferencias medias (RMSD) de 1.9 K (3 K). Sin embargo, al utilizar los datos de García-Santos et al. (2018) como referencia, obtuvieron unas diferencias medias (RMSD) de 0.1 K (1.7 K). La Tabla 15 muestra mejores resultados para este algoritmo en la zona de los arrozales, similares a los proporcionados por los algoritmos W&D_M_W y W&D_G y las diferentes versiones del algoritmo de Jiménez-Muñoz et al. (2014).

La Tabla 15 muestra también los resultados de los algoritmos SW propuestos por Yu et al. (2014) (Yu) y Rozenstein et al. (2014) (Ro). En estos algoritmos se utilizan como datos de entrada las transmisividades atmosféricas para las bandas TIRS, que se estiman a partir de W utilizando regresiones lineales. Estas regresiones se proporcionaron para dos atmósferas estándar: US Standard (US) de 1976 y Mid-Latitude Summer (MLS). Usamos transmisividades simuladas usando MODTRAN 5 para cada medida (SIM), pero también las correspondientes regresiones US y MLS para evaluar el efecto de usarlas en los resultados LST. Cabe destacar que estos dos algoritmos son mucho más complicados de aplicar que todos los anteriores algoritmos SW clásicos.

La Figura 28b muestra los resultados del algoritmo Yu con las transmisividades simuladas. Obtuvimos resultados similares para el algoritmo Ro equivalente (no se muestra gráficamente). Las incertidumbres estimadas mediante propagación de errores para los algoritmos Yu y Ro fueron 1.1 ± 0.4 K (0.8-2.2 K) y 0.9 ± 0.3 K (0.7-1.7 K), respectivamente.

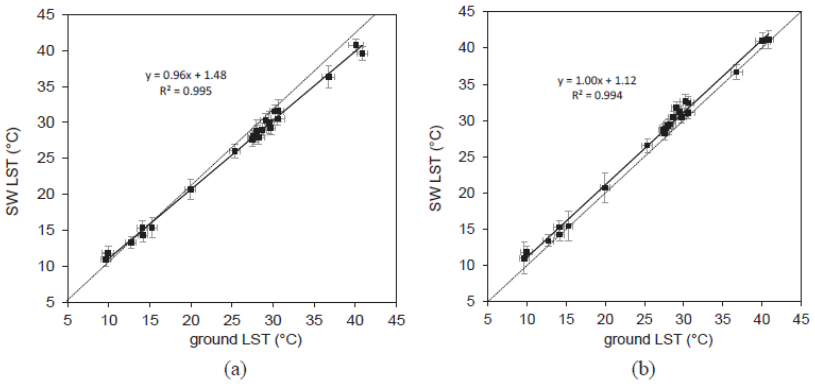

Figura 28. *LSTs de los SW frente a los datos in situ: (a) Algortimo Enterprise propuesto por Meng et al. (2019) con los coeficientes para rangos de W (Ent_M_W); y (b) algortimo propuesto por Yu et al. (2014) (Yu) usando las transmisividades simuladas como datos de entrada (Yu_SIM). (Niclòs et al., 2021).*

Los resultados fueron muy similares para ambos algoritmos (excepto en el caso de las estimaciones de transmisividad MLS), como era de esperar ya que se basaban en la misma técnica SW propuesta por Qin et al. (2001). Yu et al. (2014) mostraron resultados de validación para su algoritmo utilizando datos de cuatro estaciones SURFRAD como referencia. Utilizaron de 8 a 12 pares de datos TIRS y SURFRAD para cada estación, y se obtuvieron RMSD de 0.7 K a 1.3 K, con diferencias medias inferiores a 0.3 K. Sin embargo, para esta validación se utilizaron datos TIRS anteriores al reprocesado de 2017, y nosotros proporcionamos una validación para el algoritmo aplicado sobre datos TIRS actuales.

En Rozenstein et al. (2014) no se proporcionaron resultados de validación para su algoritmo. Guo et al. (2020) compararon los resultados de este algoritmo con datos de cinco emplazamientos SURFRAD. Sin embargo, el signo negativo en una de las ecuaciones del SW, que fue publicado como corrección en el algoritmo por los propios autores, no fue considerado por Guo et al. (2020), y esto hizo que obtuvieran unas diferencias medias (RMSD) de -0.09 K (2.5 K) al utilizar los datos de los cinco emplazamientos SURFRAD, y de 0.3 K (2.3 K) al excluir uno de los emplazamientos. A la vista de estos resultados, señalaron este algoritmo como uno de los mejores algoritmos SW para datos TIRS tras el reprocesado de 2017. Si utilizáramos un signo positivo en la ecuación como en Guo et al. (2020), se obtendría unas diferencias medias de -1.8 K, una SD de 1.9 K y una RMSD de 2.6 K, en lugar de los valores mostrados en la Tabla 15 (Ro_SIM), lo que llevaría a resultados muy diferentes para los algoritmos Ro y Yu.

A la vista de nuestros resultados, los algoritmos de Yu et al. (2014) y Rozenstein et al. (2014) proporcionaron las mayores diferencias medias, pero con SD similares a los otros algoritmos SW cuando se utilizaron estimaciones de transmisividad SIM y US (se obtuvieron SD mayores cuando se utilizaron estimaciones de transmisividad MLS en el algoritmo de Yu). Las RMSD de estos dos algoritmos son mayores que las obtenidas para los algoritmos JM, W&D_M_W, W&D_G y Ent_M_W para nuestra zona experimental. Por tanto, teniendo en cuenta este hecho, la mayor complejidad en sus aplicaciones y las estimaciones

requeridas de transmisividades atmosféricas, son los algoritmos SW menos recomendados para los usuarios de los datos TIRS tras el reprocesado de 2019.

Por último, también evaluamos la actualización de la calibración de 2019 en términos de los resultados de LST al utilizar estos diecisiete algoritmos SW. Los diferentes cambios de respuesta en cada banda TIRS debidos a esta actualización de la calibración, con cambios mayores para la banda 11 que para la banda 10, produjeron principalmente aumentos significativos en las diferencias medias (entre 1.3 K y 1.9 K para todos los algoritmos, siendo 1.4 K para JM, W&D_M_W, W&D_G y Ent_M_W) respecto al procesado de 2017. Sin embargo, no se observaron diferencias significativas en cuanto a las SD obtenidas para los algoritmos SW, con diferencias de SD que oscilan entre -0.19 K y 0.18 K para todos ellos (-0.14 K, -0.16 K, -0.14 K y -0.09 K para los algoritmos JM, W&D_M_W, W&D_G y Ent_M_W, respectivamente).

Por lo tanto, los algoritmos SW que mejores resultados mostraron cuando se aplicaron a los datos TIRS de la Colección 1 proporcionan una diferencia negativa en la zona de los arrozales con la recalibración de 2019.

c. Resultados para el sensor TIRS-2

En Niclòs et al., (2023) llevamos a cabo una evaluación de la calibración de las bandas 10 y 11 del sensor TIRS-2 y la validación de diferentes algoritmos SW propuestos por diversos autores, incluyendo los estudiados en el apartado anterior que también pueden ser aplicados a TIRS-2 debido a la semejanza que tienen las bandas de ambos sensores, además de nuevos (Wang et al., 2023). Aquí, realizamos un estudio con los datos antes y después del reprocesado de marzo de 2023. Todas las medidas utilizadas como referencia para este sensor fueron de nuevo tomadas mediante transectos sobre la superficie de los arrozales de Valencia, siguiendo la misma metodología que para el sensor TIRS.

c.1. Calibración de las bandas TIR del sensor TIRS-2

La evaluación de la calibración se realizó en términos de radiancias TOA (L_i) y temperaturas de brillo (T_i) en cada banda. Además de la diferencia media (bias), SD, y RMSD, realizamos tres pruebas estadísticas diferentes con el fin de reforzar los resultados estadísticos. La primera prueba fue el análisis de la varianza (ANOVA), utilizada para determinar si las temperaturas y las radiancias de ambos grupos, datos de satélite e in situ, tienen una media común (función «anova1» en Matlab). En segundo lugar, se utilizó una prueba T-test pareada, en la que la hipótesis nula era que las medias de los datos de satélite y de tierra eran iguales. Se utilizó la función «ttest2» de Matlab, con un nivel de significación del 5 % para rechazar la hipótesis nula. Por último, se evaluó la prueba de Levene en la que la hipótesis nula era que ambos grupos tenían varianzas iguales (función «vartestn» en Matlab utilizando la opción «LeveneAbsolute»).

La Tabla 16 muestra los valores de error medio, SD, RMSD, y los valores p para las pruebas mencionadas, para: a) el valor único del píxel que cubre la zona experimental, y b) el conjunto de 3×3 píxeles de 100 m y c) el conjunto de 5×5 píxeles de 100 m de TIRS-2 sobre la zona. La Tabla 16 muestra resultados semejantes para

todos ellos (con diferencias insignificantes, < 0.005 W/m^2·sr·µm y < 0.03 K), lo que demostró de nuevo la homogeneidad de la zona. Dado que los datos de OC (calibrado original) dejaron de proporcionarse tras el reprocesado de marzo de 2023, se evaluaron de diciembre de 2021 a febrero de 2023, con un total de 17 casos sin nubes. Los resultados de los datos RP (reprocesado de 2023) se muestran para los mismos datos que la calibración original, y para el periodo de diciembre de 2021 a agosto de 2023, con un total de 23 casos sin nubes.

Tabla 16. Resultados estadísticos de las diferencias de L_i y T_i (datos de satélite menos datos in situ) para las dos bandas TIRS-2, ambas con datos OC y RP, para 1 píxel (1p), 3×3 píxeles y 5×5 píxeles. (Niclòs et al., 2023).

		Datos OC (N = 17)			**Datos RP (N = 17)**			**Datos RP (N = 23)**		
		1p	**3×3**	**5×5**	**1p**	**3×3**	**5×5**	**1p**	**3×3**	**5×5**
L_{10}	**Bias (10^{-2}W/m^{2}srµm)**	-4	-4	-4	-4	-4	-4	-4	-4	-4
	SD (10^{-2}W/m^{2}srµm)	5	5	5	5	5	5	6	6	6
	RMSD (10^{-2}W/m^{2}srµm)	6	7	7	7	7	7	7	7	7
	ANOVA p	0.93	0.93	0.93	0.93	0.92	0.92	0.92	0.92	0.91
	T-test p	0.93	0.93	0.93	0.92	0.92	0.91	0.92	0.92	0.91
	Levene p	0.97	0.98	0.98	0.98	0.98	0.98	0.91	0.92	0.93
L_{11}	**Bias (10^{-2}W/m^{2}srµm)**	-1	-1	-1	3	2	2	2	2	2
	SD (10^{-2}W/m^{2}srµm)	6	6	6	6	6	7	7	7	7
	RMSD (10^{-2}W/m^{2}srµm)	7	7	7	7	7	7	7	7	7
	ANOVA p	0.97	0.97	0.97	0.93	0.94	0.94	0.93	0.93	0.94
	T-test p	0.97	0.96	0.96	0.93	0.93	0.94	0.93	0.93	0.94
	Levene p	0.89	0.90	0.89	0.89	0.90	0.90	0.91	0.92	0.93
T_{10}	**Bias (K)**	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3
	SD (K)	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
	RMSD (K)	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
	ANOVA p	0.93	0.93	0.93	0.92	0.92	0.92	0.92	0.92	0.91
	T-test p	0.93	0.93	0.93	0.92	0.92	0.91	0.92	0.92	0.91
	Levene p	0.96	0.97	0.96	0.97	0.97	0.97	0.91	0.93	0.94
T_{11}	**Bias (K)**	-0.1	-0.1	-0.1	0.2	0.2	0.2	0.2	0.2	0.2
	SD (K)	0.6	0.6	0.6	0.5	0.5	0.5	0.6	0.6	0.6
	RMSD (K)	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
	ANOVA p	0.97	0.96	0.96	0.94	0.94	0.94	0.93	0.93	0.94
	T-test p	0.97	0.96	0.96	0.93	0.93	0.94	0.93	0.93	0.93
	Levene p	0.88	0.89	0.89	0.90	0.91	0.90	0.93	0.94	0.94

En la Tabla 16, los valores p de las tres pruebas de hipótesis estadísticas muestran que la hipótesis nula no puede rechazarse en ningún caso, dados los elevados valores p que se muestran. Este hecho refuerza la hipótesis de que los datos de L_i y T_i de TIRS-2 y los obtenidos a partir de medidas en superficie son equivalentes en términos de medias y varianzas. Además, los estadísticos F del ANOVA (es decir, la varianza entre grupos dividida por la varianza dentro de cada grupo) eran próximos a cero (≤0.01 para todos ellos).

La Figura 29a muestra las BTs obtenidas a partir de datos in situ frente a las del satélite (valores medios para la matriz de 5×5 100-m píxeles) con datos OC para las dos bandas. La Figura 29b muestra las mismas comparaciones, pero para datos RP. En estas figuras, las barras de error de T_i muestran las incertidumbres, que oscilan entre 0.06 y 0.5 K para ambas bandas. Como se muestra en la Tabla 16 y en la Figura 29a, los datos OC de L_i y T_i de TIRS-2 coincidieron con los estimados a partir de datos in situ, con ligeras diferencias medias negativas (-0.3 K para b10 y -0.1 K para b11) y RMSDs de 0.5 K y 0.6 K para las bandas 10 y 11

respectivamente. La subestimación de TIRS-2 también fue observada por Barsi et al. (2022) utilizando datos medidos en zonas de agua como referencia para la evaluación de datos de L9 hasta abril de 2022. En ese caso, las diferencias medias fueron de -0.2 K para la banda 10 y -0.7 K para la banda 11.

En la Tabla 16 no se observa ningún cambio apreciable en la diferencia media de la banda 10 (de alrededor de -0.01 K) entre los datos de RP y OC. Sin embargo, se encontró un cambio significativo de 0.3 K en la banda 11. Por lo tanto, el reprocesado de 2023 introdujo un cambio de calibración mayor para la banda 11, aumentando la diferencia media entre ambas bandas (de -0.2 K para los datos OC a -0.6 K para los datos RP), con posibles consecuencias al aplicar los algoritmos SW. Los resultados de los datos RP son similares para el total de pares de datos (23) y los 17 datos coincidentes con OC. Al igual que la Tabla 16, la Figura 29b revela el cambio en la diferencia media entre las dos bandas para los datos RP, no observada en la Figura 29a para los datos OC.

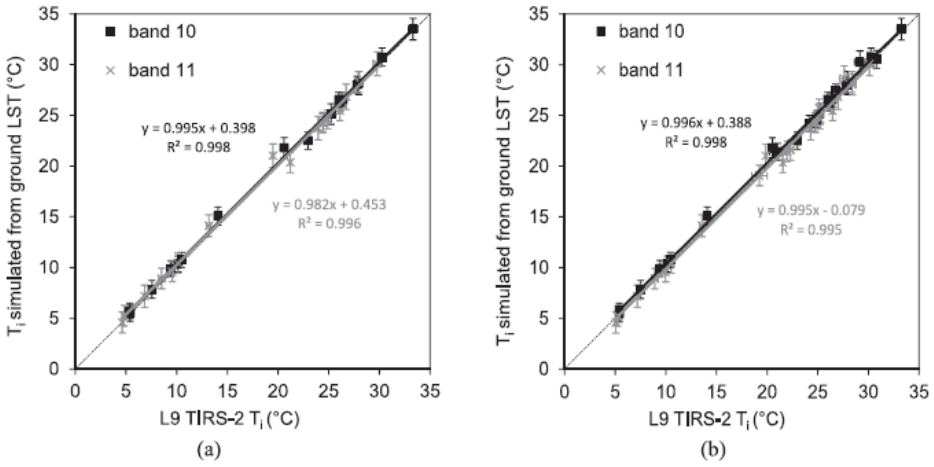

Figura 29. *BTs obtenidas a partir de datos in situ frente a los valores de TIRS-2: (a) OC (17 coincidencias); y (b) RP (23 coincidencias). Las regresiones lineales se muestran como líneas continuas. Se incluye también las ecuaciones de regresión y los coeficientes de determinación. La línea de puntos es la línea 1:1. (Niclòs et al., 2023).*

c.2. Validación con el método t-based

Por otro lado, evaluamos los valores de LST obtenidos mediante distintos algoritmos aplicados a datos TIRS-2. La Tabla 17 muestra los resultados de la validación del método SC (ecuación 2.4) a las bandas 10 y 11 por separado para el conjunto de 5×5 píxeles de 100 m. Obtuvimos resultados similares para el conjunto de 3×3 píxeles de 100 m y el píxel único que cubre la zona, con diferencias insignificantes entre ellos (<0.04 K), de forma equivalente a los mostrados en la Tabla 16. Por lo tanto, el análisis de los resultados se muestra sólo para la matriz de 5×5 píxeles de 100 m en lo sucesivo. La Figura 30 muestra las LSTs in situ frente a las LSTs obtenidas a partir del método SC.

Tabla 17. Resultados estadísticos de las diferencias de LST (satélite menos datos in situ, en K) al utilizar el método SC a cada banda, tanto con datos OC como datos RP.

	Datos OC (N = 17)			**Datos RP (N = 17)**			**Datos RP (N = 23)**		
	Bias	**SD**	**RMSD**	**Bias**	**SD**	**RMSD**	**Bias**	**SD**	**RMSD**
banda 10	-0.4	0.5	0.6	-0.4	0.5	0.6	-0.4	0.5	0.6
banda 11	-0.1	0.8	0.8	0.4	0.8	0.9	0.3	0.9	0.9

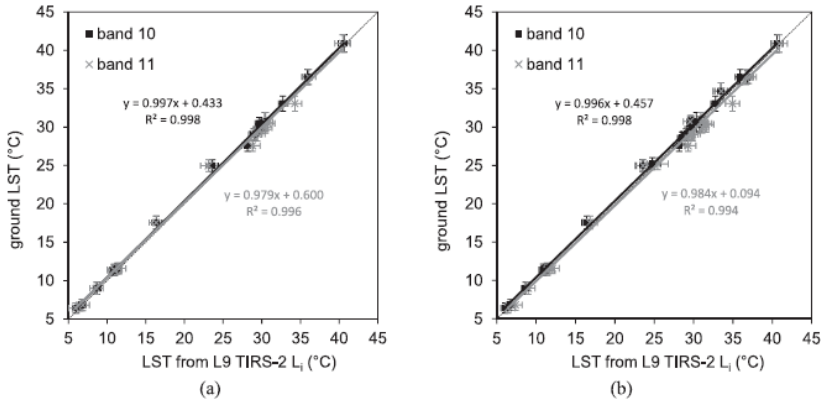

Figura 30. *LSTs in situ frente a LSTs obtenidas con el método SC, (a) datos OC y (b) datos RP. (Niclòs et al., 2023).*

La Tabla 17 y la Figura 30 muestran de nuevo la buena relación en términos de LSTs estimadas usando datos OC, con ligeras subestimaciones de -0.4 K y -0.1 K y RMSDs de 0.6 K y 0.8 K para las bandas 10 y 11, respectivamente. Se muestra una significativa disparidad en las diferencias medias entre ambas bandas cuando se utilizan datos RP (-0.4 K para b10 y 0.4 K para b11).

La Tabla 18 incluye los resultados de validación para el producto L2 de LST generado a partir de los datos de la banda 10, y para el algoritmo SC propuesto por Jiménez-Muñoz et al. (2014) aplicado a las bandas 10 y 11 por separado. La Figura 31 muestra la comparación de las LSTs in situ con los valores LST del producto L2.

Tabla 18. Resultados estadísticos de las diferencias entre el producto L2 LST y los datos LST in situ para datos OC y RP. También se muestran los resultados para el SW alternativo de SC_JM_10 y SC_JM_11.

	Datos OC (N = 17)			**Datos RP (N = 17)**			**Datos RP (N = 23)**		
	Bias	SD	RMSD	Bias	SD	RMSD	Bias	SD	RMSD
LST_{L2} (K)	0.4	1.0	1.1	0.4	1.0	1.1	0.5	1.1	1.2
LST_{sc_JM10} (K)	-0.7	0.9	1.2	-0.7	1.0	1.2	-0.9	1.1	1.4
LST_{sc_JM11} (K)	1.1	1.9	2.2	1.6	2.0	2.5	1.0	2.2	2.5

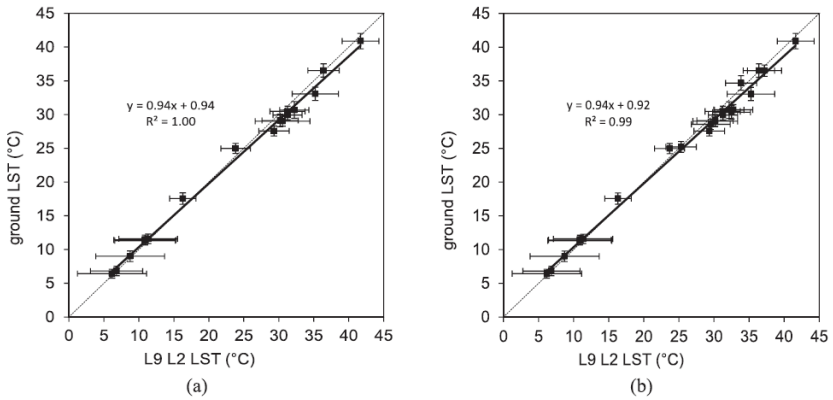

Figura 31. *LST in situ frente a LST L2 cuando se utilizan: (a) datos OC; y (b) datos RP. (Niclòs et al., 2023).*

El producto L2 proporciona la incertidumbre de la LST para cada píxel, que incluye las diferentes fuentes de error (las procedentes de la emisividad, los parámetros atmosféricos, etc.) y el efecto de la distancia a la nube más cercana (Malakar et al., 2018). Para esta última, Laraby et al. (2016) estudiaron la relación entre la proximidad de las nubes y los errores de LST obtenidos utilizando datos de boyas como referencia. Se observa que el efecto de distancia de las nubes aumenta la incertidumbre de la LST proporcionada por el producto L2 de alrededor de 2 K (para una distancia de las nubes > 10 km) a más de 5 K (para una distancia de las nubes < 1 km). Algunos pares de datos medidos en nuestra zona experimental se vieron afectados por esta última condición (véase la Figura 31), mientras que las mediciones in situ se llevaron a cabo únicamente para condiciones libres de nubes al menos a una distancia de 10 km del lugar. Por lo tanto, concluimos que las incertidumbres del producto LST se sobreestimaron en gran medida para algunas medidas en nuestra zona debido a la detección errónea de nubes.

La Tabla 18 muestra una ligera y similar sobreestimación de la LST de 0.4-0.5 K para el producto L2 tanto con datos OC como RP, aumentando la sobreestimación con la temperatura (Figura 31). Sin embargo, los resultados mostrados en la Tabla 17 para el método SC con la banda 10 revelan una subestimación (-0.4 K). Los resultados de validación para el algoritmo SC de Jiménez-Muñoz et al. (2014) (SC_JM_10) aplicado a la banda 10, con valores de emisividad medidos en el lugar y valores de W estimados para cada dato, muestran también una subestimación de alrededor de -0.7 K.

Las emisividades incluidas en el producto L2 de LST se aproximan a las medidas in situ en nuestra zona para las condiciones de suelo inundado y suelo desnudo (con pequeñas diferencias de hasta 0.003). Sin embargo, no fue así para las condiciones de cobertura vegetal completa, ya que las emisividades del producto fueron inferiores a las medidas in situ en 0.010-0.013. La subestimación de la emisividad explica en parte las diferencias positivas sistemáticas (incluso superiores a 1 K) observadas para esta cobertura, que sesga los resultados globales de L2 hacia unas diferencias positivas. Además, se observaron valores inesperadamente demasiado bajos de radiancia atmosférica descendente que incluye y utiliza el producto L2, que eran de

alrededor del 50 % de las radiancias ascendentes también incluidas. Este hecho también podría explicar en parte las diferencias positivas.

En cualquier caso, obtuvimos RMSD de 1.1-1.2 K para el producto L2, que son muy inferiores a los obtenidos en Meng et al. (2022) y Ye et al. (2022) con medidas in situ de pirgeómetros. Por otro lado, los algoritmos alternativos SC_JM_10 y SC_JM_11 proporcionaron incertidumbres mayores que el método SC (Tabla 17) y que el producto L2 LST (Tabla 18). Además, los resultados para el SC_JM_11, con coeficientes dados en Guo et al. (2020), mostraron RMSD 1 K mayores que los del SC_JM_10, por lo que no se recomienda el uso del SC_JM_11.

La Tabla 19 resume las validaciones correspondientes a los 6 tipos de algoritmos SW, y sus variantes, cuando se aplican a los datos OC y RP. Para el análisis de los resultados se utilizó como referencia los mismos datos LST in situ. La Figura 32 muestra un diagrama de cajas en el que se comparan los resultados de los distintos algoritmos.

Tabla 19. Resultados estadísticos de las diferencias en LST (datos de satélite menos in situ) obtenidos mediante algoritmos SW para TIRS-2, tanto de datos OC como de datos RP. (Niclòs et al., 2023).

		LST SW – LST in situ Bias ± SD (RMSD) (K)	
Algoritmo SW	**Versión**	Datos OC (N = 17)	Datos RP (N = 23)
(Jiménez-Muñoz et al., 2014)	JM	-0.7 ± 0.7 (1.0)	-1.3 ± 0.7 (1.4)
	JM_M_W	-0.4 ± 0.7 (0.8)	-1.0 ± 0.8 (1.2)
	JM_M	-0.4 ± 0.9 (1.0)	-1.0 ± 0.9 (1.3)
(Wan and Dozier, 1996) con el refinamiento de (Wan, 2014)	W&D_M_W	-0.3 ± 0.7 (0.7)	-0.8 ± 0.7 (1.1)
	W&D_M	-0.1 ± 0.7 (0.7)	-0.6 ± 0.6 (0.9)
	W&D_D_W	0.8 ± 0.8 (1.1)	0.2 ± 0.8 (0.8)
	W&D_D	0.8 ± 0.9 (1.2)	0.3 ± 0.8 (0.9)
	W&D_D_W_T	0.0 ± 0.8 (0.8)	-0.6 ± 0.8 (1.0)
	W&D_D_G	-0.3 ± 0.7 (0.8)	-0.9 ± 0.7 (1.1)
NOAA JPSS algoritmo Enterprise (Meng et al., 2019)	Ent_M_W	-0.3 ± 0.7 (0.8)	-0.9 ± 0.7 (1.1)
	Ent_M	-0.2 ± 0.8 (0.8)	-0.8 ± 0.9 (1.2)
(Yu et al., 2014)	Yu_SIM	0.1 ± 0.7 (0.7)	-0.5 ± 0.7 (0.9)
	Yu_US	0.0 ± 0.6 (0.6)	-0.6 ± 0.7 (0.9)
	Yu_MLS	0.9 ± 1.3 (1.6)	0.2 ± 1.3 (1.3)
(Rozenstein et al., 2014)	Ro_SIM	0.1 ± 0.7 (0.7)	-0.5 ± 0.7 (0.9)
	Ro_US	-0.2 ± 0.7 (0.7)	-0.7 ± 0.7 (1.0)
	Ro_MLS	-0.3 ± 0.7 (0.8)	-0.9 ± 0.7 (1.1)
(Wang et al., 2023)	W_SIM	-0.3 ± 0.6 (0.6)	-0.8 ± 0.7 (1.0)
	W_REG	-1.3 ± 0.7 (1.5)	-1.8 ± 0.7 (2.0)

La Tabla 19 y la Figura 32 muestran una diferencia global de alrededor de -0.6 K en las diferencias medias obtenidas para los algoritmos SW entre los datos RP y OC, lo que conduce a mayores diferencias negativas para la mayoría de los algoritmos SW con datos RP. Al utilizar datos OC, las diferencias medias variaron entre -1.3 K y 0.9 K dependiendo del algoritmo, con una media de las diferencias de -0.08 K. Con datos RP se obtuvieron diferencias entre -1.8 K y 0.3 K, y la media fue de -0.7 K. Sin embargo, las SD son bastante similares para los datos OC y RP. Como hemos comentado anteriormente, el reprocesado aumenta la diferencia entre las incertidumbres sistemáticas en las temperaturas de brillo de las bandas 10 y 11 (de -0.2 K en los datos OC a -

0.6 K en los datos RP, Tabla 16), y conduce a un cambio significativo en las diferencias obtenidas al utilizar los algoritmos SW.

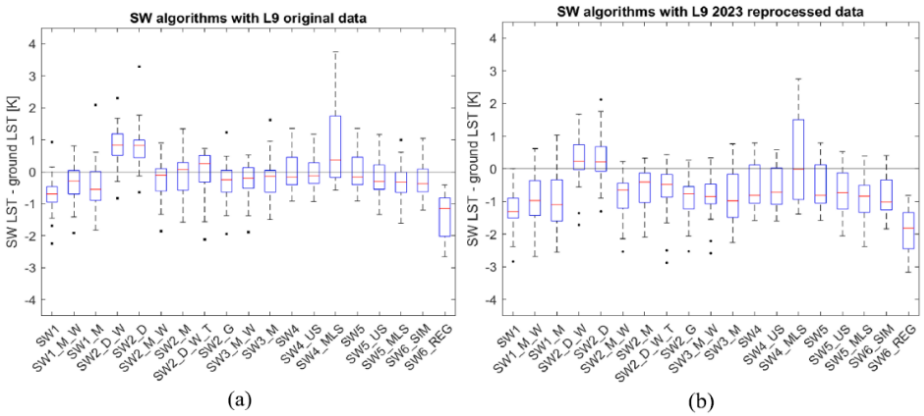

Figura 32. *Comparación de la validación del algoritmo SW para los datos TIRS-2: (a) Datos OC y (b) Datos RP. Los gráficos de cajas muestran los valores medios y los cuartiles primero y tercero. Los valores atípicos se muestran como puntos. Los bigotes corresponden aproximadamente al 99.3% de cobertura. Siendo SW1=JM, SW2=W&D, SW3=Ent, SW4=Yu, SW5=Ro y SW6=W. (Niclòs et al., 2023).*

Según los resultados al utilizar datos OC, varios algoritmos SW proporcionaron RMSD ≤ 1 K, como las diferentes versiones del algoritmo basado en el propuesto por Wan y Dozier (1996) y Wan (2014) (es decir, W&D_M_W, W&D _M, W&D _D_W_T y W&D _G). Además, también se encontraron RMSD ≤ 1 K para el resto de los tipos de algoritmos, excepto en el caso de usar regresiones entre τ_i y W para atmósferas de latitud media en Yu_MLS y en el caso de usar regresiones para τ_i y φ_i con W en W_REG.

Debido al cambio en las diferencias de la banda 11 introducido por el reprocesado de 2023, los RMSD de la mayoría de ellos superan 1 K con los datos de RP. Las diferencias aumentaron de -0.3 K a -0.9 K para el W&D_G propuesto para el futuro producto SW operativo de TIRS-2, y por tanto la RMSD aumentó de 0.8 K a 1.1 K. Tras el reprocesado de 2023, el algoritmo SW con la RMSD más baja es la versión W&D_D_W.

En cualquier caso, las diferencias medias y las SD de la mayoría de los algoritmos SW estudiados, tanto aplicados a datos OC como RP, se encuentran dentro del umbral recomendado de 1 K establecido por GCOS para los productos LST de satélite (GCOS, 2016).

d. Resultados para el sensor MODIS

Para este sensor en Coll et al., (2024) realizamos: (a) la validación de las LSTs de referencia obtenidas mediante el método r-based frente a LSTs medidas in situ en la superficie de los arrozales de Valencia, obteniendo así una estimación real de la incertidumbre del método. (b) La validación de los productos M*D11 y M*D21 mediante el método r-based para una gran base de datos sobre distintos tipos de superficies.

d.1. Validación del método r-based

Para evaluar la precisión de las LSTs de referencia del método r-based con respecto a LST in situ independientes utilizamos 65 pares de datos de medidas in situ simultáneas con las observaciones del sensor MODIS en Terra entre 2002 y 2022, y 21 pares de datos simultáneos con las observaciones del sensor MODIS Aqua entre 2016 y 2022 en la zona de los arrozales de Valencia. Todas las LST in situ fueron tomadas con varios radiómetros CIMEL312-2 a lo largo de transectos, en condiciones diurnas, sin nubes, con tiempo de sobrepaso entre 10:10 y 11:40 UTC para Terra y 12:40 y 13:45 UTC para Aqua. La mayoría de los datos corresponden a condiciones de cobertura vegetal completa (75%), pero hay unos pocos casos de agua (13%) y suelo desnudo (12%). Las medidas de campo se realizaron cerca del nadir, por lo que para validar el método nos limitamos a datos de MODIS medidos con el ángulo de observación cenital inferiores a 40º para evitar efectos de anisotropía térmica (Lagouarde et al., 2015).

Aplicamos el método r-based a los datos MODIS en las bandas 31 (11 μm) y 32 (12 μm) con las medidas de emisividad mostrada en la Tabla 7 y los perfiles atmosféricos correspondientes para calcular las LST para cada banda (T_{gi}) (ecuación 2.4). La Figura 33 representa la diferencia entre las LSTs in situ (T_g) y las LST de referencia r-based (T_{31g}) frente a la diferencia $\delta = T_{31g}-T_{32g}$ para todos los casos disponibles. Muestra una buena correlación entre ambas magnitudes, con grandes valores negativos (positivos) de δ que suelen corresponderse con grandes valores negativos (positivos) en T_g-T_{31g}, mientras que tenemos valores pequeños de T_g-T_{31g} para δ en torno a 0. La regresión lineal de la Figura 33 muestra un coeficiente de determinación aceptable ($R^2 = 0.54$) y una ordenada en el origen cercana a cero (0.15 K). La pendiente de la regresión lineal es cercana a 2, lo que sugiere que el rango en δ debería ser de ±0.5 K para mantener la incertidumbre de la LST dentro de ±1.0 K (comparable con las medidas in situ). Este umbral coincide con el considerado por Hulley et al. (2016) y Li et al. (2021) a partir del análisis teórico de sensibilidad.

En la Figura 33, unos pocos puntos con valores de δ próximos a cero tienen valores negativos de $T_g - T_{31g}$ superiores a -1.5 K. Tales casos (1 Terra y 4 Aqua) corresponden a condiciones de verano (atmósfera cálida y húmeda), con incertidumbres de la LST en tierra entre 1.0 y 1.2 K, por lo que las grandes diferencias entre las LST en tierra y r-based pueden ser compatibles con la incertidumbre combinada de $T_g - T_{31g}$.

Tomando sólo los casos que cumplen $-0.5\ K \leq \delta \leq 0.5\ K$, la diferencia media y el error cuadrático medio (RMSD) de $T_g - T_{31g}$ fueron respectivamente 0.0 y 0.8 K para Terra (48 casos válidos), -0.4 y 0.9 K para Aqua (19 casos válidos), y -0.1 y 0.8 K en total. Estos resultados muestran que la precisión que puede obtenerse con las LSTs r-based es comparable a la de las LSTs medidas in situ. Por lo tanto, las LSTs r-based se pueden utilizar como alternativa a las LSTs t-based para la validación cuando δ está dentro de ±0.5 K.

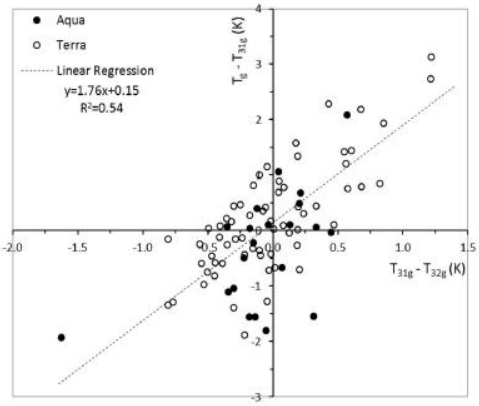

Figura 33. *Diferencia entre las LST medidas in situ y calculadas con el método r-based en la banda 31 frente a las diferencias de LST r-based en las bandas 31 y 32. Los casos de MODIS en Terra (Aqua) están marcados con círculos abiertos (cerrados). Se muestra también la regresión lineal y el coeficiente de determinación(R^2). (Coll et al., 2024).*

d.2. Validación r-based de los productos M*D11 y M*D21

Una vez comprobada la validez del método r-based, este se ha aplicado para la validación de los productos M*D11 y M*D21 de los satélites Terra y Aqua en seis zonas de validación en condiciones diurnas y nocturnas durante los años 2020 y 2021. La base de datos contiene un total de 4267 casos con aproximadamente el mismo número para Terra y Aqua, de día y de noche en cada zona (Tabla 20).

Tabla 20. Zonas para la validación r-based, número de escenas, emisividades de los componentes y su rango de fracción de cobertura vegetal (f). *suelo húmedo.

		Vegetación		Suelo desnudo		
Zona estudio	**Nº de datos**	**Banda 31**	**Banda 32**	**Banda 31**	**Banda 32**	**Rango f**
Albufera	380	N/D	N/D	N/D	N/D	N/D
Arrozales	397	0.985	0.980	0.970/0.972*	0.972/0.977*	0.03-1
Cortes	811	0.985	0.986	0.972	0.977	0.16-0.33
Soria	775	0.977	0.977	0.972	0.977	0.27-0.53
Viñedo	792	0.972	0.973	0.967	0.959	0.03-0.20
Olivar	1112	0.976	0.976	0.967	0.959	0.09-0.21

Una parte importante del método r-based son los valores de emisividad necesarios para derivar las LSTs de referencia. Para el presente estudio de validación r-based, hemos considerado dos opciones: (1) utilizar los datos de emisividad proporcionados por cada producto LST, es decir, la emisividad a priori basada en una clasificación de usos de suelo para M*D11 y la emisividad derivada con el método TES para M*D21. (2) Utilizar una base de datos de emisividad dinámica modelada para cada emplazamiento teniendo en cuenta la naturaleza del emplazamiento y las medidas de emisividad de campo o de laboratorio disponibles para el agua y las componentes de vegetación y suelo, y una estimación de la fracción de cobertura vegetal. Las emisividades del producto están estrechamente relacionadas con las LST proporcionadas por cada producto, ya sea porque se utilizan como entradas para la obtención de la LST en el producto M*D11 o porque se calculan

simultáneamente con la LST en el producto M*D21. Así pues, estos datos de emisividad pueden considerarse un caso favorable para la validación r-based. Por otra parte, dado que las emisividades modeladas son independientes de los productos, son útiles para comprobar el impacto de la emisividad en la validación.

Para las superficies parcialmente vegetadas, las emisividades se modelaron con el método de Valor y Caselles (1996) (ecuación 2.54) en términos de la fracción de cobertura de vegetación (f) y los valores de la emisividad de vegetación y suelo de la Tabla 20. Dado que MODIS no dispone de un producto operativo de fracción de cobertura vegetal, utilizamos los valores de f a 1 km que proporciona el producto LST de Sentinel-3 como datos auxiliares de su algoritmo. Las estimaciones de f de Sentinel-3 se basan en índices de vegetación obtenidos a partir de las bandas del rojo y del infrarrojo cercano (0.66 y 0.87 µm) del SLSTR y se proporcionan con una resolución temporal de 10 días adquiridos a partir de una ventana temporal móvil de compuestos de 30 días. Los valores de f para las fechas de observación de MODIS se obtuvieron por interpolación lineal a partir de las fechas más próximas de Sentinel-3 libres de nubes, normalmente en torno a los 3 días del paso de MODIS. La Tabla 20 muestra el rango aproximado de f para los zonas terrestres. En las zonas de agua, aplicamos la ecuación de Niclòs et al. (2009) (sección 2.f.4) para calcular la emisividad del agua en función del ángulo cenital de observación, de 0 a 65°, en las bandas 31 y 32.

La Figura 34 muestra una comparación de los valores de emisividad de los productos MOD11, MOD21 y la emisividad modelada en la banda 31 (resultados similares se obtuvieron para Aqua y la banda 32). Para la zona de la Albufera, las emisividades MOD11 y modeladas sólo dependen del ángulo cenital y concuerdan muy bien hasta los 40º; sin embargo, MOD11 parece sobreestimar la emisividad para ángulos cenitales mayores.

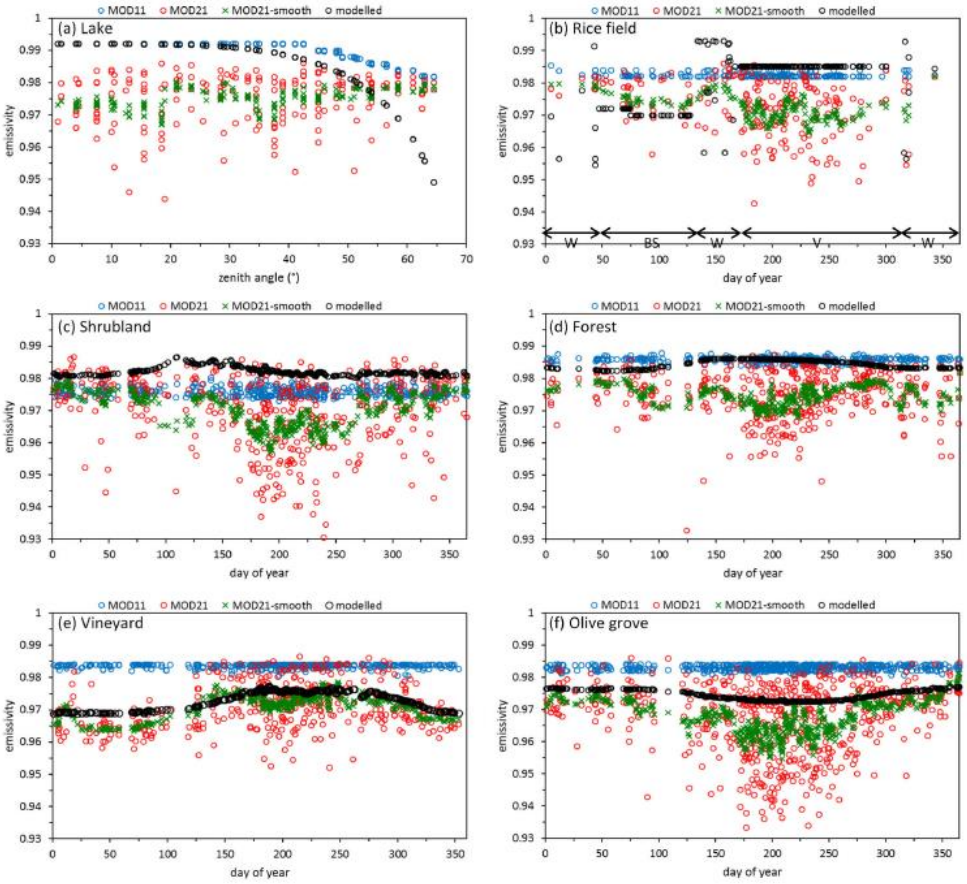

Figura 34. *Emisividad de la banda 31 de MODIS de MOD11 (azul), MOD21 (rojo), MOD21 suavizada con media móvil (verde) y modelada (negro). (a) agua en función del ángulo cenital de observación; (b) arrozales con agua (W), suelo desnudo (BS) y vegetación (V); (c) matorral; (d) bosque; (e) viñedo; (f) olivar, en función del día del año. (Coll et al., 2024).*

Para las zonas terrestres, las emisividades MOD11 y modeladas muestran una pequeña variabilidad durante el año. La emisividad modelada tiene en cuenta las diferentes etapas de la zona de los arrozales (agua, suelo desnudo y vegetación) y los cambios en la f para las demás zonas terrestres, lo que resulta en una variabilidad estacional relativamente pequeña con valores máximos en primavera-verano, mientras que MOD11 muestra valores de emisividad mayoritariamente constantes para cada zona. En general, la concordancia entre MOD11 y la emisividad modelada es razonable para las zonas con cobertura vegetal media a alta (arrozales en verano, matorral y bosque), encontrándose las mayores diferencias para los emplazamientos de viñedo y olivar con f baja. En estos casos, las emisividades MOD11 son más altas y cercanas a los valores de la vegetación, mientras que las emisividades modeladas se acercan más a las del suelo desnudo debido a la baja f. Por otra parte, las emisividades modeladas para la zona de Cortes (matorral perenne) son algo superiores a las emisividades MOD11, que se aproximan más a las del suelo desnudo.

Por el contrario, las emisividades MOD21 muestran valores inferiores a los datos de emisividad MOD11 y modelados y una gran variabilidad sin correlación con el ángulo cenital en la zona de la Albufera, ni con el día del año en los emplazamientos terrestres. La razón de la gran variabilidad es que las emisividades MOD21 se obtienen físicamente junto con la LST, por lo que varían de un caso a otro, ya que se ven afectadas por varias fuentes de incertidumbre, como el ruido instrumental, la corrección atmosférica y el error del algoritmo (Hulley et al., 2016). Por lo tanto, también se muestra en la Figura 34 el suavizado de las emisividades diarias MOD21 usando un promedio móvil sobre intervalos de 5º en ángulo de visión para la zona de la Albufera y de diez días para las zonas terrestres con el fin de igualar la resolución temporal de los datos f usados para las emisividades modeladas. Según la Figura 34, la variabilidad de la emisividad MOD21 parece ser mayor en primavera-verano (mayor humedad atmosférica) y los valores máximos suelen estar cerca de las emisividades MOD11 y modeladas. Las emisividades MOD21 suavizadas muestran una mayor coherencia con respecto a las demás bases de datos de emisividad y una dependencia estacional más clara que el producto diario MOD21.

El impacto de la emisividad en el método r-based puede evaluarse mediante un análisis de sensibilidad. La incertidumbre en T_{1g} (LST de la banda 31) correspondiente a una incertidumbre de 0.005 en ε_1 (emisividad de la banda 31) es de 0.37 K para las condiciones atmosféricas medias del conjunto de datos. Asumiendo la misma incertidumbre para ε_2 (emisividad de la banda 32) y que las incertidumbres en emisividad no están correlacionadas en ambas bandas, el impacto de la incertidumbre de emisividad por sí sola en δ es de 0.52 K. Esto muestra que la emisividad es un dato crucial, especialmente para el cálculo de δ, y que los resultados de la validación r-based dependen de la base de datos de emisividad utilizada.

El método r-based se aplicó al conjunto de datos de la Tabla 20 para la validación operativa de los productos LST de MODIS. Las Figuras 35 (MOD11) y 36 (MOD21) representan gráficamente la diferencia entre la LST del producto (T_{MOD11} o T_{MOD21}) y la LST r-based (T_{31g}) frente a la diferencia $\delta = T_{31g}-T_{32g}$ para cada zona de estudio (datos diurnos y nocturnos combinados). En cada caso, mostramos los resultados obtenidos a partir de la emisividad del producto y la emisividad modelada. Sólo se presentan los datos de Terra, siendo los resultados para Aqua similares en su mayoría. Las Figuras 35 y 36 muestran una relación entre las diferencias de temperatura similar a la mostrada en la Figura 33. Las mejores correlaciones se obtuvieron para MOD21 usando las emisividades del producto, donde los coeficientes de determinación (R^2) están en el rango 0.7-0.8 (Figura 36). Sin embargo, fueron inferiores (0.2-0.5) cuando se utilizan los datos de emisividad modelados. Para el producto MOD11 (Figura 35), los valores R^2 oscilaron aproximadamente entre 0.3 y 0.5 dependiendo de la zona, con una pequeña diferencia entre la emisividad del producto y la emisividad modelada.

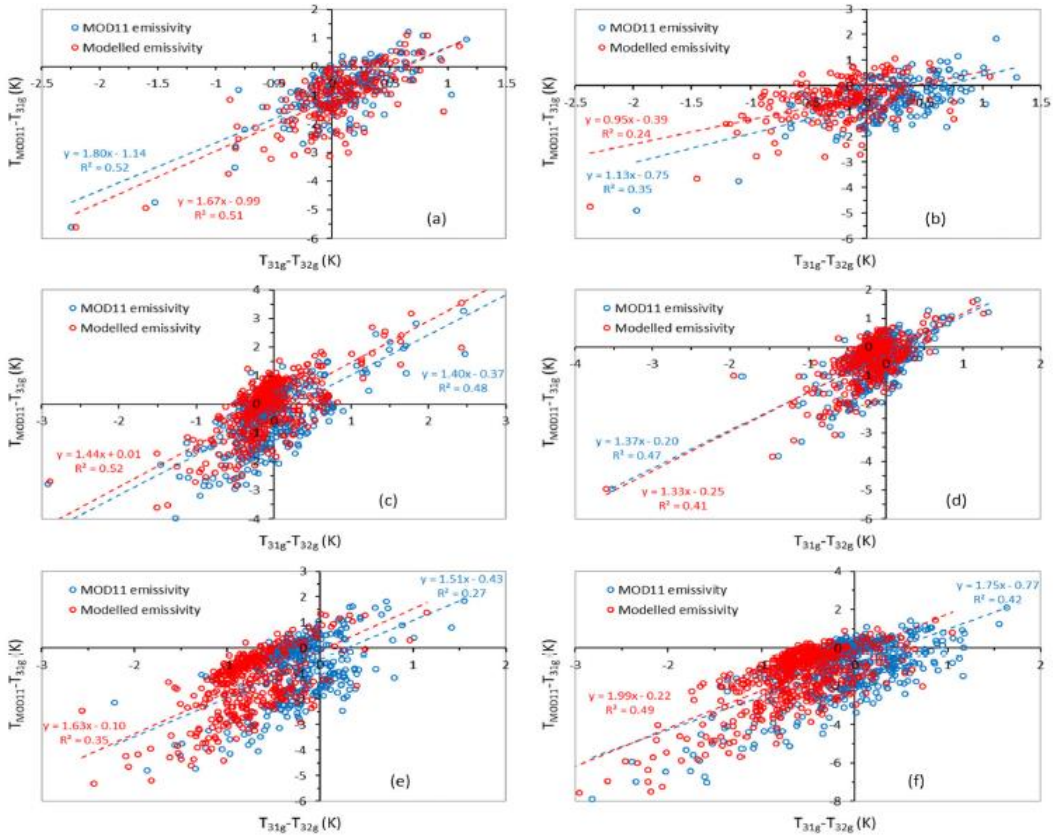

Figura 35. *Diferencia entre MOD11 y LST r-based en la banda 31 frente a la diferencia $\delta=T_{31g}-T_{32g}$ para las 6 zonas experimentales cuando se utilizan los datos de emisividad MOD11 (azul) y los valores de emisividad modelados (rojo). También se muestran la regresión lineal y el coeficiente de determinación (R^2). (a) Albufera, (b) Arrozales, (c) Cortes (matorral), (d) Soria (bosque), (e) Viña, (f) Olivar. (Coll et al., 2024).*

Teniendo en cuenta sólo los casos que superan la condición -0.5 K ≤ δ ≤ 0.5 K, calculamos las diferencias medias (bias) y el RMSD (producto menos LST r-based) para M*D11 (Tabla 21) y M*D21 (Tabla 22), utilizando las emisividades del producto y las modeladas. En estas tablas, los resultados se muestran para cada zona y en conjunto, y para el día, la noche, y en conjunto. Los datos de Terra y Aqua se consideran por separado.

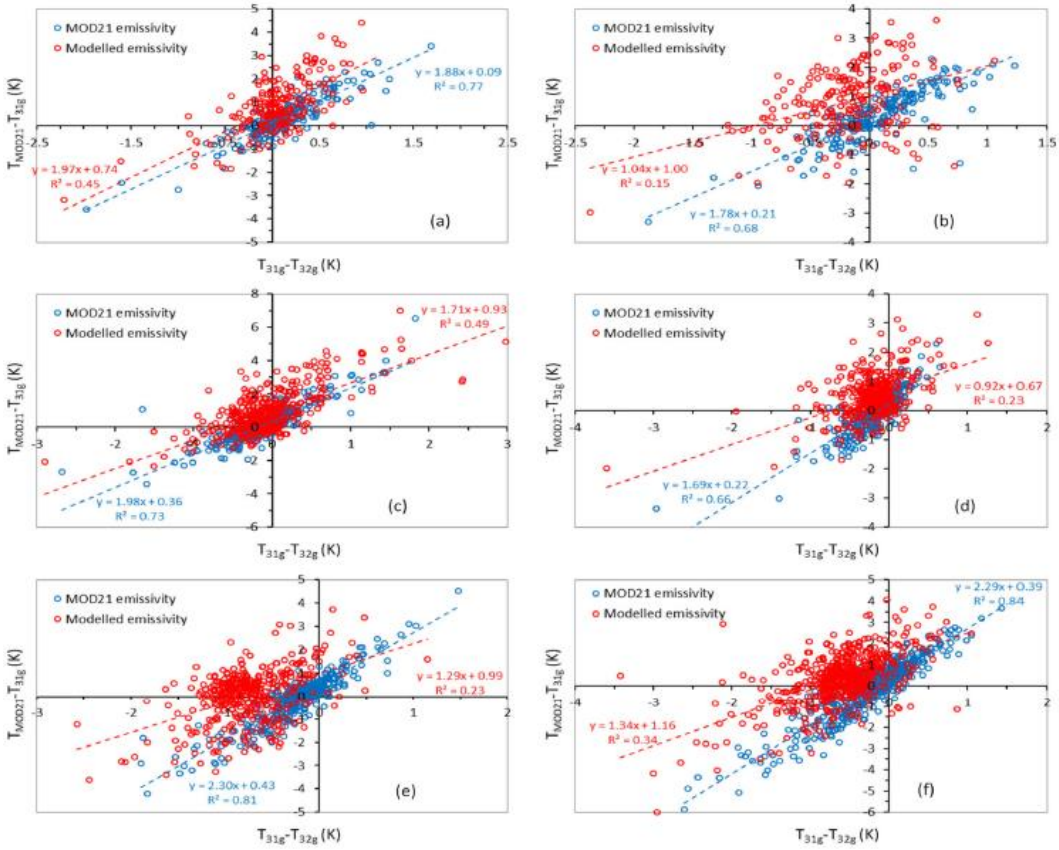

Figura 36. *Diferencia entre MOD21 y LST r-based en la banda 31 frente a la diferencia $\delta=T_{31g}-T_{32g}$ para las 6 zonas experimentales cuando se utilizan los datos de emisividad MOD21 (azul) y los valores de emisividad modelados (rojo). También se muestran la regresión lineal y el coeficiente de determinación (R^2). (a) Albufera, (b) Arrozales, (c) Cortes (matorral), (d) Soria (bosque), (e) Viña, (f) Olivar. (Coll et al., 2024).*

Las Tablas 21 y 22 muestran también la fracción (en %) de casos válidos que superan la condición δ para cada zona e instrumento (MOD y MYD) con las emisividades del producto y las modeladas. En general, los casos válidos fueron superiores al 70% para todos las zonas y bases de datos de emisividad. Las únicas excepciones fueron las zonas de viñedos y olivares con las emisividades modeladas (24%-56%). El menor número de casos válidos en dichas zonas puede deberse a la disimilitud entre las emisividades modeladas y las del producto, especialmente en lo que respecta a M*D11 (Figuras 34(e) y (f), 35(e) y (f), y 36(e) y (f)). Esto significa que los datos seleccionados podrían ser significativamente diferentes para las zonas de viñedos y olivares al comparar las distintas bases de datos de emisividad, lo que podría afectar a los resultados de la validación. No obstante, disponemos de más de 100 puntos válidos por emplazamiento e instrumento, incluso en el peor de los casos. También se observa, aunque no se muestra en la tabla que los datos nocturnos presentan en general una mayor fracción de casos válidos que los diurnos, siendo el más favorable el MYD nocturno.

La Tabla 21 muestra una buena concordancia entre las LSTs del producto M*D11 de Terra y Aqua, tanto diurnas como nocturnas. En general, Aqua proporcionó resultados ligeramente mejores que Terra, pero las diferencias en bias y RMSD estuvieron por debajo de 0.2 K en valor absoluto para la mayoría de las zonas, con diferencias máximas de 0.3 K para la zona de matorral y usualmente dentro de 0.1 K para el resto de zonas. Los resultados de la validación fueron similares tanto con el producto como con las emisividades modeladas, como se muestra en la Figura 35.

Considerando los resultados de la emisividad del producto para todas las zonas, diurnas, nocturnas y Terra y Aqua combinados, el bias global fue de -0.5 K con una RMSD de 0.9 K. Por zonas, los bias varian de -0.2 K para el bosque a -0.8 K para la Albufera (lago), mientras que los RMSDs oscilan entre 0.6 K para Soria y 1.1 K para el olivar. Considerando sólo los datos diurnos, se encuentran bias negativos (producto más frío que la LST r-based) de -1.0 K para todas las zonas juntas, oscilando entre -0.5 K para Soria el bosque y -1.4 K para el olivar. Sin embargo, el bias global fue de 0.0 K para los datos nocturnos, siendo el caso extremo -0.4 K para la Albufera (lago) y entre -0.1 y 0.2 K para las demás zonas terrestres. Los diferentes bias diurnos y nocturnos de M*D11 pueden apreciarse en la Figura 35, donde son claramente visibles dos grupos de datos correspondientes a las dos condiciones, especialmente para las zonas del viñedo y el olivar.

Para el producto M*D11, los resultados de validación fueron ligeramente mejores para las emisividades del producto que para las emisividades modeladas, pero la diferencia fue bastante pequeña, con cambios dentro de ±0.2 K en bias y ±0.1 K en RMSDs para la mayoría de las zonas, todos los datos y plataformas. Esto se debe a que las dos bases de datos de emisividad se definieron de forma similar mediante clasificación y muestran valores comparables (Figura 34).

De acuerdo con la Tabla 22, los resultados para el producto M*D21 de Terra y Aqua también fueron muy similares entre sí utilizando los datos de emisividad del producto, con bias y RMSDs dentro de ±0.1 K en promedio para todas las zonas, de día y de noche. Cuando utilizamos las emisividades modeladas, las diferencias de Terra-Aqua fueron ligeramente mayores en algunos casos, especialmente para el viñedo y el olivar por la noche. Esto puede deberse a que estos casos tenían el menor número de datos que superaban el umbral δ. Sin embargo, las diferencias entre Terra y Aqua se mantuvieron dentro de una media de ±0.3 K.

Considerando los resultados con las emisividades del producto M*D21 para todas las zonas, todos los tiempos de adquisición y plataformas combinados, el bias global fue de 0.2 K con RMSD de 0.6 K. Considerando las zonas por separado, los bias oscilan entre 0.1 K para las zonas de matorral y bosque y 0.3 K para la zona de los arrozales, mientras que los RMSD oscilaron entre 0.4 K para la zona de bosque y 0.7 K para la Albufera (lago), los arrozales y el olivar. A diferencia de M*D11, las diferencias entre los resultados nocturnos y diurnos del producto M*D21 no fueron significativas, siendo el bias diurno (nocturno) de 0.1 K (0.3 K) para todas las zonas, Terra y Aqua juntos.

Tabla 21. Bias y RMSD (todo en K) del producto M*D11 y de la LST con la emisividad modelada menos la LST r-based para cada zona y todas las zonas juntas, por día, noche y conjuntamente tanto para Terra (MOD) como para Aqua (MYD). El porcentaje de casos validos se indica entre paréntesis (en %). (Coll et al., 2024).

		Emisividad M*D11		**Emisividad modelada**	
Albufera (lago)		MOD (85)	MYD (84)	MOD (83)	MYD (86)
Día	Bias	-1.1	-0.8	-1.3	-1.0
	RMSD	1.3	1.0	1.5	1.1
Noche	Bias	-0.5	-0.4	-0.7	-0.5
	RMSD	0.7	0.7	0.8	0.8
Todo	Bias	-0.9	-0.7	-1.1	-0.8
	RMSD	1.1	0.9	1.3	1.0
Arrozales		MOD (72)	MYD (72)	MOD (77)	MYD (81)
Día	Bias	-1.0	-0.9	-0.7	-0.6
	RMSD	1.1	1.0	1.0	1.0
Noche	Bias	-0.2	0.0	0.0	-0.1
	RMSD	0.5	0.4	0.5	0.6
Todo	Bias	-0.6	-0.5	-0.5	-0.4
	RMSD	0.9	0.8	0.8	0.9
Cortes (matorral)		MOD (81)	MYD (84)	MOD (80)	MYD (85)
Día	Bias	-0.9	-0.6	-0.5	-0.2
	RMSD	1.2	0.9	0.9	0.7
Noche	Bias	0.2	0.1	0.5	0.4
	RMSD	0.4	0.4	0.6	0.6
Todo	Bias	-0.4	-0.2	-0.1	0.1
	RMSD	0.9	0.7	0.8	0.6
Soria (bosque)		MOD (87)	MYD (95)	MOD (85)	MYD (94)
Día	Bias	-0.6	-0.5	-0.6	-0.5
	RMSD	0.9	0.7	0.8	0.7
Noche	Bias	0.0	0.1	-0.1	0.1
	RMSD	0.4	0.3	0.4	0.3
Todo	Bias	-0.3	-0.1	-0.4	-0.2
	RMSD	0.7	0.5	0.6	0.5
Viña		MOD (79)	MYD (88)	MOD (24)	MYD (35)
Día	Bias	-1.4	-1.1	-1.5	-1.1
	RMSD	1.5	1.4	1.6	1.3
Noche	Bias	0.1	0.2	0.3	0.2
	RMSD	0.5	0.5	0.7	0.4
Todo	Bias	-0.6	-0.3	-0.7	-0.5
	RMSD	1.1	0.9	1.3	1.0
Olivar		MOD (76)	MYD (74)	MOD (36)	MYD (56)
Día	Bias	-1.5	-1.3	-1.5	-1.3
	RMSD	1.7	1.6	1.7	1.5
Noche	Bias	0.0	-0.1	-0.2	-0.3
	RMSD	0.4	0.4	0.4	0.5
Todo	Bias	-0.7	-0.5	-0.8	-0.9
	RMSD	1.2	1.0	1.2	1.1
Todas las zonas		MOD (80)	MYD (83)	MOD (60)	MYD (69)
Día	Bias	-1.1	-0.9	-1.0	-0.8
	RMSD	1.3	1.1	1.3	1.1
Noche	Bias	0.0	0.0	0.0	0.0
	RMSD	0.4	0.4	0.5	0.5
Todo	Bias	-0.6	-0.4	-0.6	-0.4
	RMSD	1.0	0.8	1.0	0.9

Tabla 22. Bias y RMSD (todo en K) del producto M*D21 y de la LST con la emisividad modelada menos la LST r-based para cada zona y todas las zonas juntas, por día, noche y juntos conjuntamente tanto para Terra (MOD) como para Aqua (MYD). El porcentaje de casos validos se indica entre paréntesis (en %). (Coll et al., 2024).

		Emisividad M*D21		**Emisividad modelada**	
Albufera (lago)		MOD (85)	MYD (84)	MOD (83)	MYD (86)
Día	Bias	0.2	0.3	0.8	1.0
	RMSD	0.6	0.8	1.2	1.4
Noche	Bias	0.2	0.3	0.9	0.7
	RMSD	0.6	0.7	1.2	1.0
Todo	Bias	0.2	0.3	0.8	0.9
	RMSD	0.6	0.7	1.2	1.3
Arrozales		MOD (72)	MYD (72)	MOD (77)	MYD (81)
Día	Bias	0.2	0.3	1.0	1.1
	RMSD	0.7	0.8	1.6	1.7
Noche	Bias	0.4	0.4	0.8	0.6
	RMSD	0.6	0.5	1.2	0.9
Todo	Bias	0.3	0.3	0.9	0.9
	RMSD	0.7	0.7	1.4	1.5
Cortes (matorral)		MOD (81)	MYD (84)	MOD (80)	MYD (85)
Día	Bias	0.2	0.3	1.0	1.0
	RMSD	0.7	0.6	1.4	1.3
Noche	Bias	0.0	0.0	0.2	0.3
	RMSD	0.5	0.3	0.6	0.6
Todo	Bias	0.1	0.1	0.7	0.6
	RMSD	0.6	0.5	1.1	1.0
Soria (bosque)		MOD (87)	MYD (95)	MOD (85)	MYD (94)
Día	Bias	0.0	0.1	0.5	0.9
	RMSD	0.5	0.4	0.9	1.1
Noche	Bias	0.2	0.1	0.5	0.4
	RMSD	0.4	0.2	0.7	0.5
Todo	Bias	0.1	0.1	0.5	0.6
	RMSD	0.5	0.3	0.8	0.8
Viña		MOD (79)	MYD (88)	MOD (24)	MYD (35)
Día	Bias	-0.2	0.1	-0.1	0.1
	RMSD	0.6	0.7	1.1	1.0
Noche	Bias	0.4	0.3	1.3	0.5
	RMSD	0.6	0.4	1.6	0.8
Todo	Bias	0.1	0.2	0.5	0.3
	RMSD	0.6	0.5	1.4	0.9
Olivar		MOD (76)	MYD (74)	MOD (36)	MYD (56)
Día	Bias	-0.1	0.0	0.7	0.6
	RMSD	0.7	0.8	1.3	1.2
Noche	Bias	0.5	0.4	0.9	0.4
	RMSD	0.7	0.5	1.3	0.7
Todo	Bias	0.3	0.2	0.8	0.5
	RMSD	0.7	0.6	1.3	1.0
Todas las zonas		MOD (80)	MYD (83)	MOD (60)	MYD (69)
Día	Bias	0.0	0.2	0.6	0.8
	RMSD	0.6	0.7	1.2	1.2
Noche	Bias	0.3	0.2	0.8	0.5
	RMSD	0.6	0.4	1.1	0.7
Todo	Bias	0.2	0.2	0.7	0.6
	RMSD	0.6	0.5	1.2	1.0

Por otra parte, el impacto de la base de datos de emisividad fue mayor para el producto M*D21 que para el M*D11. Al utilizar las emisividades modeladas, se observa un aumento uniforme de aproximadamente 0.6 K en el bias y los RMSD con respecto a las obtenidas con las emisividades del producto para la mayoría de las zonas. Esto demuestra la consistencia entre la LST y las emisividades del producto M*D21, que se obtienen simultáneamente a partir del método TES. Con las emisividades modeladas, el bias fue de 0.6 K y el RMSD de 1.1 K (todos los datos), lo que sigue mostrando un funcionamiento aceptable. Por zonas, los bias oscilan entre 0.4 K (viñedo) y 0.9 K (arrozales) y los RMSD entre 0.8 K (bosque) y 1.4 K (arrozales). El aumento del bias se explica porque, según la Figura 34, las emisividades modeladas suelen ser mayores que las emisividades del producto M*D21, lo que produce valores de LST r-based más pequeños. Al igual que con las emisividades del producto, no se encuentran diferencias significativas entre los bias diurnos y nocturnos.

Los productos M*D11 y M*D21 proporcionan una estimación de la incertidumbre de la LST para cada valor LST. Un ejemplo se muestra para la zona del olivar en la Figura 37(a), donde las incertidumbres de LST de MOD11 y MOD21 LSTs (Terra solo) se representan frente al día del año. La incertidumbre atribuida a M*D11 es uniforme con un valor alrededor de 0.5 K para todas las zonas, diurnas y nocturnas y datos de Terra y Aqua. La incertidumbre de M*D21 es normalmente mayor, con valores mínimos en torno a 1 K y una gran variabilidad en los días centrales del año, donde la incertidumbre puede alcanzar más de 3 K, y un valor medio de 1.3 K para todas las zonas, horas de adquisición y plataformas. En la Figura 37(b), las incertidumbres estimadas de los productos LST se representan gráficamente frente a $W/\cos(\theta)$, que representa el contenido de vapor de agua atmosférico en la trayectoria de observación del satélite. Se puede ver que la incertidumbre MOD21 aumenta claramente con $W/\cos(\theta)$ ya que las incertidumbres en el proceso de corrección atmosférica requeridas para el producto M*D21 aumentan para altas cargas de vapor de agua atmosférico y temperaturas de superficie, como es el caso del verano y principios del otoño. En el caso de la incertidumbre M*D11, sólo se produce un pequeño aumento de 0.4 a 0.7 K con la humedad atmosférica.

Las RMSD halladas para M*D11 en esta tesis son compatibles con la incertidumbre del producto sólo para los valores nocturnos, mientras que son mayores para los diurnos debido a los bias fríos relativamente grandes obtenidos en la mayoría de las zonas. En el caso del producto M*D21, las RMSD obtenidas aquí tanto con la emisividad del producto como con la emisividad modelada son menores que la incertidumbre del producto para la mayoría de las zonas tanto de día como de noche.

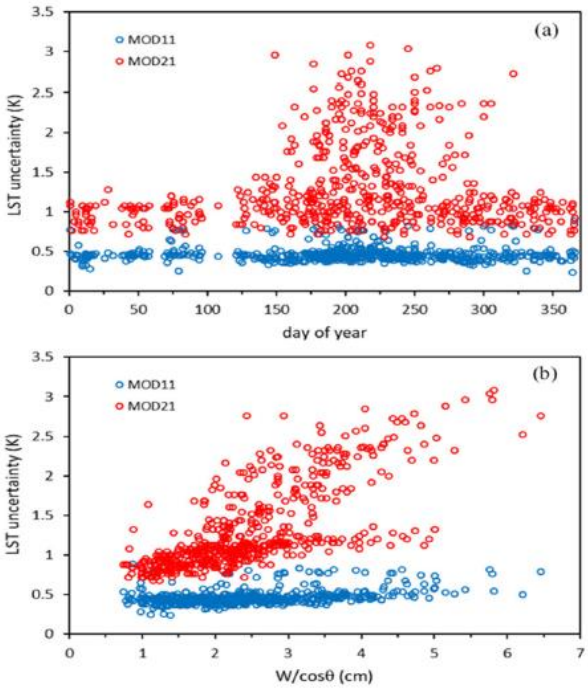

Figura 37. *Incertidumbre de los productos LST MOD11 (azul) y MOD21 (rojo) para la zona del olivar en función de: (a) el día del año y (b) el contenido de vapor de agua en la trayectoria de observación del satélite W/cos(θ). (Coll et al., 2024).*

Como mencionamos anteriormente, el producto M*D11 muestra un bias diurno de aproximadamente -1.0 K de media (Tabla 21). El bias diurno puede estar relacionado con la subestimación de las LST del producto a altas temperaturas, como se muestra en la Figura 38(a), donde la diferencia T_{MOD11}-T_{31g} (con T_{31g} obtenida a partir de las emisividades del producto) se representa gráficamente frente a la temperatura de brillo T_{31} para la zona del olivar. Los resultados para las demás zonas y los datos de MYD11 son comparables. Se puede observar que el bias negativo se produce para temperaturas de brillo superiores a 295 K, que corresponden mayoritariamente al día. En la Figura 38(b) se representan las diferencias T_{MOD11}-T_{31g} frente a las diferencias de temperatura de brillo T_{31}-T_{32}, que desempeñan un papel importante en el algoritmo SW del producto M*D11. Normalmente, los valores grandes de T_{31}-T_{32} corresponden a casos diurnos en verano. Según la Figura 38(b), la subestimación de T_{MOD11} parece aumentar linealmente con la diferencia de temperatura de brillo. Esto puede indicar un mal ajuste de los coeficientes del SW en el algoritmo M*D11 para altas temperaturas de superficie y contenidos de vapor de agua atmosférico.

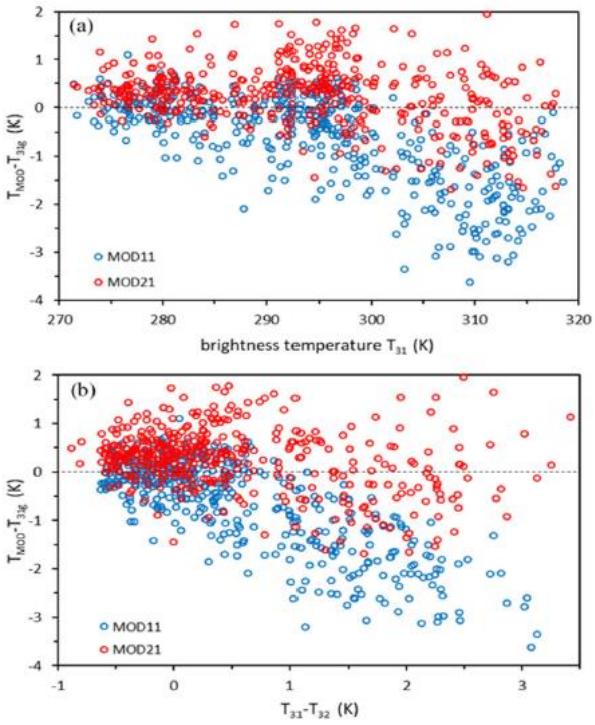

Figura 38. *Diferencia entre el producto LST de MODIS y la temperatura de referencia r-based T_{31g} para MOD11 (azul) y MOD21 (rojo) en la zona del olivar en función de, (a) la BT de MODIS en la banda 31 y (b) la diferencia de temperatura de brillo T_{31}-T_{32}. (Coll et al., 2024).*

La emisividad del agua es uniforme, constante y bien conocida, por lo que las emisividades modeladas deberían aproximarse al valor real. Sin embargo, los bias diurnos y nocturnos fueron considerables (-1 y -0.5 K, respectivamente, en la Tabla 21). Parte del bias negativo podría deberse a la sobreestimación de la emisividad M*D11 para ángulos de visión superiores a 40º mostrada en la Figura 34(a). Teniendo en cuenta sólo los casos con $\theta < 40º$, el bias diurno disminuye a -0.8 K, mientras que el bias nocturno aumenta un valor similar de -0.7 K. Estos bias negativos pueden deberse a que las temperaturas del agua se encuentran en el rango de 295-300 K para más del 50% de los casos tanto diurnos como nocturnos analizados.

Por el contrario, el producto M*D21 no mostró diferencias significativas en los bias diurnos y nocturnos. En la Figura 38, se representan también las diferencias T_{MOD21}-T_{31g} frente a T_{31} y T_{31}-T_{32} para la zona del olivar, sin mostrar ninguna correlación. La dispersión aumenta para valores más altos de T_{31} y T_{31}-T_{32}, que corresponden típicamente a primavera y verano con atmósferas húmedas, en concordancia con los patrones de incertidumbre del producto M*D21 mostrados en la Figura 37. Los resultados son similares para las demás zonas y datos MYD.

e. Resultados para el sensor AVHRR/3

Para el sensor AVHRR/3 a bordo de los satélites MetOp-A, -B y –C, hemos realizado una validación t-based del producto de LST propuesto por LSA-SAF, y de otros tres algoritmos SW de distintos autores (sección 2.e.4): el propuesto en esta tesis para cada uno de los satélites (ecuación 2.46, Tabla 2), el de Jiménez-Muñoz y Sobrino (2008) (JMS) (ecuación 2.11), y el de Parra et al (2006) (ecuación 2.47). El producto LSA-SAF (solo disponible para MetOp-B) ha sido evaluado en las versiones v100 y v110. Las diferencias entre la versión v100 y la v110 son: corrección de errores en el promedio temporal entre pasadas en la composición diaria, adición de una capa de incertidumbre de la LST, separación entre órbitas ascendentes y descendentes, corrección de errores en la conversión de radiancias a temperaturas de brillo y desenmascarado de píxeles de aguas interiores. Las emisividades en los canales 4 (10.8 μm) y 5 (12.0 μm) de AVHRR/3, necesarias para la aplicación de estos tres algoritmos, son las de la Tabla 7. Todas las medidas LST in situ que se han utilizado como referencia fueron tomadas con radiómetros Apogee SI-121 en la estación fija de los arrozales de Valencia entre 2016 y 2018.

e.1. Validación t-based para MetOp-A

Aunque el satélite MetOp-A ya ha sido decomisado, ha sido incluido en el estudio debido a que sus datos siguen estando disponibles. La Tabla 23 muestra la mediana, RSD y R-RMSD de las diferencias entre las LSTs calculadas a partir de los tres algoritmos SW mencionados y las medidas in situ. Podemos observar que el propuesto en esta tesis (ecuación 2.46) proporciona medianas cercanas a cero para el agua y la vegetación y relativamente pequeña (0.6 K) en suelos desnudos. Similarmente, el algoritmo JMS presenta medianas pequeñas, excepto en el caso de la vegetación (1.0 K), mientras que el algoritmo de Parra et al. (2006) ecuación (2.47), desviaciones sistemáticas considerables (alrededor de 1 K) en todas las cubiertas. Por otro lado, las RSD fueron similares para todos los algoritmos y cubiertas. Como consecuencia, la R-RMSD para todos los datos fue 1.5 K con una mediana cercana a cero para la ecuación (2.46) y JMS, mientras que la de la ecuación (2.47) fue 1.7 K con una sobreestimación sistemática de 0.9 K.

Tabla 23. Resultados estadísticos (en K) de las diferencias entre las LST de satélite y las LST in situ para el algoritmo SW propuesto (ecuación 2.46), el de Jiménez-Muñoz y Sobrino (2008) (JMS) (ecuación 2.11), y el de Parra et al (2006) (ecuación 2.47). Los datos corresponden al satélite MetOp-A. Presentamos los resultados por tipo de cobertura y para el total de datos, con el número de datos en cada caso.

Cobertura	Estadístico	Ec. (2.46)	Ec. (2.11)	Ec. (2.47)
Agua	Mediana	0.0	0.0	0.7
	RSD	1.6	1.5	1.6
	R-RMSD	1.6	1.5	1.7
	N de datos	252	239	245
Tierra húmeda	Mediana	0.6	-0.1	1.0
	RSD	1.2	1.2	1.3
	R-RMSD	1.3	1.2	1.6
	N de datos	85	82	81
Tierra seca	Mediana	0.6	-0.6	1.0
	RSD	1.1	1.3	1.3
	R-RMSD	1.3	1.5	1.7
	N de datos	140	141	144
Vegetación	Mediana	0.1	1.0	0.9
	RSD	1.5	1.5	1.6
	R-RMSD	1.5	1.8	1.9
	N de datos	105	105	105
Total	Mediana	0.3	-0.1	0.9
	RSD	1.4	1.5	1.5
	R-RMSD	1.5	1.5	1.7
	N de datos	582	567	575

e.2. Validación t-based para MetOp-B

En el caso del satélite MetOp-B, realizamos la validación del producto de LSA-SAF (versión 100 y versión 110), además de los tres algoritmos de SW de la sección anterior. En la Tabla 24 presentamos inicialmente los resultados de las dos versiones del producto LSA-SAF y del algoritmo propuesto en esta tesis, separados entre datos diurnos y nocturnos para todos los datos de la zona de arrozales. Como se observa en la Tabla 24 cuando se estudian todos los datos conjuntamente, la versión v100 muestra una mediana cercana a cero (-0.1 K) en los datos diurnos (R-RMSD de 1.5 K), mientras subestima las medidas in situ (-1.3 K) para los nocturnos (R-RMSD de 1.9 K). Para la versión v110 aumentan notablemente los valores de la RSD, y la mediana de los datos diurnos empeora, con una sobreestimación de 1.8 K (R-RMSD de 3.4 K), mientras que la de los datos nocturnos se hace cercana a cero (0.2 K). En general, los resultados muestran un mejor funcionamiento de la versión v100 en la zona de los arrozales. El algoritmo SW propuesto ofrece resultados más consistentes entre día y noche, tanto para las medianas (0.8 K y -0.5 K, respectivamente) como para la R-RMSD (1.8 K y 1.9 K, respectivamente). Sin embargo, cuando se separan por tipos de coberturas, en el caso del agua el que menor R-RMSD muestra para los datos diurnos es el SW propuesto con 1.7 K y para los datos nocturnos la versión v100 con 2.2 K. El SW propuesto es el que mejor rendimiento ofrece en tierra húmeda (R-RMSD de 1.6 K con datos diurnos y de 1.5 K con datos nocturnos), tierra seca (1.4 K y 0.9 K, respectivamente) y vegetación (0.0 K y -

0.2 K, respectivamente). Esto puede ser debido a la asignación constante de la emisividad a los píxeles de la zona de estudio, ya que el producto LSA-SAF tanto la versión v100 como la v110 obtiene la emisividad con la fracción de cobertura vegetal que obtienen del International Geosphere-Biosphere Programme (IGBP), el cual asigna de manera continua este valor con independencia de los cambios sobre la cobertura de la zona.

Tabla 24. Resultados estadísticos (en K) de las diferencias entre las LST de satélite y las LST in situ para el producto LSA-SAF (versiones v100 y v110) y el algoritmo SW propuesto en esta tesis (ecuación 2.46). Los datos corresponden al satélite MetOp-B en la zona de los arrozales (separando por cubiertas y todas juntas), separando los resultados diurnos y nocturnos. con el número de datos en cada caso.

		Producto v100		**Producto v110**		**Ec. (2.46)**	
Cobertura	**Estadístico**	**Día**	**Noche**	**Día**	**Noche**	**Día**	**Noche**
Agua	Mediana	0.4	-0.4	3.4	-1.1	1.1	-1.5
	RSD	2.1	2.1	1.3	2.4	1.7	2.6
	R-RMSD	2.1	2.2	3.6	2.6	1.7	2.6
	N de datos	68	68	39	84	124	134
Tierra húmeda	Mediana	-0.8	-2.0	2.2	0.4	1.1	-0.6
	RSD	2.1	1.7	2.2	2.5	1.2	1.4
	R-RMSD	2.3	2.6	3.1	2.5	1.6	1.5
	N de datos	12	14	26	27	47	70
Tierra seca	Mediana	-1.7	-1.7	-0.4	1.3	0.5	0.2
	RSD	2.4	1.2	2.6	2.1	1.3	0.9
	R-RMSD	2.9	2.1	2.6	2.4	1.4	0.9
	N de datos	39	44	64	53	42	92
Vegetación	Mediana	-0.1	-1.3	2.2	0.3	0.0	-0.2
	RSD	1.8	1.0	2.8	1.5	2.0	1.5
	R-RMSD	1.8	1.6	3.6	1.5	2.0	1.5
	N de datos	155	140	34	148	59	63
Total	Mediana	-0.1	-1.3	1.8	0.2	0.8	-0.5
	RSD	1.4	1.3	2.9	2.0	1.6	1.8
	R-RMSD	1.5	1.9	3.4	2.0	1.8	1.9
	N de datos	274	266	163	312	272	359

Por otro lado, hemos analizado la dependencia con la LST in situ de las diferencias entre la LST de satélite y la LST in situ para el producto LSA-SAF, versiones v100 y 110, y el SW propuesto (ecuación 2.46) en los datos diurnos y nocturnos por separado (Figura 39). No se muestra los resultados de los algoritmos de Jiménez-Muñoz y Sobrino (2008) (JMS) y Parra et al (2006) debido a que son similares a los de la ecuación (2.46). Como se observa, en los valores nocturnos no existe una tendencia marcada de los resultados de los algoritmos con la LST in situ. Sin embargo, los valores diurnos del producto LST en las dos versiones 100 y 110 muestran una clara disminución con la temperatura de la superficie de las diferencias entre LST de satélite e in situ, siendo positivas (negativas) para valores bajos (altos) de LST. La tendencia es similar en las dos versiones, con lo que los valores dispares de la mediana diurna en v100 y v110 (Tabla 24) se deben a la distinta distribución de temperaturas en cada base de datos. Esto también afecta a la notablemente mayor dispersión de la versión v110 comparada con la de v100 (véase RSD en la Tabla 24). Por otra parte, el algoritmo SW propuesto muestra

una tendencia negativa con la LST in situ, aunque no tan acusada como el producto LST, con lo que las medianas diurna y nocturna son más consistentes entre sí, como hemos mencionado anteriormente.

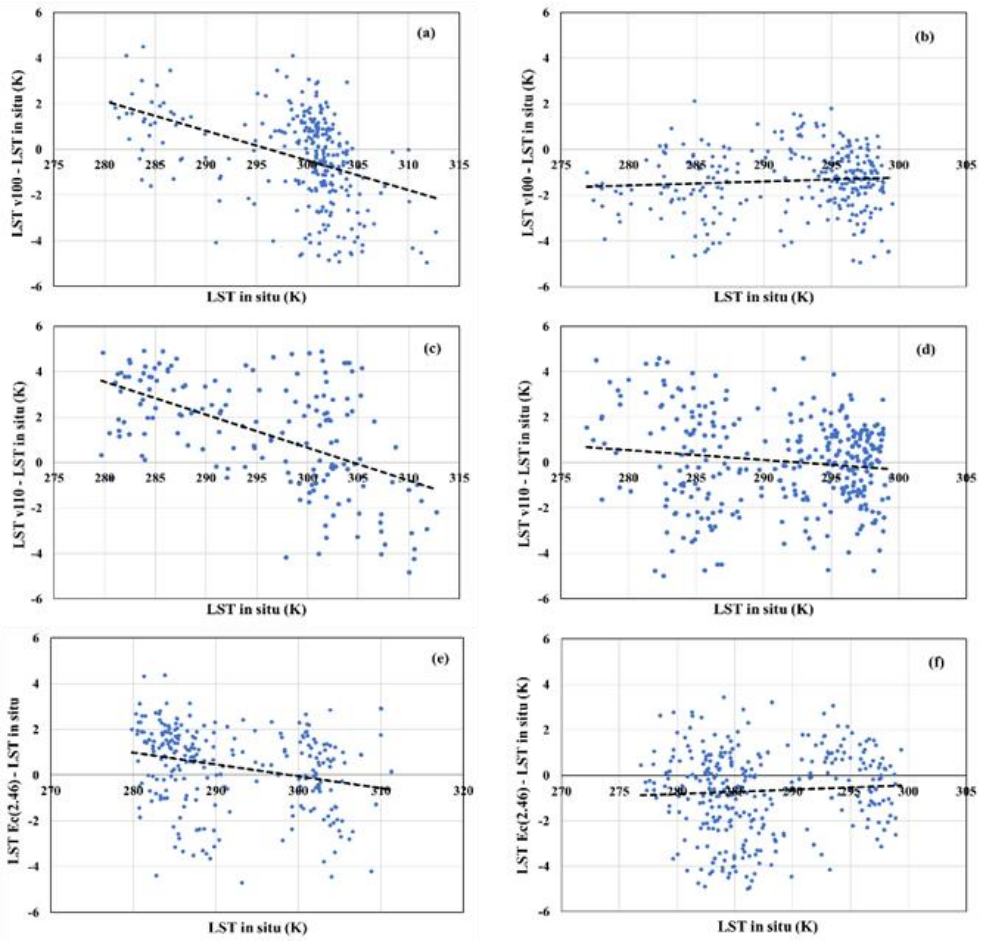

Figura 39. *Diferencias entre LST de satélite y LST in situ frente a LST in situ. La línea de tendencia está punteada en negro. Siendo (a-b): producto (versión v100) día- noche; (c-d): producto (versión v110) día-noche; (e-f): Ecuación (2.46) día-noche.*

Al igual que para el satélite MetOp-A, la Tabla 25 muestra los resultados de validación para el satélite MetOp-B del algoritmo SW propuesto, el de Jiménez-Muñoz y Sobrino (2008) (JMS) y el de Parra et al. (2006) por tipo de cubierta y para todos los datos de la zona de arrozales. Los resultados son consistentes con los de la Tabla 23 (MetOp-A), siendo las medianas para el algoritmo propuesto cercanas a cero en todas las cubiertas (entre -0.4 K y 0.3 K), mientras que el algoritmo JMS muestra un rango mayor (entre -0.7 K y 1.1 K). También se observa que la RSD es mayor en la cubierta de agua comparada con los datos de MetOp-A, lo que hace que aumente la R-RMSD. Por otro lado, el algoritmo de Parra et al. (2006) sigue proporcionando sobreestimaciones sistemáticas de alrededor de 1.0 K en todas las cubiertas.

Teniendo en cuenta todo el conjunto de datos, la ecuación (2.46) y el algoritmo JMS muestran desviaciones sistemáticas cercanas a cero, siendo de 1.0 K para la ecuación (2.47), de la misma forma que para MetOp-A. Sin embargo, las RSD son mayores para el MetOp-B, aumentando las R-RMSDs hasta alrededor de 2.0 K en los tres algoritmos.

Tabla 25. Resultados estadísticos (en K) de las diferencias entre las LST de satélite y las LST in situ para el algoritmo SW propuesto (ecuación 2.46), el de Jiménez-Muñoz y Sobrino (2008) (JMS) (ecuación 2.11), y el de Parra et al (2006) (ecuación 2.47). Los datos corresponden al satélite MetOp-B. Presentamos los resultados por tipo de cobertura y para el total de datos, con el número de datos en cada caso.

Cobertura	Estadístico	Ec. (2.46)	Ec. (2.11)	Ec. (2.47)
Agua	Mediana	-0.4	-0.1	0.9
	RSD	2.7	2.5	2.6
	R-RMSD	2.8	2.5	2.8
	N de datos	258	234	242
Tierra húmeda	Mediana	0.3	-0.4	0.8
	RSD	1.7	1.9	1.7
	R-RMSD	1.7	1.9	1.9
	N de datos	117	112	109
Tierra seca	Mediana	0.3	-0.7	1.0
	RSD	0.9	0.9	1.1
	R-RMSD	1.0	1.1	1.5
	N de datos	134	128	136
Vegetación	Mediana	-0.2	1.1	1.2
	RSD	1.8	1.8	1.8
	R-RMSD	1.8	2.1	2.2
	N de datos	122	122	125
Total	Mediana	0.0	-0.2	1.0
	RSD	1.9	2.0	2.0
	R-RMSD	1.9	2.1	2.2
	N de datos	631	596	612

e.3. Validación t-based para MetOp-C

Para el sensor AVHRR/3 a bordo del satélite MetOp-C hemos realizado la validación de los tres algoritmos SW anteriores para los distintos tipos de cobertura de los arrozales. Análogamente a las Tablas 23 y 25, la Tabla 26 muestra los resultados para MetOp-C para cada cubierta y para el total de los datos. El algoritmo SW propuesto muestra resultados consistentes con los de MetOp-A y -B, con medianas cercanas a cero en todas las cubiertas (máxima desviación de -0.6 K para tierra seca y -0.3 K para el conjunto de los datos), aunque se produce un aumento en la RSD para la superficie de agua y en la R-RMSD del total (1.9 K). El algoritmo JMS presenta sobreestimaciones sistemáticas de 2.6 K en la vegetación y 1.4 K en el agua, con 1.4 K para todas las coberturas, con lo que su R-RMSD aumenta hasta 2.2 K. El algoritmo de Parra et al. (2006) sigue mostrando sobreestimaciones sistemáticas grandes en todas las superficies (2.1 K para el conjunto de datos), que superan las de MetOp-A y –B y hacen que la R-RMSD sea mayor (2.7 K). Por tanto, podemos observar que para los

tres satélites MetOp el algoritmo SW propuesto en esta tesis es el que mejor se adapta a los distintos tipos de cobertura, tanto en términos de la incertidumbre sistemática (mediana) como de la R-RMSD.

Tabla 26. Resultados estadísticos (en K) de las diferencias entre las LST de satélite y las LST in situ para el algoritmo SW propuesto (ecuación 2.46), el de Jiménez-Muñoz y Sobrino (2008) (JMS) (ecuación 2.11), y el de Parra et al (2006) (ecuación 2.47). Los datos corresponden al satélite MetOp-C. Se presenta los resultados por tipo de cobertura y para el total de datos, con el número de datos en cada caso.

Cobertura	**Estadístico**	**Ec. (2.46)**	**Ec. (2.11)**	**Ec. (2.47)**
Agua	Mediana	-0.3	1.4	2.0
	RSD	2.1	1.7	1.9
	R-RMSD	2.1	2.2	2.7
	N de datos	174	130	138
Tierra húmeda	Mediana	0.2	0.6	1.6
	RSD	1.0	1.2	1.7
	R-RMSD	1.1	1.4	2.3
	N de datos	36	33	37
Tierra seca	Mediana	-0.6	0.3	2.1
	RSD	1.3	1.2	0.9
	R-RMSD	1.4	1.2	2.3
	N de datos	29	22	25
Vegetación	Mediana	-0.4	2.6	2.5
	RSD	1.5	1.4	1.5
	R-RMSD	1.6	3.0	2.9
	N de datos	59	52	55
Total	Mediana	-0.3	1.4	2.1
	RSD	1.8	1.7	1.6
	R-RMSD	1.9	2.2	2.7
	N de datos	298	237	255

Capítulo 5

CONCLUSIONES

En esta sección presentamos las conclusiones más relevantes de los análisis realizados en esta tesis. De la calibración del sensor SLSTR y validación de algoritmos de LST mediante el método t-based y r-based para este sensor se concluye que:

1. La calibración de las bandas del IRT del sensor SLSTR, tanto en Sentinel-3A como -3B, mostró una buena concordancia entre las BTs simuladas y las obtenidas desde satélite para ambos instrumentos, con una desviación media (R-RMSD) de -0.3 K (0.8 K) para la banda 8 y de -0.2 K (0.7) para la banda 9. Estas diferencias se encuentran dentro de los límites de incertidumbre de las BTs simuladas y muestran que tanto los sensores como las bandas funcionan de forma coherente, sin desviaciones sistemáticas y con una buena precisión.

2. La validación con el método t-based del producto operativo LST de SLSTR mostró diferencias medias que superaban el umbral de precisión de ±1.0 K recomendado por el GCOS para el desierto (-2 K) y los cultivos de regadío (1.5 K) tanto de día como de noche, mientras que ofreció desviaciones medias dentro de ±1.0 K y R-RMSD en torno a 1.5 K para el agua y el bosque. Para todas las zonas combinadas, la R-RMSD diurno (nocturno) del producto SLSTR fue de 1.5 K (1.7 K).

3. Para el algoritmo SW que propusimos en Pérez-Planells et al. (2021), solo la zona del desierto con datos diurnos muestra desviaciones medias excesivas (-1.7 K), siendo los resultados similares al producto SLSRT en los casos de agua y bosques. Para todas las zonas en conjunto, la R-RMSD tanto diurna como nocturna fue de 1.3 K.

4. En la zona de estudio de los arrozales de Valencia, el producto LST de SLSTR presenta una R-RMSD grande (2.0 K) ya que asigna un coeficiente fijo a este bioma, que no se ajusta debidamente al ser una zona donde la cobertura varía a lo largo del año.

5. La validación r-based mostró diferencias medias superiores al umbral del GCOS para el producto de SLSTR en zonas cultivo de regadío y zonas desérticas tanto de día como de noche, y para nuestro SW (Perez-Planells et al., 2021) en la zona desértica con datos diurnos (siendo cercanos a cero para todos los demás casos). Esto concuerda con la validación t-based y puede indicar problemas en los propios algoritmos para esos casos particulares.

6. Las R-RMSD totales diurnas (nocturnas) fueron 1.0 K (1.0 K) para el producto LST y 0.6 K (0.3 K) para nuestro SW (Pérez-Planells et al., 2021), que son inferiores a los de la validación t-based.

De la calibración y validación de algoritmos SC y SW mediante el método t-based con datos del sensor TIRS podemos destacar las siguientes conclusiones:

7. Hemos evaluado la calibración de las bandas 10 y 11 en los distintos reprocesados que se han llevado a cabo. Los datos originales (antes de 2014) mostraron sobreestimaciones de 0.7 K y 2.7 K en las BTs de las bandas 10 y 11, respectivamente, respecto los valores simulados a partir de datos in situ.

8. Tras el reprocesado de 2014, obtuvimos subestimaciones de -1.4 K y -1.6 K en las bandas 10 y 11, respectivamente. Sin embargo, después del reprocesado de 2017 nuestros resultados mostraron

desviaciones medias insignificantes para ambas bandas, lo que demuestra la mejora introducida por el algoritmo TIRS-on-TIRS. En la actualización de 2019, obtuvimos diferencias medias de -0.1 K y 0.5 K para las bandas 10 y 11, respectivamente, introduciendo así una diferencia de calibración apreciable entre ambas bandas.

9. La evaluación t-based de la LST obtenida mediante el método SC aplicada a los datos reprocesados de 2017 mostró desviaciones medias insignificantes, con la RMSD más baja para la banda 10 (0.7 K). Tras evaluar los resultados con el reprocesado de 2019, el RMSD se mantiene idéntico.

10. Los algoritmos SW para que proporcionaron los mejores resultados TIRS con el reprocesado de 2017 sobre la zona de validación de los arrozales fueron: el algoritmo de Jiménez-Muñoz et al. (2014) con una RMSD de 0.6 K, el algoritmo de Wan y Dozier (1996) con el refinamiento de Wan (2014) y con coeficientes constantes estimados por Gerace et al. (2020) con una RMSD de 0.8 K, y el algoritmo Enterprise con coeficientes estimados por Meng et al. (2019) por rangos W con una RMSD de 0.8 K.

11. Sin embargo, tras el reprocesado de 2019 las diferencias medias aumentaron entre 1.3 K y 1.9 K dependiendo del algoritmo, no observándose diferencias significativas en las SD. Esto podría significar un empeoramiento de los resultados de los SW aplicados a la Colección 1 en su aplicación a los datos de la Colección 2.

Al realizar la calibración de las bandas TIR al sensor TIRS-2 y validar los algoritmos con el método t-based sobre la zona de los arrozales de Valencia se concluye lo siguiente:

12. Las BTs de la calibración original (OC) de TIRS-2 coincidieron bien con las estimadas a partir de LSTs in situ, con diferencias medias y RMSDs de -0.3 K y 0.5 K, respectivamente, para la banda 10, siendo -0.1 K y 0.6 K para la banda 11, respectivamente.

13. Al aplicar el reprocesado de 2023 (datos RP) no se observa ninguna diferencia apreciable en la banda 10 (de alrededor de -0.01 K) respecto a los datos OC. Sin embargo, se encontró un cambio significativo de 0.3 K en la banda 11. El reprocesado de 2023 hizo que la diferencia media entre ambas bandas aumentara (de -0.2 K para OC a -0.6 K para RP), lo que afecta al funcionamiento de los SW.

14. Las validaciones t-based de la mayoría de los SW para el sensor TIRS-2 mostraron diferencias medias y SD dentro del umbral recomendado por el GCOS (1 K), tanto para los datos OC como los RP. Sin embargo, se observó una diferencia sistemática significativa (alrededor de -0.6 K) entre las diferencias medias al usar datos RP y OC, debido al cambio en las bandas 10 y 11 mencionado en el punto anterior.

15. Los resultados de calibración y validación obtenidos para los actuales datos TIRS-2 muestran buenas precisiones y exactitudes para monitorizar con precisión la LST en la alta resolución espacial proporcionada por el sensor. No se observaron errores de calibración significativos, a diferencia de lo ocurrido en las primeras calibraciones de TIRS.

En relación al sensor MODIS a bordo de Terra y Aqua al realizar la validación de sus productos de LST M*D11 y M*D21 mediante el método r-based podemos destacar que:

16. En primer lugar, al comparar las LSTs de referencia del método r-based con medidas de LST in situ para comprobar la validez del método r-based en un conjunto de datos diurnos de Terra y Aqua sobre la zona de los arrozales de Valencia. Los resultados mostraron una RMSD<1.0 K en RMSD cuando la diferencia de δ estaba dentro de ±0.5 K.

17. Las LSTs diurnas del producto M*D11 mostraron una subestimación media para todas las zonas (datos de Terra y Aqua combinados) de -1.0 K y RMSD de 1.2 K, tanto con las emisividades del producto como con las modeladas. Esta subestimación diurna se atribuye a problemas del algoritmo para condiciones de altas temperaturas y humedades atmosféricas. Para los datos nocturnos, la desviación media fue 0.0 K (variando de -0.4 K a 0.2 K entre zonas) y el RMSD fue 0.5 K.

18. Las LSTs del producto M*D21 mostraron una gran estabilidad entre zonas y entre el día y la noche, con una desviación media diurna (nocturna) de 0.1 K (0.3 K) y RMSD de 0.7 K (0.5 K) para todos los sitios y datos de Terra y Aqua combinados, cuando utilizamos las emisividades del producto. Usando emisividades modeladas, la desviación media y la RMSD aumentaron de forma coherente en todas las zonas, de día y de noche, y para Terra y Aqua., pero siguen proporcionando una desviación media (0.6 K) y una RMSD (1.1 K) aceptables.

19. El método t-based debe considerarse como el método de validación independiente de referencia, aunque requiere medidas precisas de LST in situ durante largos periodos de tiempo en zonas muy homogéneas térmicamente, de una sola componente, con instrumentación y recursos materiales considerables. Esto limita notablemente el número de datos y el tipo de superficies terrestres aptas para la validación t-based.

20. El método r-based proporciona una alternativa valiosa para la mayoría de las cubiertas terrestres (donde medidas de LST in situ no son factibles) y condiciones atmosféricas globales. La precisión de las LSTs r-based de referencia puede comprobarse con LSTs medidas in situ en zonas seleccionadas, como en la de los arrozales de Valencia del presente trabajo.

Por último, cuando se validan con el método t-based los algoritmos de LST para el sensor AVHRR/3 a bordo de los satélites MetOp-A, -B y -C obtenemos las siguientes conclusiones:

21. La validación t-based del producto de LST propuesto por LSA-SAF para el sensor AVHRR/3 (MetOp-B) muestra un mejor funcionamiento de la versión v100, especialmente de día con una mediana cercana a cero y R-RMSD de 1.5 K, mientras que para datos nocturnos la mediana es -1.3 K. En la versión v110, los datos diurnos aumentan notablemente la mediana (1.8 K) y la RSD (R-RMSD de 3.4 K), mientras que la mediana de los datos nocturnos se hace cercana a cero (R-RMSD de 2.0 K). Además, observamos una dependencia de las diferencias con la temperatura de la superficie, tendiendo a ser negativas para valores altos en datos diurnos.

22. El algoritmo SW propuesto en esta tesis con coeficientes particularizados para cada satélite MetOp-A, B y C presenta medianas cercanas a cero para los distintos tipos de cubierta de la zona de validación de los

arrozales, y resultados consistentes entre los tres satélites y entre datos diurnos y nocturnos, con R-RMSD total entre 1.5 K y 1.9 K, según el satélite.

23. El algoritmo de Jiménez-Muñoz y Sobrino (2008) proporciona resultados similares a los del algoritmo propuesto para MetOp-A y -B, excepto en el caso de la vegetación donde presenta sobreestimaciones de la LST de alrededor de 1.0 K. Para el MetOp-C la sobreestimación es mayor y en todas las coberturas, con una mediana total de 1.4 K. El algoritmo de Parra et al. (2006) presenta sobreestimaciones sistemáticas en todas las cubiertas, con valores medios de 1.0 K para MetOp-A y –B y de 2.1 K para MetOp-C.

BIBLIOGRAFÍA

Alfieri, J.G., Anderson, M.C., Kustas, W.P., Cammalleri, C., 2017. Effect of the revisit interval and temporal upscaling methods on the accuracy of remotely sensed evapotranspiration estimates. Hydrol. Earth Syst. Sci. 21, 83–98. https://doi.org/10.5194/hess-21-83-2017

Alien, M.R., Mutlow, C.T., Blumberg, G.M.C., Christy, J.R., McNider, R.T., Llewellyn-Jones, D.T., 1994. Global change detection. Nature 370, 24–24. https://doi.org/10.1038/370024b0

Anderson, M., 1997. A Two-Source Time-Integrated Model for Estimating Surface Fluxes Using Thermal Infrared Remote Sensing. Remote Sens. Environ. 60, 195–216. https://doi.org/10.1016/S0034-4257(96)00215-5

Anderson, M.C., Allen, R.G., Morse, A., Kustas, W.P., 2012. Use of Landsat thermal imagery in monitoring evapotranspiration and managing water resources. Remote Sens. Environ. 122, 50–65. https://doi.org/https://doi.org/10.1016/j.rse.2011.08.025

Baldridge, A.M., Hook, S.J., Grove, C.I., Rivera, G., 2009. The ASTER spectral library version 2.0. Remote Sens. Environ. 113, 711–715. https://doi.org/10.1016/j.rse.2008.11.007

Barker-Snook, I., Theocharous, E., Fox, N.P., 2017. 2016 comparison of IR brightness temperature measurements in support of satellite validation. Part 3: Sea surface temperature comparison of radiation thermometers.

Barker-Snook, I., Theocharous, E., Fox, N.P., 2016. comparison of IR brightness temperature measurements in support of satellite validation.

Barnes, W.L., Pagano, T.S., Salomonson, V. V, 1998. Prelaunch characteristics of the moderate resolution imaging spectroradiometer (MODIS) on EOS-AMI. IEEE Trans. Geosci. Remote Sens. 36, 1088–1100. https://doi.org/10.1109/36.700993

Barsi, J., Schott, J., Hook, S., Raqueno, N., Markham, B., Radocinski, R., 2014. Landsat-8 Thermal Infrared Sensor (TIRS) Vicarious Radiometric Calibration. Remote Sens. 6, 11607–11626. https://doi.org/10.3390/rs61111607

Barsi, J.A., Hook, S.J., Palluconi, F.D., Schott, J.R., Raqueno, N.G., 2006. Landsat TM and ETM+ thermal band calibration, in: Butler, J.J., Xiong, J. (Eds.), . p. 62960F. https://doi.org/10.1117/12.683212

Barsi, J.A., Montanaro, M., Thome, K., Raqueno, N.G., Hook, S.J., Anderson, C., Micijevic, E., 2022. Early radiometric performance of Landsat-9 Thermal Infrared Sensor, in: Butler, J.J., Xiong, X. (Jack), Gu, X. (Eds.), Earth Observing Systems XXVII. SPIE, p. 31. https://doi.org/10.1117/12.2634058

Barton, I.J., Zavody, A.M., O'Brien, D.M., Cutten, D.R., Saunders, R.W., Llewellyn-Jones, D.T., 1989. Theoretical algorithms for satellite-derived sea surface temperatures. J. Geophys. Res. Atmos. 94, 3365–3375. https://doi.org/10.1029/JD094iD03p03365

BECKER, F., LI, Z.-L., 1990. Towards a local split window method over land surfaces. Int. J. Remote Sens. 11, 369–393. https://doi.org/10.1080/01431169008955028

Berk, A., Anderson, G.P., Acharya, P.K., Chetwynd, J.H., Bernstein, L.S., Shettle, E.P., Matthew, M.W., Adler-Golden, S.M., 1999. MODTRAN4 USER'S MANUAL.

Berk, A., Anderson, G.P., Acharya, P.K., Shettle, E.P., 2008. MODTRAN5. 2.0. 0 user's manual. Spectr. Sci. Inc., Burlingt. MA, Air Force Res. Lab. Hanscom MA.

Bicheron, P., 2008. GLOBCOVER Products Description and Validation Report.

Bignell, K.J., 1970. The water-vapour infra-red continuum. Q. J. R. Meteorol. Soc. 96, 390–403. https://doi.org/10.1002/qj.49709640904

Bojinski, S., Verstraete, M., Peterson, T.C., Richter, C., Simmons, A., Zemp, M., 2014. The Concept of Essential Climate Variables in Support of Climate Research, Applications, and Policy. Bull. Am. Meteorol. Soc. 95, 1431–1443. https://doi.org/10.1175/BAMS-D-13-00047.1

Borbas, E.E., Seemann, S.W., Kern, A., Moy, L., Li, J., Gumley, L., Menzel, W.P., 2011. MODIS ATMOSPHERIC PROFILE RETRIEVAL ALGORITHM THEORETICAL BASIS DOCUMENT COLLECTION 6 Cooperative Institute for Meteorological Satellite Studies.

Boronina, A., Ramillien, G., 2008. Application of AVHRR imagery and GRACE measurements for calculation of actual evapotranspiration over the Quaternary aquifer (Lake Chad basin) and validation of groundwater models. J. Hydrol. 348, 98–109. https://doi.org/10.1016/j.jhydrol.2007.09.061

Brogniez, G., Pietras, C., Legrand, M., Dubuisson, P., Haeffelin, M., 2003. A High-Accuracy Multiwavelength Radiometer for In Situ Measurements in the Thermal Infrared. Part II: Behavior in Field Experiments. J. Atmos. Ocean. Technol. 20, 1023–1033. https://doi.org/10.1175/1520-0426(2003)20<1023:AHMRFI>2.0.CO;2

Calle, A., Sanz, J., Moclan, C., Casanova, J.L., Goldammer, J.G., Zengyuan, L., Xiantin, Q., 2006. Detection and monitoring of forest fires in China through the ENVISAT-AATSR sensor, in: DRAGON PROGRAMME MID-TERM RESULTS, PROCEEDINGS. ESA PUBLICATIONS DIVISION C/O ESTEC, 2200 AG NOORDWIJK, pp. 283–286.

Chang, T., Wu, X., 2012. On-board calibration radiance error correction for MetOp/AVHRR IR channel, in: Butler, J.J., Xiong, X., Gu, X. (Eds.), . p. 85100L. https://doi.org/10.1117/12.929619

Chedin, A., Scott, N.A., Berroir, A., 1982. A Single-Channel, Double-Viewing Angle Method for Sea Surface Temperature Determination from Coincident METEOSAT and TIROS-N Radiometric Measurements. J. Appl. Meteorol. 21, 613–618. https://doi.org/10.1175/1520-0450(1982)021<0613:ASCDVA>2.0.CO;2

Coll, C., Caselles, V., 1997. A split-window algorithm for land surface temperature from advanced very high resolution radiometer data: Validation and algorithm comparison. J. Geophys. Res. Atmos. https://doi.org/10.1029/97jd00929

Coll, C., Caselles, V., Galve, J.M., Valor, E., Niclòs, R., Sánchez, J.M., 2006. Evaluation of split-window and dual-angle correction methods for land surface temperature retrieval from Envisat/Advanced Along Track Scanning Radiometer (AATSR) data. J. Geophys. Res. Atmos. 111. https://doi.org/10.1029/2005JD006830

Coll, C., Caselles, V., Galve, J.M., Valor, E., Niclòs, R., Sánchez, J.M., Rivas, R., 2005. Ground measurements for the validation of land surface temperatures derived from AATSR and MODIS data. Remote Sens. Environ. 97, 288–300. https://doi.org/10.1016/j.rse.2005.05.007

Coll, C., Caselles, V., Sobrino, J.A., Valor, E., 1994. On the atmospheric dependence of the split-window equation for land surface temperature. Int. J. Remote Sens. 15, 105–122. https://doi.org/10.1080/01431169408954054

Coll, C., Caselles, V., Valor, E., Niclòs, R., 2012. Comparison between different sources of atmospheric profiles for land surface temperature retrieval from single channel thermal infrared data. Remote Sens. Environ. 117, 199–210. https://doi.org/10.1016/j.rse.2011.09.018

Coll, C., Galve, J.M., Sánchez, J.M., Caselles, V., 2010. Validation of landsat-7/ETM+ thermal-band calibration and atmospheric correction with ground-based measurements. IEEE Trans. Geosci. Remote Sens. 48, 547–555. https://doi.org/10.1109/TGRS.2009.2024934

Coll, C., Niclòs, R., Puchades, J., García-Santos, V., Galve, J.M., Pérez-Planells, L., Valor, E., Theocharous, E., 2019. Laboratory calibration and field measurement of land surface temperature and emissivity using thermal infrared multiband radiometers. Int. J. Appl. Earth Obs. Geoinf. 78, 227–239. https://doi.org/10.1016/j.jag.2019.02.002

Coll, C., Niclòs, R., Puchades, J., García-Santos, V., Perelló, M., Pérez-Planells, L., 2024. Demonstrating the Suitability of the Radiance- Based Method for Assessing the Accuracy of MODIS Land Surface Temperature Products. IEEE Trans. Geosci. Remote Sens. 62, 1–15. https://doi.org/10.1109/TGRS.2024.3454377

Coll, C., Wan, Z., Galve, J.M., 2009. Temperature-based and radiance-based validations of the V5 MODIS land surface temperature product. J. Geophys. Res. Atmos. 114, 1–15. https://doi.org/10.1029/2009JD012038

Darge, Y.M., Hailu, B.T., Muluneh, A.A., Kidane, T., 2019. Detection of geothermal anomalies using Landsat 8 TIRS data in Tulu Moye geothermal prospect, Main Ethiopian Rift. Int. J. Appl. Earth Obs. Geoinf. 74, 16–26. https://doi.org/10.1016/j.jag.2018.08.027

Dash, P., Göttsche, F.-M., Olesen, F.-S., Fischer, H., 2002. Land surface temperature and emissivity estimation from passive sensor data: Theory and practice-current trends. Int. J. Remote Sens. 23, 2563–2594. https://doi.org/10.1080/01431160110115041

Deschamps, P.Y., Phulpin, T., 1980. Atmospheric correction of infrared measurements of sea surface temperature using channels at 3.7, 11 and 12 ?m. Boundary-Layer Meteorol. 18, 131–143. https://doi.org/10.1007/BF00121320

Donlon, C., Berruti, B., Buongiorno, A., Ferreira, M.H., Féménias, P., Frerick, J., Goryl, P., Klein, U., Laur, H., Mavrocordatos, C., Nieke, J., Rebhan, H., Seitz, B., Stroede, J., Sciarra, R., 2012. The Global Monitoring for Environment and Security (GMES) Sentinel-3 mission. Remote Sens. Environ. 120, 37–57. https://doi.org/10.1016/j.rse.2011.07.024

Du, C., Ren, H., Qin, Q., Meng, J., Zhao, S., 2015. A Practical Split-Window Algorithm for Estimating Land Surface Temperature from Landsat 8 Data. Remote Sens. 7, 647–665. https://doi.org/10.3390/rs70100647

Er-Raki, S., Chehbouni, A., Hoedjes, J., Ezzahar, J., Duchemin, B., Jacob, F., 2008. Improvement of FAO-56 method for olive orchards through sequential assimilation of thermal infrared-based estimates of ET. Agric. Water Manag. 95, 309–321. https://doi.org/10.1016/j.agwat.2007.10.013

Fisher, J.B., Lee, B., Purdy, A.J., Halverson, G.H., Dohlen, M.B., Cawse-Nicholson, K., Wang, A., Anderson, R.G., Aragon, B., Arain, M.A., Baldocchi, D.D., Baker, J.M., Barral, H., Bernacchi, C.J., Bernhofer, C., Biraud, S.C., Bohrer, G., Brunsell, N., Cappelaere, B., Castro-Contreras, S., Chun, J., Conrad, B.J., Cremonese, E., Demarty, J., Desai, A.R., Ligne, A. De, Foltýnová, L., Goulden, M.L., Griffis, T.J., Grünwald, T., Johnson, M.S., Kang, M., Kelbe, D., Kowalska, N., Lim, J.H., Maïnassara, I., McCabe, M.F., Missik, J.E.C., Mohanty, B.P., Moore, C.E., Morillas, L., Morrison, R., Munger, J.W., Posse, G., Richardson, A.D., Russell, E.S., Ryu, Y., Sanchez-Azofeifa, A., Schmidt, M., Schwartz, E., Sharp, I., Šigut, L., Tang, Y., Hulley, G., Anderson, M., Hain, C., French, A., Wood, E., Hook, S., 2020. ECOSTRESS: NASA's Next Generation Mission to Measure Evapotranspiration From the International Space Station. Water Resour. Res. 56. https://doi.org/10.1029/2019WR026058

Galleguillos, M., Jacob, F., Prévot, L., French, A., Lagacherie, P., 2011. Comparison of two temperature differencing methods to estimate daily evapotranspiration over a Mediterranean vineyard watershed from ASTER data. Remote Sens. Environ. 115, 1326–1340. https://doi.org/10.1016/j.rse.2011.01.013

Galve, J.M., Coll, C., Caselles, V., Valor, E., 2008. An atmospheric radiosounding database for generating land surface temperature algorithms. IEEE Trans. Geosci. Remote Sens. 46, 1547–1557. https://doi.org/10.1109/TGRS.2008.916084

García-Llamas, P., Suárez-Seoane, S., Fernández-Guisuraga, J.M., Fernández-García, V., Fernández-Manso, A., Quintano, C., Taboada, A., Marcos, E., Calvo, L., 2019. Evaluation and comparison of Landsat 8, Sentinel-2 and Deimos-1 remote sensing indices for assessing burn severity in Mediterranean fire-prone ecosystems. Int. J. Appl. Earth Obs. Geoinf. 80, 137–144. https://doi.org/10.1016/j.jag.2019.04.006

García-Santos, V., Cuxart, J., Martínez-Villagrasa, D., Jiménez, M., Simó, G., 2018. Comparison of Three Methods for Estimating Land Surface Temperature from Landsat 8-TIRS Sensor Data. Remote Sens. 10, 1450. https://doi.org/10.3390/rs10091450

García-Santos, V., Valor, E., Caselles, V., Burgos, M.N., Coll, C., 2012. On the angular variation of thermal infrared emissivity of inorganic soils. J. Geophys. Res. Atmos. 117. https://doi.org/10.1029/2012JD017931

GCOS, 2022. The 2022 GCOS ECVs Requirements (GCOS 245).

GCOS, 2016. The Global Observing System for Climate: Implementation Needs. World Meteorol. Organ. 200, 341.

Gerace, A., Kleynhans, T., Eon, R., Montanaro, M., 2020. Towards an operational, split window-derived surface temperature product for the thermal infrared sensors onboard Landsat 8 and 9. Remote Sens. 12. https://doi.org/10.3390/rs12020224

Gerace, A., Montanaro, M., 2017. Derivation and validation of the stray light correction algorithm for the thermal infrared sensor onboard Landsat 8. Remote Sens. Environ. 191, 246–257. https://doi.org/10.1016/j.rse.2017.01.029

Gerhards, M., Schlerf, M., Mallick, K., Udelhoven, T., 2019. Challenges and future perspectives of multi-/Hyperspectral thermal infrared remote sensing for crop water-stress detection: A review. Remote Sens. 11. https://doi.org/10.3390/rs11101240

Ghent, D., Anand, J.S., Veal, K., Remedios, J., 2024. The Operational and Climate Land Surface Temperature Products from the Sea and Land Surface Temperature Radiometers on Sentinel-3A and 3B. Remote Sens. 16, 3403. https://doi.org/10.3390/rs16183403

Ghent, D., Kaduk, J., Remedios, J., Ardö, J., Balzter, H., 2010. Assimilation of land surface temperature into the land surface model JULES with an ensemble Kalman filter. J. Geophys. Res. Atmos. 115. https://doi.org/10.1029/2010JD014392

Ghent, D.J., Corlett, G.K., Göttsche, F. -M., Remedios, J.J., 2017. Global Land Surface Temperature From the Along-Track Scanning Radiometers. J. Geophys. Res. Atmos. 122. https://doi.org/10.1002/2017JD027161

Gillespie, A., Rokugawa, S., Matsunaga, T., Cothern, J.S., Hook, S., Kahle, A.B., 1998. A temperature and emissivity separation algorithm for advanced spaceborne thermal emission and reflection radiometer (ASTER) images. IEEE Trans. Geosci. Remote Sens. 36, 1113–1126. https://doi.org/10.1109/36.700995

Gillespie, A.R., 1986. Lithologic mapping of silicate rocks using TIMS, in: Workshop, T.T.D.U. (Ed.), .

Good, E.J., Aldred, F.M., Ghent, D.J., Veal, K.L., Jimenez, C., 2022. An Analysis of the Stability and Trends in the LST_cci Land Surface Temperature Datasets Over Europe. Earth Sp. Sci. 9. https://doi.org/10.1029/2022EA002317

Göttsche, F.-M., Olesen, F.-S., Bork-Unkelbach, A., 2013. Validation of land surface temperature derived from MSG/SEVIRI with in situ measurements at Gobabeb, Namibia. Int. J. Remote Sens. 34, 3069–3083. https://doi.org/10.1080/01431161.2012.716539

Göttsche, F.M., Olesen, F.S., Trigo, I.F., Bork-Unkelbach, A., Martin, M.A., 2016. Long term validation of land surface temperature retrieved from MSG/SEVIRI with continuous in-situ measurements in Africa. Remote Sens. 8. https://doi.org/10.3390/rs8050410

Goutorbe, J.-P., Lebel, T., Tinga, A., Bessemoulin, P., Brouwer, J., Dolman, A.J., Engman, E.T., Gash, J.H.C., Hoepffner, M., Kabat, P., Kerr, Y.H., Monteny, B., Prince, S., Said, F., Sellers, P., Wallace, J.S., 1994. HAPEX-Sahel: a large-scale study of land-atmosphere interactions in the semi-arid tropics. Ann. Geophys. 12, 53–64. https://doi.org/10.1007/s00585-994-0053-0

Guillevic, P., Göttsche, F., Nickeson, J., Hulley, G., Ghent, D., Yu, Y., Trigo, I., Hook, S., Sobrino, J.A., Remedios, J., Román, M., Camacho, F., 2018. Land Surface Temperature Product Validation Best Practice Protocol. Version 1.1 58. https://doi.org/10.5067/doc/ceoswgcv/lpv/lst.001

Guo, J., Ren, H., Zheng, Y., Lu, S., Dong, J., 2020. Evaluation of land surface temperature retrieval from landsat 8/TIRS images before and after stray light correction using the SURFRAD dataset. Remote Sens. 12. https://doi.org/10.3390/rs12061023

Hall, T.M., Cho, J.Y.N., Kelley, M.C., Hocking, W.K., 1992. A re-evaluation of the Stokes drift in the polar summer mesosphere. J. Geophys. Res. Atmos. 97, 887–897. https://doi.org/10.1029/91JD02835

Hollmann, R., Merchant, C.J., Saunders, R., Downy, C., Buchwitz, M., Cazenave, A., Chuvieco, E., Defourny, P., Leeuw, G. De, Forsberg, R., Holzer-Popp, T., Paul, F., Sandven, S., Sathyendranath, S., Roozendael, M. Van, Wagner, W., 2013. The ESA climate change initiative: Satellite data records for essential climate variables. Bull. Am. Meteorol. Soc. 94, 1541–1552. https://doi.org/10.1175/BAMS-D-11-00254.1

Hook, S.J., Chander, G., Barsi, J.A., Alley, R.E., Abtahi, A., Palluconi, F.D., Markham, B.L., Richards, R.C., Schladow, S.G., Helder, D.L., 2004. In-flight validation and recovery of water surface temperature with Landsat-5 thermal infrared data using an automated high-altitude lake validation site at Lake Tahoe. IEEE Trans. Geosci. Remote Sens. 42, 2767–2776. https://doi.org/10.1109/TGRS.2004.839092

Hook, S.J., Myers, J.J., Thome, K.J., Fitzgerald, M., Kahle, A.B., 2001. The MODIS/ASTER airborne simulator (MASTER) — a new instrument for earth science studies. Remote Sens. Environ. 76, 93–102. https://doi.org/10.1016/S0034-4257(00)00195-4

Hulley, G., Malakar, N., Hughes, T., Islam, T., Hook, S., 2016. Moderate Resolution Imaging Spectroradiometer (MODIS) MOD21 Land Surface Temperature and Emissivity Algorithm Theoretical Basis Document (ATBD).

Hulley, G.C., Gottsche, F.M., Rivera, G., Hook, S.J., Freepartner, R.J., Martin, M.A., Cawse-Nicholson, K., Johnson, W.R., 2021. Validation and Quality Assessment of the ECOSTRESS Level-2 Land Surface Temperature and Emissivity Product. IEEE Trans. Geosci. Remote Sens. 1–23. https://doi.org/10.1109/TGRS.2021.3079879

Hulley, G.C., Hook, S.J., 2012. A radiance-based method for estimating uncertainties in the Atmospheric Infrared Sounder (AIRS) land surface temperature product. J. Geophys. Res. Atmos. 117. https://doi.org/10.1029/2012JD018102

Irons, J.R., Dwyer, J.L., Barsi, J.A., 2012. The next Landsat satellite: The Landsat Data Continuity Mission. Remote Sens. Environ. 122, 11–21. https://doi.org/10.1016/j.rse.2011.08.026

Jacob, F., Schmugge, T., Olioso, A., French, A., Courault, D., Ogawa, K., Petitcolin, F., Chehbouni, G., Pinheiro, A., Privette, J., 2008. Modeling and Inversion in Thermal Infrared Remote Sensing over Vegetated Land Surfaces, in: Advances in Land Remote Sensing. Springer Netherlands, Dordrecht, pp. 245–291. https://doi.org/10.1007/978-1-4020-6450-0_10

JCGM, 2008. Evaluation of measurement data-Guide to the expression of uncertainty in measurement Évaluation des données de mesure-Guide pour l'expression de l'incertitude de mesure.

Jiang, G.-M., Li, Z.-L., 2008. Remote sensing and sensors; (280.4991) Passive remote sensing; (120.3940) Metrology; (120.5630) Radiometry; (100.3190) Inverse problems.

Jiménez-Muñoz, J.C., Sobrino, J.A., 2008. Split-window coefficients for land surface temperature retrieval from low-resolution thermal infrared sensors. IEEE Geosci. Remote Sens. Lett. 5, 806–809. https://doi.org/10.1109/LGRS.2008.2001636

Jiménez-Muñoz, J.C., Sobrino, J.A., Skokovic, D., Mattar, C., Cristobal, J., 2014. Land surface temperature retrieval methods from landsat-8 thermal infrared sensor data. IEEE Geosci. Remote Sens. Lett. 11, 1840–1843. https://doi.org/10.1109/LGRS.2014.2312032

Johnson, W.R., Hook, S.J., Mouroulis, P., Wilson, D.W., Gunapala, S.D., Hill, C.J., Mumolo, J.M., Realmuto, V., Eng, B.T., 2009. Towards HyTES: an airborne thermal imaging spectroscopy instrument, in: Shen, S.S., Lewis, P.E. (Eds.), . p. 745706. https://doi.org/10.1117/12.826416

Johnson, W.R., Hook, S.J., Mouroulis, P., Wilson, D.W., Gunapala, S.D., Realmuto, V., Lamborn, A., Paine, C., Mumolo, J.M., Eng, B.T., 2011. HyTES: Thermal imaging spectrometer development, in: 2011 Aerospace Conference. IEEE, pp. 1–8. https://doi.org/10.1109/AERO.2011.5747394

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Jenne, R., Joseph, D., 1996. The NCEP/NCAR 40-Year Reanalysis Project. Bull. Am. Meteorol. Soc. 77, 437–471. https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2

Koetz, B., Bastiaanssen, W., Berger, M., Defourney, P., Bello, U. Del, Drusch, M., Drinkwater, M., Duca, R., Fernandez, V., Ghent, D., Guzinski, R., Hoogeveen, J., Hook, S., Lagouarde, J.-P., Lemoine, G., Manolis, I., Martimort, P., Masek, J., Massart, M., Notarnicola, C., Sobrino, J., Udelhoven, T., 2018. High Spatio-Temporal Resolution Land Surface Temperature Mission - a Copernicus Candidate Mission in Support of Agricultural Monitoring, in: IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium. IEEE, pp. 8160–8162. https://doi.org/10.1109/IGARSS.2018.8517433

Kondratyev, K. Y., 1969, Radiation in the Atmosphere. (Academic Press, New York).

Krishnan, P., Meyers, T.P., Hook, S.J., Heuer, M., Senn, D., Dumas, E.J., 2020. Intercomparison of in situ sensors for ground-based land surface temperature measurements. Sensors (Switzerland) 20, 1–26. https://doi.org/10.3390/s20185268

Lagouarde, J.-P., Bhattacharya, B.K., Crébassol, P., Gamet, P., Adlakha, D., Murthy, C.S., Singh, S.K., Mishra, M., Nigam, R., Raju, P. V, Babu, S.S., Shukla, M. V, Pandya, M.R., Boulet, G., Briottet, X., Dadou, I., Dedieu, G., Gouhier, M., Hagolle, O., Irvine, M., Jacob, F., Kumar, K.K., Laignel, B., Maisongrande, P., Mallick, K., Olioso, A., Ottlé, C., Roujean, J.-L., Sobrino, J., Ramakrishnan, R., Sekhar, M., Sarkar, S.S., 2019. Indo-French High-resolution Thermal Infrared Space Mission for Earth Natural Resources Assessment and Monitoring - Concept and Definition of TRISHNA, in: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. ISPRS-GEOGLAM-ISRS Joint Int. Workshop on "Earth Observations for Agricultural Monitoring," pp. 403–407. https://doi.org/10.5194/isprs-archives-XLII-3-W6-403-2019

Lagouarde, J.-P., Bhattacharya, B.K., Crebassol, P., Gamet, P., Babu, S.S., Boulet, G., Briottet, X., Buddhiraju, K.M., Cherchali, S., Dadou, I., Dedieu, G., Gouhier, M., Hagolle, O., Irvine, M., Jacob, F., Kumar, A., Kumar, K.K., Laignel, B., Mallick, K., Murthy, C.S., Olioso, A., Ottle, C., Pandya, M.R., Raju, P. V., Roujean, J.-L., Sekhar, M., Shukla, M. V., Singh, S.K., Sobrino, J., Ramakrishnan, R., 2018. The Indian-French Trishna Mission: Earth Observation in the Thermal Infrared with High Spatio-Temporal Resolution, in: IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium. IEEE, pp. 4078–4081. https://doi.org/10.1109/IGARSS.2018.8518720

Lagouarde, J.-P., Irvine, M., Dupont, S., 2015. Atmospheric turbulence induced errors on measurements of surface temperature from space. Remote Sens. Environ. 168, 40–53. https://doi.org/10.1016/j.rse.2015.06.018

Lambin, E.F., Ehrlich, D., 1997. Land-cover changes in sub-saharan Africa (1982–1991): Application of a change index based on remotely sensed surface temperature and vegetation indices at a continental scale. Remote Sens. Environ. 61, 181–200. https://doi.org/10.1016/S0034-4257(97)00001-1

Laraby, K.G., Schott, J.R., Raqueno, N., 2016. Developing a confidence metric for the Landsat land surface temperature product, in: Velez-Reyes, M., Messinger, D.W. (Eds.), . p. 98400C. https://doi.org/10.1117/12.2222582

Legrand, M., Pietras, C., Brogniez, G.É.R., Abuhassan, N.K., Michae¨, M., Sicard, M., 2000. A High-Accuracy Multiwavelength Radiometer for In Situ Measurements in the Thermal Infrared. Part I: Characterization of the Instrument.

Li, H., Li, R., Yang, Y., Cao, B., Bian, Z., Hu, T., Du, Y., Sun, L., Liu, Q., 2021. Temperature-Based and Radiance-Based Validation of the Collection 6 MYD11 and MYD21 Land Surface Temperature Products Over Barren Surfaces in Northwestern China. IEEE Trans. Geosci. Remote Sens. 59, 1794–1807. https://doi.org/10.1109/TGRS.2020.2998945

Li, R., Li, H., Hu, T., Bian, Z., Liu, F., Cao, B., Du, Y., Sun, L., Liu, Q., 2023. Land Surface Temperature Retrieval From Sentinel-3A SLSTR Data: Comparison Among Split-Window, Dual-Window, Three-Channel, and Dual-Angle Algorithms. IEEE Trans. Geosci. Remote Sens. 61, 1–14. https://doi.org/10.1109/TGRS.2023.3288584

Llewellyn-Jones, D., Remedios, J., 2012. The Advanced Along Track Scanning Radiometer (AATSR) and its predecessors ATSR-1 and ATSR-2: An introduction to the special issue. Remote Sens. Environ. 116, 1–3. https://doi.org/10.1016/j.rse.2011.06.002

Malakar, N.K., Hulley, G.C., Hook, S.J., Laraby, K., Cook, M., Schott, J.R., 2018. An Operational Land Surface Temperature Product for Landsat Thermal Data: Methodology and Validation. IEEE Trans. Geosci. Remote Sens. 56, 5717–5735. https://doi.org/10.1109/TGRS.2018.2824828

Martin, M., 2020. CCI Land Surface Temperature Product Validation and Intercomparison Report (PVIR).

Martin, M.A., Ghent, D., Pires, A.C., Göttsche, F.M., Cermak, J., Remedios, J.J., 2019. Comprehensive in situ validation of five satellite land surface temperature data sets over multiple stations and years. Remote Sens. 11. https://doi.org/10.3390/rs11050479

Masuoka, E., Fleig, A., Wolfe, R.E., Patt, F., 1998. Key characteristics of MODIS data products. IEEE Trans. Geosci. Remote Sens. 36, 1313–1323. https://doi.org/10.1109/36.701081

Maul, G.A., Sidran, M., 1973. Atmospheric effects on ocean surface temperature sensing from the NOAA satellite scanning radiometer. J. Geophys. Res. 78, 1909–1916. https://doi.org/10.1029/JC078i012p01909

McMillin, L.M., 1975. Estimation of sea surface temperatures from two infrared window measurements with different absorption. J. Geophys. Res. 80, 5113–5117. https://doi.org/10.1029/JC080i036p05113

Meerdink, S.K., Hook, S.J., Roberts, D.A., Abbott, E.A., 2019. The ECOSTRESS spectral library version 1.0. Remote Sens. Environ. 230, 111196. https://doi.org/10.1016/j.rse.2019.05.015

Meng, X., Cheng, J., Guo, H., Guo, Y., Yao, B., 2022. Accuracy Evaluation of the Landsat 9 Land Surface Temperature Product. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 15, 8694–8703. https://doi.org/10.1109/JSTARS.2022.3212736

Meng, X., Cheng, J., Zhao, S., Liu, S., Yao, Y., 2019. Estimating land surface temperature from Landsat-8 data using the NOAA JPSS enterprise algorithm. Remote Sens. 11. https://doi.org/10.3390/rs11020155

Mokhtari, A., Noory, H., Pourshakouri, F., Haghighatmehr, P., Afrasiabian, Y., Razavi, M., Fereydooni, F., Naeni, A.S., 2019. Calculating potential evapotranspiration and single crop coefficient based on energy balance equation using Landsat 8 and Sentinel-2. ISPRS J. Photogramm. Remote Sens. 154, 231–245. https://doi.org/10.1016/j.isprsjprs.2019.06.011

Montanaro, M., Gerace, A., Lunsford, A., Reuter, D., 2014a. Stray Light Artifacts in Imagery from the Landsat 8 Thermal Infrared Sensor. Remote Sens. 6, 10435–10456. https://doi.org/10.3390/rs61110435

Montanaro, M., Gerace, A., Rohrbach, S., 2015. Toward an operational stray light correction for the Landsat 8 Thermal Infrared Sensor. Appl. Opt. 54, 3963. https://doi.org/10.1364/AO.54.003963

Montanaro, M., Levy, R., Markham, B., 2014b. On-Orbit Radiometric Performance of the Landsat 8 Thermal Infrared Sensor. Remote Sens. 6, 11753–11769. https://doi.org/10.3390/rs61211753

Montanaro, M., McCorkel, J., Tveekrem, J., Stauder, J., Mentzell, E., Lunsford, A., Hair, J., Reuter, D., 2022. Landsat 9 Thermal Infrared Sensor 2 (TIRS-2) Stray Light Mitigation and Assessment. IEEE Trans. Geosci. Remote Sens. 60, 1–8. https://doi.org/10.1109/TGRS.2022.3177312

Niclòs, R., Caselles, V., Valor, E., Coll, C., Sánchez, J.M., 2009. A simple equation for determining sea surface emissivity in the 3-15 μm region. Int. J. Remote Sens. 30, 1603–1619. https://doi.org/10.1080/01431160802541523

Niclos, R., Dona, C., Valor, E., Bisquert, M., 2014. Thermal-infrared spectral and angular characterization of crude oil and seawater emissivities for oil slick identification. IEEE Trans. Geosci. Remote Sens. 52, 5387–5395. https://doi.org/10.1109/TGRS.2013.2288517

Niclòs, R., Galve, J.M., Valiente, J.A., Estrela, M.J., Coll, C., 2011. Accuracy assessment of land surface temperature retrievals from MSG2-SEVIRI data. Remote Sens. Environ. 115, 2126–2140. https://doi.org/10.1016/j.rse.2011.04.017

Niclòs, R., Perelló, M., Puchades, J., Coll, C., Valor, E., 2023. Evaluating Landsat-9 TIRS-2 calibrations and land surface temperature retrievals against ground measurements using multi-instrument spatial and temporal sampling along transects. Int. J. Appl. Earth Obs. Geoinf. 125, 103576. https://doi.org/10.1016/j.jag.2023.103576

Niclòs, R., Pérez-Planells, L., Coll, C., Valiente, J.A., Valor, E., 2018. Evaluation of the S-NPP VIIRS land surface temperature product using ground data acquired by an autonomous system at a rice paddy. ISPRS J. Photogramm. Remote Sens. 135, 1–12. https://doi.org/10.1016/j.isprsjprs.2017.10.017

Niclòs, R., Puchades, J., Coll, C., Barberà, M.J., Pérez-Planells, L., Valiente, J.A., Sánchez, J.M., 2021. Evaluation of Landsat-8 TIRS data recalibrations and land surface temperature split-window algorithms over a homogeneous crop area with different phenological land covers. ISPRS J. Photogramm. Remote Sens. 174, 237–253. https://doi.org/10.1016/j.isprsjprs.2021.02.005

Niclòs, R., Valiente, J.A., Barberà, M.J., Coll, C., 2015. An autonomous system to take angular thermal-infrared measurements for validating satellite products. Remote Sens. 7, 15269–15294. https://doi.org/10.3390/rs71115269

Niclòs, R., Valor, E., Caselles, V., Coll, C., Sánchez, J.M., 2005. In situ angular measurements of thermal infrared sea surface emissivity - Validation of models. Remote Sens. Environ. 94, 83–93. https://doi.org/10.1016/j.rse.2004.09.002

Olioso, A., Soria, G., Sobrino, J., Duchemin, B., 2007. Evidence of Low Land Surface Thermal Infrared Emissivity in the Presence of Dry Vegetation. IEEE Geosci. Remote Sens. Lett. 4, 112–116. https://doi.org/10.1109/LGRS.2006.885857

Palluconi, F.D., Hoover, G., Alley, R.E., Jentoft-Nilsen, M., Thompson, T., 1999. Advanced Spaceborne Thermal Emission \& Reflection Radiometer Algorithm Theoretical Basic Document for: An Atmospheric Correction Method for ASTER Thermal Radiometry Over Land.

Parra, J., Sobrino, J., Morales, L., Castellaro, G., Uribe, J., Gaete, N., 2006. Aplicación de un Algoritmo de Split-Window para la Estimación de la Temeratura de la Superficie Terrestre desde Datos AVHRR-NOAA. Agric. Técnica 66. https://doi.org/10.4067/S0365-28072006000400007

Pearlman, A., Efremova, B., Montanaro, M., Lunsford, A., Reuter, D., McCorkel, J., 2022. Landsat 9 Thermal Infrared Sensor 2 On-Orbit Calibration and Initial Performance. IEEE Trans. Geosci. Remote Sens. 60, 1–8. https://doi.org/10.1109/TGRS.2022.3183551

Peres, L. de F., Lucena, A.J. de, Rotunno Filho, O.C., França, J.R. de A., 2018. The urban heat island in Rio de Janeiro, Brazil, in the last 30 years using remote sensing data. Int. J. Appl. Earth Obs. Geoinf. 64, 104–116. https://doi.org/10.1016/j.jag.2017.08.012

Pérez-Planells, L., Niclòs, R., Puchades, J., Coll, C., Göttsche, F.M., Valiente, J.A., Valor, E., Galve, J.M., 2021. Validation of sentinel-3 slstr land surface temperature retrieved by the operational product and comparison with explicitly emissivity-dependent algorithms. Remote Sens. 13. https://doi.org/10.3390/rs13112228

Perez-Planells, L., Niclos, R., Valor, E., Gottsche, F.-M., 2022. Retrieval of Land Surface Emissivities Over Partially Vegetated Surfaces From Satellite Data Using Radiative Transfer Models. IEEE Trans. Geosci. Remote Sens. 60, 1–21. https://doi.org/10.1109/TGRS.2022.3224639

Pinheiro, A.C.T., Privette, J.L., Mahoney, R., Tucker, C.J., 2004. Directional effects in a daily AVHRR land surface temperature dataset over Africa. IEEE Trans. Geosci. Remote Sens. 42, 1941–1954. https://doi.org/10.1109/TGRS.2004.831886

Plummer, S., Lecomte, P., Doherty, M., 2017. The ESA Climate Change Initiative (CCI): A European contribution to the generation of the Global Climate Observing System. Remote Sens. Environ. 203, 2–8. https://doi.org/10.1016/j.rse.2017.07.014

Prabhakara, C., Dalu, G., Kunde, V.G., 1974. Estimation of sea surface temperature from remote sensing in the 11- to 13-μm window region. J. Geophys. Res. 79, 5039–5044. https://doi.org/10.1029/JC079i033p05039

Prata, A., 2003. Land surface temperature measurement from space: Validation of the AATSR land surface temperature product. CSIRO.

Prata, A.J., 1994. Land surface temperatures derived from the advanced very high resolution radiometer and the along-track scanning radiometer: 2. Experimental results and validation of AVHRR algorithms. J. Geophys. Res. Atmos. 99, 13025–13058. https://doi.org/10.1029/94JD00409

Prata, F., 2002. Land Surface Temperature Measurement from Space: AATSR Algorithm Theoretical Basis Document. Contract Rep. to ESA, CSIRO Atmos. Res. Aspendale, Victoria, Aust. 2002, 1–34.

Price, J.C., 1983. Estimating surface temperatures from satellite thermal infrared data—A simple formulation for the atmospheric effect. Remote Sens. Environ. 13, 353–361. https://doi.org/10.1016/0034-4257(83)90036-6

Qin, Z., Karnieli, A., Berliner, P., 2001. A mono-window algorithm for retrieving land surface temperature from Landsat TM data and its application to the Israel-Egypt border region. Int. J. Remote Sens. 22, 3719–3746. https://doi.org/10.1080/01431160010006971

Roy, P., Guha, A., Kumar, K.V., 2015. An approach of surface coal fire detection from ASTER and Landsat-8 thermal data: Jharia coal field, India. Int. J. Appl. Earth Obs. Geoinf. 39, 120–127. https://doi.org/10.1016/j.jag.2015.03.009

Rozenstein, O., Qin, Z., Derimian, Y., Karnieli, A., 2014. Derivation of Land Surface Temperature for Landsat-8 TIRS Using a Split Window Algorithm. Sensors 14, 5768–5780. https://doi.org/10.3390/s140405768

Rubio, E., Caselles, V., Badenas, C., 1997. Emissivity measurements of several soils and vegetation types in the 8-14 μm wave band: Analysis of two field methods. Remote Sens. Environ. 59, 490–521. https://doi.org/10.1016/S0034-4257(96)00123-X

Rubio, E., Caselles, V., Coll, C., Valour, E., Sospedra, F., 2003. Thermal-infrared emissivites of natural surfaces: Improvements on the experimental set-up and new measurements. Int. J. Remote Sens. 24, 5379–5390. https://doi.org/10.1080/0143116031000102412

Salisbury, J.W., D'Aria, D.M., 1992. Emissivity of terrestrial materials in the 8–14 μm atmospheric window. Remote Sens. Environ. 42, 83–106. https://doi.org/10.1016/0034-4257(92)90092-X

Sánchez, J.M., Kustas, W.P., Caselles, V., Anderson, M.C., 2008. Modelling surface energy fluxes over maize using a two-source patch model and radiometric soil and canopy temperature observations. Remote Sens. Environ. 112, 1130–1143. https://doi.org/10.1016/j.rse.2007.07.018

Schmugge, T., Hook, S.J., Coll, C., 1998. Recovering Surface Temperature and Emissivity from Thermal Infrared Multispectral Data. Remote Sens. Environ. 65, 121–131. https://doi.org/10.1016/S0034-4257(98)00023-6

Schott, J.R., Hook, S.J., Barsi, J.A., Markham, B.L., Miller, J., Padula, F.P., Raqueno, N.G., 2012. Thermal infrared radiometric calibration of the entire Landsat 4, 5, and 7 archive (1982-2010). Remote Sens. Environ. 122, 41–49. https://doi.org/10.1016/j.rse.2011.07.022

Seemann, S., Borbas, E., Li, J., Menzel, W., Gumley, L., 2006. MODIS Atmospheric Profile Retrieval Algorithm Theoretical Basis Document.

Sellers, P.J., Heiser, M.D., Hall, F.G., Goetz, S., Strebel, D.E., Verma, S.B., Desjardins, R.L., Schuepp, P.M., MacPherson, J.I., 1995. Effects of spatial variability in topography, vegetation cover and soil moisture on area-averaged surface fluxes: A case study using the FIFE 1989 data. J. Geophys. Res. 100, 25607–25629.

Sismanidis, P., Bechtel, B., Perry, M., Ghent, D., 2022. The Seasonality of Surface Urban Heat Islands across Climates. Remote Sens. 14. https://doi.org/10.3390/rs14102318

Slater, P.N., Biggar, S.F., Thome, K.J., Gellman, D.I., Spyak, P.R., 1996. Vicarious Radiometric Calibrations of EOS Sensors. J. Atmos. Ocean. Technol. 13, 349–359. https://doi.org/10.1175/1520-0426(1996)013<0349:VRCOES>2.0.CO;2

Smith, D., Hunt, S.E., Etxaluze, M., Peters, D., Nightingale, T., Mittaz, J., Woolliams, E.R., Polehampton, E., 2021. Traceability of the Sentinel-3 SLSTR Level-1 Infrared Radiometric Processing. Remote Sens. 13, 374. https://doi.org/10.3390/rs13030374

Snyder, W.C., Wan, Z., Zhang, Y., Feng, Y.-Z., 1998. Classification-based emissivity for land surface temperature measurement from space. Int. J. Remote Sens. 19, 2753–2774. https://doi.org/10.1080/014311698214497

Sobrino, J.A., Jiménez-Muñoz, J.C., Sòria, G., Ruescas, A.B., Danne, O., Brockmann, C., Ghent, D., Remedios, J., North, P., Merchant, C., Berger, M., Mathieu, P.P., Göttsche, F.M., 2016. Synergistic use of MERIS and AATSR as a proxy for estimating Land Surface Temperature from Sentinel-3 data. Remote Sens. Environ. 179, 149–161. https://doi.org/10.1016/j.rse.2016.03.035

Sobrino, J.A., Romaguera, M., 2004. Land surface temperature retrieval from MSG1-SEVIRI data. Remote Sens. Environ. 92, 247–254. https://doi.org/10.1016/j.rse.2004.06.009

Sobrino, J.A., Zhao-Liang Li, Stoll, M.P., 1993. Impact of the atmospheric transmittance and total water vapor content in the algorithms for estimating satellite sea surface temperatures. IEEE Trans. Geosci. Remote Sens. 31, 946–952. https://doi.org/10.1109/36.263765

Strahler, A., Gopal, S., Lambin, E., Moody, A., 1999. MODIS Land Cover Product Algorithm Theoretical Basis Document (ATBD) MODIS Land Cover and Land-Cover Change.

Sun, D., Pinker, R.T., 2003. Estimation of land surface temperature from a Geostationary Operational Environmental Satellite (GOES-8). J. Geophys. Res. Atmos. 108. https://doi.org/10.1029/2002JD002422

Theocharous, E., Fox, N.P., Barker-Snook, I., Niclòs, R., Santos, V.G., Minnett, P.J., Göttsche, F.M., Poutier, L., Morgan, N., Nightingale, T., Wimmer, W., Høyer, J., Zhang, K., Yang, M., Guan, L., Arbelo, M., Donlon, C.J., 2019. The 2016 CEOS infrared radiometer comparison: Part II: Laboratory comparison of radiation thermometers. J. Atmos. Ocean. Technol. 36, 1079–1092. https://doi.org/10.1175/JTECH-D-18-0032.1

Theocharous, E., Snook, I.B., Fox, N.P., 2017. NPL REPORT ENV 12 2016 comparison of IR brightness temperature measurements in support of satellite validation. Part 1: Blackbody Laboratory comparison.

Theocharous, E., Usadi, E., Fox, N.P., 2010. CEOS comparison of IR brightness temperature measurements in support of satellite validation. Part I: Laboratory and Ocean surface temperature comparison of radiation thermometers.

Tonooka, H., Palluconi, F.D., Hook, S.J., Matsunaga, T., 2005. Vicarious calibration of ASTER thermal infrared bands. IEEE Trans. Geosci. Remote Sens. 43, 2733–2746. https://doi.org/10.1109/TGRS.2005.857885

Trigo, I.F., Boussetta, S., Viterbo, P., Balsamo, G., Beljaars, A., Sandu, I., 2015. Comparison of model land skin temperature with remotely sensed estimates and assessment of surface-atmosphere coupling. J. Geophys. Res. Atmos. 120. https://doi.org/10.1002/2015JD023812

Turpie, K., Turpie, K.R., Casey, K., Crawford, C.J., Guild, L.S., Kieffer, H., Lin, G., Kokaly, R., Shrestha, A., Anderson, C., Chandra, S.N., Green, R., Hook, S., Lukashin, C., Thome, K., Wolfe, R.E., 2023. Calibration and Validation for the Surface Biology and Geology (SBG) Mission Concept: Challenges of a Multi-Sensor System for Imaging Spectroscopy and Thermal Imagery. https://doi.org/10.22541/essoar.167397459.95085418/v1

Valor, E., Caselles, V., 2005. Validation of the vegetation cover method for land surface emissivity estimation. Recent Adv. Therm. Remote Sens.

Valor, E., Caselles, V., 1996. Mapping land surface emissivity from NDVI: Application to European, African, and South American areas. Remote Sens. Environ. 57, 167–184. https://doi.org/10.1016/0034-4257(96)00039-9

Vidal, T.H.G., Gamet, P., Olioso, A., Jacob, F., 2022. Optimizing TRISHNA TIR channels configuration for improved land surface temperature and emissivity measurements. Remote Sens. Environ. 272, 112939. https://doi.org/10.1016/j.rse.2022.112939

Wan, Z., 2014. New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product. Remote Sens. Environ. 140, 36–45. https://doi.org/10.1016/j.rse.2013.08.027

Wan, Z., 1999. MODIS Land-Surface Temperature Algorithm Theoretical Basis Document (LST ATBD).

Wan, Z., Dozier, J., 1996. A generalized split-window algorithm for retrieving land-surface temperature from space. IEEE Trans. Geosci. Remote Sens. 34, 892–905. https://doi.org/10.1109/36.508406

Wan, Z., Li, Z.L., 2008. Radiance-based validation of the V5 MODIS land-surface temperature product. Int. J. Remote Sens. 29, 5373–5395. https://doi.org/10.1080/01431160802036565

Wan, Z., Zhang, Y., Zhang, Q., Li, Z., 2002. Validation of the land-surface temperature products retrieved from Terra Moderate Resolution Imaging Spectroradiometer data. Remote Sens. Environ. 83, 163–180. https://doi.org/10.1016/S0034-4257(02)00093-7

Wang, M., Li, M., Zhang, Zhengjia, Hu, T., He, G., Zhang, Zhaoming, Wang, G., Li, H., Tan, J., Liu, X., 2023. Land Surface Temperature Retrieval From Landsat 9 TIRS-2 Data Using Radiance-Based Split-Window Algorithm. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 16, 1100–1112. https://doi.org/10.1109/JSTARS.2022.3232621

Wang, W., Liang, S., Meyers, T., 2008. Validating MODIS land surface temperature products using long-term nighttime ground measurements. Remote Sens. Environ. 112, 623–635. https://doi.org/10.1016/j.rse.2007.05.024

Weng, Q., 2009. Thermal infrared remote sensing for urban climate and environmental studies: Methods, applications, and trends. ISPRS J. Photogramm. Remote Sens. 64, 335–344. https://doi.org/10.1016/j.isprsjprs.2009.03.007

Wilrich, P.T., 2007. Robust estimates of the theoretical standard deviation to be used in interlaboratory precision experiments. Accredit. Qual. Assur. 12, 231–240. https://doi.org/10.1007/s00769-006-0240-7

Yamada, Y., Harris, S., Hayes, M., Simpson, R., Wimmer, W., Holmes, R., Nightingale, T., Lee, A., Jepsen, N., Morgan, N., Göttsche, F.-M., Niclòs, R., Perelló, M., Donlon, C., Fox, N., 2023a. CEOS International Thermal Infrared RadiometerComparison: Part I: Laboratory Comparison of Radiometers and Blackbodies. Submitted toJournal of Atmospheric and Oceanic Technology.

Yamada, Y., Harris, S., Hayes, M., Simpson, R., Wimmer, W., Holmes, R., Nightingale, T., Lee, A., Jepsen, N., Morgan, N., Göttsche, F.-M., Niclòs, R., Perelló, M., Donlon, C., Fox, N., 2023b. Results from CEOS International Thermal Infrared Radiometer Inter-comparison (CRIC) Part 2 of 3: Laboratory Comparison of Radiometers.

Ye, X., Ren, H., Zhu, J., Fan, W., Qin, Q., 2022. Split-Window Algorithm for Land Surface Temperature Retrieval From Landsat-9 Remote Sensing Images. IEEE Geosci. Remote Sens. Lett. 19, 1–5. https://doi.org/10.1109/LGRS.2022.3184980

Yu, X., Guo, X., Wu, Z., 2014. Land Surface Temperature Retrieval from Landsat 8 TIRS—Comparison between Radiative Transfer Equation-Based Method, Split Window Algorithm and Single Channel Method. Remote Sens. 6, 9829–9852. https://doi.org/10.3390/rs6109829

Zhang, S., Duan, S.B., Li, Z.L., Huang, C., Wu, H., Han, X.J., Leng, P., Gao, M., 2019. Improvement of split-window algorithm for land surface temperature retrieval from sentinel-3A SLSTR data over barren surfaces using ASTER GED product. Remote Sens. 11. https://doi.org/10.3390/rs11243025

Zheng, Y., Ren, H., Guo, J., Ghent, D., Tansey, K., Hu, X., Nie, J., Chen, S., 2019. Land Surface Temperature Retrieval from Sentinel-3A Sea and Land Surface Temperature Radiometer, Using a Split-Window Algorithm. Remote Sens. 11, 650. https://doi.org/10.3390/rs11060650

Zhou, J., Zhang, X., Zhan, W., Gottsche, F.-M., Liu, S., Olesen, F.-S., Hu, W., Dai, F., 2017. A Thermal Sampling Depth Correction Method for Land Surface Temperature Estimation From Satellite Passive Microwave Observation Over Barren Land. IEEE Trans. Geosci. Remote Sens. 55, 4743–4756. https://doi.org/10.1109/TGRS.2017.2698828

ANEXOS

ANEXO A

Laboratory calibration and field measurement of land surface temperature and emissivity using thermal infrared multiband radiometers

Int J Appl Earth Obs Geoinformation 78 (2019) 227–239

Contents lists available at ScienceDirect

Int J Appl Earth Obs Geoinformation

journal homepage: www.elsevier.com/locate/jag

Laboratory calibration and field measurement of land surface temperature and emissivity using thermal infrared multiband radiometers

César Coll[a,*], Raquel Niclòs[a], Jesús Puchades[a], Vicente García-Santos[b], Joan Miquel Galve[c], Lluís Pérez-Planells[a], Enric Valor[a], Evangelos Theocharous[d]

[a] Department of Earth Physics and Thermodynamics, Faculty of Physics, University of Valencia. C/Dr. Moliner, 50. 46100, Burjassot (Valencia), Spain
[b] EOLAB SPAIN S.L., Parc Científic de la Universitat de València, C/Catedrático Agustín Escardino, 9. 46980, Paterna (Valencia), Spain
[c] Applied Physics Department, Regional Development Institute, University of Castilla-La Mancha, Campus Universitario s/n, 02071, Albacete, Spain
[d] National Physical Laboratory, Hampton Road, Teddington, Middlesex, TW11 0LW, United Kingdom

ARTICLE INFO

Keywords:
Land surface temperature
Ground measurements
Emissivity
Radiometer calibration

ABSTRACT

Accurate ground measurements of land surface temperature (LST) are necessary for validating satellite LST products. In order to provide reliable data, ground radiometers must be calibrated with reference to an international standard, and radiometric temperatures must be corrected for land surface emissivity. As opposed to water, land surface emissivity is not usually known for many ground covers, so an emissivity value has to be assumed, assigned from spectral emissivity libraries or measured for each land cover and spectral band considered. The aim of this study is to show the laboratory calibration and the methodology for simultaneous field measurements of LST and emissivity employed in the comparison experiment held in the project Fiducial Reference Measurements for validation of Surface Temperature from Satellites (FRM4STS) funded by the European Space Agency. We used multiband CE-312 radiometers (five narrow bands plus one broad band in the 8–13 μm window) to simultaneously retrieve LST and band emissivities by means of the temperature-emissivity separation (TES) method for different ground covers (clover, tarmac, soil, gravel and sand). The TES method requires near-simultaneous measurements of ground-leaving radiances and sky downwelling radiances; the latter being measured in the field using a gold reflectance panel. For each surface cover, TES provided band emissivities in the CE-312 bands and LST from continuous radiance measurements performed over time. As a result of the experiment, we present the laboratory calibration of the CE-312 radiometers carried out against a traceable, reference blackbody, and the LST series and band emissivity values for the ground covers considered, together with a detailed LST uncertainty analysis including the uncertainties associated to the calibration of ground radiometers, the emissivity estimation by means of the TES method, and the sky radiance measurements, among others. According to these results, the total LST uncertainty was estimated as 0.4-0.6 K for the ground covers measured during the experiment.

1. Introduction

In recent years, a number of land surface temperature (LST) datasets derived from thermal infrared (TIR) satellite measurements have been provided to the user for a variety of applications in many environmental sciences (Li et al., 2013). It is therefore essential to carry out an independent validation of the satellite-derived LSTs and a reliable estimation of the LST uncertainty with reference to traceable, international standards. Recently, international initiatives have arisen with the aim of defining the best practices for ground measurements and validation of satellite LST products at global scale (Guillevic et al., 2017; Martin and Göttsche, 2016). Direct comparison of satellite LST products with ground measurements performed with TIR radiometers in thermally homogenous test sites is generally considered as the reference validation method (Guillevic et al., 2017). With this aim, several research groups have established validation sites in different environments and surface types where in situ LSTs are measured with TIR radiometers (Hook et al., 2007; Göttsche et al., 2016; Tang et al., 2015; Niclòs et al., 2018). Apart of the representativeness of in situ measurements with regard to satellite data due to the dissimilarity in the spatial scale (from typically less than 1 m^2 to 1 km^2 or more), two other problems must be addressed in order to provide accurate ground LST

* Corresponding author.
E-mail addresses: cesar.coll@uv.es (C. Coll), raquel.niclos@uv.es (R. Niclòs), jesus.puchades@uv.es (J. Puchades), vicente.garcia-santos@uv.es (V. García-Santos), joanmiquel.galve@uclm.es (J.M. Galve), Lluis.Perez@uv.es (L. Pérez-Planells), enric.valor@uv.es (E. Valor), theo.theocharous@npl.co.uk (E. Theocharous).

https://doi.org/10.1016/j.jag.2019.02.002
Received 10 July 2018; Received in revised form 6 February 2019; Accepted 11 February 2019
Available online 23 February 2019
1569-8432/ © 2019 Elsevier B.V. All rights reserved.

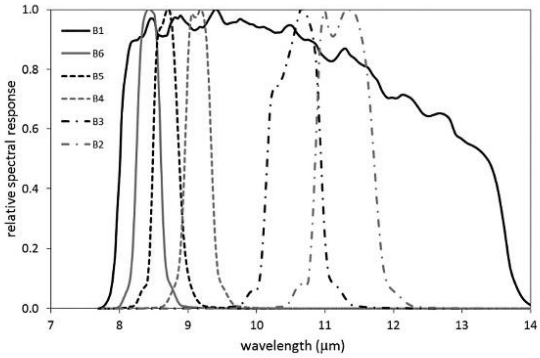

Fig. 1. Relative spectral response of the six CE-312 bands as a function of wavelength.

measurements: (i) the calibration of the field radiometers with reference to SI (*Système International*) standards, and (ii) the measurement of the land surface emissivity.

In 2009, the Committee on Earth Observation Satellites (CEOS) organised a comparison of TIR radiometric instrumentation (radiometers and blackbodies) used to support calibration and validation of satellite-borne sensors with emphasis on sea surface temperature (Theocharous et al., 2010a, b). Stage 1 of this comparison included laboratory measurements at the National Physical Laboratory (NPL), which provided the reference transfer radiometer AMBER (Theocharous et al., 1998) to calibrate the participant's blackbodies, and the NPL ammonia heat-pipe reference blackbody (Chu and Machin, 1999) to calibrate the participant's radiometers. In Stage 2 of the 2009 CEOS comparison, similar experiments were conducted at the Rosenstiel School of Marine and Atmospheric Science (RSMAS), University of Miami, with the National Institute of Standards and Technology (NIST) providing traceability to SI units (Theocharous and Fox, 2010b). As a follow up, analogous comparisons were held up at NPL in 2016, namely the Fiducial Reference Measurements for validation of Surface Temperature from Satellites (FRM4STS, http://www.frm4sts.org), funded by the European Space Agency (ESA), which also focused on land and ice surface temperature measurement. The Thermal Remote Sensing Group of the University of Valencia (TRSG-UV) participated in both the 2009 and 2016 experiments using TIR multiband radiometers CIMEL CE-312 (Sicard et al., 1999; Legrand et al., 2000), which have several bands in the 8–13 μm range, and the Landcal P80 P blackbody source (www.landinst.nl/wordpress/calibration/p80p).

The objective of the present paper is to show the methodology followed by TRSG-UV for the laboratory calibration of radiometers and blackbodies, and the simultaneous measurement of LST and emissivity in field conditions, including a detailed uncertainty analysis. Data from the AMBER radiometer and the NPL blackbody measured in the FRM4STS laboratory comparison (Theocharous et al., 2017a; Barker-Snook et al., 2017) were used in this paper. We also show and discuss the field LST and emissivity measurements over five different land covers considered in the FRM4STS field LST comparison (Theocharous et al., 2017b). As opposed to water, land surface emissivity is usually unknown for many ground covers, so it has to be either assumed, assigned from spectral emissivity libraries or, ideally, measured for each land cover and spectral band considered.

One possibility is the use of the box method for field emissivity measurement in either the two-lid or one-lid version (Rubio et al., 2003); however it is time-consuming and has to be implemented independently of the LST measurements in a separate experiment. Another possibility is the application of the Temperature-Emissivity Separation (TES) method (Gillespie et al., 1998) to multiband CE-312 radiometer measurements for simultaneously retrieving LST and emissivity in the radiometer bands. The method is based on an empirical relationship between the apparent band emissivity contrast and the minimum band emissivity, which breaks down the well-known under determination problem in TIR measurements. The TES method, originally proposed for the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) on board the Terra satellite, can also be applied to field multiband measurements (Mira et al., 2009) and was selected by our group for the FRM4STS field LST comparison.

The paper follows with the description of the field radiometer and the results of the laboratory calibration, including the FRM4STS experiment and the characterisation of the radiometer uncertainty. Section 3 describes the TES method and its application with the field radiometer in the FRM4STS LST comparison. In Section 4, the LST and emissivity results are shown and discussed and the total LST uncertainty is estimated. Finally, Section 5 presents the main conclusions of the study.

2. Instrument description and laboratory calibration

2.1. CIMEL CE-312 TIR multiband radiometer

The CIMEL-Electronique CE-312 radiometer (Sicard et al., 1999; Legrand et al., 2000) has a field of view of 10° and six spectral bands within 8 and 13 μm spectral range (Fig. 1): one wide band B1 (8.0–13.3 μm) and five narrow bands B2 (10.9–11.7 μm), B3 (10.2–11.0 μm), B4 (9.0–9.3 μm), B5 (8.5–8.9 μm) and B6 (8.3–8.6 μm) similar to those of the space borne ASTER sensor. The broad band is composed of a Germanium window with zinc sulphide filters and the narrow bands are interference filters. Measurements can be taken for a single selected band or for all bands consecutively (B1 to B6).

The detector is a thermopile at ambient temperature. The instrument has a built-in radiance reference made of a concealable gold-coated mirror occupying the same field of view as the target measurements which enables comparison between the target radiance and the reference radiation from inside the detector cavity. The temperature of the detector is measured with a calibrated platinum resistance thermometer (PRT), thus allowing compensation for the cavity radiation. The relevant outputs of the radiometer are the detector temperature and the difference in digital counts between the signals from the target and the detector cavity.

In order to measure the radiance coming from a target in a given

band, two consecutive radiance measurements are performed. First, with the concealable gold mirror closed, the detector measures the radiance at the detector's cavity (L_i(cav); unknown) plus the radiance coming from the detector itself which is reflected off the gold mirror, so we can write for band i

$$L_i(\text{closed}) = L_i(\text{cav}) + B_i(T_d) \quad (1)$$

where B_i is the Planck's radiance function integrated for band i, and T_d is the detector temperature measured with the PRT. Then, the gold mirror is open so the detector measures the target's radiance (L_i) plus the detector's cavity radiance, that is

$$L_i(\text{open}) = L_i + L_i(\text{cav}) \quad (2)$$

The unknown value of L_i(cav) can be thus eliminated by subtracting the open and closed radiance measurements so that the radiance coming for the target can be expressed as

$$L_i = B_i(T_d) + \Delta L_i \quad (3)$$

where ΔL_i is the difference between the open and closed radiance measurements. The radiometer radiance outputs are given in terms of digital numbers (DN_i) in the range 0–60000, and ΔL_i is obtained from the difference ΔDN_i between the digital numbers for the open and closed measurements as

$$\Delta L_i = \Delta DN_i / S_i \quad (4)$$

where S_i are the band sensitivity coefficients provided by the manufacturer based on laboratory calibration measurements in 2008. This calibration procedure will be hereafter referred to as standard calibration (sc). According to the manufacturer, the linear calibration (Eq. (4)) holds for target temperatures ranging between -40 and 50 °C, with the detector being at laboratory temperature.

On the other hand, the Planck's radiance function for band i of the CE-312 radiometer can be accurately expressed as

$$B_i(T) = \frac{a_i}{\exp\left(\frac{b_i}{T^{n_i}}\right) - d_i} \quad (5)$$

where coefficients a_i, b_i, n_i and d_i were provided by the manufacturer. When L_i is obtained from the radiometer outputs (T_d, ΔDN_i) following the above procedure, the equivalent radiometric temperature for band i (T_i) can be calculated from L_i by inversion of Planck's radiance function (that is, taking $B_i(T_i) = L_i$ and inverting Eq. (5)).

2.2. Laboratory calibration

Due to the decrease of the radiometer detector's sensitivity with time, the standard calibration must be corrected for the calibration drift using laboratory measurements against a reference blackbody source for a range of temperatures of interest. The calibration drift particularly has an effect on the calculation of ΔL_i (Eq. (4)), which in turns is a function on the temperature or radiance difference between the target and the detector. Since the relationship between ΔL_i and ΔDN_i is linear, we can assume a linear recalibration equation as

$$L_i(\text{cal}) = L_i(\text{sc}) + \alpha_i[L_i(\text{sc}) - B_i(T_d)] + \beta_i \quad (6)$$

where L_i(cal) is the re-calibrated radiance (that is, the radiance corresponding to the reference blackbody temperature), L_i(sc) is the radiance measured with CE-312 using the standard calibration, and α_i and β_i are band-dependent coefficients obtained from linear regression over a set of calibration radiance measurements. Alternatively, Eq. (6) can be written in terms of temperature if the difference between the target and detector temperature is not excessive (as shown below in this section and Section 2.3):

$$T_i(\text{cal}) = T_i(\text{sc}) + A_i[T_i(\text{sc}) - T_d] + B_i \quad (7)$$

where $T_i(\text{cal}) = T_{BB}$ is the reference blackbody temperature, T_i(sc) is the brightness temperature with standard calibration, and A_i and B_i are band-dependent coefficients obtained from linear regression over a set of calibration temperatures. Note that coefficients β_i and B_i in Eqs. (6) and (7), respectively, should be close to zero since there is no independent term in Eq. (4).

Two CE-312 radiometer units (hereafter referred to as CE1 and CE2) were calibrated in the laboratory in May 2016 (before the FRM4STS comparison at NPL in July 2016), for which we used a Landcal P80 P blackbody source (www.landinst.nl/wordpress/calibration/p80p). The blackbody cavity is made of aluminum with black, high temperature refractory coating and has dimensions of 50 mm (diameter) × 155 mm (length), with a 120° cone at the closed end. The nominal temperature range is from -10 to 80 °C, and the emissivity of the cavity is > 0.995 according to the manufacturer. The instrument is equipped with two thermometers: one internal PRT connected to digital display with resolution of 0.01 K and one external PRT with uncertainty of 0.1 K traceable to National Standards (calibration certificate). Previously to the calibration measurements, we checked that both PRTs yielded the same temperatures (within their uncertainty limits) for all the temperature range.

Since the cavity emissivity is not exactly equal to one, the blackbody temperatures measured by the PRTs must be corrected for emissivity effects (including the reflection of ambient radiation) in order to be compared with radiometric temperature measurements. Thus, the radiometric temperature of the blackbody, T_{BB}, is given by

$$B_i(T_{BB}) = \varepsilon B_i(T'_{BB}) + (1-\varepsilon)L_i(\text{amb}) \quad (8)$$

where ε is the cavity emissivity, T'_{BB} is the PRT blackbody temperature, and L_i(amb) is the ambient radiance, which can be approximated as $L_i(\text{amb}) = B_i(T_a)$ with T_a being the ambient (laboratory) temperature when a radiometer operating at ambient temperature (as CE-312) is placed in front of the blackbody cavity to take measurements. The assumption that the detector and the ambient temperature are close is reasonable when the radiometer is kept in the laboratory at least one hour before the measurements start. In such conditions, differences between detector and ambient temperatures never exceed ± 2 K during the experiment. Since ε is high (> 0.995), the differences between T_{BB} and T'_{BB} are small, so Eq. (8) can be approximated with sufficient accuracy regardless of the radiometer band as

$$T_{BB} = T'_{BB} - (1-\varepsilon)(T'_{BB}-T_a) \approx T'_{BB} - 0.005(T'_{BB}-T_a) \quad (9)$$

where in practice we take the lower emissivity bound (0.995) thus yielding the upper limit (in absolute value) for the emissivity correction. Note that, in this case, the difference T_{BB}-T'_{BB} can reach -0.1 K (+ 0.1 K) for T'_{BB}-T_a = + 20 K (-20 K).

The P80 P blackbody was calibrated in the CEOS 2009 comparison against the NPL AMBER radiometer, showing agreement within ± 0.15 K in the 20–30 °C temperature range (Theocharous and Fox, 2010b). The emissivity correction of Eq. (9) was not applied at that time. In the FRM4STS 2016 experiment, the emissivity correction was applied and, for temperatures from 0 to 50 °C, the P80 P blackbody minus AMBER temperatures ranged from -0.04 to 0.08 K, with a mean bias of 0.02 K, standard deviation of 0.04 K and root mean squared difference (rmsd) of 0.05 K (Theocharous et al., 2017a). The AMBER radiometer is characterized with an uncertainty of 53 mK (Theocharous et al., 2017a), so the agreement is good within the combined uncertainty limits. If the emissivity correction was not applied, the mean bias, standard deviation and rmsd would be 0.07, 0.05 and 0.09 K respectively. Temperatures below 0 °C are problematic because ice formation in the blackbody cavity and they were not considered in this experiment. The good calibration results in both the CEOS 2009 comparison and the FRM4STS 2016 experiment show the long-term stability and the accuracy of the P80 P blackbody.

For the laboratory calibration of the CE-312 radiometers, the P80 P

blackbody source was set at six fixed temperatures (0 to 50 °C) in two different runs with instrument re-aligning. Enough time (about 1 h) was allowed for the blackbody to reach equilibrium at each temperature. The blackbody temperature was corrected for emissivity (Eq. (9)) taking T_a as the laboratory temperature. Radiometers were aligned with the black body cavity, and placed at a distance so that the field of view was smaller than the cavity diameter. Taking into account the radiometer nominal field of view (10°) and the length of the cavity (155 mm), we estimated that placing the radiometer at 20 mm from the cavity entrance yields a diameter of 30 mm on the cavity closed end, which is considerably smaller than the cavity diameter (50 mm). The standard calibration was applied to the radiometer readouts to calculate the equivalent brightness temperature. For each temperature and band, we made 15 radiance measurements for which the average value and standard deviation were calculated. For a given band, the difference between the average values for the same temperature in the two runs (with re-alignment) gives an estimate of the reproducibility of measurements, whereas the standard deviation of 15 measurements for a given temperature (without re-alignment) gives the repeatability. The detector temperature ranged between 24 and 28 °C during the experiment. Fig. 2 shows the calibration results for the six bands of the CE2 unit (similar results were obtained for CE1) together with the linear regression according to Eq. (7). Table 1 gives the calibration coefficients A_i and B_i of Eq. (7), the coefficient of determination and the standard error of the regression for each band and radiometer. These results show that the linear calibration equation is accurate for all bands in the temperature range considered and it is necessary to apply it to the standard calibration due to the detector's sensitivity drift, especially when the target temperature is not close to the detector temperature. Moreover, B_i values are close to zero in all cases as anticipated.

Table 1
Calibration coefficients A_i and B_i of Eq. (7) derived from the laboratory measurements in Valencia, May 2016, using the P80 P blackbody. The coefficient of determination (r^2) and the standard regression error (σ) are also given. (a) CE1. (b) CE2.

(a) CE1 band	A_i	B_i (K)	r^2	σ (K)
1	0.01613	−0.01	0.964	0.06
2	0.00981	−0.06	0.931	0.05
3	0.00496	−0.08	0.810	0.04
4	0.00875	−0.13	0.930	0.04
5	0.01512	−0.08	0.971	0.05
6	0.00747	−0.16	0.898	0.05
(b) CE2 band	A_i	B_i (K)	r^2	σ (K)
1	0.01674	0.02	0.986	0.04
2	0.00955	0.00	0.968	0.03
3	0.00645	−0.03	0.927	0.03
4	0.01157	−0.07	0.934	0.06
5	0.01511	−0.01	0.988	0.03
6	0.01115	−0.07	0.947	0.05

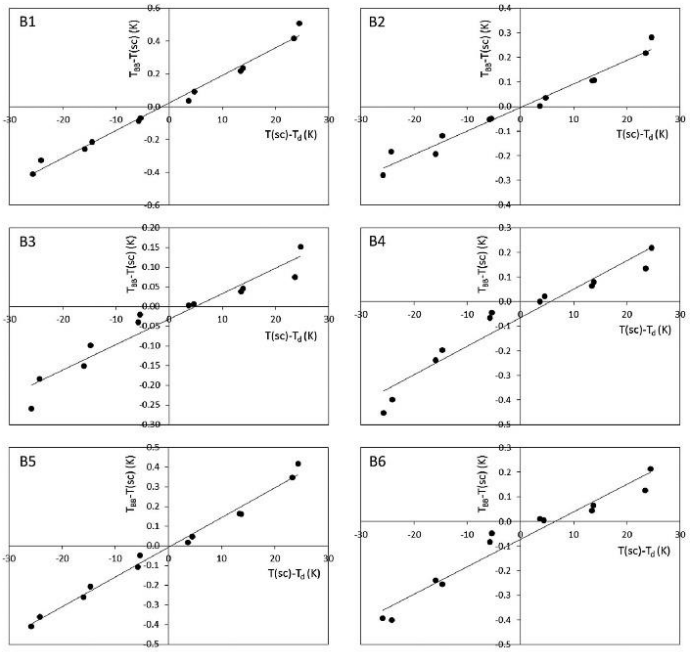

Fig. 2. Laboratory calibration measurements performed in May 2016 for the six bands of CE2 using the P80 P blackbody. The straight lines show the linear regression according to Eq. (7). For each band, differences between the blackbody temperatures, T_{BB}, and radiometric temperatures with standard calibration, T(sc), are plotted against differences between the latter and the detector temperatures, T_d.

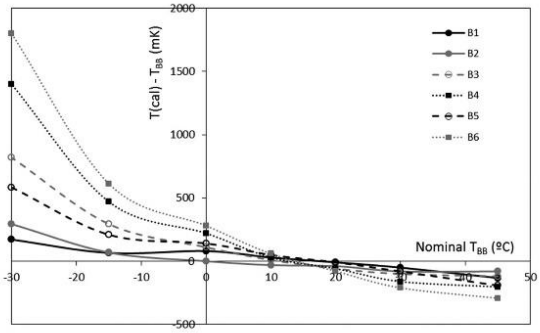

Fig. 3. Differences between the CE-312 (CE1) temperature measurements calibrated with Eq. (7) for the six bands and the NPL blackbody temperature against the nominal NPL blackbody temperature.

2.3. FRM4STS calibration

The two CE-312 radiometers were calibrated against the NPL ammonia heat-pipe reference blackbody in the FRM4STS experiment in June 2016, so the previous laboratory calibration can be compared with a reference standard. The NPL blackbody is characterized with an uncertainty of 0.04 K (Barker-Snook et al., 2017) and it was set at nominal temperatures of -30, -15, 0, 10, 20, 30 and 45 °C (the exact values were unknown at the time of measurements). Note that the temperature range covered temperatures well below 0 °C, unlike the P80 P blackbody that cannot operate at such low temperatures due to ice formation in the cavity. For each blackbody temperature, we performed 12 consecutive measurements for each band of each CE-312 radiometer, for which the average and standard deviation (repeatability) was calculated and compared with the blackbody temperatures. The detector temperature of both radiometers ranged between 21 and 24 °C during the experiment.

Fig. 3 shows the results of the FRM4STS comparison for the six bands of the CE1 instrument (similar results were obtained for CE2). It is apparent that there were important differences for bands 3–6 at the lower temperature values, which were out of the calibration range in the previous experiment with the P80 P blackbody. For such bands, we can also observe that the difference between the CE-312 and the blackbody temperatures does not hold a linear relationship with temperature in the full temperature range considered, as opposite to the linear standard calibration (Eq. (4)) and the laboratory re-calibration (Eq. (7)). However, the agreement between the measured and blackbody temperatures is good and approximately linear in the 0–45 °C range, with linear correlation coefficients ranging between 0.938 (B3) and 0.998 (B1 and B5) for CE1, and between 0.931 (B4) and 0.998 (B2 and B5) for CE2.

In Table 2, we show the basic statistics of the differences T_i(cal)-T_{BB} for the six bands of the two radiometers in the 0–45 °C interval, which is the most useful for LST measurements. According to these results, the rmsd was between 0.06 and 0.10 K for bands B1-B3 and between 0.13 and 0.23 K for bands B4-B6 of both instruments. The best (worst) accuracy was obtained for band B2 (B6). If the laboratory re-calibration was not applied (standard calibration only), the rmsd would yield 0.2 K for bands B2 and B3, and 0.3-0.4 K for bands B1 and B4-B6.

In Fig. 3 we can observe that the differences between measured and blackbody temperatures are close to zero around 20 °C (close laboratory temperature), where differences change sign. Besides, the largest differences (in absolute value) between measured and blackbody temperatures occur for the largest differences (in absolute value) between measured and detector temperatures. This suggests that the above FRM4STS calibration data can be used to derive new coefficients A_i and B_i for the re-calibration of the CE-312 radiometers according to Eq. (7). In order to keep the linearity of the re-calibration, and because there were no data enough below 0 °C we only considered the 0–45 °C temperature range to obtain the calibration coefficients from linear regression. This is shown in Fig. 4 where, as examples, we plot the differences between the blackbody temperatures and the measured temperatures with standard calibration against the differences between that temperature and the detector's temperature for band 1 (the most accurate) and band 6 (the less accurate) of CE1 considering the full temperature range of the experiment (-30 to 45 °C). Fig. 4 also shows the linear regression for the 0–45 °C range only, which were used to derive the new coefficients A_i and B_i.

Table 2
Basic statistics of the differences between the CE-312 measurements calibrated with Eq. (7) for the six bands and the NPL reference blackbody temperatures in the 0–45 °C range from the FRM4STS laboratory comparison. (a) CE1. (b) CE2.

(a) Radiometer CE1	B1	B2	B3	B4	B5	B6
minimum difference (K)	−0.13	−0.08	−0.12	−0.20	−0.19	−0.29
maximum difference (K)	0.08	0.00	0.11	0.22	0.14	0.28
bias (K)	−0.02	−0.05	−0.03	−0.03	−0.02	−0.05
standard deviation (K)	0.08	0.03	0.09	0.17	0.13	0.23
rmsd (K)	0.08	0.06	0.10	0.17	0.13	0.23
(b) Radiometer CE2	B1	B2	B3	B4	B5	B6
minimum difference (K)	−0.15	−0.10	−0.13	−0.17	−0.22	−0.22
maximum difference (K)	0.08	0.04	0.14	0.26	0.19	0.35
bias (K)	−0.02	−0.03	−0.01	−0.01	−0.01	0.01
standard deviation (K)	0.09	0.05	0.10	0.17	0.16	0.22
rmsd (K)	0.09	0.06	0.10	0.17	0.16	0.22

Table 3 shows the calibration coefficients to be applied to the standard calibration derived for the six bands of the two radiometers, together with the coefficient of determination and the standard regression error. Results show high correlation and small uncertainty in the linear calibration for all bands in the considered temperature range. Note that the calibration coefficients were derived in laboratory conditions with quite stable ambient (and detector) temperatures (21–24 °C). However, the calibration of the CE-312 radiometers is based on the radiance difference between the target and the detector (Eq. (4)) and then, in principle, it depends only on the difference between T_d and the target temperature and not on T_d itself, provided that the sensitivity of the detectors is independent of temperature. This assumption could

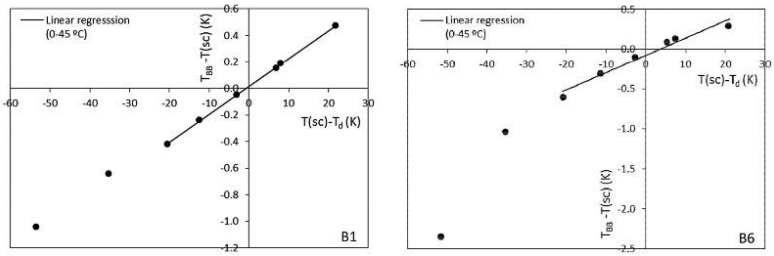

Fig. 4. FRM4STS calibration measurements performed at NPL, June 2016 for bands 1 (left) and 6 (right) of CE1 using the NPL reference blackbody. T_{BB} is the blackbody temperature, T(sc) is the radiometric temperature with standard calibration, and T_d is the detector temperature. The straight lines show the linear regression for the 0–45 °C interval according to Eq. (7).

Table 3
Calibration coefficients A_i and B_i of Eq. (7) derived from the FRM4STS measurements at NPL, June 2016 using the NPL reference blackbody for the 0–45 °C interval. The coefficient of determination (r^2) and the standard regression error (σ) are also given. (a) CE1. (b) CE2.

(a) CE1 band	A_i	B_i (K)	r^2	σ (K)
1	0.02103	0.02	0.9997	0.01
2	0.01195	−0.01	0.9955	0.01
3	0.01043	−0.03	0.9515	0.04
4	0.01854	−0.07	0.9634	0.06
5	0.02288	−0.04	0.9978	0.02
6	0.02177	−0.08	0.9645	0.07
(b) CE2 band	A_i	B_i (K)	r^2	σ (K)
1	0.02200	0.04	0.9995	0.01
2	0.01275	0.03	0.9994	0.01
3	0.01253	0.00	0.9759	0.03
4	0.02129	−0.04	0.9649	0.07
5	0.02455	0.02	0.9990	0.01
6	0.02411	−0.05	0.9596	0.08

only be checked by performing calibration experiments for different values of T_d covering a wide range of ambient temperatures in field conditions, for what a climatic chamber with controlled and variable ambient temperature would be required. Unfortunately, this was not possible in the FRM4STS laboratory comparison. With this caveat in mind, we assumed that the re-calibration coefficients of Table 3 can be applicable regardless the detector's temperature if the difference between the detector and target temperatures is within ± 25 °C as in Fig. 4. Out of this range (e.g., sub-zero temperatures corresponding to measurement of atmospheric radiances), there could be large discrepancies for most of the CE-312 bands.

2.4. Uncertainty budget

As a summary of this section, we present a detailed uncertainty analysis of the temperature measurements of the CE-312 radiometers, which was carried out following the guidelines of the Joint Committee for Guides in Metrology (JCGM, 2008). According to it, we distinguished two types of uncertainty contribution: Type A, referring to the evaluation of a component of measurement uncertainty by a statistical analysis of measured quantity values obtained under defined measurement conditions (e.g., random), and Type B, that is an evaluation of a component of measurement uncertainty determined from non-statistical calculations, including uncertainties taken from calibration certificates or derived from the characterisation of the performance of the instruments being used (e.g., systematic). Results shown here correspond to the laboratory calibration experiment performed in May 2016 prior to the FRM4STS experiment.

For Type A, we considered the repeatability and reproducibility derived from multiple measurements as mentioned in Section 2.2. For Type B, several error sources were considered:

- The uncertainty of the primary calibration, which was taken as the total uncertainty of the P80 P blackbody (0.05 K) according to the FRM4STS blackbody calibration results (Theocharous et al., 2017b).
- The linearity of radiometer, which was taken as 0.06 K as the typical value for all bands in the temperature range 0–40 °C (Legrand et al., 2000).
- Correction of calibration drift: It was corrected for by means of linear re-calibration equations derived from P80 P blackbody measurements for each band and radiometer (Section 2.2). The uncertainty for this correction is the root sum square (RSS) of the typical estimation uncertainty of the linear regression (0.05 K for CE1 and 0.04 K for CE2) and the uncertainties resulting from the propagation of input temperature errors (standard deviations for 15 measurements at a fixed temperature) in the linear correcting equation. The resulting uncertainty in the correction of calibration drift is 0.07 K for CE1 and 0.05 K for CE2.
- Ambient temperature fluctuations: This effect is compensated by the CE312 radiometers by measuring the detector cavity temperature by means of a calibrated PRT as explained in Section 2.1. The uncertainty in this process is the uncertainty of the internal PRT, which is 0.04 K according to Legrand et al. (2000).
- Finally, the effect of atmospheric absorption and emission was estimated as negligible (< 0.01 K) due to very short path length and radiometers working in the atmospheric window.

Table 4 summarizes the different uncertainty estimates for both

Table 4
Uncertainty analysis of the two CE-312 radiometers considering the different error sources (see text). RSS is the root sum square, and the total RSS represents the standard uncertainty of the radiometer, with coverage factor k = 1.

Uncertainty Contribution	Radiometer CE1		Radiometer CE2	
	Type A (K)	Type B (K)	Type A (K)	Type B (K)
Repeatability of measurement	0.05		0.03	
Reproducibility of measurement	0.08		0.06	
Primary calibration		0.05		0.05
Linearity of radiometer		0.06		0.06
Correction of calibration drift		0.07		0.05
Ambient temperature fluctuations		0.04		0.04
RSS	0.09	0.11	0.07	0.10
Total RSS	**0.15**		**0.12**	

radiometers, as well as the RSS for Types A and B, and the total RSS considered as the standard uncertainty of the radiometers (that is, with coverage factor k = 1). The uncertainty estimations shown here are compatible with most of the differences between the measured and the NPL blackbody temperatures shown in section 2.3 for the FRM4STS comparison in the 0–45 °C range. These results show the accuracy and precision of the CE-312 radiometers, which is suitable for ground measurements for validating satellite LSTs.

3. Simultaneous measurement of LST and emissivity in field conditions

In order to measure LST in field conditions, an estimate of the surface emissivity is necessary. This is not always an easy task, since land surface emissivity can be rather variable depending on the cover type, the spectral band used, the soil moisture, and the observation angle, among others. The best option is usually to measure emissivity at the same field conditions as LST. To this end we propose the use of the Temperature-Emissivity Separation (TES) method (Gillespie et al., 1998) to simultaneously retrieve LST and emissivity using the five narrow bands (B2-B6) of the CE-312 radiometer.

3.1. The TES method

The TES method starts from the calibrated radiance measured at surface level in band i of the ground radiometer, which can be written as

$$L_i^{surf} = \varepsilon_i B_i(T) + (1-\varepsilon_i)L_i^{sky} \quad (10)$$

where T is the LST, ε_i is the surface emissivity in band i and L_i^{sky} is the sky downwelling radiance in band i, which needs to be measured in the field near-simultaneously to the target radiance. If Eq. (10) is applied to the N spectral bands of a radiometer, we have N + 1 unknowns (one LST and N band emissivities ε_i), so the problem is underdetermined. In the TES method, the under-determination is solved with an empirical relationship between the apparent spectral contrast in emissivity, or maximum-minimum emissivity difference (MMD) and the minimum spectral emissivity (Gillespie et al., 1998). The application of the TES method involves the following steps:

1. Normalized emissivity method (NEM): We assume an emissivity value $\varepsilon_{NEM} = 0.98$ in Eq. (10) for all bands and calculate T_{NEMi} for each band; that is

$$B_i(T_{NEMi}) = \frac{L_i^{surf} - (1-\varepsilon_{NEM})L_i^{sky}}{\varepsilon_{NEM}} \quad (11)$$

2. Select the maximum value of T_{NEMi} ($T_{max} = \max(T_{NEMi})$, i = 1,.., N) and use T_{max} to obtain an estimate of the N NEM emissivities according to

$$\varepsilon_{NEMi} = \frac{L_i^{surf} - L_i^{sky}}{B_i(T_{max}) - L_i^{sky}} \quad (12)$$

3. Calculate the β_i spectrum from the NEM emissivities as

$$\beta_i = \frac{\varepsilon_{NEMi}}{\bar{\varepsilon}} \quad (13)$$

where $\bar{\varepsilon}$ is the average NEM emissivity for the five bands. From the β_i spectrum, the MMD is defined as the difference between the maximum and minimum β_i values (MMD = max(β_i)-min(β_i)).

4. Calculate the minimum band emissivity (ε_{min}) from the MMD using the empirical relationship

$$\varepsilon_{min} = 0.9951 - 0.7264 \times MMD^{0.7873} \quad (14)$$

which was derived by Hulley and Hook (2009) from laboratory emissivity spectra taken from the ASTER spectral library (ASL, Baldridge et al., 2009). They selected 49 soil spectra from ASL in terms of taxonomy (14 aridisols, 9 alfisols, 10 entisols, 7 inceptisols and 9 mollisols) plus nine additional dune sand samples, which were linearly combined with the conifer vegetation emissivity spectrum (with reflectance reduced by a factor of two to account for canopy scattering effects) by varying the fractional vegetation cover from 0 to 100% in steps of 5% (Hulley and Hook, 2009). The resulting spectra were convolved to the spectral response functions of the ASTER bands and then the ε_{min}-MMD relationship was obtained by regression. The numerical coefficients of Eq. (14) differ from the coefficients of the original ε_{min}-MMD relationship derived by Gillespie et al. (1998) and, according to Hulley and Hook (2009), they improve the estimations for vegetated surfaces characterized by low values of MMD and high values of ε_{min}.

Although Eq. (14) was obtained originally for the ASTER bands, it can be applied to the similar CE-312 narrow bands as well (Mira et al., 2009). We checked this assumption using the same emissivity spectra from ASL (except for the nine sand dune spectra which were not available) and the linear mixture of soil and vegetation spectra mentioned above to calculate the emissivities in the CE-312 and ASTER bands. We found that the CE-312 emissivities were very close to the ASTER emissivities in the equivalent bands, with rmsd of 0.0003 for CE-312 band B2 (equivalent to ASTER band 14), 0.0010 for B3 (band 13), 0.0020 for B4 (band 12), 0.0010 for B5 (band 11) and 0.0083 for B6 (band 10). Differences were larger for band B6 because its center (8.45 μm) is slightly shifted with respect to ASTER band 10 (center at 8.29 μm). Fig. 5 shows the ε_{min}-MMD relationship for both CE-312 and ASTER using the calculated emissivities, together with the Hulley and Hook (2009) regression curve (Eq. (14)). We can see that there is a good agreement between the CE-312 and ASTER for the whole range of the data, with rmsd of 0.004 for both MMD and ε_{min}. Equation (14) also agrees well with ε_{min} values for both CE-312 and ASTER (rmsd of 0.005 for CE-312 and 0.007 for ASTER for the full data range), although the agreement is better for MMD values lower than 0.15. For larger MMD values, Eq. (14) slightly overestimates ε_{min} by about 0.01 for both CE-312 and ASTER. This is because Eq. (14) was optimized for a wider range of samples, with higher (lower) MMD (ε_{min}) values corresponding to the sand dune spectra used by Hulley and Hook (2009) but not considered in the present study. Therefore, we consider that the ASTER ε_{min}-MMD empirical relationship can be also applied to CE-312 data, as it provides an accurate estimation of ε_{min} (rmsd of 0.005) within the typical estimation error of the ε_{min}-MMD relationship (about 0.01 according to Gillespie et al., 1998).

5. The final emissivity values are obtained by scaling the β_i spectrum with ε_{min} according to

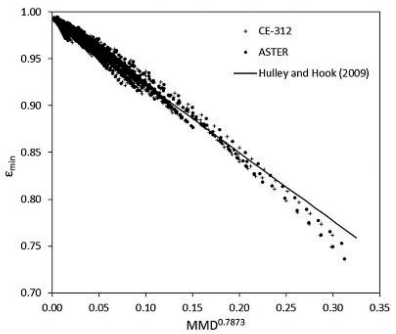

Fig. 5. Empirical relationship between ε_{min} and MMD for CE-312 (crosses) and ASTER (dots) using the emissivities calculated form ASL spectra with linear combinations of conifer vegetation and 49 soil samples. The continuous line shows the ε_{min}-MMD curve from Hulley and Hook (2009).

$$\varepsilon_i = \varepsilon_{min}[\beta_i/\min(\beta_i)] \quad (15)$$

and they are used to re-calculate the surface temperature for each band as in Eq. (11). Ideally, such temperatures should be band-independent; however, they may differ slightly (< 0.1 K) in practice. Finally, the LST is taken as the maximum band temperature.

3.2. Application in field conditions

In the FRM4STS LST comparison, the CE2 radiometer was used for continuous multiband measurement of the radiance from the surface (L_i^{surf}) because its accuracy was slightly better than CE1. We planned to use CE1 for continuous measurements of sky downwelling radiance (L_i^{sky}) using a diffuse reflectance gold panel (García-Santos et al., 2012) simultaneously to the surface measurements of CE2. However, CE1 showed electronical and power problems during the campaign and did not provide useful data for most of the time. Therefore, sky radiance measurements were performed with CE2 by placing the diffuse reflectance gold panel in the field of view of the radiometer at regular, short intervals during the measurement of surface radiances. The measured sky radiances were then linearly interpolated to the time of the surface radiance measurements. During the FRM4STS field comparison, environmental conditions were not ideal for LST measurements, with partial and rather variable cloud cover. Therefore L_i^{sky} varied considerably and the interpolation procedure was subject to significant error. Besides, the CE-312 measurements were calibrated in the 0–50 °C range (Section 2.2, Table 1), so L_i^{sky} measurements that correspond to low temperatures may be affected by severe errors depending on the band (Section 2.3). Fortunately, the impact of L_i^{sky} uncertainties on LST is small compared to other error sources, as discussed in Section 4.3.

The multiband CE-312 measurements are not really simultaneous but successive, from B1 to B6 with a gap of 6 s between two consecutive bands and a total of 39 s for a six-band measurement series including filter rearrangement. This means that there is a time gap of 24 s between measurements in bands B2 and B6 in a series and a gap of 39 s for a given band between two consecutive series. In the TES method, it is required that LST is the same for all bands what may be a problem if surface temperature changes significantly during the multiband measurement. This was the case in FRM4STS where variable cloud cover caused rapid changes in temperature, so the sample was not isothermal during the 24-s multiband measurement. In order to mitigate this problem, we applied a correction method based on linear temporal interpolation between consecutive multiband series. To do that, we selected the measurement time of band B3 as the reference time for each series so B3 measurements remain unchanged and the measurements for all other bands are reduced to that time by means of linear interpolation between the two closest measurements, one from the same measurement series and the other from either the preceding (bands B1 and B2) or the next series (bands B4-B6). While this procedure may not be required in ideal field conditions where surface temperatures may remain constant during the multiband measurement, the particular conditions during the FRM4STS experiment made the correction necessary.

4. Results and discussion

The field LST measurements were performed at the NPL campus in July 2016 (Theocharous et al., 2017b). Five different samples were considered (see Fig. 6): clover, tarmac, soil (dark, with high content of organic matter), gravel and sand (yellowish, typical construction sand). The last three samples were placed in open containers with dimensions of 1.5 m × 1.1 m and the surface was tilled flat and uniform. The CE2 radiometer was placed at a height of 2.5 m from the sample and looking at 25° from nadir, thus covering a quasi-elliptic area with semi axes of 23 and 27 cm over the surface. Measurements were taken off-nadir to prevent reflections from the base of the radiometer mount but keeping the viewing angle as small as possible to avoid directional effects on LST and make sure that the field of view of the radiometers fitted within the target.

For each sample, continuous multiband measurements were performed during different time intervals, ranging from 40 min for clover to more than 2 h for gravel. That sample was measured in two different days (July 6 and 7). As mentioned before, the CE2 measurements were calibrated with coefficients of Table 1, the temporal interpolation was applied to the band brightness temperatures and the TES method was used. We present here the results obtained in terms of LST and band emissivities, including a detailed uncertainty analysis.

4.1. LST and emissivity

In the TES method, LST is obtained for the band where surface emissivity is maximum, which was B2 (10.9–11.7 μm) for the five samples considered. Fig. 7 shows the LST values obtained as a function of time for each sample. We can observe rapid changes in LST for some cases, which are indicative of variable cloud cover. In the cases of sand, soil and gravel there were variations of about 10 K in few minutes, with soil, gravel and tarmac LSTs reaching about 50 °C around noon. The dynamic range of LST was about 10 K for sand and gravel, while it reached 20 K for soil. LSTs were more constant for clover and tarmac, with variations of about 5 K during the measurement period in both cases. It should be also noted that the difference between the sample and detector's temperature reached 20–25 °C for soil, tarmac and gravel thus being close to the upper limit of the calibration range but still within it (e.g., Fig. 2). A good calibration of the instrument is important for such cases, otherwise biases up to -0.5 K could be obtained. In the case of sand, the difference between sample and detector's temperature reached 15 °C only because measurements were made early in the morning. For clover, both temperatures were quite similar, with differences ranging from -4 to 1 °C.

For each LST retrieval, TES provides an estimation of the B2-B6 band emissivities. However, emissivity must be constant for a given sample during the measurement time, so we calculated the mean value and the standard deviation of the TES derived emissivities for each band during the measurement period corresponding to each sample. The mean values were taken as the band emissivities and the uncertainty in the retrieved emissivity was calculated as the RSS of the standard deviation of band emissivity, the standard estimation error of the empirical ε_{min}-MMD relationship (Eq. (14)), which is 0.010 (Gillespie et al., 1998), and the rmsd of the ASTER ε_{min}-MMD empirical relationship (Eq. (14)) with regard to CE-312 (0.005; Section 3.1). We also checked the influence of the initial ε_{NEM} assumption (taken as 0.98 in the present study) on the TES method. To do this, we varied ε_{NEM} by ± 0.01 and compared the resulting TES emissivities and LSTs with those obtained with original ε_{NEM} value. Results showed that the impact of ε_{NEM} was very small for all targets considered. The highest (lowest) impact was found for clover (sand), with variations of 9×10^{-5} (1.3×10^{-5}) in maximum emissivity (band B2), 0.003 (0.003) in minimum emissivity (band B6) and rmsd of 0.07 K (0.001 K) in LST. Such variations can be considered as negligible when compared to the other uncertainty sources considered above.

Fig. 8 shows the emissivity spectra obtained for the five samples, and Table 5 gives the emissivity for band B2 corresponding to the maximum emissivity value and for band B6 usually corresponding to the minimum emissivity.

The CE-312 wideband B1 (8.0–13.3 μm) is not used in the TES method, however the radiances measured in such band can be used to retrieve emissivity in band B1 according to

$$\varepsilon_1 = \frac{L_1^{surf} - L_1^{sky}}{B_1(T) - L_1^{sky}} \quad (16)$$

Fig. 6. Pictures of the five samples considered in the FRM4STS field LST measurements.

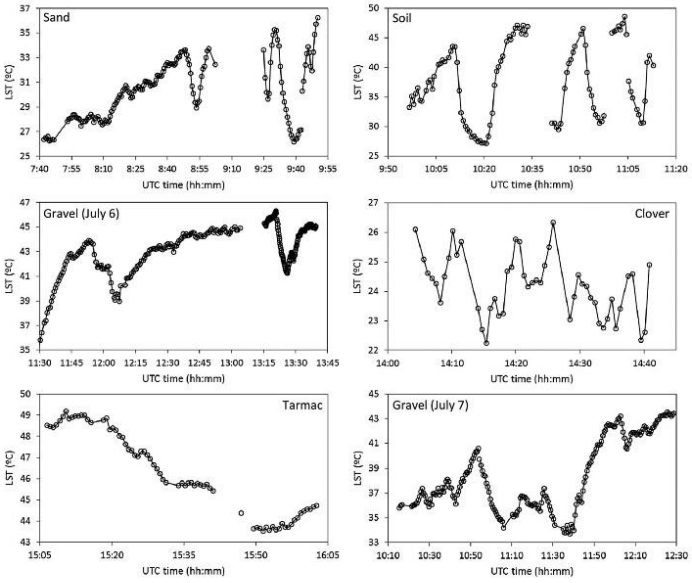

Fig. 7. LST values as a function of time for the samples measured during the FRM4STS field experiment. Gaps in the LST series correspond to power and/or electronic problems with the radiometer.

where T is the LST retrieved from TES. Table 5 also gives the emissivity in band B1 for the five samples.

4.2. Discussion

Results show that emissivity in long wavelength bands B2 and B3 was high and with small variability among the different covers considered ranging roughly from 0.95 to 0.98, which is an optimal case for radiometric LST measurements. Differently, short wavelength bands B4-B6 showed much larger variability and lower values for sand (≈ 0.80) and gravel (≈ 0.85) which show the characteristic features due to quartz absorption at 8–9 μm (*reststrahlen* bands; Rowan and Mars, 2002). This also translates into low emissivity values for broad band B1 in such cases. Besides, we can note that gravel emissivity estimates for two different days were in very good agreement within the uncertainty limits. For clover and soil with high content in organic matter, we

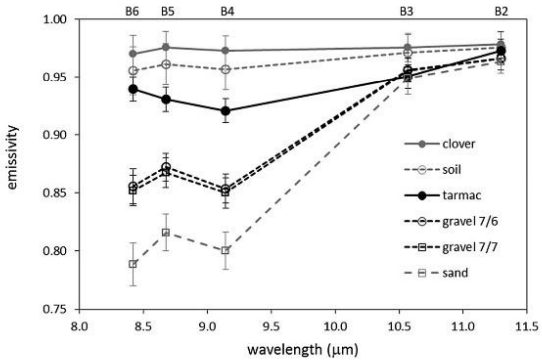

Fig. 8. Emissivity of the samples measured in FRM4STS as a function of the centre wavelength of the narrow bands of CE-312. Results for the gravel sample measured in two different days are presented.

obtained flat, near grey body spectra with relatively high values as expected (Salisbury and D'Aria, 1992).

Unfortunately, independent emissivity measurements were not performed for the samples considered in the FRM4STS field LST comparison, so we cannot validate the TES derived emissivities and estimate their uncertainty with regard to an absolute reference. To this end, emissivity spectra obtained from field Fourier Transform Infrared (FTIR) spectrometer data taken in the same conditions as the LST measurements would be required. In the present FRM4STS comparison, participants were encouraged to make their own best estimate of the target emissivities for their radiometer under the conditions which prevailed during the LST measurements, but such measurements cannot be compared with the TES emissivity estimates because the different spectral bands of the radiometers.

As pointed out above, emissivities in band B2 (B6) were the highest (lowest) for each cover and also showed the lowest (highest) variability along the multiple TES measurements performed for a given target and thus the lowest (highest) uncertainty as shown in Table 5. This implies that B2 (B6) is the most (less) appropriate band for LST retrieval. We can show this calculating the LST for a given band from the at-surface and sky radiances according to

$$B_i(T) = \frac{L_i^{surf} - (1-\varepsilon_i)L_i^{sky}}{\varepsilon_i} \quad (17)$$

where ε_i is the average band emissivity in band i. For a given target, such LSTs would differ from the LSTs derived from TES due to the variability of TES-retrieved emissivities during the measurement period. We compared the TES LSTs with those derived from Eq. (17) in bands B2, B6 and B1 using the average emissivities of Table 5. Fig. 9 shows the LST differences between TES and Eq. (17) (ΔT_i) for two targets (sand and gravel on July, 6), and Table 6 gives the rmsd with respect to TES LSTs for the three bands and all targets considered in the study. We can see that, as expected, B2 yielded the smallest rmsd compared to the other bands for all targets, with values of 0.01-0.02 K except for the soil and clover targets, which were the most affected by variable atmospheric conditions prevailing during the campaign and thus yielded more variable TES emissivity estimations. In any case, B2 showed very good stability during the measurement periods, B6 being less stable in time with respect to TES LSTs owing to more variable emissivity values obtained from TES. B1 showed intermediate results between B2 and B6.

Finally, we show the effect of the temporal interpolation method described in Section 3.2 taking as an example the sand emissivity results for band B6 (8.3–8.6 μm). Fig. 10 shows the emissivity values obtained for each multiband series during the sand measurement period, with comparison to the case when temporal interpolation was not applied. We can see that the dispersion in emissivity is larger for the latter case, especially at times around 8:55 and from 9:25, when LST showed rapid variations due to variable cloud cover (see Fig. 7 for sand). When applying the temporal interpolation method, the retrieved emissivity values were more uniform for the whole measurement period, yielding a mean value of 0.787 with standard deviation of 0.008 (0.788 and 0.014 respectively without temporal interpolation).

We can also observe in Fig. 10 an increase of emissivity with time (or LST) of the sand sample in the first part of the measurement period. We do not believe that the emissivity variation is due to the LST variation since the retrieved emissivities remained constant for gravel and tarmac while LST varied considerably during their measurement periods (see Fig. 7). The observed increase of sand emissivity could be due to an increase of the real sky downwelling radiance which was not detected by our L_i^{sky} measurements due to rapid changes in cloud cover and insufficient sampling of L_i^{sky} during the measurement period. In

Table 5
Emissivity values and uncertainties for the five FRM4STS samples in narrow bands B2 and B6 and in wideband B1.

Target	B2 (10.9-11.7 μm)		B6 (8.3-8.6 μm)		B1 (8.0-13.3 μm)	
	Emissivity	Uncertainty	Emissivity	Uncertainty	Emissivity	Uncertainty
Clover	0.978	0.012	0.970	0.017	0.973	0.014
Soil	0.976	0.014	0.955	0.021	0.970	0.015
Tarmac	0.973	0.011	0.939	0.012	0.951	0.010
Gravel	0.966	0.011	0.854	0.015	0.925	0.010
Sand	0.963	0.011	0.789	0.019	0.900	0.011

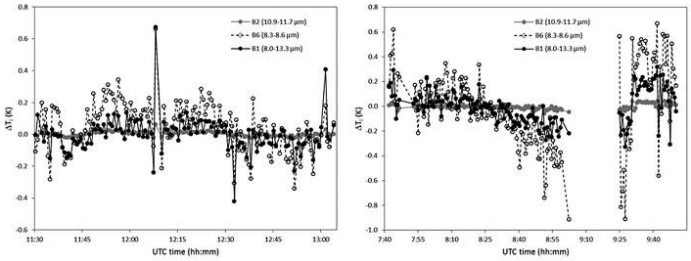

Fig. 9. LST difference between TES and Eq. (17) (ΔT_i) for bands B2, B6 and B1 as a function of time for gravel, July, 6 (left) and sand (right).

Table 6
Root mean squared differences (K) between LSTs derived from TES and Eq. (17) (ΔT_i) for bands B2, B6 and B1 and the targets considered in the study.

Target	B2 (10.9-11.7 μm)	B6 (8.3-8.6 μm)	B1 (8.0-13.3 μm)
clover	0.19	0.34	0.28
soil	0.36	0.66	0.44
tarmac	0.02	0.14	0.07
gravel (July, 6)	0.01	0.16	0.11
gravel (July, 7)	0.02	0.30	0.11
sand	0.02	0.32	0.14

fact we used mostly constant values of L_i^{sky} for the sand sample while there was variability in cloud cover during the measurements. As pointed out before, atmospheric conditions were not ideal during most of the campaign so it is difficult to assess the reason for the emissivity increase shown in Fig. 10. One option would be to repeat the experiment in ideal atmospheric conditions (i. e., with completely clear skies) but with variable LST to check if the emissivity remains constant with time. However, it is out of our possibilities in the present study and it is considered as a future work.

4.3. LST uncertainty analysis

An uncertainty analysis considering all the error sources affecting the LST retrieval was conducted following the same guidelines as in Section 2.4. Part of the LST uncertainty comes from the radiometer itself, therefore the error sources from Table 4 were taken into account again. In addition, the following error sources were considered.

- Target emissivity and sky downwelling radiance: We used the emissivity uncertainty values for B2 (the band where LST is obtained) given in Table 5. As mentioned before, the sky radiance was rather variable during the experiment, showing sky brightness temperatures in B2 ranging from -7 to -27 °C, which is roughly equivalent to a variation of ± 15% in sky downwelling radiance. This uncertainty bound also includes radiometer calibration errors at low temperatures, and the effect of self-emission of the gold panel used for the measurement of L_i^{sky}, which is lower than 5% according to Garcia-Santos et al. (2013). The impact of both emissivity and sky radiance uncertainties on LST was assessed through Eq. (10).

- Angle of view from nadir: An uncertainty of 2.5° in zenith angle was considered for the analysis, which results in emissivity uncertainty of 0.0003 at observation angle of 25° from nadir (García-Santos et al., 2012).

- Atmospheric absorption and emission: Atmospheric transmittances and upwelling radiances in the path length between the sample and the radiometer (height of 2.5 m and observation angle of 25° from nadir) were simulated with the MODTRAN 5 radiative model (Berk et al., 2006) using standard summer atmospheric conditions.

Results for each error source are shown in Table 7 for the soil sample, which yielded the largest total LST uncertainty (0.64 K). Clearly, the total LST error was practically determined by the uncertainty in the target emissivity and sky radiance that was the largest error source (0.63 K for soil). The magnitude of this effect depends on the sample, the minimum value being 0.46 K for clover (total LST uncertainty of 0.47 K) due to lower emissivity uncertainty and LST values in this case. Moreover, the impact of emissivity uncertainty is much larger than that of the sky radiance due to the high emissivity values in band B2 for all samples considered. Taking into account the uncertainty in the sky radiance alone (taken as 15% here), the resulting LST

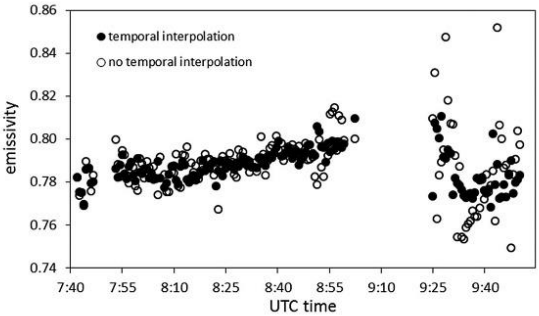

Fig. 10. Comparison of emissivity values retrieved in narrow band B6 for the sand sample when the temporal interpolation was applied and when it was not applied.

Table 7
Uncertainty analysis of LST measured by the CE2 radiometer for the soil sample considering the different error sources (see text). RSS is the root sum square, and the total RSS represents the standard uncertainty of LST, with coverage factor k = 1.

Uncertainty Contribution	Type A (K)	Type B (K)
Repeatability of measurement	0.03	
Reproducibility of measurement	0.06	
Primary calibration		0.05
Linearity of radiometer		0.06
Correction of calibration drift		0.05
Ambient temperature fluctuations		0.04
Target emissivity and sky radiance		0.63
Angle of view from nadir		0.02
Atmospheric absorption and emission		0.02
RSS	0.07	0.56
Total RSS	**0.64**	

uncertainty was only between 0.1 and 0.2 K depending on the sample, whereas the impact of emissivity error alone ranged between 0.4 and 0.6 K.

5. Conclusions

The validation of satellite LST products requires accurate and precise ground measurements performed in thermally homogeneous land covers. In this paper we addressed the two main error sources in ground LST measurements, which are radiometer calibration and uncertainty in target emissivity. Ground radiometers must be calibrated with reference to traceable SI standards in the temperature range of interest. The CEOS 2009 and ESA's FRM4STS 2016 experiments provided the framework for such calibration experiments against NPL reference radiometers and blackbodies, as well as an opportunity for harmonising methods and protocols for international groups involved in ground LST measurements. On the other hand, emissivity is not usually known in advance for land surface targets; therefore it should be measured in the field for the spectral band used by the radiometer for each cover type considered. A detailed uncertainty analysis including all error sources in the LST measurement should be provided together with the ground LST.

The detector of the CE-312 radiometer operates at ambient temperature, but the effect of the radiance from the detector's cavity on the target radiance is corrected by measuring the detector's temperature with a PRT and taking two radiance measurements, one with the radiometer aperture closed with a gold-coated mirror and the other with the aperture open to the target. However, due to the decrease of the detector's sensitivity with time, it is necessary to re-calibrate the CE-312 measurements; otherwise there could be errors up to ± 0.5 K for target versus detector temperature differences of ± 20 K. Linear re-calibration equations (e.g., Eq. 7) were used to this end, with coefficients obtained from laboratory measurements against the P80 P blackbody source in the temperature range 0–50 °C. The P80 P blackbody was in turn calibrated against the NPL reference AMBER radiometer in the 2009 and 2016 comparisons, showing long-term stability and good accuracy (0.05 K) in the aforementioned range when the correction for non-unity in blackbody emissivity (Eq. 9) was applied. Using the re-calibration equations derived with the P80 P blackbody shortly before the FRM4STS 2016 comparison, the two CE-312 radiometers showed rmsd ranging between 0.1 and 0.2 K (depending on the band) with respect to the NPL reference blackbody in the 0–45 °C range, which is the most useful for LST measurements. If only the standard calibration was applied, rmsd values would be between 0.2 and 0.4 K. It is therefore recommended to perform calibration experiments yearly to correct for calibration drifts. Outside of the above mentioned temperature range, the CE-312 radiometers could show larger biases (in excess of 1 K for certain bands at -30 °C) and non-linearity effects, especially for sub-zero temperatures corresponding to sky radiance measurements. Fortunately, the impact of sky radiance errors on the retrieved LST is rather small for samples with high emissivity.

Thanks to the multiband capability of the CE-312 radiometers, the TES method can be applied to ground measurements to simultaneously retrieve LST and band emissivities in the field. This was done in the FRM4STS LST comparison for five different samples. Although conditions were not ideal for field LST measurements due to changing cloud cover, the derived emissivity values were reasonable according to the nature of the covers considered (Salisbury and D'Aria, 1992; Rowan and Mars, 2002). However, it was not possible to validate the TES emissivities against independent emissivity measurements for the FRM4STS samples in the same conditions. Results showed that long wavelength TIR bands (i.e., 10–12 μm) yielded the highest emissivity values with small variation among samples, thus being the most suitable for LST retrieval. The uncertainty of retrieved emissivities in band B2 (10.9–11.7) ranged from 0.011 to 0.014 for the five samples, being larger for bands B4-B6 located in the 8.3–9.3 μm range (up to 0.02 in band B6, 8.3–8.6 μm) were emissivity varied considerably among samples (roughly from 0.80 for sand to 0.97 for clover). A detailed LST uncertainty analysis was performed, showing that emissivity was the largest source of error and total LST uncertainties ranging between 0.5 K (clover) and 0.6 K (soil). This stresses the need for accurate emissivity estimations for the different ground covers used for field LST measurements.

Acknowledgments

Financial support was obtained from research projects CGL2015-64268-R and CGL2013-46862-C2-1-P (Ministerio de Economía y Competitividad), the "Juan de la Cierva" (FJCI-2015-24876) Research Contract of Dr. Galve and the "Torres Quevedo" program through the Spanish Government under Contract PTQ-16-08578 of Dr. Garcia-Santos. We thank the anonymous reviewers, who helped to improve the manuscript.

References

Baldridge, A.M., Hook, S.J., Grove, C.I., Rivera, G., 2009. The ASTER spectral library version 2.0. Remote Sens. Environ. 113, 711–715.
Barker-Snook, I., Theocharous, E., Fox, N.P., 2017. 2016 Comparison of IR Brightness Temperature Measurements in Support of Satellite Validation. Part 2: Laboratory Comparison of Radiation Thermometers. NPL REPORT ENV 14. National Physical Laboratory, Teddington, UK.
Berk, A., Anderson, G.P., Acharya, P.K., Bernstein, L.S., Muratov, L., Lee, J., Fox, M., Adler-Golden, S.M., Chetwynd, J.H., Hoke, M.L., Lockwood, R.B., Gardner, J.A., Cooley, T.W., Borel, C.C., Lewis, P.E., Shettle, E.P., 2006. MODTRAN5: 2006 update. Proceedings of SPIE, vol. 6233.
Chu, B., Machin, G., 1999. A low-temperature blackbody reference source to -40 °C. Meas. Sci. Technol. 10, 1–6.
Garcia-Santos, V., Valor, E., Caselles, V., Mira, M., Galve, J.M., Coll, C., 2013. Evaluation of different methods to retrieve the hemispherical downwelling irradiance in the thermal infrared region for field measurements. Ieee Trans. Geosci. Remote. Sens. 51 (4), 2155–2165 Apr 2013.
García-Santos, V., Valor, E., Caselles, V., Burgos, M.Á., Coll, C., 2012. On the angular variation of thermal infrared emissivity of inorganic soils. J. Geophys. Res. 117, D19116. https://doi.org/10.1029/2012JD017931. 2012.
Gillespie, A.R., Matsunaga, T., Rokugawa, S., Hook, S.J., 1998. Temperature and emissivity separation from advanced spaceborne thermal emission and reflection radiometer (ASTER) images. Ieee Trans. Geosci. Remote. Sens. 36, 1113–1125.
Göttsche, F.-M., Olesen, F.S., Trigo, I.F., Bork-Unkelbach, A., Martin, M.A., 2016. Long term validation of land surface temperature retrieved from MSG/SEVIRI with continuous in-situ measurements in Africa. Remote Sens. (Basel) 8 (5), 1–27 n° 410.
Guillevic, P., Göttsche, F., Nickeson, J., Hulley, G., Ghent, D., Yu, Y., Trigo, I., Hook, S., Sobrino, J.A., Remedios, J., Román, Mi, Camacho, F., 2017. Land Surface Temperature Product Validation Best Practice Protocol, Version 1.0 - October, 2017. Committee on Earth Observation Satellites Working Group on Calibration and Validation, Land Product Validation Subgroup.
Hook, S.J., Vaughan, R.G., Tonooka, H., Schladow, S.G., 2007. Absolute radiometric in-flight validation of mid infrared and thermal infrared data from ASTER and MODIS on the Terra spacecraft using the Lake Tahoe, CA/NV, USA, automated validation site. IEEE Trans. Geosci. Remote Sens. 45, 1798–1807.
Hulley, G., Hook, S., 2009. The North American ASTER land surface emissivity database (NAALSED) version 2.0. Remote Sens. Environ. 113, 1967–1975.

JCGM, 2008. Evaluation of measurement data – guide to the expression of uncertainty in measurement, Joint Committee for Guides in Metrology. Report JCGM 100 2008.
Legrand, M., Pietras, C., Brogniez, G., Haeffelin, M., Abuhassan, N.K., Sicard, M., 2000. A high-accuracy multiwavelength radiometer for in situ measurements in the thermal infrared. Part I: characterization of the instrument. J. Atmos. Ocean Techn. 17, 1203–1214.
Li, Z.-L., Tang, B.-H., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I.F., Sobrino, J.A., 2013. Satellite-derived land surface temperature: current status and perspectives. Remote Sens. Environ. 131, 14–37.
Martin, M., Göttsche, F., 2016. Satellite LST Validation Report. Available online at. DUE Globtemperature Project. http://www.globtemperature.info/index.php/public-documentation/deliverables-1/117-validation-report-del12.
Mira, M., Schmugge, T.J., Valor, E., Caselles, V., Coll, C., 2009. Comparison of thermal emissivities retrieved with the two-lid box and TES methods with laboratory spectra. Ieee Trans. Geosci. Remote. Sens. 47 (4), 1012–1021.
Niclòs, R., Pérez-Planells, L., Coll, C., Valiente, J.A., Valor, E., 2018. Evaluation of the S-NPP VIIRS land surface temperature product using ground data acquired by an autonomous system at a rice paddy. Isprs J. Photogramm. Remote. Sens. 135, 1–12 2018.
Rowan, L.C., Mars, J.C., 2002. Lithologic mapping in the Mountain Pass, California area using advanced spaceborne thermal emission and reflection radiometer (ASTER) data. Remote Sens. Environ. 84 (3), 350–366.
Rubio, E., Caselles, V., Coll, C., Valor, E., Sospedra, F., 2003. Thermal-infrared emissivities of natural surfaces: improvements on the experimental set-up and new measurements. Int. J. Remote Sens. 24 (24), 5379–5390.
Salisbury, J.W., D'Aria, D.M., 1992. Emissivity of terrestrial materials in the 8-14 μm atmospheric window. Remote Sens. Environ. 42, 83–106.
Sicard, M., Spyak, P.R., Brogniez, G., Legrand, M., Abuhassan, N.K., Pietras, C., Buis, J.P., 1999. Thermal infrared field radiometer for vicarious cross-calibration: characterization and comparisons with other field instruments. Opt. Eng. 38 (2), 345–356.
Tang, B.-H., Shao, K., Li, Z.-L., Wu, H., Nerry, F., Zhou, G., 2015. Estimation and validation of land surface temperatures from Chinese second-generation polar orbitFY-3A VIRR data. Rem. Sens. 7, 3250–3273. https://doi.org/10.3390/rs70303250.
Theocharous, E., Fox, N.P., 2010. CEOS Comparison of IR Brightness Temperature Measurements in Support of Satellite Validation. Part II: Laboratory Comparison of the Brightness Temperature of Blackbodies. NPL REPORT OP4. National Physical Laboratory, Teddington, UK.
Theocharous, E., Fox, N.P., Sapritsky, V.I., Mekhontsev, S.N., Morozova, S.P., 1998. Absolute measurements of black-body emitted radiance. Metrologia 35, 549–554.
Theocharous, E., Usadi, E., Fox, N.P., 2010. CEOS Comparison of IR Brightness Temperature Measurements in Support of Satellite Validation. Part I: Laboratory and Ocean Surface Temperature Comparison of Radiation Thermometers. NPL REPORT OP3. National Physical Laboratory, Teddington, UK.
Theocharous, E., Barker Snook, I., Fox, N.P., 2017a. 2016 Comparison of IR Brightness Temperature Measurements in Support of Satellite Validation. Part 1: Blackbody Laboratory Comparison. NPL REPORT ENV 12. National Physical Laboratory, Teddington, UK.
Theocharous, E., Barker Snook, I., Fox, N.P., 2017b. 2016 Comparison of IR Brightness Temperature Measurements in Support of Satellite Validation. Part 4: Land Surface Temperature Comparison of Radiation Thermometers. NPL REPORT ENV 13. National Physical Laboratory, Teddington, UK.

ANEXO B

Validation of Sentinel-3 SLSTR Land Surface Temperature Retrieved by the Operational Product and Comparison with Explicitly Emissivity-Dependent Algorithms

remote sensing

MDPI

Article

Validation of Sentinel-3 SLSTR Land Surface Temperature Retrieved by the Operational Product and Comparison with Explicitly Emissivity-Dependent Algorithms

Lluís Pérez-Planells [1,2,*], Raquel Niclòs [1], Jesús Puchades [1], César Coll [1], Frank-M. Göttsche [2], José A. Valiente [3], Enric Valor [1] and Joan M. Galve [4]

[1] Department of Earth Physics and Thermodynamics, Faculty of Physics, University of Valencia, 50, Dr. Moliner, E-46100 Burjassot, Spain; raquel.niclos@uv.es (R.N.); jesus.puchades@uv.es (J.P.); cesar.coll@uv.es (C.C.); enric.valor@uv.es (E.V.)
[2] IMK-ASF, Karlsruhe Institute of Technology (KIT), Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany; frank.goettsche@kit.edu
[3] Fundación Centro de Estudios Ambientales del Mediterráneo (CEAM), 14 Charles R. Darwin, E-46980 Paterna, Spain; valiente_jospar@gva.es
[4] GIS and Remote Sensing Group, Institute for Regional Development, University of Castilla-La Mancha, Campus Universitario SN, E-02071 Albacete, Spain; joanmiquel.galve@uclm.es
* Correspondence: lluis.perez@uv.es; Tel.: +34-963-543-249

Abstract: Land surface temperature (LST) is an essential climate variable (ECV) for monitoring the Earth climate system. To ensure accurate retrieval from satellite data, it is important to validate satellite derived LSTs and ensure that they are within the required accuracy and precision thresholds. An emissivity-dependent split-window algorithm with viewing angle dependence and two dual-angle algorithms are proposed for the Sentinel-3 SLSTR sensor. Furthermore, these algorithms are validated together with the Sentinel-3 SLSTR operational LST product as well as several emissivity-dependent split-window algorithms with in-situ data from a rice paddy site. The LST retrieval algorithms were validated over three different land covers: flooded soil, bare soil, and full vegetation cover. Ground measurements were performed with a wide band thermal infrared radiometer at a permanent station. The coefficients of the proposed split-window algorithm were estimated using the Cloudless Land Atmosphere Radiosounding (CLAR) database: for the three surface types an overall systematic uncertainty (median) of −0.4 K and a precision (robust standard deviation) 1.1 K were obtained. For the Sentinel-3A SLSTR operational LST product, a systematic uncertainty of 1.3 K and a precision of 1.3 K were obtained. A first evaluation of the Sentinel-3B SLSTR operational LST product was also performed: systematic uncertainty was 1.5 K and precision 1.2 K. The results obtained over the three land covers found at the rice paddy site show that the emissivity-dependent split-window algorithms, i.e., the ones proposed here as well as previously proposed algorithms without angular dependence, provide more accurate and precise LSTs than the current version of the operational SLSTR product.

Keywords: emissivity; LST; SLSTR; split-window algorithm; in-situ validation

check for updates

Citation: Pérez-Planells, L.; Niclòs, R.; Puchades, J.; Coll, C.; Göttsche, F.-M; Valiente, J.A.; Valor, E.; Galve, J.M. Validation of Sentinel-3 SLSTR Land Surface Temperature Retrieved by the Operational Product and Comparison with Explicitly Emissivity-Dependent Algorithms. *Remote Sens.* **2021**, *13*, 2228. https://doi.org/10.3390/rs13112228

Academic Editor: Itamar Lensky

Received: 28 April 2021
Accepted: 2 June 2021
Published: 7 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Land surface temperature (LST)—like near-surface air temperature—is a key variable in a wide variety of studies, since it is linked to land–atmosphere energy transfer and flux balances [1,2]. Thus, it is required for monitoring evapotranspiration and climate change [3,4], as well as for providing estimates of fire size and temperature [5,6], volcanoes and lava flow [7,8], and vegetation health [9–11]. According to the Global Climate Observing System [12], the World Meteorological Organization (WMO) considers LST as one of the essential climate variables (ECVs). The Climate Change Initiative (CCI) was launched by the European Space Agency (ESA) for improving the prediction of climate change

trends by means of satellite data [13]. The CCI considers LST an important variable for monitoring the Earth climate system; therefore, they included it in the list of ECVs required for understanding and predicting the evolution of climate (http://cci.esa.int/ (accessed on 1 March 2021)). Consequently, the validation of satellite derived LSTs against independent references is crucial for assessing their accuracy and precision. For LST retrieval from satellite data, the GCOS set the recommended thresholds on accuracy (bias, defined as the systematic uncertainty by the Joint Committee for Guides in Metrology [14], JCGM) and precision (standard deviation, SD, defined as the random uncertainty by the JCGM [14]) to 1 K [12].

The Sea and Land Surface Temperature Radiometer (SLSTR) on board the Sentinel-3A and 3B spacecrafts is a follow-on instrument of the Advanced Along-Track Scanning Radiometer (AATSR). The two sensors have similar characteristics, including their thermal channels at 11 and 12 μm, with double view capability, and allow us to apply split-window algorithms (SWAs) and dual-angle algorithms (DAAs). In this paper, the SWA proposed by Niclòs et al. in [15] and the DAA proposed by Coll et al. in [16] were adapted to SLSTR's thermal bands. The SWA proposed by Niclòs et al. in [15] was developed for the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) onboard Meteosat Second Generation (MSG) and depends explicitly on emissivity and view zenith angle. SLSTR has view zenith angles up to 60° [17] and, thus, angular anisotropy may have an important impact on LST retrieval, which was noticed when analyzing the angular dependence of the SWA's regression coefficients. For the SEVIRI sensor, over the rice paddy site, the SWA proposed by Niclòs et al. in [15] provided an accuracy (bias) and precision (SD) of 0.5 and 0.8 K, respectively. The capability of the AATSR sensor to apply the DAA was previously analyzed in [16] over full vegetation cover. These authors proposed and validated a SWA and a DAA, obtaining a higher standard deviation for the DAA, with accuracy (precision) of 0.0 K (1.0 K). They concluded that the DAA performed worse than the SWA, mainly due to differences between the nadir and oblique footprints [16].

The operational LST level 2 (L2) product for the SLSTR sensor is generated with a SWA whose coefficients depend on surface biome, water vapor content (WVC) in the atmosphere, and vegetation fraction cover [18,19]. Previous studies validated the Sentinel-3A SLSTR operational LST product over a variety of surfaces, but not over a rice paddy. In the ESA validation report, 11 sites were used to validate the SLSTR LST product over different land covers [20]: seven were stations of the SURFace RADiance (SURFRAD) network, which uses pyrgeometers (3–50 μm), three were stations of the Karlsruhe Institute Technology (KIT) equipped with narrow band radiometers (9.6–11.5 μm), and one was the U.S. Department of Energy's Atmospheric Radiation Measurement (ARM) station equipped with narrow band radiometers.

In this paper, phenological changes of a rice paddy during the growing period were used to validate the SLSTR LST product over three different surfaces: bare soil (wet and dry), water (flooded surfaces), and full vegetation cover. A permanent station with a wide band Thermal Infrared (TIR) radiometer continuously recorded ground measurements, which were then compared with concurrent satellite LST values.

The main objective of this paper is to validate the results of the proposed SWAs and the operational SLSTR LST product. Additionally, three explicitly emissivity-dependent SWAs proposed by Sobrino et al. [21], Zhang et al. [22], and Zheng et al. [23] (hereafter called Sobrino16, Zhang19, and Zheng 19 SWAs, respectively) were evaluated under the same conditions. The main goal of proposing an explicitly angular and emissivity-dependent SWA for SLSTR is to provide a better-performing alternative to the biome-dependent (i.e., implicitly emissivity-dependent) SWA used for generating the operational product, but also to Sobrino16, Zhang19 and Zheng 19. Building on these works, this paper presents the adaptation of an SWA with explicit angular dependence, which was previously successfully applied to SEVIRI data, to SLSTR; the validation of the adapted SWA and its comparison with other SWAs with an explicit emissivity dependence; the adaptation of a DAA to SLSTR and its validation. The validation results presented here are based on in-situ LST

obtained from wide band radiometers (8–14 μm; more similar to satellite TIR observations and more accurate than pyrgeometers), which are installed at a permanent station located in a rice paddy (i.e., the Valencia LST Validation site). Despite being limited to a single site, the phenological changes over the year allowed us to validate the LST retrieved from Sentinel-3A and Sentinel-3B over three, previously unrepresented, homogeneous land cover types.

Section 2 describes the validation site and the in-situ LST and emissivity data. The SLSTR LST operational product algorithm and the different emissivity-dependent algorithms evaluated in this study are described in Section 3. Section 4 presents the validation results for each algorithm, and a discussion is provided in Section 5. Conclusions are drawn in Section 6.

2. Study Site and Ground Data

2.1. Site

The study site is a 100-km^2 rice paddy area located near Valencia, Spain (39.274°N, −0.317°E; WGS-84). This extensive area is bordered by the city of Valencia in the north, the Mediterranean Sea in the east, and tree crops and small urban areas in the south and west. Due to rice phenology, over the year, three different homogeneous land covers alternate (Figure 1). Full vegetation covers July to mid-September; flooded surface (i.e., water) in December, January, and June; and bare soil from February to May, which is wet during February, and dry from March to May. These seasonal changes allow us to validate over three different homogeneous land covers at a single site (i.e., as if we were observing three different sites). The SLSTR L2 fraction of vegetation cover data in Figure 2 show the typical seasonal changes. The composition of the bare soil found at the rice paddy site is: 14% sand, 50% silt, and 37% clay, with 4.5 % of organic matter (further soil details are provided in [24]). Based on SLSTR Level 1 (L1) auxiliary data (See Section 2.3), over the year the atmospheric WVC at the study site varies between 0.5 and 4 cm.

This site has been extensively used for LST validation purposes [15,25–28]. Previous studies demonstrated a high thermal homogeneity for this site at different spatial resolutions [27,29–31] and concluded that it is suitable for validating satellite LST with in-situ measurements. For full vegetation cover, these studies found a standard deviation (SD) lower than 0.5 K for 33 × 33 ASTER pixels (~9 km^2) centered on the study area and for a Landsat TM5 scene (~16 km^2). In [30], the authors analyzed the variability of 11 × 11 ASTER pixels (1 km^2) centered on the study area, and obtained a SD < 0.3 K. In [27], the thermal variability of the area was studied for the three land covers present at the site with hand-held radiometer measurements along transects (~300 m long) through the station parcel on different dates: the SD values obtained were 0.5 K, 0.4 K, and 0.9 K for full vegetation, flooded soil, and bare soil, respectively.

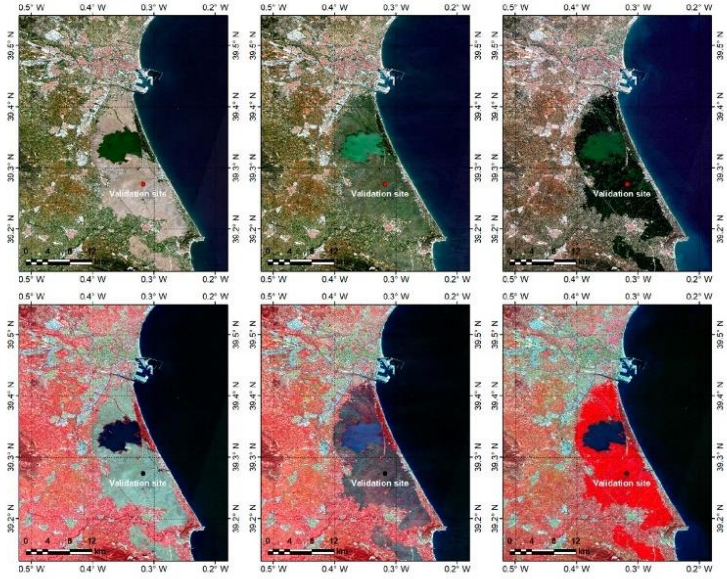

Figure 1. RGB true color compositions (R-G-B 4-3-2; **top**) and false color compositions (R-G-B 8-4-3; **bottom**) for three Sentinel-2 Multispectral Instrument (MSI) scenes. The three land covers at the site are: bare soil (April, **left**), flooded soil, i.e., water (May, **center**), full vegetation (August, **right**). The location of the validation site is shown in the composition.

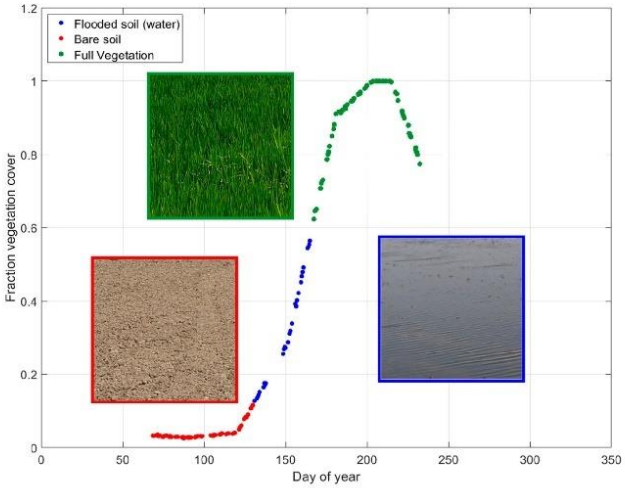

Figure 2. Fraction of vegetation cover given by the SLSTR L2 product as a function of day of year. A representative photo for each land cover is also shown.

2.2. Ground Data

2.2.1. SI-121 Radiometer

The Apogee SI-121 radiometer of the LST validation station took measurements during five periods: 5 days in 2016; July 2017 (full month); from April to August 2018; August 2019 (full month), and from November 2019 to April 2020. This instrument measures radiance in the TIR spectral region (8–14 μm) and has a field of view of 36° and an uncertainty of 0.2 K (manufacturer specification, www.apogeeinstruments.com (accessed on 1 March 2021)). The SI-121 was installed at three meter height and observed the ground at nadir view, which resulted in a footprint of ~3 m^2. A second SI-121 radiometer was set up at 53° from zenith to provide measurements representative of the downwelling hemispheric radiance [32]. Measurements were taken from both SI-121 radiometers every 4 s; the two radiometers were periodically cleaned and calibrated against a Landcal blackbody source P80P for temperatures ranging between 273 K and 313 K. The uncertainty obtained for both SI-121 radiometers was less than ±0.1 K. The manufacturer specification uncertainty (±0.2 K) was used instead the calibration uncertainty. During the Fiducial Reference Measurements for validation of surface temperature from satellites (FRM4STS) experiment in June 2016, the blackbody source was calibrated against the National Physics Laboratory (NPL) reference radiometer (AMBER), characterized with an uncertainty of 0.053 K [25]. The blackbody showed good agreement in the temperature range from 273 to 323 K with a root mean square difference (RMSD) of 0.05 K [33,34].

Only measurements of the SI-121 radiometers acquired 3 min before and after a satellite overpass were retained to have enough measurements (i.e., 90) for statistical analyses, but avoiding significant changes due to the daily trends in the LSTs within the temporal acquisition window [32]. The SD of the measurements within the 3 min was used in the estimation of the in-situ LST uncertainty. Then, the brightness temperatures (Ti) were corrected for emissivity and reflected sky radiance (atmospheric transmittance and path radiance were negligible). The sky radiance was measured with the radiometer pointing to sky, which approximates downwelling atmospheric irradiance divided by π. The emissivity values were known from previous characterizations of the site (see Section 2.2.3). The above corrections are described by Equation (1):

$$\mathrm{B}_i(T) = \frac{L_i - (1-\varepsilon_i)L^{\downarrow}_{i,\,a}}{\varepsilon_i} \quad (1)$$

where T is land surface temperature, B_i is Planck function integrated with the channel i filter function of the radiometer, L_i is the radiance measured by the sensor and estimated from Ti as Li = Bi(Ti), ε_i is surface emissivity in channel i, and $L^{\downarrow}_{i,\,a}$ is the sky radiance in this channel. After retrieving T via inverting the Planck function, the in-situ LST used for validation was estimated as the average of the T values acquired concurrently to the SLSTR overpasses. The final dataset selected for validation was obtained by removing cloudy data with the cloud mask of the SLSTR LST product.

2.2.2. CIMEL Electronique CE-312 Radiometers

Two multiband CIMEL Electronique CE-312 radiometers [35] were used to acquire daytime LST concurrently to Sentinel-3A satellite overpasses and SI-121 radiometer measurements. The CE-312 radiometer has a field of view of 10° and six channels in the 8–13 μm TIR spectral range, i.e., one wide channel and five narrow bands (channel 1: 8–13.3 μm; channel 2: 10.9–11.7 μm; channel 3: 10.2–11.0 μm; channel 4: 9.0–9.3 μm; channel 5: 8.5–8.9 μm; channel 6: 8.3–8.6 μm). During the FRM4STS calibration campaign [33], both CE-312 radiometers were calibrated against the NPL ammonia heat-pipe reference blackbody. For a range of temperatures between 273 and 318 K, a RMSD between 0.06 and 0.1 K was obtained for channels 1 to 3 and between 0.13 and 0.23 K for channels 4 to 6 [34].

The two handheld instruments were carried along ~300 m transects (150 m in opposite directions starting from the SI-121 station radiometer position) over the site on twelve

cloudless days: four days corresponded to flooded soil and eight to full vegetation. Sky radiance was measured just at the beginning and end of the transects. For flooded soil, the sky radiance was directly measured at zenith due to the specular reflectance feature of the water. Vegetated and bare soil covers were considered as near-Lambertian surfaces and sky radiance for these land covers was measured using an Infragold Reflectance Target (IRT-94-100) made by Labsphere [32], which is a highly diffuse gold panel with a reflectivity close to 0.92 in the 8–14 μm region [36].

Ground LSTs were estimated from the CE-312 radiometer measurements using Equation (1). Finally, average LSTs for the transect measurements, three minutes before and after the satellite overpasses, were calculated.

2.2.3. In-Situ Land Surface Emissivity

Surface emissivity is a key parameter for accurate LST retrievals [37–39]. For the studied land covers, in-situ emissivity values were obtained with different techniques (i.e., temperature-emissivity separation (TES) method, box method, and relative emissivity measurements) for the CE-312 radiometers.

The TES method [40] was used to obtain water and bare soil emissivity. The TES method requires at-surface radiances of the five CE-312 radiometer narrow bands (see [34] for details). These at-surface radiances were then used to obtain the relative spectral contrast. Minimum emissivity is retrieved via an empirical relationship between maximum-minimum difference (MMD) of relative emissivity and absolute minimum emissivity. Minimum absolute emissivity was used to obtain absolute emissivity of the other four channels using the temperature-independent index. LST can then be retrieved using any of the channel-specific emissivity values. The retrieved LST was used to obtain the emissivity of the CE-312's broadband channel. Bare soil TES measurements from [41] were used to obtain wet and dry bare soil emissivity values. For obtaining the wet bare soil emissivity, the emissivities of a soil sample collected at the site, with different moisture contents, with an average value of 0.41 $m^3\ m^{-3}$, were used. For dry bare soil, emissivity values of the soil sample with soil moisture of 0.03 $m^3\ m^{-3}$ were used.

In the case of vegetation, emissivity was estimated with the box method [42,43]. The box consisted of four inner aluminum walls and three different lids: two aluminum cold lids (one of them with a little hole for the radiometer measurements to be used as a top lid and the other one to be used as bottom lid), and a third non-reflecting hot lid with a temperature of around 60 °C (also used as a top lid). Moreover, the outside of the box walls and the lids were covered by a thermally insulating material. Emissivity could then be obtained by combining four radiance measurements: (1) cold top lid—sample at bottom; (2) hot top lid—sample at bottom; (3) hot top lid—cold bottom lid; (4) cold top lid—cold bottom lid [43].

The emissivities used for each land cover at the site are provided in Table 1 along with the associated uncertainty for each spectral channel of the CE-312 radiometers. The emissivities used for the SI-121 radiometer were the same as those for the broadband channel 1 (8–13.3 μm) of the CE-312.

Table 1. Emissivity values for the three CE-312 channels used.

Land Cover	8–13.3 μm	10.9–11.7 μm	10.2–11.0 μm
Flooded soil	0.986 ± 0.005	0.991 ± 0.004	0.990 ± 0.004
Wet bare soil	0.973 ± 0.012	0.977 ± 0.008	0.972 ± 0.011
Dry bare soil	0.967 ± 0.016	0.972 ± 0.004	0.970 ± 0.005
Full vegetation soil	0.983 ± 0.004	0.980 ± 0.005	0.985 ± 0.004

Since SLSTR's view zenith angle can reach up to 60°, the angular variation of emissivity was taken into account: for flooded soil, i.e., water, the emissivity relationship in [37] was used, which directly estimates the emissivity in MODIS spectral channels 31 and 32 (11 μm and 12 μm). Due to the similarity between MODIS and SLSTR spectral channels, the same

relationship could be used here. For wet and dry bare soil, the emissivity values were measured with two CE-312 radiometers under view angles from 0° to 70° in steps of 10° in order to obtain relative to nadir measurements [41]. These values were interpolated to SLSTR sensor view zenith angles. The emissivity values for flooded soil and bare soil at different view angles are shown in Figure 3.

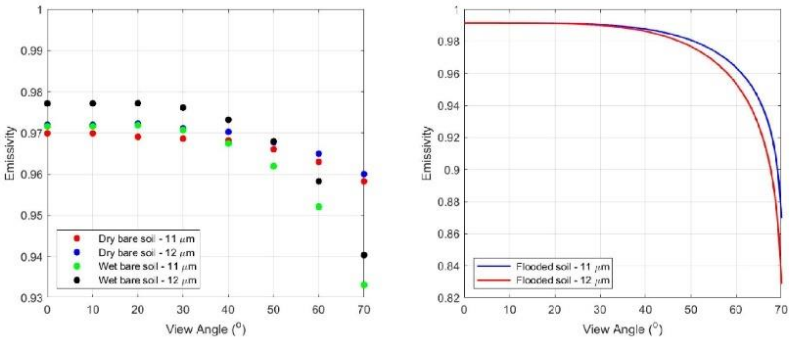

Figure 3. Angular emissivity variation of the bare soil (**left**) and flooded soil (**right**) for the CE-312 channels centered on 11 and 12 μm.

2.3. SLSTR Level 1 Data

The SLSTR onboard Sentinel-3A (launched in February 2016) and Sentinel-3B (launched in April 2018) have nine spectral channels between 0.5 and 12 μm (three visible and near infrared channels, VNIR; three short wave infrared channels, SWIR; three TIR channels). The SLSTR L1 product (baseline 003) was used in this study for the period from August 2016 to January 2020. LST was retrieved from SLSTR's TIR channels located at 11 and 12 μm (SLSTR channels 8 and 9, respectively). Brightness temperatures for these channels were provided by the SLSTR L1 product in K; auxiliary data were also provided, e.g., cloud information or pixels filled with cosmetic values (i.e., copies of the closest adjacent valid pixels). All SLSTR level 1 pixels were analyzed with different cloud tests (i.e., VNIR and SWIR thresholds tests, and TIR histogram tests). Cloudy and cosmetic pixels were filtered out from the SLSTR dataset used in this study.

WVC is an input parameter of SWAs and is included in SLSTR L1 auxiliary data (obtained from the European Centre for Medium-Range Weather Forecasts (ECWMF) analysis data). Different authors compared the WVC from ECWMF analyses with WVC obtained from Global Positioning System (GPS) data, radiosonde data, and flight measurements [44–46]. The studies showed a good performance of ECWMF WVC, although it was reported to overestimate WVC over dry areas [45] and underestimate it over humid areas [44,46]. In this paper, the ECMWF WVC provided in the SLSTR L1 data was used as input for the SWAs. In order to check its consistency over our study site, the WVC obtained from 12 SLSTR scenes concurrent with the two radiometer transects (CE312) and permanent station acquisitions (Apogee SI-121) were compared with those obtained from National Center for Environmental Prediction (NCEP) atmospheric profiles. Since NCEP atmospheric profiles are provided every 6 h on a grid of $1° \times 1°$, the four closest profiles before and after a Sentinel-3 overpass were interpolated temporally and spatially to the time of the SLSTR data acquisition and site coordinates. The comparison showed that the bias between NCEP and ECWMF WVC was 0.26 cm and the SD was 0.22 cm. The corresponding RMSD was 0.34 cm, which was lower than the uncertainty associated with

the WVC (±0.5 cm; [15,21]). The mean WVC of the ECWMF WVC for the twelve days coincident with the transects measurements was 2.4 cm and had an SD of 0.7 cm.

3. LST Retrieval Algorithms

3.1. Operational SLSTR LST Product

The operational SLSTR LST L2 product is retrieved with the SWA described by Equation (2) [11]:

$$T = a_{f,i,wvc} + b_{f,i}(T_{11} - T_{12})^{\sec(\theta/m)} + \left(b_{f,i} + c_{f,i}\right)T_{12} \quad (2)$$

where T is the LST, T_{11} and T_{12} are the brightness temperatures at 11 and 12 µm, respectively, θ is the satellite viewing angle, m is a parameter related to the view angle, and $a_{f,i,wvc}$, $b_{f,i}$, $c_{f,i}$ are algorithm coefficients, which depend on vegetation fraction (f), surface biome (i), WVC and day/night time. Algorithm coefficients are given for the 27 land cover classes of the Globcover classification scheme, which provides global classification maps with a resolution of 300 m [47]. Each coefficient is subdivided into a vegetation and a soil coefficient, which are weighted by vegetation cover fraction. However, for some biomes, these vegetation and bare soil coefficients have the same values, e.g., for irrigated cropland (biome 1), which is the biome assigned to the study area, but also for rainfed cropland (biome 2), needle leaved evergreen forest (biome 8), grassland (biome 14), sparse vegetation (biome 15), vegetation on regularly flooded or waterlogged soil (biome 18), urban areas (biome 19), bare areas (biomes from 20 to 25), water bodies (biome 26), and permanent snow and ice (biome 27). Moreover, day and night coefficients are equal for most of the biomes, except for those of water or flooded surfaces, as it is the case of forests regularly flooded (biomes 16 and 17) and biomes 1, 18, and 26. The study area consists exclusively of biome 1, which corresponds to a post-flooding or irrigated croplands land classification. While this (constant) classification of the station pixel is correct for the full vegetation period, it does not account for changes of surface type; therefore, the flooded and bare soil land covers encountered during other parts of the year are misclassified.

The SLSTR LST L2 product (baseline collection 003) was used for the period from August 2016 to 13 January 2020. From the latter date onwards, the SLSTR LST product baseline collection changed to version 004 (changes in product data format and re-gridding). Version 004 was to complete the Sentinel-3B database with bare soil covers.

The SLSTR operational product was validated in previous studies. The ESA validation report [20] showed that over most sites the accuracy threshold was achieved at daytime and nighttime. However, fewer sites met the precision threshold, especially in the case of the SURFRAD stations, likely mainly due to the heterogeneity of the surroundings [20]. In contrast to the SURFRAD stations, KIT's stations are located in specifically selected, homogenous areas, and use narrow band Heitronics KT15.85 IIP (9.6–11.5 µm) radiometers. For KIT's Evora site (Portugal, temperate evergreen vegetation) an SLSTR LST accuracy of −0.8 K and precision of 0.7 K was obtained for daytime, and an accuracy of −0.4 K and a precision of 0.3 K was found for nighttime. For KIT's Kalahari site (Namibia, Kalahari bush), an accuracy of 0.7 K (1.1 K) and a precision of 0.7 K (0.3 K) for daytime (nighttime) were obtained. For KIT's Gobabeb site (Namibia, gravel plains), an accuracy of 1.8 K (−0.9) and a precision of 0.8 K (1.1 K) for daytime (nighttime) were obtained. For the ARM station (cattle pasture), a high accuracy for both daytime (0.17 K) and nighttime (−0.02 K) was obtained. However, precision was low, with values of 1.9 K and 2.1 K for daytime and nighttime, respectively. In [48], a pyrgeometer and a thermal infrared (TIR) wide band radiometer for validating the SLSTR LST product over a forest site in the Amazon basin were used. An accuracy of −0.1 K and a precision of 0.6 K were estimated from the comparison with the wide band radiometer, while an accuracy and precision of 1.0 K were estimated from the comparison with the pyrgeometer, thereby reaching the GCOS thresholds. In [22], the SLSTR LST product over two desert sites (Dalad Banner and Wuhai, China) was validated using wide band radiometers. The accuracies obtained at these

sites were 1.0 and 1.1 K, with precisions of 1.7 and 0.9 K for Dalad Banner and Wuhai, respectively. In [49], the SLSTR product was validated against in-situ LST from two KIT sites (Namib gravel plains near Gobabeb and Lake Constance): the product achieved an accuracy (RMSD) of 1.6 K (2.4 K) and 0.4 K (0.7 K) over the Namib gravel plains and Lake Constance, respectively.

3.2. Proposal of Two Algorithms Adapted to SLSTR

We propose two alternative SLSTR algorithms that are based on the split-window and the dual angle technique, respectively. The three main differences between the SWA proposed here and the Sobrino16, Zhang19, and Zheng19 SWAs are: (1) the use of the Cloudless Land Atmosphere Radiosounding (CLAR) database to calculate the coefficients of the proposed algorithms [50]; (2) the dependence of the LST retrieval algorithm on view angle; and (3) the independence of its coefficients from emissivity.

3.2.1. CLAR Database and Simulation Dataset

The CLAR database is composed of 382 clear-sky atmospheric profiles selected from radiosoundings compiled by the University of Wyoming [50]. These atmospheric profiles are relatively evenly distributed over the latitudes and, therefore, well suited to generate global algorithms: 40% of the radiosoundings belong to latitudes between 0°and 30°, 40% belong to latitudes between 30° and 60°, and 20% to latitudes higher than 60°. The WVC values of these profiles are distributed between nearly 0 and 7 cm. The temperatures of the lowest layer of the atmosphere range from 253 to 313 K.

Gaussian angles from 0° to 65° (0°, 11.6°, 26.1°, 40.3°, 53.7°, and 65°) were chosen to generate the dataset for training the SWA. Input T values were set to: $T_0 - 6$ K, $T_0 - 2$ K, $T_0 + 1$ K, $T_0 + 3$ K, $T_0 + 5$ K, $T_0 + 8$ K, and $T_0 + 12$ K following the global analysis performed in [50]. The dataset contained a total of 16,044 different cases and was used to obtain the algorithm coefficients.

For a SLSTR dual-angle algorithm (DAA), we used the same range of input temperatures T and two pairs of viewing angles: 0–53.7° and 11.6–53.7°. In this case, the total number of simulations in the dataset was 5348.

3.2.2. Split-Window Algorithm

The SWA presented in this work is based on the algorithm of Niclòs et al. in [15] for the Spinning Enhanced Visible and Infrared (SEVIRI) sensor on board METEOSAT Second Generation 2 (MSG-2), which is given by Equation (3):

$$T = T_{11} + a_0 + a_1(\sec(\theta) - 1) + (a_2 + a_3(\sec(\theta) - 1))(T_{11} - T_{12}) + (a_4 + a_5(\sec(\theta) - 1))(T_{11} - T_{12})^2 + \alpha(1 - \varepsilon) - \beta\Delta\varepsilon \quad (3)$$

where T is LST and T_{11} and T_{12} are at-sensor brightness temperatures in K for the SLSTR channels at 11 µm and 12 µm, respectively; $\varepsilon = 0.5(\varepsilon_{11} + \varepsilon_{12})$ is the mean emissivity for the SLSTR channels at 11 µm and 12 µm and $\Delta\varepsilon = \varepsilon_{11} - \varepsilon_{12}$ is the difference between them; θ is the sensor viewing angle; and $\alpha = (a_6 + a_7 W + a_8 W^2)$ and $\beta = (a_9 + a_{10} W)$ determine the emissivity correction term, with W defined as the WVC divided by the cosine of the viewing angle. The values of the algorithm coefficients a_0 to a_{10} are given in Table 2. Emissivities obtained for each land cover (Table 1) and WVC from SLSTR L1 auxiliary data were used for the application of the algorithm.

Table 2. Coefficients of the proposed split-window algorithm (Equation (3)) and the dual-angle algorithms (Equation (7)).

Coefficient	Split-Window	Coefficient	Dual-Angle 11 μm	Dual-Angle 12 μm
a_0 (K)	0.052 ± 0.013	c_0 (K)	-0.18 ± 0.02	-0.27 ± 0.04
a_1 (K)	0.15 ± 0.02	c_1	2.03 ± 0.02	2.28 ± 0.04
a_2	0.95 ± 0.02	c_2 (K^{-1})	0.114 ± 0.005	0.198 ± 0.007
a_3	-0.30 ± 0.03	c_3 (K)	57.56 ± 0.15	66.02 ± 0.19
a_4 (K^{-1})	0.305 ± 0.004	c_4 (K cm^{-1})	1.85 ± 0.11	-4.35 ± 0.14
a_5 (K^{-1})	0.202 ± 0.007	c_5 (K cm^{-2})	-1.278 ± 0.018	-0.81 ± 0.02
a_6 (K)	52.51 ± 0.18	c_6 (K)	132.2 ± 0.3	139.4 ± 0.4
a_7 (K cm^{-1})	-0.11 ± 0.12	c_7 (K cm^{-1})	-21.80 ± 0.07	-26.05 ± 0.11
a_8 (K cm^{-2})	-1.004 ± 0.018	–	–	–
a_9 (K)	75.7 ± 0.2	–	–	–
a_{10} (K cm^{-1})	-11.21 ± 0.06	–	–	–

The atmospheric coefficients (from a_0 to a_5) in Equation (3) were obtained from regression analyses between LST–T_{11} and T_{11}–T_{12} in Figure 4, using the blackbody approach ($\varepsilon = 1$ and $\Delta\varepsilon = 0$ [51]) and, therefore, the obtained coefficients are independent from emissivity. The emissivity correction term is controlled by α and β [52], which depend on atmospheric parameters (i.e., atmospheric transmissivity, at-surface brightness temperature, and effective atmospheric temperature).

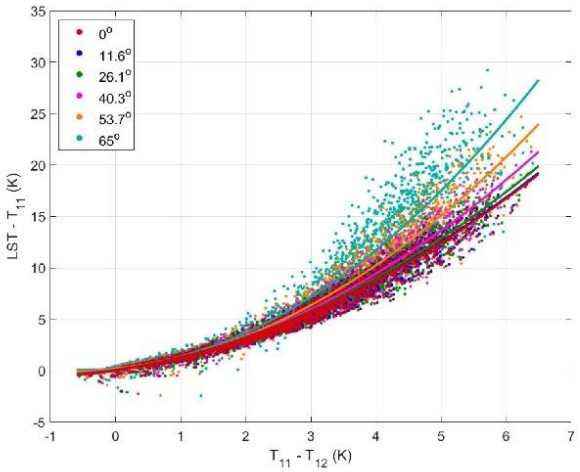

Figure 4. LST–T_{11} against T_{11}–T_{12} simulated from the CLAR database at the different view angles for the SLSTR SWA atmospheric coefficients retrieval. The regression functions corresponding to each angular dataset are plotted as lines in the same color as their corresponding data.

The uncertainties of the LST retrieved with the algorithm were obtained as the square root of the quadratic sum of model fitting uncertainty, $\delta(T)_M$, and propagated input parameter uncertainties, $\delta(T)_p$ [15]:

$$\delta(T) = \left[\delta(T)_M^2 + \delta(T)_p^2\right]^{1/2} \quad (4)$$

where both error sources are considered independent and were defined as follows:

$$\delta(T)_M = \left[\sigma_{AC}^2 + [(1-\varepsilon)\sigma_\alpha]^2 + [\Delta\varepsilon\,\sigma_\beta]^2\right]^{1/2} \quad (5)$$

$$\delta(T)_p = \left[\sum_i \left[\frac{\partial T}{\partial x_i}\delta x_i\right]^2\right]^{1/2} \quad (6)$$

where σ_{AC} is the fitting error associated with the atmospheric coefficients (from a_0 to a_5) and σ_α and σ_β are the fitting errors associated with α and β, respectively. The fitting error was defined as the standard error obtained from the regression analyses for each set of coefficients. The regression standard error was estimated by minimizing the sum of squared deviations from the predictions over the simulation dataset. The propagation uncertainty of the input parameters is expressed by Equation (6), where the partial derivative of T with respect to each input parameter x_i (i.e., emissivity, WVC, brightness temperatures) is estimated and multiplied by uncertainty δx_i. The experimental emissivity uncertainties in Table 1 were assigned and WVC uncertainty was assumed to be ±0.5 cm, which is considered to be a representative value [15,21]. Brightness temperature uncertainty is the noise equivalent error of the instrument, which is about ±0.05 K for the SLSTR thermal bands at 11 and 12 µm for a temperature of 270 K [53]. As the latter is a random uncertainty element, it must be divided by the square root of the number of pixels used to average the LST [54]. The mean and SD of the LST uncertainty contributions from each parameter are given in Table 3. Full vegetation and flooded soil were grouped together due to their similar emissivity values and were assigned the same emissivity uncertainties. For all cases, the main uncertainty sources were modeling and emissivity.

Table 3. Mean and SD of the uncertainty contributions obtained for the simulation dataset. The different uncertainty sources (modeling uncertainty and input parameters: emissivity, $\delta(T)_\varepsilon$; WVC, $\delta(T)_W$; brightness temperature, $\delta(T)_{BT}$), and total SLSTR LST retrieval uncertainty are shown.

Surface	δ	Split-Window		Dual-Angle 11 µm		Dual-Angle 12 µm	
		Mean (K)	**SD (K)**	**Mean (K)**	**SD (K)**	**Mean (K)**	**SD (K)**
Dry/Wet Bare Soil	$\delta(T)_\varepsilon$	0.50	0.14	0.74	0.10	0.76	0.12
	$\delta(T)_W$	0.09	0.04	0.03	0.02	0.04	0.02
	$\delta(T)_{BT}$	0.08	0.02	0.101	0.010	0.114	0.013
	$\delta(T)_p$	0.52	0.13	0.75	0.10	0.77	0.11
	$\delta(T)_M$	1.4441	0.0009	0.9203	0.0002	1.4996	0.0002
	$\delta(T)$	**1.54**	**0.05**	**1.19**	**0.06**	**1.69**	**0.05**
Water / Full vegetation	$\delta(T)_\varepsilon$	0.32	0.09	0.53	0.10	0.48	0.12
	$\delta(T)_W$	0.04	0.03	0.05	0.02	0.10	0.02
	$\delta(T)_{BT}$	0.09	0.02	0.109	0.006	0.131	0.010
	$\delta(T)_p$	0.36	0.08	0.54	0.10	0.51	0.11
	$\delta(T)_M$	1.4362	0.0012	0.909	0.003	1.492	0.009
	$\delta(T)$	**1.49**	**0.02**	**1.06**	**0.05**	**1.58**	**0.04**

3.2.3. Dual-Angle Algorithm

SLSTR's dual view also allows retrieving LSTs with DAAs. However, over land surfaces DAAs perform worse than SWAs, which is mainly due to differences in footprints and observation geometries between the two views [16]. Here, we analyzed the specific dual-view capability of the SLSTR TIR channels to retrieve LST. The DAA used here was adapted from [9] and is given by Equation (7):

$$T = T_n + c_1(T_n - T_b) + c_2(T_n - T_b)^2 + c_0 + \alpha(1-\varepsilon) - \beta\Delta\varepsilon \quad (7)$$

where T is the LST, c_0 to c_2 are the atmospheric coefficients, T_n and T_b are the brightness temperatures corresponding to nadir view (n) and backward view (b). ε is the mean emissivity for SLSTR nadir and backward views ($\varepsilon = 0.5(\varepsilon_n + \varepsilon_b)$) and $\Delta\varepsilon$ is the emissivity difference between nadir and backward views ($\Delta\varepsilon = \varepsilon_n - \varepsilon_b$); $\alpha = \left(c_3 + c_4 W + c_5 W^2\right)$) and $\beta = (c_6 + c_7 W)$ are functions modifying the impact of emissivity on LST retrieval, where W is the water vapor content. The coefficients determined for the two DAAs (one for each channel) are given in Table 2.

LST uncertainty for the DAA was estimated in analogy to the SWA with Equations (4)–(6) for the simulation dataset. The same input parameters uncertainties were used to estimate the dual-angle LST uncertainty. The mean uncertainty contribution of each input parameter, the algorithm fitting errors, and the mean LST uncertainty for the two DAAs are shown in Table 3. For the DAA at 12 µm (DAA12), the main uncertainty sources are the fitting error and the emissivity, as for the SWA. However, for the dual-angle algorithm at 11 µm (DAA11), the fitting error is lower than for the SWA, as it was found for AATSR [38].

3.3. Alternative Split-Window Algorithms

Various SWAs with explicit emissivity dependence were proposed as alternatives to the operational AATSR/SLSTR LST product algorithm. These alternative SWAs used the same input parameters (i.e., emissivity, WVC, and brightness temperatures) as the adapted SWA (Section 3.2.2).

3.3.1. Sobrino16 Split-Window Algorithm

The Sobrino16 SWA [21] employed the algorithm given by Equation (8) for the retrieval of LST from AATSR:

$$T = T_{11} + d_1(T_{11} - T_{12}) + d_2(T_{11} - T_{12})^2 + d_0 + (d_3 + d_4 W)(1 - \varepsilon) + (d_5 + d_6 W)\Delta\varepsilon \quad (8)$$

where T is LST and T_{11} and T_{12} are at-sensor brightness temperatures in K for the SLSTR channels at 11 and 12 µm, respectively; W is the WVC divided by the cosine of the viewing angle, $\varepsilon = 0.5(\varepsilon_{11} + \varepsilon_{12})$ is the mean emissivity for the SLSTR channels at 11 and 12 µm, and $\Delta\varepsilon = \varepsilon_{11} - \varepsilon_{12}$ is the corresponding difference between them; d_i, for i from 0 to 6, are the coefficients of the Sobrino16 SWA.

In order to obtain coefficients d_i, a broad range of T_{11} and T_{12} were simulated with the MODTRANv4 radiative transfer code [55] for 61 atmospheric profiles selected from the Thermodynamic Initial Guess Retrieval version 1 (TIGR-1) database and 108 emissivity spectra obtained from the ASTER Spectral Library [56]. For each atmospheric profile, five T values were simulated: T_0 − 5 K, T_0, T_0 + 5 K, T_0 + 10 K, T_0 + 20 K, where T_0 is the air temperature of the lowest level of the atmospheric profile. Additionally, five viewing angles (0°, 10°, 20°, 30°, and 40°) were simulated. Based on the uncertainties of the input parameters (emissivity, WVC and brightness temperatures) and model regression uncertainty, a final algorithm uncertainty of ±1.6 K was estimated [21].

3.3.2. Zhang19 Split-Window Algorithm

The Zhang19 SWA [22] was developed to improve LST retrieval over barren surfaces. This algorithm is given by Equation (9):

$$T = d_1 T_{11} + d_2(T_{11} - T_{12}) + d_3(T_{11} - T_{12})^2 + d_0 + (d_4 - d_5 W)(1 - \varepsilon) + (d_6 - d_7 W)\Delta\varepsilon \quad (9)$$

where the variables represent the same quantities as in Equation (8). For the simulation dataset, 60 clear-sky atmospheric profiles were selected from the TIGR2000 database. These atmospheric profiles were used as input to the MODTRANv5.2 code to obtain simulated values of T_{11} and T_{12}. For each atmospheric profile, the input T varied with T_0 as T_0 − 5 K, T_0, T_0 + 5 K, T_0 + 10 K, T_0 + 20 K when T_0 > 280 K, and T_0 − 5 K, T_0, T_0 + 5 K when $T_0 \leq 280$ K. Average emissivity (ε) was varied from 0.9 to 1.0 in steps of 0.02, and the difference in emissivity ($\Delta\varepsilon$) varied from 0.02 to −0.02 in steps of 0.005. Simulations were

performed for two viewing angles and yielded a total dataset of 30,456 simulated cases. According to [22], the SWA uncertainty ranges between ±0.5 K and ±1 K, depending on WVC. These values represent the uncertainty of the algorithm and do not consider input parameters uncertainties.

3.3.3. Zheng19 Split-Window Algorithm

Based on the refined form of the generalized split-window algorithm [57] proposed in [58], Zheng et al. in [15] adjusted the algorithm to match the spectral channels of SLSTR. The Zheng19 algorithm is described by Equation (10):

$$T = d_0 + \left(d_1 + d_2\frac{1-\varepsilon}{\varepsilon} + d_3\frac{\Delta\varepsilon}{\varepsilon^2}\right)\frac{T_{11}+T_{12}}{2} + \left(d_4 + d_5\frac{1-\varepsilon}{\varepsilon} + d_6\frac{\Delta\varepsilon}{\varepsilon^2}\right)\frac{T_{11}-T_{12}}{2} + d_7(T_{11}-T_{12})^2 \quad (10)$$

where the variables are the same as in Equation (8). MODTRANv5.2 [55] was used to obtain a simulated dataset of T_{11} and T_{12} from 946 clear-sky atmospheric profiles selected from the TIGR-3 database. The input T values varied from T_0, ranging from $T_0 - 10$ K to $T_0 + 30$ K in steps of 5 K. Sixty emissivity spectra from the ASTER spectral library and the University of California–Santa Barbara Emissivity Library were used to simulate a dataset under five viewing angles (0°, 15°, 25°, 35°, and 45°), which resulted in a total of 2,550,200 different cases.

The simulation dataset was divided into 160 groups to obtain coefficients stratified by WVC, brightness temperature T_{11} and five viewing angles. However, only the coefficients for nadir view and four brightness temperature and WVC subranges were published in [23]. For these subranges, an algorithm uncertainty ranging from ±0.6 K to ±2.1 K was estimated by propagating model regression uncertainty and emissivity uncertainty.

4. Validation of Satellite LST Products

4.1. Analysis of In-Situ Measurements

The number of transect measurements (12) was too limited for a statistical analysis while we had a sufficient number of permanent station matchups for validating SLSTR LST retrieved with the proposed adapted algorithms (201), the operational LST product (194), and LST retrieved with the other emissivity-dependent algorithms (201) from SLSTR L1 data. Based on the above, it was decided to use the transect measurements to analyze the spatial representativeness of the station measurements. Figure 5 compares the simultaneous measurements obtained with the (mobile) CE-312 and the (fixed) SI-121 radiometers.

The comparison between fixed station measurements and transect measurements shown in Figure 5 yielded a RMSD less than 0.4 K. These results indicate that the permanent station LST values are representative of the site, since they are in good agreement with the LST along the transects. The homogeneity of the area allows us to use the permanent station measurements for validating satellite LST, which is in agreement with previous studies (e.g., [27,29]).

The uncertainties shown in Figure 5 were estimated from the average of the propagated uncertainty for each measured variable and its standard deviation (i.e., within 3 min before and after the satellite overpasses). Mean uncertainty values of ±0.7 K and ±0.3 K were obtained from the CE-312 measurements and the SI-121 measurements, respectively; the values for the CE-312 are larger due to the spatial variability along the transects.

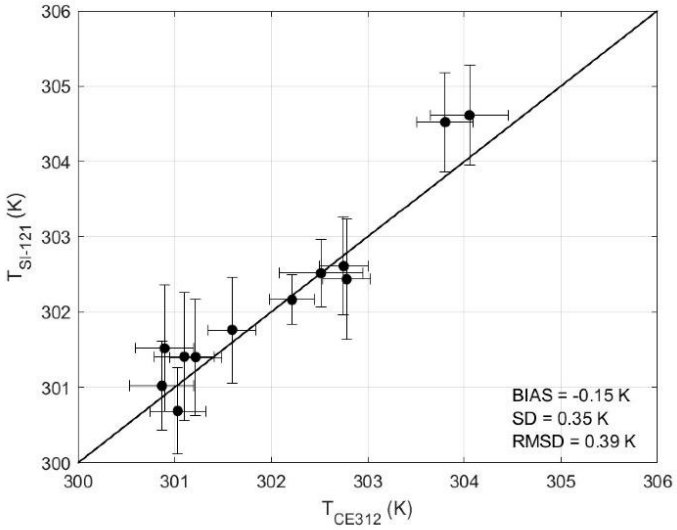

Figure 5. LST obtained with the fixed SI-121 radiometer compared to LST obtained along the transects with mobile CE-312 radiometers.

4.2. Operational SLSTR LST Product

The operational SLSTR LST product was evaluated against in-situ LSTs. Average LST weighted by the inverse of the squared distance to the site coordinates was obtained for the 2×2 closest satellite pixels. The statistical parameters are the median of the differences ($T_{SLSTR} - T_{ground}$), robust standard deviation (RSD, given by Equation (11)), and robust root mean squared difference (R-RMSD), which is obtained as the square root of the quadratic sum of the median and the RSD.

$$\mathrm{RSD} = \mathrm{median}\left|\left(T_{SLSTR} - T_{ground}\right)_i - \mathrm{median}\left(T_{SLSTR} - T_{ground}\right)_i\right| \cdot 1.483 \quad (11)$$

Figure 6 shows the comparison of the operational Sentinel-3A SLSTR LST product on cloudless days against the corresponding in-situ LST obtained from SI-121 measurements (daytime and nighttime data; land cover types in different colors). Robust statistics were used in this analysis to avoid outlier effects [59].

as for Sentinel-3A. The statistical analysis yielded a median of 1.5 K and an RSD of 1.2 K for all surfaces combined. The validation statistics are summarized in Table 5 for all data as well as separately for daytime and nighttime.

Table 5. Validation statistics for the operational Sentinel-3B SLSTR LST product against in-situ LST for the three land covers at the Valencia rice paddy site. All the statistics are in Kelvin (K) and N is the number of data points.

	All Data				Daytime				Nighttime			
	MEDIAN	RSD	R-RMSD	N	MEDIAN	RSD	R-RMSD	N	MEDIAN	RSD	R-RMSD	N
All Surfaces	1.5	1.2	1.9	107	1.6	1.3	2.0	41	1.3	1.0	1.7	66
Flooded soil	2.1	0.6	2.2	48	2.5	1.1	2.7	15	1.9	0.7	2.0	33
Bare soil	0.8	1.1	1.3	31	0.8	1.7	1.8	13	0.8	1.0	1.3	18
Vegetation	1.0	1.3	1.6	28	1.4	0.9	1.7	13	0.5	1.2	1.3	15

4.3. LST Retrieved with Explicit Emissivity-Dependent Algorithms

The proposed SWA with angular dependence and explicit dependence on surface emissivity (adapted from [15]; see Section 3.2) was analyzed using the larger dataset of Sentinel-3A SLSTR measurements. Additionally, the alternative SWAs discussed in Section 3.3 were evaluated.

Figure 7 shows the LSTs obtained with the SWAs against the in-situ LSTs obtained from the SI-121 measurements at the permanent validation station.

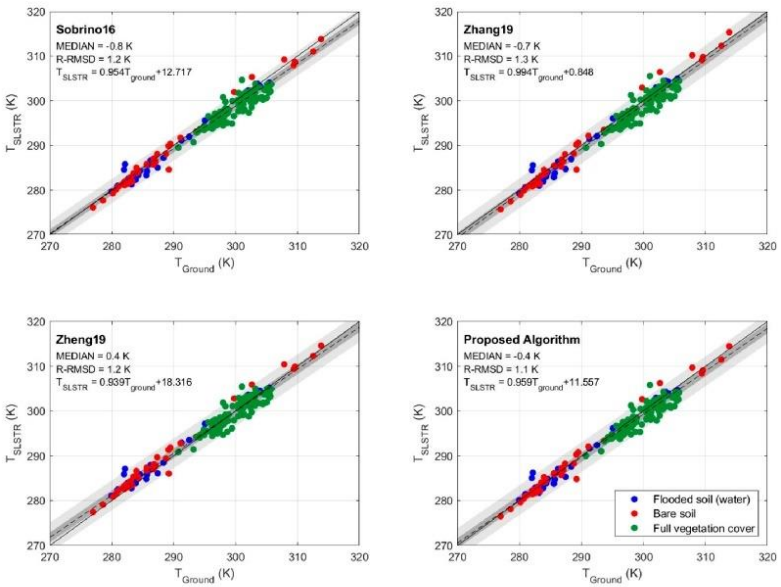

Figure 7. LST retrieved from Sentinel-3A with emissivity-dependent algorithms against in-situ LST obtained from the SI-121 radiometer. Top left: Sobrino16. Top right: Zhang19. Bottom left: Zheng19. Bottom right: the proposed algorithm.

In Figure 7, LSTs range from 277 to 315 K, covering a wide range of values. Data for bare soil and flooded soil cover larger LST ranges, while full vegetation covers a smaller range (i.e., between 290 and 306 K). A median (RSD) of −0.4 K (1.1 K) was obtained for the proposed SWA. Similar statistical results were obtained for the other emissivity-dependent SWA: median (RSD) of −0.8 K (0.9 K) for Sobrino16, −0.7 K (1.1 K) for Zhang19, and 0.4 K (1.1 K) for Zheng19. In total 198 points were used (32 flooded soil, 38 bare soil and 128 full vegetation). The statistics for all validation results are summarized in Table 6.

Table 6. Validation statistics for the four emissivity-dependent split-window algorithms for the three land covers at the Valencia rice paddy site. All values are in Kelvin (K) and N is the number of data points.

		MEDIAN	RSD	R−RMSD	N
All Surfaces	Sobrino16	−0.8	0.9	1.2	198
	Zhang19	−0.7	1.1	1.3	198
	Zheng19	0.4	1.1	1.2	198
	Proposed SWA	−0.4	1.1	1.1	198
Flooded Soil	Sobrino16	−0.4	0.6	0.7	32
	Zhang19	−0.5	1.0	1.1	32
	Zheng19	1.0	0.7	1.2	32
	Proposed SWA	0.0	0.6	0.6	32
Bare Soil	Sobrino16	−0.4	0.9	0.9	38
	Zhang19	−0.5	0.6	0.8	38
	Zheng19	0.9	0.7	1.2	38
	Proposed SWA	−0.2	0.9	0.9	38
Full Vegetation	Sobrino16	−1.0	1.0	1.4	128
	Zhang19	−0.9	1.2	1.5	128
	Zheng19	−0.1	1.3	1.3	128
	Proposed SWA	−0.7	1.2	1.4	128

Medians are lower than the RSDs, except for a few cases. The results obtained over flooded and bare soils are slightly better than those over full vegetation, in terms of both bias and RSD, for all algorithms. Considering all surfaces, the SWA proposed here obtains the lowest R-RMSD.

The same statistical analysis was repeated for the full dataset for daytime and nighttime cases separately. Table 7 shows the corresponding median, RSD, and R-RMSD. The total data used for each surface at daytime (nighttime) are 99 (99) points for all surfaces, 16 (16) points for flooded soil, 14 (24) for bare soil, and 69 (59) points for full vegetation. Better results are obtained in general for nighttime cases than for daytime, especially over bare soil and full vegetation surfaces.

Table 7. Validation statistics for the different emissivity-dependent split-window algorithms. Results are shown for all data and separately for flooded soil, bare soil, and full vegetation cover. All values are in Kelvin (K).

		Daytime			Nighttime		
		MEDIAN	RSD	R−RMSD	MEDIAN	RSD	R−RMSD
All Surfaces	Sobrino16	−0.8	1.2	1.5	−0.9	0.8	1.2
	Zhang19	−0.5	1.3	1.4	−0.9	0.8	1.2
	Zheng19	0.5	1.4	1.5	0.2	1.1	1.1
	Proposed SWA	−0.3	1.5	1.5	−0.5	0.8	0.9
Flooded Soil	Sobrino16	−0.3	0.7	0.7	−0.5	0.6	0.8
	Zhang19	−0.4	0.8	0.9	−0.9	0.9	1.2
	Zheng19	1.1	0.6	1.3	0.9	0.9	1.3
	Proposed SWA	0.2	0.7	0.8	−0.1	0.7	0.7

Table 7. *Cont.*

		Daytime			Nighttime		
		MEDIAN	**RSD**	**R−RMSD**	**MEDIAN**	**RSD**	**R−RMSD**
Bare Soil	Sobrino16	−0.4	1.5	1.5	−0.4	0.8	0.9
	Zhang19	0.6	1.5	1.6	−0.6	0.7	1.0
	Zheng19	0.8	1.4	1.6	1.0	0.5	1.1
	Proposed SWA	0.2	1.5	1.5	−0.2	0.6	0.6
Full Vegetation	Sobrino16	−1.3	1.5	1.9	−1.0	0.7	1.2
	Zhang19	−0.7	1.5	1.7	−0.9	0.8	1.2
	Zheng19	0.0	1.4	1.4	−0.1	0.9	0.9
	Proposed SWA	−0.8	1.6	1.8	−0.7	0.8	1.0

4.4. Proposed Dual-Angle Algorithms

Two DAAs with coefficients generated using the CLAR database were proposed for the SLSTR TIR channels at 11 µm and 12 µm. Figure 8 shows the LST retrieved by DAA11 and DAA12 (for the station pixel) against the in-situ LST obtained from the SI-121 measurements at the station.

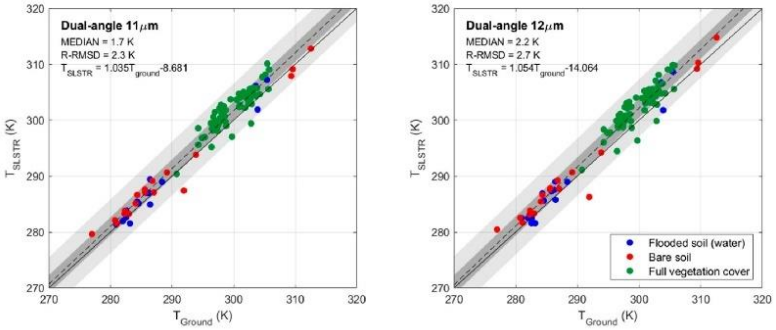

Figure 8. SLSTR LST retrieved with the dual-angle algorithms for the 11 µm channel (left; DAA11) and 12 µm channel (right; DAA12) against in-situ LST for the three seasonal land covers at the Valencia rice paddy site.

Both DAAs overestimated in-situ LST: DAA11 yielded better statistics for all surfaces combined with median (RSD) of 1.7 K (1.6 K) than DAA12, for which a median (RSD) of 2.2 K (1.7 K) was obtained. The validation statistics for both DAAs are summarized in Table 8.

Table 8. Validation statistics of the dual-angle algorithms for SLSTR 11 and 12 μm channels at the Valencia rice paddy site. All statistics are in Kelvin (K) and N is the number of data points.

	Dual-Angle 11 μm				Dual-Angle 12 μm			
	MEDIAN	**RSD**	**R-RMSD**	**N**	**MEDIAN**	**RSD**	**R-RMSD**	**N**
All Surfaces	1.7	1.6	2.3	102	2.2	1.7	2.7	102
Flooded Soil	0.6	1.0	1.1	15	1.0	2.3	2.5	15
Bare soil	1.0	1.3	1.7	18	1.4	1.2	1.8	18
Vegetation	2.1	1.2	2.5	69	2.6	1.5	3.0	69

5. Discussion

An explicit emissivity and angle dependent SWA and two DAAs for Sentinel-3 SLSTR (using the channels centered at 11 μm and 12 μm) were proposed and validated. The SWA was adapted from the SWA proposed in [15] for SEVIRI sensor, while the DAAs were adapted from an algorithm developed for AATSR in [16]. Although [16] determined a better performance for the AATSR SWA, the double view capability of SLSTR (i.e., its nadir and backward views) for LST retrieval should be analyzed in order to identify possible differences to AATSR. Furthermore, the operational Sentinel-3A and Sentinel-3B SLSTR L2 LST products and three explicit emissivity-dependent SWAs (i.e., Sobrino16 [21], Zhang19 [22], and Zheng19 [23] SWAs) were validated.

The validation used in-situ LSTs from a rice paddy site close to Valencia, Spain, which represents three seasonal homogeneous land cover types with different spectral features. These in-situ data were collected by two Apogee SI-121 wideband (8–14 μm) radiometers installed on a permanent station at the site. The narrower viewing geometry and spectral range makes TIR radiometers (e.g., Apogee SI-121, Heitronics KT15.85) more suitable for LST validation purposes than broadband hemispherical pyrgeometers (3–50 μm), which are commonly used [32]. Additionally, the uncertainty of typically used radiometers (e.g., ±0.2 K for Apogee SI-121) is lower than for pyrgeometers, which is around 1 K [60]. When considering the uncertainties in upwelling and downwelling radiance measurements and emissivity, the uncertainty of in-situ LST obtained with pyrgeometers results in a typical uncertainty of ±1 to ±2 K [61]. In [62], compared simultaneous measurements with wideband radiometers and a pyrgeometer over asphalt and four grassland sites. From this comparison, they observed a standard deviation of up to 2 K at the grassland sites and a general underestimation for the pyrgeometer data. This is in agreement with LST validations performed for various satellite sensors, e.g., MODIS [63], VIIRS [60], and Landsat-8 [64], which used pyrgeometer measurements as reference: especially at daytime, these studies obtained similar standard deviations of around 2 K at grassland sites.

The GlobCover classification map, which is based on a static global classification, is used for generating the SLSTR LST product. In order to consider surface changes due to vegetation, seasonal changes or cropland harvest, each coefficient of the operational SLSTR LST algorithm is obtained as a combination of a vegetation coefficient and a bare soil coefficient, weighted by their cover fractions. However, for flooded soil at the study site, the vegetation fraction is higher than 0.3 (Figure 2): while this may be plausible for the last few days considered as flooded soil, when the rice starts growing, it is implausible at the beginning of the flooding, when there is only water. According to agricultural laborers, changes on the surface should be more marked, since the site is flooded in a few days and is then covered entirely by water. However, for 15 out of 27 land cover types, the vegetation and bare soil coefficients provided in the SLSTR auxiliary data are the same, as is the case for the biome assigned to the study site (i.e., weighting by cover fraction has no effect).

Different coefficients for daytime and nighttime are provided only for water and flooded surface biomes (i.e., post-flooding or irrigated cropland). However, for most land cover types, e.g., bare soils, non-flooded forests, scrubland or grassland areas, the coefficients are the same for daytime and nighttime.

In the SLSTR LST algorithm, coefficients for irrigated cropland areas were obtained as an average of the coefficients for water, winter wheat, and broadleaf-deciduous trees according to the land cover classification given in [65]. Since the land cover of the study site changes over the year, only the period of full vegetation matches with the assigned biome. However, the best validation results were obtained for the bare soil cover at daytime (R-RMSD = 1.6 K) and nighttime (R-RMSD = 1.0 K). Similar results were obtained for the SLSTR LST product over arid areas by other authors. In [20], a RMSD of 1.9 K at the Gobabeb (Namibia) station was obtained, with a bias of 1.8 K and a SD of 0.8 K. In [23], a bias of 1.1 K and a SD of 0.9 K were obtained, leading to a RMSD of 1.4 K. In these two cases, as well as in this study, SLSTR LST had a good precision, i.e., lower than or equal to 1.0 K, but an accuracy larger than the GCOS threshold (>1 K). Yang et al. in [49] obtained a systematic uncertainty of 1.6 K and a RMSD of 2.4 K for Sentinel-3 SLSTR LST at the Gobabeb (Namibia) site. It should be noted that the biomes assigned to each validation site differ, so discrepancies due to different coefficients are possible.

For the Valencia rice paddy site, the validation over full vegetation cover shows considerably better results at nighttime, with median and RSD around 1K. However, for the daytime data, the median and RSD increase to 1.7 and 1.5 K, respectively. Due to the higher thermal heterogeneity at daytime, a slight increase in RSD is expected, but not the large increase observed for the median difference, which causes the daytime accuracy to miss the GCOS threshold. It is suspected that the increased median difference is mainly caused by different day and night retrieval coefficients. These results cannot be directly compared with results obtained over other vegetated areas, e.g., the Amazon site [48] and Evora [20]. The Amazon site [48] was classified as closed to open broadleaved evergreen and/or semi-deciduous forest (biome 5) and yielded a SLSTR LST bias of −0.1 and a SD of 0.6 K for daytime and nighttime data. For Evora [20], the assigned biome was rainfed croplands (biome 2) and the SLSTR LST bias was −0.8 (−0.4) K and SD 0.7 (0.3) K for daytime (nighttime). The biases obtained for these validation sites were relatively small and showed a slight LST underestimation, while an overestimation was found at the Valencia site, which was misclassified as biome 1 (irrigated cropland) with very different characteristics to a rice paddy.

For the evaluation of the operational Sentinel-3B SLSTR product, a total of 107 scenes (43 over flooded soil, 31 over bare soil, and 28 over full vegetation) were used. Compared to the validation results for Sentinel-3A, the obtained accuracy for full vegetation and bare soil was slightly better, while the precision was similar for full vegetation and worse for bare soil. For flooded soil, the validation results for Sentinel-3B were less accurate and more precise than for Sentinel-3A. As for Sentinel-3A, better results were observed for Sentinel-3B nighttime data, mainly because of the higher thermal homogeneity. For both sensors, large systematic uncertainty was observed over flooded soil (around 2 K for both daytime and nighttime). In contrast, over the deep and large water body of Lake Constance (classified as water body, biome 26), Yang et al. in [49] reported a considerably smaller systematic uncertainty of 0.4 K and a RMSD of 0.7 K for the operational Sentinel-3 SLSTR LST product.

The proposed SWA with explicit emissivity and angular dependence and the three published emissivity-dependent SWAs were validated under identical conditions. Generally, all investigated algorithms performed well, with median and RSD lower than 1.5 K over all surfaces. For all surfaces combined, the proposed algorithm yielded median (RSD) values of −0.4 K (1.1 (K): together with the Zheng19 SWA, it showed the lowest median (best accuracy). However, all SWAs obtained similar RSD values between 0.9 and 1.1 K. The better accuracy of the Zheng19 algorithm is mainly linked to its exceptionally low median over full vegetation cover, which also represented most data; the other SWA proposed here showed more consistently low median values for all three land covers.

The coefficients of the proposed SWA were based on a simulated dataset produced for LST ranging between −6 K and +12 K around the lowest level of air temperature (T_0). These values were determined in [50] from statistical analysis of MODIS products MOD08

and MOD11 for air temperature and LST values, respectively. This statistical analysis showed that the range of temperatures used for the simulation dataset covers most of the cases found over natural surfaces [50]. A maximum increment of up to +20 K was used to produce Sobrino16 and Zhang19, although these increments were only for $T_0 < 280$ K on the latter. The Zheng19 SWA was produced with even larger increments of up to +30 K: this can be interesting for some applications (e.g., urban heat island, analyses of extreme temperatures), but can also cause an overfitting of retrieval coefficients, which in turn can increase retrieval uncertainty, particularly over the most common natural surfaces [64].

The similarity of the results could be linked to the moderate WVCs at the site (ranging from 0.5 to 4.4 cm, with a mean value of 2.4 ± 0.9 cm and only 3% of data >4 cm), which implies small atmospheric effects and a small dependence on viewing angle. The effect of the differential absorption in the atmosphere in the regression of the proposed SWA per viewing angle was shown in Figure 4. For low to moderate brightness temperature differences, there is a minor angular dependence of the regression coefficients. However, for high brightness temperature differences, there is considerable angular dependence of the coefficients, corresponding to high WVCs (up to 7 cm) in the CLAR atmospheric database used for the regressions (brightness temperature differences were up to 6.4 K). For comparison, the largest values in the database for the brightness temperature difference in SLSTR channels 8 and 9 were around 4.0 K (with a mean of 1.5 ± 0.8 K). Thus, further validation experiments in tropical atmospheres and over regions with WVCs exceeding 4 cm should be performed to evaluate the algorithms in such extreme cases. Although there is a slight WVC seasonality (i.e., higher in the summer, lower in the winter), no significant differences observed in the results were unrelated to WVC, since no extreme WVC values were found at the site. Moreover, the uncertainty introduced by WVC (~0.1 K) on the SWA is negligible compared to the uncertainty introduced by emissivity (~0.5 K) or the retrieval algorithm (~1.4 K), as shown in Table 3. The difference in the accuracy obtained with the proposed SWA for viewing angles lower and higher than 40° was 0.3 K, with an associated difference in precision of 0.6 K. Based on the simulations shown in Figure 4, a decrease of precision with viewing angle was expected, since the atmospheric absorption increases considerably with viewing angle and, thus, also, the regression error. However, the relatively small change in accuracy indicates a good performance of the algorithm also at larger viewing angles.

At nighttime, for the three land covers at the rice paddy site, the algorithms showed good performance with accuracies and precisions better than the GCOS threshold (<1 K). In contrast, at daytime, the larger thermal heterogeneity caused an increase of RSDs for bare soil and full vegetation covers, with values of about 1.5 K. However, generally, similar accuracies were obtained at daytime and nighttime over bare soil and full vegetation cover and most values met the GCOS accuracy threshold.

For all algorithms the results were in agreement with previous validations performed by other authors at different sites: in [21] obtained a bias of −1.4 K and a SD of 1.2 K for a cropland area in Oklahoma, which is comparable to our site with full vegetation cover. Although they had few data points, their results showed a similar precision to that obtained for the rice paddy site. Similar results were also obtained for SLSTR LST at an Amazon site in [48], who obtained a bias of −1.3 K and a SD of 0.9 K. Zhang et al. in [22] obtained a bias of −0.4 K and a SD of 0.9 K for a desert area in Wuhai, which are similar results to those obtained here with the same algorithm for bare soil. Yang et al. in [49] trained nine SWAs to retrieved SLSTR LST. These SWAs were evaluated over the gravel plains at Gobabeb (Namibia) and Lake Constance (Germany, Switzerland and Austria); a bias from −0.2 K to −0.3 K (from −0.2 K to 0.3 K) and an RMSD of 1.6 K (0.5 K) were obtained at Gobabeb (Lake Constance). Finally, Zheng et al. in [23] validated their proposed SWA using pyrgeometers and radiometers over cropland and grassland sites. Their overall results showed a bias of 0.6 K and SD of 2.2 K, which is higher than the corresponding values obtained at our study site for all surfaces combined and for full vegetation cover. The underestimation reported for a station at Henan Hebi, China, with daytime data from a radiometer over cropland,

was similar to that for vegetation cover at our site, under similar conditions. The bias obtained in [23] was the same as the median obtained here, while SD deviation was 2.4 K at the Henan Hebi site. The RSD found here was 1.3 K and, therefore, the Zheng19 algorithm performed much better at our study site. The proposed DAAs, which use SLSTR's nadir and backward views, showed better results for the version applied to the 11 µm channel (DAA11), which over flooded soil yielded an accuracy and precision better than the GCOS threshold. For bare soil, the accuracy and precision were also close to the GCOS threshold. However, the accuracy was worse for full vegetation cover. DAA12 yielded R-RMSD values between 1.8 and 3 K for all land covers. These findings are in agreement with results for previous sensors (i.e., AATSR, [16,50]), where, regardless of land cover, DAAs also performed worse than SWAs, probably due to differences in sensor footprint between the views and directional effects on radiometric temperatures [16].

6. Conclusions

The operational SLSTR LST algorithm depends on biome, day/nighttime, vegetation fraction, and viewing zenith angle. From the validation results it is concluded that the operational Sentinel-3A SLSTR LST product is accurate for nighttime data, with an accuracy (systematic uncertainty, i.e., median) of 1.0 K and a precision (random uncertainty, i.e., RSD) of 1.0 K for the three investigated surfaces combined. In contrast, for daytime data an accuracy of 1.8 K and precision 1.2 K was determined. The increase in daytime RSD is attributed to the typically larger thermal heterogeneity of the land surface. In contrast, the increase in bias is thought to be caused by wrongly assigned biomes, i.e., the same coefficients were used for the three investigated land cover types. Additionally, the validation for the Sentinel-3B SLSTR LST product is of relevance since no robust validations were published for this platform. An accuracy of 1.5 K and a precision of 1.2 K were obtained, yielding to similar results to those obtained for the Sentinel-3A SLSTR LST product for all data combined.

The angular and emissivity-dependent algorithm proposed by Niclòs et al. in [15] for MSG SEVIRI was adapted to Sentinel-3 SLSTR. The adapted SLSTR SWA was evaluated together with three emissivity-dependent algorithms proposed by Sobrino et al. in [21], Zhang et al. in [22] and Zheng et al. in [23] using Sentinel-3A SLSTR L1 data. For all data combined (i.e., the three land cover types), the differences between LST obtained with the proposed algorithm and in-situ LST had a median (RSD) of −0.4 K (1.1 K); the respective values were −0.8 K (0.9 K) for Sobrino16, −0.7 K (1.1 K) for Zhang19, and 0.4 K (1.1 K) for Zheng19. While Zheng19 and the SWA proposed here achieved the overall best accuracies, the latter showed a more consistent performance for the three investigated land covers. These cover a wide range of natural surface emissivities, i.e., from low values for dry bare soil, to medium values for wet bare soil, and high emissivity values for vegetation and water surfaces. Additionally, the explicit angular dependence of the proposed SWA will have higher benefits over areas with higher WVC, which is also illustrated by simulation data).

The overall accuracy improvements of the proposed SWA compared to the operational product is of 0.9 K, while it is 0.4 and 0.3 K compared to Sobrino16 and Zhang19 SWAs, respectively. The achieved improvements are highly significant, e.g., for climatological studies: when performing LST trend analyses, a global LST increase of 0.27 K/decade was observed from satellite data [66], i.e., the observed trends per decade are still smaller than the accuracy improvement achieved by the proposed algorithm.

Furthermore, a DAA was proposed to investigate the usefulness of SLSTR's dual-view capability for LST retrieval and separate sets of coefficients were determined for the 11 and 12 µm channels. While DAA11 performed better than DAA12, the dual-view algorithms still performed worse than the SWAs. However, an acceptable accuracy and precision of DAA11 was found over flooded soil and bare soil at the Valencia rice paddy site.

Over the rice paddy site, the explicitly emissivity-dependent SWAs were found to perform better than the operational Sentinel-3 SLSTR algorithm with biome-dependent

coefficients. Among the emissivity-dependent SWAs, the proposed algorithm with explicit angular dependence showed a slightly better performance at the three land covers. The results of this algorithm are expected to improve for more humid atmospheres (i.e., WCV > 4 cm), where the impact of the angular effect is higher due to the increased atmospheric absorption.

Author Contributions: Conceptualization, R.N. and L.P.-P.; methodology, R.N. and L.P.-P.; validation, L.P.-P.; formal analysis, R.N. and L.P; resources and data curation, J.A.V. and J.P.; software, J.P., J.M.G. and L.P.-P.; writing—original draft preparation, L.P.-P. and R.N.; writing—review and editing, C.C., F.-M.G. and E.V.; funding acquisition, R.N., E.V., C.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Spanish Ministry of Economy and Competitiveness and the European Regional Development Fund (FEDER) through project CGL2015-64268-R (MINECO/FEDER, UE), by the grant supported by the regional program of training of technicians for R&D&I—Youth Guarantee GJIDI-2018-A-142, and by the Spanish Ministry of Science and Innovation funded contract Torres Quevedo PTQ2018-010068. The validation campaign of 2016 was also funded by the Spanish Ministry of Economy and Competitiveness under the project CGL2013-46862-C2-1-P.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sánchez, J.M.; López-Urrea, R.; Dona, C.; Caselles, V.; González-Piqueras, J.; Niclos, R. Modeling evapotranspiration in a spring wheat from thermal radiometry: Crop coefficients and E/T partitioning. *Irrig. Sci.* **2015**, *33*, 399–410. [CrossRef]
2. Mokhtari, A.; Noory, H.; Pourshakouri, F.; Haghighatmehr, P.; Afrasiabian, Y.; Razavi, M.; Fereydooni, F.; Naeni, A.S. Calculating potential evapotranspiration and single crop coefficient based on energy balance equation using Landsat 8 and Sentinel. *ISPRS J. Photogramm. Remote Sens.* **2019**, *154*, 231–245. [CrossRef]
3. Bian, Z.; Roujean, J.-L.; Lagouarde, J.-P.; Cao, B.; Li, H.; Du, Y.; Liu, Q.; Xiao, Q.; Liu, Q. A semi-empirical approach for modeling the vegetation thermal infrared directional anisotropy of canopies based on using vegetation indices. *ISPRS J. Photogramm. Remote Sens.* **2020**, *160*, 136–148. [CrossRef]
4. Li, Z.-L.; Tang, B.-H.; Wu, H.; Ren, H.; Yan, G.; Wan, Z.; Trigo, I.; Sobrino, J.A. Satellite-derived land surface temperature: Current status and perspectives. *Remote Sens. Environ.* **2013**, *131*, 14–37. [CrossRef]
5. Csiszar, I.; Schroeder, W.; Giglio, L.; Ellicott, E.; Vadrevu, K.P.; Justice, C.O.; Wind, B. Active fires from the Suomi NPP Visible Infrared Imaging Radiometer Suite: Product status and first evaluation results. *J. Geophys. Res. Atmos.* **2014**, *119*, 803–816. [CrossRef]
6. Wang, J.; Roudini, S.; Hyer, E.J.; Xu, X.; Zhou, M.; Garcia, L.C.; Reid, J.S.; Peterson, D.A.; da Silva, A.M. Detecting nighttime fire combustion phase by hybrid application of visible and infrared radiation from Suomi NPP VIIRS. *Remote Sens. Environ.* **2020**, *237*, 111466. [CrossRef]
7. Cigna, F.; Tapete, D.; Lu, Z. Remote Sensing of Volcanic Processes and Risk. *Remote Sens.* **2020**, *12*, 2567. [CrossRef]
8. Nádudvari, Á.; Abramowicz, A.; Maniscalco, R.; Viccaro, M. The Estimation of Lava Flow Temperatures Using Landsat Night-Time Images: Case Studies from Eruptions of Mt. Etna and Stromboli (Sicily, Italy), Kīlauea (Hawaii Island), and Eyjafjallajökull and Holuhraun (Iceland). *Remote Sens.* **2020**, *12*, 2537. [CrossRef]
9. Gerhards, M.; Schlerf, M.; Mallick, K.; Udelhoven, T. Challenges and Future Perspectives of Multi-/Hyperspectral Thermal Infrared Remote Sensing for Crop Water-Stress Detection: A Review. *Remote Sens.* **2019**, *11*, 1240. [CrossRef]
10. Hatton, N.; Sharda, A.; Schapaugh, W.; van der Merwe, D. Remote thermal infrared imaging for rapid screening of sudden death syndrome in soybean. *Comput. Electron. Agric.* **2020**, *178*, 105738. [CrossRef]
11. Chang, S.; Chen, H.; Wu, B.; Nasanbat, E.; Yan, N.; Davdai, B. A Practical Satellite-Derived Vegetation Drought Index for Arid and Semi-Arid Grassland Drought Monitoring. *Remote Sens.* **2021**, *13*, 414. [CrossRef]
12. GCOS. The Global Observing System for Climate: Implementation Needs. *World Meteorol. Organ.* **2016**, *200*, 341. Available online: https://library.wmo.int/opac/doc_num.php?explnum_id=3417 (accessed on 1 May 2021).
13. Hollmann, R.; Merchant, C.J.; Saunders, R.; Downy, C.; Buchwitz, M.; Cazenave, A.; Chuvieco, E.; Defourny, P.; De Leeuw, G.; Forsberg, R.; et al. The ESA Climate Change Initiative: Satellite Data Records for Essential Climate Variables. *Bull. Am. Meteorol. Soc.* **2013**, *94*, 1541–1552. [CrossRef]
14. Lequin, R.M. Guide to the Expression of Uncertainty of Measurement: Point/Counterpoint. *Clin. Chem.* **2004**, *50*, 977–978. [CrossRef] [PubMed]
15. Niclòs, R.; Galve, J.M.; Valiente, J.A.; Estrela, M.J.; Coll, C. Accuracy assessment of land surface temperature retrievals from MSG2-SEVIRI data. *Remote Sens. Environ.* **2011**, *115*, 2126–2140. [CrossRef]

16. Coll, C.; Caselles, V.; Galve, J.M.; Valor, E.; Niclòs, R.; Sánchez, J.M. Evaluation of split-window and dual-angle correction methods for land surface temperature retrieval from Envisat/Advanced Along Track Scanning Radiometer (AATSR) data. *J. Geophys. Res. Space Phys.* **2006**, *111*, 1–12. [CrossRef]
17. Fisher, D.; Wooster, M.J. Multi-decade global gas flaring change inventoried using the ATSR-1, ATSR-2, AATSR and SLSTR data records. *Remote Sens. Environ.* **2019**, *232*, 111298. [CrossRef]
18. Ghent, D.J.; Corlett, G.K.; Göttsche, F.-M.; Remedios, J.J. Global Land Surface Temperature from the Along-Track Scanning Radiometers. *J. Geophys. Res. Atmos.* **2017**, *122*, 12–167. [CrossRef]
19. Sentinel-3 Optical Products and Algorithm Definition: SLSTR Land Surface Temperarure Algorithm Theoretical Basis Document (ATBD). 2012. Available online: https://sentinel.esa.int/documents/247904/349589/SLSTR_Level-2_LST_ATBD.pdf (accessed on 1 May 2021).
20. Ghent, D. S3 Validation Report—SLSTR. Internal Publication, S3MPC.UOL.VR.029 Issue 1.0, 65p. Available online: https://sentinels.copernicus.eu/documents/247904/3320896/Sentinel-3-SLSTR-Level-2-Land-Validation-Report (accessed on 1 May 2021).
21. Sobrino, J.; Jiménez-Muñoz, J.; Sòria, G.; Ruescas, A.; Danne, O.; Brockmann, C.; Ghent, D.; Remedios, J.; North, P.; Merchant, C.; et al. Synergistic use of MERIS and AATSR as a proxy for estimating Land Surface Temperature from Sentinel-3 data. *Remote Sens. Environ.* **2016**, *179*, 149–161. [CrossRef]
22. Zhang, S.; Duan, S.-B.; Li, Z.-L.; Huang, C.; Wu, H.; Han, X.-J.; Leng, P.; Gao, M. Improvement of Split-Window Algorithm for Land Surface Temperature Retrieval from Sentinel-3A SLSTR Data Over Barren Surfaces Using ASTER GED Product. *Remote Sens.* **2019**, *11*, 3025. [CrossRef]
23. Zheng, Y.; Ren, H.; Guo, J.; Ghent, D.; Tansey, K.; Hu, X.; Nie, J.; Chen, S. Land Surface Temperature Retrieval from Sentinel-3A Sea and Land Surface Temperature Radiometer, Using a Split-Window Algorithm. *Remote Sens.* **2019**, *11*, 650. [CrossRef]
24. Miralles, V.C.; Valor, E.; Boluda, R.; Caselles, V.; Coll, C. Influence of soil water content on the thermal infrared emissivity of bare soils: Implication for land surface temperature determination. *J. Geophys. Res. Space Phys.* **2007**, *112*, 1–11. [CrossRef]
25. Coll, C.; Galve, J.M.; Sanchez, J.M.; Caselles, V. Validation of Landsat-7/ETM+ Thermal-Band Calibration and Atmospheric Correction With Ground-Based Measurements. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 547–555. [CrossRef]
26. Coll, C.; Garcia-Santos, V.; Niclos, R.; Caselles, V. Test of the MODIS Land Surface Temperature and Emissivity Separation Algorithm With Ground Measurements Over a Rice Paddy. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 3061–3069. [CrossRef]
27. Niclòs, R.; Pérez-Planells, L.; Coll, C.; Valiente, J.A.; Valor, E. Evaluation of the S-NPP VIIRS land surface temperature product using ground data acquired by an autonomous system at a rice paddy. *ISPRS J. Photogramm. Remote Sens.* **2018**, *135*, 1–12. [CrossRef]
28. Niclòs, R.; Puchades, J.; Coll, C.; Barberà, M.J.; Pérez-Planells, L.; Valiente, J.A.; Sánchez, J.M. Evaluation of Landsat-8 TIRS data recalibrations and land surface temperature split-window algorithms over a homogeneous crop area with different phenological land covers. *ISPRS J. Photogramm. Remote Sens.* **2021**, *174*, 237–253. [CrossRef]
29. Coll, C.; Caselles, V.; Galve, J.; Valor, E.; Niclos, R.; Sanchez, J.; Rivas, R. Ground measurements for the validation of land surface temperatures derived from AATSR and MODIS data. *Remote Sens. Environ.* **2005**, *97*, 288–300. [CrossRef]
30. Coll, C.; Caselles, V.; Valor, E.; Niclos, R.; Sánchez, J.M.; Galve, J.M.; Mira, M. Temperature and emissivity separation from ASTER data for low spectral contrast surfaces. *Remote Sens. Environ.* **2007**, *110*, 162–175. [CrossRef]
31. Niclòs, R.; Valiente, J.A.; Barberà, M.J.; Coll, C. An Autonomous System to Take Angular Thermal-Infrared Measurements for Validating Satellite Products. *Remote Sens.* **2015**, *7*, 15269–15294. [CrossRef]
32. Guillevic, P.; Göttsche, F.; Nickeson, J.; Hulley, G.; Ghent, D.; Yu, Y.; Trigo, I.; Hook, S.; Sobrino, J.A.; Remedios, J.; et al. *Land Surface Temperature Product Validation Best Practice Protocol, Version 1.1.*; National Aeronautics and Space Administration: Washington, DC, USA, 2018; p. 58. [CrossRef]
33. Theocharous, E.; IBarker Snook, I.; Fox, N.P. *2016 Comparison of IR Brightness Temperature Measurements in Support of Satellite Validation Part 1*; Blackbody Laboratory Comparison: Teddington, UK, 2017.
34. Coll, C.; Niclòs, R.; Puchades, J.; García-Santos, V.; Galve, J.M.; Pérez-Planells, L.; Valor, E.; Theocharous, E. Laboratory calibration and field measurement of land surface temperature and emissivity using thermal infrared multiband radiometers. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *78*, 227–239. [CrossRef]
35. Legrand, M.; Pietras, C.; Brogniez, G.; Haeffelin, M.; Abuhassan, N.K.; Sicard, M. A High-Accuracy Multiwavelength Radiometer for In Situ Measurements in the Thermal Infrared. Part I: Characterization of the Instrument. *J. Atmos. Ocean. Technol.* **2000**, *17*, 1203–1214. [CrossRef]
36. Garcia-Santos, V.; Valor, E.; Caselles, V.; Mira, M.; Galve, J.M.; Coll, C. Evaluation of Different Methods to Retrieve the Hemispherical Downwelling Irradiance in the Thermal Infrared Region for Field Measurements. *IEEE Trans. Geosci. Remote Sens.* **2012**, *51*, 2155–2165. [CrossRef]
37. Vanhellemont, Q. Combined land surface emissivity and temperature estimation from Landsat 8 OLI and TIRS. *ISPRS J. Photogramm. Remote Sens.* **2020**, *166*, 390–402. [CrossRef]
38. Cao, B.; Guo, M.; Fan, W.; Xu, X.; Peng, J.; Ren, H.; Du, Y.; Li, H.; Bian, Z.; Hu, T.; et al. A New Directional Canopy Emissivity Model Based on Spectral Invariants. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 6911–6926. [CrossRef]
39. Ren, H.; Yan, G.; Chen, L.; Li, Z. Angular effect of MODIS emissivity products and its application to the split-window algorithm. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 498–507. [CrossRef]

40. Gillespie, A.; Rokugawa, S.; Matsunaga, T.; Cothern, J.; Hook, S.; Kahle, A. A temperature and emissivity separation algorithm for Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) images. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 1113–1126. [CrossRef]
41. Garcia-Santos, V.; Valor, E.; Caselles, V.; Coll, C.; Burgos, M.A. Effect of Soil Moisture on the Angular Variation of Thermal Infrared Emissivity of Inorganic Soils. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 1091–1095. [CrossRef]
42. Rubio, E. Emissivity measurements of several soils and vegetation types in the 8–14, µm Wave band: Analysis of two field methods. *Remote Sens. Environ.* **1997**, *59*, 490–521. [CrossRef]
43. Rubio, E.; Caselles, V.; Coll, C.; Valour, E.; Sospedra, F. Thermal–infrared emissivities of natural surfaces: Improvements on the experimental set-up and new measurements. *Int. J. Remote Sens.* **2003**, *24*, 5379–5390. [CrossRef]
44. Bock, O.; Keil, C.; Richard, E.; Flamant, C.; Bouin, M.-N. Validation of precipitable water from ECMWF model analyses with GPS and radiosonde data during the MAP SOP. *Q. J. R. Meteorol. Soc.* **2005**, *131*, 3013–3036. [CrossRef]
45. Dyroff, C.; Zahn, A.; Christner, E.; Forbes, R.; Tompkins, A.M.; Van Velthoven, P.F.J. Comparison of ECMWF analysis and forecast humidity data with CARIBIC upper troposphere and lower stratosphere observations. *Q. J. R. Meteorol. Soc.* **2015**, *141*, 833–844. [CrossRef]
46. Ovarlez, J.; Van Velthoven, P.; Sachse, G.; Vay, S.; Schlager, H.; Ovarlez, H. Comparison of water vapor measurements from POLINAT 2 with ECMWF analyses in high-humidity conditions. *J. Geophys. Res. Space Phys.* **2000**, *105*, 3737–3744. [CrossRef]
47. Bicheron, P.; Defourny, P.; Brockmann, C.; Schouten, L.; Vancutsem, C.; Huc, M.; Bontemps, S.; Leroy, M.; Achard, F.; Herold, M.; et al. *GLOBCOVER 2009 Products Description and Validation Report*; MEDIAS-France: Touluse, France, 2011. [CrossRef]
48. Jimenez, J.C.; Gomis-Cebolla, J.; Sabrino, J.A.; Soria, G.; Skokovic, D.; Julien, Y.; Garcia-Monteiro, S.; Mattar, C.; Santamaria-Artigas, A.; Pasapera-Gonzales, J.J. Sentinel 2 and 3 for Temperature Monitoring Over the Amazon. *IEEE Int. Geosci. Remote Sens. Sympos.* **2018**, 2–3, 5925–5928. [CrossRef]
49. Yang, J.; Zhou, J.; Göttsche, F.-M.; Long, Z.; Ma, J.; Luo, R. Investigation and validation of algorithms for estimating land surface temperature from Sentinel-3 SLSTR data. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *91*, 102136. [CrossRef]
50. Galve, J.M.; Coll, C.; Caselles, V.; Valor, E. An Atmospheric Radiosounding Database for Generating Land Surface Temperature Algorithms. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 1547–1557. [CrossRef]
51. Coll, C.; Caselles, V.; Sobrino, J.A.; Valor, E. On the atmospheric dependence of the split-window equation for land surface temperature. *Int. J. Remote Sens.* **1994**, *15*, 105–122. [CrossRef]
52. Coll, C.; Caselles, V. A split-window algorithm for land surface temperature from advanced very high resolution radiometer data: Validation and algorithm comparison. *J. Geophys. Res. Space Phys.* **1997**, *102*, 16697–16713. [CrossRef]
53. Donlon, C.; Berruti, B.; Buongiorno, A.; Ferreira, M.-H.; Féménias, P.; Frerick, J.; Goryl, P.; Klein, U.; Laur, H.; Mavrocordatos, C.; et al. The Global Monitoring for Environment and Security (GMES) Sentinel-3 mission. *Remote Sens. Environ.* **2012**, *120*, 37–57. [CrossRef]
54. Ghent, D.; Veal, K.; Trent, T.; Dodd, E.; Sembhi, H.; Remedios, J. A New Approach to Defining Uncertainties for MODIS Land Surface Temperature. *Remote Sens.* **2019**, *11*, 1021. [CrossRef]
55. Berk, A.; Anderson, G.P.; Acharya, P.K.; Shettle, E.P. *MODTRAN5. 2.0.0 User's Manual*; Spectral Science Inc.: Burlingt, MA, USA; Air Force Res. Lab.: Hanscom, MA, USA, 2008.
56. Baldridge, A.M.; Hook, S.J.; Grove, C.I.; Rivera, G. The ASTER spectral library version 2. *Remote Sens. Environ.* **2009**, *113*, 711–715. [CrossRef]
57. Wan, Z.; Dozier, J. A generalized split-window algorithm for retrieving land-surface temperature from space. *IEEE Trans. Geosci. Remote Sens.* **1996**, *34*, 892–905. [CrossRef]
58. Wan, Z. New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product. *Remote Sens. Environ.* **2014**, *140*, 36–45. [CrossRef]
59. Wilrich, P.-T. Robust estimates of the theoretical standard deviation to be used in interlaboratory precision experiments. *Accredi. Qual. Assur.* **2007**, *12*, 231–240. [CrossRef]
60. Guillevic, P.C.; Biard, J.C.; Hulley, G.C.; Privette, J.L.; Hook, S.J.; Olioso, A.; Göttsche, F.M.; Radocinski, R.; Román, M.O.; Yu, Y.; et al. Validation of Land Surface Temperature products derived from the Visible Infrared Imaging Radiometer Suite (VIIRS) using ground-based and heritage satellite measurements. *Remote Sens. Environ.* **2014**, *154*, 19–37. [CrossRef]
61. Martin, M.A.; Ghent, D.; Pires, A.C.; Göttsche, F.-M.; Cermak, J.; Remedios, J.J. Comprehensive In Situ Validation of Five Satellite Land Surface Temperature Data Sets over Multiple Stations and Years. *Remote Sens.* **2019**, *11*, 479. [CrossRef]
62. Krishnan, P.; Meyers, T.P.; Hook, S.J.; Heuer, M.; Senn, D.; Dumas, E.J. Intercomparison of In Situ Sensors for Ground-Based Land Surface Temperature Measurements. *Sensors* **2020**, *20*, 5268. [CrossRef] [PubMed]
63. Duan, S.-B.; Li, Z.-L.; Li, H.; Göttsche, F.-M.; Wu, H.; Zhao, W.; Leng, P.; Zhang, X.; Coll, C. Validation of Collection 6 MODIS land surface temperature product using in situ measurements. *Remote Sens. Environ.* **2019**, *225*, 16–29. [CrossRef]
64. Gerace, A.; Kleynhans, T.; Eon, R.; Montanaro, M. Towards an Operational, Split Window-Derived Surface Temperature Product for the Thermal Infrared Sensors Onboard Landsat 8 and 9. *Remote Sens.* **2020**, *12*, 224. [CrossRef]
65. Dorman, J.L.; Sellers, P.J. A Global Climatology of Albedo, Roughness Length and Stomatal Resistance for Atmospheric General Circulation Models as Represented by the Simple Biosphere Model (SiB). *J. Appl. Meteorol.* **1989**, *28*, 833–855. [CrossRef]
66. Sobrino, J.; García-Monteiro, S.; Julien, Y. Surface Temperature of the Planet Earth from Satellite Data over the Period 2003–2019. *Remote Sens.* **2020**, *12*, 2036. [CrossRef]

ANEXO C

Evaluation of Landsat-8 TIRS data recalibrations and land surface temperature split-window algorithms over a homogeneous crop area with different phenological land covers

ISPRS Journal of Photogrammetry and Remote Sensing 174 (2021) 237–253

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Evaluation of Landsat-8 TIRS data recalibrations and land surface temperature split-window algorithms over a homogeneous crop area with different phenological land covers

Raquel Niclòs [a,*], Jesús Puchades [a], César Coll [a], María J. Barberà [a], Lluís Pérez-Planells [a], José A. Valiente [b], Juan M. Sánchez [c]

[a] Department of Earth Physics and Thermodynamics, Faculty of Physics, University of Valencia, 50, Dr. Moliner, E 46100 Burjassot, Spain
[b] Instituto Universitario Centro de Estudios Ambientales del Mediterráneo–CEAM-UMH, 14 Charles Darwin, E 46980 Paterna, Spain
[c] Regional Development Institute, University of Castilla La Mancha, Campus Universitario s/n, E 02071 Albacete, Spain

ARTICLE INFO

Keywords:
Landsat-8 TIRS
Calibration
LST algorithms
Validation
In-situ measurements
Thermal-infrared

ABSTRACT

Successive re-calibrations were implemented in Landsat-8 TIRS data since launch. This paper evaluates the performances of both: (1) these re-calibrations, up to the last calibration update announced for TIRS data in the next Landsat Collection 2; and (2) single-channel (SC) corrections and split-window (SW) algorithms to retrieve land surface temperature (LST) from TIRS data. A robust and accurate multi-year (2014–2019) set of reference ground data were used, which included thermal infrared (TIR) radiance measurements taken along transects in a uniform and thermally homogeneous rice paddy area, but also emissivity measurements for the different ground covers at the site through the year. The calibration results showed significant biases at the site for data after the 2014 reprocessing, but negligible biases and root-mean-square differences (RMSDs) <1.5 K were obtained when using the current TIRS data in Collection 1 (i.e., data after the 2017 reprocessing). The last announced calibration update mainly introduced differences in biases, improving slightly the results for band 10 and presenting a calibration response difference between bands. The SC corrections showed negligible LST biases for both bands and the lowest RMSDs (<1.6 K) when using the band 10 data in the current Collection 1, and the bias disappeared for this band after applying the calibration update. Three of the seventeen different SW equations evaluated in the paper showed negligible biases and LST RMSDs lower than or equal to 0.8 K. These three SW algorithms are mainly recommended for users of the current TIRS data in Collection 1; one of them being that proposed to generate a SW LST product in the future Collection 3. Finally, bias differences around 1.4 K were shown in the results of the SW algorithms after applying the calibration update announced for Collection 2.

1. Introduction

Land surface temperature (LST) is an Essential Climate Variable (ECV) for the monitoring of the Earth climate system (Bojinski et al., 2014; Guillevic et al., 2018), as defined by the Global Climate Observing System (GCOS, GCOS-200) and the Climate Change Initiative (CCI) programme of the European Space Agency (ESA). LST is required for studies in meteorology, climatology and hydrology (Weng, 2009; Ghent et al., 2010; Anderson et al., 2012; Trigo et al., 2015; Mokhtari et al., 2019).

NASA and the United States Geological Survey (USGS) launched the Landsat-8 (L8) on February 11, 2013 to keep on providing Earth surface data at high spatial resolution like the previous missions of the Landsat series. The L8 Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) started the data acquisition in April 2013. The L8 TIRS is the first Landsat sensor with two Thermal Infrared (TIR) bands, 10 (10.60–11.19 μm) and 11 (11.50–12.51 μm), at a nominal spatial resolution of 100 m. These two TIR bands allow the use of split-window (SW) algorithms to obtain LST from atmospheric correction through the differential atmospheric absorption between the two bands.

In November 2013, the L8 Calibration Team warned of the need to reprocess the TIRS data due to discrepancies observed when comparing

* Corresponding author.
E-mail addresses: raquel.niclos@uv.es (R. Niclòs), jesus.puchades@uv.es (J. Puchades), cesar.coll@uv.es (C. Coll), m.jesus.barbera@uv.es (M.J. Barberà), lluis.perez@uv.es (L. Pérez-Planells), valiente_jospar@gva.es (J.A. Valiente), JuanManuel.Sanchez@uclm.es (J.M. Sánchez).

https://doi.org/10.1016/j.isprsjprs.2021.02.005
Received 8 August 2020; Received in revised form 27 January 2021; Accepted 2 February 2021
Available online 4 March 2021
0924-2716/© 2021 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

the L8 TIRS data with equivalent top-of-atmosphere (TOA) data retrieved from in-situ data measurements acquired at reference water sites. They proposed subtracting 0.29 ± 0.12 $Wm^{-2}sr^{-1}\mu m^{-1}$ (2.1 ± 0.8 K at 300 K) and 0.5 ± 0.2 $Wm^{-2}sr^{-1}\mu m^{-1}$ (4.4 ± 1.8 K at 300 K) from radiances (brightness temperatures) measured by TIRS band 10 and 11, respectively. The USGS Earth Resource Observation Science (EROS) processing system implemented the bias error correction on February 3, 2014 (thus hereafter called 2014 reprocessing), but the residual variability for the recalibrated data was still larger than desired for the TIRS bands (Barsi et al., 2014). The research conducted after these first calibration activities indicated that the calibration errors were dependent on the scene due to radiation incursion from outside the Instantaneous Field of View (IFOV) of the TIRS, and thus they were inherent to each TIRS scene. Thus, the calibration based on local bias assessments could be just valid for particular subsets of data. The out-of-field stray light effect was shown to be larger for band 11 than for band 10, which made the use of band 11 data not recommended where absolute calibration was required (Barsi et al., 2014). Additionally, users were advised not to apply SW atmospheric and emissivity correction techniques, given the large uncertainties of the TIRS band 11 values.

Montanaro et al. (2014a; 2015) proposed an algorithm for modeling and correcting the radiance introduced from outside the IFOV to reduce the calibration observed discrepancies to attain the required absolute error limit of 2% for TIRS. The out-of-field radiance produced, additionally to a banding observed in uniform and homogeneous areas, differences between the detected signals and the modeled TOA radiances from target temperatures up to 4–5% in band 10 and up to 8–9% in band 11 (Gerace and Montanaro, 2017). Maps of the out-of-field radiation locations were made for each TIRS detector, which were used in the correction algorithm. It was known as "TIRS-on-TIRS" algorithm, since the correction is estimated with TIRS data within the acquisition swath to avoid using external data (Montanaro et al., 2015). The algorithm implementation in the TIRS data was done in 2017 (hereafter called 2017 reprocessing), as a result of the successive improvements implemented. Gerace and Montanaro (2017) showed evaluation results when using independent scenes of EOS-Terra MODIS as reference data. They showed that the magnitude of the banding artifact was reduced by half on average over the uncorrected product and that the absolute radiometric error was reduced to approximately 0.5% in both TIRS bands on average, i.e., below the 2% requirement. Guo et al. (2020) observed that the relationship between the difference between TIRS TOA brightness temperatures at bands 10 and 11 and in-situ LSTs was closer to the atmospheric differential absorption theory after the stray-light correction. They found this fact implied that 2007-corrected data were more proper to obtain suitable results when using SW algorithms than before the correction.

In December 2019, the L8 Calibration Team announced a TIRS absolute radiometric calibration update in preparation for the upcoming Landsat Collection 2 processing. According to them, once the stray-light correction was in place, it became possible to work on reducing remaining radiometric calibration errors, as those that they observed from vicarious calibration results using instrumented buoys in various water bodies. The new calibration update proposes two different corrections: one from launch to March 2, 2015 (with TIRS instrument operating on side-A electronics); and another one after March 2, 2015 (with TIRS operating on side-B electronics, switched over from nominal side-A electronics for correcting the anomaly observed in TIRS data operation since November 2014 (England, 2019)). The corrections use linear gain and bias adjustments. The notice on the U.S. Geological Survey website (usgs.gov/land-resources/nli/landsat/landsat-8-oli-and-tirs-calibration-notices) provides a table that lists the estimated change in TIRS bands response (in K) as a result of the calibration update for a range of temperatures that is being corrected and for both operational modes of the instrument (sides A and B). It also shows figures of the estimated change both in radiance and temperature per band.

The first objective of the paper is the evaluation of the abovementioned recalibrations implemented in the TIRS data in 2014 and 2017, and also when using the last calibration update proposed in 2019, not yet evaluated with independent reference ground data over land sites. The second objective of the paper is to check different methods for atmospheric and emissivity corrections of TIRS brightness temperatures for LST retrievals. We evaluated single-channel (SC) corrections and five different SW algorithms. Different sets of regression coefficients for these SW algorithms were estimated in previous papers (Du et al., 2015; Guo et al., 2020; Gerace et al., 2020; Jimenez-Muñoz et al., 2014; Meng et al., 2019; Rozenstein et al.,2014; Yu et al., 2014), and thus finally this paper validates a set of seventeen SW equations against ground data.

Previous studies (Guo et al., 2020; Gerace et al., 2020; Meng et al., 2019; Yu et al., 2014) evaluated either the change in TIRS TOA radiances after the stray-light correction in the 2017 reprocessing or the results of specific SW algorithms using Surface Radiation Budget Monitoring (SURFRAD) data as reference. The SURFRAD stations measure upwelling and downwelling irradiances using pyrgeometers in the spectral range from 3.5 to 50 μm, from which estimates of LSTs can be derived. However, according to Guillevic et al. (2012), the SURFRAD instrumental uncertainty alone give a LST uncertainty of around 1 K. Krishnan et al. (2020) reported average differences up to 2 K between LST derived from TIR radiometers and those obtained from pyrgeometer data for grassland sites, depending on the heterogeneity of the site and season, and even larger for an asphalt sample. Gerace et al. (2020) indicated that standard deviations (SDs) of approximately 2 K were observed for the differences between LSTs retrieved from spaceborne TIR instruments and LSTs estimated from SURFRAD data. They concluded that given that this 2 K residual error was consistently observed and that the pyrgeometers are sensitive from 4 to 50 μm, this fact could indicate that reflected solar radiation might be contributing the measurements and that the uncertainty in the required broadband emissivity was potentially a limiting factor for applications requiring high accuracy. Guo et al. (2020) indicated that, besides SURFRAD sites, it was necessary to validate using more robust and homogeneous ground-measured datasets to help clarify the effect of the recalibrations on LST retrieval.

We evaluated all the TIRS recalibrations, and also LSTs retrieved with SC corrections and a complete set of SW algorithms, using accurate radiance data measured by hand-held radiometers with narrow-bands in the TIR, comparable to those of L8 TIRS, instead of pyrgeometer data. These TIR radiance measurements were acquired along transects over a uniform and thermally-homogeneous crop area, with very different phenological land covers, concurrently with the registration of TIRS images in the years 2014 and 2016–2019. Additionally, emissivities for these land covers were measured at the site, and not assumed or estimated from a mean emissivity climatology or threshold methods like in previous papers (Gerace et al. 2020; Yu et al., 2014; Rozenstein et al., 2014; Meng et al., 2019; Guo et al., 2020).

To ensure the high quality and accuracy of the data used for both evaluations and of the results, we also did a comprehensive uncertainty analysis and applied a test based on the R-based validation method (Wan and Li, 2008) to assess the quality of the atmospheric profiles used in the radiative transfer calculations.

Therefore, the main strengths and novelties of this paper can be summarized as: (1) Multi-year evaluation of L8 TIRS calibrations for very different ground covers and atmospheric conditions through the year. (2) Evaluation of the successive re-calibration procedures implemented for L8 TIRS data since launch and up to the present, i.e., including the last calibration update announced for the Landsat Collection 2. (3) Ground LST measurements, used as reference data, acquired along transects by multiband TIR radiometers (instead of pyrgeometers at fix stations) over a thermally homogeneous site at the spatial scales of both ground radiometric measurements and the L8 TIRS pixel. (4) In situ measurements of narrow-band TIR emissivities for the different covers of the site. (5) Comprehensive uncertainty analysis of

the ground-measured and Landsat-derived LSTs used in the comparison, taking into account all possible sources of error involved in the measurements. (6) Evaluation of a complete set of SC and SW methods for atmospheric and emissivity corrections of L8 TIRS brightness temperatures for LST retrievals. (7) Application of a test in order to assess the quality of the atmospheric profiles used in the required radiative transfer calculations.

The following section describes the site and the acquisition of ground data. Section 3 explains the required atmospheric radiative transfer simulations, and the SC corrections. Section 4 describes the evaluated SW techniques and the sets of coefficients. Section 5 includes the results and discussion, together with the uncertainty analysis and the atmospheric profile quality test. The calibration results are shown and discussed in 5.1; the results of the evaluation of the LSTs retrieved using SC corrections are included in 5.2; and the results obtained for the different SW techniques are described and discussed in 5.3. Finally, section 6 summarizes the main conclusions of the paper.

2. Experimental site and ground data

The experimental site is a uniform and thermally-homogeneous rice paddy area located near the city of Valencia, previously used in calibration/validation (CAL/VAL) studies for other satellite sensors, and called the Valencia Test Site (Coll et al., 2005, 2012; Niclòs et al., 2018). This area is covered by L8 scenes acquired at paths 198 and 199 (Fig. 1). The rice paddy has land cover changes along the year due to the phenology of the rice crop, with nearly full vegetation cover (fraction of vegetation cover close to one) in July-September, flooded soils (water surfaces) in December-January and also in May-June, and bare soils in February-April, with a trend from wet to dry soils (Fig. 2). The soil at the site is a silty clay loam soil (textural classification: 14% sand, 50% silt and 35% clay; mineral classification: 19.3% quartz, 3.5% feldspar, 6% filosilicate, 8.9% hematite, 62.3% calcite; organic matter: 4.5% (Niclòs et al., 2018)). The land cover variations, with three totally different but uniform land covers at the same site, make it interesting for calibrations and validations under very different surface conditions and temperatures along the year (Niclòs et al., 2018). In previous papers, we analyzed the thermal homogeneity of the site at different scales to evaluate possible dissimilarities between the footprint of satellite pixels and the areas covered by in-situ measurements (Coll et al., 2005, 2012; Niclòs et al., 2018). These analyses used data acquired over the site (with full vegetation cover) by Landsat-5 TM and EOS ASTER and showed SDs<0.5 K for an area of 16 km^2. They also used TIR radiance measurements along transects covering areas of 250 m by 250 m and showed SDs<0.6 K. The latter was repeated by Niclòs et al. (2018) for the three land cover conditions in the site, and SDs<0.5 K, 0.4 K, 0.9 K were obtained for full vegetation cover, flooded soil, and bare soil, respectively, which proved that the site was adequately thermally homogeneous, even for the bare soil conditions.

In this paper, unlike previous studies (Guo et al., 2020, Gerace et al., 2020, Meng et al., 2019, Yu et al. 2014) that use pyrgeometer data of SURFRAD stations for TIRS data validations over land sites, ground TIR radiances were measured along transects in our experimental area using hand-held Cimel Electronique multiband radiometers model CE-312 (Legrand et al., 2000, Brogniez et al., 2003, Coll et al., 2019). These ground measurements were acquired, concurrently with L8 overpasses, following the methodology detailed described in Coll et al. (2005) for a total of 25 cloud-free days from 2014 to 2019. The CE-312 radiometers measured the surface radiance within a spectral band i, $L_{surf,i}$, which depends on the surface emissivity, ε_i, as follows:

$$L_{surf,i} = \varepsilon_i B_i(T) + \left(1 - \varepsilon_i\right) L_{i\ a,hem}\downarrow \quad (1)$$

where $B_i(T)$ is the channel Planck's function for a temperature T (here T being the LST). $L_{i\ a,hem}\downarrow$ is the atmospheric downwelling irradiance divided by π (Niclòs et al., 2018). $L_{i\ a,hem}\downarrow$ was measured using an Infragold Reflectance Target (IRT-94–100) made by Labsphere (Guillevic et al., 2018), which is a highly diffuse gold panel with a reflectivity close to 0.92 in the 8 – 14 μm region.

Emissivity measurements were taken for the different land covers at the site using the Temperature-Emissivity Separation method (TES) (Gillespie et al., 1998; Hulley & Hook, 2009) applied to the ground data measured by the CE-312 radiometers and also the Box Method (Niclòs et al., 2018). The TES method was used for the bare soils at the site, since this method works better for surfaces with significant emissivity spectral variations (Coll et al., 2005; Niclòs et al., 2018). This method requires $L_{surf,i}$ measured over a specific sample with the different bands of the CE-312 radiometer, the $L_{i\ a,hem}\downarrow$ in each band and an initial ε_i guess to retrieve the maximum-minimum emissivity difference (MMD), i.e., the emissivity spectral contrast, among the CE-312 bands. Then, an empirical relationship between the MMD and the minimum ε_i is used to rescale these band ε_i to the actual values. The empirical relationship was obtained by integrating the spectra contained in the ASTER spectral library (Baldridge et al., 2009) with the response functions of the CE-312 bands. Using the TES method, ε_i values for the bare soil in the site were previously measured for different soil moisture contents, and an increase in ε_i with the soil moisture was observed (Niclòs et al., 2018).

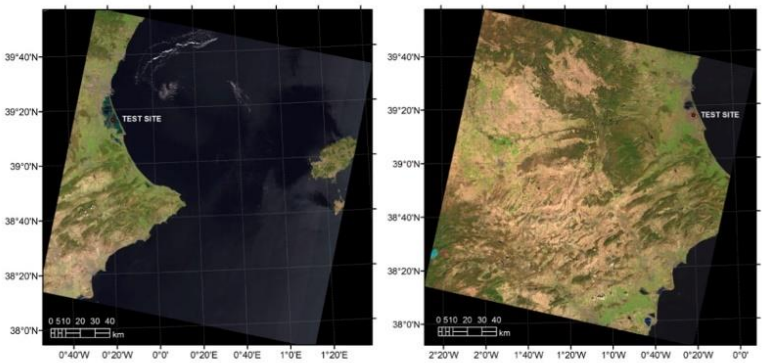

Fig. 1. L8 scenes (false color composites of L8 OLI bands; RGB 654) acquired by the paths 198 and 199 (left and right figures, respectively) over the rice paddy site.

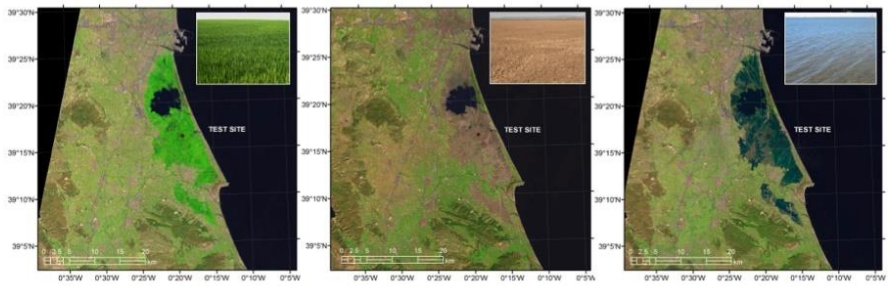

Fig. 2. Views of the different land covers at the site: full vegetation cover (10 July 2018; left), bare soil (27 March 2018; middle), and flooded soil (2 June 2016; right). False color composites of L8 OLI (RGB 654).

The Box Method was used for the full vegetation cover, with a near-grey body spectrum (Coll et al., 2005; Niclòs et al., 2018). This method uses a thermally insulated box with inner walls made of highly reflecting polished aluminum and two interchangeable top lids. One lid is made of the same aluminum (ideally with $\varepsilon = 0$) and the other lid is made of highly emitting material (corrugated aluminum painted in Parson's black, ideally $\varepsilon = 1$). The latter can be heated to about 330 K and has a thermostat. Both lids have a small aperture for the radiometer to observe the radiance coming from the bottom of the box, which can be open to the sample or closed with another aluminum lid without aperture. The sample ε_i measurement requires a series of radiance measurements for the system of sample and box lids in different configurations: 1) aluminum top lid – sample at bottom, 2) hot top lid – sample at bottom, 3) hot top lid – aluminum lid at bottom, and 4) aluminum top lid – aluminum lid at bottom. Further details on this procedure and the correction required for a not ideal box are given in Rubio et al. (2003).

For water surfaces, ε_i values were retrieved with eq. (1) from CE-312 $L_{surf,i}$ measurements and T values measured by contact probes (Niclòs et al., 2014).

The ε_i values obtained from the measurements for the different land cover conditions at the site were around 0.990, 0.985, and 0.970 for the CE-312 band centered in 10.9 μm for flooded soils (water), full vegetation cover, and dry bare soils, respectively. Uncertainties ranging from 0.004 to 0.006 were obtained for the ε_i values measured over flooded soils and full vegetation cover, and from 0.004 to 0.011 for ε_i over bare soils. The uncertainty values were in agreement with those obtained for the TES method over other surfaces in Coll et al. (2019). These emissivity measurements acquired in field conditions were used to correct the ground-measured TIR radiances.

Finally, the uncertainties in ground LSTs were estimated taking into account three error sources added together in quadrature (Coll et al. 2005, 2019). The error sources are: (i) the calibration uncertainty of the ground radiometers, estimated from 0.12 K to 0.15 K in Coll et al. (2019); (ii) the spatial variability within the test site taken as the SD of all the LST measured in the transects within five minutes of each overpass (lower than the maximum values mentioned above for the different covers); and (iii) the uncertainty resulting from the propagation of the uncertainties in the input variables used to retrieve the LSTs through eq. (1); the emissivity one producing the major effect in terms of LST (from 0.4 K to 0.6 K) (Coll et al., 2019). We obtained LSTs from 282 K to 316 K for the whole period, with uncertainties ranging from 0.6 to 0.9 K (with an average LST uncertainty of 0.67 ± 0.08 K).

3. Atmospheric radiative transfer and single-channel correction

The evaluation of the successive recalibrations consisted of comparing L8 TIRS TOA radiances / brightness temperatures with those simulated from ground LSTs, and the evaluation of the LSTs retrieved from L8 TIRS data consisted of comparing them with ground LSTs, as summarized in the evaluation flow chart shown in Fig. 3.

To obtain TOA radiances, L_i, from ground-measured LSTs, or inversely LSTs from L_i, the radiative transfer equation (RTE) is used. The RTE is written as follows:

$$L_i = \tau_i L_{surf,i} + L_{i\,a}\uparrow \quad (2)$$

where $L_{surf,i}$ is dependent on T and ε_i as described in eq. (1), and τ_i and $L_{i\,a}\uparrow$ are the atmospheric transmittance and upwelling radiance, respectively. A single-channel (SC) correction can be applied to retrieve LSTs from L_i inverting eq. (2) for $L_{surf,i}$ and then eq. (1) for $B_i(T)$. The error introduced by using eqs. (1) - (2) with band magnitudes instead of the spectral integration in terms of radiance was evaluated in previous papers (Coll et al., 2012), and errors lower than 0.1 K were shown for satellite TIR bandwidths lower than 1 μm. TOA brightness temperatures, T_i, are obtained from $L_i = B_i(T_i)$ in eq. (2).

In addition to ε_i values (measured for each land cover at the site as described in section 2), the atmospheric τ_i, $L_{i\,a}\uparrow$ and $L_{i\,a,hem}\downarrow$ variables should be estimated for each atmospheric condition. The thermal window band (10–12.5 μm) is affected by absorption and emission from atmospheric water vapor, whose effect in terms of T_i (and also in LST retrieval from them) depends on the vertical profiles of air humidity and temperature (Coll et al., 2012). Spatial and temporal interpolations of the National Centers for Environmental Prediction (NCEP) atmospheric profiles sampled on a 1° x 1° grid and generated every six hours were used as input data to a radiative transfer code, the MODerate resolution atmospheric TRANsmission (MODTRAN) 5 (Berk et al., 2006; Guillevic et al., 2018), to simulate the required τ_i, $L_{i\,a}\uparrow$ and $L_{i\,a,hem}\downarrow$.

A sensitivity analysis was performed to assess the uncertainty in the LSTs retrieved with the SC method, i.e., inverting eq. (1)-(2). We have followed the procedure proposed by Ghent et al. (2019) to define the different uncertainties affecting the satellite LST retrieval. We have considered four uncertainty sources. (i) Uncertainties in radiative transfer simulations including possible errors in the atmospheric profiles and the radiative transfer model used (Hook et al., 2004; Coll et al., 2012), obtaining values of 0.6 K and 1.1 K for TIRS bands 10 and 11, respectively for variations of 10% in relative humidity and 1 K in air temperature at each atmospheric profile level. (ii) Uncertainties in the measured emissivity values, that propagate through the inverse of eq. (1). (iii) Uncertainties in TOA T_i that contribute to the LST uncertainty by propagation in eq. (2). The T_i uncertainty was taken as the noise-equivalent-error (NEΔT) specified for the TIRS bands, i.e., 0.4 K at 300 K for both bands (Irons et al., 2012; Meng et al., 2019). Although the actual TIRS performance was shown to be closer to 0.05 K by Montanaro et al. (2014b), this value of 0.4 K was considered for the sensitivity analysis to be consistent with those in previous studies (Jimenez-Muñoz

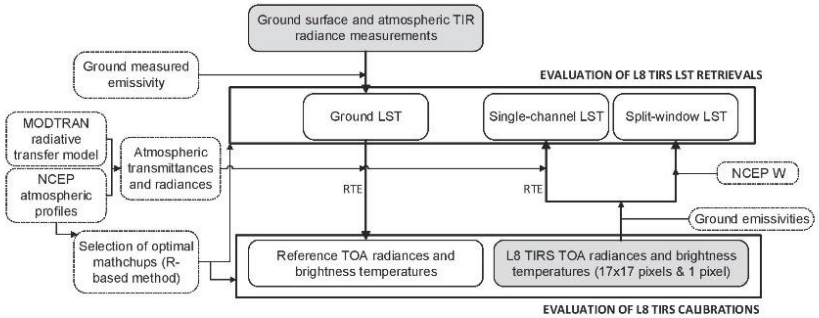

Fig. 3. Flow chart of the procedure followed for evaluating both the successive L8 TIRS recalibrations and the LSTs retrieved from L8 TIRS data.

et al., 2014; Du et al., 2015; Meng et al., 2019). For the evaluations, we considered the average LST value for an array of 17 × 17 TIRS pixels resampled to 30-m centered on the test site, that corresponds to around 500 × 500 m^2 or 5 × 5 native 100-m pixels. The array size was selected as a compromise between having sufficient number of pixels and preserving the thermal homogeneity of the selected area. Since the NEΔT is purely random, its uncertainty contribution to the averaged LST must by divided by $\sqrt{N}$ (Ghent et al., 2019), with N = 17 × 17 being the number of samples (SD of the mean). Note that error sources (i) and (ii) are not random but locally correlated, and thus the uncertainties are not divided by $\sqrt{N}$ (Ghent et al., 2019). (iv) Finally, a fourth error contribution was taken as the SD of the calculated LSTs over the 17 × 17 pixel array to account for the LST spatial variability. The four error contributions were added together in quadrature to obtain the total SC LST uncertainty, which ranged from 0.6 to 0.8 K for band 10, and from 1.2 to 1.3 K for band 11, and with average uncertainties of 0.69 ± 0.05 K and 1.24 ± 0.03 K for bands 10 and 11, respectively.

Guo et al. (2020) used the SC algorithm proposed by Jimenez-Muñoz et al. (2009; 2014) to retrieve LSTs from TOA L_i and T_i in L8 TIRS band 10 and 11 and evaluate the stray-light correction effect in terms of LSTs. This SC algorithm uses parametric atmospheric terms as function of the total water vapor content in the atmosphere (W), which implies a simplification for users and has been used for multiple applications (e.g., for estimating evapotranspiration using L8 data (Mokhtari et al., 2019; Olivera-Guerra et al., 2017)). However, it involves higher uncertainties, which are not proper for calibration studies. Yu et al. (2014) compared the results obtained with the RTE (i.e., inverting eq (1)-(2)) and those when using this SC algorithm, concluding that the LSTs estimated from the former was more accurate than the latter. Jiménez-Muñoz et al. (2014) quantified the effect of the uncertainty in W, as one of the main input variables in their SC algorithm (together with emissivity), in terms of the retrieved LSTs and estimated differences up to 1 K for band 10, which implies much larger LST uncertainties than those above estimated for the RTE, since LST uncertainties are due to both input data uncertainties and those from the SC algorithm regression. They also tested the algorithm results for band 10 with an independent simulated database and obtained a RMSD of 2 K (1.5 K for W < 3 cm).

4. Split-window algorithms for L8 TIRS

Several split-window (SW) algorithms were proposed or adjusted for the L8 TIRS spectral bands in the literature (Jimenez-Muñoz et al., 2014; Yu et al. 2014; Rozenstein et al. 2014; Du et al., 2015; Meng et al. 2019; Guo et al. 2020; Gerace et al. 2020). In this section, we describe five SW algorithms with different estimates of the algorithm regression coefficients, leading a total of seventeen different SW equations, which will be later validated.

Jimenez-Muñoz et al. (2014) followed the expression (hereafter referred to as JM):

$$T = T_{10} + c_0 + c_1(T_{10} - T_{11}) + c_2(T_{10} - T_{11})^2 + (c_3 + c_4 W)(1 - \varepsilon) + (c_5 + c_6 W)\Delta\varepsilon \quad (3)$$

where ε and Δε are the average and the difference of ε_i for the TIRS bands 10 and 11, respectively, and T, T_{10} and T_{11} are the LST and TOA T_i for the TIRS bands. W is again the total water vapor content in the atmosphere. The c_j coefficients were obtained by regression of the synthetic TOA T_i simulated with the MODTRAN 4 radiative transfer code (Beck et al., 1999) using the Global Atmospheric Profiles from Reanalysis Information (GAPRI). The GAPRI database is a compilation of selected atmospheric profiles at a global scale derived from ERA-Interim reanalysis data during 2011 (Mattar et al., 2015). Similarly, Meng et al. (2019) estimated the c_j coefficients in eq. (3) using the MODTRAN 5 code and the GAPRI database. In this case, the coefficients were estimated per W subranges, and also globally for any W between 0 cm and 7 cm. Table 1 shows the c_j in eq. (3) obtained by Jimenez-Muñoz et al. (2014) (JM) and Meng et al. (2019) (JM_M) for the subranges of W that cover the actual atmospheric conditions in our site. The algorithm regression uncertainties, δ_{alg}, provided by the corresponding papers and defined as standard errors of estimate (Jimenez-Muñoz et al., 2014) or root-mean-square errors of the predicted T (Meng et al., 2019), are also shown in the last column of Table 1. Thus, δ_{alg} is the uncertainty obtained with algorithm regression, which minimizes the sum of squared deviations of predictions from actual scores on the simulation database, and takes into account the variance of the residuals.

Meng et al. (2019) also derived the coefficients in other two SW algorithms for the TIRS bands using MODTRAN 5 and the GAPRI database. They calculated the c_j in the generalized SW algorithm designed by Wan and Dozier (1996) and refined by Wan (2014) by adding an additional term related to the quadratic difference between both TOA T_i. This algorithm is given as follows:

$$T = c_0 + \left(c_1 + c_2\frac{1-\varepsilon}{\varepsilon} + c_3\frac{\Delta\varepsilon}{\varepsilon^2}\right)\frac{T_{10} + T_{11}}{2} + \left(c_4 + c_5\frac{1-\varepsilon}{\varepsilon} + c_6\frac{\Delta\varepsilon}{\varepsilon^2}\right)\frac{T_{10} - T_{11}}{2} + c_7(T_{10} - T_{11})^2 \quad (4)$$

The algorithm given in eq. (4) will be hereafter referred to as W&D, and W&D_M when the c_j are those derived by Meng et al. (2019) for TIRS. Du et al. (2015) previously also derived the c_j in eq. (4) for TIRS (hereafter W&D_D) using MODTRAN 5 and the Thermodynamic Initial

Table 1
Coefficients in the SW algorithm given in eq. (3) provided by Jimenez-Muñoz et al. (2014) (JM) and Meng et al. (2019) (JM_M) for different W subranges.

Reference	W (cm)	c_0	c_1	c_2	c_3	c_4	c_5	c_6	δ_{alg} (K)
Jimenez-Muñoz et al. (2014) **(JM)**	0.0–6.0	−0.268	1.378	0.183	54.30	−2.238	−129.20	16.40	0.6
Meng et al. (2019) **(JM_M)**	0.0–2.5	−0.39	2.116	−0.045	64.386	−3.7	−147.522	21.065	0.431
	2.0–3.5	−1.631	2.681	−0.054	67.827	−3.213	−204.953	41.441	0.503
	0.0–7.0	−0.717	1.988	0.121	70.148	−7.006	−143.246	19.247	0.72

Guess Retrieval (TIGR) atmospheric profile database. Guo et al. (2020) refined the coefficients as Zheng et al. (2019) did for Sentinel-3A SLSTR, dividing the simulation data into temperature subranges. In this case, data were divided into several subranges in terms of T_{10}, additionally to the W subranges. This version will be referred to as W&D_D_W_T. They also increased the LST minus bottom air temperature maximum difference in the simulations from 20 K (in Du et al. (2015)) to 35 K, considering that it was more probable for barren or desert surfaces. Gerace et al. (2020) derived the c_j in eq. (4) again (hereafter as W&D_G). Their purpose was proposing a SW algorithm for a future operational SW LST product generated from L8 and L9 TIRS image data in the Collection 3 processing. They used 1393 TIGR atmospheric profiles (after filtering those where relative humidity were higher than 90%), 7 surface temperatures for each of them, 113 emissivity values obtained from the MODIS UCSB database (not using the man-made materials, unlike Du et al. (2015)), and MODTRAN simulations. Table 2 includes the coefficients c_j in eq. (4) obtained by Du et al. (2015), Meng et al. (2019) and Guo et al. (2020) for the different subranges of W content, and temperatures for the latter, that cover the actual conditions in the validation site. The constant c_j derived by Gerace et al. (2020) are also included in Table 2. The corresponding δ_{alg} values (K), defined as the root-mean-square errors in these papers, are also shown.

Meng et al. (2019) also derived the coefficients in the NOAA JPSS Enterprise algorithm for TIRS and W subranges (hereafter Ent_W) (Guillevic et al., 2018; Yu et al., 2017):

$$T = c_0 + c_1 T_{10} + c_2 (T_{10} - T_{11}) + c_3 \varepsilon + c_4 \varepsilon (T_{10} - T_{11}) + c_5 \Delta\varepsilon \quad (5)$$

Table 3 includes the c_j in eq. (5), and the δ_{alg} values, obtained for the subranges of W that cover the atmospheric conditions in the validation site.

Qin et al. (2001) presented a SW technique for the AVHRR dependent on what they defined as two essential variables: ε_i and τ_i. Yu et al. (2014) proposed an algorithm for TIRS based on that of Qin et al. (2001). It relied on the determination of a parameter p_i (with dimensions of temperature, in K) for the TIRS spectral bands using linear regression equations with the TIRS TOA T_i:

$$p_i = a_i + b_i \cdot T_i \quad (6)$$

where a_i and b_i are the regression coefficients. The following equations (7)-(13) shows the Yu et al. (2014) algorithm (hereafter referred to as Yu):

$$T = T_{10} + B_1 \cdot (T_{10} - T_{11}) + B_0 \quad (7)$$

$$B_0 = \frac{C_{11} \cdot (1 - A_{10} \cdot C_{10}) \cdot p_{10} - C_{10} \cdot (1 - A_{11} - C_{11}) \cdot p_{11}}{C_{11} \cdot A_{10} - C_{10} \cdot A_{11}} \quad (8)$$

$$B_1 = \frac{C_{10}}{C_{11} \cdot A_{10} - C_{10} \cdot A_{11}} \quad (9)$$

$$A_{10} = \varepsilon_{10} \cdot \tau_{10} \quad (10)$$

$$A_{11} = \varepsilon_{11} \cdot \tau_{11} \quad (11)$$

$$C_{10} = (1 - \tau_{10}) \cdot (1 + (1 - \varepsilon_{10}) \cdot \tau_{10}) \quad (12)$$

$$C_{11} = (1 - \tau_{11}) \cdot (1 + (1 - \varepsilon_{11}) \cdot \tau_{11}) \quad (13)$$

Term B_0 in eq. (8) is explicitly dependent on p_i (obtained with the regression coefficients a_i and b_i given in Table 4). Yu et al. (2014) also provided quadratic regressions between τ_i and W for two standard atmospheres: a mid-latitude summer profile and a 1976 US Standard profile.

Rozenstein et al. (2014) proposed another SW algorithm based on that of Qin et al. (2001) for TIRS bands. Table 4 shows the regression coefficients in eq. (6) provided by Rozenstein et al. (2014) for the ranges of temperature in the study database. However, the algorithm proposed by Rozenstein et al. (2014) do not use the p_i parameter, but just the a_i and b_i coefficients. The following equations (14)-(23) provide the Rozenstein et al. (2014) algorithm (hereafter referred to as Ro):

$$T = A_0 + A_1 \cdot T_{10} - A_2 \cdot T_{11} \quad (14)$$

$$A_0 = E_1 \cdot b_{10} - E_2 \cdot b_{11} \quad (15)$$

$$A_1 = 1 + A + E_1 \cdot a_{10} \quad (16)$$

$$A_2 = A + E_2 \cdot a_{11} \quad (17)$$

$$A = D_{10} / E_0 \quad (18)$$

$$E_1 = D_{11} \cdot (1 - C_{10} - D_{10}) / E_0 \quad (19)$$

$$E_2 = D_{10} \cdot (1 - C_{11} - D_{11}) / E_0 \quad (20)$$

$$E_0 = D_{11} \cdot C_{10} - D_{10} \cdot C_{11} \quad (21)$$

$$C_i = \varepsilon_i \cdot \tau_i(\theta) \quad (22)$$

Table 2
Coefficients in the SW algorithm given in eq. (4) provided by Du et al. (2015), Meng et al. (2019), Guo et al. (2020) and Gerace et al. (2020).

Reference	W (cm)	T_{10} (K)	c_0	c_1	c_2	c_3	c_4	c_5	c_6	c_7	δ_{alg} (K)
Du et al. (2015) **(W&D_D)**	0.0–2.5	–	−2.78009	1.01408	0.15833	−0.34991	4.04487	3.55414	−8.88394	0.09152	0.34
	2.0–3.5	–	11.00824	0.95995	0.17243	−0.28852	7.11492	0.42684	−6.62025	−0.06381	0.6
	0.0–6.3	–	−0.41165	1.00522	0.14543	−0.27297	4.06655	−6.92512	−18.27461	0.24468	0.87
Meng et al. (2019) **(W&D_M)**	0.0–2.5	–	−1.56	1.007	0.162	−0.288	3.179	6.864	−11.209	0.165	0.44
	2.0–3.5	–	−0.099	0.998	0.148	−0.252	5.236	5.488	−5.455	0.02	0.57
	0.0–7.0	–	−2.64	1.012	0.142	−0.201	2.844	−0.569	−7.6	0.263	0.844
Guo et al. (2020) **(W&D_D_W_T)**	0.0–2.5	270–300	1.6214	0.9968	0.1739	−0.3965	4.3444	5.6164	12.8573	−0.1175	0.30
	0.0–2.5	300–330	7.3937	0.9788	0.1917	−0.3384	3.0247	3.2533	−14.4977	0.1291	0.30
	2.0–3.5	<300	24.913	0.911	0.174	−0.299	6.351	3.92	−5.582	−0.0640	0.54
	2.0–3.5	≥300	27.467	0.904	0.187	−0.349	5.675	2.842	−7.853	0.023	0.58
Gerace et al. (2020) **((W&D_G)**	–	–	2.2925	0.9929	0.1545	−0.3122	3.7186	0.3502	−3.5889	0.1825	0.73

Table 3
Coefficients in the SW algorithm given in eq. (5) provided by Meng et al. (2019) for different W subranges (Ent_M).

Reference	W (cm)	c_0	c_1	c_2	c_3	c_4	c_5	δ_{bias} (K)
Meng et al. (2019) **(Ent_M)**	0.0–2.5	54.95	1.01	1.557	−57.805	0.147	−103.52	0.481
	2.0–3.5	50.035	1.006	5.377	−52.801	−3.16	−87.906	0.589
	0.0–7.0	67.297	0.985	−6.916	−63.855	9.548	−90.919	1.075

Table 4
Regression coefficients in eq. (6) to obtain the parameter p_i according to Yu et al. (2014) and Rozenstein et al. (2014), per ranges of temperature.

Reference	T_i range (°C)	a_{10} (K)	b_{10}	a_{11} (K)	b_{11}
Yu et al. (2014)	−10–20	−55.58	0.4087	−59.85	0.4442
	20–50	−66.61	0.4464	−71.23	0.4831
Rozenstein et al. (2014)	10–40	−62.8065	0.4338	−67.1728	0.4694

$$D_i = [1 - \tau_i(\theta)] \cdot [1 + (1 - \varepsilon_i) \cdot \tau_i(\theta)] \quad (23)$$

The negative sign in eq. (15) was considered following the correction published by the authors. Rozenstein et al. (2014) also provided linear regressions between τ_i and W for two standard atmospheres: a mid-latitude summer profile and a 1976 US Standard profile, and recommend general users to use the regression coefficients for the second one.

Estimates of W are required for the algorithms and they were obtained from the four nearest NCEP profiles by interpolating them spatially for the site and temporally for each L8 TIRS scene acquisition time. The ε_i measured for each land cover condition at the site (described in section 2) were used as input values in the SW algorithms. Previous studies used pixel ε_i estimates based mainly on the ASTER global emissivity dataset (ASTER-GED) (Hulley et al., 2015; García-Santos, et al., 2018; Gerace et al. 2020), which includes a worldwide mean emissivity climatology for the period 2000–2008, or threshold methods in terms of the Normalized-Difference Vegetation Index (Yu et al., 2014; Rozenstein et al., 2014; Meng et al., 2019; Guo et al., 2020) for validating some of the SW algorithms. Recently, an alternative method for estimating TIRS ε_i using neural networks trained on OLI surface reflectances and the ECOSTRESS spectral library (Meerdink et al., 2019) was proposed by Vanhellemont (2020), although they did not evaluate the method in estimating input ε_i in SW algorithm results. However, when possible, a good knowledge of the study site and in-situ ε_i measurements of representative ground covers that concur with satellite overpasses are desired (Rozenstein et al., 2014), since land ε_i are variable due to seasonal effects, land use variability, and changes in soil moisture content (Vanhellemont, 2020; Hongtao et al., 2019; Koley and Jeganathan, 2020), and they also depend on the site specific surface elements (e.g., vegetation species and soil types).

5. Results and discussion

5.1. Evaluation of TIRS calibrations

This section shows the results of the evaluations of the successive TIRS calibrations, i.e., using data after the 2014 and 2017 reprocessings and with the changes in TIRS bands response introduced by the 2019 calibration update. The evaluation consisted of comparing L8 TIRS TOA L_i and T_i with those retrieved using the RTE (eq. (1)-(2)) from ground T and ε_i (as shown in Fig. 3).

First, and for comparison purposes, the calibration of the TIRS data acquired and downloaded in 2014 before the 2014 reprocessing (i.e., original or pre-correction data) was evaluated. The reference ground LSTs acquired this year covered a significant range (from around 285 K to 314 K) due to land cover changes at the site. The first row in Table 5 shows the statistical results of the TIRS minus ground data differences

Table 5
Statistical results, in terms of bias ± SD (RMSD), of the TIRS minus ground data differences per band in terms of TOA L_i and T_i when using the original TIRS data acquired before reprocessing, and those after the 2014 and 2017 reprocessings. Results are shown both for an array of 17 × 17 pixels (resampled to 30-m) centered on the test site and only 1 pixel at the site.

		L_{10} (W m^{-2} sr^{-1} μm^{-1})	T_{10} (K)	L_{11} (W m^{-2} sr^{-1} μm^{-1})	T_{11} (K)
Original data	array	0.09 ± 0.10 (0.13)	0.7 ± 0.7 (1.0)	0.33 ± 0.13 (0.35)	2.7 ± 1.0 (2.9)
	pixel	0.08 ± 0.10 (0.13)	0.6 ± 0.7 (0.9)	0.32 ± 0.12 (0.34)	2.7 ± 1.0 (2.8)
2014 reprocessed data	array	−0.20 ± 0.10 (0.22)	−1.4 ± 0.7 (1.6)	−0.18 ± 0.13 (0.22)	−1.6 ± 1.1 (2.0)
	pixel	−0.21 ± 0.10 (0.24)	−1.5 ± 0.7 (1.7)	−0.19 ± 0.12 (0.23)	−1.7 ± 1.1 (2.0)
2017 reprocessed data	array	0.02 ± 0.17 (0.17)	0.2 ± 1.2 (1.2)	0.01 ± 0.18 (0.18)	0.1 ± 1.6 (1.6)
	pixel	0.00 ± 0.17 (0.17)	0.1 ± 1.2 (1.2)	0.01 ± 0.18 (0.18)	0.1 ± 1.5 (1.5)

per band in terms of TOA L_i and T_i for the array of 17 × 17 pixels and only 1 pixel over the experimental site. Biases and SD of the differences, and root-mean-square differences (RMSD), are shown. Close results were obtained for 17 × 17 pixels and 1 pixel due to the thermal homogeneity of the experimental site, which was also proven in previous papers as described in Section 2 (Coll et al., 2005; Niclòs et al., 2018). Fig. 4a shows the TOA T_i simulated from ground LSTs against the TIRS TOA T_i (average values for 17 × 17 pixels) for bands 10 and 11. The uncertainties in TIRS TOA L_i and T_i were calculated by adding in quadrature the SD of TIRS data in 17 × 17 pixels and the TIRS NEΔT of 0.4 K for both bands, which was divided by $\sqrt{N}$, N = 17 × 17 (Ghent et al., 2019). The resulting TOA T_i uncertainties ranged between 0.1 and 0.4 K for both bands. Such small uncertainty values, which account mainly for the spatial variability of TIRS T_i, show the thermal homogeneity of the test site at the satellite scale. The uncertainties in the TOA L_i and T_i simulated from ground LSTs were obtained by addition in quadrature of the three uncertainty components resulting from the propagation in eq. (1)-(2) of the input variables. They were: (i) the uncertainty in the ground-measured temperatures, (ii) the emissivity uncertainty for each surface in the experimental site, and (iii) uncertainties in atmospheric radiances and transmittances as obtained from a radiative transfer model, in a similar way as in section 3. According to this, uncertainties in simulated T_i ranged between 0.7 and 0.9 K for band 10 and between 1.0 and 1.2 K for band 11.

As shown in Fig. 4a and Table 5, TOA L_i and T_i originally provided by TIRS overestimated those simulated from ground LSTs at the experimental site, confirming the calibration problem mainly for band 11. However, the observed biases at the rice paddy site (in the first row of Table 5) were much lower than those proposed to be subtracted by the L8 calibration team in 2014 (0.29 ± 0.12 Wm^{-2}sr^{-1}μm^{-1} (2.1 ± 0.8 K) and 0.5 ± 0.2 Wm^{-2}sr^{-1}μm^{-1} (4.4 ± 1.8 K) for bands 10 and 11, respectively).

Fig. 4b and the second row in Table 5 show the evaluation results when using TIRS data after the 2014 reprocessing downloaded for the same acquisition dates in 2014. Since the overestimation of the originally acquired TIRS data was lower than that observed by the L8 Calibration Team using reference data at water sites, the reprocessing applied in 2014 on the TIRS data was not proper for our experimental

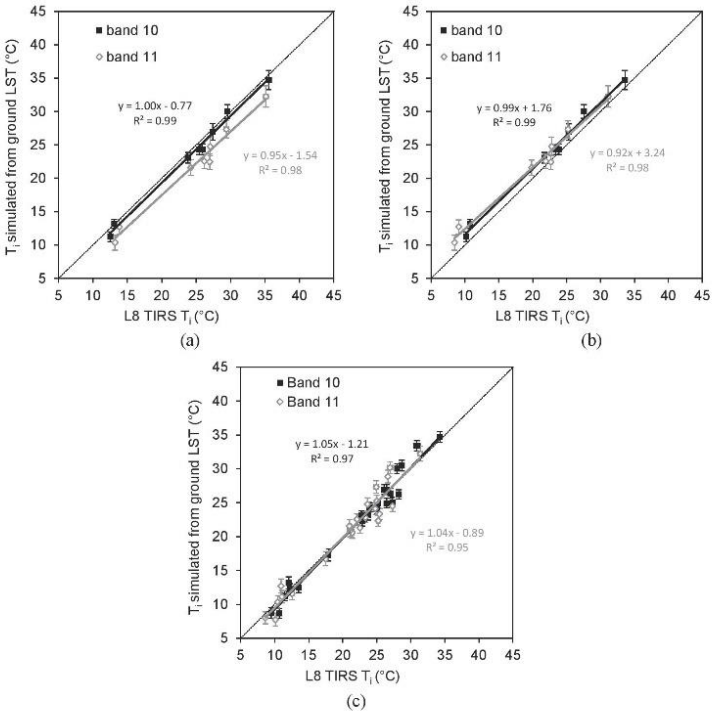

Fig. 4. Comparison of TOA T_i simulated from ground LST and acquired by L8 TIRS to evaluate the successive recalibrations, using TIRS data: a) before reprocessing (original data), b) after the 2014 reprocessing, and c) after the 2017 reprocessing (currently available data in the Landsat Collection 1 L8 database). a) and b) with the available data of 2014, and c) with matchups from 2014 to 2019. The straight lines show the linear regression of the data, with their fitting equations and coefficient of determination (R^2).

site and yielded similar underestimations (of around −1.5 K) for both bands. Results for band 10, used for SC LST retrievals by users, were worse than before this reprocessing.

Fig. 4c and the third row of Table 5 show the results when data after the reprocessing in 2017 were used. The 2014 scenes were again downloaded after this new reprocessing, but additionally we used scenes acquired concurrently with ground data measured along transects in 2016, 2017, 2018 and 2019 to extend the evaluation of the currently available L8 TIRS data. A total of 25 scenes acquired at the site for different seasonal land covers were used for it, covering ground LSTs ranging from 282 K to 316 K. The resulting TIRS TOA T_i uncertainties ranged between 0.1 and 0.5 K for both bands and the complete period, and the uncertainties in TOA T_i simulated from ground LSTs ranged again between 0.7 and 0.9 K for band 10 and between 1.0 and 1.2 K for band 11 for this dataset.

Unlike in the 2014 reprocessing, when just bias subtracting was applied, the reprocessing in 2017 used the TIRS-on-TIRS algorithm that allowed modeling and eliminating the out-of-field radiance, thus using a custom correction for each scene and pixel. Results in Fig. 4c and Table 5 shows that this correction improved the results in the currently available TIRS data. Negligible bias and SDs (RMSDs) of 0.17 W m^{-2} sr^{-1} μm^{-1} and 0.18 W m^{-2} sr^{-1} μm^{-1} (1.2 K and 1.6 K) were obtained for bands 10 and 11, respectively.

Fig. 5a-b show the 2017 calibration results for both TIRS bands by type of surface cover at the experimental site in order to analyze possible land cover dependences. The largest differences were mainly observed for the dry bare soil condition. This fact could be due to the lowest homogeneity observed in previous studies for this surface cover at the site (Niclòs et al., 2018). However, these largest differences corresponds to the highest LSTs (>307 K) and they could be also a consequence of this fact.

Results were also analyzed by scene, since the TIRS-on-TIRS algorithm used in the 2017 reprocessing should not be dependent on scene. Fig. 5c-d shows the results split in L8 paths 198 and 199. Path 198 covers a larger proportion of sea than path 199, as shown in Fig. 1. There are differences in the results between paths 198 and 199 of around 0.8 K in terms of SDs, but not in terms of bias. However, since the number of matchups are quite different for both paths, with just 8 matchups for the 198 and 17 matchups for the 199, the observed SD differences can be just consequence of this fact.

For some cases in Figs. 4-5, the comparison between simulated and satellite T_i shows differences larger than the combined uncertainty of both measurements. The largest source of error in the simulated T_i is the radiative transfer calculations. In some cases, the atmospheric profiles

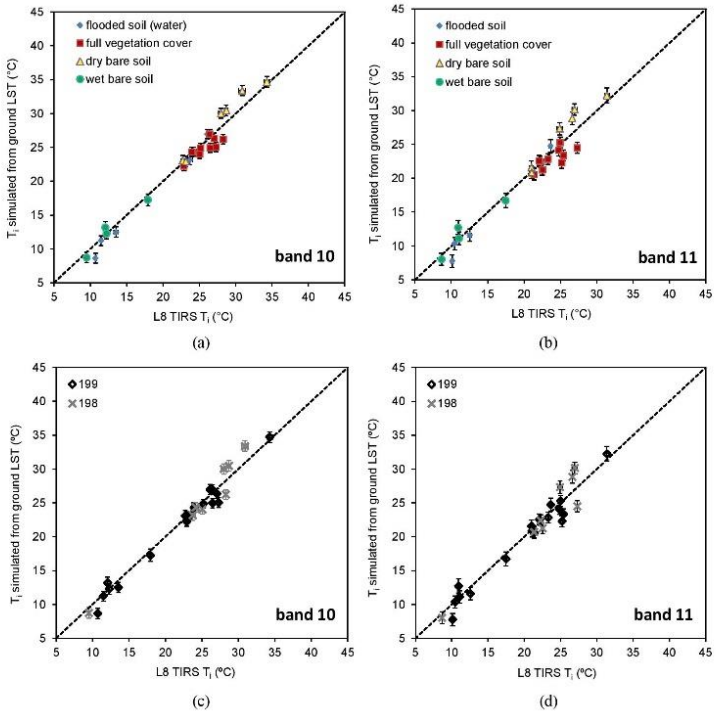

Fig. 5. Comparison of TOA brightness temperatures (for bands 10 and 11) simulated from ground LST and acquired by L8 TIRS after the reprocessing in 2017 split by: (a)-(b) land cover at the site, and (c)-(d) acquisition path (198 and 199).

used may differ from the actual atmospheric profile over the site at the overpass time more than we assumed in the sensitivity analysis. We can check the suitability of the atmospheric profiles by comparing the L8 TIRS observed T_i differences, $(T_{10}\text{-}T_{11})_{obs}$ with the same T_i difference as obtained from the simulation from ground data, $(T_{10}\text{-}T_{11})_{sim}$. Then the difference $\delta(T_{10}\text{-}T_{11}) = (T_{10}\text{-}T_{11})_{obs} - (T_{10}\text{-}T_{11})_{sim}$ can be analyzed; a concept that is related with the R-based validation method (Wan and Li, 2008; Coll et al., 2012). Note that in the two bands, atmospheric absorption is mostly due to the same mechanism (water vapor continuum), which is stronger in band 11, so $T_{10}\text{-}T_{11}$ is usually positive and increases with W. Then if the atmospheric profile is close to the actual profile, the difference $\delta(T_{10}\text{-}T_{11})$ should be close to zero. However, if the profile used is overestimating (underestimating) the atmospheric effect, the value of $(T_{10}\text{-}T_{11})_{sim}$ will be larger (smaller) than $(T_{10}\text{-}T_{11})_{obs}$ and therefore $\delta(T_{10}\text{-}T_{11}) < 0$ (>0). This can be seen in Fig. 6, where the differences between T_{10} simulated from ground data and TIRS observed T_{10} are plotted against $\delta(T_{10}\text{-}T_{11})$ for the whole dataset. The cases where $\delta(T_{10}\text{-}T_{11})$ is small (in absolute value) are associated with small absolute values of temperature differences, the opposite being also true. Thus, we can use the $\delta(T_{10}\text{-}T_{11})$ difference as a check of the atmospheric profile quality and consequently of the accuracy of the T_i simulation. To do that, we can set a threshold on $\delta(T_{10}\text{-}T_{11})$ based on an uncertainty analysis applied to the temperature differences plotted in Fig. 6 using the results of the preceding uncertainty analyses for the individual temperatures

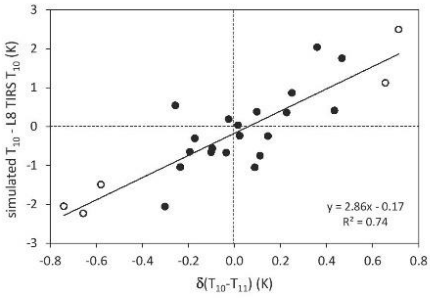

Fig. 6. Differences between T_i simulated from ground data and TIRS observed T_i at band 10 against the differences between the observed and simulated T_i difference for both bands, $\delta(T_{10}\text{-}T_{11})$. The straight line is the linear regression of the data, with the fitting equation and coefficient of determination (R^2) shown in the graph. Open symbols represent cases not satisfying the condition -0.5 K $< \delta(T_{10}\text{-}T_{11}) < 0.5$ K.

involved. Since both temperatures are independent on each other, the combined uncertainty in the difference between simulated and observed T_{10} is the addition in quadrature of the individual uncertainties, yielding an average value of 0.9 K for all cases. For the difference $\delta(T_{10}\text{-}T_{11})$, we have to consider both $(T_{10}\text{-}T_{11})_{obs}$ and $(T_{10}\text{-}T_{11})_{sim}$. For the first, the uncertainty is the addition in quadrature of the uncertainties of the TIRS T_i. For the second, the uncertainties in the simulated T_{10} and T_{11} due to emissivity, NEΔT and spatial variability are added in quadrature. However, the error in the radiative transfer simulations has the same sign in both bands, regardless we are overestimating or underestimating the real atmospheric effect, so they cancel in part when calculating the difference $\delta(T_{10}\text{-}T_{11})$. Adding the five sources of error in quadrature, we obtained an average uncertainty of 0.5 K for $\delta(T_{10}\text{-}T_{11})$ for all the dataset.

Therefore, according to the preceding uncertainty analysis and Fig. 6, we can see that when $\delta(T_{10}\text{-}T_{11})$ is within $\pm$ 0.5 K, the uncertainty in T_i differences will be within $\pm$ 0.9 K (SD of the differences), which suggests taking the condition $-0.5\ K < \delta(T_{b10}\text{-}T_{b11}) < 0.5\ K$ as a threshold to select cases where the atmospheric profile data are confidently accurate. For the dataset studied here, 5 out of 25 cases did not meet the condition (see Fig. 6). Taking out the 5 cases from the comparison, the bias in T_{10} differences was finally 0.08 K (instead of 0.15 K obtained previously), and the SD and RMSD were 1.0 K (instead of 1.2 K; see in Table 5).

Finally, we evaluated the last calibration update proposed in December 2019 in preparation for Collection 2. Fig. 7 shows the results after correcting the current TIRS data with the update proposed in the L8 Calibration Team notice, both for all the matchups together and they divided into side-A and side-B periods. The results showed that the overall biases in T_i differences for bands 10 and 11 were -0.08 K and 0.5 K, respectively (instead of 0.15 K and 0.14 K obtained for the 2017-reprocessed data), without significant changes in terms of SDs and RMSDs. Therefore, for our site, the calibration update slightly improves the results in terms of T_{10}, to which a SC algorithm will be applied for providing a LST product in the next Collection 2 (Malakar et al., 2018; Gerace et al., 2020). The bias difference introduced by the calibration update was -0.2 K at band 10 for both periods with TIRS instrument operating on side-A and side-B electronics. The bias differences introduced at band 11 were 0.6 K for the side-A period and 0.3 K for the side-B period, respectively. Thus, the response change due to calibration update was larger for band 11 than for band 10 at the rice paddy site, increasing the bias difference between bands. Again the RMSDs were slightly reduced (to 0.9 K for band 10) when taking out the five cases mentioned in the previous paragraph.

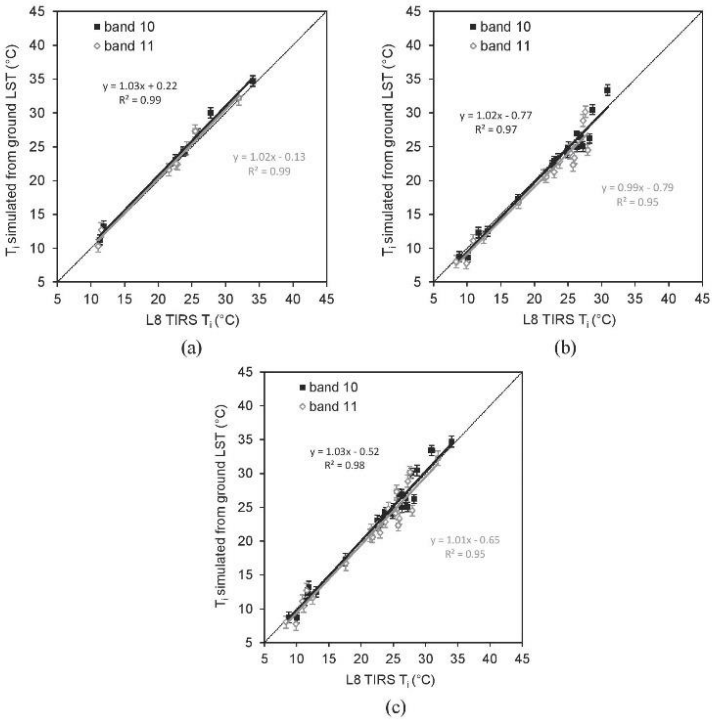

Fig. 7. Comparison of TOA T_{10} and T_{11} simulated from ground LSTs with the L8 TIRS data corrected with the calibration updated proposed in December 2019: a) for the period with TIRS instrument operating on side-A; b) for the period with TIRS instrument operating on side-B; and c) both together.

5.2. Evaluation of single-channel retrievals

Similarly, the comparison was also made in terms of LST to show the results when using SC corrections on the TIRS band 10 and band 11 data (see section 3). The RTE was applied to the TIRS TOA radiances (now inverting eq. (1)-(2)), using simulated atmospheric parameters and suitable ground-measured ε_i values, to retrieve LSTs from data acquired in both spectral bands (see Fig. 3).

Fig. 8 shows the results when using the original data, and data after the 2014 and 2017 reprocessings. Table 6 shows the statistical differences of the TIRS-retrieved LSTs (using band 10 and band 11, respectively) minus ground LSTs.

The results show that using the TIRS data after the reprocessing in 2017 with a SC correction provides LSTs with reduced biases, and equal for both bands, in relation to those obtained with the previous TIRS datasets. SDs are the highest for the 2017 reprocessed data likely due to the largest number of matchups used in this case (i.e., 25 cases in the 2014–2019 period instead of 8 cases in 2014). The LST results for band 10 are shown to have lower RMSDs than those for band 11, which suggest a better performance of SC correction for band 10 with the current data.

Guo et al. (2020) obtained the same conclusion using the SC algorithm proposed by Jimeñez-Muñoz et al. (2014) and data from five SURFRAD sites as reference, but with higher uncertainties. Using TIRS data after the 2017 reprocessing, they obtained biases (RMSDs) of 0.5 K (2.5 K) and 1.6 K (3.6 K) for bands 10 and 11, respectively. The results were different when they excluded a SURFRAD site (DRA) due to heterogeneity issues and high temperatures, with biases (RMSDs) of 0.10 K (2.1 K) and 1.0 K (2.8 K) for bands 10 and 11, respectively. However, Guo et al. (2020) obtained lower biases (but larger RMSDs) for the 2014 reprocessing data when using the five SURFRAD data: −0.17 K (2.7 K) and 1.0 K (4.3 K) for bands 10 and 11, respectively. After excluding the same SURFRAD site from the 2014 reprocessing data analysis, biases (RMSDs) of −0.8 K (2.4 K) and −0.17 K (3.2 K) were obtained for bands 10 and 11, respectively. The bias increase observed by Guo et al. (2020) with the 2017 reprocessing correction when using data from the five SURFRAD sites (and also for band 11 when excluding DRA data) is not in accordance with our SC results in Table 6 nor with the calibration results

Table 6
Statistical results, in terms of bias ± SD (RMSD), of the differences between LSTs obtained with a SC correction of TIRS radiances per band and ground LSTs (17 × 17 pixels) when using: the original TIRS data, those after the 2014 reprocessing, and those after the 2017 reprocessing.

	LST_{10} (K)	LST_{11} (K)
Original data	0.9 ± 1.0 (1.3)	4.0 ± 2.2 (4.6)
2014 reprocessed data	−1.7 ± 0.8 (1.9)	−2.1 ± 1.4 (2.5)
2017 reprocessed data	0.2 ± 1.5 (1.6)	0.3 ± 2.3 (2.3)

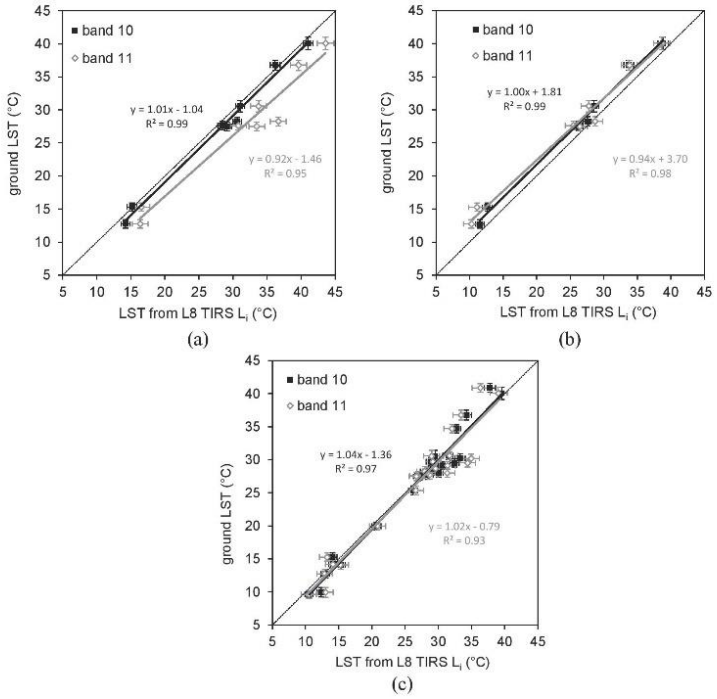

Fig. 8. Comparison of ground LST against LSTs obtained from TIRS radiances (17 × 17 pixels) using SC corrections for bands 10 and 11, respectively, with: a) original TIRS data (2014 scenes), b) data after the reprocessing in 2014 (2014 scenes), and c) data after the reprocessing in 2017 (2014–2019 scenes).

shown in Fig. 4 and Table 5.

As seen in Fig. 8, the difference between ground and satellite LSTs is larger than the combined uncertainty for both measurements in some cases. As for the comparison in terms of T_i (section 5.1), this may be due to the quality of the atmospheric profiles used to calculate the atmospheric radiances and transmittances. Similarly, we can use the difference between the SC LSTs derived from bands 10 and 11, $T_{SC}(10)$-$T_{SC}(11)$, as a test. $T_{SC}(10)$-$T_{SC}(11)$ should be close to zero if the atmospheric profiles used in the simulation are close to the real conditions at the L8 overpass. Oppositely, if $T_{SC}(10)$-$T_{SC}(11) > 0$ (<0) the calculated atmospheric effect appears to underestimate (overestimate) the real atmospheric effect and thus the SC derived LSTs should underestimate (overestimate) the real LST. Fig. 9 shows the differences between the ground and SC LSTs obtained from band 10, T_g-$T_{SC}(10)$, plotted against $T_{SC}(10)$-$T_{SC}(11)$ for the whole dataset. We can see that the cases where the absolute value of $T_{SC}(10)$-$T_{SC}(11)$ is small (large) correspond with small (large) absolute values of T_g-$T_{SC}(10)$. As in section 5.1, we can set a threshold on $T_{SC}(10)$-$T_{SC}(11)$ based on an uncertainty analysis of the temperature differences plotted in Fig. 9 using the results of the preceding uncertainty analyses for the individual temperatures. The combined uncertainty in T_g-$T_{SC}(10)$ is simply the addition in quadrature of the T_g and $T_{SC}(10)$ uncertainties, yielding an average value of 1.0 K for all cases. The uncertainty for $T_{SC}(10)$-$T_{SC}(11)$ was calculated in an analogous way as for $(T_{10}$-$T_{11})_{sim}$ in section 5.1, yielding an average uncertainty of 0.6 K for $T_{SC}(10)$-$T_{SC}(11)$ for all the dataset. Therefore, when $T_{SC}(10)$-$T_{SC}(11)$ is within ± 0.6 K, the uncertainty in T_g-$T_{SC}(10)$ will be within ± 1.0 K, which suggests taking the condition -0.6 K < $T_{SC}(10)$-$T_{SC}(11)$ < 0.6 K as a threshold to select cases where the SC method is confidently accurate. Nine out of 25 cases did not meet the condition (see Fig. 9), and when taking out them from the comparison, the average bias remains equal as in Table 6 (0.2 K) and SD and RMSD were 0.7 K (instead of 1.5 K and 1.6 K).

Finally, we also evaluated the 2019 calibration update in terms of the results when using the SC correction. The bias of the differences between LSTs obtained with a SC correction from TIRS band 10 radiances and ground LSTs was 0.0 K (instead of 0.2 K shown in Table 6), without significant changes in terms of SD and RMSD. This result confirms, for our site and temperature range, the slight improvement introduced with the calibration update for correcting band 10 data with a SC algorithm, as planned for the next Collection 2. Again the RMSD was reduced (to 0.7 K for band 10) when taking out the nine cases mentioned in the previous paragraph.

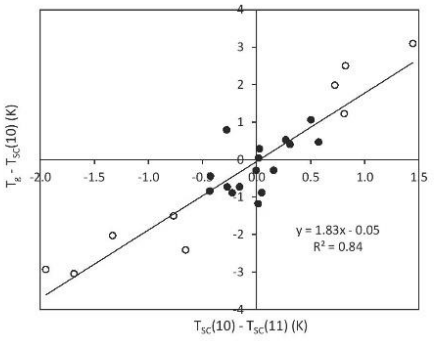

Fig. 9. Differences between ground and SC LST from band 10 against the differences between the SC LST from band 10 and 11. The straight line is the linear regression of the data, with the linear fit equation and coefficient of determination (R^2). Open symbols represent cases not satisfying the condition -0.6 K < $T_{SC}(10)$-$T_{SC}(11)$ < 0.6 K.

5.3. Evaluation of split-window techniques

This section shows the validation results in terms of LSTs for the different SW algorithms proposed for L8 TIRS, previously described in section 4. The 25 matchups of TIRS data (average of 17 × 17 pixels) after the reprocessing of 2017 were used to retrieve LSTs for the validation against the ground LST data acquired in the site (Fig. 3).

Fig. 10 and Table 7 (first to third rows) show the validation results for the SW algorithm described in Jiménez-Muñoz et al. (2014) when using the coefficients proposed in this paper (JM), but also with the coefficients obtained for the same algorithm by Meng et al. (2019). In the latter case, results are shown in Table 7 both when using coefficients per W ranges (JM_M_W) and when using general coefficients for any W between 0.0 cm and 7.0 cm (JM_M). Fig. 10a and 10b shows the results for JM and JM_M_W, respectively. Similar results (not shown in Fig. 10) were obtained when using the general coefficients (JM_M). Error bars in Fig. 10 are the uncertainties estimated for the ground LST data in section 2 (in x-axis) and for the Jiménez-Muñoz et al. (2014) algorithm (in y-axis). For the latter, we considered the following error sources, which were added together in quadrature. (i) δ_{alg} (K) shown in the last column of Table 1. (ii) The propagation in eq. (3) of the uncertainties in band emissivities as obtained for the ground measured data for each surface conditions (section 2). (iii) The propagation in eq. (3) of the uncertainty in W (of 0.5 cm according to Jiménez-Muñoz et al. (2015) and Meng et al. (2019)). Here we also considered the impact of W uncertainty in the selection of SW coefficients, when applicable. (iv) The propagation in eq. (3) of band NEΔTs, which were divided by $\sqrt{N}$ when calculating the average LST in the N = 17 × 17 pixel array (Ghent et al., 2019). Finally, (v) the SD of LST in the 17 × 17 pixel array (spatial variability in LST within the test site). As a result, the average LST uncertainties (and uncertainty range) were 1.0 ± 0.2 K (0.9–1.7 K), 1.0 ± 0.2 K (0.8–1.8 K) and 1.2 ± 0.2 K (1.1–1.9 K) for the JM, JM_M_W and JM_M algorithms, respectively.

Results in Fig. 10 and Table 7 show an accurate performance for the Jimeñez-Muñoz et al. (2014) SW algorithm (JM), with a negligible systematic uncertainty and an accuracy (RMSD) of 0.6 K. The same SD values, but slightly higher biases (0.5 K and 0.6 K), were obtained when using the coefficients provided by Meng et al. (2019) for the algorithm (JM_M_W and JM_M), with RMSDs of 0.8 K. Thus, the Meng et al. (2019) coefficient modification did not introduce improvements for this SW algorithm at our site. Meng et al. (2019) validated the JM_M_W results using data from several SURFRAD sites as reference, and obtained a bias of 2 K and a RMSD of 3 K. Guo et al. (2020) validated the JM algorithm using five SURFRAD sites and obtained a bias of 1.4 K and a RMSD of 2.5 K. They also provided results after removing the SURFRAD DRA data from the analysis and showed a bias of 1.2 K and a RMSD of 2.3 K. Meng et al. (2019) also used the data acquired by a broad-band TIR radiometer (IR120, Campbell Scientific) installed at a fixed tower in a wild grass site with a footprint of 2 m (García-Santos et al., 2018) and obtained biases (RMSDs) of − 0.4 K (1.6 K) for the JM algorithm and of − 0.2 K (1.7 K) for the JM_M_W. The results obtained for our experimental site, using the accurate CE-312 radiometer measurements along transects over different seasonal land covers, show much better results for these SW algorithms, with the best results for the JM version.

Table 7 (fourth to fifth rows) shows the results for the generalized SW algorithm proposed by Wan and Dozier (1996), and refined by Wan (2014) (eq. (4)), with the coefficients estimated for the TIRS by Meng et al. (2019) per W ranges (W&D_M_W) and for any W between 0 and 7 cm (W&D_M). Fig. 11a shows graphically the W&D_M_W results. Table 7 shows also the results for the same algorithm but with the coefficients estimated previously in Du et al. (2015), both per W ranges (W&D_D_W) and for any W between 0.0 and 6.3 cm (W&D_D), using the TIGR atmospheric profile database instead of the GAPRI used by Meng et al.

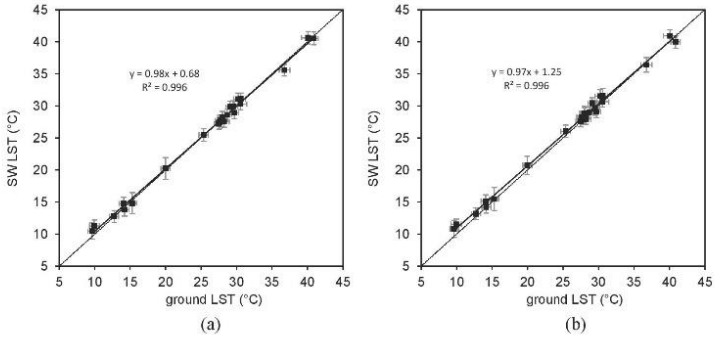

Fig. 10. Comparison of LSTs obtained with the Jiménez-Muñoz et al. (2014) SW algorithm applied to TIRS data (after the 2017 reprocessing) against ground LST: (a) using the original coefficients (JM), and (b) using the coefficients provided by Meng et al. (2019) per W ranges (JM_M_W).

Table 7
Statistical results, in terms of bias ± SD (RMSD), of the differences between LSTs obtained with the SW algorithms applied to TIRS brightness temperatures (with data after the 2017 reprocessing) and ground LSTs. See text for algorithm versions.

SW algorithm	Version	SW LST − ground LST (K)
Jiménez-Muñoz et al. (2014)	JM	0.1 ± 0.6 (0.6)
	JM_M_W	0.5 ± 0.6 (0.8)
	JM_M	0.6 ± 0.6 (0.8)
Wan and Dozier (1996) with the refinement in Wan (2014)	W&D_M_W	0.4 ± 0.7 (0.8)
	W&D_M	0.7 ± 0.7 (1.0)
	W&D_D_W	1.6 ± 0.6 (1.8)
	W&D_D	1.7 ± 0.6 (1.8)
	W&D_D_W_T	0.6 ± 0.9 (1.1)
	W&D_G	0.4 ± 0.6 (0.8)
NOAA JPSS Enterprise algorithm in Meng et al. (2019)	Ent_M_W	0.5 ± 0.7 (0.8)
	Ent_M	0.8 ± 0.6 (1.0)
Yu et al. (2014)	Yu SIM	1.1 ± 0.7 (1.3)
	Yu US	0.8 ± 0.7 (1.1)
	Yu MLS	2.1 ± 1.2 (2.4)
Rozenstein et al. (2014)	Ro SIM	1.1 ± 0.7 (1.3)
	Ro US	0.6 ± 0.8 (1.0)
	Ro MLS	0.4 ± 0.8 (0.9)

(2019). Fig. 11b shows the W&D_D_W performance. Results of the same algorithm but with the W-non-dependent coefficients proposed by Gerace et al. (2020) (W&D_G) and modified by Guo et al. (2020) per subranges of W and temperature (W&D_D_W_T) were also included in Table 7 and in Fig. 11c-d. Error bars in Fig. 11 are the uncertainties estimated for the ground LST data in section 2 (in x-axis) and for the SW algorithms (in y-axis). For the latter, the same five error sources previously described for the Jiménez-Muñoz et al. (2014) algorithm were assumed, which were added together in quadrature. The uncertainties estimated for this SW algorithm were in average 1.0 ± 0.3 K (0.7–1.6 K) and 1.02 ± 0.06 K (0.9–1.2 K) for the W&D_M_W and W&D_M versions, respectively, and 1.1 ± 0.3 K (0.8–1.8 K), 1.17 ± 0.11 K (1.0–1.5 K), 1.1 ± 0.3 K (0.8–1.7 K), and 1.07 ± 0.13 K (0.9–1.5 K) for the W&D_D_W, W&D_D, W&D_D_W_T and W&D_G versions, respectively.

The results in Table 7 for the algorithms based on that of Wan and Dozier (1996) show that algorithm coefficients non-dependent on W and temperature, like those proposed by Gerace et al. (2020) for a future SW LST product in Collection 3, provided the best results for our site. The suitable selection of input data to generate the simulation database used for algorithm regression is essential to obtain accurate LST results. Gerace et al. (2020) did it by filtering TIGR atmospheric profiles and MODIS UCSB surface emissivities and using surface temperatures appropriate for natural materials.

Meng et al. (2019) showed validation results for the W&D_M_W using SURFRAD data as reference. A bias of 1.9 K and a RMSD of 3 K were obtained. However, when using the García-Santos et al. (2018) data as reference, the following biases (RMSDs) were obtained: 1.4 K (2 K) for the W&D_D_W and −0.2 K (1.7 K) for the W&D_M_W. Thus, the overestimation of the W&D_D_W algorithm observed in Table 7 was in accordance with the previously observed results. However, better results were obtained here for the W&D_M_W than in the previous papers, being close to those obtained with the JM algorithm. Guo et al. (2020) evaluated the Du et al. (2015) algorithm with coefficients per W subranges but also per temperature subranges (W&D_D_W_T) using data from five SURFRAD sites as reference and showed a bias of 2 K and a RMSD close to 3 K. Excluding DRA data, both bias and RMSD were reduced in 0.1 K. Much better results were obtained also for the W&D_D_W_T version against our site measurements, likely due to the higher accuracy of our ground data. Gerace et al. (2020) evaluated both the W&D_D and W&D_G algorithms over land sites using data from three SURFRAD sites and one FLUXNET site, i.e., using pyrgeometers. A bias (SD) of −1.3 K (2.4 K) was obtained for the W&D_D algorithm, and of −0.4 (2.1 K) for the W&D_G version when the distance to nearest cloud was >10 km (condition accomplished for the acquisition of ground data along transects at our site). Gerace et al. (2020) also used thirty-six NOAA buoys in the near-shore of the US coastline (with bulk to surface temperature adjustment) and the NASA's Jet Propulsion Laboratory's (JPL) instrumented buoys in Lake Tahoe and Salton Sea as reference over water sites. In this case, a bias (SD) of 0.0 K (1.5 K) was obtained for the W&D_D algorithm, and of 0.2 (0.8 K) for the W&D_G version. On the view of those results, they indicated the abovementioned limitations of using pyrgeometer data as reference for validations (see section 1). The results in Table 7 for the W&D_G algorithm are in accordance with those obtained by Gerace et al. (2020) over water sites, proving the good performance of this algorithm but also the high accuracy of our ground data over a land site.

Additionally, Table 7 shows the results for the Enterprise SW algorithm with coefficients derived for TIRS by Meng et al. (2019), both per W ranges (Ent_M_W) and for any W from 0 to 7 cm (Ent_M). Fig. 12a shows the Ent_M_W results against ground data. The uncertainties estimated for the Enterprise SW algorithm, following the same estimation procedure described for the Jiménez-Muñoz et al. (2014) algorithm, were 1.1 ± 0.3 K (0.8–1.6 K) and 1.31 ± 0.09 K (1.2–1.6 K) for the Ent_M_W and Ent_M versions, respectively.

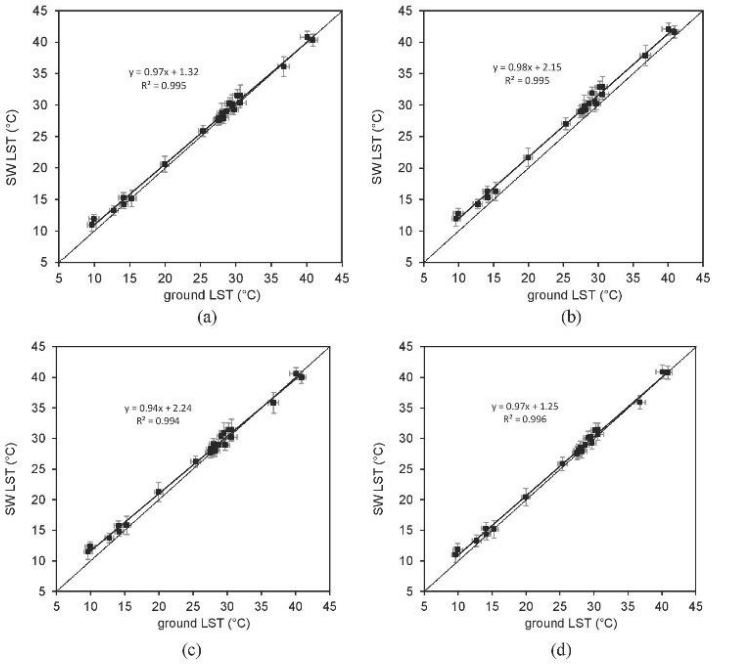

Fig. 11. LST results of the Wan and Dozier (1996) SW algorithm (refined by Wan (2014)) against ground data when using the algorithm coefficients estimated for TIRS by: (a) Meng et al. (2019) per W ranges (W&D_M_W); (b) Du et al. (2015) per W ranges (W&D_D_W); (c) Guo et al. (2020) per W and temperature ranges (W&D_D_W_T); and (d) Gerace et al. (2020) (W&D_G).

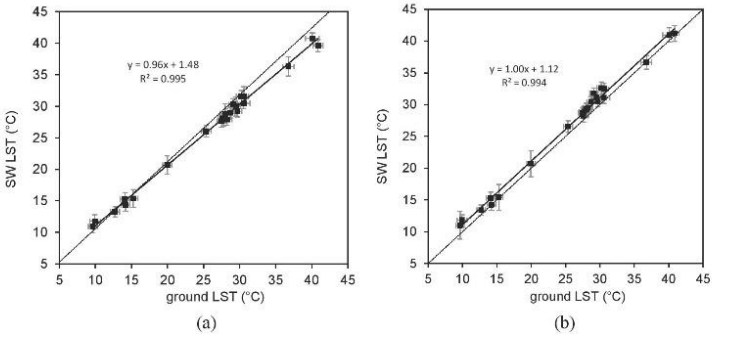

Fig. 12. SW LST results against ground data for the: (a) Enterprise algorithm proposed by Meng et al. (2019) with coefficients per W ranges (Ent_M_W); and (b) the algorithm proposed by Yu et al. (2014) (Yu) using simulated transmittances as input data (Yu SIM). Close results to the latter were provided by the equivalent algorithm proposed by Rozenstein et al. (2014).

Meng et al. (2019) also showed validation results for the Ent_M_W algorithm when using SURFRAD data as reference. A bias (RMSD) of 1.9 K (3 K) was obtained. However, when using the García-Santos et al. (2018) data as reference, they obtained a bias (RMSD) of 0.1 K (1.7 K). Table 7 shows better results for this algorithm at our site, similar to those provided by the W&D_M_W and W&D_G algorithms and the different versions of the algorithm in Jiménez-Muñoz et al. (2014). Although the Enterprise algorithm is the only one of them without a quadratic difference between both TOA T_i, this fact seemed not to have effect for the site conditions.

Finally, Table 7 shows also the results for the SW algorithms proposed by Yu et al. (2014) (Yu) and Rozenstein et al. (2014) (Ro). In these algorithms, atmospheric transmittances for the TIRS bands are used as input data, and they are estimated from W using linear regressions. These regressions were provided for two standard atmospheres: 1976 US Standard (US) and Mid-Latitude Summer (MLS). We used transmittances simulated using MODTRAN 5 as described in section 3 for each matchup (SIM), but also the corresponding US and MLS regressions to evaluate the effect of using them in the LST results. Fig. 12b shows the results for the Yu algorithm with simulated transmittances. Close results were obtained for the equivalent Ro algorithm (not shown graphically). The LST uncertainties estimated for these algorithms (as described before, but with W uncertainty taken into account in the estimate of atmospheric transmittances) were 1.1 ± 0.4 K (0.8–2.2 K) and 0.9 ± 0.3 K (0.7–1.7 K) for Yu and Ro algorithms.

Very close results were obtained for both algorithms (except in the case of MLS transmittance estimates), as it was expected since they were based on the same SW technique proposed by Qin et al. (2001). Yu et al. (2014) showed validation results for their algorithm using data of four SURFRAD stations as reference. They used from 8 to 12 matchups of TIRS and SURFRAD data for each station, and RMSDs from 0.7 K to 1.3 K were obtained, with biases lower than 0.3 K. However, TIRS data before the 2017 reprocessing were used for this validation, and we provide a validation for the algorithm applied on current TIRS data. No validation results were provided in Rozenstein et al. (2014) for their algorithm. Guo et al. (2020) compared the results of the Rozenstein et al. (2014) algorithm with data from five SURFRAD sites. However, the negative sign in eq. (15), which was published as a correction in the algorithm by the own authors, was not considered by Guo et al. (2020). They showed a bias (RMSD) of −0.09 K (2.5 K) when using data from the five SURFRAD sites, and of 0.3 K (2.3 K) when excluding the DRA site. On view of these results, they pointed out this algorithm as one of the best SW algorithms for TIRS data after the 2017 reprocessing correction. If we used a positive sign in eq. (15), like in Guo et al. (2020), bias of −1.8 K, SD of 1.9 K and RMSD of 2.6 K would be obtained, instead of the values shown in Table 7 (in the Ro SIM row), which would lead to very different results for the Ro and Yu algorithms.

In view of our results, the algorithms of Yu et al. (2014) and Rozenstein et al. (2014) provided the largest biases, but with similar SDs to the other SW algorithms when SIM and US transmittance estimates were used (larger SDs were obtained when using MLS transmittance estimates in the Yu algorithm). RMSDs for these two algorithms are larger than those obtained for the JM, W&D_M_W, W&D_G, and Ent_M_W algorithms for the site. Thus, taking into account this fact, the higher complexity in their applications and the required estimates of atmospheric transmittances, they are the least recommended SW algorithms for users of the current TIRS data.

Finally, we also evaluated the 2019 calibration update in terms of the LST results when using these seventeen SW equations. The different response changes in both TIRS bands due to this calibration update, with larger changes for band 11 than for band 10 (i.e., biases in brightness temperature differences were −0.08 K and 0.5 K for bands 10 and 11, respectively), produced mainly significant differences in the biases obtained in the results of the SW algorithms. A quite constant positive difference was observed between the reported biases for the SW algorithms in Table 7 (i.e., applied to TIRS data with the reprocessing of 2017) and those biases obtained for the algorithms when they were applied to TIRS brightness temperatures modified accordingly to the calibration update proposed in 2019. The biases differences were from 1.3 K to 1.9 K for all the algorithms, with an equal difference of 1.4 K for the JM, W&D_M_W, W&D_G and Ent_M_W algorithms, with the best validation results in Table 7. However, we did not observed significant differences in terms of the SDs obtained for the SW algorithms, with SD differences ranging from −0.19 K to 0.18 K for all them. Differences between the SW LST SDs obtained with the 2017 and 2019 calibrations were −0.14 K, −0.16 K, −0.14 K and −0.09 for the JM, W&D_M_W, W&D_G, and Ent_M_W algorithms, respectively. Therefore, the SW algorithms that best results shown when applied on the current TIRS data in Collection 1 will likely provide a negative bias at our site when applied to TIRS data in Collection 2, if the 2019 calibration update is finally incorporated as currently notified into the USGS EROS processing system.

6. Summary and conclusions

This paper shows a robust and accurate evaluation of the successive re-calibrations implemented in L8 TIRS data since launch and up to the last update proposed in preparation for Collection 2, and also an evaluation of the LSTs retrieved from them by using SC corrections and a complete set of SW algorithms. Unlike in previous papers that evaluated some of these calibrations and LST algorithms against station pyrgeometer data, in this case evaluations were performed using a multi-year reference database of ground radiance measurements acquired along transects by accurate TIR hand-held radiometers in a uniform and homogeneous rice paddy area with different surface conditions. The drastic crop changes of land cover (i.e., periods with bare soil, full vegetation cover, and water surface) due to the crop phenology at the experimental area provide data which otherwise would require at least three static sites. Moreover, ground emissivity measurements were acquired for the different covers at the site, and not estimated like in previous studies.

TOA T_i provided by the original TIRS data, before any reprocessing, overestimated in 0.7 K and 2.7 K at bands 10 and 11, respectively, those simulated from the ground LSTs measured in 2014. Although these overestimations were not negligible, they were much lower than the constant values subtracted in the TIRS data reprocessing in 2014, which leaded to underestimations of −1.4 K and −1.6 K for the 2014-recalibrated TOA T_i at bands 10 and 11, respectively (with RMSDs of 1.6 and 2 K). However, our calibration results showed negligible biases for both bands after the 2017 reprocessing (with RMSDs of 1.2 K and 1.6 K for bands 10 and 11, respectively), proving the enhancement introduced by the scene-dependent TIRS-on-TIRS algorithm in the current dataset (Collection 1). The RMSDs were reduced slightly after removing possible atmospherically inaccurate matchups (with a final RMSD of 1.0 K for band 10). The evaluation of the SC correction applied to the 2017-reprocessed data showed negligible biases in terms of the LSTs retrieved from both TIRS bands, with the lowest RMSD for band 10 (of 1.6 K; or 0.7 K after removing the doubtful matchups).

Additionally, seven of the seventeen SW equations applied to the TIRS data after the reprocessing in 2017 provided LST RMSDs lower than 1 K for our experimental site, which are totally within the uncertainty thresholds established for satellite LST retrievals by GCOS. The best versions of these algorithms were: a) the Jimenez-Muñoz et al. (2014) algorithm with the coefficients estimated in that paper (JM) (with RMSD of 0.6 K), b) the Wan and Dozier (1996) algorithm with the Wan (2014) refinement and with constant coefficient estimated by Gerace et al. (2020) (W&D_G) (RMSD of 0.8 K), and c) the Enterprise algorithm with coefficients estimated by Meng et al. (2019) per W subranges (Ent_M_W) (RSMD of 0.8 K). The refined Wan and Dozier (1996) algorithm with coefficients estimated by Meng et al. (2019) per W subranges (W&D_M_W) obtained the same RSMD of 0.8 K, but this algorithm requires W and did not show an improvement with respect to

W&D_G for our site. Therefore, the abovementioned three SW algorithms are recommended for the users of current TIRS data to retrieve LST; being the one proposed by Gerace et al. (2020), for an operational SW product in the future Collection 3, the only one without W as input variable and thus more simple to apply.

Finally, the results when the calibration update proposed in December 2019 was applied to the TIRS data showed biases in TOA T_i of −0.08 K and 0.5 K for bands 10 and 11, respectively, thus introducing a calibration response difference between both bands. The calibration update improved slightly the LSTs obtained for our site with the SC correction applied to band 10 data, for which the RMSD was the lowest (1.5 K; or 0.7 K after removing doubtful matchups). A positive difference was observed between the reported biases for the SW algorithms applied to 2017-reprocessed TIRS data and those biases obtained for the algorithms when using the TIRS dataset after the 2019 update correction. The differences in biases were from 1.3 K to 1.9 K for all the algorithms, with a difference of 1.4 K for the JM, W&D_G and Ent_M_W algorithms, and no significant differences were observed in SDs. These results could mean that the best SW algorithms shown for the current TIRS data in Collection 1 would likely provide a negative bias at our site when applied to TIRS data in Collection 2. In any case, a new evaluation would be advisable when the Collection 2 data will be finally available for users, i.e., after the operational application of the calibration update in the L8 TIRS data processing system.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was funded by the Spanish Ministry of Economy and Competitiveness and the European Regional Development Fund (FEDER) through the project CGL2015-64268-R (MINECO/FEDER, UE). A researcher was hired for it with the support of a grant of the regional program for training of technicians for R&D&i (Youth Guarantee GJIDI-2018-A-142). The validation campaigns in 2014 and 2016 were also funded by the Spanish Ministry of Economy and Competitiveness under the projects CGL2011-30433-C02-02 and CGL2013-46862-C2-1-P.

References

Anderson, M.C., Allen, R.G., Morse, A., Kustas, W.P., 2012. Use of Landsat thermal imagery in monitoring evapotranspiration and managing water resources. Rem. Sens. Environ. 122, 50–65.

Baldridge, A.M., Hook, S.J., Grove, C.I., Rivera, G., 2009. The ASTER spectral library version 2.0. Remote Sens. Environ. 113, 711–715. https://doi.org/10.1016/j.rse.2008.11.007.

Barsi, J.A., Schott, J.R., Hook, S.J., Raqueno, N.G., Markham, B.L., Radocinski, R.G., 2014. Landsat-8 Thermal Infrared Sensor (TIRS) Vicarious Radiometric Calibration. Remote Sens. 6 (11), 11607–11626.

Beck, A., Anderson, G.P., Acharya, P.K., Chetwynd, J.H., Bernstein, L.S., Shettle, E.P., Matthew, M.W., Adler-Golden, S.M., 1999. MODTRAN4 User's Manual. Air Force Res. Lab, Hanscom AFB, MA, USA.

Berk, A., Anderson, G.P., Acharya, P.K., Bernstein, L.S., Muratov, L., Lee, J., Fox, M., Adler-Golden, S.M., Chetwynd, J.H., Hoke, M.L., Lockwood, R.B., Gardner, J.A., Cooley, T.W., Borel, C.C., Lewis, P.E., Shettle, E.P., 2006. MODTRAN5: 2006 Update. Proc. SPIE 6233 (62331F).

Bojinski, S., Verstraete, M., Peterson, T., Richter, C., Simmons, A., Zemp, M., 2014. The Concept of Essential Climate Variables in Support of Climate Research, Applications, and Policy. Bull. Amer. Meteor. Soc. 95 (9), 1431–1443.

Brogniez, G., Pietras, C., Legrand, M., Dubuisson, P., Haeffelin, M., 2003. A high accuracy multiwavelength radiometer for in situ measurements in the thermal infrared. Part II: Behavior in field experiments. J. Atmos. Oceanic Technol. 20, 1023–1033.

Coll, C., Caselles, V., Galve, J.M., Valor, E., Niclòs, R., Sánchez, J.M., Rivas, R., 2005. Ground measurements for the validation of land surface temperatures derived from AATSR and MODIS data. Rem. Sens. Environ. 97, 288–300.

Coll, C., Caselles, V., Valor, E., Niclòs, R., 2012. Comparison between different sources of atmospheric profiles for land surface temperature retrieval from single channel thermal infrared data. Rem. Sens. Environ. 117, 199–210.

Coll, C., Niclòs, R., Puchades, J., García-Santos, V., Galve, J.M., Pérez-Planells, L.I., Valor, E., Theocharous, E., 2019. Laboratory calibration and field measurement of land surface temperatura and emissivity using thermal infrared multiband radiometers. Int. J. Appl. Earth Obs. Geoinformation 78, 227–239.

Du, C., Ren, H., Qin, Q., Meng, J., Zhao, S., 2015. A practical split-window algorithm for estimating land surface temperature from landsat 8 data. Remote Sens. 7, 647–665.

England, M.N., 2019. Landsat 8: TIRS SSM Encoder Current Anomaly Resolution. In: 2019 IEEE Aerospace Conference, Big Sky, pp. 1–10. https://doi.org/10.1109/AERO.2019.8741692.

García-Santos, V., Cuxart, J., Martínez-Villagrasa, D., Jiménez, M., Simó, G., 2018. Comparison of three methods for estimating land surface temperature from Landsat 8-TIRS sensor data. Remote Sens. 10, 1450.

GCOS-200. The Global Observing System for Climate: Implementation Needs, 342PP. WMO, 2016. https://library.wmo.int/opac/doc_num.php?explnum_id=3417 (Accessed 3 August, 2020).

Gerace, A., Montanaro, M., 2017. Derivation and validation of the stray light correction algorithm for the thermal infrared sensor onboard Landsat 8. Rem. Sens. Environ. 191, 246–257.

Gerace, A., Kleynhans, T., Eon, R., Montanaro, M., 2020. Towards an Operational, SplitWindow-Derived Surface Temperature Product for the Thermal Infrared Sensors Onboard Landsat 8 and 9. Remote Sens. 12, 224.

Ghent, D., Kaduk, J., Remedios, J., Ardö, J., Balzter, H., 2010. Assimilation of land surface temperature into the land surface model JULES with an ensemble Kalman filter. J. Geophys. Res. 115 (D19), 112.

Ghent, D., Veal, K., Trent, T., Dodd, E., Sembhi, H., Remedios, J., 2019. A new approach to defining uncertainties for MODIS land surface temperature. Remote Sens. 11, 1–24.

Gillespie, A.R., Matsunaga, T., Rokugawa, S., Hook, S.J., 1998. Temperature and emissivity separation from Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) images. IEEE Trans. Geosci. Rem. Sens. 36, 1113–1125.

Guillevic, P.C., Privette, J.L., Coudert, B., Palecki, M.A., Demarty, J., Ottlé, C., Augustine, J.A., 2012. Land Surface Temperature product validation using NOAA's surface climate observation networks – scaling methodology for the Visible Infrared Imager Radiometer Suite (VIIRS). Rem. Sens. Environ. 124, 282–298.

Guillevic, P., Göttsche, F., Nickeson, J., Hulley, G., Ghent, D., Yu, Y., Trigo, I., Hook, S., Sobrino, J.A., Remedios, J., Román, M., Camacho, F., 2018. Land surface temperature product validation best practice protocol version 1.1. Best Pract. Satell. L. Prod. Valid. (p. 60) L. Prod. Valid. Subgr. doi, 58. https://doi.org/10.5067/doc/ceoswgcv/lpv/lst.001.

Guo, J., Ren, H., Zheng, Y., Lu, S., Dong, J., 2020. Evaluation of Land Surface Temperature Retrieval from Landsat 8/TIRS Images before and after Stray Light Correction Using the SURFRAD Dataset. Remote Sens. 12, 1023.

Hongtao, J., Huanfeng, S., Xinghua, L., Chao, Z., Huiqin, L., Fangni, L., 2019. Extending the SMAP 9-km soil moisture product using a spatio-temporal fusion model. Rem. Sens. Environ. 231, 111224.

Hook, S.J., Chander, G., Barsi, J.A., Alley, R.E., Abtahi, A., Palluconi, F.D., Markham, B. L., 2004. In-flight validation and recovery of water surface temperature with Landsat-5 thermal infrared data using an automated high-altitude lake validation site at Lake Tahoe. IEEE Trans. Geosci. Remote Sens. 42 (12), 2767–2776.

Hulley, G.C., Hook, S.J., 2009. The North American ASTER Land Surface Emissivity Database (NAALSED). Version 2.0. Remote Sens. Environ. 113, 1967–1975.

Hulley, G.C., Hook, S.J., Abbott, E., Malakar, N., Islam, T., Abrams, M., 2015. The ASTER Global Emissivity Dataset (ASTER GED): Mapping Earth's emissivity at 100 meter spatial scale. Geophys. Res. Lett. 42 (19), 7966–7976.

Irons, J.R., Dwyer, J.L., Barsi, J.A., 2012. The next Landsat satellite: The Landsat data continuity mission. Remote Sens. Environ. 122, 11–21.

Jiménez-Muñoz, J.-C., Cristóbal, J., Sobrino, J.A., Soria, G., Ninyerola, M., Pons, X., 2009. Revision of the Single-Channel Algorithm for Land Surface Temperature Retrieval From Landsat Thermal-Infrared Data. IEEE Trans. Geosci. Remote Sens. 47, 339–349.

Jiménez-Muñoz, J.-C., Sobrino, J.A., Skoković, D., Mattar, C., Cristóbal, J., 2014. Land Surface Temperature Retrieval Methods From Landsat-8 Thermal Infrared Sensor Data. IEEE Geosci. Remote Sens. Lett. 11, 1840–1843.

Koley, S., Jeganathan, C., 2020. Estimation and evaluation of high spatial resolution surface soil moisture using multi-sensor multi-resolution approach. Geoderma 378, 114618.

Krishnan, P., Meyers, T.P., Hook, S.J., Heuer, M., Senn, D., Dumas, E.J., 2020. Intercomparison of In Situ Sensors for Ground-Based Land Surface Temperature Measurements. Sensors 20, 5268.

Legrand, M., Pietras, C., Brogniez, G., Haeffelin, M., Abuhassan, N.K., Sicard, M., 2000. A high-accuracy multiwavelength radiometer for in situ measurements in the thermal infrared. Part I: characterization of the instrument. J. Atmos. Ocean Techn. 17, 1203–1214.

Malakar, N., Hulley, G., Hook, S., Laraby, K., Cook, M., Schott, J., 2018. An Operational Land Surface Temperature Product for Landsat Thermal Data: Methodology and Validation. IEEE Trans. Geosci. Remote Sens. 2018 (56), 5717–5735.

Mattar, C., Durán-Alarcón, C., Jiménez-Muñoz, J.C., Santamaría-Artigas, A., Olivera-Guerra, L. Sobrino, J.A., 2015. Global Atmospheric Profiles from Reanalysis Information (GAPRI): a new database for earth surface temperature retrieval. Int. J. Remote Sens. 36, 5045–5060.

Meerdink, S.K., Hook, S.J., Roberts, D.A., Abbott, E.A., 2019. The ECOSTRESS spectral library version 1.0. Remote Sens. Environ. 230, 111196.

Meng, X., Cheng, J., Zhao, S., Liu, S., Yao, Y., 2019. Estimating Land Surface Temperature from Landsat-8 Data using the NOAA JPSS Enterprise Algorithm. Remote Sens. 11, 155.
Mokhtari, A., Noory, H., Pourshakouri, F., Haghighatmehr, P., Afrasiabian, Y., Razavi, M., Fereydooni, F., Naeni, A.S., 2019. Calculating potential evapotranspiration and single crop coefficient based on energy balance equation using Landsat 8 and Sentinel-2. ISPRS J. Photogramm. Remote Sens. 154, 231–245.
Montanaro, M., Gerace, A., Lunsford, A., Reuter, D., 2014. Stray Light Artifacts in Imagery from the Landsat 8 Thermal Infrared Sensor. Remote Sens. 11, 10435–10456.
Montanaro, M., Levy, R., Markham, B., 2014b. On-Orbit Radiometric Performance of the Landsat 8 Thermal Infrared Sensor. Remote Sens. 6, 11753–11769.Montanaro, M., Gerace, A., Rohrbach, S, 2015. Toward an operational stray light correction for the Landsat 8 thermal infrared sensor. Appl. Opt. 54 (13), 3963–3978.
Montanaro, M., Gerace, A., Rohrbach, S., 2015. Toward an operational stray light correction for the Landsat 8 Thermal Infrared Sensor. Appl. Opt. 54 (13), 3963–3978.
Niclòs, R., Doña, C., Bisquert, M., Valor, E., 2014. Thermal-infrared spectral and angular characterization of crude oil and seawater emissivities for oil slick identification. IEEE Trans. Geosci. Remote Sens. 52, 5387–5395.
Niclòs, R., Pérez-Planells, Ll, Valiente, J.A., Coll, C., Valor, E., 2018. Evaluation of the S-NPP VIIRS Land Surface Temperature product using ground data acquired by an autonomous system at a rice paddy. ISPRS J. Photogramm. Remote Sens. 135, 1–12.
Olivera-Guerra, L., Mattar, C., Merlin, O., Duran-Alarcon, C., Santamaria-Artigas, A., Fuster, R., 2017. An operational method for the disaggregation of land surface temperature to estimate actual evapotranspiration in the arid region of Chile. ISPRS J. Photogramm. Remote Sens. 128, 170–181.
Qin, Z., Karnieli, A., Berliner, P., 2001. A mono-window algorithm for retrieving land surface temperature from Landsat TM data and its application to the Israel-Egypt border region. Int. J. Remote Sens. 22, 3719–3746.
Rozenstein, O., Qin, Z., Derimian, Y., Karnieli, A., 2014. Derivation of land surface temperature for Landsat-8 TIRS using a split window algorithm. Sensors. 14, 5768–5780.
Rubio, E., Caselles, V., Coll, C., Valour, E., Sospedra, F., 2003. Thermal-infrared emissivites of natural surfaces: Improvements on the experimental set-up and new measurements. Int. J. Remote Sens. 24, 5379–5390. https://doi.org/10.1080/0143116031000102412.
Trigo, I.F., Boussetta, S., Viterbo, P., Balsamo, G., Beljaars, A., Sandu, I., 2015. Comparison of model land skin temperature with remotely sensed estimates and assessment of surface-atmosphere coupling. J. Geophys. Res. 120 https://doi.org/10.1002/2015JD023812.
Vanhellemont, Q., 2020. Combined land surface emissivity and temperature estimation from Landsat 8 OLI and TIRS. ISPRS J. Photogramm. Remote Sens. 166, 390–402.
Wan, Z., 2014. New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product. Remote Sens. Environ. 140, 36–45.
Wan, Z., Dozier, J., 1996. A generalized split-window algorithm for retrieving land-surface temperature from space. IEEE Trans. Geosci. Remote Sens. 34, 892–905.
Wan, Z., Li, Z.L., 2008. Radiance-based validation of the V5 MODIS land-surface temperature product. Int. J. Remote Sens. 29 (17–18), 5373–5395.
Weng, Q., 2009. Thermal infrared remote sensing for urban climate and environmental studies: methods, applications, and trends. ISPRS J. Photogramm. Rem. Sens. 64 (4), 335–344.
Yu, X., Guo, X., Wu, Z., 2014. Land surface temperature retrieval from Landsat 8 TIRS—Comparison between radiative transfer equation-based method, split window algorithm and single channel method. Remote Sens. 6, 9829–9852.
Yu, Y., Liu, Y., Yu, P., Wang, H., 2017. Enterprise Algorithm Theoretical Basis Document for VIIRS Land Surface Temperature Production. NOAA, Silver Spring, MD, USA.
Zheng, Y., Ren, H., Guo, J., Ghent, D., Tansey, K., Hu, X., Nie, J., Chen, S., 2019. Land Surface Temperature Retrieval from Sentinel-3A Sea and Land Surface Temperature Radiometer. Using a Split-Window Algorithm. Remote Sens. 11, 650.

ANEXO D

Evaluating Landsat-9 TIRS-2 calibrations and land surface temperature retrievals against ground measurements using multi-instruments spatial and temporal sampling along transects

International Journal of Applied Earth Observation and Geoinformation 125 (2023) 103576

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

Evaluating Landsat-9 TIRS-2 calibrations and land surface temperature retrievals against ground measurements using multi-instrument spatial and temporal sampling along transects

Raquel Niclòs *, Martín Perelló, Jesús Puchades, César Coll, Enric Valor

Thermal Remote Sensing Group, Department of Earth Physics and Thermodynamics, Faculty of Physics, University of Valencia, 50, Dr. Moliner, E-46100 Burjassot, Spain

ARTICLE INFO

Keywords:
Landsat 9 TIRS-2
Calibration
Land surface temperature
Validation
Ground TIR data
Multi-instrument spatio-temporal sampling

ABSTRACT

High-resolution TIR data are essential for a wide range of studies, e.g., in environmental monitoring, urban effects, water stress, agricultural productivity, and land/water resources. Landsat series provides high-resolution TIR data, but data accuracy must be ensured to contribute to these fields. After the calibration problems of Landsat 8 (L8) TIRS due to a stray light effect, efforts were made in the design of Landsat 9 (L9) TIRS-2 to mitigate this effect. In any case, field calibration and validation activities are necessary to evaluate the accuracy of recently launched sensors. These activities require high-quality ground TIR radiance data that are representative at satellite spatial resolution and that quantify spatial–temporal fluctuations at the experimental sites concurrently to satellite overpasses. However, these requirements are not always accomplished in the literature.

This paper provides vicarious calibrations of the L9 TIRS-2 data with accurate ground measurements acquired using multi-instrument spatial and temporal sampling in a rice paddy area from December 2021 to August 2023, in order to contribute to the calibration and validation of L9 TIRS-data.

In March 1, 2023 L9 TIRS-2 scenes started to be reprocessed and calibration was updated. Both the original calibration and that after reprocessing were evaluated. Close results were obtained for band 10 (b10). When comparing satellite brightness temperatures and those determined from in situ LSTs, systematic differences of −0.3 K were shown, with root-mean-square differences (RMSDs) of 0.5 K. Results for band 11 showed a bias of −0.1 K for the original data and of 0.2 K after the reprocessing, with RMSDs of 0.6 K.

The L2 LST product generated operationally from b10 data was also evaluated, together with alternative single-channel (SC) algorithms. A bias (RMSD) of 0.4 K (1.1 K) was obtained for the product. However, negative biases were shown both with the radiative transfer equation used on b10 radiances and the alternative SC algorithm.

Finally, we validated six types of split-window (SW) algorithms applied to L9 TIRS-2 data before and after the reprocessing. A systematic difference of about −0.6 K between SW LST biases when using the reprocessed and original data was observed. In any case, for most of the algorithms with original and reprocessed data, LST biases and standard deviations were within the recommended threshold of 1 K, which proves the good performance of L9 TIRS-2 data to monitor LST at high spatial resolution.

1. Introduction

LANDSAT 9 (L9) has been in orbit since late September 2021, and operational since February 2022. Nowadays L9 together with Landsat 8 (L8) provide a full coverage every 8 days. Thermal Infrared Sensor 2 (TIRS-2) is onboard L9 together to the Operational Land Imager 2 (OLI-2). It has two thermal-infrared (TIR) bands like the TIRS in L8. A primary radiometric calibration change was introduced with the addition of baffling in the L9 TIRS-2 telescope to minimize the stray light effect found in the case of L8 TIRS (Barsi et al., 2022; Montanaro et al., 2022). On-orbit calibrations have been performed, and according to Barsi et al. (2022), TIRS-2 has shown excellent radiometric performance, with noise level below 0.07 K for both bands. Additionally, lunar scans and vicarious calibration data showed that the stray light effect has been mitigated (Barsi et al., 2022; Montanaro et al., 2022). Up to March 1, 2023, the L9 TIRS-2 absolute calibration come from the prelaunch one, with a

* Corresponding author.
E-mail address: raquel.niclos@uv.es (R. Niclòs).

https://doi.org/10.1016/j.jag.2023.103576
Received 29 August 2023; Received in revised form 2 November 2023; Accepted 18 November 2023
1569-8432/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

small response change introduced after the March 2022 cryocooler electronics reset (Barsi et al., 2022). However, L9 collection 2 (C2) data started to be reprocessed on March 1, 2023 based on the first calibration results obtained by the USGS/NASA Calibration and Validation team using ground data acquired in inland waters. According to them, the updates after TIRS-2 data reprocessing were: (i) updates to the absolute radiances, with the largest adjustment of around 0.7 K in band 11 (b11). (ii) Adjustments to improve TIRS-2 to OLI-2 co-instrument alignment. (iii) 16 detectors in TIRS-2 band 10 (b10) and 3 detectors in b11 were replaced with on-board redundant detectors to reduce striping due to detector response changes (Barsi et al., 2022).

L9 TIRS-2 can provide valuable land surface temperature (LST) records at high spatial resolution (100 m, resampled to 30 m). LST is a key variable to understand land–atmosphere physical processes, and thus for climatological and hydrological applications (Anderson et al., 2012; Good et al., 2022; Li et al., 2023). It was defined as an essential climate variable by the Global Climate Observing System (GCOS) (GCOS, 2016), and also by space agencies such as NASA and ESA, the latter through the Climate Change Initiative. Additionally, high-resolution LST retrievals are valuable in the study of: natural hazards, thermal anomalies, fire detection and severity, urban heat island effects, crop irrigation needs, and natural resources (Darge et al., 2019; García-Llamas et al., 2019; Li et al., 2023; Peres et al., 2018; Roy et al., 2015). All these applications are the objectives of the large number of dedicated satellite missions approved for the near future, i.e., TRISHNA, SBG-TIR, LSTM, and Landsat Next (Vidal et al., 2022).

Suitable calibration and validation (CAL/VAL) activities based on independent ground-data (T-based method) are required to evaluate the accuracy of both the TIRS-2 calibration and the Level 2 (L2) single-channel (SC) derived LST product, and also the suitability of split-window (SW) algorithms to be used on TIRS-2 data. Significant systematic and/or random uncertainties in the calibration of the TIRS-2 bands and in the LSTs retrieved from TIRS-2 radiance data would prevent their used for the mentioned applications. This was the case of L8 TIRS band 11 data before the systematic application of the stray light correction method proposed by Montanaro et al. (2015) in view of the CAL/VAL results. The Landsat calibration team advised against its use in SC and SW algorithms to retrieve LSTs.

First calibration evaluations for L9 TIRS-2 were performed over water surfaces in Barsi et al. (2022) between November 2021 and April 2022, which indicated possible small underestimations of less than 0.2 K in b10 and 0.7 K in b11. They used single-band TIR radiometers set at buoys in Lake Tahoe and Salton Sea, but also mined data from the NOAA National Data Buoy Center. The results were used for the reprocessing starting on March 1, 2023. However, further CAL/VAL experiments carried out over land surfaces, using ground data acquired at homogeneous sites, would allow more reliable evaluation results to be obtained. Meng et al. (2022) and Ye et al. (2022) evaluated L9 TIRS-2 LSTs retrieved before reprocessing using a dataset of temperatures derived from pyrgeometer data as reference. Meng et al. (2022) indicated root-mean-square errors (RMSEs) $\geq$ 3 K in L2 LSTs, even after filtering outliers. Ye et al. (2022) showed a RMSE of 1.8 K for L2 LSTs. However, previous studies indicated that pyrgeometer data were not accurate enough for CAL/VAL evaluations, since: (i) only the instrument uncertainty lead to a 1 K uncertainty in LST, (ii) substantial differences were shown when comparing LSTs determined from data measured by TIR radiometers and by pyrgeometers, and (iii) residual errors were consistently observed when comparing satellite-retrieved LSTs and those obtained from pyrgeometer data (Gerace et al., 2020; Guillevic et al., 2012; Krishnan et al., 2020; Niclòs et al., 2021). Thus, more robust ground-measured datasets were recommended for validations (Guo et al., 2020; Niclòs et al., 2021). Ground datasets should be acquired over thermally-homogeneous sites with TIR radiometers (Hulley et al., 2021), and possible spatial and temporal fluctuations should be evaluated during satellite overpasses (Lagouarde et al., 2019).

The objective of the paper is to provide CAL/VAL results for the L9 TIRS-2 data using accurate ground TIR data acquired with calibrated TIR radiometers carried on along transects in a rice paddy area, known as Valencia test site (Coll et al., 2012, 2005; Niclòs et al., 2021, 2018). The land surface at the site changes across the year, which is an advantage for these activities since allow us to measure in a wide range of temperatures and conditions. This site was formerly used to evaluate the performances of several satellite sensors (Coll et al., 2012, 2005; Niclòs et al., 2021, 2018), e.g., L8 TIRS bands when the stray light effect was a problem and after the correction method (Niclòs et al., 2021).

The contributions and novelties of the present paper include: (1) Vicarious calibrations of the new L9 TIRS-2 over homogeneous land covers, both for the original calibrated data (hereafter OC data; with absolute calibration based on the prelaunch calibration) and the March 2023 reprocessed data (hereafter RP data; with the calibration adjustment). (2) Use of accurate ground TIR data acquired with multi-instrument spatial and temporal sampling to quantify the site thermal spatial variability and possible high-frequency temporal fluctuations induced by atmospheric turbulence during each L9 overpass. (3) Validation of the operational L2 product, together to alternative SC methods, both for OC and RP data. (4) Evaluation of six types of SW algorithms applied to L9 TIRS-2 data in order to show the change in uncertainties introduced by the reprocessing, and test whether the accuracies of the algorithms are within recommended thresholds for climate studies.

The next section describes the experimental site, the calibration of the TIR radiometers used to acquire measurements and the procedure followed to derive LSTs from the radiance data measured. Section 3 describes the methodology followed to evaluate the calibrations of TIRS-2 b10 and b11 and the derived L2 LST product, including atmospheric radiative transfer simulations and emissivities. This section also includes a list of SW algorithms additionally validated. The calibration results both for OC and RP data are shown and discussed in Section 4.1. The validation results of the L2 LST product are shown in Section 4.2, and the results when using SW algorithms are analyzed in Section 4.3. Lastly, we summarize the most important findings in Section 5.

2. Ground data

2.1. Test site

Ground TIR radiance data were acquired in the Valencia Test Site concurrently to cloud-free L9 overpasses from December 2021 to August 2023. The site is a rice paddy area with several land surfaces as a consequence of the rice crop phenology, varying from flooded surface and wet and dry bare soils to full vegetation cover, as described in Niclòs et al. (2021). The site is covered by WRS-2 paths 198 and 199 of row 33. Fig. 1 shows L9 OLI-2 scenes at the site on May 26, 2022 (ploughed dry bare soil), August 14, 2022 (full vegetation), and January 21, 2023 (flooded soil). Fig. 2 shows photographs for the three surface conditions at the site taken in measurement days. All land covers were shown to be thermally homogeneous at various spatial scales (Coll et al., 2012, 2005; Niclòs et al., 2018). LST standard deviations (SDs) obtained from radiometer measurements along 200 m transects in the site were generally lower than 0.5 K, and up to 0.9 K in some cases for bare soils (Niclòs et al., 2021, 2018). Therefore, the site is interesting for CAL/VAL purposes under three very different but appropriate surface conditions, which permit to cover a wide range of temperatures throughout the year.

2.2. Ground TIR radiance measurements

Four Cimel Electronique TIR multichannel radiometers, model CE-312 (Brogniez et al., 2003), were used to acquire TIR radiance data. These radiometers were hand-held, distributed over the area and carried on along transects, covering the maximum possible extension within 5 min around the L9 overpass time. Each radiometer took measurements each 5–10 m, with about 25–30 measurements within the 5 min per

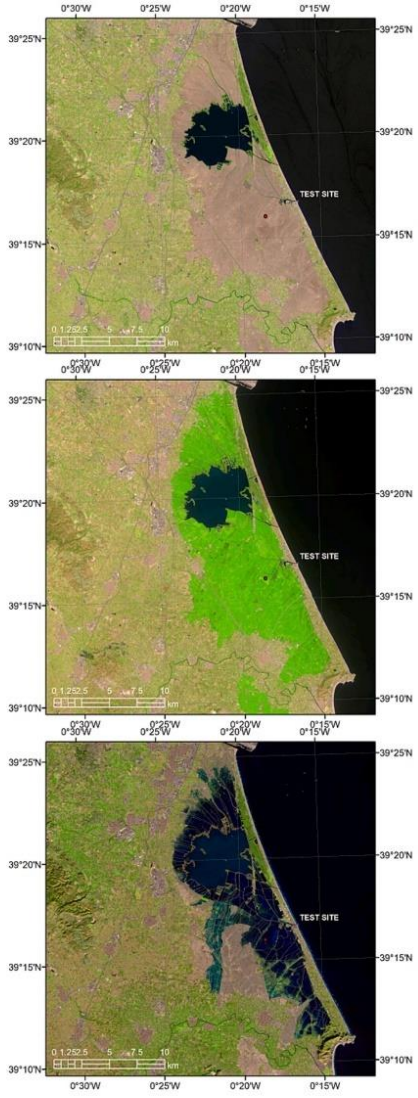

Fig. 1. L9 OLI-2 color composites (RGB 654) at the site for three of the measurement days: May 26, 2022 (ploughed dry bare soil), August 14, 2022 (full vegetation), and January 21, 2023 (flooded soil).

radiometer. Thus, using 4 radiometers, the total number of measurements taken was 100–120 per L9 overpass. This sampling was done to assess the thermal variability, which is one of the main sources of uncertainty for ground-measured LSTs. According to Lagouarde et al. (2019), the spatial thermal variability of land surfaces and temporal fluctuations induced by atmospheric turbulence can only be assessed whether measurements of several radiometers, acquiring data sufficiently spaced apart to avoid correlations, are averaged over the measurement time. Thus, such measurement protocol applied here increases the accuracy of the evaluation results.

Ground measurements were done following this procedure concurrently with the acquisition of 23 L9 scenes with cloud-free conditions over the site within the abovementioned period.

The ground radiometers acquired surface-leaving radiance in each band i as follows:

$$L_{surf,i} = \varepsilon_i B_i(LST) + (1 - \varepsilon_i) L_{ia,hem}\downarrow \quad (1)$$

where B_i is the band-averaged Planck's radiance function, ε_i is emissivity, and $L_{a,hem}^{\downarrow}$ is the atmospheric downwelling irradiance divided by π in the case of Lambertian surfaces (i.e., bare soil and full vegetation). $L_{i\,a,hem}{}^{\downarrow}$ was measured at the site using a Labsphere Infragold Reflectance Target IRT-94–100 with a high reflectivity in the TIR (Niclòs et al., 2021). However, specular reflection was considered for flooded surface, and $L_{i\,a,hem}^{\downarrow}$ was replaced by $L_{i\,a}{}^{\downarrow}$ in Eq. (1), which was measured directly at zenith. The relative percentage difference between $L_{i\,a,hem}^{\downarrow}$ and $L_{i\,a}{}^{\downarrow}$ (when divided by the former) is from 36 % to 42 % at 10.9 μm for the days we measured flooded soils, with a ratio between them from 1.5 to 1.7.

2.3. Calibration of ground radiometers

The six TIR bands of the CE-312 radiometers were calibrated once per year, and before and after experimental campaigns, using a Landcal P80P calibration source. In June 2022, our radiometer calibrations were also evaluated in an international TIR radiometer inter-comparison funded by ESA within the FRM4STS project (Yamada et al., 2023a, 2023b). During the inter-comparison experiment, TIR radiometers usually used for validation purposes were compared with reference blackbodies (BB), with uncertainties of 0.04–0.13 K for the measurement range (Yamada et al., 2023a, 2023b). They were set at nominal temperatures from −30 to 50 °C. The uncertainty budget of our radiometers was assessed as described in Coll et al. (2019) for the previous experiment in 2016, and combined uncertainties (i.e., total root sum squares) lower than 0.2 K were obtained from the 2022 calibration data. Calibration equations were also estimated from the experiment results for each radiometer and band. Since the relationship between radiance differences and radiometer digital number differences is linear, we used linear recalibration equations between BB minus radiometer-measured brightness temperatures ($T_{i,BB}$-$T_{i,m}$) and radiometer-measured minus internal detector temperature ($T_{i,m}$-T_d) (Coll et al., 2019). $T_{i,m}$ were measured with each band calibration provided originally by the manufacturer. Fig. 3 plots the results for one of the radiometers and band as an example, the solid line being the regression equation line.

2.4. Ground emissivity measurements

Unlike some previous studies, ε_i were neither estimated from emissivity datasets (e.g., ASTER-GED) or spectral libraries (e.g., ECOSTRESS) nor from threshold methods in terms of spectral vegetation indexes (Chatterjee et al., 2017; Gerace et al., 2020; Guo et al., 2020; Meng et al., 2019; Rozenstein et al., 2014; Wang et al., 2023; Yu et al., 2014). When possible, surface emissivity data acquired concurrently with a satellite sensor overpass is the best option (Rozenstein et al., 2014), as emissivity depends on the specific element types at the site and also varies with seasons and modifications in land use and soil moisture (Mira et al.,

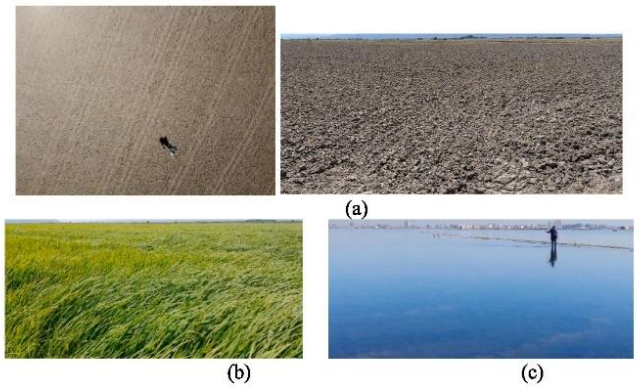

Fig. 2. Photographs taken in the experimental area with the three totally disparate surface conditions: (a) bare soil, (b) complete rice coverage, and (c) flooded soil.

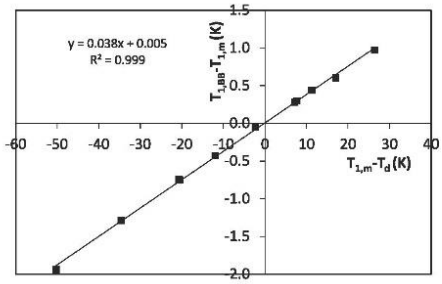

Fig. 3. Example of recalibration for band 1 of one of the radiometers used.

2010). Therefore, ε_i were measured at the Valencia test site to correct the ground-measured TIR radiances. Different measurement methods were used. The Temperature-Emissivity Separation method (Hulley and Hook, 2009) was used for wet to dry bare soils, since it works more accurately due to the higher emissivity spectral variations of this kind of surfaces. With vegetation, ε_i were measured with the Box Method since the emissivity spectrum was flatter (Niclòs et al., 2021, 2018). For water surfaces, ε_i values were obtained as described in Niclòs et al. (2014). Average ε_i values of 0.99 and 0.985 were measured (at 10.9 μm) for flooded soil and full vegetation, respectively, at the site. However, values of 0.972 were obtained for dry bare soil. ε_i uncertainties from 0.004 to 0.006 were determined for the two first types of covers, and from 0.004 to 0.011 for bare soils.

2.5. Ground LSTs and uncertainties

Eq. (1) was inverted in order to obtain the LSTs from the ground-measured radiances and corresponding measurements of surface emissivities and atmospheric downwelling radiances. Finally, we calculated average values of the LSTs retrieved from the TIR radiances measured by the different CE-312 radiometers for each L9 overpass. The average ground LSTs ranged from 279 K to 314 K for the whole study period. Additionally, the uncertainty budget was assessed for the LST retrieval, and the different uncertainty sources were added together in quadrature (Niclòs et al., 2021). An uncertainty lower than 0.2 K was estimated for the calibration of the radiometer bands used. A spatial and temporal variability usually from 0.3 to 0.5 K was determined from the LSTs measured by the different radiometers along transects within 5 min of each L9 overpass. The largest uncertainty propagated through Eq. (1) was the emissivity one, of around 0.4–0.6 K in LSTs. Finally, ground LST total uncertainties of 0.6–1.0 K were determined for the study period.

The same dataset of ground-measured LSTs was used for evaluating the calibration of L9 TIR bands (Section 3.1) and for the validation of SC and SW LST products (Sections 3.2 and 3.3).

3. Methodology

3.1. Evaluation of the L9 TIRS-2 calibrations

The vicarious calibration involved the evaluation of L9 TIRS-2 data against those determined from ground data. TOA radiances, L_i, were simulated as follows. First, we calculated surface radiances from ground-measured LSTs using Eq. (1) with the band-integrated magnitudes for TIRS-2 bands 10 and 11. Second, simulated L9 TOA radiances were obtained from the surface radiances with the radiative transfer equation (RTE):

$$L_i = \tau_i L_{surf,i} + L_{ia}\uparrow \quad (2)$$

where τ_i and $L_{ia}^{\uparrow}$ are the transmittance and upwelling radiance for the whole atmosphere, respectively. They, together with $L_{i\ a,hem}^{\downarrow}$, were estimated for the atmospheric condition of each L9 overpass and then integrated for TIRS-2 bands. For that, we used MODTRAN (Berk et al., 2008) simulations with NCEP atmospheric profiles interpolated for the site and overpass time. We extracted the four NCEP grid corners surrounding the site location for the two times before and after the overpass time. First, the corner profiles were spatially interpolated to obtain the profiles at the site location for each time, and then the resulting profiles were temporally interpolated for the acquisition time. The interpolated profiles were then completed from their maximum heights (of around 30 km above sea level) to 100 km from the MODTRAN standard atmospheres most appropriate for the site location and L9 overpass time (i.e., midlatitude summer or winter profiles). An inverse distance weighted interpolation with $1/d_j^2$ as weighting function was used for the spatial interpolation, d_j being the distance from the site location to the NCEP grid corners (Coll et al., 2012). Additionally, ε_i values required in Eq. (1)

were those measured in the experimental area (Section 2). The calibration was evaluated in terms of L_i, but also brightness temperatures, T_i. A flowchart of the vicarious calibration procedure is shown in Fig. 4a.

L9 TIRS-2 scenes (C2) were downloaded, both before and after the March 2023 reprocessing (i.e., OC and RP data), using the tool provided in earthexplorer.usgs.gov. We used the raster tool of the Sentinel Application Platform (SNAP) to extract pixel values of these scenes over the experimental site. The pixel values were then analyzed using a Matlab code developed to assess band 10 and 11 radiances by applying the radiometric calibration to convert raw digital numbers to radiances with scale and offset values provided in the metadata of the scenes. Brightness temperatures were calculated from radiances using the Planck's function with band-dependent coefficients. Then, evaluations against ground data were done using the single value of the pixel covering the site. We also calculated average values and SDs for a pixel array covering 5 × 5 100-m pixels (~17 × 17 30-m pixels) (Niclòs et al., 2021). This array size implies a good compromise between a substantial number of 30 m pixels to compute SDs and keeping a homogeneous footprint for the array. The rice paddy parcels have usually distances between narrow adjacent paths or irrigation ditches from 500 m to 750 m. Thus, 5 x 5 100-m pixels are similar size, and the array footprint is not covering these paths and ditches. In any case, we also tested the use of averages and SDs for 3x3 100-m pixels to compare results.

After the L_i and T_i calibration, we also used the RTE (with Eqs. (1) and (2)) to assess LST from b10 and b11 L_i, separately, and then they were evaluated directly against ground LSTs.

Four uncertainty sources were considered to determine the uncertainties in the LSTs retrieved in this way (Niclòs et al., 2021): (i) those of measured ε_i (from 0.004 to 0.011), (ii) uncertainties of the atmospheric variables both due to those in NCEP profiles and MODTRAN simulations (0.4 K and 0.8 K for TIRS-2 b10 and b11) (Coll et al., 2012), (iii) uncertainties in T_i, as $NE\Delta T/n.pixels^{1/2}$, where n.pixels is the number of pixels in the array (Ghent et al., 2019), and (iv) the array variability, as the SD of the pixels considered. $NE\Delta T < 0.1$ K was shown for both TIRS-2 bands (Barsi et al., 2022; Pearlman et al., 2022). These four uncertainty contributions led to LST uncertainties from 0.5 to 0.8 K for b10 and from 0.8 to 1.0 K for b11. These uncertainty sources were also considered to determine uncertainties in L_i and T_i assessed from in situ LSTs using Eqs. (1) and (2) directly, thus obtaining T_{10} uncertainties from 0.7 to 1.0 K and T_{11} uncertainties from 0.9 to 1.1 K.

3.2. L2 LST product validation

The L9 L2 LST product available in C2 makes use of the RTE, i.e., uses a SC technique on TIRS-2 b10 radiances. It uses ε_{10} values estimated from the ASTER-GED and MODTRAN simulations of the atmospheric parameters (τ_{10}, $L^{\uparrow}_{10\ a}$, $L_{10\ a,\overline{hem}}$) as input data (Malakar et al., 2018; Meng et al., 2022). The latter were performed using the GEOS-5 FP-IT atmospheric data (https://gmao.gsfc.nasa.gov/GMAO_products/NRT_products.php). ASTER-GED emissivities, i.e., average emissivities generated from cloud-free ASTER images between 2000 and 2008 at ~ 100 m (3 arc sec) resolution, are temporally and spectrally adjusted for TIRS-2 data. Temporal adjustments are based on changes in vegetation and snow covers by using related spectral indexes (Malakar et al., 2018). L9 scenes of the product were also downloaded before and after the March 2023 reprocessing, i.e., for OC and RP data.

Additionally, an alternative SC method suggested by Jiménez-Muñoz et al. (2014) was used for comparison. We applied this method to at-sensor L_i and T_i data in b10 but also in b11 (hereafter referred as SC_JM_10 and SC_JM_11, respectively). Additional input values were ε_i measured at the site and the atmospheric total water vapor content, W, estimated for each scene from interpolated NCEP profiles (Fig. 4a). The method provides second degree polynomial regression equations between atmospheric variables and W, which facilitate the algorithm application for potential users (i.e., without needing atmospheric simulations of τ_i, $L^{\uparrow}_{i\ a}$, $L_{i\ a,\overline{hem}}$). Coefficients for these equations were only given for b10 in Jiménez-Muñoz et al. (2014), but Guo et al. (2020) provided coefficients for b11. We used both of them. However, this approximation can involve higher uncertainties. Yu et al. (2014) compared the L8 results of this SC algorithm with those when using directly the RTE, and concluded that LSTs estimated from direct RTE calculations were more accurate. Jiménez-Muñoz et al. (2014) evaluated the input W uncertainty and the impact in the LST uncertainty with the algorithm, of up to 1 K. This input variable together with the emissivity one and the algorithm regression uncertainty lead to considerable LST uncertainties compared to those of the RTE.

3.3. Evaluation of SW algorithms

As summarized in Niclòs et al. (2021), five types of SW algorithms were fit for L8, each of them with several versions due to regression coefficients obtained from different simulation datasets and different

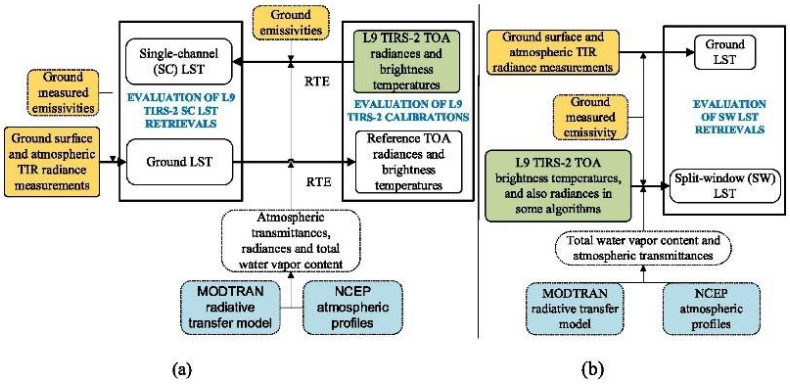

Fig. 4. Flowcharts of the procedure followed for evaluating: a) both the L9 TIRS-2 calibrations and the LSTs retrieved from L9 TIRS-2 data inverting the RTE, but also SC algorithms, and b) SW algorithms.

parametric equations. As indicated in Ye et al. (2022), b10 and b11 of TIRS-2 and TIRS are spectrally extremely close, and the identical algorithm is used to generate the L2 LST product from L9 and L8 data. Thus, we first evaluated these five SW algorithms for L9 TIRS-2.

(i) The first algorithm, proposed by Jiménez-Muñoz et al. (2014), is quadratic on brightness temperature differences, T_{10} - T_{11} (see in Table 1). We applied this algorithm both with the regression coefficients provided in Jiménez-Muñoz et al. (2014) (hereafter referred to as SW1), and also with the regression coefficients obtained in Meng et al. (2019) per subranges of W (hereafter SW1_M_W) and also for any W ≤ 7 cm (SW1_M).

(ii) The second class of algorithms is based on that designed by Wan and Dozier (1996) with refinements in Wan (2014), and coefficients obtained from regression by:

1) Meng et al. (2019) for any W ≤ 7 cm (hereafter SW2_M) and per subranges of W (SW2_M_W).

2) Du et al. (2015) for any W ≤ 6.3 cm (SW2_D) and per subranges of W (SW2_D_W), and also refined by Guo et al. (2020) using different coefficients depending on W and temperature (SW2_D_W_T).

3) Gerace et al. (2020) (SW2_G) for any W and temperature. This algorithm was proposed to generate operationally a new and alternative LST product in the next Landsat collection 3 (C3).

(iii) The third algorithm is the NOAA JPSS Enterprise SW equation with coefficients regressed per W subranges (hereafter SW3_M_W) and also for W ≤ 7.0 cm (SW3_M) by Meng et al. (2019).

(iv) The fourth algorithm was proposed by Yu et al. (2014) for TIRS (hereafter SW4) similarly to that in Qin (2001) for AVHRR, which is based on assessing parameters p_i using linear regressions with TIRS TOA T_i.

(v) The fifth algorithm is that of Rozenstein et al. (2014) (SW5) also based on that in Qin et al. (2001), but p_i parameters are unused in this case and it only considers the corresponding regression coefficients.

Both SW4 and SW5 algorithm use atmospheric transmittances, τ_i, and authors provided regressions that characterize the τ_i dependence on W for a midlatitude summer profile (hereafter SW4_MLS and SW5_MLS) and the US standard atmosphere (hereafter SW4_US and SW5_US), and they recommend the US regressions to users. The regressions are quadratic in SW4 and linear in SW5. However, we also evaluated both algorithms using the τ_i values that we simulated with MODTRAN from NCEP profiles interpolated for the L9 overpasses at the site (hereafter SW4_SIM and SW5_SIM, respectively).

(vi) Here, we also considered a recently proposed method by Wang et al. (2023) (hereafter SW6) to retrieve LSTs using L9 TIRS-2 data based on the radiance-based SW algorithm proposed for MODIS (Wang et al., 2019). The algorithm uses TIRS-2 L_i instead of T_i values. The black-body equivalent surface emittance at b10, B_{10}, is first estimated using a regression equation with L_{10} and L_{11}. Then the LST is estimated from B_{10} with the RTE. We applied the algorithm using τ_i and $\varphi_i = L^{\downarrow}_{i\,a,hem} / L^{\downarrow}_{i\,a}$ values simulated with the interpolated NCEP profiles and MODTRAN (hereafter SW6_SIM), but also those estimated from W with regression equations proposed in Wang et al. (2023) using a selection of profiles from the TIGR2000 dataset (hereafter SW6_REG).

The six SW algorithms are summarized in Table 1. These algorithms, using different coefficients or τ_i estimates, were applied to L9 TIRS-2 data (both with OC and RP data) and evaluated against ground data, through Matlab code developed. As inputs in the SW algorithms, we used the ε_i values measured at the site, and the W estimates obtained from interpolated NCEP profiles. Atmospheric parameters simulated with MODTRAN were also used for the SW4_SIM, SW5_SIM and SW6_SIM versions. Fig. 4b shows the flowchart used for the validation of the SW algorithms.

To determine the uncertainties of the different algorithms, we took the next five uncertainty sources into account: the algorithm regression uncertainty provided by the authors, the propagation in the algorithm of measured ε_i uncertainties, those of W uncertainty (0.5 cm (Niclòs et al., 2021)) and band NEΔTs (as NEΔT/n.pixels$^{1/2}$), and the SD of LSTs in the array. The total LST uncertainties varied from 0.7 K to 2.0 K for the different versions of algorithm SW1, with average uncertainties around 1.2 K. Uncertainties were from 0.4 K to 1.9 K for the different SW2 versions, with averages around 0.8–1.1 K. They varied from 0.6 K to 2.0 K for SW3 versions, with averages of 1.1 K. Uncertainties from 0.4 K to 2.1 K were determined for the SW4_SIM, SW5_SIM, SW5_US and SW5_MLS, with average uncertainties of 1.0–1.1 K (with maximum uncertainties up to nearly 3.0 K for the SW4_US and SW4_MLS, which increased the averages up to 1.5 K). Finally, they were in the range of 0.6–3.0 K for the SW6_SIM and SW6_REG versions, with averages of 1.5 K.

Table 1
Summary of the SW algorithms used, including names assigned, equations for each algorithm, and references in which they were proposed.

Name	SW algorithm	Reference
SW1	$LST = T_{10} + c_0 - c_1\,(T_{10} - T_{11}) + c_2\,(T_{10} - T_{11})^2 + (c_3 + c_4\,W)\,(1-\varepsilon) + (c_5 - c_6\,W)\,\Delta\varepsilon$	(Jimenez-Munoz et al., 2014)
SW2	$LST = c_0 + \left(c_1 + c_3\frac{1-\varepsilon}{\varepsilon} + c_5\frac{\Delta\varepsilon}{\varepsilon^2}\right)\frac{T_{10}+T_{11}}{2} + \left(c_2 + c_4\frac{1-\varepsilon}{\varepsilon} + c_6\frac{\Delta\varepsilon}{\varepsilon^2}\right)\frac{T_{10}-T_{11}}{2} + c_7(T_{10}-T_{11})^2$	(Wan and Dozier, 1996) & Wan (2014)
SW3	$LST = c_0 + c_1\,T_{10} + c_2\,(T_{10} - T_{11}) + c_3\,\varepsilon + c_4\,\varepsilon\,(T_{10} - T_{11}) - c_5\,\Delta\varepsilon$	(Meng et al., 2019)
SW4	$LST = T_{10} + B_1\cdot(T_{10} - T_{11}) + B_0$	(Yu et al., 2014)
	$B_0 = \frac{C_{11}\cdot(1 - A_{10} - C_{10})\cdot p_{10} - C_{10}\cdot(1 - A_{11} - C_{11})\cdot p_{11}}{C_{11}\cdot A_{10} - C_{10}\cdot A_{11}}$	
	$B_1 = \frac{C_{10}}{C_{11}\cdot A_{10} - C_{10}\cdot A_{11}}$	
	$A_{10} = \varepsilon_{10}\cdot\tau_{10}$	
	$A_{11} = \varepsilon_{11}\cdot\tau_{11}$	
	$C_{10} = (1-\varepsilon_{10})\cdot(1 - (1-\varepsilon_{10})\cdot\tau_{10})$	
	$C_{11} = (1-\varepsilon_{11})\cdot(1 - (1-\varepsilon_{11})\cdot\tau_{11})$	
SW5	$LST = A_0 - A_1\cdot T_{10} - A_2\cdot T_{11}$	(Rozenstein et al., 2014)
	$A_0 = E_1\cdot b_{10} - E_2\cdot b_{11}$	
	$A_1 = 1 - A + E_1\cdot a_{10}$	
	$A_2 = A + E_2\cdot a_{11}$	
	$A = D_{10}/E_0$	
	$E_1 = D_{11}\cdot(1 - C_{10} - D_{10})/E_0$	
	$E_2 = D_{10}\cdot(1 - C_{11} - D_{11})/E_0$	
	$E_0 = D_{11}\cdot C_{10} - D_{10}\cdot C_{11}$	
	$C_i = \varepsilon_i\cdot\tau_i(\theta)$	
	$D_i = [1 - \tau_i(\theta)]\cdot[1 + (1-\varepsilon_i)\cdot\tau_i(\theta)]$	
SW6	$LST = \frac{c_2\cdot\lambda_{10}^{-5}}{\ln\left(\frac{c_1\cdot\lambda_{10}^{-5}}{B_{10}(LST)} + 1\right)}$	(Wang et al., 2023)
	$B_{10}(LST) = A_0\cdot L_{10} + A_1\cdot L_{11} + A_2$	
	$A_0 = \frac{D_{11}}{C_{10}\cdot D_{11} - C_{11}\cdot D_{10}}$	
	$A_1 = \frac{-D_{10}}{K\cdot(C_{10}\cdot D_{11} - C_{11}\cdot D_{10})}$	
	$A_2 = \frac{b\cdot D_{10}\cdot(C_{11} + D_{11})}{k\cdot(C_{10}\cdot D_{11} - C_{11}\cdot D_{10})}$	
	$C_{10/11} = \varepsilon_{10/11}\cdot\tau_{10/11}$	
	$D_{10/11} = (1 - \tau_{10/11})\cdot[1 + (1-\varepsilon_{10/11})\cdot\tau_{10/11}\cdot\varphi_{10/11}]$	
	$K = \frac{c_1^2\cdot\lambda_{10}^{-4}\cdot\lambda_{11}^{-5}\cdot(c_1\cdot\lambda_{10}^{-5}\cdot L_{10}^{-1} + 1)^{\left(c_2\lambda_{10}\cdot\lambda_{11}^{-1}\right)-1)}}{\left[(c_1\cdot\lambda_{10}^{-5}\cdot L_{10}^{-1} + 1)^{\left(\lambda_{10}\cdot\lambda_{11}^{-1}\right)} - 1\right]^2\cdot L_{10}^2}$	
	$b = \frac{c_1\cdot\lambda_{11}^{-5}}{(c_1\cdot\lambda_{10}^{-5}\cdot L_{10}^{-1} - 1)^{\left(\lambda_{10}\cdot\lambda_{11}^{-1}\right)} - 1} - k\cdot L_{10}$	
	$\tau_{10/11} = a_0\cdot\omega + a_1$	
	$\varphi_{10/11} = a_2\cdot\ln(\omega) + a_3$	

parametric equations. As indicated in Ye et al. (2022), b10 and b11 of TIRS-2 and TIRS are spectrally extremely close, and the identical algorithm is used to generate the L2 LST product from L9 and L8 data. Thus, we first evaluated these five SW algorithms for L9 TIRS-2.

(i) The first algorithm, proposed by Jiménez-Muñoz et al. (2014), is quadratic on brightness temperature differences, T_{10} - T_{11} (see in Table 1). We applied this algorithm both with the regression coefficients provided in Jiménez-Muñoz et al. (2014) (hereafter referred to as SW1), and also with the regression coefficients obtained in Meng et al. (2019) per subranges of W (hereafter SW1_M_W) and also for any W ≤ 7 cm (SW1_M).

(ii) The second class of algorithms is based on that designed by Wan and Dozier (1996) with refinements in Wan (2014), and coefficients obtained from regression by:

1) Meng et al. (2019) for any W ≤ 7 cm (hereafter SW2_M) and per subranges of W (SW2_M_W).

2) Du et al. (2015) for any W ≤ 6.3 cm (SW2_D) and per subranges of W (SW2_D_W), and also refined by Guo et al. (2020) using different coefficients depending on W and temperature (SW2_D_W_T).

3) Gerace et al. (2020) (SW2_G) for any W and temperature. This algorithm was proposed to generate operationally a new and alternative LST product in the next Landsat collection 3 (C3).

(iii) The third algorithm is the NOAA JPSS Enterprise SW equation with coefficients regressed per W subranges (hereafter SW3_M_W) and also for W ≤ 7.0 cm (SW3_M) by Meng et al. (2019).

(iv) The fourth algorithm was proposed by Yu et al. (2014) for TIRS (hereafter SW4) similarly to that in Qin (2001) for AVHRR, which is based on assessing parameters p_i using linear regressions with TIRS TOA T_i.

(v) The fifth algorithm is that of Rozenstein et al. (2014) (SW5) also based on that in Qin et al. (2001), but p_i parameters are unused in this case and it only considers the corresponding regression coefficients.

Both SW4 and SW5 algorithm use atmospheric transmittances, τ_i, and authors provided regressions that characterize the τ_i dependence on W for a midlatitude summer profile (hereafter SW4_MLS and SW5_MLS) and the US standard atmosphere (hereafter SW4_US and SW5_US), and they recommend the US regressions to users. The regressions are quadratic in SW4 and linear in SW5. However, we also evaluated both algorithms using the τ_i values that we simulated with MODTRAN from NCEP profiles interpolated for the L9 overpasses at the site (hereafter SW4_SIM and SW5_SIM, respectively).

(vi) Here, we also considered a recently proposed method by Wang et al. (2023) (hereafter SW6) to retrieve LSTs using L9 TIRS-2 data based on the radiance-based SW algorithm proposed for MODIS (Wang et al., 2019). The algorithm uses TIRS-2 L_i instead of T_i values. The black-body equivalent surface emittance at b10, B_{10}, is first estimated using a regression equation with L_{10} and L_{11}. Then the LST is estimated from B_{10} with the RTE. We applied the algorithm using τ_i and $\varphi_i = L_{i_a}^{\downarrow}{}_{hem} / L_{i_a}^{\uparrow}$ values simulated with the interpolated NCEP profiles and MODTRAN (hereafter SW6_SIM), but also those estimated from W with regression equations proposed in Wang et al. (2023) using a selection of profiles from the TIGR2000 dataset (hereafter SW6_REG).

The six SW algorithms are summarized in Table 1. These algorithms, using different coefficients or τ_i estimates, were applied to L9 TIRS-2 data (both with OC and RP data) and evaluated against ground data, through Matlab code developed. As inputs in the SW algorithms, we used the ε_i values measured at the site, and the W estimates obtained from interpolated NCEP profiles. Atmospheric parameters simulated with MODTRAN were also used for the SW4_SIM, SW5_SIM and SW6_SIM versions. Fig. 4b shows the flowchart used for the validation of the SW algorithms.

To determine the uncertainties of the different algorithms, we took the next five uncertainty sources into account: the algorithm regression uncertainty provided by the authors, the propagation in the algorithm of measured ε_i uncertainties, those of W uncertainty (0.5 cm (Niclòs et al., 2021)) and band NEΔTs (as NEΔT/n.pixels$^{1/2}$), and the SD of LSTs in the array. The total LST uncertainties varied from 0.7 K to 2.0 K for the different versions of algorithm SW1, with average uncertainties around 1.2 K. Uncertainties were from 0.4 K to 1.9 K for the different SW2 versions, with averages around 0.8–1.1 K. They varied from 0.6 K to 2.0 K for SW3 versions, with averages of 1.1 K. Uncertainties from 0.4 K to 2.1 K were determined for the SW4_SIM, SW5_SIM, SW5_US and SW5_MLS, with average uncertainties of 1.0–1.1 K (with maximum uncertainties up to nearly 3.0 K for the SW4_US and SW4_MLS, which increased the averages up to 1.5 K). Finally, they were in the range of 0.6–3.0 K for the SW6_SIM and SW6_REG versions, with averages of 1.5 K.

Table 1
Summary of the SW algorithms used, including names assigned, equations for each algorithm, and references in which they were proposed.

Name	SW algorithm	Reference
SW1	$LST = T_{10} + c_0 + c_1 (T_{10} - T_{11}) + c_2 (T_{10} - T_{11})^2 + (c_3 + c_4 W)(1 - \varepsilon) + (c_5 - c_6 W)\,\Delta\varepsilon$	(Jimenez-Munoz et al., 2014)
SW2	$LST = c_0 + \left(c_1 + c_2\frac{1-\varepsilon}{\varepsilon} + c_3\frac{\Delta\varepsilon}{\varepsilon^2}\right)\frac{T_{10}+T_{11}}{2} + \left(c_4 + c_5\frac{1-\varepsilon}{\varepsilon} + c_6\frac{\Delta\varepsilon}{\varepsilon^2}\right)\frac{T_{10}-T_{11}}{2} + c_7(T_{10}-T_{11})^2$	(Wan and Dozier, 1996) & Wan (2014)
SW3	$LST = c_0 + c_1 T_{10} + c_2 (T_{10} - T_{11}) + c_3\,\varepsilon + c_4\,\varepsilon\,(T_{10} - T_{11}) + c_5\,\Delta\varepsilon$	(Meng et al., 2019)
SW4	$LST = T_{10} + B_1\hat{A}\cdot(T_{10} - T_{11}) + B_0$	(Yu et al., 2014)
	$B_0 = \frac{C_{11}\hat{A}\cdot(1 - A_{10} - C_{10})\hat{A}\cdot p_{10} - C_{10}\hat{A}\cdot(1 - A_{11} - C_{11})\hat{A}\cdot p_{11}}{C_{11}\hat{A}\cdot A_{10} - C_{10}\hat{A}\cdot A_{11}}$	
	$B_1 = \frac{C_{10}}{C_{11}\hat{A}\cdot A_{10} - C_{10}\hat{A}\cdot A_{11}}$	
	$A_{10} = \varepsilon_{10}\hat{A}\cdot\tau_{10}$	
	$A_{11} = \varepsilon_{11}\hat{A}\cdot\tau_{11}$	
	$C_{10} = (1 - \tau_{10})\hat{A}\cdot(1 - (1 - \varepsilon_{10})\hat{A}\cdot\tau_{10})$	
	$C_{11} = (1 - \tau_{11})\hat{A}\cdot(1 - (1 - \varepsilon_{11})\hat{A}\cdot\tau_{11})$	
SW5	$LST = A_0 - A_1\hat{A}\cdot T_{10} - A_2\hat{A}\cdot T_{11}$	(Rozenstein et al., 2014)
	$A_0 = E_1\hat{A}\cdot b_{10} - E_2\hat{A}\cdot b_{11}$	
	$A_1 = 1 - A + E_1\hat{A}\cdot a_{10}$	
	$A_2 = A + E_2\hat{A}\cdot a_{11}$	
	$A = D_{10}/E_0$	
	$E_1 = D_{11}\hat{A}\cdot(1 - C_{10} - D_{10})/E_0$	
	$E_2 = D_{10}\hat{A}\cdot(1 - C_{11} - D_{11})/E_0$	
	$E_0 = D_{11}\hat{A}\cdot C_{10} - D_{10}\hat{A}\cdot C_{11}$	
	$C_i = \varepsilon_i\hat{A}\cdot\tau_i(\theta)$	
	$D_i = [1 - \tau_i(\theta)]\hat{A}\cdot[1 + (1 - \varepsilon_i)\hat{A}\cdot\tau_i(\theta)]$	
SW6	$LST = \frac{c_2\hat{A}\cdot\lambda_{10}^{-1}}{\ln\left(\frac{c_1\hat{A}\cdot\lambda_{10}^{-5}}{B_{10}(LST)} + 1\right)}$	(Wang et al., 2023)
	$B_{10}(LST) = A_0\hat{A}\cdot L_{10} + A_1\hat{A}\cdot L_{11} + A_2$	
	$A_0 = \frac{D_{11}}{C_{10}\hat{A}\cdot D_{11} - C_{11}\hat{A}\cdot D_{10}}$	
	$A_1 = \frac{-D_{10}}{K\hat{A}\cdot(C_{10}\hat{A}\cdot D_{11} - C_{11}\hat{A}\cdot D_{10})}$	
	$A_2 = \frac{bA\cdot D_{10}\hat{A}\cdot(C_{11} + D_{11})}{k\hat{A}\cdot(C_{10}\hat{A}\cdot D_{11} - C_{11}\hat{A}\cdot D_{10})}$	
	$C_{10/11} = \varepsilon_{10/11}\hat{A}\cdot\tau_{10/11}$	
	$D_{10/11} = (1 - \tau_{10/11})\hat{A}\cdot[(1 - \varepsilon_{10/11})\hat{A}\cdot\tau_{10/11}\hat{A}\cdot\varphi_{10/11} + 1]$	
	$K = \frac{c_1^2\hat{A}\cdot\lambda_{10}^{-4}\hat{A}\cdot\lambda_{11}^{-5}\hat{A}\cdot\left(c_1\hat{A}\cdot\lambda_{10}^{-5}\hat{A}\cdot L_{10}^{-1} + 1\right)^{\left(\lambda_{10}\hat{A}\cdot\lambda_{11}^{-1}\right)}}{\left[\left(c_1\hat{A}\cdot\lambda_{10}^{-5}\hat{A}\cdot L_{10}^{-1} + 1\right)^{\left(\lambda_{10}\hat{A}\cdot\lambda_{11}^{-1}\right)} - 1\right]^2\hat{A}\cdot L_{10}^2}$	
	$b = \frac{c_1\hat{A}\cdot\lambda_{11}^{-5}}{\left(c_1\hat{A}\cdot\lambda_{10}^{-5}\hat{A}\cdot L_{10}^{-1} - 1\right)^{\left(\lambda_{10}\hat{A}\cdot\lambda_{11}^{-1}\right)} - 1} - k\hat{A}\cdot L_{10}$	
	$\tau_{10/11} = a_0\hat{A}\cdot\omega + a_1$	
	$\varphi_{10/11} = a_2\hat{A}\cdot\ln(\omega) + a_3$	

4. Results

4.1. L9 TIRS-2 calibration results

Statistics of the differences between L9 TIRS-2 L_i and T_i values and those determined from ground LSTs were assessed to obtain systematic (mean error or bias) and random (SD) uncertainties, and root-mean-square differences (RMSD). We also computed three different statistical tests to compare them in order to reinforce the statistical results. First, we used a paired T-test, in which the null hypothesis was that the means of satellite and ground data were equal. The function "ttest2" in Matlab was used, with a 5 % significance level for rejecting the null hypothesis. The second test was the analysis of variance (ANOVA), used to determine whether temperatures and radiances from both groups have a common mean ("anova1" function in Matlab). Finally, we assessed the Levene's test in which the null hypothesis was that both groups had equal variances ("vartestn" function in Matlab using "LeveneAbsolute" option). Table 2 shows bias, SD and RMSD values, and also p-values for the abovementioned tests, for: a) the single value of the pixel covering the site, and b) 3x3 100-m pixel array and c) 5x5 100-m pixel array over the site in L9 TIRS-2 images. Table 2 shows close results for all of them (with negligible differences, < 0.005 W/m^2·sr·μm and < 0.03 K), which proved again the site homogeneity. Results for TIRS-2 OC and RP data are shown in Table 2. Since OC data were no longer provided after the reprocessing in March 2023, they were evaluated from December 2021 to February 2023, with a total of 17 cloud-free matchups of the 23 measurement days in the study period. Results for RP data are shown both for the same period (17 matchups), to permit a proper comparison of calibration results, and from December 2021 to August 2023, with the total of 23 cloud-free matchups in this case.

Fig. 5a shows T_i retrieved from ground data against satellite T_i (average values for the 5x5 100-m pixel array) with OC data for b10 and b11. Fig. 5b shows the same comparison but for RP data. In these figures, errors bars of TIRS-2 TOA T_i show the uncertainties calculated as the square sum of the SD of the pixels in the array and $NE\Delta T/n.pixels^{1/2}$. These uncertainties ranged from 0.06 to 0.5 K for both bands. Error bars of T_i retrieved from ground data were determined as detailed in Section 3.1.

In Table 2, the p-values of the three statistical hypothesis tests show that the null hypothesis cannot be rejected in any case, giving the high p-values shown. This fact reinforces the hypothesis that the L9 L_i and T_i data and those obtained from ground measurements are equivalent in terms of means and variances. Moreover, the ANOVA F statistics (i.e., between-group variance divided by within-group variance) were close to zero (≤ 0.01 for all of them).

As shown in Table 2 and in Fig. 5a, TIRS-2 OC L_i and T_i data agreed with those estimated from ground data, with slight negative biases (-0.3 K for b10 and -0.1 K for b11) and RMSDs of 0.5 K for b10 and 0.6 K for b11. The TIRS-2 underestimation was also observed by Barsi et al. (2022) using data measured in inland waters as reference for the L9 data evaluation up to April 2022. In that case, bias was of -0.7 K for b11 and -0.2 K for b10.

No appreciable T_{10} bias difference (of around -0.01 K) between RP and OC data was shown in Table 2. However, a significant bias difference of 0.37 K in T_{11} was found. Therefore, the 2023 reprocessing introduced a larger calibration change for b11, increasing the difference between biases in b10 and b11 (from -0.17 K for OC data to -0.55 K for RP data), with potential consequences when applying SW algorithms. The last three columns of Table 2 shows the results for RP data with the total number of matchups (23), yielding close results to those for 17 matchups. Like Table 2, Fig. 5b reveals the band difference in terms of biases for RP data, not observed in Fig. 5a for OC data.

Similarly, the evaluation was done in terms of LST. Table 3 shows the validation results for the TIRS-2 retrieved LSTs using the RTE on b10 and b11 data, as explained in Section 3.1, for the 5x5 100-m pixel array. Close results were obtained in terms of LSTs for the 3x3 100-m pixel array and the single pixel covering the site, with negligible differences among them (<0.04 K), similarly to those shown in Table 2. Therefore, the analysis of results is shown only for the 5x5 100-m pixel array hereafter. Fig. 6 shows in situ LSTs against LSTs retrieved from TIRS-2 data.

Table 3 and Fig. 6 show again the good agreement in terms of LSTs estimated using OC data, with slight underestimates of -0.4 K and -0.1 K and RMSDs of 0.6 K and 0.8 K for b10 and b11. A notable disparity in biases is shown between both bands when using RP data (-0.4 K for b10 and 0.4 K for b11).

Table 2
Statistics of the L_i and T_i differences (satellite minus reference data) for the two TIRS-2 bands, both with OC and RP data, for a single pixel (1p) and 3x3 100-m array (3x3) and 5x5 100-m array (5x5).

		OC data (N = 17)			RP data (N = 17)			RP data (N = 23)		
		1p	3x3	5x5	1p	3x3	5x5	1p	3x3	5x5
L_{10} (10^{-2} W/m^2 sr μm)	Bias	−4	−4	−4	−4	−4	−4	−4	−4	−4
	SD	5	5	5	5	5	5	6	6	6
	RMSD	6	7	7	7	7	7	7	7	7
	ANOVA p	0.93	0.93	0.93	0.93	0.92	0.92	0.92	0.92	0.91
	T-test p	0.93	0.93	0.93	0.92	0.92	0.91	0.92	0.92	0.91
	Levene p	0.97	0.98	0.98	0.98	0.98	0.98	0.91	0.92	0.93
L_{11} (10^{-2} W/m^2 sr μm)	Bias	−1	−1	−1	3	2	2	2	2	2
	SD	6	6	6	6	6	7	7	7	7
	RMSD	7	7	7	7	7	7	7	7	7
	ANOVA p	0.97	0.97	0.97	0.93	0.94	0.94	0.93	0.93	0.94
	T-test p	0.97	0.96	0.96	0.93	0.93	0.94	0.93	0.93	0.94
	Levene p	0.89	0.90	0.89	0.89	0.90	0.90	0.91	0.92	0.93
T_{10} (K)	Bias	−0.3	−0.3	−0.3	−0.3	−0.3	−0.3	−0.3	−0.3	−0.3
	SD	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
	RMSD	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
	ANOVA p	0.93	0.93	0.93	0.92	0.92	0.92	0.92	0.92	0.91
	T-test p	0.93	0.93	0.93	0.92	0.92	0.91	0.92	0.92	0.91
	Levene p	0.96	0.97	0.96	0.97	0.97	0.97	0.91	0.93	0.94
T_{11} (K)	Bias	−0.1	−0.1	−0.1	0.2	0.2	0.2	0.2	0.2	0.2
	SD	0.6	0.6	0.6	0.5	0.5	0.5	0.6	0.6	0.6
	RMSD	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
	ANOVA p	0.97	0.96	0.96	0.94	0.94	0.94	0.93	0.93	0.94
	T-test p	0.97	0.96	0.96	0.93	0.93	0.94	0.93	0.93	0.93
	Levene p	0.88	0.89	0.89	0.90	0.91	0.90	0.93	0.94	0.94

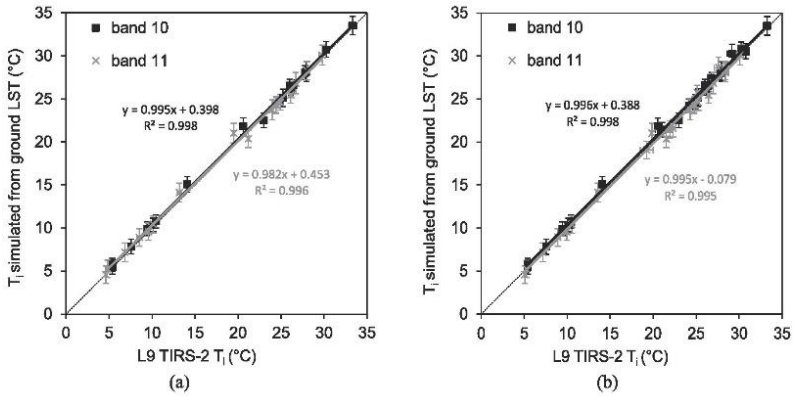

Fig. 5. T_i retrieved from ground data against L9 TIRS-2 T_i values when using TIRS-2 (a) OC data (17 matchups), and (b) RP data (23 matchups). Linear regressions are shown as solid lines. Regression equations and coefficients of determination are included. The dotted line is the 1:1 identity line.

Table 3
Statistics of the LST differences (satellite minus reference data) when using the RTE both with OC and RP data.

	OC data (N = 17)			RP data (N = 17)			RP data (N = 23)		
Bias	**SD**	**RMSD**	**Bias**	**SD**	**RMSD**	**Bias**	**SD**	**RMSD**	
LST_{10} (K)	−0.4	0.5	0.6	−0.4	0.5	0.6	−0.4	0.5	0.6
LST_{11} (K)	−0.1	0.8	0.8	0.4	0.8	0.9	0.3	0.9	0.9

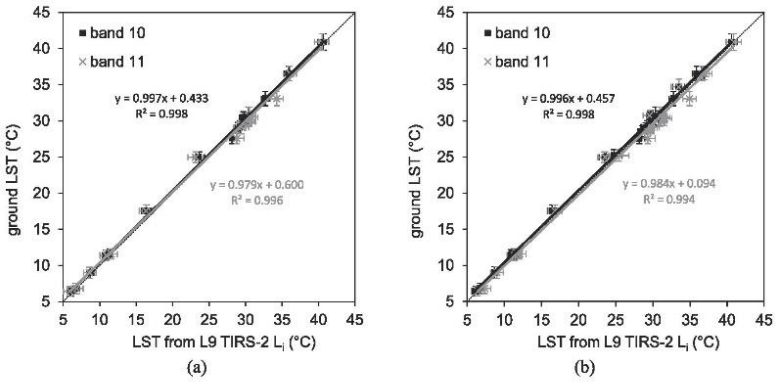

Fig. 6. Reference LST versus LSTs retrieved from TIRS-2 L_i with (a) OC data, and (b) RP data.

Table 4
Statistics of the differences between the L2 LST product and ground LST for OC and RP data. The same is shown for the alternative SC_JM_10 and SC_JM_11 algorithms.

	OC data (N = 17)			RP data (N = 17)			RP data (N = 23)		
Bias	**SD**	**RMSD**	**Bias**	**SD**	**RMSD**	**Bias**	**SD**	**RMSD**	
LST_{L2} (K)	0.4	1.0	1.1	0.4	1.0	1.1	0.5	1.1	1.2
LST_{SC_JM10} (K)	−0.7	0.9	1.2	−0.7	1.0	1.2	−0.9	1.1	1.4
LST_{SC_JM11} (K)	1.1	1.9	2.2	1.6	2.0	2.5	1.0	2.2	2.5

4.2. Validation results for the L2 LST product and alternative SC algorithms

Table 4 includes validation results for the L2 LST product generated operationally from TIRS-2 b10 data, and the SC algorithm proposed in Jiménez-Muñoz at al. (2014) applied to TIRS-2 data in b10 and b11 separately. Fig. 7 plots the comparison of the ground LSTs with L2 LSTs (from OC and RP data). X-axis error bars shows the square sum of the average uncertainty for the LSTs over the 5 x5 100-m pixel array and the LST spatial variability. L2 product provides the LST uncertainty for each pixel, which includes the different error sources (those from emissivity, atmospheric parameters, etc.) and the nearest cloud distance effect (Malakar et al., 2018). For the latter, Laraby et al. (2016) studied the relationship between cloud proximity and LST errors reported using buoy data as reference. We observed that the cloud distance effect increases the LST uncertainty from around 2 K (for a cloud distance > 10 km) to more than 5 K (for a cloud distance < 1 km). Some cases in our data set were affected by the latter condition (see Fig. 7), while the ground measurements were carried out only for cloud-free conditions, at least 10 km away from the site. Thus, LST product uncertainties were largely overestimated for some cases in our site because of wrong cloud detection.

Table 4 shows a slight and similar LST overestimation of 0.4–0.5 K for the L2 product both with OC and RP data, with the overestimation increasing with temperature (Fig. 7). However, the results shown in Table 3 for the RTE LST for b10 reveal an underestimation (−0.4 K). The validation results for the SC_JM_10 algorithm applied to b10, with ε_{10} values measured at the site and W estimated for each L9 scene, show also an underestimation (of around −0.7 K).

The emissivities included in the L2 product were close to those measured at the site for flooded soil and bare soil conditions (with differences up to 0.003). However, this was not the case for the full vegetation conditions, the emissivities in the product being lower than those measured at the site by 0.010–0.013. The emissivity underestimation explains in part the systematic positive bias (even larger than 1 K) observed for full vegetation cover, which skew the overall L2 results to a positive bias. Additionally, we observed unexpected too low atmospheric downwelling radiance values in the L2 product, which are of around 50 % of the upwelling radiances included. This fact could also explain in part the positive bias.

In any case, RMSDs of 1.1–1.2 K were shown for the L2 product, which are much lower than those obtained in Meng et al. (2022) and Ye et al. (2022) with pyrgeometer data. On the other hand, the alternative SC_JM_10 and SC_JM_11 algorithms provided larger uncertainties than the RTE (Table 3). This fact was expected, as mentioned in Section 3.2, according to the findings of Jiménez-Muñoz et al. (2014) and Yu et al. (2014). The alternative algorithms also provided larger uncertainties than the L2 LST product (Table 4). Furthermore, results for the SC_JM_11, with coefficients given in Guo et al. (2020), showed RMSD 1 K larger than those for the SC_JM_10, and thus the use of SC_JM_11 is not recommended.

4.3. Validation results for the SW algorithms

Table 5 summarizes the validations corresponding to the 6 types of SW algorithms, and their variants (Section 3.3), when they were applied to OC and RP T_{10} and T_{11} data. The same ground LST data were used as reference for the analysis of results. Results for RP data are given for the total of 23 matchups, with close results for 17 matchups. Fig. 8 shows a box plot comparing the results for the different algorithms.

Table 5
Statistics of the LST differences (satellite minus ground data) obtained for the SW techniques when using TIRS-2 T_{10} and T_{11} as input variables, both from OC and RP data. See Section 3.3 for the description of the SW versions.

SW equation	SW LST − ground LST (K)					
	OC data (N = 17)			RP data (N = 23)		
	Bias	SD	RMSD	Bias	SD	RMSD
SW1	−0.7	0.7	1.0	−1.3	0.7	1.4
SW1_M_W	−0.4	0.7	0.8	−1.0	0.8	1.2
SW1_JM_M	−0.4	0.9	1.0	−1.0	0.9	1.3
SW2_M_W	−0.3	0.7	0.7	−0.8	0.7	1.1
SW2_M	−0.1	0.7	0.7	−0.6	0.6	0.9
SW2_D_W	0.8	0.8	1.1	0.2	0.8	0.8
SW2_D	0.8	0.9	1.2	0.3	0.8	0.9
SW2_D_W_T	0.0	0.8	0.8	−0.6	0.8	1.0
SW2_G	−0.3	0.7	0.8	−0.9	0.7	1.1
SW3_M_W	−0.3	0.7	0.7	−0.9	0.7	1.1
SW3_M	−0.2	0.8	0.8	−0.8	0.9	1.2
SW4_SIM	0.1	0.7	0.7	−0.5	0.7	0.9
SW4_US	0.0	0.6	0.6	−0.6	0.7	0.9
SW4_MLS	0.9	1.3	1.6	0.2	1.3	1.3
SW5_SIM	0.1	0.7	0.7	−0.5	0.7	0.9
SW5_US	−0.2	0.7	0.7	−0.7	0.7	1.0
SW5_MLS	−0.3	0.7	0.8	−0.9	0.7	1.1
SW6_SIM	−0.3	0.6	0.6	−0.8	0.7	1.0
SW6_REG	−1.3	0.7	1.5	−1.8	0.7	2.0

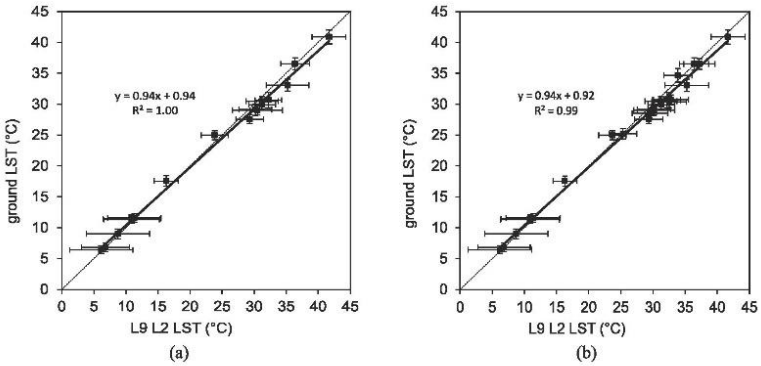

Fig. 7. Ground LSTs against L2 LSTs when using (a) OC data, and (b) RP data.

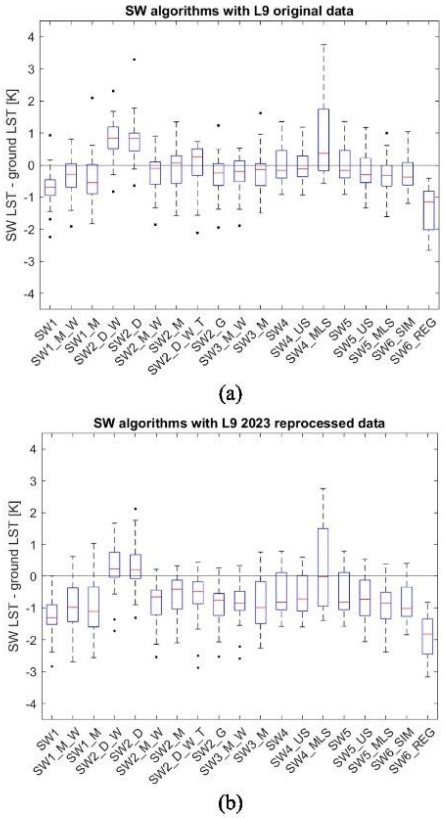

Fig. 8. Comparison of SW algorithm validation for the TIRS-2 (a) OC data and (b) RP data. Box plots show median values, and first and third quartiles. Outliers are shown as dots. Whiskers corresponds to approximately 99.3% coverage.

Table 5 and Fig. 8 show an overall difference of around −0.6 K in the biases obtained for the SW algorithms between RP and OC data, which leads to larger negative biases for most SW algorithms with RP data. When using OC data, biases varied from −1.3 K to 0.9 K depending on the algorithm, with an average of the biases of −0.08 K. Biases from −1.8 K to 0.3 K were shown with RP data, and the average was −0.7 K. However, SDs are quite similar for OC and RP data. The reprocessing increases the difference between the systematic uncertainties in T_{10} and T_{11} (from −0.17 K in OC data to −0.55 K in RP data, Table 2), and leads to a significant change in the biases obtained when using the SW algorithms.

According to the results when using OC data, several SW algorithms provided RMSDs ≤ 1 K, such as different versions of the algorithm based on that proposed by Wan and Dozier (1996) and Wan (2014) (i.e., SW2_M_W, SW2_M, SW2_D_W_T, and SW2_G). Additionally, RMSDs ≤ 1 K were also found for the rest of algorithm types, except in the case of using τ_i regressions on W for Midlatitude atmospheres in SW4_MLS and in the case of using τ_i and φ_i regressions on W in SW6_REG.

Due to the change in T_{11} bias introduced by the 2023 reprocessing, the RMSDs of most of them exceed 1 K with RP data. Bias increased from −0.3 K to −0.9 K for the SW2_G proposed for the future operational SW LST product, and thus RMSD increased from 0.8 K to 1.1 K. After the 2023 reprocessing, the SW algorithm with the lowest RMSD is the SW2_D_W version.

In any case, biases and SDs for most of the SW algorithms studied, both applied to OC and RP data, are within the recommended threshold of 1 K established by the GCOS for LST satellite-retrieved products (GCOS, 2016).

5. Summary and conclusions

This paper provides, first, vicarious calibrations of the L9 TIRS-2 radiances and brightness temperatures in b10 and b11 against ground TIR radiance data measured in a thermally homogeneous site using multi-instrument, spatial and temporal sampling along transects. The vicarious calibrations were performed using L1 data from the original calibration and after the March 2023 reprocessing (OC and RP data, respectively), in which absolute radiances were updated (with the largest adjustment in b11) and some detectors were replaced in both bands to reduce striping.

According to our evaluations at the Valencia test site, close results were obtained for OC and RP data for b10, with radiance biases (RMSDs) of −0.04 (0.07) W/m^2·sr·μm, and brightness temperature biases (RMSDs) of −0.3 (0.5) K. However, results for b11 were different for OC and RP data. While a negligible underestimation was shown for OC data (with radiance bias (RMSD) of −0.01 (0.07) W/m^2·sr·μm, and −0.1 (0.6) K in brightness temperatures), a slight overestimation was observed with RP data (of 0.02 (0.07) W/m^2·sr·μm and 0.2 (0.6) K, respectively). Therefore, the 2023 reprocessing led to an increase in the difference between the systematic uncertainties in T_{10} and T_{11}, from −0.17 K in OC data to −0.55 K in RP data, which was reflected in the results of SW algorithms. Thus, based on our results, the change introduced in the calibration of b11 by the reprocessing seems to be overestimated, accordingly to the low T_{11} bias obtained for OC data at the site.

Secondly, we evaluated the results in LSTs when using the RTE on the L9 TIRS-2 radiances in b10 and b11, i.e., a SC method per band, with emissivity values measured at the site and simulated atmospheric parameters. In that case, LST biases (RMSDs) of −0.4 (0.6) K were obtained when using b10 data, both for OC and RP data. However, for b11, a bias (RMSD) of −0.1 (0.8) K was obtained with OC data, but they were 0.3 K (0.9 K) with RP data. In contrast, positive biases of 0.4–0.5 K, and RMSDs of 1.1–1.2 K, were obtained when validating the L2 product generated operationally from b10 data in C2. An emissivity underestimation detected in the L2 product at the site in full vegetation cover conditions could explain, in part, the systematic positive bias. Additionally, unexpected too low atmospheric downwelling radiances were observed in the L2 product at the site, and this fact could also contribute to the positive bias. However, the results of the SC algorithm proposed in Jiménez-Muñoz et al. (2014) applied to b10 data gave a negative bias from −0.7 K to −0.9 K, with RMSD from 1.2 K to 1.4 K.

Finally, the validations for most of the SW equations tested for L9 data showed biases (accuracy) and SDs (precision) within the recommended GCOS threshold of 1 K, for both OC and RP data, being one of them the SW algorithm provided in Gerace et al. (2020) (SW2_G) for a new operational LST product in the next C3. However, a significant systematic difference (about −0.6 K) was observed between biases when using RP data and biases for OC data, due to the change in brightness temperature systematic uncertainties between both bands with the reprocessing. Thus, systematic uncertainties for most of the SW algorithms, like that proposed in Gerace et al. (2020), were increased with the reprocessing, and RMSDs also increased accordingly.

In any case, the CAL/VAL results reported here for L9 TIRS-2 OC data

and the current RP data show good accuracies and precisions to monitor accurately the LST at the high spatial resolution provided by the sensor and to use it for the applications mentioned in Section 1. No significant calibration errors were observed, unlike in the first calibrations of L8 TIRS due to the stray light effect, and both SC and SW methods can provide LSTs from L9 TIRS-2 data with uncertainties around 1 K.

The methodology of using multi-instrument spatial and temporal TIR sampling along transects provided highly accurate and representative ground data. These data yielded high quality CAL/VAL results for L9 TIRS-2 at the Valencia test site. However, more ground data will be acquired following the methodology to allow evaluating possible further reprocessings of TIRS-2 data. Next high-resolution thermal missions will also require high-quality data such as these, and thus the methodology will be applicable at the site for them and it could be extended to other sites for CAL/VAL purposes.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This study was carried out in the framework of the project PID2020-118797RB-I00 (Tool4Extreme) funded by the Spanish Ministry of Science and Innovation and the Spanish State Research Agency (MCIN/AEI/10.13039/501100011033). We also acknowledge the project PROMETEO/2021/016 funded by Generalitat Valenciana. Authors would like to thank the useful information about the L9 status provided by Dr. J. Barsi, from NPL-NASA.

References

Anderson, M.C., Allen, R.G., Morse, A., Kustas, W.P., 2012. Use of Landsat thermal imagery in monitoring evapotranspiration and managing water resources. Remote Sens. Environ. 122, 50–65. https://doi.org/10.1016/j.rse.2011.08.025.

Barsi, J.A., Montanaro, M., Thome, K., Raqueno, N.G., Hook, S.J., Anderson, C., Micijevic, E., 2022. Early radiometric performance of Landsat-9 Thermal Infrared Sensor, in: Butler, J.J., Xiong, X. (Jack), Gu, X. (Eds.), Earth Observing Systems XXVII. SPIE, p. 31. 10.1117/12.2634058.

Berk, A., Anderson, G.P., Acharya, P.K., Shettle, E.P., 2008. MODTRAN5. 2.0. 0 user's manual. Spectr. Sci. Inc., Burlingt. MA, Air Force Res. Lab. Hanscom MA.

Brogniez, G., Pietras, C., Legrand, M., Dubuisson, P., Haeffelin, M., 2003. A High-Accuracy Multiwavelength Radiometer for In Situ Measurements in the Thermal Infrared. Part II: Behavior in Field Experiments. J. Atmos. Ocean. Technol. 20, 1023–1033. https://doi.org/10.1175/1520-0426(2003)20<1023:AHMRFI>2.0.CO;2.

Chatterjee, R.S., Singh, N., Thapa, S., Sharma, D., Kumar, D., 2017. Retrieval of land surface temperature (LST) from landsat TM6 and TIRS data by single channel radiative transfer algorithm using satellite and ground-based inputs. Int. J. Appl. Earth Obs. Geoinf. 58, 264–277. https://doi.org/10.1016/j.jag.2017.02.017.

Coll, C., Caselles, V., Galve, J.M., Valor, E., Niclòs, R., Sánchez, J.M., Rivas, R., 2005. Ground measurements for the validation of land surface temperatures derived from AATSR and MODIS data. Remote Sens. Environ. 97, 288–300. https://doi.org/10.1016/j.rse.2005.05.007.

Coll, C., Caselles, V., Valor, E., Niclòs, R., 2012. Comparison between different sources of atmospheric profiles for land surface temperature retrieval from single channel thermal infrared data. Remote Sens. Environ. 117, 199–210. https://doi.org/10.1016/j.rse.2011.09.018.

Coll, C., Niclòs, R., Puchades, J., García-Santos, V., Galve, J.M., Pérez-Planells, L., Valor, E., Theocharous, E., 2019. Laboratory calibration and field measurement of land surface temperature and emissivity using thermal infrared multiband radiometers. Int. J. Appl. Earth Obs. Geoinf. 78, 227–239. https://doi.org/10.1016/j.jag.2019.02.002.

Darge, Y.M., Hailu, B.T., Muluneh, A.A., Kidane, T., 2019. Detection of geothermal anomalies using Landsat 8 TIRS data in Tulu Moye geothermal prospect, Main Ethiopian Rift. Int. J. Appl. Earth Obs. Geoinf. 74, 16–26. https://doi.org/10.1016/j.jag.2018.08.027.

Du, C., Ren, H., Qin, Q., Meng, J., Zhao, S., 2015. A Practical Split-Window Algorithm for Estimating Land Surface Temperature from Landsat 8 Data. Remote Sens. (Basel) 7, 647–665. https://doi.org/10.3390/rs70100647.

García-Llamas, P., Suárez-Seoane, S., Fernández-Guisuraga, J.M., Fernández-García, V., Fernández-Manso, A., Quintano, C., Taboada, A., Marcos, E., Calvo, L., 2019. Evaluation and comparison of Landsat 8, Sentinel-2 and Deimos-1 remote sensing indices for assessing burn severity in Mediterranean fire-prone ecosystems. Int. J. Appl. Earth Obs. Geoinf. 80, 137–144. https://doi.org/10.1016/j.jag.2019.04.006.

GCOS, 2016. The Global Observing System for Climate: Implementation Needs. World Meteorol. Organ. 200, 341.

Gerace, A., Kleynhans, T., Eon, R., Montanaro, M., 2020. Towards an operational, split window-derived surface temperature product for the thermal infrared sensors onboard Landsat 8 and 9. Remote Sens. 12 https://doi.org/10.3390/rs12020224.

Ghent, D., Veal, K., Trent, T., Dodd, E., Sembhi, H., Remedios, J., 2019. A new approach to defining uncertainties for MODIS land surface temperature. Remote Sens. 11 https://doi.org/10.3390/rs11091021.

Good, E.J., Aldred, F.M., Ghent, D.J., Veal, K.L., Jimenez, C., 2022. An Analysis of the Stability and Trends in the LST_cci Land Surface Temperature Datasets Over Europe. Earth Sp. Sci. 9 https://doi.org/10.1029/2022EA002317.

Guillevic, P.C., Privette, J.L., Coudert, B., Palecki, M.A., Demarty, J., Ottlé, C., Augustine, J.A., 2012. Land surface temperature product validation using NOAA's surface climate observation networks-scaling methodology for the visible infrared imager radiometer suite (VIIRS). Remote Sens. Environ. 124, 282–298. https://doi.org/10.1016/j.rse.2012.05.004.

Guo, J., Ren, H., Zheng, Y., Lu, S., Dong, J., 2020. Evaluation of land surface temperature retrieval from landsat 8/TIRS images before and after stray light correction using the SURFRAD dataset. Remote Sens. 12 https://doi.org/10.3390/rs12061023.

Hulley, G.C., Hook, S.J., 2009. The North American ASTER Land Surface Emissivity Database (NAALSED) Version 2.0. Remote Sens. Environ. 113, 1967–1975. https://doi.org/10.1016/j.rse.2009.05.005.

Hulley, G.C., Gottsche, F.M., Rivera, G., Hook, S.J., Freepartner, R.J., Martin, M.A., Cawse-Nicholson, K., Johnson, W.R., 2021. Validation and Quality Assessment of the ECOSTRESS Level-2 Land Surface Temperature and Emissivity Product. IEEE Trans. Geosci. Remote Sens. 1–23 https://doi.org/10.1109/TGRS.2021.3079879.

Jimenez-Munoz, J.C., Sobrino, J.A., Skokovic, D., Mattar, C., Cristobal, J., 2014. Land surface temperature retrieval methods from landsat-8 thermal infrared sensor data. IEEE Geosci. Remote Sens. Lett. 11, 1840–1843. https://doi.org/10.1109/LGRS.2014.2312032.

Krishnan, P., Meyers, T.P., Hook, S.J., Heuer, M., Senn, D., Dumas, E.J., 2020. Intercomparison of in situ sensors for ground-based land surface temperature measurements. Sensors (switzerland) 20, 1–26. https://doi.org/10.3390/s20185268.

Lagouarde, J.-P., Irvine, M., Guillevic, P., 2019. New Insights of Ground-Based Land Surface Temperature Measurements Protocols for Improving Validation of Thermal Infrared Satellite Data, in: IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium. IEEE, pp. 9184–9187. 10.1109/IGARSS.2019.8898312.

Laraby, K.G., Schott, J.R., Raqueno, N., 2016. Developing a confidence metric for the Landsat land surface temperature product, in: Velez-Reyes, M., Messinger, D.W. (Eds.). p. 98400C. 10.1117/12.2222582.

Li, Z., Wu, H., Duan, S., Zhao, W., Ren, H., Liu, X., Leng, P., Tang, R., Ye, X., Zhu, J., Sun, Y., Si, M., Liu, M., Li, J., Zhang, X., Shang, G., Tang, B., Yan, G., Zhou, C., 2023. Satellite Remote Sensing of Global Land Surface Temperature: Definition, Methods, Products, and Applications. Rev. Geophys. 61 https://doi.org/10.1029/2022RG000777.

Malakar, N.K., Hulley, G.C., Hook, S.J., Laraby, K., Cook, M., Schott, J.R., 2018. An Operational Land Surface Temperature Product for Landsat Thermal Data: Methodology and Validation. IEEE Trans. Geosci. Remote Sens. 56, 5717–5735. https://doi.org/10.1109/TGRS.2018.2824828.

Meng, X., Cheng, J., Zhao, S., Liu, S., Yao, Y., 2019. Estimating land surface temperature from Landsat-8 data using the NOAA JPSS enterprise algorithm. Remote Sens. 11 https://doi.org/10.3390/rs11020155.

Meng, X., Cheng, J., Guo, H., Guo, Y., Yao, B., 2022. Accuracy Evaluation of the Landsat 9 Land Surface Temperature Product. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 15, 8694–8703. https://doi.org/10.1109/JSTARS.2022.3212736.

Mira, M., Valor, E., Caselles, V., Rubio, E., Coll, C., Galve, J.M., Niclòs, R., Sánchez, J.M., Boluda, R., 2010. Soil moisture effect on thermal infrared (813-μm) emissivity. IEEE Trans. Geosci. Remote Sens. 48, 2251–2260. https://doi.org/10.1109/TGRS.2009.2039143.

Montanaro, M., Gerace, A., Rohrbach, S., 2015. Toward an operational stray light correction for the Landsat 8 Thermal Infrared Sensor. Appl. Opt. 54, 3963–3978. https://doi.org/10.1364/AO.54.003963.

Montanaro, M., McCorkel, J., Tveekrem, J., Stauder, J., Mentzell, E., Lunsford, A., Hair, J., Reuter, D., 2022. Landsat 9 Thermal Infrared Sensor 2 (TIRS-2) Stray Light Mitigation and Assessment. IEEE Trans. Geosci. Remote Sens. 60, 1–8. https://doi.org/10.1109/TGRS.2022.3177312.

Niclòs, R., Doña, C., Valor, E., Bisquert, M., 2014. Thermal-infrared spectral and angular characterization of crude oil and seawater emissivities for oil slick identification. IEEE Trans. Geosci. Remote Sens. 52, 5387–5395. https://doi.org/10.1109/TGRS.2013.2288517.

Niclòs, R., Pérez-Planells, L., Coll, C., Valiente, J.A., Valor, E., 2018. Evaluation of the S-NPP VIIRS land surface temperature product using ground data acquired by an autonomous system at a rice paddy. ISPRS J. Photogramm. Remote Sens. 135, 1–12. https://doi.org/10.1016/j.isprsjprs.2017.10.017.

Niclòs, R., Puchades, J., Coll, C., Barberà, M.J., Pérez-Planells, L., Valiente, J.A., Sánchez, J.M., Juan, M.S., 2021. Evaluation of Landsat-8 TIRS data recalibrations and land surface temperature split-window algorithms over a homogeneous crop

area with different phenological land covers. ISPRS J. Photogramm. Remote Sens. 174, 237–253. https://doi.org/10.1016/j.isprsjprs.2021.02.005.
Pearlman, A., Efremova, B., Montanaro, M., Lunsford, A., Reuter, D., McCorkel, J., 2022. Landsat 9 Thermal Infrared Sensor 2 On-Orbit Calibration and Initial Performance. IEEE Trans. Geosci. Remote Sens. 60, 1–8. https://doi.org/10.1109/TGRS.2022.3183551.
Peres, L. de F., Lucena, A.J. de, Rotunno Filho, O.C., França, J.R. de A., 2018. The urban heat island in Rio de Janeiro, Brazil, in the last 30 years using remote sensing data. Int. J. Appl. Earth Obs. Geoinf. 64, 104–116. 10.1016/j.jag.2017.08.012.
Qin, Z., Karnieli, A., Berliner, P., 2001. A mono-window algorithm for retrieving land surface temperature from Landsat TM data and its application to the Israel-Egypt border region. Int J Remote Sens 22, 3719–3746. https://doi.org/10.1080/01431160010006971.
Roy, P., Guha, A., Kumar, K.V., 2015. An approach of surface coal fire detection from ASTER and Landsat-8 thermal data: Jharia coal field, India. Int. J. Appl. Earth Obs. Geoinf. 39, 120–127. https://doi.org/10.1016/j.jag.2015.03.009.
Rozenstein, O., Qin, Z., Derimian, Y., Karnieli, A., 2014. Derivation of Land Surface Temperature for Landsat-8 TIRS Using a Split Window Algorithm. Sensors 14, 5768–5780. https://doi.org/10.3390/s140405768.
Vidal, T.H.G., Gamet, P., Olioso, A., Jacob, F., 2022. Optimizing TRISHNA TIR channels configuration for improved land surface temperature and emissivity measurements. Remote Sens. Environ. 272, 112939 https://doi.org/10.1016/j.rse.2022.112939.
Wan, Z., 2014. New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product. Remote Sens. Environ. 140, 36–45. https://doi.org/10.1016/j.rse.201.
Wan, Z., Dozier, J., 1996. A generalized split-window algorithm for retrieving land-surface temperature from space. IEEE Trans. Geosci. Remote Sens. 34, 892–905. https://doi.org/10.1109/36.5084063.08.027.
Wang, M., He, G., Zhang, Z., Wang, G., Wang, Z., Yin, R., Cui, S., Wu, Z., Cao, X., 2019. A radiance-based split-window algorithm for land surface temperature retrieval: Theory and application to MODIS data. Int. J. Appl. Earth Obs. Geoinf. 76, 204–217. https://doi.org/10.1016/j.jag.2018.11.015.
Wang, M., Li, M., Zhang, Z., Hu, T., He, G., Zhang, Z., Wang, G., Li, H., Tan, J., Liu, X., 2023. Land Surface Temperature Retrieval From Landsat 9 TIRS-2 Data Using Radiance-Based Split-Window Algorithm. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 16, 1100–1112. https://doi.org/10.1109/JSTARS.2022.3232621.
Yamada, Y., Harris, S., Hayes, M., Simpson, R., Wimmer, W., Holmes, R., Nightingale, T., Lee, A., Jepsen, N., Morgan, N., Göttsche, F.-M., Niclòs, R., Perelló, M., Donlon, C., Fox, N., 2023a. Results from CEOS International Thermal Infrared Radiometer Inter-comparison (CRIC) Part 2 of 3: Laboratory Comparison of Radiometers.
Yamada, Y., Harris, S., Hayes, M., Simpson, R., Wimmer, W., Holmes, R., Nightingale, T., Lee, A., Jepsen, N., Morgan, N., Göttsche, F.-M., Niclòs, R., Perelló, M., Donlon, C., Fox, N., 2023. CEOS International Thermal Infrared RadiometerComparison: Part I: Laboratory Comparison of Radiometers and Blackbodies. Submitted J. Atmos. Oceanic Technol.
Ye, X., Ren, H., Zhu, J., Fan, W., Qin, Q., 2022. Split-Window Algorithm for Land Surface Temperature Retrieval From Landsat-9 Remote Sensing Images. IEEE Geosci. Remote Sens. Lett. 19, 1–5. https://doi.org/10.1109/LGRS.2022.3184980.
Yu, X., Guo, X., Wu, Z., 2014. Land Surface Temperature Retrieval from Landsat 8 TIRS—Comparison between Radiative Transfer Equation-Based Method, Split Window Algorithm and Single Channel Method. Remote Sens. 6, 9829–9852. https://doi.org/10.3390/rs6109829.

ANEXO E

Demonstrating the suitability of the Radiance-Based Method for Assessing the Accuracy of MODIS Land Surface Temperature Products

IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, VOL. 62, 2024 4412915

Demonstrating the Suitability of the Radiance-Based Method for Assessing the Accuracy of MODIS Land Surface Temperature Products

César Coll, Raquel Niclòs, Jesús Puchades, Vicente García-Santos, Martín Perelló, and Lluís Pérez-Planells

***Abstract*— Validation of satellite land surface temperature (LST) products is usually performed through direct comparison with ground-measured LSTs (temperature-based or T-b validation) and with the radiance-based (R-b) method where reference ground LSTs are obtained from at-sensor radiances, atmospheric temperature and water vapor profiles, and emissivity measurements. While T-b is the preferred method, it is only applicable to a few, single-component covers that are thermally homogeneous from the ground measurement scale (a few m) to the satellite scale (usually 1 km). The indirect R-b method is an alternative to extend the LST validation at a global scale over cover types where the T-b method is not feasible. In this study, we used ground LST measurements taken in a thermally homogeneous site coincident with moderate resolution imaging spectroradiometer (MODIS) Terra and Aqua overpasses (86 matchups) to assess the accuracy of the R-b LSTs derived from MODIS data. Mean bias (ground minus R-b LSTs) of −0.1 K and root mean square error (RMSE) of 0.8 K were obtained, showing good performance against ground measurements. Then, we apply operationally the R–b method for validating MODIS Level 2 LST products M*D11 and M*D21 over six test sites comprising a varied range of surfaces (4267 cases). The emissivity values necessary for the R-b calculations were obtained from 1) the M*D11 and M*D21 emissivity product and 2) an independently modeled emissivity using ground measurements and vegetation cover fraction estimates. For the M*D11 product, we obtained an overall bias (product minus R-b LST) of −0.5 K and RMSE of 0.9 K with either the product or the modeled emissivities. However, a significant difference was found between daytime and nighttime bias (−1.0 and 0.0 K, respectively). The daytime cold bias of M*D11 was attributed to an ill-tuning of the algorithm for high surface temperature and atmospheric humidity conditions. For the M*D21 product, the overall bias (RMSE) was 0.2 K (0.6 K) with the product emissivity, and 0.6 K (1.1 K) with the modeled emissivity, both in nighttime and daytime. These results compare well with recent studies and contribute to the global assessment of MODIS LST uncertainty.**

Manuscript received 12 April 2024; revised 26 June 2024; accepted 26 August 2024. Date of publication 4 September 2024; date of current version 17 September 2024. This work was supported in part by the Generalitat Valenciana under Project PROMETEO/2021/016 and in part by the Spanish Ministry of Science and Innovation and the Spanish State Research Agency (MCIN/AEI/10.13039/501100011033) under Project PID2020-118797RB100 (Tool4Extreme). *(Corresponding author: César Coll.)*

César Coll, Raquel Niclòs, Jesús Puchades, Vicente García-Santos, and Martín Perelló are with the Department of Earth Physics and Thermodynamics, Faculty of Physics, University of Valencia, 46100 Burjassot, Spain (e-mail: cesar.coll@uv.es; raquel.niclos@uv.es; jesus.puchades@uv.es; Vicente.Garcia-Santos@uv.es; Martin.Perello@uv.es).

Lluís Pérez-Planells is with IMK-ASF, Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany (e-mail: Lluis.Perez@uv.es).

Digital Object Identifier 10.1109/TGRS.2024.3454377

***Index Terms*— Emissivity, land surface temperature (LST), M*D11 and M*D21 products, radiance-based (R-b) validation.**

I. Introduction

LAND surface temperature (LST) is a key magnitude for many of Earth's sciences. The Global Climate Observing System and the European Space Agency's Climate Change Initiative program have recognized LST as an Essential Climate Variable for the monitoring of the Earth's climate [1], [2]. The availability of global, highly repetitive, long-term, and consistent LST data provided by thermal infrared (TIR) satellite sensors has increased their use in different applications in recent years, including estimation of evapotranspiration and soil moisture, monitoring of agricultural drought, urban heat island, thermal anomalies, and climate change indexes [3]. One of the key points conditioning the use of satellite LSTs is the necessity of a consistent assessment of the LST uncertainty. In addition, continuous evaluation of satellite LST estimates may help to improve and refine existing LST retrieval algorithms.

The committee on Earth observation satellites (CEOSs) compiled a best practice protocol for satellite LST validation [4] where two main validation methods are generally considered: The temperature-based (T-b) and the radiance-based (R-b) methods. T-b is the preferred validation method [4] and consists of a direct comparison of satellite LST with concurrent ground LST measurements in thermally homogenous test sites. The T-b method requires ground TIR radiometers calibrated against traceable standards [5], in situ measurements of surface emissivity and downwelling sky radiance for emissivity correction of radiometric temperatures, and especially that the ground measurements, which correspond to a small field of view (typically few m^2), represent the satellite-scale LST (usually 1 km^2).

Lagouarde et al. [6], [7] showed that atmospheric turbulence might induce high-frequency temporal fluctuations of surface temperature at the ground instrument scale even for single-component homogeneous surfaces. They proposed frequent sampling using 2–3 instruments per site at distances enough to avoid correlation in surface temperatures and averaging during a few minutes around the satellite overpass to reduce the uncertainty in the ground-measured LSTs. Moreover, a comprehensive uncertainty assessment of the ground

© 2024 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

LSTs must be provided [8], showing the significance of the ground data with regard to the satellite LSTs. The uncertainty of ground LST measurements includes calibration and emissivity correction errors, plus spatial-temporal variability of LST within the satellite pixel area at the acquisition time, which can only be accounted for by multiinstrument and frequent sampling and it is usually the largest source of error. The above issues make the collection of ground LSTs a highly demanding task in terms of resources. It is limited to a few cover types, with single-component temperature, spatially uniform from a few meters to 1 km such as lakes, dense vegetation, and bare surfaces [9], [10], [11]. This limits the applicability of the T-b method for global scale validation of satellite LSTs.

As an alternative, the R-b method [12] does not rely on ground LST measurements. Instead, it uses the brightness temperatures in two bands in the 10–12.5 μm atmospheric window, coincident atmospheric profiles of pressure, temperature, and water vapor, and emissivity in the considered bands to simulate the LST through radiative transfer modeling. Strictly, the R-b method cannot be considered as a direct, independent validation in comparison with the T-b method. However, it can be a useful alternative when ground measurements of LST are not viable. The R-b method is affected by several sources of uncertainty, especially the atmospheric profiles used for radiative transfer simulation and the emissivity values. Such uncertainties can be assessed through the difference between LSTs derived in two bands of the 10–12.5 μm window (δ), which should be close to zero in ideal cases. Therefore, we can define a δ threshold value to filter out cases with excessive error [12]. Since emissivity in the 10–12.5 μm window shows small variations among most natural land covers, sites with uniform emissivity are more frequent than thermally homogeneous sites, so the R-b method can be potentially applied to a larger number of sites than the T-b method.

In this article, we focused on the moderate resolution imaging spectroradiometer (MODIS) onboard Terra and Aqua satellites, which currently provides two global, long-term, four times a day LST and emissivity products at 1 km resolution: M*D11_L2 and M*D21_L2 (here M*D11 and M*D21, with MOD for Terra and MYD for Aqua), Collection C6.1. M*D11 is probably the most extensively used and validated satellite LST product and is usually considered a reference for intercomparison with other satellite LST products [3]. The more recent M*D21 LST product has been validated in comparatively fewer studies. The CEOS LST validation protocol [4] defined four validation stages. Stage 1 means that only a small (<30) set of sites and times have been evaluated. Stage 2 implies that a significant set of locations and times have been evaluated as well as the spatial and temporal consistency against similar products at a global scale. Stage 3 is achieved when uncertainties are well quantified and statistically characterized in a robust way over global conditions. Finally, stage 4 requires that stage 3 be updated for new versions and time series. According to the MODIS data web page (https://modis.gsfc.nasa.gov/), the M*D11 LST product is currently in validation stage 2, while M*D21 was recently promoted from stages 1 to 2. Thus, it is worthwhile pursuing the validation of the M*D11 and M*D21 LST products for different biomes and atmospheric conditions in order to promote the validation stage of the products. In order to accomplish this goal, the R-b validation method appears as the most suitable and operational possibility for global LST validation.

The objective of this article is twofold. First, demonstrating the feasibility of the R-b method applied to MODIS data through a comparison with ground LSTs measured in a homogeneous site in order to assess the uncertainty of the method and the threshold for the δ difference. To our knowledge, this is the first study where T-b and R-b LSTs are compared in an experimental test site. Second, showing the operational application of the R-b method for M*D11 and M*D21 LST validation in different test sites encompassing various surface types and atmospheric conditions during a two-year period. Novelties of the article include: 1) a simplified formulation of the R-b method and the δ difference; 2) validation of R-b LSTs with ground-measured LSTs in a thermally homogeneous site; and 3) operational R-b validation of M*D11 and M*D21 LST products, and evaluation of the impact of emissivity inputs.

The article continues with a description of the simplified R-b method and the datasets used in the present study. Section III shows the results of the comparison of R-b LSTs against ground-measured LSTs and the operational validation of M*D11 and M*D21 LSTs. Section IV discusses the results obtained and Section V gives the main conclusions of the study.

II. Method and Materials

A. Simplified R-b Method

The R-b method [12] starts from the top-of-atmosphere (TOA) thermal radiance measured in band i of a satellite sensor, L_i^{sen}, which can be written as

$$L_i^{\text{sen}} = \left[\varepsilon_i B_i\left(T_g\right) + (1-\varepsilon_i)L_i^{\text{sky}}\right]\tau_i(\theta) + L_i^{\text{atm}}(\theta) \quad (1)$$

where subindex i denotes a band-averaged magnitude according to the band's response function, ε_i is the emissivity, T_g is the LST, B_i is the Planck function for blackbody radiance, L_i^{sky} is the downwelling sky irradiance divided by π, $\tau_i(\theta)$, and $L_i^{\text{atm}}(\theta)$ are respectively the atmospheric transmittance and upwelling radiance at the zenith angle θ. If surface emissivity is known and atmospheric transmittance and radiances can be calculated from atmospheric profiles and a radiative transfer model, the reference LST can be obtained from (1) according to

$$T_{ig} = B_i^{-1}\left[\frac{L_i^{\text{sen}} - L_i^{\text{atm}}}{\varepsilon_i\tau_i(\theta)} - \frac{1-\varepsilon_i}{\varepsilon_i}L_i^{\text{sky}}\right] \quad (2)$$

where B_i^{-1} is the inverse Planck function. The LST obtained from (2) is represented by T_{ig} since it may depend on the band used for the retrieval. Typically, the R-b method is applied separately to two bands splitting the 10–12.5 μm atmospheric window (e.g., MODIS bands 31 and 32), hereafter referred to as bands $i = 1$ and 2, respectively.

In an ideal case, T_{ig} should be identical in both bands but they are affected by uncertainties in the emissivity and atmospheric profiles. For most natural surfaces such as water,

vegetation, and soils, emissivity in the 10–12.5 μm window are close to unity and show small variability among surface types [13], [14], [15]. Therefore, emissivity uncertainties can be kept within reasonable limits. More critical is the need for accurate concurrent atmospheric profiles. The accuracy of the radiative transfer calculations can be assessed through the difference in the LST obtained from the two window bands, $\delta = T_{1g} - T_{2g}$. Atmospheric effects in the window region are mainly due to the continuum absorption of water vapor, so they are highly correlated between bands and they are larger at 12 μm than at 11 μm. So, when the atmospheric profile used is overcorrecting the real atmospheric effect in band 1, that is, the difference between the actual and derived LST, $T_g - T_{1g}$, is negative, the overcorrection in T_{2g} will be larger and then $\delta < 0$. The reverse is true in the case of under correction in band 1 ($T_g - T_{1g} > 0$), which will yield to a larger underestimation in band 2 and thus $\delta > 0$. Therefore, the difference $\delta = T_{1g} - T_{2g}$ should be within a small range around zero when atmospheric profiles are accurate and uncertainties in surface emissivities are small.

The δ threshold that determines the valid cases can be obtained through an uncertainty analysis of the various factors involved in the derivation of T_{ig}, so a range of δ values can be defined in order to keep the uncertainty in T_{1g} within certain limits [12], [16]. In this article, we adopted a different strategy using ground-measured LSTs for validating the derived temperatures and establishing the valid δ range.

The simplified R-b method can be summarized in three steps as follows.

1) Using the at-sensor radiances in bands 1 and 2, atmospheric parameters, and emissivity data, the surface temperature is obtained for both bands, T_{1g} and T_{2g} (2).
2) The difference $\delta = T_{1g} - T_{2g}$ is calculated and only cases for which δ is within a small interval around zero are selected as valid cases.
3) For the valid cases, the calculated temperature in band 1, T_{1g}, is considered the reference LST for satellite LST validation. Band 1 is selected because atmospheric correction errors are smaller than in band 2.

In the original formulation [12], δ was defined as the brightness temperature difference ($T_1 - T_2$) obtained from the actual satellite data minus the value calculated from direct (upward) simulation using T_{1g} as input LST in (1). The new δ definition seems more straightforward and relies on the same rationale behind the original definition. In addition, we used an approximation for the band-averaged Planck function, B_i, as an analytical function of temperature (e.g., [8]) so T_{ig} can be readily inverted from (2) and the iterative procedure to calculate T_{ig} [12] is not necessary.

Fig. 1 shows a flow diagram of the methodology applied in this study. Sections II-B–II-D describe the different datasets used.

B. MODIS Data and Atmospheric Profiles

MODIS products were downloaded from https://modis.gsfc.nasa.gov/data/ and include M*D21KM (Level 1B calibrated TOA radiances), M*D03 (Geolocation), and M*D11 and M*D21 LST products, all at 1 km resolution. The M*D11 product uses the generalized split-window method [17] with the brightness temperatures in bands 31 (10.8–11.3 μm) and 32 (11.8–12.3 μm) and a priori classification-based band emissivities. The M*D21 product [18], [19] uses a physics-based temperature-emissivity separation algorithm [20] with the three MODIS TIR bands 29 (8.4–8.7 μm), 31 and 32. The M*D21 algorithm simultaneously retrieves LST and emissivities in the three bands involved. Both products provide the LST, the band emissivities, and a theoretical estimation of the LST uncertainty on a pixel basis.

Two different MODIS datasets were used in this study. First, a limited dataset corresponds to a large and homogeneous rice field in Valencia, Spain where simultaneous ground LST measurements were performed (see Section II-D). The dataset includes 65 Terra cases and 21 Aqua cases from 2002 to 2022, all daytime. They were used to compare the R-b LSTs against independent T-b LSTs to assess the uncertainty of the method and select the δ threshold. Second, an extended dataset covering six selected test sites (see Table I) in the years 2020 and 2021 was used for operational R-b validation with a total of 4267 cases with approximately the same number for Terra and Aqua, daytime, and nighttime in each site. We checked that all pixels covering the test sites were classified as cloud-free and flagged with the highest quality in the auxiliary data.

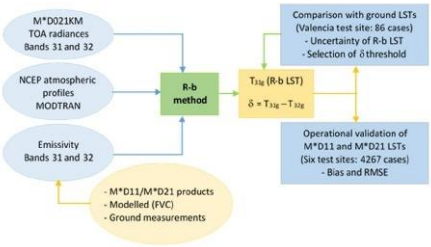

Fig. 1. Flow diagram showing the steps of the methodology used.

TABLE I
TEST SITES FOR R-b VALIDATION, SHOWING THE LOCATION, NUMBER OF SCENES, COMPONENT EMISSIVITIES, AND RANGE OF FVC. *WET BARE SOIL

Site	Longitude, Latitude	scenes	vegetation band 31	vegetation band 32	soil band 31	soil band 32	FVC range
Lake	39°20'24''N, 0°21'00''W	380	N/A	N/A	N/A	N/A	N/A
Rice field	39°16'12'', 0°19'12''W	397	0.985	0.980	0.970 0.972*	0.972 0.977*	0.03-1
Shrubland	39°13'30"N, 0°53'42"W	811	0.985	0.986	0.972	0.977	0.16-0.33
Forest	41°35'31"N, 2°33'00"W	775	0.977	0.977	0.972	0.977	0.27-0.53
Vineyard	39°34'48"N, 1°16'12"W	792	0.972	0.973	0.967	0.959	0.03-0.20
Olive grove	37°49'30"N, 3°58'12"W	1112	0.976	0.976	0.967	0.959	0.09-0.21

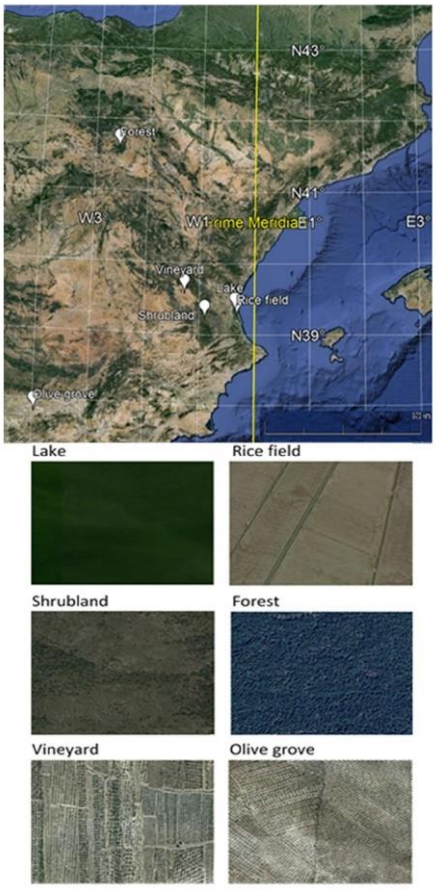

Fig. 2. (Top) Location and (bottom) detailed views of the test sites. The rice field is shown in conditions of dry bare soil.

The test sites are located within large uniform areas containing at least 5 × 5 pixels (3 × 3 for the lake site). Fig. 2 shows the locations of the sites and detailed views of the different covers. For each site and scene, we extracted the four pixels closest to the coordinates of the test site to obtain the satellite value by weighted spatial interpolation.

While the test sites of Table I are restricted to a limited geographical area, they encompass a significant variety of surface covers, from water and nearly full vegetation conditions to partial vegetation covers and practically bare soils. As an indicator of the variety of conditions included, brightness temperatures in band 31 ranged from approximately 280–320 K in daytime, and from 270 to 290 K in nighttime.

We used atmospheric profiles from the National Centers for Environmental Prediction (NCEP) global model. It provides atmospheric profiles over a 1° × 1° grid at 00:00, 06:00, 12:00, and 18:00 UTC, which were spatially and temporally interpolated to the coordinates of each site and the time of sensor overpass. The atmospheric profiles were input in the radiative transfer code MODTRAN 5 [21] to obtain spectral transmittances and radiances. As NCEP profiles only arrive up to 29 km over ground level, upper atmospheric levels were added from the mid-latitude summer and winter standard atmospheres, depending on the date. We added the vertical distribution of CO_2, O_3, and other minor gases included in the MODTRAN code. The atmospheric spectral transmittances and radiances were integrated into the MODIS bands 31 and 32 using the corresponding response functions. The total column water vapor (W) ranged from 0.3 to 3.5 cm in the study period, showing a wide range of mid-latitude atmospheric conditions.

C. Emissivity Data

An important issue of the R-b method is the emissivity inputs required to derive the reference LSTs. For the operational R-b validation, we have considered two options: 1) the emissivity data provided by each LST product, that is, a priori classification-based emissivity for M*D11 and the physics-based derived emissivity for M*D21 and 2) a dynamic emissivity database defined for each site taking into account the nature of the site and available ground or laboratory emissivity measurements for water, vegetation, and soil components.

For partially vegetated surfaces, emissivities were modeled with the vegetation cover method (VCM, [22]) which relies on the fractional vegetation cover (FVC). Since MODIS does not provide an operational FVC product, we used the Level 2 LST product of Sentinel-3, which includes FVC at 1 km as an auxiliary dataset. Sentinel-3 FVC estimates are based on vegetation indexes obtained from the near-infrared bands (0.66 and 0.87 μm) of the Sea and LST Radiometer. FVC values are provided with a temporal resolution of ten days acquired from a moving temporal window of 30-day composites. The FVC values for the MODIS observation dates were obtained by linear interpolation from the closest cloud-free Sentinel-3 dates, usually around three days before the MODIS overpass. Table I shows the approximate range of FVC for the land sites.

The product emissivities are closely related to the LSTs provided by each product, either because they are used as inputs for LST derivation in the M*D11 product or calculated simultaneously with LST in the M*D21 product. Thus, such emissivity data may be considered a favorable case for R-b validation. Since the modeled emissivities are independent of the products, they are useful to check the impact of emissivity on the validation.

Water emissivity and its angular dependence are well known (e.g., [23]), and it is spatially uniform and constant in time.

In the lake site, we applied the parametrization of [24] to calculate water emissivity as a function of observation zenith angle from 0° to 65° in bands 31 and 32. The rice field site is the same as used for ground measurements (Section II-D) showing different covers through the year from water, wet and dry bare soil, and rice crops in various stages with FVC ranging roughly from 0 to 1. For this site, we measured emissivity values in the field for full vegetation and wet and dry bare soil conditions (Table I) using the box method and the TES method applied to multiband field radiometer CIMEL CE-312 [25] as described in [26]. Emissivity measurements were repeated during the campaigns, Table I shows the average emissivity values at nadir for CE-312 bands B3 (10.2–11.0 μm) and B2 (10.9–11.7 μm), which are comparable to MODIS bands 31 and 32, respectively. For water-flooding conditions, we modeled the emissivity in the same way as for the lake site. Field emissivity measurements are not usually available for most land sites, so we can consider the rice field site as an ideal case for R-b validation.

The shrubland and forest are natural evergreen vegetation showing seasonal variability in FVC, with maximum values of 0.3 for the shrubland in spring and 0.5 for the forest in summer, and minimum values over 0.2. The vineyard and olive grove are dryland crops, with minimum FVC values lower than 0.1 and not exceeding 0.2 during the year, so they can be considered representative of arid or semi-arid conditions. For these sites, the band i emissivity ε_i was modeled from VCM according to [15]

$$\varepsilon_i = \varepsilon_{vi} f + \varepsilon_{si}(1-f) + 4(-0.435\varepsilon_{si} + 0.4343)(1-f)f \tag{3}$$

where f is FVC and ε_{vi} (ε_{si}) is the vegetation (soil) component emissivity. Component emissivities are shown in Table I for each site. Emissivity values for the vineyard and shrubland sites were obtained from ground measurements [15], [27], whereas they were taken from [28] for the olive grove sites and from [29] for the forest. Given the emissivity and FVC values of the different test sites (Table I), we have estimated that an uncertainty of 0.1 in FVC leads to an average uncertainty of 0.005 in emissivity according to (3).

The emissivity model of (3) is simple and easy to apply, only depending on FVC and vegetation and soil emissivities that are considered static for each site during the year. However, leaf emissivities depend on the growing conditions, usually decreasing from fresh to dry and senescent stages. Emissivity variations during the growing season depend on the vegetal species and are difficult to know, so they have not been considered in the model. Vegetation emissivities of Table I correspond to full development conditions so the resulting modeled emissivities can be overestimated in some sites during dry and senescent periods.

Fig. 3 shows a comparison of the MOD11, MOD21, and the modeled emissivity values in band 31 (similar results were obtained for Aqua and band 32). For the lake site, MOD11 and modeled emissivities only depend on the zenith angle and agree very well up to 40°; however, MOD11 appears to overestimate emissivity for larger zenith angles. For the land sites, the MOD11 and modeled emissivities show small variability during the year. The modeled emissivity takes into account the different stages of the rice field site (water, bare soil, and vegetation) and changes in FVC for the other land sites resulting in relatively small seasonal variability with maximum values in spring-summer, while MOD11 shows mostly constant emissivity values for each site. In general, the agreement between MOD11 and the modeled emissivity is reasonable for sites with medium to high vegetation cover (rice fields in summer, shrubland, and forest), the largest differences being found for the vineyard and olive grove sites with small FVC. In these cases, MOD11 emissivities are higher and close to vegetation values, whereas the modeled emissivities are closer to those of bare soil due to the low FVC. On the other hand, modeled emissivities for the evergreen shrubland site are somewhat higher than MOD11 emissivities, which are closer to bare soil.

Oppositely, MOD21 emissivities show values smaller than the MOD11 and modeled emissivity data and a large variability with no correlation with the zenith angle in the lake site nor with day of year in the land sites. The reason for the large variability is that MOD21 emissivities are physically retrieved together with LST so they vary from case to case as they are affected by several sources of uncertainty such as instrumental noise, atmospheric correction, and algorithm error [19]. Therefore, we also show in Fig. 3(a) the smoothing of the daily MOD21 emissivities using a moving average over intervals of 5° in viewing angle for the lake site and of ten days for the land sites in order to match the temporal resolution of the FVC data used for the modeled emissivities. According to Fig. 3, the MOD21 emissivity variability appears to be larger in spring-summer (larger atmospheric humidity) and the maximum values are usually close to MOD11 and modeled emissivities. The smoothed emissivities show better consistency with regard to the other emissivity databases and a clearer seasonal dependence than the daily MOD21 product.

The impact of the emissivity inputs on the R-b method can be assessed through a sensitivity analysis of (2). The uncertainty in T_{1g} corresponding to an uncertainty of 0.005 in ε_1 is 0.37 K for the average atmospheric conditions of the dataset. Assuming the same uncertainty for ε_2 and that emissivity uncertainties are uncorrelated in both bands, the impact of emissivity uncertainty alone on δ is 0.52 K. This shows that emissivity is a crucial input, especially, for the calculation of δ, and that R-b validation results depend on the emissivity database used.

D. Ground LST Measurements

The rice field site has been used as a T-b validation site for different TIR sensors since 2002 [10], [16], [26], [27], [30], [31], [32], [33]. In order to assess the accuracy of the R-b derived LSTs with regard to independent ground data, we used 65 cases of ground measurements concurrent with Terra in 2002–2022, and 21 cases concurrent with Aqua in 2016–2022. All cases correspond to daytime, cloud-free conditions with overpass time between 10:10 and 11:40 UTC for Terra and 12:40 and 13:45 UTC for Aqua. Most of the data correspond to full vegetation cover conditions (75%), but there are a few cases for water (13%) and bare soil (12%).

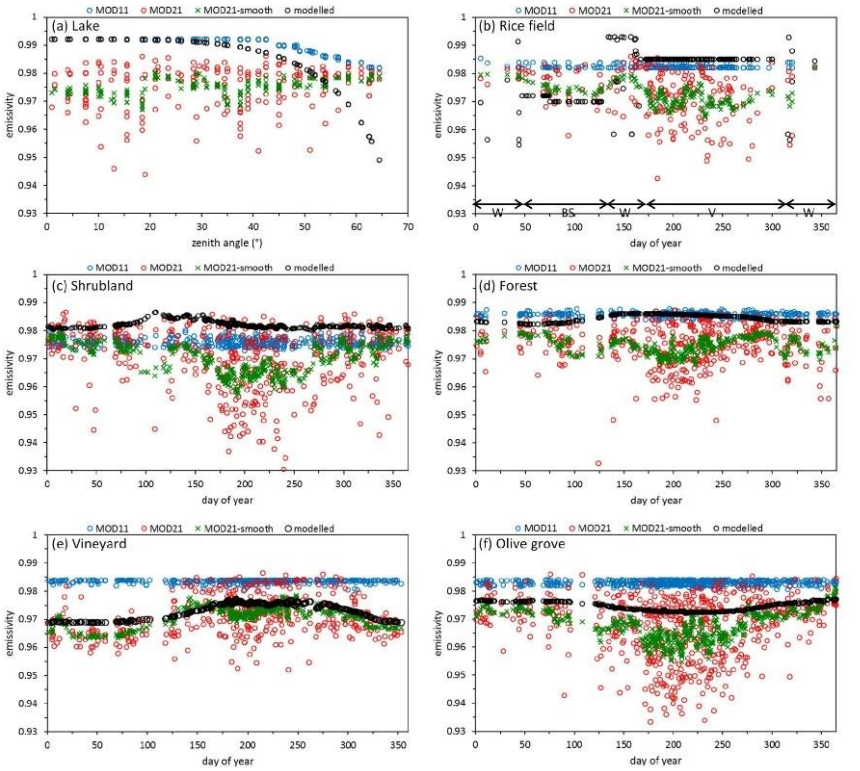

Fig. 3. Emissivity in MODIS band 31 from MOD11 (blue), MOD21 (red), and MOD21 smoothed with moving average (green) and modeled (black) for (a) water as a function of zenith observation angle, (b) rice field with water (W), bare soil (BS) and vegetation (V) stages indicated, (c) shrubland, (d) forest, (e) vineyard, and (f) olive grove as a function of the day of year.

Ground LST measurements were performed using two to four portable radiometers CE-312 [25]. Instruments were calibrated against absolute standards in the TIR instrumentation comparisons held by CEOS in the National Physical Laboratory (NPL) in 2009, 2016 [8], and more recently in 2022 [5]. According to these results, the calibration uncertainty of the CE-312 radiometers was better than 0.20 K in the 273–323 K range. In the NPL comparisons, we also checked the accuracy of the Landcal P80P blackbody (https://www.ametek-land.com), which showed good stability and an error smaller than 0.15 K compared with the NPL standard. The Landcal P80P blackbody source is used for CE-312 calibration in our laboratory on a yearly basis, so the calibration of the radiometers can be regularly updated and reliably estimated for each measurement campaign.

Field measurements were made close to the nadir, so we checked that the corresponding MODIS zenith observation angle was lower than 40° in order to avoid thermal anisotropy effects [6]. The field of view of CE-312 is 10°, which corresponds to a sampling area of about 30 cm in diameter. To assess the variability of the ground temperatures within the satellite pixel, radiometers were assigned to different parts of the test site and carried on along predetermined transects taking measurements each 5–10 m and at a rate of five measurements/min approximately [10], [32].

We corrected radiometric temperatures for surface emissivity effects using ground measurements of emissivity (Table I) and downwelling sky radiance. The final ground LST was calculated as the mean value of the individual LSTs taken by all ground radiometers during 5 min around the satellite

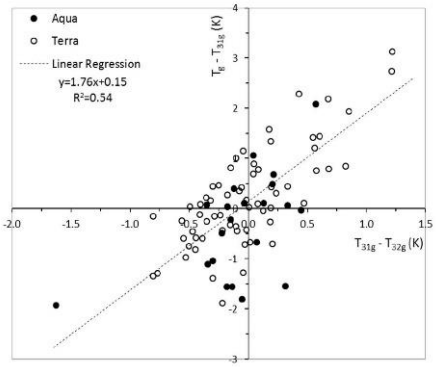

Fig. 4. Difference between measured and R-b surface temperatures in band 31 against the difference in R-b surface temperatures in bands 31 and 32. Terra (Aqua) cases are marked with open (closed) circles. The linear regression and determination coefficient (R^2) are shown.

overpass. The uncertainty of the ground LSTs was estimated considering the following sources of error: 1) calibration of ground radiometers; 2) emissivity measurements; 3) downwelling sky radiance measurements; and 4) spatial-temporal LST variability. The latter was taken as the standard deviation of the LSTs measured by the different radiometers during the 5 min period. Sources (1–3) typically accounted for an error of 0.5 K [8], while the total uncertainty in the ground-measured LST dataset ranged from 0.7 to 1.4 K, the largest source being the spatial and temporal variability.

III. Results

A. Comparison of R-b and Ground-Measured LSTs

We applied the R-b method to the dataset with concurrent ground LST measurements to compare both LSTs. Fig. 4 plots the difference between the ground-measured LSTs (T_g) and R-b reference LSTs (T_{31g}) against the difference $\delta = T_{31g} - T_{32g}$ for all available cases. It shows a good correlation between both magnitudes, with large negative (positive) δ values usually corresponding to large negative (positive) values in $T_g - T_{31g}$, while we have small $T_g - T_{31g}$ values for δ around 0. The linear regression in Fig. 4 yields an acceptable determination coefficient ($R^2 = 0.54$) and an intercept close to zero (0.15 K) as expected. The slope of the linear regression is close to 2, which suggests that the range in δ should be ± 0.5 K to keep the LST uncertainty within ± 1.0 K (comparable with ground measurements). This threshold coincides with that considered by [19] and [34] from theoretical sensitivity analysis.

In Fig. 4, a few points with δ values close to zero have negative values of $T_g - T_{31g}$ in excess of -1.5 K. Such cases (1 Terra and 4 Aqua) correspond to summer conditions (warm and wet atmosphere), with ground LST uncertainties between 1.0 and 1.2 K, so the large differences between the ground and R-b LSTs may be compatible with the combined uncertainty of $T_g - T_{31g}$.

Taking only the cases meeting $-0.5 \text{ K} \leq \delta \leq 0.5$ K, the mean value or bias and root mean square error (RMSE) of $T_g - T_{31g}$ were respectively 0.0 and 0.8 K for Terra (48 valid cases), -0.4 and 0.9 K for Aqua (19 valid cases), and -0.1 and 0.8 K overall. These results show that the accuracy that can be obtained with R-b LSTs is comparable to that of ground-measured LSTs. Therefore, the R-b ground temperatures can be used as an alternative to T-b LSTs for validation when δ is within ± 0.5 K.

B. Application of the R-b Method for Operational Validation

The R-b method was applied to the extended dataset for operational validation of MODIS LST products. Figs. 5 (MOD11) and 6 (MOD21) plot the difference between the product LST (T_{MOD11} or T_{MOD21}) and the R-b LST (T_{31g}) against the difference $\delta = T_{31g} - T_{32g}$ for each test site (daytime and nighttime data combined). In each case, we show the results obtained from the product emissivity and the modeled emissivity. Only Terra data are presented, results for Aqua were mostly similar. Figs. 5 and 6 show a similar relationship between the temperature differences as displayed in Fig. 4. The best correlations were obtained for MOD21 using the product emissivities, where coefficients of determination (R^2) were in the 0.7–0.8 range (Fig. 6). However, they were lower (0.2–0.5) when we used the modeled emissivity data. For the MOD11 product (Fig. 5), R^2 values ranged roughly between 0.3 and 0.5 depending on the site, with a small difference between the product emissivity and the modeled emissivity.

Taking into account only the cases passing the δ condition, we calculated the mean bias and RMSE of the product minus R-b LST differences for M*D11 (Table II) and M*D21 (Table III). Results are shown for each site and all sites combined, and for daytime, nighttime, and all together. Terra and Aqua data are considered separately. Results obtained when using the product and the modeled emissivities are shown.

Tables II and III also show the fraction (in %) of valid cases passing the δ condition for each site and instrument (MOD and MYD) with the product and modeled emissivities. In general, valid cases were more than 70% for all sites and emissivity databases. The only exceptions were the vineyard and olive grove sites with the modeled emissivities (24%–56%). The lower number of valid cases in such sites may be due to the dissimilarity between modeled and product emissivities, especially with regard to M*D11 [Figs. 3(e) and (f), 5(e) and (f), and 6(e) and (f)]. This means that the validation database could be significantly different for the vineyard and olive grove sites when comparing the different emissivity databases, which may affect the validation results. However, the validation databases contain more than 100 data points per site and instrument even in the worst case. It is also noticeable that nighttime datasets generally had a larger fraction of valid cases than daytime datasets, the most favorable being MYD at nighttime.

In Sections III-C and III-D, we describe the validation results of Tables II and III.

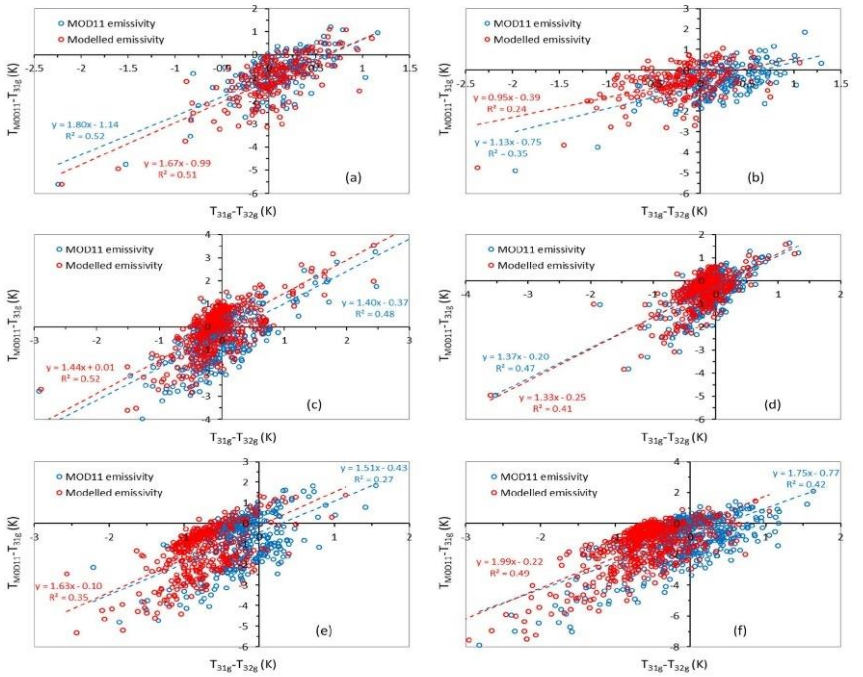

Fig. 5. Difference between MOD11 and R-b LSTs in band 31 against the difference $\delta = T_{31g} - T_{32g}$ for the six test sites when the MOD11 emissivity data (blue) and the modeled emissivity values (red) are used. The linear regression and determination coefficient (R^2) are shown for each case. (a) Lake. (b) Rice field. (c) Shrubland. (d) Forest. (e) Vineyard. (f) Olive grove.

C. M*D11 Validation

Table II shows a good consistency between Terra and Aqua M*D11 LSTs both in the daytime and nighttime. Generally, Aqua provided slightly better results than Terra, but differences in bias and RMSE were below 0.2 K in absolute value for most of the sites, with maximum differences of 0.3 K for the shrubland site and usually within 0.1 K for the other sites. Validation results were mostly similar with either the product or the modeled emissivities, as anticipated in Fig. 5. Considering the results from the product emissivity for all sites, daytime, nighttime, and Terra and Aqua combined, the overall bias was −0.5 K with RMSE of 0.9 K. Per site, biases varied from −0.2 K for the forest to −0.8 K for the lake, while RMSEs ranged from 0.6 K for the forest to 1.1 K for the olive grove. Considering the daytime data only, we found negative biases (product colder than R-b LST) of −1.0 K for all sites combined, ranging between −0.5 K for the forest and −1.4 K for the olive grove. However, the overall bias was 0.0 K for nighttime data, the extreme case being −0.4 K for the lake site and ranging between −0.1 and 0.2 K for the other land sites. The different daytime and nighttime biases of M*D11 can be noticed in Fig. 5, where two clusters of data corresponding to the two conditions are clearly visible, especially for the vineyard and olive grove sites.

D. M*D21 Validation

According to Table III, Terra and Aqua results were very similar as well for the M*D21 using the product emissivity data, with biases and RMSEs within ±0.1 K on average for all sites, in daytime and nighttime. When the modeled emissivities were used, Terra-Aqua differences were slightly larger in some cases, especially for the vineyard and the olive grove at nighttime. It may be because such cases had the lowest number of data points passing the δ threshold as mentioned in Section III-B. However, Terra-Aqua differences were still within ±0.3 K on average. On the other hand, the impact of the emissivity database was larger for the M*D21 product than for M*D11, with the lowest biases and RMSEs being obtained from the product emissivities.

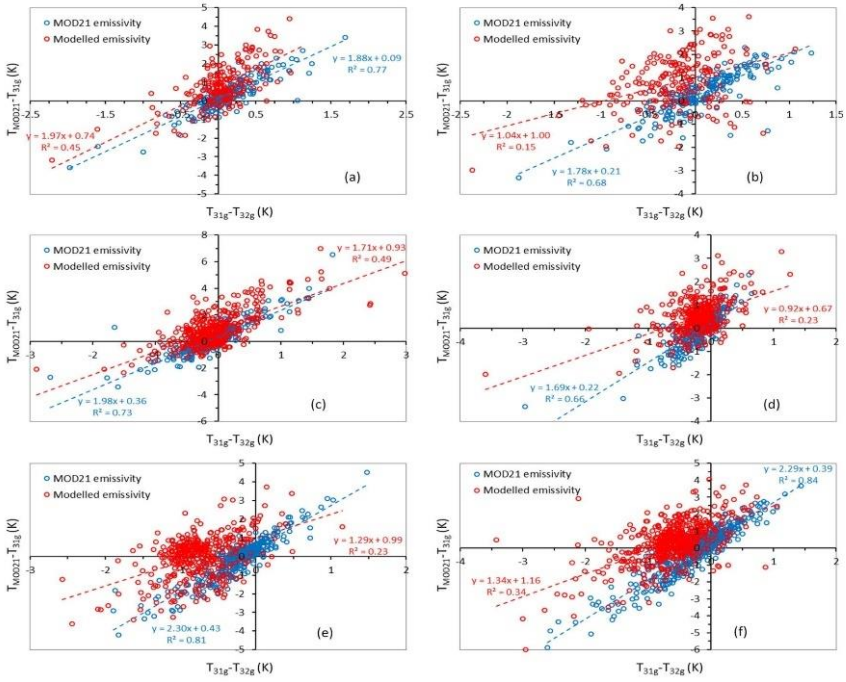

Fig. 6. Difference between MOD21 and R-b LSTs in band 31 against the difference $\delta = T_{31g} - T_{32g}$ for the six test sites when the MOD21 emissivity data (blue) and the modeled emissivity values (red) are used. The linear regression and determination coefficient (R^2) are shown for each case. (a) Lake. (b) Rice field. (c) Shrubland. (d) Forest. (e) Vineyard. (f) Olive grove.

Considering the results with the product emissivities for all sites, all acquisition times and platforms combined, the overall bias was 0.2 K with RMSE of 0.6 K. Considering the sites separately, biases ranged from 0.1 K for the shrubland and forest sites to 0.3 K for the rice field site, while RMSEs ranged between 0.4 K for the forest site and 0.7 K for the lake, rice field and olive grove site. Unlike M*D11, differences between nighttime and daytime results were not significant, the daytime (nighttime) bias being 0.1 K (0.3 K) for all sites, Terra and Aqua combined. When using the modeled emissivities, we can see a uniform increase of about 0.6 K in bias and RMSEs with regard to the product emissivities for most sites.

IV. Discussion

A. Comparison With LST Product Uncertainty

The M*D11 and M*D21 products provide an estimate of the LST uncertainty for each LST value. An example is shown for the olive grove site in Fig. 7(a), where the LST uncertainties of MOD11 and MOD21 LSTs (Terra only) are plotted against the day of the year. The M*D11 LST uncertainty was uniform with a value around 0.5 K for all test sites, daytime and nighttime, Terra, and Aqua data. The M*D21 LST uncertainty was typically larger, with minimum values around 1 K and high variability at the central days of the year where the uncertainty may reach over 3 K, and an average value of 1.3 K for all sites, acquisition times, and platforms.

In Fig. 7(b), the product LST uncertainties are plotted against $W/\cos\theta$, which represents the atmospheric water vapor content in the satellite-viewing path. We can see that the MOD21 LST uncertainty increases clearly with $W/\cos\theta$ since the uncertainties in the atmospheric correction process required for the M*D21 product increase for high atmospheric water vapor loads and surface temperatures, as is the case for summer and early autumn. For the M*D11 uncertainty, there is only a small increase from 0.4 to 0.7 K with atmospheric humidity.

The RMSEs found for M*D11 (Section III-C) are compatible with the product uncertainty for the nighttime data only, while they are larger for daytime due to the relatively large

TABLE II

MEAN BIAS AND RMSE (ALL IN K) OF THE M*D11 PRODUCT MINUS R-b LSTS FOR EACH SITE AND ALL SITES COMBINED WITH THE PRODUCT EMISSIVITY AND THE MODELED EMISSIVITY, FOR DAYTIME, NIGHTTIME AND ALL COMBINED, AND FOR TERRA (MOD) AND AQUA (MYD). THE FRACTION (IN %) OF MOD AND MYD VALID CASES IN EACH SITE IS GIVEN IN PARENTHESES

		M*D11 emissivity		modelled emissivity	
lake		MOD (85)	MYD (84)	MOD (83)	MYD (86)
day	bias	-1.1	-0.8	-1.3	-1.0
	RMSE	1.3	1.0	1.5	1.1
night	bias	-0.5	-0.4	-0.7	-0.5
	RMSE	0.7	0.7	0.8	0.8
all	bias	-0.9	-0.7	-1.1	-0.8
	RMSE	1.1	0.9	1.3	1.0
rice field		MOD (72)	MYD (72)	MOD (77)	MYD (81)
day	bias	-1.0	-0.9	-0.7	-0.6
	RMSE	1.1	1.0	1.0	1.0
night	bias	-0.2	0.0	0.0	-0.1
	RMSE	0.5	0.4	0.5	0.6
all	bias	-0.6	-0.5	-0.5	-0.4
	RMSE	0.9	0.8	0.8	0.9
shrubland		MOD (81)	MYD (84)	MOD (80)	MYD (85)
day	bias	-0.9	-0.6	-0.5	-0.2
	RMSE	1.2	0.9	0.9	0.7
night	bias	0.2	0.1	0.5	0.4
	RMSE	0.4	0.4	0.6	0.6
all	bias	-0.4	-0.2	-0.1	0.1
	RMSE	0.9	0.7	0.8	0.6
forest		MOD (87)	MYD (95)	MOD (85)	MYD (94)
day	bias	-0.6	-0.5	-0.6	-0.5
	RMSE	0.9	0.7	0.8	0.7
night	bias	0.0	0.1	-0.1	0.1
	RMSE	0.4	0.3	0.4	0.3
all	bias	-0.3	-0.1	-0.4	-0.2
	RMSE	0.7	0.5	0.6	0.5
vineyard		MOD (79)	MYD (88)	MOD (24)	MYD (35)
day	bias	-1.4	-1.1	-1.5	-1.1
	RMSE	1.5	1.4	1.6	1.3
night	bias	0.1	0.2	0.3	0.2
	RMSE	0.5	0.5	0.7	0.4
all	bias	-0.6	-0.3	-0.7	-0.5
	RMSE	1.1	0.9	1.3	1.0
olive grove		MOD (76)	MYD (74)	MOD (38)	MYD (56)
day	bias	-1.5	-1.3	-1.5	-1.3
	RMSE	1.7	1.6	1.7	1.5
night	bias	0.0	-0.1	-0.2	-0.3
	RMSE	0.4	0.4	0.4	0.5
all	bias	-0.7	-0.5	-0.8	-0.9
	RMSE	1.2	1.0	1.2	1.1
all sites		MOD (80)	MYD (83)	MOD (60)	MYD (69)
day	bias	-1.1	-0.9	-1.0	-0.8
	RMSE	1.3	1.1	1.3	1.1
night	bias	0.0	0.0	0.0	0.0
	RMSE	0.4	0.4	0.5	0.5
all	bias	-0.6	-0.4	-0.6	-0.4
	RMSE	1.0	0.8	1.0	0.9

TABLE III

MEAN BIAS AND RMSE (ALL IN K) OF THE M*D21 PRODUCT MINUS R-b LSTS FOR EACH SITE AND ALL SITES COMBINED WITH THE PRODUCT EMISSIVITY AND THE MODELED EMISSIVITY, FOR DAYTIME, NIGHTTIME AND ALL COMBINED, AND FOR TERRA (MOD) AND AQUA (MYD). THE FRACTION (IN %) OF MOD AND MYD VALID CASES IN EACH SITE IS GIVEN IN PARENTHESES

		M*D21 emissivity		modelled emissivity	
lake		MOD (79)	MYD (78)	MOD (83)	MYD (86)
day	bias	0.2	0.3	0.8	1.0
	RMSE	0.6	0.8	1.2	1.4
night	bias	0.2	0.3	0.9	0.7
	RMSE	0.6	0.7	1.2	1.0
all	bias	0.2	0.3	0.8	0.9
	RMSE	0.6	0.7	1.2	1.3
rice field		MOD (81)	MYD (80)	MOD (77)	MYD (81)
day	bias	0.2	0.3	1.0	1.1
	RMSE	0.7	0.8	1.6	1.7
night	bias	0.4	0.4	0.8	0.6
	RMSE	0.6	0.5	1.2	0.9
all	bias	0.3	0.3	0.9	0.9
	RMSE	0.7	0.7	1.4	1.5
shrubland		MOD (85)	MYD (90)	MOD (80)	MYD (85)
day	bias	0.2	0.3	1.0	1.0
	RMSE	0.7	0.6	1.4	1.3
night	bias	0.0	0.0	0.2	0.3
	RMSE	0.5	0.3	0.6	0.6
all	bias	0.1	0.1	0.7	0.6
	RMSE	0.6	0.5	1.1	1.0
forest		MOD (89)	MYD (98)	MOD (85)	MYD (94)
day	bias	0.0	0.1	0.5	0.9
	RMSE	0.5	0.4	0.9	1.1
night	bias	0.2	0.1	0.5	0.4
	RMSE	0.4	0.2	0.7	0.5
all	bias	0.1	0.1	0.5	0.6
	RMSE	0.5	0.3	0.8	0.8
vineyard		MOD (80)	MYD (91)	MOD (24)	MYD (35)
day	bias	-0.2	0.1	-0.1	0.1
	RMSE	0.6	0.7	1.1	1.0
night	bias	0.4	0.3	1.3	0.5
	RMSE	0.6	0.4	1.6	0.8
all	bias	0.1	0.2	0.5	0.3
	RMSE	0.6	0.5	1.4	0.9
olive grove		MOD (75)	MYD (87)	MOD (38)	MYD (56)
day	bias	-0.1	0.0	0.7	0.6
	RMSE	0.7	0.8	1.3	1.2
night	bias	0.5	0.4	0.9	0.4
	RMSE	0.7	0.5	1.3	0.7
all	bias	0.3	0.2	0.8	0.5
	RMSE	0.7	0.6	1.3	1.0
all sites		MOD (81)	MYD (89)	MOD (60)	MYD (69)
day	bias	0.0	0.2	0.6	0.8
	RMSE	0.6	0.7	1.2	1.2
night	bias	0.3	0.2	0.8	0.5
	RMSE	0.6	0.4	1.1	0.7
all	bias	0.2	0.2	0.7	0.6
	RMSE	0.6	0.5	1.2	1.0

cold biases obtained in most sites. For the M*D21 product, the RMSEs obtained here with both the product and the modeled emissivity are smaller than the product uncertainty for most sites in daytime and nighttime.

B. Impact of Emissivity on R-b Validation

For M*D11 (Table II), validation results were slightly better for the product emissivities than for the modeled emissivities, but the difference was rather small, with changes within ± 0.2 K in biases and ± 0.1 K in RMSEs for most of the sites, all acquisition times, and platforms. This is because the two emissivity databases were similarly defined through classification and show comparable values (Fig. 3). The impact of emissivity was larger for M*D21, the best validation results being obtained from the product emissivity (overall RMSE of 0.6 K). This shows the consistency of the LST and emissivities

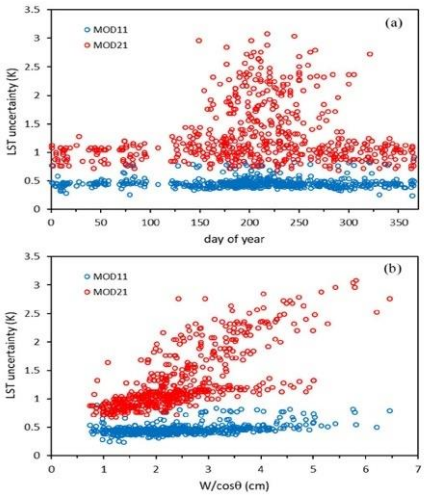

Fig. 7. Uncertainty of the LST products MOD11 (blue) and MOD21 (red) for the olive grove site as a function of (a) day of the year and (b) water vapor content in the satellite viewing path $W/\cos\theta$. All data points available for the site are included.

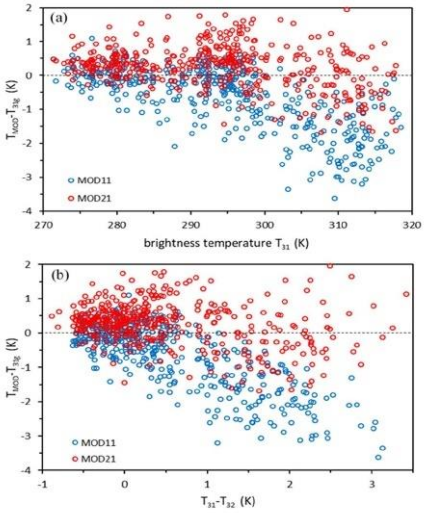

Fig. 8. Difference between the MODIS LST product and the R-b reference temperature T_{31g} for MOD11 (blue) and MOD21 (red) in the olive grove site as a function of (a) MODIS brightness temperature in band 31 and (b) brightness temperature difference $T_{31} - T_{32}$. Only data points passing the δ condition are included.

of the M*D21 product, which are simultaneously obtained from the physics-based TES method. As noted before, the bias and RMSE increased when using the independently modeled emissivities. For the overall dataset, the bias was 0.6 K and RMSE was 1.1 K, still showing an acceptable performance. Per site, biases ranged from 0.4 K (vineyard) to 0.9 K (rice field) and RMSEs from 0.8 K (forest) to 1.4 K (rice field). The increase in bias is explained since, according to Fig. 3, the modeled emissivities are typically larger than M*D21 emissivities, thus yielding smaller R-b LSTs. As with the product emissivities, no significant differences were found between daytime and nighttime biases.

*C. M*D11 Daytime Bias*

Section III-C shows a cold daytime bias of about −1.0 K on average for M*D11. The daytime bias may be related to the underestimation of product LSTs at high temperatures, as shown in Fig. 8(a) where the difference $T_{\mathrm{MOD11}} - T_{31g}$ (with T_{31g} obtained from the product emissivities) is plotted against the brightness temperature T_{31} for the olive grove site. Results for the other sites and MYD11 data were comparable. We can see that the negative bias occurs for brightness temperatures higher than 295 K, which correspond mostly to daytime. In Fig. 8(b) the differences $T_{\mathrm{MOD11}} - T_{31g}$ are plotted against the brightness temperature differences $T_{31} - T_{32}$, which play an important role in the M*D11 split-window algorithm. Typically, large values of $T_{31} - T_{32}$ correspond to daytime cases in summer. According to Fig. 8(b), the underestimation of T_{MOD11} appears to increase linearly with the brightness temperature difference. This may indicate an ill-tuning of the split-window coefficients in the M*D11 algorithm for high surface temperatures and atmospheric water vapor contents.

Water emissivity is uniform, constant, and well-known, so the modeled emissivities should be close to the real value. However, daytime and nighttime biases were considerable (−1 and −0.5 K, respectively in Table II). Part of the negative bias could be due to the overestimation of M*D11 emissivity for viewing angles higher than 40° shown in Fig. 3(a). Taking into account only the cases with $\theta < 40°$, the daytime bias decreased to −0.8 K, while the nighttime bias increased to −0.7 K. The similar negative biases can be due to water temperatures being in the 295–300 K range for more than 50% of both daytime and nighttime cases analyzed here.

Oppositely, the M*D21 product showed no significant difference in daytime and nighttime biases. In Fig. 8, we plotted as well the differences $T_{\mathrm{MOD21}} - T_{31g}$ against T_{31} and $T_{31} - T_{32}$ for the olive grove site showing no correlation. Dispersion increased for higher values of T_{31} and $T_{31} - T_{32}$, which typically correspond to spring and summer with humid atmospheres, in concordance with the M*D21 product uncertainty patterns shown in Fig. 7. Results were mostly similar for the other sites and MYD data.

D. Comparison of R-b Validation Results With Previous Studies

Wan [35] validated the product M*D11 using the R-b method and found biases mostly within ±0.6 K for several desert sites in North Africa for daytime, nighttime, Terra, and Aqua data. While daytime cases correspond to high temperatures (up to 320 K), low atmospheric humidity conditions make the brightness temperature difference $T_{31} - T_{32}$ negative (see [35, Tables 3 and 4]) so LST biases should be close to zero according to Fig. 8(b). Duan et al. [36] applied the R-b method to M*D11 data showing mean biases between −0.4 and 0.7 K and RMSEs smaller than 1.0 K. They found similar daytime and nighttime biases, oppositely to the results shown here. R-b validation of MYD11 over sand dunes in China [34] showed negative biases in daytime and nighttime (−1.3 and −0.7 K, respectively for all sites combined). However, biases varied from site to site, with daytime (nighttime) values about −3 and −4 K (−1 and −3 K) for three sites and better than −1.0 K (−0.5 K) for the other sites. The large cold bias of MYD11 LST was attributed to an overestimation of emissivity for the three sites in the MYD11 product. However, emissivity in MODIS bands 31 and 32 shows small variability among different types of soils [13], [14], [15] and cannot totally account for the large observed biases. Perez-Planells et al. [15] analyzed MYD11 daytime LSTs for shrubland and vineyard sites using different sets of emissivity inputs and showed a similar temperature-depending bias as in Fig. 8(a), regardless of emissivity. In addition, the difference between nighttime and daytime bias in [34] is similar to that reported here and may be related to high temperature and atmospheric humidity conditions prevailing in daytime.

In [19], the R-b method was applied to the M*D21 product over desert dunes, vegetated surfaces, and water, using emissivity data from ground measurements. They showed biases within ±0.9 K and RMSEs from 0.5 to 1.5 K depending on the site for daytime and nighttime data combined. More recently, [34] showed R-b results for MYD21 over sand dunes using ground-measured emissivities. Biases and RMSEs were quite consistent among sites and between daytime and nighttime, with a mean bias of 0.5 K (0.7 K) and RMSE of 1.1 K (0.9 K) for all daytime (nighttime) cases. Dune sites in desert atmospheric conditions are likely an optimal case for the MOD21 algorithm because of the high spectral contrast of sand emissivity between bands 31–32 and 29, and the low impact of atmospheric water vapor [19]. The validation sites of the article include cases of vegetated and water surfaces with near gray-body spectrum in humid atmospheres, which are the less favorable cases for M*D21. Nevertheless, the results shown in Table III with the modeled emissivities are comparable to those of [19] and [34] for all sites considered.

E. Comparison of R-b and T-b Validation

The R-b results obtained here compare well with previous T-b validation results of [19] with daytime and nighttime ground measurements in Lake Tahoe and Salton Sea. For MYD11, they found a bias (RMSE) of 0.2 K (0.5 K) in Lake Tahoe (low-temperature range) and −1.1 K (1.5 K) in Salton Sea (high-temperature range). For MYD21, results were 0.7 K (0.9 K) and 0.2 K (1.2 K), respectively. Such error figures are comparable as well to the R-b validation results of [19] mentioned earlier. In the same way, [30] showed T-b validation for MOD11 and MOD21 daytime data over the same rice field site used here for R-b validation, with bias (RMSE) of 0.1 K (0.6 K) for MOD11 and 0.5 (0.6 K) for MOD21.

Several studies on T-b validation of satellite LSTs rely on ground measurements performed by single, tower-mounted pyrgeometers (flux-meters) in meteorological stations instead of radiometers (directional radiance-meters) [34], [37], [38]. Pyrgeometer measurements represent a broadband, hemispherically integrated value while satellite measurements are narrow-band and of directional nature, depending on the zenith and azimuth viewing and solar angles. Therefore, they are not necessarily comparable, especially in the case of sparse vegetation areas. Moreover, climatological stations are not usually located in thermal homogeneous areas so it is not straightforward to assign the ground-measured LST to the whole satellite pixel.

The effect of thermal heterogeneity at the subpixel scale was analyzed in [38] through the standard deviation of 11 × 11 pixels of 90-m ASTER temperatures coincident with a MODIS pixel. However, single-point ground measurements correspond to areas of a few meters and there may be still a large heterogeneity from the ground scale to the 90-m scale. Daytime results showed M*D11 RMSEs larger than 2 K for most sites and biases from −0.1 to −4.7 K, while nighttime RMSEs were <2 K for all sites. Duan et al. [38] attributed the big daytime LST inaccuracies to the thermal heterogeneity of the sites. The selected cases where the ASTER LST standard deviation was <1 K provided the best results, with close to zero bias and RMSE < 1.3 K. For comparison, typical daytime standard deviations of ASTER LSTs for the rice field site used here were about 0.5 K [10].

Li et al. [34] showed T-b validation for MYD11 and MYD21 LSTs using ground measurements in four desert sites carried out by single, mast-mounted instruments with no information provided about the thermal heterogeneity at the sub-pixel scale. Daytime MYD11 LSTs showed large cold biases close to −3 K and RMSE close to 4 K for the four sites on average. For nighttime MYD11 LSTs, the average bias reduced to −1.4 K and the RMSE to 1.7 K. Li et al. [34] attributed such cold biases to emissivity problems with MYD11 in barren surfaces, where they found MYD11 emissivities overestimating MYD21 emissivities by 0.01. While this may account for a part of the nighttime bias, it cannot explain alone the large daytime bias and RMSEs observed. Such errors include a considerable component due to sub-pixel thermal heterogeneity, possible algorithm problems with high temperatures mentioned before, and probably thermal directional effects from comparing nadir ground measurements with satellite measurements at different zenith angles. The T-based validation of MYD21 showed daytime bias of −0.5 K and RMSE of 2.5 K and nighttime bias of 0.2 K and RMSE of 1.0 K. We can see that the daytime bias was reduced considerably, but still, the RMSE is large and may contain heterogeneity and anisotropy issues. The nighttime MYD21 results were

optimum and only comparable to the R-b validation results of [34] discussed in Section IV-D.

V. Conclusion

The R-b method was used for the validation of M*D11 and M*D21 LST products. First, we applied the R-b method to a set of Terra and Aqua daytime data to derive the R-b reference LSTs that were compared with a dataset of coincident ground-measured LSTs in a homogeneous rice field. The comparison showed an agreement between R-b and T-b LSTs better than 1.0 K in RMSE (i.e., similar to typical uncertainty in ground measurements) when the δ difference was within ± 0.5 K. Then, we applied the R-b method to six test sites including water, homogeneous rice paddy, and partially vegetated areas with FVC ranging from nearly zero to 0.5. All sites were at mid-latitudes but still covering a wide range of LST (270–320 K), and atmospheric humidity (0.5–3.5 cm). For each LST product, two emissivity datasets were used for R-b calculations: the corresponding product emissivity and the modeled emissivity using component emissivity measurements and FVC.

According to the results shown here, daytime M*D11 LSTs showed an average bias of -1.0 K (product LST colder than reference LST) and RMSE of 1.2 K for all sites, Terra and Aqua data combined, with either the product or the modeled emissivities. The daytime cold bias was attributed to the unfitting of the MOD11 LST algorithm for high temperatures and atmospheric humidities. For nighttime data, the overall M*D11 LST bias was 0.0 K (ranging from -0.4 to 0.2 K among sites) and RMSE was 0.5 K. M*D21 LSTs showed great stability among sites and between daytime and nighttime, with daytime (nighttime) bias of 0.1 K (0.3 K) and RMSE of 0.7 K (0.5 K) for all sites, Terra and Aqua data combined, when the product emissivities were used. For the modeled emissivities, bias, and RMSE increased consistently for all sites, times of day, and platforms, but still provided acceptable overall bias (0.6 K) and RMSE (1.1 K).

For a meaningful validation of LST products, the uncertainty of the LSTs used as a reference (either T-b or R-b) must be fully characterized and reliably estimated. In the R-b method, uncertainties mainly include atmospheric correction errors coming from inaccurate atmospheric profiles, radiative transfer model errors, and emissivity input uncertainties. Emissivity has an important impact, especially on the δ calculations. For bare surfaces (rock, soil, sand), emissivity and its spectral variation within the 10–12.5 μm band depend on many factors such as composition, moisture, and texture, so uncertainties may be larger than for vegetated or partially vegetated surfaces. Atmospheric correction errors also have an effect on the R-b LSTs, which tend to be larger for warm, humid atmospheres, and high surface temperatures. However, the impact of the error sources can be limited to a useful range through the δ condition. In the T-b method, sources of error include calibration of ground radiometers, emissivity measurement, and most importantly, the LST spatial-temporal variability within the pixel scale. The sub-pixel LST variability can only be fully assessed through multispatial and frequent temporal sampling in single-component thermal homogeneous areas. On the other hand, ground LST measurements are usually performed at the nadir while satellite viewing zenith angles may exceed 60°, so there may be directional effects, especially for partially vegetated surfaces. These issues limit the application of the T-b method, especially in the daytime.

The T-b method must be considered as the reference, independent validation method for satellite LSTs, so long-term and accurate ground-LST measurements in highly homogeneous test sites are necessary. However, the R-b method provides a valuable alternative for most worldwide surface covers and atmospheric conditions, and the accuracy of the R-b derived LSTs can be checked against ground-measured LSTs in selected test sites as in the present article. However, more studies are required for different ground covers in warm and humid tropical environments where atmospheric effects may be larger thus being a challenging case for satellite LST retrieval.

References

[1] *The 2022 GCOS ECVs Requirements (GCOS 245)*, GCOS, Geneva, Switzerland, 2022, pp. 1–244.

[2] R. Hollmann et al., "The ESA climate change initiative: Satellite data records for essential climate variables," *Bull. Amer. Meteorolog. Soc.*, vol. 94, no. 10, pp. 1541–1552, Oct. 2013, doi: 10.1175/BAMS-D-11-00254.1.

[3] Z. Li et al., "Satellite remote sensing of global Land Surface Temperature: Definition, methods, products, and applications," *Rev. Geophys.*, vol. 61, no. 1, pp. 1–77, Mar. 2023, doi: 10.1029/2022rg000777.

[4] P. Guillevic et al., "Land Surface Temperature product validation best practice protocol. Version 1.1," in *Good Practices for Satellite-Derived Land Product Validation*, P. Guillevic, F. Göttsche, J. Nickeson, and M. Román, Eds., Land Product Validation Subgroup (WGCV/CEOS), 2018, p. 58, doi: 10.5067/doc/ceoswgcv/lpv/lst.001.

[5] Y. Yamada et al., "CEOS international thermal infrared radiometer comparison: Part I: Laboratory comparison of radiometers and blackbodies," *J. Atmos. Oceanic Technol.*, vol. 41, pp. 295–307, Mar. 2024.

[6] J.-P. Lagouarde, M. Irvine, and S. Dupont, "Atmospheric turbulence induced errors on measurements of surface temperature from space," *Remote Sens. Environ.*, vol. 168, pp. 40–53, Oct. 2015, doi: 10.1016/j.rse.2015.06.018.

[7] J.-P. Lagouarde, M. Irvine, and P. Guillevic, "New insights of ground-based Land Surface Temperature measurements protocols for improving validation of thermal infrared satellite data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2019, pp. 9184–9187, doi: 10.1109/IGARSS.2019.8898312.

[8] C. Coll et al., "Laboratory calibration and field measurement of Land Surface Temperature and emissivity using thermal infrared multiband radiometers," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 78, pp. 227–239, Jun. 2019, doi: 10.1016/j.jag.2019.02.002.

[9] S. J. Hook, R. G. Vaughan, H. Tonooka, and S. G. Schladow, "Absolute radiometric in-flight validation of mid infrared and thermal infrared data from ASTER and MODIS on the Terra spacecraft using the lake tahoe, CA/NV, USA, automated validation site," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1798–1807, Jun. 2007, doi: 10.1109/TGRS.2007.894564.

[10] C. Coll et al., "Ground measurements for the validation of Land Surface Temperatures derived from AATSR and MODIS data," *Remote Sens. Environ.*, vol. 97, no. 3, pp. 288–300, Jun. 2005, doi: 10.1016/j.rse.2005.05.007.

[11] F.-M. Göttsche, F.-S. Olesen, I. Trigo, A. Bork-Unkelbach, and M. Martin, "Long term validation of Land Surface Temperature retrieved from MSG/SEVIRI with continuous in-situ measurements in Africa," *Remote Sens.*, vol. 8, no. 5, p. 410, May 2016, doi: 10.3390/rs8050410.

[12] Z. Wan and Z.-L. Li, "Radiance-based validation of the V5 MODIS land-surface temperature product," *Int. J. Remote Sens.*, vol. 29, nos. 17–18, pp. 5373–5395, Sep. 2008, doi: 10.1080/01431160802036565.

[13] W. C. Snyder, Z. Wan, Y. Zhang, and Y.-Z. Feng, "Classification-based emissivity for Land Surface Temperature measurement from space," *Int. J. Remote Sens.*, vol. 19, no. 14, pp. 2753–2774, Jan. 1998, doi: 10.1080/014311698214497.

[14] A. C. T. Pinheiro, J. L. Privette, R. Mahoney, and C. J. Tucker, "Directional effects in a daily AVHRR Land Surface Temperature dataset over Africa," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 9, pp. 1941–1954, Sep. 2004, doi: 10.1109/TGRS.2004.831886.

[15] L. Pérez-Planells, R. Niclòs, E. Valor, and F.-M. Göttsche, "Retrieval of land surface emissivities over partially vegetated surfaces from satellite data using radiative transfer models," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5003821, doi: 10.1109/TGRS.2022.3224639.

[16] C. Coll, Z. Wan, and J. M. Galve, "Temperature-based and radiance-based validations of the V5 MODIS Land Surface Temperature product," *J. Geophys. Res., Atmos.*, vol. 114, no. 20, pp. 1–15, Oct. 2009, doi: 10.1029/2009jd012038.

[17] Z. Wan and J. Dozier, "A generalized split-window algorithm for retrieving land-surface temperature from space," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 4, pp. 892–905, Jul. 1996, doi: 10.1109/36.508406.

[18] G. C. Hulley and S. J. Hook, "Generating consistent Land Surface Temperature and emissivity products between ASTER and MODIS data for Earth science research," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 4, pp. 1304–1315, Apr. 2011, doi: 10.1109/TGRS.2010.2063034.

[19] G. Hulley, N. Malakar, T. Hughes, T. Islam, and S. Hook, "Moderate resolution imaging spectroradiometer (MODIS) MOD21 Land Surface Temperature and emissivity algorithm theoretical basis document (ATBD)," JPL, Pasadena, CA, USA, Tech. Rep. JPL Publication 12-17, Mar. 2016.

[20] A. Gillespie, S. Rokugawa, T. Matsunaga, J. S. Cothern, S. Hook, and A. B. Kahle, "A temperature and emissivity separation algorithm for Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) images," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 4, pp. 1113–1126, Jul. 1998, doi: 10.1109/36.700995.

[21] A. Berk, G. P. Anderson, P. K. Acharya, and E. P. Shettle, "MODTRAN5. 2.0. 0 user's manual," Air Force Res. Lab., Spectral Sci., Burlington, MA, USA, 2008.

[22] E. Valor and V. Caselles, "Mapping land surface emissivity from NDVI: Application to European, African, and South American areas," *Remote Sens. Environ.*, vol. 57, no. 3, pp. 167–184, Sep. 1996, doi: 10.1016/0034-4257(96)00039-9.

[23] X. Wu and W. L. Smith, "Emissivity of rough sea surface for 8–13 μm: Modeling and verification," *Appl. Opt.*, vol. 36, no. 12, p. 2609, Apr. 1997, doi: 10.1364/ao.36.002609.

[24] R. Niclòs, V. Caselles, E. Valor, C. Coll, and J. M. Sánchez, "A simple equation for determining sea surface emissivity in the 3–15 μm region," *Int. J. Remote Sens.*, vol. 30, no. 6, pp. 1603–1619, Mar. 2009, doi: 10.1080/01431160802541523.

[25] M. Legrand, C. Pietras, G. Brogniez, M. Haeffelin, N. K. Abuhassan, and M. Sicard, "A high-accuracy multiwavelength radiometer for in situ measurements in the thermal infrared. Part I: Characterization of the instrument," *J. Atmos. Ocean. Technol.*, vol. 17, no. 9, pp. 1203–1214, Sep. 2000.

[26] R. Niclos et al., "Evaluation of Landsat-8 TIRS data recalibrations and Land Surface Temperature split-window algorithms over a homogeneous crop area with different phenological land covers," *ISPRS J. Photogramm. Remote Sens.*, vol. 174, pp. 237–253, Apr. 2021, doi: 10.1016/j.isprsjprs.2021.02.005.

[27] R. Niclòs, J. M. Galve, J. A. Valiente, M. J. Estrela, and C. Coll, "Accuracy assessment of Land Surface Temperature retrievals from MSG2-SEVIRI data," *Remote Sens. Environ.*, vol. 115, no. 8, pp. 2126–2140, Aug. 2011, doi: 10.1016/j.rse.2011.04.017.

[28] G. Sepulcre-Cantó, P. J. Zarco-Tejada, J. C. Jiménez-Muñoz, J. A. Sobrino, E. D. Miguel, and F. J. Villalobos, "Detection of water stress in an olive orchard with thermal remote sensing imagery," *Agricult. Forest Meteorol.*, vol. 136, nos. 1–2, pp. 31–44, Jan. 2006, doi: 10.1016/j.agrformet.2006.01.008.

[29] S. B. Idso, R. D. Jackson, W. L. Ehrler, and S. T. Mitchell, "A method for determination of infrared emittance of leaves," *Ecology*, vol. 50, no. 5, pp. 899–902, Sep. 1969, doi: 10.2307/1933705.

[30] C. Coll, V. García-Santos, R. Niclòs, and V. Caselles, "Test of the MODIS Land Surface Temperature and emissivity separation algorithm with ground measurements over a rice paddy," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 3061–3069, May 2016, doi: 10.1109/TGRS.2015.2510426.

[31] R. Niclòs, L. Pérez-Planells, C. Coll, J. A. Valiente, and E. Valor, "Evaluation of the S-NPP VIIRS Land Surface Temperature product using ground data acquired by an autonomous system at a rice paddy," *ISPRS J. Photogramm. Remote Sens.*, vol. 135, pp. 1–12, Jan. 2018, doi: 10.1016/j.isprsjprs.2017.10.017.

[32] R. Niclòs, M. Perelló, J. Puchades, C. Coll, and E. Valor, "Evaluating Landsat-9 TIRS-2 calibrations and Land Surface Temperature retrievals against ground measurements using multi-instrument spatial and temporal sampling along transects," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 125, Dec. 2023, Art. no. 103576, doi: 10.1016/j.jag.2023.103576.

[33] L. Pérez-Planells et al., "Validation of Sentinel-3 SLSTR Land Surface Temperature retrieved by the operational product and comparison with explicitly emissivity-dependent algorithms," *Remote Sens.*, vol. 13, no. 11, p. 2228, Jun. 2021, doi: 10.3390/rs13112228.

[34] H. Li et al., "Temperature-based and radiance-based validation of the collection 6 MYD11 and MYD21 Land Surface Temperature products over barren surfaces in northwestern China," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 2, pp. 1794–1807, Feb. 2021, doi: 10.1109/TGRS.2020.2998945.

[35] Z. Wan, "New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product," *Remote Sens. Environ.*, vol. 140, pp. 36–45, Jan. 2014, doi: 10.1016/j.rse.2013.08.027.

[36] S.-B. Duan, Z.-L. Li, H. Wu, P. Leng, M. Gao, and C. Wang, "Radiance-based validation of Land Surface Temperature products derived from collection 6 MODIS thermal infrared data," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 70, pp. 84–92, Aug. 2018, doi: 10.1016/j.jag.2018.04.006.

[37] H. Li et al., "Evaluation of the VIIRS and MODIS LST products in an arid area of Northwest China," *Remote Sens. Environ.*, vol. 142, pp. 111–121, Feb. 2014, doi: 10.1016/j.rse.2013.11.014.

[38] S.-B. Duan et al., "Validation of Collection 6 MODIS Land Surface Temperature product using in situ measurements," *Remote Sens. Environ.*, vol. 225, pp. 16–29, May 2019.

César Coll received the B.Sc., M.Sc., and Ph.D. degrees in physics from the University of Valencia, Valencia, Spain, in 1988, 1992, and 1994, respectively.

He has been a Full Professor of earth physics with the Department of Earth Physics and Thermodynamics, Faculty of Physics, University of Valencia, since 2010, where he is currently the Director of the remote sensing master's degree. He has published more than 70 scientific articles in international journals and 100 conference papers. His research interest focuses on the physical processes and applications of thermal-infrared (TIR) remote sensing and ground validation of land surface temperature (LST) and emissivity.

Raquel Niclòs received the B.Sc., M.Sc., and Ph.D. degrees in physics from the University of Valencia (UV), Valencia, Spain, in 2000, 2002, and 2005, respectively.

She has more than 20 years of research experience in Earth physics. She is currently an Associate Professor with the Department of Earth Physics and Thermodynamics, UV, since 2017. She has been a Supervisor of seven Ph.D. theses, the Principal Investigator of four research and development and innovation (R&D&i) projects in competitive calls and five contracts with outstanding companies, and a member of five international scientific committees. She has published more than 45 articles in journals including the Journal Citation Reports (JCR), with 90% (50%) in the top tercile (top decile) positions. She has also published more than 100 works as non-JCR papers, book chapters, and conference proceedings. She was an Editor of an international book about thermal remote sensing advances and holds a patent. Her research interests mainly focus on thermal-infrared (TIR) remote sensing, atmospheric and emissivity corrections, calibration of TIR sensors, and validation of satellite products of the land and sea surface temperatures, but also other related climate essential variables, meteorological and climate applications, and trends and impacts associated with climate change.

Jesús Puchades received the B.Sc. degree in electronic telecommunications engineering and the M.Sc. degree in remote sensing from the University of Valencia, Valencia, Spain, in 2015 and 2020, respectively, where he is currently pursuing the Ph.D. degree in remote sensing.

As a member of the Thermal Remote Sensing Group, University of Valencia, he was involved in several Spanish research projects and land surface temperature (LST) validation campaigns. His research interests include the retrieval and validation of land surface temperature/emissivity from satellite data.

Martín Perelló received the B.Sc., M.Sc., and Ph.D. degrees in physics from the University of Valencia (UV), Valencia, Spain, in 2014, 2015, and 2020, respectively.

He is a member of the Thermal Remote Sensing Group, University of Valencia, where he has participated in international and Spanish land surface temperature (LST) validation campaigns. His research interests include the calibration and validation of LST from satellite data and the study of LST trends associated with climate change.

Vicente García-Santos received the B.Sc., M.Sc., and Ph.D. degrees in physics from the University of Valencia, Valencia, Spain, in 2007, 2010, and 2013, respectively.

He is currently an Assistant Lecturer at the University of Valencia. He has authored more than 27 articles published in international journals and more than 50 conference papers. His research interests include the physical processes of thermal infrared remote sensing with special attention to the retrieval and validation of land surface temperature and emissivity products, characterization and correction of atmospheric effects, and estimation of energy fluxes exchange between surface and atmosphere.

Lluís Pérez-Planells received the B.Sc. degree in physics, the M.Sc. degree in remote sensing, and the Ph.D. degree in remote sensing from the University of Valencia, Valencia, Spain, in 2014, 2015, and 2021, respectively.

He is a member of the Thermal Remote Sensing Group, University of Valencia. He is currently a Post-Doctoral Researcher with Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, where he is the Main Validation Scientist for the LST_cci project of the European Space Agency (ESA), responsible for validating and intercomparing a broad range of land surface temperature (LST) satellite products generated within the project. He is also a Team Member of the Copernicus LAW project, for which he developed software for continuously analyzing and monitoring in situ LST from five new LST validation stations, i.e., by evaluating the associated uncertainties and the calibration and stability of the deployed field radiometers. He was involved in several Spanish research projects and land surface temperature (LST) validation campaigns. His research focused on the modeling of thermal-infrared (TIR) radiative transfer and the emissivity of vegetated surfaces.

ANEXO F

Ground Calibration of Senintel-3 SLSTR Thermal Infrared Bands and Validation of LST Algorithms

El siguiente trabajo ha sido enviado para su revisión por pares a la revista ISPRS Journal of Photogrammetry and Remote Sensing. Esta revista tiene un factor de impacto de 10.6 según Journal Citation Reports en la edición de 2024.

Manuscript - Word version is mandatory in the absence of Latex file

Click here to view linked References

Ground Calibration of Sentinel-3 SLSTR Thermal Infrared Bands and Validation of LST Algorithms

AUTHORS: Jesús Puchades[1], Raquel Niclòs[1], Lluís Pérez-Planells[2], César Coll[1], Frank-M. Göttsche[2], Jose A. Valiente[3] and Enric Valor[1]

E-mail: jesus.puchades@uv.es; raquel.niclos@uv.es; lluis.perez@kit.edu; cesar.coll@uv.es; frank.goettsche@kit.edu; valiente_jospar@gva.es; enric.valor@uv.es

1) Thermal Remote Sensing Group, Department of Earth Physics and Thermodynamics, Faculty of Physics, University of Valencia, 50, Dr. Moliner, E-46100 Burjassot, Spain
2) IMK-ASF, Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany
3) Fundación Centro de Estudios Ambientales del Mediterráneo (CEAM), 14 Charles R. Darwin, E-46980 Paterna, Spain

Corresponding author: Jesús Puchades (E-mail: jesus.puchades@uv.es)

Abstract: Land Surface Temperature (LST) is an important quantity for studying environmental changes and monitoring the Earth climate system. The SLSTR on board Sentinel-3A and 3B satellites is a key sensor to provide LST on a global scale with daily revisit. In order to assure good data accuracy, it is necessary to check the sensors on-board calibration and to validate LST retrieval algorithms. This work presents the first assessment of the calibration of bands 8 and 9 of the SLSTR sensor with ground LSTs obtained from narrow-band, directional radiometer measurements from nine sites representing four different biomes. The operational SLSTR LST product and an emissivity-dependent split-window algorithm (SWA) were validated using the same in-situ LST database (t-based method). Additionally, the radiance-based (r-based) validation method was applied to both LST algorithms over the same sites and other five homogeneous sites without ground LST measurements. R-RMSDs (Robust Root Mean Square Difference) between simulated and satellite brightness temperatures for both sensors (3A and 3B) were 0.7 K and 0.6 K for band 8 and 9, respectively, with small bias, which are well within the uncertainty limits of the simulated brightness temperatures. The validation of the operational SLSTR product and the SWA yielded consistent results when comparing Sentinel-3A and 3B. For the t-based method, the SLSTR product showed overall daytime (nighttime) R-RMSD of 1.5 K (1.7 K), with mean biases (satellite minus ground) exceeding the recommended ±1.0 K threshold for desert (-2.0 K) and irrigated croplands (1.5 K) at daytime and nighttime. For the SWA, R-RMSDs were 1.3 K both for daytime and nighttime, the desert site at daytime being the only case with excessive bias (-1.5 K). The r-based results confirmed the large biases in the above mentioned cases, showing overall daytime (nighttime) R-RMSDs of 1.0 K (1.0 K) for the LST product and of 0.6 K (0.3 K) for the SWA, which are lower than for the t-based validation. The

35 results presented here contribute to the global validation of LSTs derived from the SLSTR sensor
36 onboard Sentinel-3A and 3B satellites.