Combined land surface emissivity and temperature estimation from Landsat 8 OLI and TIRS

Quinten Vanhellemont

Royal Belgian Institute of Natural Sciences, Operational Directorate Natural Environments, Vautierstraat 29, 1000 Brussels, Belgium

ARTICLE INFO

ABSTRACT

Remote sensing of Land Surface Temperature (LST) generally requires atmospheric parameters and the emissivity (ϵ) of the target to be estimated. The atmospheric up- and downwelling radiances and transmittance can be accurately modelled using radiative transfer models and profiles of relative humidity and temperature, either measured by radiosonde probes or retrieved from assimilating weather models. The estimation of ϵ is a large source of uncertainty in the resulting LST product, and there are various approaches using multi-angle observations, multispectral optical or multispectral thermal infrared imagery. In this paper, the estimation of LST from the Thermal InfraRed Sensor (TIRS) on board Landsat 8 is evaluated using more than 6 years of in situ temperature measurements from a network of 14 Autonomous Weather Stations (AWS) in Belgium. ϵ is estimated from concomitant atmospherically corrected imagery from the Operational Land Imager (OLI) using two new neural network approaches trained on ECOSTRESS spectra, and an established NDVI based method. Results are compared to using $\epsilon = 1$ and the ASTER Global Emissivity Dataset. LST retrievals from L8/TIRS perform well for all emissivity data sources for > 500 matchups with AWS subsoil temperature measurements: Mean Differences 0.8–3.7 K and unbiased Root Mean Squared Differences of 2.9–3.5 K for both B10 and B11. The use of unity emissivity gives the best results in terms of MD (0.8 K) and unb-RMSD (3 K). Similar ranges of unb-RMSD are found for > 500 matchups with broadband radiometer temperatures (2.6–3.1 K), that have lower absolute MD values (-2.2–0.6 K). For the radiometer temperatures, both the neural net approaches gave lowest MD, in the best case ± 0.1 K. The present investigation can hence recommend the neural nets to derive ϵ for the retrieval of LST over the AWS in Belgium. Using published matchup results from other authors however, no single source of ϵ data performed better than $\epsilon = 1$, but this could be due to their low number of matchups. Further efforts for estimating representative pixel average emissivities are needed, and establishing a denser in situ measurement network over varied land use, with rather homogeneous land cover within a TIRS pixel, may aid further validation of a per pixel and per scene ϵ estimates from multispectral imagery. AWS data seems valuable for evaluation of satellite LST, with the advantage of a much lower cost and higher potential matchup density compared to conventional radiometers.

1. Introduction

Satellite observations in the thermal infrared (TIR) part of the spectrum, with typical wavelengths in the atmospheric windows between 3–5 and 8–15 μm are used to retrieve surface temperatures (ST) for land (LST) and water (WST) targets (Wan et al., 2004; Wan, 2014; Sekertekin and Bonafoni, 2020; Vanhellemont, 2020). As surface temperature drives heat exchange between the surface and the atmosphere the remote estimation of LST can be a very powerful tool in understanding climate and ecological processes. In the aquatic environment, WST can be an important driver of nutrient mixing and can structure phytoplankton communities (Wiltshire et al., 2008; Trombetta et al.,

2019). Water temperature in lakes and reservoirs may be a driver for the occurrence of harmful algal blooms and WST could be used in predictive models (Schaeffer et al., 2018; Wynne et al., 2013). In terrestrial applications, soil moisture and evapotranspiration for example can be inferred from satellite estimated LST (Sandholt et al., 2002; Anderson et al., 2012), and the remotely sensed LST could hence be crucial for synoptic monitoring of drought conditions (Karnieli et al., 2010; Wan et al., 2004) and crop health (Anderson and Kustas, 2008; Sholihah et al., 2016). Thermal imagery from MODIS is routinely used for active fire detection (Justice et al., 2002; Giglio et al., 2003), and higher resolution data e.g. from Landsat has been used for the monitoring of underground coal fires (Mansor et al., 1994; Kuenzer et al.,

2007) and geothermal activity (Lagios et al., 2007; Mia et al., 2014; Mia et al., 2017; Chan et al., 2018). High resolution satellite derived LST can also be used to study the effects of urban heat islands (Li et al., 2011; Liu and Zhang, 2011) and heatwave risks (Dousset et al., 2011; Buscail et al., 2012).

The TIR radiance observed by satellite sensors is a function of atmospheric and surface (emissivity and kinetic temperature) properties. Parameters needed for the atmospheric correction of top-of-atmosphere TIR (atmospheric transmittance, τ, downward and upward atmospheric radiances, L_d and L_u) can be reliably estimated using radiative transfer simulations and atmospheric profiles of relative humidity and temperature, either from radiosondes or assimilated weather models (Barsi et al., 2005; Pérez-Planells et al., 2015; Vanhellemont, 2020). Alternative methods using a split window approach (Caselles et al., 1998; Wan, 2014) require multiple TIR bands and knowledge of the water vapour content of the atmosphere. To retrieve the surface temperature, the emissivity needs to be estimated, either simultaneously or from external data sources. The emissivity of a surface is defined as the ratio of the radiance emitted by a surface to the radiance emitted by a black body at the same temperature. The accurate estimation of emissivity is essential, as a 1% error on the emissivity can lead to errors up to 1 K in the derived LST product (e.g. Wan (2014, 2017)). Typically, emissivity is modeled using an image classification method (Snyder et al., 1998), the Normalised Difference Vegetation Index (NDVI) and/or fraction of vegetation cover (FVC) if concurrent multispectral optical imagery is available (Sobrino and Raissouni, 2000; Chen et al., 2016; Sekertekin, 2019; Ren et al., 2017) or imposed from external data sources, such as the MODIS (Wan et al., 2004) or ASTER (Hulley et al., 2015) emissivity products. A review and sensitivity analysis of different methods of emissivity estimation from NDVI is given in recent publications (Gao et al., 2020; Sekertekin and Bonafoni, 2020). MODIS and ASTER have multispectral TIR instruments that can be used to retrieve emissivity using the Temperature Emissivity Separation (TES, Gillespie et al., 1998) or Temperature Independent Spectral Indices of Emissivity (TISIE, Li and Becker, 1993) approaches. A review of TES and TISIE methods is given in Li et al. (2013), but are not discussed further due to the focus on Landsat type instruments in the present paper. There can be significant anisotropy of the emissivity especially for viewing angles > 30° (Ren et al., 2011; Li et al., 2013; García-Santos et al., 2015), which can be ignored for largely nadir-viewing instruments such as those on Landsat.

Water targets are generally used for calibration and validation of atmospheric correction methods and surface temperature retrieval, due to the constant emissivity of liquid water and the generally low spatial heterogeneity of temperature within a satellite pixel (e.g. Barsi et al. (2005)). For water targets, the high resolution thermal imagers on the Landsat series of satellites have generally a better than 1 K Root Mean Squared Difference (RMSD) compared to in situ temperature measurements (Cook et al., 2014; Malakar et al., 2018; Vanhellemont, 2020).

Land pixels generally aggregate information for a large variation of heterogeneous materials (buildings, trees, roads, fields, water). Because the content of a TIR satellite pixel is heterogeneous, the emissivity for a given pixel is an aggregation of the different pixel contents and geometry (Yang et al., 2015), and it may be a challenge to establish an appropriate pixel specific emissivity. Furthermore, land targets are typically much more variable in space and time due to seasonal effects (growth, snow cover, flooding), land use variability (ploughing, crop growth and harvesting), and the emissivity of soils can increase based on moisture content, leading to up to 2 K errors on LST estimates (Mira et al., 2007; García-Santos et al., 2013). As a result of these effects, validation over land sites typically gives larger RMSD compared to in situ measurements compared to water sites, ranging from 2–4 K for Landsat type sensors (García-Santos et al., 2018; Sekertekin, 2019). In a comprehensive validation exercise of LST derived from multiple coarser resolution sensors (AATSR, GOES, MODIS, and SEVIRI), Martin et al. (2019) found daytime accuracy within 4 K, and nighttime accuracy within 2 K.

This paper presents an alternative emissivity estimation from multispectral imagery, using a machine learning approach (neural network) and a spectral library. The method is demonstrated using Landsat 8 (L8) which has on board a 9 band visible to shortwave infrared imager, OLI, and a two band thermal infrared sensor, TIRS. Atmospherically corrected OLI imagery acquired at the same time as the TIRS imagery is used in combination with the neural network to estimate emissivity in both TIRS bands. Emissivity estimates are compared to other emissivity sources, i.e. the ASTER GDEM dataset and an NDVI based approach, and are used to retrieve single band LST from both Bands 10 and 11 in the L8 "Collection 1" Level 1 data using the TIR atmospheric correction method presented in Vanhellemont (2020). Split channel algorithms were not used due to the poor performance found for the coefficients provided by Du et al. (2015) as used by Gerace and Montanaro (2017) after their stray light correction (Vanhellemont, 2020). LST for all emissivity sources are compared to in situ subsoil temperature measurements at 14 sites operated by the Royal Meteorological Institute of Belgium (RMI), and to temperatures derived from broadband radiometers at 12 of the 14 sites. The use of these regular weather monitoring sites, which can provide many more matchups, as well as matchups from radiometers presented by previous studies are evaluated for satellite validation. The varied locations of the different sites (Bertrand et al., 2013) may be differently impacted by seasonal variability in the emissivity estimations.

2. Data and methods

2.1. In situ data

In situ temperature measurements were obtained from 14 AWS operated by the RMI across Belgium (Table 1 and Fig. 1). This dataset

Table 1
Locations of the 14 AWS sites and measured parameters.

| Number	Name	Latitude (°N)	Longitude (°E)	T°C Air	T°C Soil	T°C Soil-5 cm	T°C Grass	T°C Rad
1	BEITEM	50.9050	3.1231	Y	Y	Y	Y	Y
2	ZEEBRUGGE	51.3486	3.1956	Y	N	N	N	N
3	ZELZATE	51.1814	3.8053	N	Y	N	N	N
4	MELLE	50.9808	3.8175	Y	Y	Y	Y	Y
5	STABROEK	51.3256	4.3650	Y	Y	Y	Y	Y
6	ST-KATELIJNE-WAVER	51.0675	4.5306	Y	Y	Y	Y	Y
7	UCCLE-UKKEL	50.7986	4.3581	Y	Y	Y	Y	Y
8	DOURBES	50.0964	4.5947	Y	Y	Y	Y	Y
9	ERNAGE	50.5828	4.6908	Y	Y	Y	Y	Y
10	RETIE	51.2222	5.0283	Y	Y	Y	Y	Y
11	HUMAIN	50.1947	5.2567	Y	Y	Y	Y	Y
12	DIEPENBEEK	50.9164	5.4517	Y	Y	Y	Y	Y
13	BUZENOL	49.6211	5.5886	Y	Y	Y	Y	Y
14	MONT RIGI	50.5117	6.0747	Y	Y	Y	N	Y
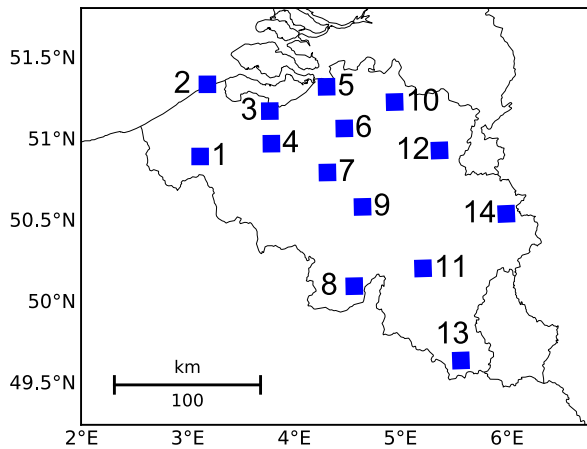
Fig. 1. Locations of the 14 AWS stations operated by RMI in Belgium. The boxes show the approximately 12 × 12 km ROI around the station used in satellite image processing. Station names and coordinates are given in Table 1.

covers the full Landsat 8 mission and includes hourly measurements of air temperature at 1.5 m above the surface, and ground temperatures (1) above grass, (2) above bare soil, and (3) 5 cm below bare soil. Details on the automated quality control of the air and soil temperatures can be found in Bertrand et al. (2013, 2015). Temperatures are measured with a PT100 probe, which has an accuracy of < 0.1 K or < 0.1% of span, and the total accuracy of the system is estimated to be 0.3 K. 12 out of the 14 sites used here also have a Campbell Scientific CNR-1 broadband (0.3–3 μm and 5–50 μm) Net Radiometer with up- and downward facing instruments installed. The irradiance measurement from the downward facing thermal instrument (E, W/m²) can be used to compute the ground surface temperature, T (K), according to the Stefan–Boltzmann law:

$$E = \sigma \cdot T^4 \tag{1}$$

with σ the Stefan–Boltzmann constant, taken here as $5.67 \cdot 10^{-8}$ W/m² K⁴.

Temporal coverage of the individual measurements vary per site, and the closest bounding temperature measurements within one hour of the satellite overpass were linearly interpolated to the overpass time. Soil and air temperatures are measured using contact probes, and those in situ temperature measurements are perhaps not directly comparable to satellite-measured skin temperature, and a better performance is expected for the broadband radiometer-derived temperatures. From the contact thermometer data, the measurements below bare soil are likely most comparable to the satellite measured skin temperature, if the soil is not frozen (Jin and Mullens, 2014), as they are less sensitive to small scale air fluctuations and direct insolation. The air temperature usually differs from the skin temperature due to surface and atmospheric conditions (Jin et al., 1997), and generally a lag is observed between skin temperatures and soil and air temperatures (Jin and Mullens, 2014). The air, grass and soil temperatures are quite sensitive to direct solar heating and are hence not used here. The presence of buildings and trees, and their cast shadows, in the satellite pixel may cause significant differences between the AWS temperatures and satellite measured skin temperatures, as the 120 × 120 m TIRS pixel may not be homogeneous around the site. The landscape in Belgium is very fragmented as a result of its long history and (lax) spatial planning policies (Albrechts et al., 2003), which results in a changing of emissivity at short spatial scales, and likely significant sub-pixel variability, and complex thermal adjacency effects. This provides a challenge for the validation of satellite imagery, but the present investigation may provide performance indicators in the typical use of satellite LST in such a fragmented landscape, even though the AWS are installed on long-term stable meteorological observation sites. The dense time series in time and space provides excellent opportunities for large scale matchup analysis, and

can give an indication whether these types of measurements can be used for evaluation of satellite derived LST. For water applications for example, bulk temperatures are often used for validation, and corrected with a simple offset to obtain equivalent skin temperatures (Donlon et al., 2002; Cook et al., 2014). Furthermore, the installation and maintenance of radiometer systems is more expensive than these AWS or automated contact thermometers, and many national institutes already have an established network of soil and subsoil temperature and irradiance measurements. The use of these AWS is here suggested as addition to, and not a replacement for, validation sites with high performance thermal radiometers such as those in SURFRAD.

2.2. Satellite data

The available "Collection 1" daytime data combining both sensors on Landsat 8, OLI and TIRS, were obtained from Google Earth Engine (GEE, Gorelick et al., 2017). Imagery was provided in standard USGS L1T format, i.e. terrain corrected Top-of-Atmosphere (TOA) radiances/reflectances in GeoTIFF format as produced from the Landsat Product Generation System (LPGS) version 2.7.0 (imagery before October 2017), 13.0.0 (October 2017 - April 2018), and 13.1.0 (imagery after April 2018). OLI is a push broom imager with 9 spectral bands (B1–9) in the visible to near-infrared part of the spectrum, with 8 bands at 30 m and 1 panchromatic band at 15 m spatial resolution. TIRS is a two band push broom imager with bands B10 and B11 centred on 10.9 (10.6–11.2) and 12.0 (11.5–12.5) μm, and records data at 100 m spatial resolution. TIRS data is resampled to 30 m and collocated on the OLI grid. OLI data from the 30 m bands excluding the cirrus band at 1.37 μm were atmospherically corrected using the Dark Spectrum Fitting Atmospheric Correction (DSF, Vanhellemont and Ruddick, 2018; Vanhellemont, 2019) using a fixed path reflectance for a 12 by 12 km Region Of Interest (ROI) centred on the measurement sites. Only scenes with > 75% ROI coverage were retained, i.e. removing edge-of-swath scenes where the station is not in the scene. Quality control was performed separately for the computation of the ROI mean and median reflectance datasets and the matchups. Mean and median reflectance datasets (e.g. Fig. 2) were computed per ROI based on quasi cloud-free images, where a simple cloud filtering was performed using thresholds on the blue and cirrus bands: scenes with > 10% of the ROI with ρ_t 1373 nm > 0.015 or ρ_s 443 nm > 0.15 were removed. Data filtering for the matchups was performed using an additional quality filtering; matchups were excluded when the mean of a 11 × 11 pixel box centred on the station had ρ_t 1373 nm > 0.01 (to exclude cirrus contamination), a maximum visible band (443, 483, 561, 655 nm) ρ_s > 0.12 (to exclude direct cloud cover over the site), and the subscene interquartile range of ρ_s 443 nm > 0.035 (to exclude scenes with scattered clouds).

TOA thermal radiances (L_t) were also extracted for TIRS B10 and B11. Parameters required for the atmospheric correction of the thermal channels were derived using the method presented in Vanhellemont (2020). Subscene center ERA5 reanalysis profiles from the 0.25x0.25 degree data provided by the European Centre for Medium-Range Weather Forecasts (ECMWF, Malardel et al., 2016) and the libRadtran (version 2.0.2) radiative transfer code (Mayer and Kylling, 2005; Emde et al., 2016) were used to compute the atmospheric transmittance (τ), and up- (L_u) and downwelling (L_d) radiances. The TOA thermal radiance is a measurement of the combined surface-emitted radiance (L_s) and atmosphere:

$$L_t = \tau \cdot (\epsilon \cdot L_s + (1 - \epsilon) \cdot L_d) + L_u, \tag{2}$$

and in order to derive land surface temperature, the emissivity (ϵ) per pixel needs to be estimated, which is detailed in the next section.

2.3. Emissivity data

Emissivity was estimated from OLI surface reflectances (ρ_s) using a neural network approach. Reference spectra were obtained from the

Fig. 2. L8/OLI ρ_s RGB composites for the BUZENOL ROI (site 13), cloud-free mission median (left) and images for 2014–05-05 (middle) and 2018–08-04 (right). All three images have the ρ_s R (655 nm), G (561 nm) and B (483 nm) linearly scaled from 0–0.15 to the 8 bit RGB channels. Scene dimensions are approximately 12 by 12 km. The left plot shows the locations of the emissivity time-series in Fig. 8, (1) Field, (2) Pine, (3) Pine cut in 2018.

ECOSTRESS spectral library (Baldridge et al., 2009; Meerdink et al., 2019). 1072 out of the 3403 ECOSTRESS spectra with complete spectral coverage from the visible to thermal infrared (wavelengths 0.3–15 μm) were resampled to the OLI and TIRS relative spectral response functions. A subset of 839 spectra (Fig. 3) from the "manmade", "rock", "soil", "vegetation", and "water" classes (excluding "meteorites", "nonphotosyntheticvegetation", and "mineral") were finally retained to train a neural network (Net1) using the Keras (Chollet, 2015) library and TensorFlow (Abadi et al., 2015) backend. A second network (Net2) was trained using a subset of 561 spectra, also excluding the "rock" class. The input reflectance spectra were normalised using the spectral mean and standard deviation in each OLI band. The neural network has n input neurons, for the ρ_s in the VNIR-SWIR OLI bands, and output neurons, for the ϵ in both TIRS bands. Input neurons either included (n = 8) or excluded (n = 7) the panchromatic channel on OLI, band 9 (cirrus at 1.37 μm) was excluded due to the low atmospheric transmittance. The mean squared error loss function and adam optimiser were used for the output neurons. Various combinations of deep (4–8) and wide (16–256) hidden layers were tested. 1000 training epochs with a batch size of 20 were ran per set up, with datasets randomly split in 90% training and 10% validation data in each epoch.

For each ROI, emissivity was estimated using the selected neural networks from the OLI ρ_s imagery on a scene by scene basis, as well as for the cloud filtered per-pixel median ρ_s over the mission lifetime. Pixels with ρ_s 1609 nm < 0.02 are considered to be water, and their emissivity is fixed to band averaged values from Vanhellemont (2020), 0.9926 and 0.9877 for B10 and B11 respectively. The scene specific

estimates are included in order to accommodate for seasonal (e.g. snow, droughts) or rapid (e.g. crop harvest, ploughing, flooding) changes where an average dataset may be insufficient. The mission lifetime median dataset is likely biased to clear weather (i.e. summer) conditions. In Fig. 2 the differences between the median and scene specific coverage for BUZENOL are illustrated. There are marked differences in field state, where in the scene specific RGB composites many more fields have been recently ploughed and are hence bare soil. The forest appears darker in the median composite than in the spring image, due to the bias to summer condition due to cloud coverage. The differences between pine and deciduous stands are more pronounced in the spring and median images than in the summer image. Fields appear to be more yellow or brown in the summer image than in the spring image, which could indicate changes in soil water conditions. Emissivity data for each ROI were also obtained from the ASTER Global Emissivity Dataset 100-meter V003 (AG100, Hulley et al., 2015), resampled to the OLI 30 m grid using cubic interpolation. The simplified Normalised Difference Vegetation Index (NDVI) approach (Sobrino and Raissouni, 2000; Sobrino et al., 2008; Jiménez-Muñoz et al., 2008) was also used to estimate emissivity:

$$\epsilon = \epsilon_s \cdot (1 - FVC) + \epsilon_v \cdot FVC \qquad (3)$$

with ϵ_s the average soil emissivity (0.971 and 0.977 in B10 and B11) and ϵ_v the average vegetation emissivity (0.987 and 0.989 in B10 and B11) taken from Skoković et al. (2014). FVC is the Fraction of Vegetation Cover (Carlson and Ripley, 1997):
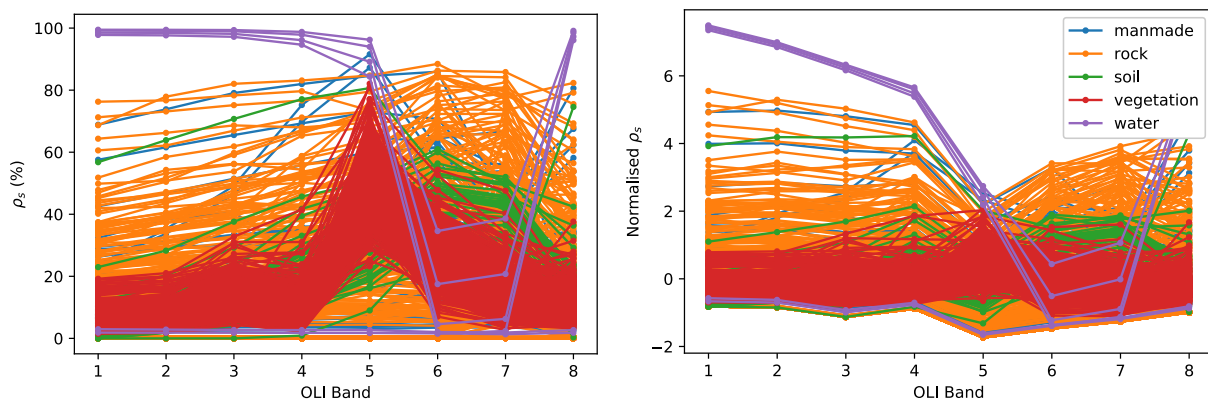

Fig. 3. The 839 spectra from the ECOSTRESS library resampled to OLI bands coloured by class, without (left) and with (right) normalisation. Note that the water class contains both liquid (highly absorbing) and solid (highly reflective) water. Band 8 is the panchromatic band (centre wavelength 592 nm). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

$$FVC = \left[\frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}} \right]^2. \tag{4}$$

where $NDVI_{min}$ and $NDVI_{max}$ give the NDVI ranges used between bare soil (< 0.18) and fully vegetated (> 0.85), and NDVI is computed from OLI red (655 nm) and NIR (865 nm) surface reflectances:

$$NDVI = \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red}}. \tag{5}$$

FVC is set to 0 for $NDVI < NDVI_{min}$ and to 1 for $NDVI > NDVI_{max}$. For FVC $= 0$ the ϵ is estimated separately for B10 and B11 according to Skoković et al. (2014):

$$\epsilon(B10) = 0.979 - 0.046 \cdot \rho_{red}, \tag{6}$$

and

$$\epsilon(B11) = 0.982 - 0.027 \cdot \rho_{red}. \tag{7}$$

Finally, temperatures were also computed using unity emissivity.

2.4. Statistics

Linear correlation statistics and Reduced Major Regression (RMA) lines were computed for the matchups, and the Root Mean Squared Difference (RMSD), Mean Difference (MD), and unbiased RMSD (unb-RMSD) are given as indication of the total error, systematic error (bias), and precision:

$$RMSD = \sqrt{\sum_{i=0}^{n} \frac{(y_i - x_i)^2}{n}}, \tag{8}$$

$$MD = \sum_{i=0}^{n} \frac{y_i - x_i}{n}, \tag{9}$$

$$unb - RMSD = \sqrt{RMSD^2 - MD^2}. \tag{10}$$

3. Results and discussion

3.1. Emissivity estimates

The hidden layer construction of the neural network was varied between widths of 16, 32, 64, 128, and 256 neurons, and depths of 4, 5, 6, 7 and 8 layers. The performance of the neural net was evaluated against the expected reflectance from the ECOSTRESS library, with the 64×4 net giving highest r^2 (87% and 86%) and lowest RMSD (1.19 and 1.02) in both TIRS bands for the first subset of spectra (Fig. 4). The second subset of ECOSTRESS spectra (excluding the "rock" spectra) gave similar performances (r^2 82%, RMSD 1.12–1.13, Fig. 5). The inclusion of the panchromatic band only marginally changed the model performance using the ECOSTRESS library, and hence was excluded from the application to satellite imagery due to the larger uncertainty on the atmospheric correction of the panchromatic band.

The emissivity estimated from the median ROI ρ_s compares well to the emissivity from the ASTER GED (an example for BUZENOL is shown in Fig. 6). Although a larger spread is found for the OLI derived emissivity (e.g. Fig. 7) compared to the ASTER GED, the Mean Average Relative Differences (MARD) are about 1.3–2.6% in B10 and 0.8–2.0% in B11 for Net1 and 0.8–2.2% in B10 and 0.6–1.8% in B11 for Net2, with slight variability between sites. Thanks to the higher OLI resolution (30 m) compared to the ASTER GED (100 m), individual fields or tree stands, and manmade materials may be better resolved spatially. The OLI data range is more recent (2013–2019) than the one used for the ASTER GED (2000–2008), a decade during which land use has changed significantly. There are some differences between the ASTER and TIRS spectral responses that are not taken into account. The OLI derived emissivity is generally a bit lower than the ASTER GED data (maximal MD -0.024 in B10 and -0.013 in B11 for Net1), likely as a

direct result of the spectral response differences and the neural network performance. Compared to the ASTER GED dataset, the different nets performed differently across the sites, depending on the ROI scene contents. Generally the second net gave lower differences with the ASTER GED compared to the first net, except for regions with large forest coverage (e.g. BUZENOL). The lower emissivity estimated from the neural nets will lead to higher LST estimates compared to using unity or ASTER emissivity.

A time-series of B10 emissivity for several pixels in the BUZENOL site is shown in Fig. 8. The median emissivities from both nets are similar to those from the ASTER GED, with Net2 giving emissivities about 0.01 lower than ASTER. The two pine stand pixels show a slight variability of emissivity through the year, with highest values reached during spring and early summer. The second pine stand was cut down in 2018, and the emissivity drops closer to values associated with bare soils or barren terrain. Net1 gives much lower emissivity for the barren pixels, likely since it confuses bare soil with the rock spectra (with generally lower, or more variable emissivity) also included in the model. The NDVI based model generally fluctuates between the imposed soil (0.971) and vegetation (0.987) emissivities, and also clearly shows the bare soil period for the second pine stand. A larger standard deviation in the 3×3 pixel boxes are found for the neural nets, likely due to the fixed emissivity range of the NDVI method, and the coarser resolution of the ASTER GED. For the Field pixel larger variability is found, cycling between tilled and planted fields. The NDVI and Neural Net methods give different results for separate tilling operations: going from crop to tilled in April 2015 decreases the NDVI ϵ but increases the estimate from both Nets, while the tilling operation in April 2017 has the NDVI ϵ going from low to high, and the nets (esp. Net2) give the inverse pattern.

3.2. Surface temperature

Surface temperatures are computed using the simulated up- and down-welling radiance and transmittance of the atmosphere, and six estimates of emissivity: (1) unity, $\epsilon = 1$, (2) ASTER GDEM, (3) per-scene OLI NDVI derived, and (4–7) per scene and ROI-average OLI derived from both neural nets. An example of an LST product for BUZENOL site is shown in Fig. 9. When using unity and ASTER GED emissivity, lower temperatures are found compared to the OLI derived emissivity, which is caused by the OLI products having an overall lower emissivity. The higher resolution of the OLI product compared to ASTER GED or the TIRS native resolution gives a sharper derived temperature product.

Matchups with the AWS data from 12 and 14 sites operated by the RMI were identified, and quality controlled using the OLI reflectance data. The number of potential matchups was decreased from 1972/1848 (scenes with bounding in situ measurements) to 537/501 (scenes passing automated quality control) for the radiometer and subsoil in situ measurements. A summary of the performance is given in Table 2 and 3, with scatter plots in Fig. 10 and 11. A strong linear relationship is found between satellite and in situ measurements, with high correlation coefficients (r^2 0.90–0.95), and near unity slope (0.94–1.04).

For the broadband radiometer derived temperatures, the RMSD are quite close between all the datasets (2.8–3.5 K), with a cold bias (negative MD) for unity, ASTER and NVDI emissivity (-0.9 to -2.2 K). Net1 and Net2 give the lowest absolute MD, with Net1 biased 0.6 K warm for B10 and $\pm$ 0.1 K for B11. For Net2 the MD is $\pm$ 0.1 K for B10 and -0.4 K for B11. The RMSD and MD are directly related to the range in the emissivity dataset used, as the emissivity is the only change between the comparisons. For these matchups, the MD seems to be the distinguishing statistic, as the performance for other statistics is very similar. The unb-RMSD can be slightly higher for the Nets than for the other datasets, due to the spatial variability inherent in the OLI data, and the larger range of emissivities from the Net outputs. Both nets give different emissivity estimates in both bands, and it seems like Net1
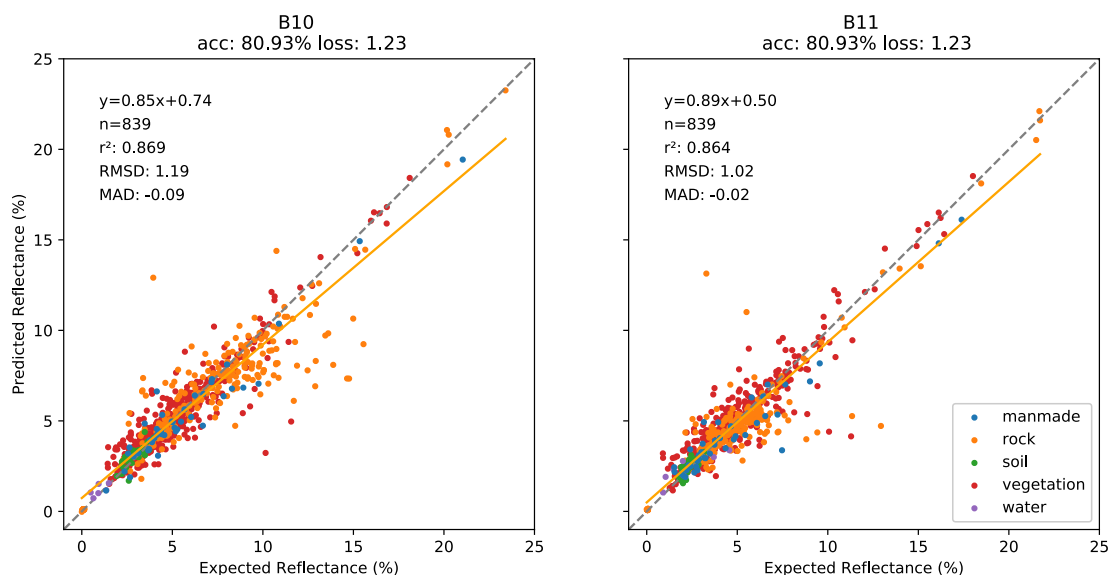
Fig. 4. A comparison of the predicted and expected reflectance outputs from the 7 band 64 × 4 neural network for both TIRS bands, B10 (left) and B11 (right), using the first subset of spectra (Net1). These plots show the results after 1000 training epochs with 90–10% training-validation split in each epoch.

gives a better B11 estimate than Net2 (MD ± 0.1 K compared to −0.4 K), and Net2 gives a better B11 estimate than Net1 (MD ± 0.1 K compared to 0.6 K).

For the subsoil temperature, the RMSD across all emissivity datasets is quite large (3–5 K). The MD (0.8–3.7 K) and hence difference between unb-RMSD and RMSD (generally 0.1–1.5 K) are also larger for the subsoil temperatures than for the radiometers. In terms of scatter, overall performances are quite similar for the radiometer (unb-RMSD 2.6–3.1 K) and subsoil (unb-RMSD 2.9–3.5 K) temperatures. For MD, the performance of the former is better, with the bias with subsoil temperatures about 3 K.

In order to evaluate the performance of the emissivity estimates for in situ measurements with high performance radiometers that measure the skin temperature, the results from Sekertekin (2019) who used twenty cloud-free matchups for four sites in the Surface Radiation Budget Network (SURFRAD) established by NOAA, and the results from García-Santos et al. (2018) who used twenty-one matchups at the University of the Balearic Islands (UIB) are here analysed. Both Sekertekin (2019) and García-Santos et al. (2018) used the Atmospheric

Correction Parameter Calculator (ACPC, Barsi et al., 2003; Barsi et al., 2005) for the atmospheric correction. The former used an NDVI based estimate for emissivity, while the latter used the ASTER GED. Table 4 replicates the results by Sekertekin (2019) for all twenty matchups and for the subset of fifteen they deemed more reliable, and includes results from the present study for unity and scene based emissivity for both Nets. The results for the per scene estimated emissivity compare well with those from Sekertekin (2019), with unb-RMSD of 2.2 K for the full matchup dataset and 1.6–1.7 K for the subset. Surprisingly, the unity emissivity gives the best performance, outperforming both other methods that have a variable emissivity, especially in terms of slope and MD. The lowest unb-RMSD is found for the Net1 estimated emissivity. Overall results from the RMI AWS sites give similar slopes, but with higher unb-RMSD and MD for the subsoil matchups (both by about 1 K). Matchups with AWS radiometer temperature show slightly higher unb-RMSD, but lower absolute MD. Matchups from García-Santos et al. (2018) processed using the present study for unity and scene based emissivity for both Nets are shown in Table 5. García-Santos et al. (2018) found the Single Channel and Split Window Algorithms (SCA
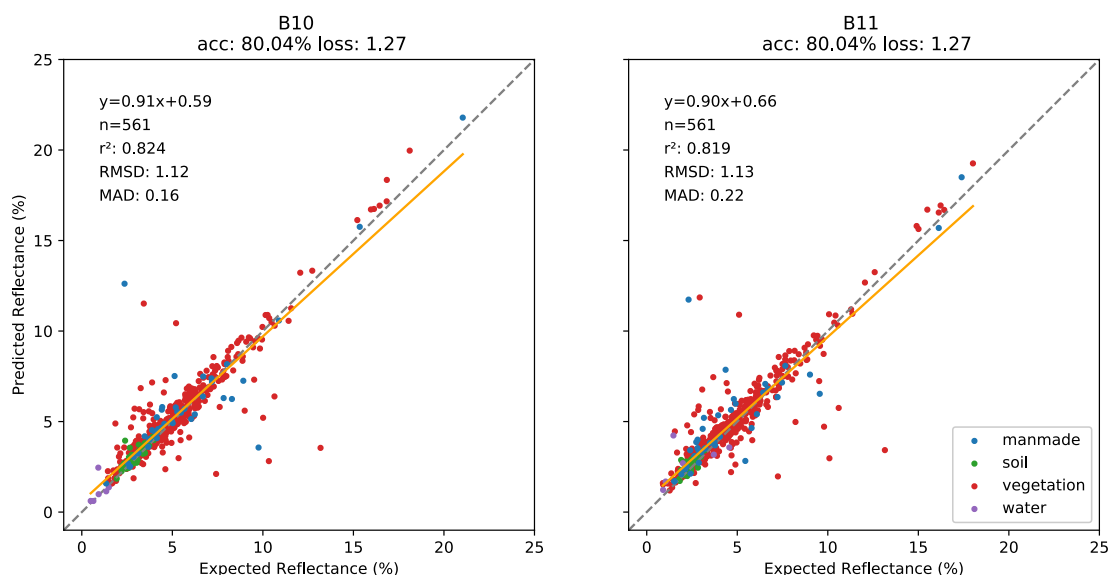

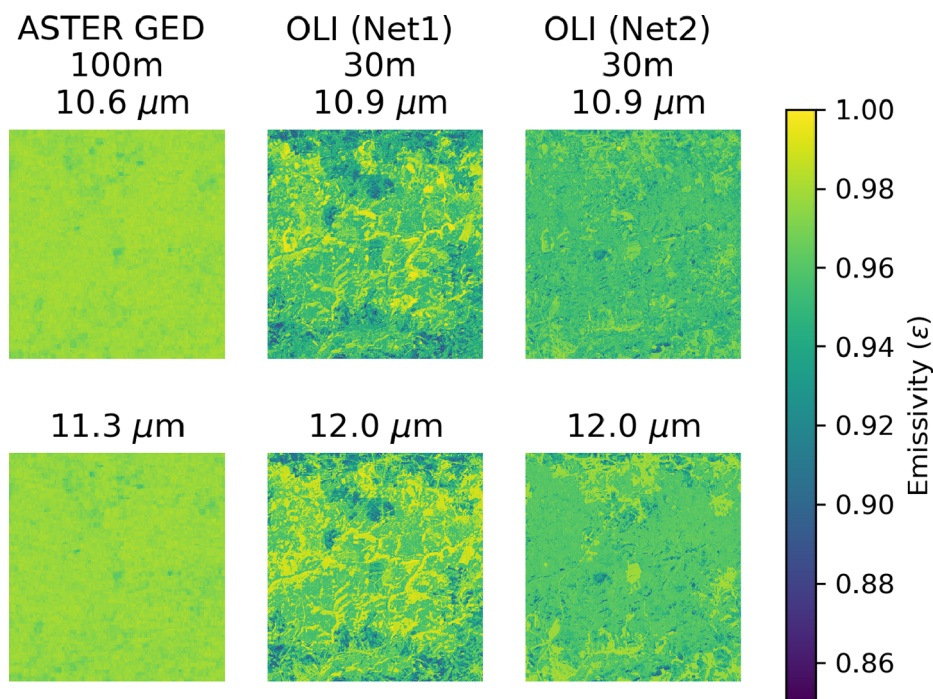**Fig. 5.** Same as Fig. 4 but for the second subset of spectra (Net2), excluding. the rock class.

Fig. 6. A comparison of ASTER GED (left) and OLI median reflectance derived emissivities for Net1 (middle) and Net2 (right) for the BUZENOL site.
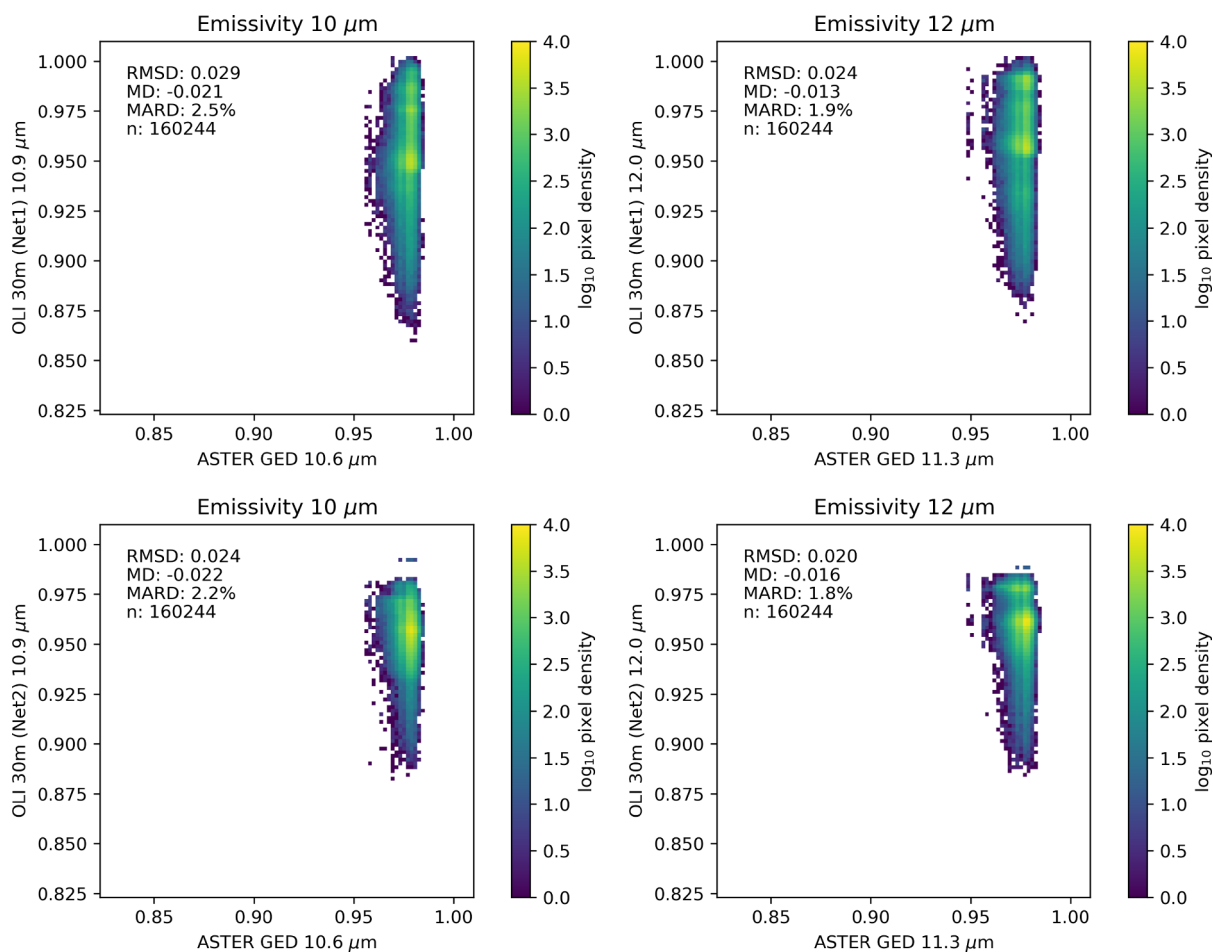

Fig. 7. Density scatter plots of the OLI median surface reflectance derived emissivities as function of ASTER GED emissivities for the BUZENOL study site. Top and bottom rows show the results from Net1 and Net2. Left and right columns show the band 10 and 11 emissivities.
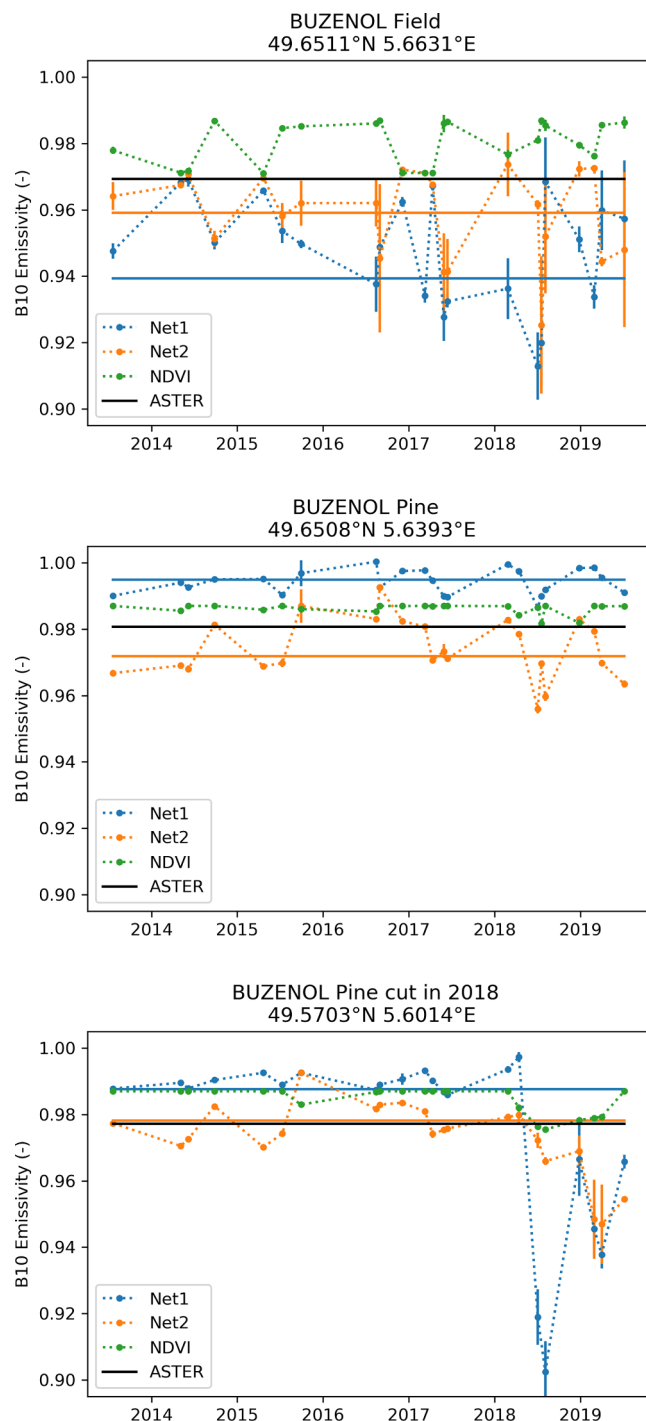
Fig. 8. Timeseries of emissivity from three locations within the BUZENOL site. (Top) Field, (Middle) Pine forest, and (Bottom) Pine forest, cut in 2018. Dashed lines connect the per scene estimates, solid lines give the mission median results. Vertical error bars show the standard deviation within a 3 × 3 30 m pixel box around the selected pixel.
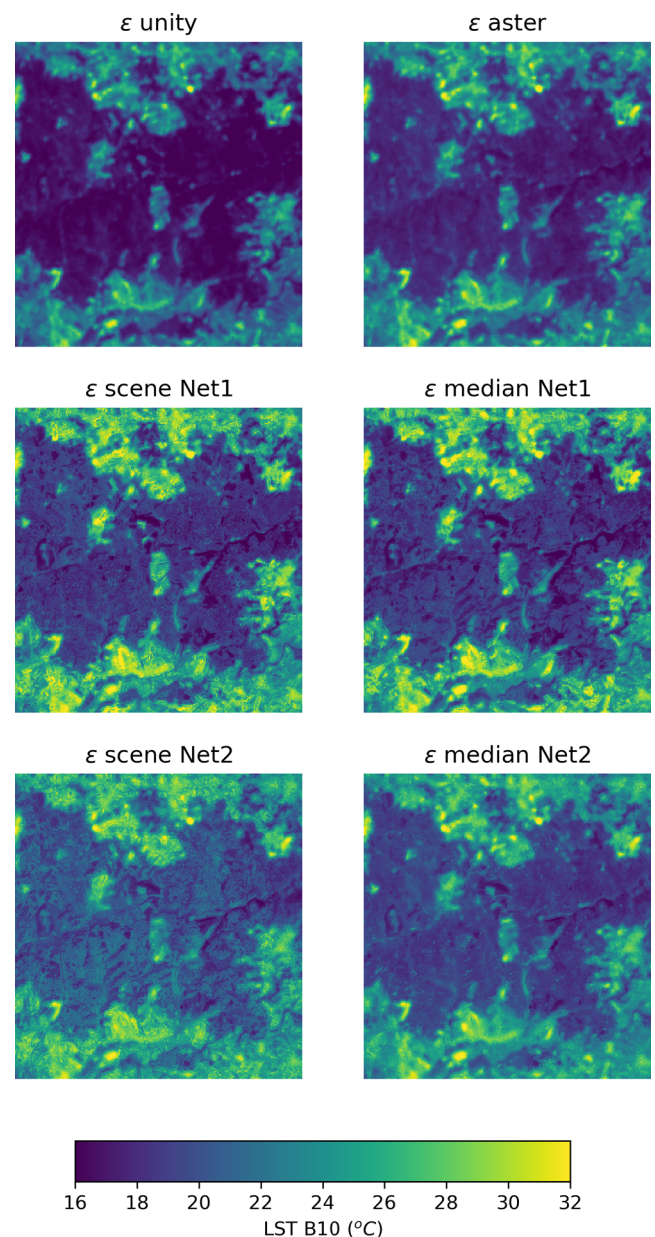

Fig. 9. B10 LST for BUZENOL derived from the L8/TIRS image of 2014–05–05 according to different emissivity settings.

< 1 for these matchups, largely due to an underestimation of high temperatures for unity emissivity, and an overestimation of low temperatures for the neural nets (not shown). Differences in unb-RMSD are quite low, indicating that the different methods have a similar relative noise level, especially methods using a fixed or unity emissivity. The differences between the unb-RMSD and the RMSD reflect the impact of bias originating from the use of different emissivity datasets, also represented in the variability of the MD.

4. Perspectives

Emissivity can be estimated from multispectral imagery and a neural network approach trained on a spectral library. There are clear conceptual advantages of estimating the emissivity per scene, but some contradictory patterns are found between the NDVI approach and both nets, and may depend on scene specific characteristics or surface properties, or the imposed limits of the NDVI approach (Gao et al., 2020). The NDVI approach has been previously found to be unsuitable

and SWA) to outperform the results obtained from the ACPC in terms of RMSD (1.6–2 K and 2.0–2.3 K respectively), but the ACPC giving the lowest bias (-0.1 K). With the results from the present study, lowest unb-RMSD is found again for unity emissivity, but the MD from both neural nets is significantly lower (0.7 K for Net1 and 0.02 K for Net2), similar to what was found for the AWS matchups. Using the ASTER GED emissivity computed by García-Santos et al. (2018), the performance is practically the same for the present method and the ACPC (similar to findings in Vanhellemont (2020)). The linear regression slopes are

Table 2

Summary of the matchups with the broadband radiometer temperature measurements at the Belgian AWS. n is the number of matchups, m and b the slope and offset of the Reduced Major Axis regression, r^2 is the square of Pearsons linear correlation coefficient. RMSD and MD are the Root Mean Squared and Mean Differences between satellite and in situ measurements (satellite - in situ), with unb-RMSD the unbiased RMSD.

Band	Emissivity	n	m	b (°C)	r^2	RMSD (°C)	unb-RMSD (°C)	MD (°C)
LST10	unity	537	0.955	−1.332	0.947	3.431	2.617	−2.219
LST10	aster	537	0.954	−0.030	0.945	2.819	2.661	−0.931
LST10	ndvi	537	0.956	−0.219	0.945	2.877	2.664	−1.088
LST10	scene Net1	537	0.982	1.002	0.926	3.178	3.109	0.656
LST10	median Net1	537	0.955	1.499	0.944	2.772	2.705	0.607
LST10	scene Net2	537	0.950	0.883	0.942	2.757	2.755	−0.095
LST10	median Net2	537	0.957	0.940	0.941	2.757	2.755	0.101
LST11	unity	537	0.945	−1.148	0.939	3.589	2.810	−2.231
LST11	aster	537	0.941	0.296	0.936	3.014	2.887	−0.864
LST11	ndvi	537	0.943	−0.217	0.937	3.151	2.856	−1.331
LST11	scene Net1	537	0.964	0.640	0.920	3.223	3.222	−0.067
LST11	median Net1	537	0.940	1.269	0.934	2.946	2.944	0.088
LST11	scene Net2	537	0.932	0.932	0.933	2.983	2.955	−0.406
LST11	median Net2	537	0.943	0.675	0.933	2.990	2.956	−0.450

for impervious materials (Chen et al., 2016). The performance of the neural nets may be impacted by the imbalance between the number of spectra in the different classes in the training dataset, and both the nets and the NDVI based method can be impacted by errors in the atmospheric correction of OLI imagery. The inclusion of rock spectra in Net1 may be less appropriate for Belgian sites, as they are typically a patchwork of fields, bare soils, trees, and man-made materials. Further sensitivity analyses need to be performed using long term soil temperature measurements in other countries with more varied surface characteristics.

Single band LST derived from L8/TIRS using atmospheric parameters computed using ERA5 reanalysis profiles of relative humidity and temperature was evaluated for 14 sites across Belgium that autonomously measure air, soil, grass and subsoil temperatures, and 12 sites that measure broadband irradiance sky- and ground irradiances. Strong linear correlations were found between LST and all measured temperatures, indicating that satellite-derived LST could be used to model any of the parameters. Excellent performance was found for the matchups of LST with the broadband radiometer data, with for B10 the neural nets giving unb-RMSD of 2.7–3.1 K, and MD 0.6 K for Net1 and ±0.1 K for Net2. Good performance was also found for the subsoil temperatures which gave near unity slopes, and an unb-RMSD of 2.9–3.5 K across bands and emissivity sources, with a 0.8–3.7 K warm bias. (Caused by the in situ subsoil temperatures being colder than the in situ radiometer temperatures.) B10 and B11 matchups gave very similar performances, and it seems the calibration and stray light correction efforts (Gerace and Montanaro, 2017) into the "Collection 1" data make both bands suitable for LST estimation. These results are on

par with other recent validation results from L8/TIRS using various correction methods an in situ thermal radiometers (García-Santos et al., 2018; Sekertekin, 2019; Meng et al., 2019), and the present study illustrates how global networks of AWS could aid in validation of satellite derived LST, by complementing radiometer measurements with higher associated costs.

Although the performance of the LST products compared to in situ measurements was similar across various emissivity datasets: (1) unity emissivity, (2) ASTER GED, (3) NDVI based, and the ECOSTRESS based neural networks developed in the present study, (4) Net1, scene ρ_s, (5) Net1, median ρ_s, (6) Net2, scene ρ_s, and (7) Net2 ρ_s, median, the best performance (lowest MD, unb-RMSD, offsets and unity slopes, with highest r^2) was found for the neural nets in comparison with the AWS radiometer data (with MD ± 0.1 K for Net2), and for unity emissivity for the AWS subsoil data. Similarly, with the matchup dataset with SURFRAD data published by Sekertekin (2019) and UIB data by (García-Santos et al., 2018), and summarised in Tables 4 and 5, the best performance was found for unity emissivity, although they provided a more limited number of matchups. The RMSD values are lower for unity emissivity likely due to the lack of noise introduced by the spatial variability present in the other datasets. Of the neural networks developed in the present paper the net excluding the "rock" spectra gave slightly better performances, i.e. the absolute MD for Net2 is lower than that for Net1 for B10. It should be noted that the AWS are situated in long-term meteorological sites, that have a largely unchanging coverage through the year, which may not fully demonstrate the power of a scene by scene emissivity estimate. The study sites in the present paper have significant vegetation coverage inside and surrounding the TIRS

Table 3

Same as Table 2 but for matchups with the in situ subsoil temperature measurements.

| Band	Emissivity	n	m	b (°C)	r^2	RMSD (°C)	unb-RMSD (°C)	MD (°C)
LST10	unity	501	1.011	0.586	0.927	3.030	2.929	0.777
LST10	aster	501	1.011	1.877	0.927	3.587	2.939	2.057
LST10	ndvi	501	1.013	1.656	0.925	3.520	2.978	1.878
LST10	scene Net1	501	1.040	3.046	0.903	5.086	3.464	3.725
LST10	median Net1	501	1.011	3.412	0.916	4.792	3.159	3.604
LST10	scene Net2	501	1.008	2.802	0.917	4.284	3.121	2.935
LST10	median Net2	501	1.013	2.984	0.922	4.411	3.038	3.197
LST11	unity	501	1.001	0.740	0.919	3.163	3.070	0.758
LST11	aster	501	0.997	2.197	0.918	3.760	3.092	2.140
LST11	ndvi	501	1.000	1.634	0.917	3.518	3.114	1.636
LST11	scene Net1	501	1.024	2.642	0.898	4.652	3.509	3.054
LST11	median Net1	501	0.993	3.320	0.903	4.649	3.367	3.206
LST11	scene Net2	501	0.989	2.815	0.909	4.183	3.254	2.628
LST11	median Net2	501	0.998	2.651	0.913	4.124	3.186	2.619
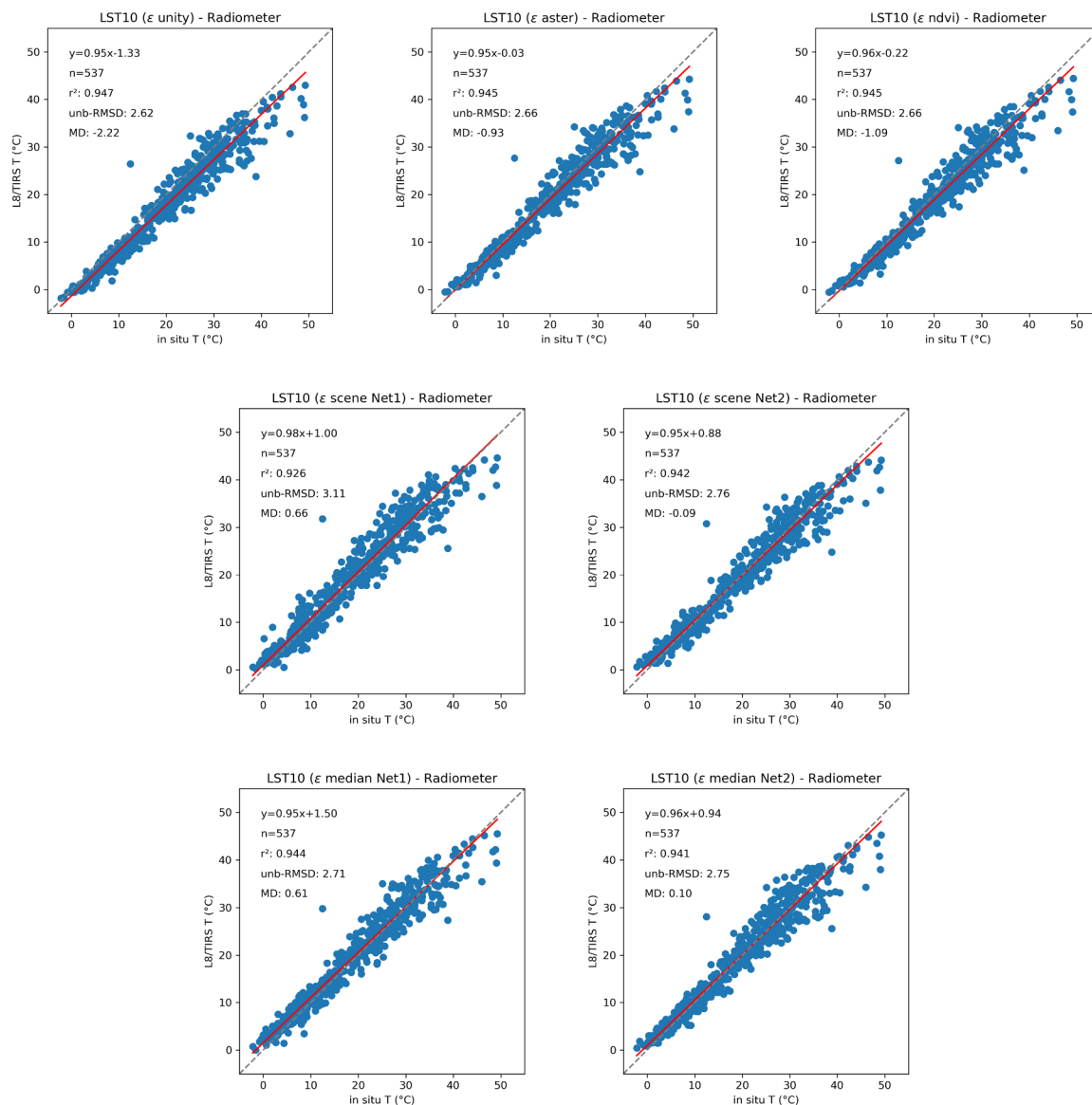
Fig. 10. Matchups for Band 10 with temperatures measured using broadband radiometers at the Belgian AWS: (top): unity, ASTER GED, and NDVI approach, (middle) scene specific Net1 and Net2, and (bottom) region median Net1 and Net2.

pixel, and the comparison should be extended to non-vegetated and homogeneous sites. The MD for the neural nets is about 1 K higher (and hence closer to 0 for the broadband radiometer data) than those when using the NDVI and ASTER GDEM emissivity estimates, although giving approximately the same unb-RMSD. This higher MD is consistent with the generally lower emissivities retrieved from the neural nets.

5. Conclusions

- Neural networks trained on the ECOSTRESS spectral library can reliably (1% RMSD) estimate emissivity in thermal infrared bands from multispectral visible to shortwave infrared reflectance, with a low negative bias. The estimation of emissivity in two thermal channels on TIRS from concomitant 7 band OLI ρ_s imagery is demonstrated.
- The OLI and neural network derived emissivity gives low mean differences compared to the ASTER Global Emissivity Dataset (0.6–2.6% MARD), but a larger emissivity range, potentially due to the higher spatial resolution of the OLI dataset. TIRS derived surface temperatures using the OLI emissivities have a crisper appearance when using the OLI derived emissivity, and the OLI derived

emissivity may lead to a 'sharpening' of the TIRS data.

- TIRS derived LST compare very well to in situ measurements of broadband radiometer temperatures at 12 Automated Weather Stations in fragmented landscapes across Belgium. Matchup results for B10 show near unity slopes, with an unb-RMSD of 2.6–3.1 K and a MD of −2.2–0.6 K, with the best performing emissivity source (Net2) giving an unb-RMSD around 2.8 K and MD of ±0.1 K, depending on whether the scene-specific or mission median emmisivity was used. TIRS derived LST also compare well to in situ measurements of subsoil temperatures at 14 AWS. Matchup results with subsoil measurements show an unb-RMSD around 3 K and MD from 1–4 K with near unity slopes.
- There were only small differences between using the neural network derived emissivity from the per-scene or the mission median reflectance, with the per-scene emissivity generally having a lower MD and slightly higher RMSD compared to the mission median. This indicates that the results from the mission median data could also be used for processing night-time data in the absence of optical data.
- Better performances may be achieved using high performance in situ radiometers which measure the land surface skin temperature, even though the results achieved here are comparable to matchup
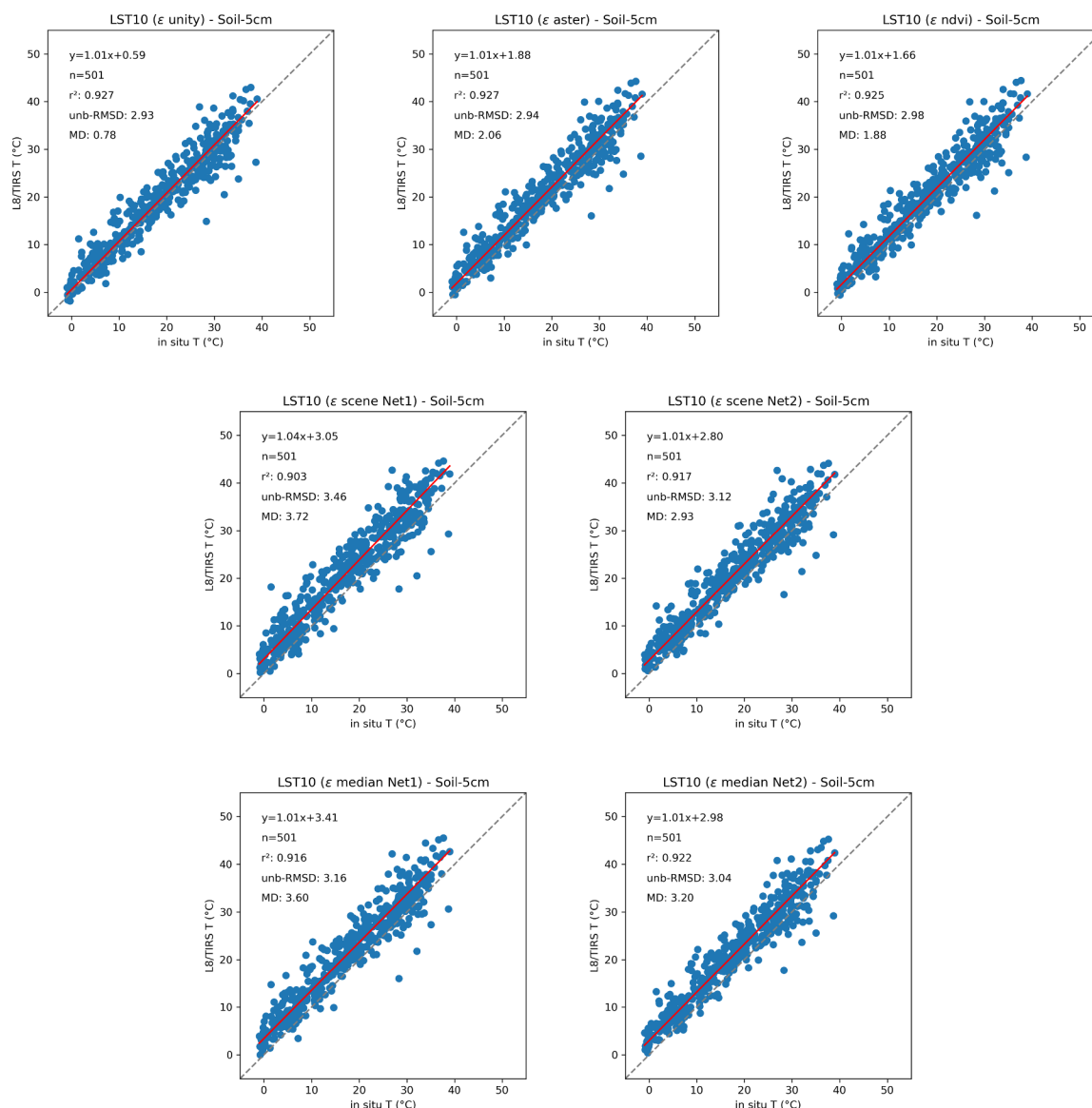
Fig. 11. Same as Fig. 10 but for subsoil temperatures.

analyses done with in situ radiometers (García-Santos et al., 2018; Sekertekin, 2019). The use of temperature measurements from dense networks of meteorological observations does significantly increase the number of matchups, and can complement the observations from sites equipped with thermal radiometers. > 2000 potential matchups of satellite overpasses with bounding in situ measurements were identified, with > 500 remaining after automated quality control (cloud masking).

• Although good performance was found for the broadband

radiometer matchups and the neural network derived emissivity, for the datasets presented by Sekertekin (2019) and García-Santos et al. (2018), and the subsoil temperature matchups presented here, there is no obvious performance improvement of using a non-unity emissivity value. Further investigations could be the measurement of temperature in contrasting land covers within the same region of interest, to establish the importance of a spatially and temporally variable emissivity. The impacts of the number of matchups and range in the validation data need to be further investigated, and

Table 4

The matchups with SURFRAD data as presented by Sekertekin (2019), and results from the present study.

Band	Study	in situ	n	m	b (°C)	r^2	RMSD (°C)	unb-RMSD (°C)	MD (°C)
LST10	Sekertekin (2019)	SURFRAD	20	1.172	−2.696	0.960	3.114	2.248	2.155
LST10	present - ∈ unity	SURFRAD	20	1.093	−2.181	0.946	2.172	2.125	0.449
LST10	present - ∈ scene Net1	SURFRAD	20	1.128	−0.757	0.952	3.592	2.162	2.869
LST10	present - ∈ scene Net2	SURFRAD	20	1.110	−0.738	0.960	3.058	1.934	2.369
LST10	Sekertekin (2019)	SURFRAD (sub)	15	1.127	−2.130	0.968	2.168	1.774	1.247
LST10	present - ∈ unity	SURFRAD (sub)	15	1.040	−1.348	0.946	1.887	1.865	−0.292
LST10	present - ∈ scene Net1	SURFRAD (sub)	15	1.059	0.405	0.962	2.542	1.614	1.964
LST10	present - ∈ scene Net2	SURFRAD (sub)	15	1.081	−0.406	0.960	2.452	1.729	1.738

Table 5

Matchups of the present study processing with UIB radiometer data from García-Santos et al. (2018) for various settings of emissivity.

Band	Emissivity	in situ	n	m	b (°C)	r^2	RMSD (°C)	unb-RMSD (°C)	MD (°C)
LST10	García-Santos et al. (2018) (ϵ = 0.968)	UIB	21	0.903	3.594	0.974	2.083	2.077	− 0.157
LST10	present - ϵ unity	UIB	21	0.903	1.744	0.975	2.854	2.031	− 2.005
LST10	present - ϵ scene Net1	UIB	21	0.821	7.681	0.961	2.984	2.895	0.726
LST10	present - ϵ scene Net2	UIB	21	0.861	5.395	0.973	2.343	2.343	0.023

perhaps the study could be extended to include the meteorological measurements from other countries.

- Quality control of the imagery was automated due to the large amount (> 2000) of potential matchups. Some scenes that would not pass manual quality assessment, e.g. presence of very thin clouds that cannot easily be detected spectrally, are still included in the matchup exercise. Further improvement of the validation results are expected if automated quality control can be improved.

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

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X., 2015. TensorFlow: Large-scale machine learning on heterogeneous systems. Software available from tensorflow.org. URL https://www.tensorflow.org/.

Albrechts, L., Healey, P., Kunzmann, K.R., 2003. Strategic spatial planning and regional governance in Europe. J. Am. Plan. Assoc. 69 (2), 113–129.

Anderson, M., Kustas, W., 2008. Thermal remote sensing of drought and evapo-transpiration. Eos, Trans. Am. Geophys. Union 89 (26), 233–234.

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.

Baldridge, A.M., Hook, S., Grove, C., Rivera, G., 2009. The ASTER spectral library version 2.0. Remote Sens. Environ. 113 (4), 711–715.

Barsi, J.A., Barker, J.L., Schott, J.R., 2003. An atmospheric correction parameter calculator for a single thermal band earth-sensing instrument. In: IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No. 03CH37477). Vol. 5. IEEE, pp. 3014–3016.

Barsi, J.A., Schott, J.R., Palluconi, F.D., Hook, S.J., 2005. Validation of a web-based atmospheric correction tool for single thermal band instruments. In: Earth Observing Systems X. Vol. 5882. International Society for Optics and Photonics, p. 58820E.

Bertrand, C., Gonzalez Sotelino, L., Journée, M., 2013. Quality control of 10-min air temperature data at RMI. AdSR 10, 1–5.

Bertrand, C., González Sotelino, L., Journée, M., 2015. Quality control of 10-min soil temperatures data at RMI. Adv. Sci. Res. 12 (1), 23–30.

Buscail, C., Upegui, E., Viel, J.-F., 2012. Mapping heatwave health risk at the community level for public health action. Int. J. Health Geogr. 11 (1), 38.

Carlson, T.N., Ripley, D.A., 1997. On the relation between NDVI, fractional vegetation cover, and leaf area index. Remote Sens. Environ. 62 (3), 241–252.

Caselles, V., Rubio, E., Coll, C., Valor, E., 1998. Thermal band selection for the PRISM instrument: 3. Optimal band configurations. J. Geophys. Res.: Atmos. 103 (D14), 17057–17067.

Chan, H.-P., Chang, C.-P., Dao, P.D., 2018. Geothermal anomaly mapping using landsat etm + data in ilan plain, northeastern taiwan. Pure Appl. Geophys. 175 (1), 303–323.

Chen, F., Yang, S., Su, Z., Wang, K., 2016. Effect of emissivity uncertainty on surface temperature retrieval over urban areas: Investigations based on spectral libraries. ISPRS J. Photogram. Remote Sens. 114, 53–65.

Chollet, F., et al., 2015. Keras. https://keras.io.

Cook, M., Schott, J., Mandel, J., Raqueno, N., 2014. Development of an operational calibration methodology for the Landsat thermal data archive and initial testing of the atmospheric compensation component of a Land Surface Temperature (LST) product from the archive. Remote Sens. 6 (11), 11244–11266.

Donlon, C., Minnett, P.J., Gentemann, C., Nightingale, T., Barton, I., Ward, B., Murray, M., 2002. Toward improved validation of satellite sea surface skin temperature measurements for climate research. J. Clim. 15 (4), 353–369.

Dousset, B., Gourmelon, F., Laaidi, K., Zeghnoun, A., Giraudet, E., Bretin, P., Mauri, E., Vandentorren, S., 2011. Satellite monitoring of summer heat waves in the Paris metropolitan area. Int. J. Climatol. 31 (2), 313–323.

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 (1), 647–665.

Emde, C., Buras-Schnell, R., Kylling, A., Mayer, B., Gasteiger, J., Hamann, U., Kylling, J., Richter, B., Pause, C., Dowling, T., et al., 2016. The libRadtran software package for radiative transfer calculations (version 2.0. 1). Geoscient. Model Develop. 5, 1647–1672.

Gao, L., Wang, X., Johnson, B.A., Tian, Q., Wang, Y., Verrelst, J., Mu, X., Gu, X., 2020. Remote sensing algorithms for estimation of fractional vegetation cover using pure vegetation index values: A review. ISPRS J. Photogram. Remote Sens. 159, 364–377.

García-Santos, V., Coll, C., Valor, E., Niclòs, R., Caselles, V., 2015. Analyzing the anisotropy of thermal infrared emissivity over arid regions using a new MODIS land surface temperature and emissivity product (MOD21). Remote Sens. Environ. 169, 212–221.

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 (9), 1450.

García-Santos, V., Valor, E., Caselles, V., Coll, C., Burgos, M.Á., 2013. Effect of soil moisture on the angular variation of thermal infrared emissivity of inorganic soils. IEEE Geosci. Remote Sens. Lett. 11 (6), 1091–1095.

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.

Giglio, L., Descloitres, J., Justice, C.O., Kaufman, Y.J., 2003. An enhanced contextual fire detection algorithm for MODIS. Remote Sens. Environ. 87 (2–3), 273–282.

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 (4), 1113–1126.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27.

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.

Jiménez-Muñoz, J.C., Cristóbal, J., Sobrino, J.A., Sòria, G., Ninyerola, M., Pons, X., 2008. Revision of the single-channel algorithm for land surface temperature retrieval from Landsat thermal-infrared data. IEEE Trans. Geosci. Remote Sens. 47 (1), 339–349.

Jin, M., Dickinson, R., Vogelmann, A., 1997. A comparison of ccm2–bats skin temperature and surface-air temperature with satellite and surface observations. J. Clim. 10 (7), 1505–1524.

Jin, M., Mullens, T., 2014. A study of the relations between soil moisture, soil temperatures and surface temperatures using ARM observations and offline CLM4 simulations. Climate 2 (4), 279–295.

Justice, C., Giglio, L., Korontzi, S., Owens, J., Morisette, J., Roy, D., Descloitres, J., Alleaume, S., Petitcolin, F., Kaufman, Y., 2002. The MODIS fire products. Remote Sens. Environ. 83 (1–2), 244–262.

Karnieli, A., Agam, N., Pinker, R.T., Anderson, M., Imhoff, M.L., Gutman, G.G., Panov, N., Goldberg, A., 2010. Use of NDVI and land surface temperature for drought assessment: Merits and limitations. J. Clim. 23 (3), 618–633.

Kuenzer, C., Zhang, J., Li, J., Voigt, S., Mehl, H., Wagner, W., 2007. Detecting unknown coal fires: synergy of automated coal fire risk area delineation and improved thermal anomaly extraction. Int. J. Remote Sens. 28 (20), 4561–4585.

Lagios, E., Vassilopoulou, S., Sakkas, V., Dietrich, V., Damiata, B.N., Ganas, A., 2007. Testing satellite and ground thermal imaging of low-temperature fumarolic fields: The dormant Nisyros Volcano (Greece). ISPRS J. Photogram. Remote Sens. 62 (6), 447–460.

Li, J., Song, C., Cao, L., Zhu, F., Meng, X., Wu, J., 2011. Impacts of landscape structure on

surface urban heat islands: A case study of Shanghai, China. Remote Sens. Environ. 115 (12), 3249–3263.

Li, Z.-L., Becker, F., 1993. Feasibility of land surface temperature and emissivity determination from thermal infrared data. Remote Sens. Environ. 43 (1), 67–85.

Li, Z.-L., Wu, H., Wang, N., Qiu, S., Sobrino, J.A., Wan, Z., Tang, B.-H., Yan, G., 2013. Land surface emissivity retrieval from satellite data. Int. J. Remote Sens. 34 (9–10), 3084–3127.

Liu, L., Zhang, Y., 2011. Urban heat island analysis using the Landsat TM data and ASTER data: A case study in Hong Kong. Remote Sens. 3 (7), 1535–1552.

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 (10), 5717–5735.

Malardel, S., Wedi, N., Deconinck, W., Diamantakis, M., Kühnlein, C., Mozdzynski, G., Hamrud, M., Smolarkiewicz, P., 2016. A new grid for the IFS. ECMWF Newsletter 146, 23–28.

Mansor, S.B., Cracknell, A.P., Shilin, B., Gornyi, V., 1994. Monitoring of underground coal fires using thermal infrared data. Int. J. Remote Sens. 15 (8), 1675–1685.

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 (5), 479.

Mayer, B., Kylling, A., 2005. The libRadtran software package for radiative transfer calculations-description and examples of use. Atmos. Chem. Phys. 5 (7), 1855–1877.

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 (2), 155.

Mia, M., Fujimitsu, Y., Nishijima, J., 2017. Thermal activity monitoring of an active volcano using Landsat 8/OLI-TIRS sensor images: A case study at the Aso volcanic area in southwest Japan. Geosciences 7 (4), 118.

Mia, M.B., Nishijima, J., Fujimitsu, Y., 2014. Exploration and monitoring geothermal activity using Landsat ETM+ images: A case study at Aso volcanic area in Japan. J. Volcanol. Geoth. Res. 275, 14–21.

Mira, M., Valor, E., Boluda, R., Caselles, V., Coll, C., 2007. Influence of soil water content on the thermal infrared emissivity of bare soils: Implication for land surface temperature determination. J. Geophys. Res.: Earth Surf. 112 (F4).

Parastatidis, D., Mitraka, Z., Chrysoulakis, N., Abrams, M., 2017. Online global land surface temperature estimation from landsat. Remote Sens. 9 (12), 1208.

Pérez-Planells, L., García-Santos, V., Caselles, V., 2015. Comparing different profiles to characterize the atmosphere for three MODIS TIR bands. Atmos. Res. 161, 108–115.

Ren, H., Liu, R., Qin, Q., Fan, W., Yu, L., Du, C., 2017. Mapping finer-resolution land surface emissivity using Landsat images in China. J. Geophys. Res.: Atmos. 122 (13), 6764–6781.

Ren, H., Yan, G., Chen, L., Li, Z., 2011. Angular effect of MODIS emissivity products and its application to the split-window algorithm. ISPRS J. Photogram. Remote Sens. 66 (4), 498–507.

Sandholt, I., Rasmussen, K., Andersen, J., 2002. A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status. Remote Sens. Environ. 79 (2–3), 213–224.

Schaeffer, B.A., Iiames, J., Dwyer, J., Urquhart, E., Salls, W., Rover, J., Seegers, B., 2018. An initial validation of Landsat 5 and 7 derived surface water temperature for US lakes, reservoirs, and estuaries. Int. J. Remote Sens. 39 (22), 7789–7805.

Sekertekin, A., 2019. Validation of Physical Radiative Transfer Equation-Based Land Surface Temperature Using Landsat 8 Satellite Imagery and SURFRAD In-situ Measurements. J. Atmos. Solar Terr. Phys. 105161.

Sekertekin, A., Bonafoni, S., 2020. Land Surface Temperature Retrieval from Landsat 5, 7, and 8 over Rural Areas: Assessment of Different Retrieval Algorithms and Emissivity Models and Toolbox Implementation. Remote Sensing 12 (2), 294.

Sholihah, R.I., Trisasongko, B.H., Shiddiq, D., La Ode, S.I., Kusdaryanto, S., Panuju, D.R., et al., 2016. Identification of agricultural drought extent based on vegetation health indices of landsat data: case of Subang and Karawang, Indonesia. Proc. Environ. Sci. 33, 14–20.

Skoković, D., Sobrino, J., Jimenez-Munoz, J., Soria, G., Juşien, Y., Mattar, C., Cristóbal, J., 2014. Calibration and validation of land surface temperature for landsat8-tirs sensor. LPVE (Land Product Validation and Evolution.

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 (14), 2753–2774.

Sobrino, J., Raissouni, N., 2000. Toward remote sensing methods for land cover dynamic monitoring: Application to morocco. Int. J. Remote Sens. 21 (2), 353–366.

Sobrino, J.A., Jiménez-Muñoz, J.C., Sòria, G., Romaguera, M., Guanter, L., Moreno, J., Plaza, A., Martínez, P., 2008. Land surface emissivity retrieval from different VNIR and TIR sensors. IEEE Trans. Geosci. Remote Sens. 46 (2), 316–327.

Trombetta, T., Vidussi, F., Mas, S., Parin, D., Simier, M., Mostajir, B., 2019. Water temperature drives phytoplankton blooms in coastal waters. PloS one 14 (4), e0214933.

Vanhellemont, Q., 2019. Adaptation of the dark spectrum fitting atmospheric correction for aquatic applications of the Landsat and Sentinel-2 archives. Remote Sens. Environ. 225, 175–192.

Vanhellemont, Q., 2020. Automated Water Surface Temperature retrieval from Landsat 8/TIRS. Remote Sens. Environ. 237, 111518.

Vanhellemont, Q., Ruddick, K., 2018. Atmospheric correction of metre-scale optical satellite data for inland and coastal water applications. Remote Sens. Environ. 216, 586–597.

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., Wang, P., Li, X., 2004. Using MODIS land surface temperature and normalized difference vegetation index products for monitoring drought in the southern Great Plains, USA. Int. J. Remote Sens. 25 (1), 61–72.

Wiltshire, K.H., Malzahn, A.M., Wirtz, K., Greve, W., Janisch, S., Mangelsdorf, P., Manly, B.F., Boersma, M., 2008. Resilience of North Sea phytoplankton spring bloom dynamics: An analysis of long-term data at Helgoland Roads. Limnol. Oceanogr. 53 (4), 1294–1302.

Wynne, T.T., Stumpf, R.P., Tomlinson, M.C., Fahnenstiel, G.L., Dyble, J., Schwab, D.J., Joshi, S.J., 2013. Evolution of a cyanobacterial bloom forecast system in western Lake Erie: Development and initial evaluation. J. Great Lakes Res. 39, 90–99.

Yang, J., Wong, M.S., Menenti, M., Nichol, J., 2015. Modeling the effective emissivity of the urban canopy using sky view factor. ISPRS J. Photogram. Remote Sens. 105, 211–219.