Automated water surface temperature retrieval from Landsat 8/TIRS
Quinten Vanhellemont
Royal Belgian Institute of Natural Sciences, Operational Directorate Natural Environments, Vautierstraat 29, 1000, Brussels, Belgium

ABSTRACT
Satellite remote sensing of Land and Water Surface Temperature (L/WST) has many applications in studies of terrestrial and aquatic ecology. Retrieval of L/WST requires a well calibrated radiometer and an accurate atmospheric correction. In the present study, the performance of the Thermal InfraRed Sensor (TIRS) on board Landsat 8 is evaluated for the retrieval of L/WST. libRadtran is used to retrieve atmospheric correction parameters based on atmospheric profiles of relative humidity and temperature from three global atmospheric models. Performance of single band retrievals is compared to typical MODTRAN results from the Atmospheric Correction Parameter Calculator (ACPC) and a split-window approach. A multi-temporal land masking method using imagery from the Operational Land Imager (OLI) on board Landsat 8 is demonstrated, and is used to automatically classify imagery in the matchup dataset in three classes of cloud cover. Two sources of in situ data covering the Belgian Coastal Zone (BCZ) are used for validation of the L/WST product: (1) fixed locations in the Flemish Banks measurement network and (2) underway data from regular RV Belgica campaigns. In the present study the single band methods outperformed the split-window approach, and consistent retrievals are found for the MODTRAN and libRadtran simulations. Typical single band surface temperature retrievals in quasi cloud-free conditions have Root Mean Squared Differences (RMSD) of 0.7 K and 1 K for Bands 10 and 11 with low bias, depending on the method and atmospheric profile source. For imagery with scattered clouds, RMSD values increase to 1 K and 2 K respectively with an approximately 0.5 K cold bias, likely caused by cloud proximity. The calibration efforts combined into Collection 1 allows for accurate absolute surface temperature retrievals from B10 on Landsat 8/TIRS for homogeneous targets with known emissivity, such as liquid water. The method is adapted to global processing and can be used for Land Surface Temperature retrieval with a suitable source of emissivity data.

1. Introduction
Land and Water Surface Temperature (L/WST) are essential variables for understanding the earth's climate and surface temperature is an important driver in ecology, biodiversity, and species distribution (Burrows et al., 2011; Doney et al., 2012; Cheung et al., 2009; Yang et al., 2013; Pachauri et al., 2014; Neukermans et al., 2018). LST retrieval is essential for understanding the evapotranspiration budget and management of freshwater sources (Anderson et al., 2012), and can be used to study water use change in agricultural areas (Anderson et al., 2018). WST in the dynamic near-shore coastal zone drives biogeochemical processes and its monitoring is essential, but difficult using traditional methods (Brewin et al., 2017, 2018). WST can be linked to water stratification and the occurrence of cyanobacterial blooms, and a positive relationship between WST and cyanobacterial growth rate can be found (Paerl, 2014). Buoyant cyanobacterial colonies may even enhance WST leading to a positive feedback mechanism (Paerl and Paul, 2012). Often satellite data with kilometre scale resolution (e.g. from MODIS and AVHRR) is used for WST retrieval, which may perform poorly at the land-water border due to mixed pixels and land influence (Smit et al., 2013; Brewin et al., 2018). WST was derived from higher resolution Landsat 5 and 7 records for US reservoirs and estuaries in order to improve harmful algal bloom forecasting capabilities (Schaeffer et al., 2018). To complement the high performance of the Operational Land Imager (OLI) on Landsat 8 over waters (Vanhellemont and Ruddick, 2014; Franz et al., 2015), several researchers combined OLI derived water colour data with Thermal InfraRed Sensor (TIRS) derived WST for river plume (Manzo et al., 2018; Brando et al., 2015), coastal and estuarine (Trinh et al., 2017; Snyder et al., 2017) studies. Landsat derived WST was used for monitoring bathing waters (Cherif et al., 2019), and high resolution thermal satellite imagery has been used to validate hydrodynamical models, e.g. for modelling thermal effluent from power plants (Salgueiro et al., 2015). There clearly is an interest in operational use of Landsat data for WST monitoring in coastal and inland waters, and the TIRS on Landsat 8, is an ideal candidate for such a high resolution WST product. One of the candidate
missions in the European Commission’s Copernicus programme is the Land Surface Temperature Monitoring (LSTM) mission which will carry a high resolution multi-band thermal sensor to provide land and water surface temperatures. Water targets are often used for validation of LST products (Barsi et al., 2005; Cook et al., 2014; Malakar et al., 2018), due to the known spectral emissivity and the generally homogeneous temperature of water bodies within a Landsat pixel. Landsat 8 derived L/WST products are generally not yet readily available for study sites worldwide, e.g. as a result of the use of localised ancillary datasets of atmospheric profiles or the lack of freely available software, and few studies compare the performance of different methods. Recently, the United States Geological Survey (USGS), which operates Landsat and distributes its data, announced that a standard LST product based on the work of Cook et al. (2014) will be distributed starting with Collection 2, at the earliest in late 2019.

TIRS was designed with two thermal infrared channels, B10 (10.6–11.2 μm) and B11 (11.5–12.5 μm), with the aim of using a split window algorithm (Caselles et al., 1998) for atmospheric correction of the imagery (Irons et al., 2012). Initial vicarious calibration results showed the TIRS channels to be too warm by several K in both channels (Barsi et al., 2014; Montanaro et al., 2014b), and significant stray light issues were found that limited the performance of split-window approaches (Montanaro et al., 2014a). USGS/Earth Explorer recommended not to use split-window algorithms in absence of a full stray light correction. After the development of a stray light correction scheme by Gerace and Montanaro (2017), the split window approach of Du et al. (2015) reportedly reached high fidelity comparable to that of MODIS. This stray light correction is now integrated in the Landsat Collection 1 dataset that is widely distributed by USGS. In the meantime, several single band algorithms have been developed (Cook et al., 2014; Wang et al., 2015), both in order to minimize the impacts of the stray light effect and for transferability to older Landsat sensors with only a single thermal channel. Cook et al. (2014) used modelled atmospheric profiles of relative humidity and temperature, and found B10 and B11 LST being respectively 0.5 K and 2 K too warm compared to buoy temperatures. Wang et al. (2015) established an algorithm using surface measured temperature and columnar water vapour to estimate atmospheric temperature and transmittance, and found a B10 derived LST RMSD of 0.8 K compared to modelled results. It should be noted that these results were based on older processing and calibration versions.

In most studies of LST retrieval and calibration of thermal bands (Schott et al., 2012; Cook et al., 2014), the radiative transfer modelling is performed using the MODerate resolution atmospheric TRANsmission program (MODTRAN) developed by Spectral Sciences, Inc. and the US Air Force (Berk et al., 1999). MODTRAN is not free or open source libRadtran (Mayer and Kylling, 2005; Emde et al., 2016) radiative transfer code is used for retrieval of atmospheric correction parameters, together with atmospheric profiles from three different gridded atmospheric model datasets, similar to the work of Cook et al. (2014) and Barsi et al. (2003). Atmospheric correction performance for these single band libRadtran runs is compared to two other algorithms: (1) the commonly used Atmospheric Correction Parameter Calculator (Barsi et al., 2003, 2005), and (2) the split-window algorithm by Wan (2014) as calibrated by Du et al. (2015). Multi-temporal land masking is developed for improved automated cloud-filtering of matchups, and L/WST retrievals are compared to in situ temperature measurements from the Flemish Banks measurement network and underway data from Research Vessel Belgica collected in the Belgian Coastal Zone (BCZ).

2. Data and methods

2.1. In situ data

In situ measurements of water temperature were obtained from 19 measurement sites in the Flemish Banks Coastal Monitoring network operated by the Agency for Maritime Services and Coast of the Flemish Government (https://meetnetvlaamsebanken.be/). The locations of these sites in the Belgian Coastal Zone (BCZ) are provided in Fig. 1. 15 sites are buoy deployments using Datawell Directional Wave riders, and 4 are measuring piles with an Aanderaa CT4120AW sensor. Temperature is measured in the surface waters every 30 min with reported 0.2 °C measurement accuracy for the buoys and 0.1 °C for the piles (see manufacturer documentation for details). Underway measurements were also obtained from the Onboard Data Acquisition System (ODAS) on Research Vessel Belgica (https://odnature.naturalsciences.be/belgica/en/odas). ODAS temperature is measured with a Sea-Bird Scientific SBE-38 Digital Oceanographic Thermometer with initial accuracy of 0.001 °C in lab conditions. Temperature is measured close to the water inlet located on the bow 3 m below the waterline. Underway data is provided every 10 min, which corresponds to an approximately 3 km spacing between measurements for typical cruising speeds. ODAS data was screened for quality flagged geolocation or temperature data. In situ data with unrealistic temperatures for the BCZ were filtered out, i.e. by removing $T \leq 0^\circ$ and $T > 25^\circ$.

Due to radiative cooling of the water surface skin, the skin temperature measured by satellite is generally a few tenths of a degree cooler than the temperature measured at depth (Donlon et al., 1999, 2002), and the measured temperature needs to be adjusted to skin temperature for validation purposes. At high wind speeds generally a constant correction factor can be used (Alappattu et al., 2017; Donlon et al., 2002) to adjust bulk to skin temperature. Winds in the BCZ are generally > 5 m/s (Ruddick and Lacroix, 2006; Baeye et al., 2011), which is confirmed here using wind speed measured at 10 m above the surface for three measurement sites. In the period 07/2012–07/2019, winds at MP0, MP4, and MP7 (see Fig. 1) were > 4 m/s respectively 83%, 81%, and 86% of the time, and hence in situ measured bulk temperature was directly adjusted to skin temperature by a constant...
offset of $-0.17$ K (Donlon et al., 2002) The tidal waters of the BCZ are well mixed and rarely stratified (Ruddick and Lacroix, 2006), the buoy measured temperature is the equilibrium temperature of the submerged bottom part of the buoy hull, and the water pumped through the ship's underway system is considered to be well mixed. Using $> 4000$ CTD casts collected by the Flemish Marine Institute (VLIZ) in the BCZ between 2001 and 2019 (available at http://www.vliz.be/vmdcdata/midas/), temperature variability in the upper 3 m was found to be low, with 95/90/50 percentiles of $< 0.06$ K/ $< 0.03$ K/ $< 0.01$ K, and hence there is no thermal gradient to the measurement depth (Cook et al., 2014; Zeng et al., 1999) that needs to be corrected. Diurnal temperature variability in the BCZ is also low; the interquartile range from the 24 h moving standard deviation in the Flemish Banks measurements was 0.1°C–0.2 °C, i.e. within the reported measurement accuracy (see further). TIRS is a two band 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. OLI data was used for quality control of the imagery - mainly for cloud and object screening (see further). TIRS was limited to a common coverage of the full BCZ (51.13–51.76 N, 2.27–3.37 E) covering the Flemish Banks in the period from end of March 2013 to the end of June 2019, from the two WRS-2 path/row tile sets covering the BCZ: 199/024 (139 tiles) and 200/024 (141 tiles). For six relatively clear images with matchup underway data a slightly larger region of interest was processed, covering nearly the full BCZ (51.13–51.76 N, 2.27–3.37 E). Nighttime L8 imagery was available for a limited period (22 scenes in path/row 056/220 between September 2016 and August 2017), but was excluded from the present study due to the lack of OLI data for automated quality assessment.

2.3. Surface temperature retrieval

Top-of-Atmosphere (TOA) radiance ($L_t$) is a measurement of the combined surface-emitted radiance ($L_s$) and atmospheric up- ($L_u$) and downwelling ($L_d$) radiances transmitted through the atmosphere ($\epsilon$): $L_t = \epsilon \cdot (L_s + (1 - \epsilon) \cdot L_u) + L_u$, with $\epsilon$ the target emissivity, and 1-$\epsilon$ the target albedo. To convert $L_t$, or equivalent at-sensor Brightness Temperature (BT), to surface temperature, the atmospheric effects need to be corrected, for which the $L_u$, $L_d$, $\tau$, and $\epsilon$ need to be estimated. $L_u$ and $\tau$ can be determined from radiative transfer simulations for $\epsilon = 1$ at two different surface temperatures $L_{s1}$ and $L_{s2}$, as the offset and slope of a linear fit of the surface and TOA signals:

$$
\tau = \frac{L_{s2} - L_{s1}}{L_{d2} - L_{d1}}
$$

$$
L_u = L_{s1} - \tau \cdot L_{d1}
$$

With the knowledge of $L_u$ and $\tau$, $L_d$ can then be computed from a simulation with $\epsilon < 1$ and $L_u = 0$:

$$
L_d = \frac{L_t - L_u}{\tau + \epsilon}
$$

2.2. Satellite data

Collection 1 imagery from the Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) on board Landsat 8 were obtained from Google Earth Engine (GEE, Gorelick et al. (2017)). GEE distributes the standard Level 1T (terrain corrected) Top-of-Atmosphere (TOA) images in GeoTiff format as produced by USGS from the Landsat Product Generation System (LPGS) from 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. OLI data was used for quality control of the imagery – mainly for cloud and object screening (see further).
Table 1. Landsat & TIRS bands and used constants.

| Band  | Wavelength | Bandwidth | fwater | K1     | K2     |
|-------|------------|-----------|---------|--------|--------|
|       | (µm)       | (µm)      |         |        |        |
| B10   | 10.9       | 10.6-11.2 | 0.9926  | 774.8853 | 1321.0789 |
| B11   | 12.0       | 11.5-12.5 | 0.9877  | 480.8883 | 1201.1442 |

$L_d$ can then be estimated from Eq. (4), and the surface temperature in K retrieved from $L_d$:

$$T = \frac{K2}{\ln \left( \frac{\tau_2}{\tau_1} + 1 \right)}$$

with K1 and K2 included in the Landsat metadata file, here reproduced in Table 1. The distilled water emissivity spectrum from the MODIS (Moderate Resolution Imaging Spectroradiometer) Emissivity Library was resampled to the TIRS spectral bands to retrieve $e$ of water in Bands 10 and 11 (Table 1).

The libRadtran (version 2.0.2) radiative transfer code (Mayer and Kylling, 2005) was used to perform these three simulations and derive hyperspectral values of $\tau$, $L_d$, and $L_d$ which were then resampled to the relative spectral response of B10 and B11 on Landsat-8/TIRS. These parameters and the surface emitted radiance depend on the observation zenith angle (Cao et al., 2019), which is here simplified to $\theta = 0^\circ$ for the largely nadir viewing Landsat. Three sources of atmospheric profiles of temperature and relative humidity were used: (1) 6 hourly $1 \times 1$ degree data from the Global Data Assimilation System (GDAS, https://www.emc.ncep.noaa.gov/gmb/gdas/), and (2) 6 hourly $1 \times 1$ degree data from the Climate Forecast System version 2 (CFSv2, Saha et al., 2014), both from the National Centers for Environmental Prediction (NCEP), and (3) hourly $0.25 \times 0.25^\circ$ ERA5 Reanalysis Data from the European Centre for Medium-Range Weather Forecasts (ECMWF, Malarde et al. (2016)). The atmospheric profiles from the grid cells bounding the region of interest in space and the overpass in time were extracted and interpolated to form a single profile per image. This profile was then input into libRadtran to retrieve the required parameters. Due to the global nature of the datasets, this method could be extended to any scene globally, and for full scene processing the use of a spatially resolved atmosphere is recommended (Appendix A). $\tau$, $L_d$, and $L_d$ were also obtained for B10 using the ‘Atmospheric Correction Parameter Calculator’ (https://atmcorr.gsfc.nasa.gov/) which uses NCEP/GDAS profiles and the MODTRAN radiative transfer model (Barsi et al., 2003, 2005). Additionally, the split-window algorithm of Wan (2014) was used:

$$LST = b_0 + \left( b_1 + b_2 \frac{\tau}{\bar{\tau}} \right) BT_{10} + BT_{11} + \left( b_3 + b_4 \frac{1 - \tau}{\bar{\tau}} \right) + b_5 \frac{\Delta T}{\bar{T}} BT_{10} - BT_{11} + b_6 (BT_{10} - BT_{11})^2$$

where $BT_{10}$ and $BT_{11}$ are the at-sensor Brightness Temperatures, computed by substituting $L_d$ by $L_d$ in Eq. (8), $\Delta T$ and $\bar{\tau}$ are the difference and average emissivity in Bands 10 and 11 (taken from Table 1), and the b coefficients derived from simulations by Du et al. (2015) based on ranges of Columnar Water Vapour (CWV, in cm). Two sets of coefficients are compared: (1) using the entire CWV range (0-6.5 cm) and (2) selecting the appropriate coefficient based on the CWV provided with the NCEP/GDAS profiles. An overview of the compared set ups is given in Table 2.

2.4. Image quality control

With the aim of image quality control, surface reflectances ($g_r$) were computed from the Operational Land Imager (OLI) using the Dark Spectrum Fitting (DSF) algorithm (Vanhellemont, 2019). Pixels were identified as non-water when $g_r$ 1609 nm > 0.05 or $g_r$ 1373 nm > 0.01, and a dilation operation with 10 iterations was performed to get a final non-water mask. Contamination by small clouds and objects (ships and offshore constructions) was filtered out by applying a threshold of 0.01 on the standard deviation of $g_r$ 560 nm in a $11 \times 11$ pixel box around the matchup location. These quality control steps will filter out thick clouds covering or close to the matchup station, but do not take into account cloud proximity and the presence of thin clouds, which may also negatively impact the matchup performance. A multi-temporal land masking method was developed to retrieve cloud fraction over water pixels for additional quality control of the matchups. For each image, a mask was computed based on $g_r$ 1609 nm > 0.05, and the number of masked pixels was recorded. The images were sorted according to the fraction of masked pixels, from low to high (Fig. 3). Land pixels are expected to be masked on most of these masks (excluding cloud shadows, intertidal and flooded areas, and scene edges), and a suitable threshold on image count needs to be determined to establish a land mask. A scaled logistic function of sorted image index (x) and number of masked pixels (y) can be fitted:

$$y = c + \frac{d - c}{1 + e^{a(x+b)}}$$

where a and b are the logistic growth rate and mid-point, and the asymptotic values $c$ and $d$ represent the number of masked pixels in cloud-free and fully cloudy conditions. The discrete difference (as % coverage) between the number of masked pixels was computed for the sorted image list. The sorted image index larger than the inflection point of the logistic function giving the longest stretch of discrete differences less than 0.5% of image pixels was selected as threshold. With the current dataset, this method gives image counts of 96 and 93 for path/row 199/024 and 200/024 respectively, i.e. pixels that are not water on at least that many images are classified as land. This method is generally applicable to coastal regions where the areas with low SWIR reflectance are water, i.e. most coastal regions. It may not perform optimally in the presence of ice, and has not been tested in extremely cloudy regions. At lower latitudes the SWIR reflectance may be contaminated by sun glint, and an appropriate sun glint correction may need to be applied to the $g_r$. 

3. Results and discussion

3.1. Land masking

The automated multi-temporal land masking generates a sensible land mask in a sensor-specific configuration, even in the presence of cloud shadows and intertidal zones, that generally complicate water/non-water masking. At the edge of the satellite swath, some issues were encountered due to the variable extent of the swath in geographic coordinates. In the present dataset, the 200/024 path/row combination has the edge of swath just east of Zeebrugge and hence variable coverage of the region by individual images. This effect causes the slight inclination in the minimum and maximum fractional coverage and slightly noisier discrete difference in the sorted image list (right hand site plots in Fig. 3). The land mask generation could be restricted to pixels with coverage in every scene, excluding a small part of the region covered by the variable swath extent. In the present situation, the mask generated from the 199/024 path/row combination, which always fully covers the study area, could however be used for both tilesets. This land mask allows for an estimation of cloudiness of the water area in the scene, which is essential for quality control of matchups. Images were sorted in three categories: (1) clear images, with < 5% of the water masked, (2) images with scattered clouds, 5-50% of the water masked and (3) cloudy images, with > 50% masked. Images from class (3) were automatically rejected from the validation exercise.

3.2. Surface temperature

Matchups of WST with the Flemish Bank measurements are given in
Tables 3 and 4, respectively, for the clear images (329 matchups for 15/19 sites) and the images with scattered clouds (294 matchups for 16/19 sites). Scatter plots for the clear image matchups are shown in Fig. 4 for the libRadtran runs and Fig. 5 for the ACPC and Split-Window approaches. Overall, the single band methods perform similarly, giving very good linear correlation statistics with near-unity slope and a low offset (< 0.2 K for B10). For the libRadtran results, the Root Mean Squared Difference (RMSD) are 0.64 K–0.88 K and the Mean Difference (MD) of 0.01 K–0.28 K for B10. Unbiased RMSD results are comparable to the RMSD due to the low MD. For B11 the differences are higher,
with RMSD ranges of 0.95 K–1.50 K and MD of 0.30 K–0.78 K. ACPC performs very well also, and on average the only method giving lower temperatures (MD -0.19 K) by Landsat. For the dataset presented here, neither of the Split Window approaches give as good results as the single channel approaches, with RMSD > 1.7 K and MD > 1.5 K. The SW approaches have unbiased RMSD results similar to the single band (B10) retrievals, indicating a similar level of precision, but a large systematic bias, likely due to the calibration of the algorithm. The unbiased RMSD is however slightly higher for SW than for the B10 approaches, perhaps due to the inclusion of B11 in the SW algorithm. The overall best performance was found for the B10 libRadtran runs and NCEP/GDAS and ERA5 reanalysis profiles with a MD of 0.01 K and 0.05 K. For the images with scattered clouds (Table 4) errors are noticeably larger, with about double the RMSD and a MD up to −0.5 K. This cold bias indicates impact of cloud proximity and contamination of undetected (thin) clouds. However, for certain applications these errors can be considered acceptable. By combining the 329 matchups from the clear and the 294 matchups from the scattered clouds datasets, RMSD and MD are about 0.8 K and −0.3 K respectively for the libRadtran results (not shown). Performance for the near-shore (< 1 km) stations RAV, ONS, OST, OOS, MP2 and ZHG (Fig. 1) was not found to be different from the more offshore stations.

### 3.3. Spatial consistency

For the underway data similar results are found, with the matchups from the libRadtran ERA5 dataset shown in Fig. 6 and Table 5, giving an RMSD of 0.6 K and a MD of −0.04 to 0.03 depending on time difference with the satellite overpass. Despite the largest time window including nearly a full tidal cycle which can modify the thermal distribution of the water mass, and heating of the top layer by the sun, errors are of the same magnitude as using the buoy matchups. The spatial patterns observed by Landsat correspond well to the in situ variability measured by ship (Figs. 7 and 8). Increase of the matchup time window does increase the RMSD and the offset of the regression line, and in general moves points away from the 1:1 line (Fig. 6), likely due to tidal displacement of water masses and heating/cooling of the surface.

---

Fig. 4. Matchups for the retrievals using the libRadtran model and three atmospheric profile sources for images with < 5% masked water pixels. Top and bottom rows show the B10 and B11 results. The red lines are the Reduced Major Axis regression lines. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Fig. 5. Same as Fig. 4 but for ACPC (Barsi et al., 2003) and the Split Window algorithm (Du et al., 2015) with two CWV ranges.
3.4. Perspectives

Combining libRadtran and ERA5 atmospheric profiles can provide single band WST products from L8/TIRS at high precision (<0.7 K unbiased RMSD) with low bias (near-zero) in cloud-free conditions. The performance of the processing presented here, specifically the atmospheric profiles, should be investigated in other sites worldwide. In situ bulk water temperature data is widely available, but for better understanding the link between bulk and skin temperature in situ thermal radiometers should be installed at validation sites, and perhaps also linked with anemometers and measurements of long wave radiation flux. The good performance of the WST product presented here opens possibilities for applications of high resolution near-shore water temperature retrieval, e.g. for monitoring river plumes, thermal pollution by factories and power plants, the assessment of coral bleaching and harmful algal bloom risks, and validation of hydrodynamical models. The SW approach provided similar unbiased RMSD results to the single band approach in the present study, but a much larger MD, indicating a large systematic bias which may be due to the calibration of the SW algorithm over man-made (urban) targets by Du et al. (2015). A recalibration of the SW algorithm for natural materials (specifically water for deriving WST) may improve the SW results, but is out of scope for the present paper. SW approaches may also be negatively affected by calibration issues in both bands, in the case of L8/TIRS specifically the stray light issues experienced with B11, and future multi-band missions should be carefully designed and characterised.

The present study shows the performances (better than 1 K) that can be achieved with a single band on L8/TIRS, and a method that may be extended to single band missions past and future. If the single band performance presented here is sufficient for the intended application, perhaps future missions can be made cheaper by including only a single band. Reliable atmospheric profiles can be retrieved from the ERA5 global reanalysis dataset at 0.25°, which is currently available from 1979 to the present. This opens up the applicability of the single band processing to older Landsat sensors, the Thematic Mapper on Landsat 5 (1984–2011) and the Enhanced Thematic Mapper on Landsat 7 (1999–present) for long term change detection. The methods presented in Section 2.3 and Appendix A are generic and can be applied to any mission, and can be adapted to Land Surface Temperature retrieval with a suitable source of per-pixel emissivity. As demonstrated in this paper, the combination of the VNIR and SWIR data from L8/OLI allows for a robust automated land and cloud masking. The use of coincident optical multi-spectral data could be further extended from quality control to emissivity estimation, e.g. based on existing spectral libraries.

4. Conclusions

- Water Surface Temperature can be accurately retrieved from single band Landsat 8/TIRS imagery. In cloud-free conditions, the RMSD between in situ buoy and satellite measurements are 0.7 K for Band 10 and 1 K for Band 11, with low biases of < 0.1 K and < 0.3 K respectively. For images with scattered clouds, the RMSD increases to 1 K and 2 K and the bias for both bands increases to −0.5 K, likely due to cloud proximity. A similar performance was found using underway shipborne measurements of temperature, comprehensively demonstrating Landsat's capability of tracking smaller scale temperature variability.
- The open source libRadtran model can be used as a robust alternative to MODTRAN for radiative transfer in the thermal domain with the aim of retrieving surface temperatures. Three sources of atmospheric profiles give comparable results, although lowest errors are found for NCEP/GDAS and ECMWF/ERA5. ACPC gave similar RMSD but larger MD than libRadtran using the same NCEP/GDAS

![Fig. 6. Matchups of the libRadtran/ERA5 setup with Belgica/ODAS underway data for different time ranges. Matchups are from 6 images, and the gap between warm and cold waters is caused by the seasonal distribution of the matchup data. Statistics are given for the full dataset + - 12 h from the overpass, and in cumulative form in Table 5.](image)

| AT (h) | n | m | b (°C) | r² | RMSD (°C) | MD (°C) | unb-RMSD (°C) |
|-------|---|---|--------|----|-----------|--------|-------------|
| 0.25  | 15| 0.990| 0.086| 0.987| 0.549 | −0.042| 0.547      |
| 0.5   | 32| 0.995| 0.029| 0.987| 0.549 | −0.034| 0.548      |
| 1     | 64| 0.995| 0.026| 0.987| 0.541 | −0.036| 0.540      |
| 2     | 125| 1.000| 0.040| 0.987| 0.551 | −0.038| 0.549      |
| 4     | 226| 1.016| 0.194| 0.983| 0.631| 0.020| 0.630      |
| 8     | 406| 1.030| 0.359| 0.983| 0.630| 0.025| 0.630      |
| 12    | 619| 1.027| 0.345| 0.982| 0.661| 0.009| 0.661      |

![Fig. 7. Two of the more complete tracks of Belgica/ODAS underway data and L8/TIRS temperature, with temperature difference (ΔT = satellite-in situ) plotted at the sampling locations. White areas denote missing data due to non-water masking and swath edge. Note the undetected thin clouds (cold areas) on the 2015-03-10 image.](image)
profiles. With the datasets used here, the single band methods outperformed a split-window approach, which had more than double the RMSD and MD in comparison with in situ measurements. Since the performance of B11 is here found to be acceptable, the calibration coefficients of the split-window approach may need to be revised.

- A multi-temporal land masking method using OLI data is demonstrated in order to estimate image cloud fraction over water pixels. An overall better matchup performance is found for scenes with low cloud cover percentage over water. This method can be used for generating sensor and projection specific land masks suitable for other applications such as turbidity and chlorophyll a mapping from Landsat or Sentinel-2 imagery.

- The method presented here is also suited for temperature retrieval over non-water targets, including e.g. urban, agricultural and forested areas, but also the Antarctic ice sheet and ice-free regions, if suitable methods to determine per-pixel emissivity can be established e.g. from the concomitant OLI data. The method is highly automated, globally applicable, and the processing software will be made freely available.

**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.

**Acknowledgments**

This work was funded by the Federal Belgian Science Policy Office (BELSPO) under the BRAIN-be programme MICROBIAN project (BR/165/A1/MICROBIAN). NASA and USGS are thanked for the acquisition and free distribution of Landsat imagery. The Agency for Maritime Services and Coast of the Flemish Government is thanked for setting up the Flemish Banks Coastal Monitoring network and free distribution of the data. The crew of RV Belgica and the personnel at Measurement Services Oostende are thanked for providing access to their CTD data. Kevin Ruddick and Wim Vyverman are kindly thanked for feedback on an initial draft of this manuscript. Two anonymous reviewers are thanked for their insights which helped improve this paper.

**Appendix A. Full tile processing**

Landsat scenes are distributed in approximately 185 x 180 km tiles in the WRS-2 grid, and for full or merged tile processing, a spatially resolved atmosphere may be more appropriate. In this appendix, the use of ERA5 0.25° grids for resolved full tile processing is examined (ERA5-R). The hourly ERA5 grids bounding the satellite overpass time are extracted for all grid cells required for a smooth interpolation over the scene (Fig. A1, left panel), and for each grid cell the three required libRadtran runs are performed to derive \( \tau \), \( \text{Lu} \), and \( L_d \). The two hourly grids of these parameters are then interpolated to the satellite overpass time (Fig. A1, right panel). The per-cell parameters are then linearly interpolated to form a per-pixel full-tile dataset (Fig. A2). The matchup exercise is repeated using the ERA5-R processing. The performance between the single profile runs (ERA5, Table 3) and the resolved atmosphere (ERA5-R, A1) is practically the same due to the small size of the ROI - about 0.4° in latitude, and 1° in longitude.

**Fig. A1.** A WRS-2 path/row 199/024 tile over the BCZ on 2019-02-26 with (left) the tile edges in red and ERA5 grid cells in blue, and (right) the temporally interpolated B10 \( \text{Lu} \) results.
Table A1

Same as Table 3 but for the ERAS-R processing with spatially resolved atmosphere, for the datasets with < 5% and 5–50% masked water pixels.

| Dataset | Band | n | m | b (°C) | r² | RMSD (°C) | MD (°C) | unb-RMSD (°C) |
|---------|------|---|---|-------|----|----------|--------|--------------|
| < 5%    | 10   | 329|   | 1.012 | -0.090 | 0.983    | 0.699  | 0.071        |
|         | 11   | 329|   | 1.035 | -0.123 | 0.968    | 1.038  | 0.360        |
| 5–50%   | 10   | 294|   | 1.010 | -0.661 | 0.969    | 0.901  | -0.512       |
|         | 11   | 294|   | 1.027 | -0.817 | 0.949    | 1.054  | -0.433       |

References

Alappattu, D.P., Wang, Q., Yamaguchi, R., Lind, R.J., Reynolds, M., Christman, A.J., 2017. Warm layer and cool skin corrections for bulk water temperature measurements for air-sea interaction studies. J. Geophys. Res.: Oceans 122 (8), 6470–6481.

Anderson, M., Gao, F., Knipper, K., Hain, C., Dulanye, W., Baldocchi, D., Eichelmann, E., Hemes, K., Yang, Y., Medellin-Azuara, J., et al., 2018. Field-scale assessment of land and water use change over the California Delta using remote sensing. Remote Sens. 10 (6), 889.

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.

Barey, M., Fettweis, M., Voulgaris, G., Van Lancker, V., 2011. Sediment mobility in response to tidal and wind-driven flows along the Belgian inner shelf, southern North Sea. Ocean Dyn. 61 (5), 611–622.

Barsi, J.A., Barker, J.J., 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, pp. 58820E.

Bark, A., Anderson, G.P., Bernstein, L.S., Acharya, P.K., Dothe, H., Mathew, M.W., Adler-Golden, S.M., Chetwynd Jr., J.H., Richtsmeier, S.C., Pukall, B., et al., 1999. MODTRAN4 radiative transfer modeling for atmospheric correction. In: Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research III, vol. 3756. International Society for Optics and Photonics, pp. 348–353.

Brando, V., Braga, F., Zaggia, L., Giardino, C., Bresciani, M., Matta, E., Bellafurore, D., Ferrari, C., Maiu, F., Benettazz, A., et al., 2015. High-resolution satellite turbidity and sea surface temperature observations of river plume interactions during a significant flood event. Ocean Sci. 11 (6), 909.

Brewin, R., Smale, D., Moore, P., Dall’Olmo, G., Miller, P., Taylor, B., Smyth, T., Fishwick, J., Yang, M., 2018. Evaluating operational AVHRR sea surface temperature data at the coastline using benthic temperature loggers. Remote Sens. 10 (6), 925.

Brewin, R.J., de Mora, L., Billson, O., Jackson, T., Russell, P., Brewin, T.G., Shutter, J.D., Miller, P.L., Taylor, B.H., Smyth, T.J., et al., 2017. Evaluating operational AVHRR sea surface temperature data at the coastline using surfers. Estuar. Coast Shelf Sci. 196, 276–289.

Burrows, M.T., Schoeman, D.S., Buckley, L.B., Moore, P., Poloczanska, E.S., Brande, K.M., Brown, C., Bruno, J.F., Duarte, C.M., Halpern, B.S., et al., 2011. The pace of significant flood event. Ocean Sci. 11 (6), 909.

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

Cheikh, E.K., Salmoun, F., Mesas-Carrascosa, F.J., et al., 2019. Determination of bathing water quality using thermal images Landsat 8 on the west coast of tangier: preliminary results. Remote Sens. 11 (8), 972.

Cheng, W.W., Lam, V.W., Sarmiento, J.L., Kearney, K., Watson, R., Pauly, D., 2009. Projecting global marine biodiversity impacts under climate change scenarios. Fish Fish. 10 (3), 235–251.

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.

Doney, S.C., Ruckelshaus, M., Emmett Duffy, J., Barry, J.P., Chan, F., English, C.A., Galindo, H.M., Grebeneiz, J.M., Hollowed, A.B., Knowlton, N., et al., 2012. Climate change impacts on marine ecosystems. Ann. Rev. Mar. Sci. 4, 11–37.

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.

Donlon, C., Nightingale, T., Shearby, T., Turner, J., Robinson, I., Emergy, W., 1999. Implications of the oceanic thermal skin temperature deviation at high wind speed. Geophys. Res. Lett. 26 (16), 2505–2508.

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., Burch-Schnell, R., Kylling, A., Mayer, B., Gasteiger, J., Harnann, U., Kylling, R., Richter, B., Pause, C., Dowling, T., et al., 2016. The libRadtran software package for radiative transfer calculations (version 2.0.1). Geosci. Model Dev. (GMD) (5), 1647–1672.

Franz, B.A., Bailey, S.W., Kuring, N., Werdel, P.J., 2015. Ocean color measurements with the operational land imager on landsat-8: implementation and evaluation in SeaDAS. J. Appl. Remote Sens. 9 (1) 096070–096070.

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.

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

Irons, J.R., Dwyer, J.L., Barst, J.A., 2012. The next Landsat satellite: the Landsat data continuity mission. Remote Sens. Environ. 122, 11–21.

Malakar, N.K., Hulley, G.C., Hook, S.J., Laramy, 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.

Mandal, S., Wedi, N., Deconinck, W., Diamantakis, M., Kuhlne, C., Modzynski, G., Hamrud, M., Smolarzewicz, P., 2016. A new grid for the IFS. ECMWF Newsletter. 146, 25–26.

Manzo, C., Federica, B., Luca, Z., Ernesto, B.V., Claudia, G., Mariano, B., Cristiana, B., 2018. Spatio-temporal analysis of prodelta dynamics by means of new satellite imagery observations and modeling: historical development, current status and perspectives. Remote Sens. Environ. 232 111304.
calculations—description and examples of use. Atmos. Chem. Phys. 5 (7), 1855–1877.
Montanaro, M., Gerace, A., Lunsford, A., Reuter, D., 2014a. Stray light artifacts in imagery from the Landsat 8 thermal infrared sensor. Remote Sens. 6 (11), 10435–10456.
Montanaro, M., Levy, R., Markham, B., 2014b. On-orbit radiometric performance of the Landsat 8 thermal infrared sensor. Remote Sens. 6 (12), 11753–11769.
Neukermans, G., Otziel, L., Babin, M., 2018. Increased intrusion of warming Atlantic water leads to rapid expansion of temperate phytoplankton in the Arctic. Glob. Chang. Biol. 24 (6), 2545–2553.
Pachauri, R.K., Allen, M.R., Barros, V.R., Broome, J., Cramer, W., Christ, R., Church, J.A., Clarke, L., Dahe, Q., Dasgupta, P., et al., 2014. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Ipcc.
Paerl, H., 2014. Mitigating harmful cyanobacterial blooms in a human-and climatically-impacted world. Life 4 (4), 988–1012.
Paerl, H.W., Paul, V.J., 2012. Climate change: links to global expansion of harmful cyanobacteria. Water Res. 46 (5), 1349–1363.
Ruddick, K., Lacroix, G., 2006. Hydrodynamics and meteorology of the Belgian coastal zone. In: Current Status Of Eutrophication In The Belgian Coastal Zone, pp. 1–15.
Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.-T., Chuang, H.-y., Iredell, M., et al., 2014. The ncep climate forecast system version 2. J. Clim. 27 (6), 2185–2208.
Salgueiro, D., de Pablo, H., Nevesa, R., Mateus, M., 2015. Modelling the thermal effluent of a near coast power plant (Sines, Portugal). Revista de Gestão Costeira Integrada J. Integr. Coast. Zone Manag. 15 (4), 533–544.
Schaeffer, B.A., Liames, J., Dwyer, J., Urquhart, B., 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.
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.
Smit, A.J., Roberts, M., Anderson, R.J., Dufois, F., Dudley, S.F., Bornman, T.G., Olbers, J., Bolton, J.J., 2013. A coastal seawater temperature dataset for biogeographical studies: large biases between in situ and remotely-sensed data sets around the coast of South Africa. PLoS One 8 (12), e81944.
Snyder, J., Ross, E., Weatherbee, R., Thomas, A., Brady, D., Newell, C., 2017. Oyster aquaculture site selection using Landsat 8-derived Sea surface temperature, turbidity, and chlorophyll a. Front. Mar. Sci. 4, 190.
Trinh, R.C., Fichot, C.G., Gierach, M.M., Holt, B., Malakar, N.K., Hulley, G., Smith, J., 2017. Application of Landsat 8 for monitoring impacts of wastewater discharge on coastal water quality. Front. Mar. Sci. 4, 329.
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., Ruddick, K., 2014. Turbid wakes associated with offshore wind turbines observed with Landsat 8. Remote Sens. Environ. 145, 105–115.
Wan, Z., 2014. New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product. Remote Sens. Environ. 140, 36–45.
Wang, F., Qin, Z., Song, C., Tu, L., Kariuki, A., Zhao, S., 2015. An improved monowindow algorithm for land surface temperature retrieval from Landsat 8 thermal infrared sensor data. Remote Sens. 7 (4), 4268–4289.
Yang, J., Gong, P., Fu, R., Zhang, M., Chen, J., Liang, S., Xu, B., Shi, J., Dickinson, R., 2013. The role of satellite remote sensing in climate change studies. Nat. Clim. Chang. 3 (10), 875.
Zeng, X., Zhao, M., Dickinson, R.E., He, Y., 1999. A multiyear hourly sea surface skin temperature data set derived from the YOGA TAO bulk temperature and wind speed over the tropical Pacific. J. Geophys. Res.: Oceans 104 (C1), 1525–1536.