

Comparison of Sea Surface Temperature from SEVIRI and SLSTR satellite instruments over collocations with drifting buoys

Visiting Scientist Report

Christoforos Tsamalis

Met Office Hadley Centre, UK



Version : 1.0

Date : 02/10/2019

Document Change record

Document version	Software version	Date	Author	Change description
1.0			CT	

Table of contents

1.	Executive Summary	4
1.1.	Disclaimer	4
2.	Introduction	5
3.	Datasets and construction of the collocation database	6
3.1.	SEVIRI SST	6
3.2.	SLSTR SST	7
3.3.	OSTIA	9
3.4.	Common collocation database	10
4.	Results	11
4.1.	Breakdown by quality level	11
4.2.	Statistics and histograms	14
4.3.	Time series	17
4.4.	Geographical distribution	19
4.5.	Variation with satellite zenith angle	23
4.6.	Dependence on meteorological variables	26
4.7.	Impact of dust aerosols on the SST retrieval	28
4.8.	Triple collocation approach	30
5.	Conclusions	31
6.	Acknowledgements	33
7.	References	34

1. Executive Summary

The intercomparison between sensors is crucial for assessing the quality of Sea Surface Temperature (SST) datasets. Here, SSTs from SEVIRI, SLSTR and drifting buoys are compared using common collocations of all three instruments. The common collocations have been found taking advantage of the pre-existing collocation databases for SEVIRI and SLSTR, developed by OSI SAF and EUMETSAT, respectively. The drifting buoys have been treated as the reference, while the OSTIA SST analysis have also been examined during night.

The results show that SLSTR SST (quality level 5) compares really well with drifting buoys, although there are some remaining dependencies with total column water vapour during day and probably with dust aerosols. There is also indication of cloud contamination, while according to expectations dual view retrievals perform better than single view. However, the current SLSTR SST does not match yet the quality of AATSR, despite being significantly improved in comparison to previous releases. Not surprising SEVIRI results against drifting buoys (quality levels 3 to 5) are worse than the respective for SLSTR, although in line with previous studies. More specifically, there is clear dependence on zenith angle, indicating that the retrieval is suboptimal especially at high angles, probably not correcting entirely for water vapour at high optical paths. This agrees with the dependence that SEVIRI shows on total column water vapour. In addition SEVIRI SST quality decreases as the Saharan Dust Index becomes more positive, indicating an impact of dust load, although this is less obvious when using the SLSTR Aerosol Dynamic Indicator as metric of dust presence. The statistics of OSTIA are similar to SLSTR, especially the standard deviations. Nevertheless, OSTIA has a stronger dependence on dust aerosols than SEVIRI or SLSTR and it depends on total column water vapour. It is hypothesised that both these dependencies are inherited from VIIRS, which is used as reference in OSTIA.

1.1. Disclaimer

All intellectual property rights of the OSI SAF products belong to EUMETSAT. The use of these products is granted to every interested user, free of charge. If you wish to use these products, EUMETSAT's copyright credit must be shown by displaying the words "Copyright © <2019> EUMETSAT" on each of the products used.

2. Introduction

Sea Surface Temperature (SST) is an important geophysical variable for both Numerical Weather Prediction (NWP) and Climate science. Satellite retrievals of SST are nowadays the main source of observations, because they provide far superior coverage than in situ instruments. However, the quality of the satellite observations needs to be assessed in order to verify that they are fit-for-purpose. Usually, this is done by validating them against SST observations from drifting and moored buoys (e.g. Marsouin et al., 2015; Tsamalis and Saunders, 2018). While also the intercomparison with other satellite SST products is quite common (e.g. Le Borgne et al., 2012; Reynolds et al., 2010).

European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Ocean and Sea Ice Satellite Application Facility (OSI SAF) provides SST observations from the Spinning Enhanced Visible and Infrared Imager (SEVIRI) instrument on board the METEOSAT geostationary satellites positioned above 0 degrees (for the primary satellite) as an operational service in near real time (NRT). Other METEOSAT satellites provide the Rapid Scanning Service (RSS) and the Indian Ocean Data Coverage (IODC), with OSI SAF also producing SST observations for the Indian Ocean region from SEVIRI as a demonstrational product (which is not considered in this study). SEVIRI is a 12 channel imager observing the earth's full disk with a 15 min repeat cycle and a sampling distance of 3 km for nadir view (Schmetz et al., 2002). The OSI SAF SST retrieval for SEVIRI is based on the non-linear split window algorithm using the brightness temperatures at 10.8 and 12 μm channels and a correction based on radiative transfer simulations of NWP profiles (Le Borgne et al., 2011). The SEVIRI SST is provided at hourly frequency with a spatial resolution of 0.05 degrees.

EUMETSAT supplies SST observations from the Sea and Land Surface Temperature Radiometer (SLSTR) on board the polar satellites Sentinel-3 built in the framework of the EU Copernicus program. Currently, two identical (i.e. carrying the same suite of instruments) Sentinel-3 platforms are flying with Sentinel-3A launched on 16th February 2016 and Sentinel-3B launched the 25th April 2018. Note that SLSTR on Sentinel-3B is not considered in this study as it is launched more recently. SLSTR is an evolution of the AATSR instrument on board the ENVISAT satellite and it is currently the only satellite instrument flying with two blackbodies for accurate calibration in the SST range (Donlon et al., 2012b). Another unique aspect of SLSTR is the dual view capability (for part of its swath), which offers improved atmospheric correction to SST retrievals. The SST observations from SLSTR are using either three channels (3.7, 10.8 and 12 μm) during night time or only two (10.8 and 12 μm) during daytime, by applying a linear retrieval to measured brightness temperatures with coefficients estimated from radiative transfer simulations (Merchant, 2012). Given that the SLSTR oblique view swath is about 740 km (centred at the SLSTR nadir point), while the nadir swath is about 1400 km (offset in a westerly direction in order to provide identical coverage with OLCI – the visible/near-infrared imager on Sentinel-3), the SST retrievals are using both views when available. The SLSTR SST is provided at a spatial resolution of 1 km.

The main objective of this study is to intercompare SEVIRI and SLSTR SSTs in order to establish their quality under different observation and meteorological conditions. The intercomparison will be performed over collocations with drifting buoys, which will act as an independent in situ reference. For night time conditions, also OSTIA analysis SST analysis is used in the intercomparison, which assimilates SST observations from drifting buoys and SEVIRI among other in situ and satellite instruments (Donlon et al., 2012a). Also SLSTR SST are assimilated in OSTIA, but only since March 2019. Although OSTIA does not constitute an independent dataset from the satellite instruments or the drifting buoys, it is included in the intercomparison as it is a critical element of NWP systems at the Met Office and ECMWF. The fact that all the observational datasets are collocated permits the application of the triple collocation approach, thus offering an estimation of the random uncertainty for all three instruments (O'Carroll et al., 2008; Zwieback et al., 2012). It is useful to note that the spatial resolution of both SEVIRI and SLSTR is quite similar (0.05° vs 1 km at nadir, respectively), which minimizes as far as possible the collocation uncertainty. This is also the

case in the temporal dimension taking advantage of the high frequency observations of SEVIRI (every 15 min), an advantage of being on a geostationary platform.

In the proposal of this study, there was a second objective about the determination of the dust aerosols impact on SST retrievals using the aerosol optical depth (AOD) from MODIS instrument on board Terra satellite. However, due to the significant effort spent for the construction of the common collocation database, there was not sufficient time left to realise this second objective and thus it is not considered here.

This study covers the time period from 1st August 2016 to 31st July 2019, i.e. 3 years, with the starting date imposed by the availability of the reprocessed SST from SLSTR. The rest of the document is structured in the following way. Firstly, the datasets are described, followed by the approach to create the common collocation database. Then, the results are presented with the conclusions afterwards.

3. Datasets and construction of the collocation database

The first task in order to perform the intercomparison is the construction of a database of collocations between SEVIRI, SLSTR and drifting buoys together with OSTIA for night time conditions. Given the existence of a database already for SEVIRI and drifting buoys developed and maintained by Meteo-France (Marsouin et al., 2015), it makes sense to use it as a starting point. Also, EUMETSAT Secretariat has created a database of collocations between SLSTR and drifting buoys, the initial version of which was built for the Federated Activity between OSI SAF and EUMETSAT Secretariat (2014-2018) (Dybkjaer et al., 2018). This database is now regularly updated and contains also the OSTIA SST observations within. Looking for collocations of satellite with in situ instruments is a rather computationally expensive task, thus in order to accelerate the construction of the common database for the collocations among the datasets and take advantage of existing work, the abovementioned collocation databases for SEVIRI and SLSTR with drifting buoys have been joined. Before describing the merging of the two databases, more detailed information is provided for SEVIRI and SLSTR SST and their collocation databases.

3.1. SEVIRI SST

SEVIRI SST is calculated for all clear-sky water pixels as identified from the Nowcasting (NWC) SAF cloud mask. Given that the measurement repeat cycle of SEVIRI is 15 min, the hourly SST is produced by remapping over a 0.05° regular grid all available SST observations within the hourly time slot, with priority given to the observation closest in time to the product nominal hour. SEVIRI SST follows the GHRSSST (Donlon et al., 2007) nomenclature, which specifies six quality levels for each satellite pixel as a simple mean of filtering the data. The quality level is provided with increasing reliability from 2 (meaning "bad") to 5 (meaning "excellent"), while 0 stands for unprocessed and 1 for cloudy conditions. OSI SAF advises users to choose quality levels 3 to 5 for quantitative applications.

From the OSI SAF SEVIRI collocation database with in situ SST observations, collocations with the following criteria are extracted for this study:

- Selection of drifting buoys only (i.e. no moored boys or ships) located over sea (i.e. filtered out those drifting buoys erroneously located over land).
- Drifting buoys with quality location 0 or 1 (when this information is available), eliminating buoys with location which is characterised as dubious or with missing values.
- Drifting buoys with location quality class 0 to 4 (when this information is available), thus keeping only buoys for which the location radius is known with confidence level 88%.
- Drifting buoys for which the last known position is reported within a time difference of equal or less to 1 h from the position under consideration.

- Removal of drifting buoys with gross errors using the OSI SAF blacklist and removal of duplicate observations.
- Selection of SEVIRI pixel with a drifting buoy located within.
- SEVIRI SST quality level 2 to 5.

In the rest of the study SEVIRI SST with quality level (QL) 3 to 5 will be considered mainly, i.e. as the baseline SEVIRI SST dataset for comparison with SLSTR. However, for further investigations SSTs with QL equal to 2 have also been included in the construction of the common collocation database. SEVIRI is using the Saharan Dust Index (SDI) (Merchant et al., 2006b) in order to mask out dust contaminated pixels or to degrade the SST quality level when the SDI is not too high.

It is important to note that during the period of the 3 years of this study the geostationary satellite changed (and thus the SEVIRI instrument). Until 19th February 2018, the SST observations are obtained from SEVIRI on board METEOSAT-10 and since 20th February 2018 from METEOSAT-11, as this day METEOSAT-11 has been declared EUMETSAT's prime operational geostationary satellite.

Also, it is useful to note that there are some days/periods without collocations at all between SEVIRI and drifting buoys. The reason for the lack of collocations is not clear and it is beyond the scope of this study to establish it for each individual day/period, but it could be due to special activities or issues of the respective satellite or SEVIRI or with the collection of drifting buoy data. These are: 16th October 2016, 16th April 2017, 16th December 2017, 21st December 2017, 7th May 2018, 24th May 2018, 6th November 2018, 8th November 2018, 22nd January to 3rd February 2019, 11th February 2019, 14th to 25th February 2019 and 6th to 9th March 2019. From the abovementioned list, it can be seen that there are many missing days from the SEVIRI-drifting buoys collocation database during the end of January and February 2019.

3.2. SLSTR SST

In contrast to SEVIRI, which uses one type of retrieval algorithm for all SST observations, SLSTR has four algorithms based on the number of channels and views. For the central part of the swath, where the oblique view is available retrievals using both views are applied and these are symbolised as D3 and D2 (D for dual), with the number indicating the number of thermal IR channels used in the SST retrieval. For the rest of the swath, where only the nadir view is available the retrieval SST algorithm is symbolised as N3 and N2 (N for nadir). 3 channels are used for night time conditions (i.e. when solar zenith angle is greater than 90°), while 2 channels are used during daytime (i.e. when solar zenith angle is lower or equal to 90°). This notation of the SLSTR SST retrievals will be used hereafter in the document, as needed. The cloud mask is based on the probabilistic (Bayesian) approach of Merchant et al. (2005), but applied separately for each view.

SLSTR also follows the GHRSSST nomenclature using six quality levels for classifying the SST retrievals. However, the criteria for assigning a SST retrieval to a specific quality level are different from SEVIRI, although their interpretation is the same. EUMETSAT (Gary Corlett, personal communication, 2019) advises users currently to choose SST retrievals with quality level 5, especially for applications that are going to use SLSTR SST as a reference similar to what have been done for the precursor instruments ATSRs (e.g. Blackmore et al., 2012; Le Borgne et al., 2012).

For the moment, there is not a unique SLSTR collocation database with in situ SST instruments covering the whole period of this study, but three slightly different databases had to be used. These are termed as reprocessed, OSI SAF (developed within the study of Dybkjaer et al., 2018) and NRT (near real time) SLSTR collocation databases here. They cover the following periods:

- Reprocessed: 1/8/2016-4/4/2018
- OSI SAF: 5/4/2018-30/11/2018
- NRT: 1/12/2018-31/7/2019

There is no scientific difference among the three databases, except from the fact that the reprocessed database is using analysis meteorological fields from ECMWF, while the other two databases forecast meteorological fields. However, the technical implementation is slightly different. The reprocessed database has one daily file, while the other two databases four files per day (i.e. each one covering 6 hours). Another technical difference is that the prefix of SLSTR variables in the reprocessed database is 's3a_sl_2_wst____r_nt_006__', in the OSI SAF database is 'S3A_SL_2_WST__' (being 'S3A_SL_2_WCT__' only for the total column water vapour) and in the NRT database is 's3a_sl_2_wst____o_nr__'. This latter difference is mentioned here because the combined collocation database of SEVIRI, SLSTR and drifting buoys (see Section 3.4 below), which is a deliverable of this study, has inherited these names for the SLSTR variables.

The SLSTR collocation databases with in situ SST instruments are constructed by extracting 21 by 21 SLSTR pixels centred over the in situ location within a time window of 2 hours. It is useful to note that there is possibility of no valid SLSTR SST retrievals over these 21 by 21 pixels, as no filtering has been applied to them. This provide the opportunity to users to apply the filtering criteria of their choice. In this study, the following criteria have been applied (in the following order) for the collocations of SLSTR with drifting buoys:

- Selection of drifting buoys only (as the SLSTR collocation databases are built around different files for each in situ instrument, the files containing only drifting buoys have been selected).
- Selection only of SLSTR pixels out of 21 by 21 pixels from each collocation with SLSTR SST quality level 3, 4, or 5.
- Selection of pixels with cloud probability for nadir view less than 0.1. The same criterion is applied for the oblique view when such probability is provided (i.e. in the dual view part of the swath).
- Finally, select the pixel which is closer to the drifting buoy position, but considering that the distance between the two should be less or equal to 5 km.

Regarding the distance criterion, 5 km has been chosen in order to be similar to the distance used in SEVIRI-drifting buoys collocations. For SEVIRI the requirement for collocation with drifting buoys was within the SST pixel, which is 0.05° , i.e. roughly 5 km. There are some days or periods in the SLSTR-drifting buoys collocation databases for which there are no collocations at all. The reason behind these missing days is not known and it is out of the scope of this study to identify it. However, certainly some of them are due to missing data from SLSTR, for example due to outgassing. For the three-year period of this study the missing days from the SLSTR-drifting buoys collocation databases are: 1st to 3rd October 2016, 12th November 2016, 17th November 2016, 23rd November 2016, 16th February 2017, 1st to 3rd August 2017, 16th to 17th January 2018, 16th to 18th February 2018, 30th June 2018, 5th August 2018, 14th to 16th September 2018, 24th to 25th September 2018, 3rd to 17th October 2018, 10th to 13th November 2018, 23rd November 2018, 12th to 18th January 2019, 2nd to 5th March 2019, 25th March 2019, 21st to 24th May 2019, 10th June 2019 and 17th June 2019. It can be noted that in October 2018 and January 2019, there are about two and one weeks, respectively, of missing data.

It is useful to note that currently there is an artefact in the reprocessed SLSTR-drifting buoys collocation database, which prohibits collocations over some regions. These areas are depicted in Figure 1, roughly indicated by the red lines for daytime (left) and night time (right) conditions. The areas with no collocations are easily discerned by comparing between the two conditions in Figure 1, as there is gap in one and collocations in the other. EUMETSAT has been informed about the artefact in the reprocessed SLSTR-drifting buoys collocation database and confirmed its existence. However, neither the OSI SAF nor the NRT collocation databases are affected by this artefact.

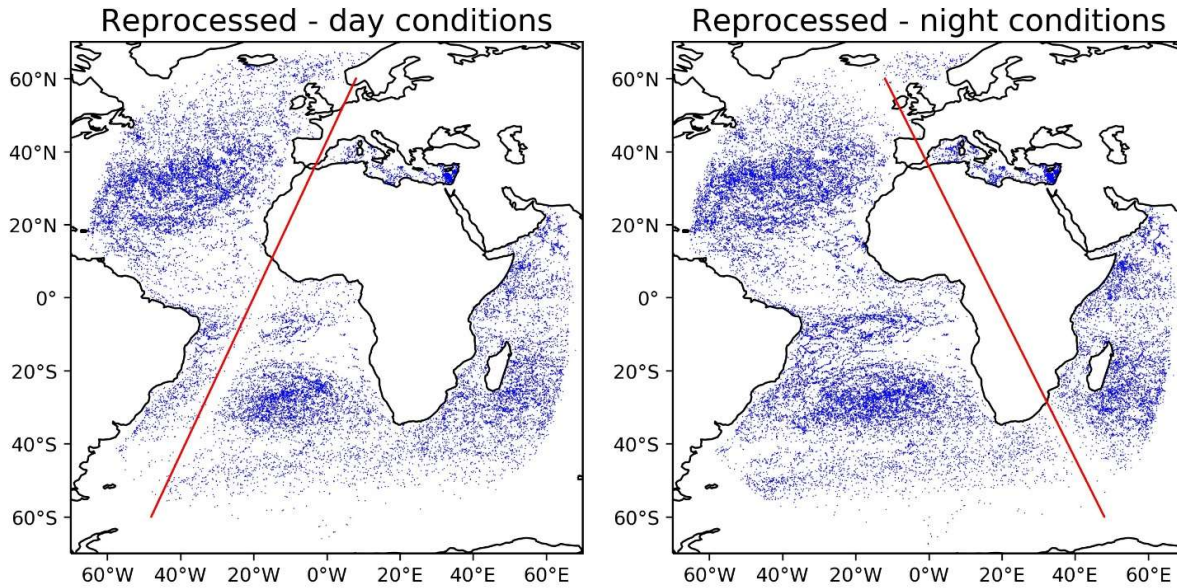


Figure 1: The geographical distribution of collocations from the common database (Section 3.4) for the period 1st August 2016 to 31st March 2018 (i.e. covering the SLSTR reprocessed database only). SEVIRI SSTs with quality level 2 to 5 and SLSTR SSTs with quality level 3 to 5 are used for daytime [left] (defined here as SEVIRI solar zenith angle less than 90°) and night time [right] (defined here as SEVIRI solar zenith angle greater than 90°) conditions. The red lines indicates roughly, where collocations between SLSTR and drifting buoys are missing in the reprocessed database.

3.3. OSTIA

OSTIA stands for Operational SST and Sea Ice Analysis. The OSTIA system uses SST observations from IR and MW satellite instruments, together with drifting buoys to produce a SST analysis by merging all the SST products (Donlon et al., 2012a). The SSTs from the satellite instruments are bias-corrected against an IR satellite reference and drifting buoys. The blending of the observations is performed using a multi-scale optimal interpolation method, which later has been replaced by NEMOVAR. OSTIA provides a foundation SST with a spatial resolution of 0.05°. Because foundation SST is free of diurnal variability and representative of a depth around 10 m, comparisons with SEVIRI, SLSTR and drifting buoys are examined only during night, when their interpretation is more straightforward. Indeed, SEVIRI and SLSTR using IR channels are sensitive to the skin SST (representative of a depth around 10 μ m), while drifting buoys are measuring the depth SST at a depth of around 20 cm. It is important to note that the reported SST from SEVIRI is adjusted to be on average equal to the depth SST during night-time, as measured from drifting buoys, while SLSTR SST is the skin SST. The OSTIA SST is included in the SLSTR-drifting buoys collocation databases (Section 3.2) and use of this parameter will be done in this study, as it is already collocated with SLSTR. The OSTIA system evolved over time, thus it is important to mention its evolutions during the 3-year period considered here as it could be helpful for the interpretation of the results. These are:

- September 2016 – Change of the reference satellite instrument from AVHRR on MetOp-B to VIIRS on S-NPP.
- February 2018 – Change of the assimilation system from multi-scale optimal interpolation scheme to NEMOVAR.
- March 2019 – Start assimilating SLSTR on Sentinel-3A and improvement of the SST feature resolution.

3.4. Common collocation database

As mentioned above given the existence of collocation databases for both SEVIRI and SLSTR with drifting buoys and the effort put to create these, it makes totally sense to merge them. As the common element in both databases is the observations of the drifting buoys, the following criteria have been applied in order to find the common observations of drifting buoys in both databases:

- Longitude and latitude absolute difference less or equal to 0.01 degrees.
- The same drifting buoy identification number.
- SST absolute difference less than 0.1 K.
- Time absolute difference less or equal to 1 min.

By applying the abovementioned criteria the relevant parameters from the SLSTR databases (mainly related to SLSTR, but also OSTIA SST and the meteorological fields from ECMWF) have been added to SEVIRI-drifting buoys daily collocation files. Thus, the common collocation database contains all the SEVIRI-drifting buoys parameters and observations (briefly presented in Section 3.1) and additional parameters from the SLSTR-drifting buoys collocation databases. However, there are valid entries from the SLSTR database only when there are common collocations in both original databases. The software that has been used to create the common collocation database is written in Python and it has been transferred to OSI SAF together with the netCDF files. The daily filename of the common collocation database is following the structure: “sstmbw_meteosatSS_YYYYMMDD_YYYYMMDD.nc”, where SS is either 10 or 11 (for METEOSAT10 or METEOSAT11), YYYY is the year, MM is the month and DD is day.

Figure 2 presents the collocations of SEVIRI and SLSTR with drifting buoys for two days in January 2017 as an example. It can be observed that some collocations from SEVIRI (blue dots) are not common with SLSTR (orange square), as the common collocations are marked with a red circle. There are two reasons for this. Firstly, only SEVIRI observations that have been processed and are not cloudy (i.e. with quality level at least 2) are considered for collocations with drifting buoys, while in the SLSTR database all possible collocations with drifting buoys are included initially even if all SLSTR pixels are cloudy. Secondly, SLSTR is on a polar satellite with a swath of about 1400 km, which means that it does not cover fully the SEVIRI disk even by combining all orbits within a day. This can be seen more clearly by comparing the SEVIRI only collocations between the two days in Figure 2.

The common database contains 105803 collocations of SEVIRI, SLSTR and drifting buoys. However, some of those drifting buoys SST observations have gross errors. In order to eliminate them a simple filter has been applied (Marsouin et al., 2015), removing drifting buoys with absolute SST difference from the OSTIA climatology (Roberts-Jones et al., 2012) larger or equal to 5 K. This filter removes 165 collocations, which are not considered further in this study, but do exist in the common collocation database.

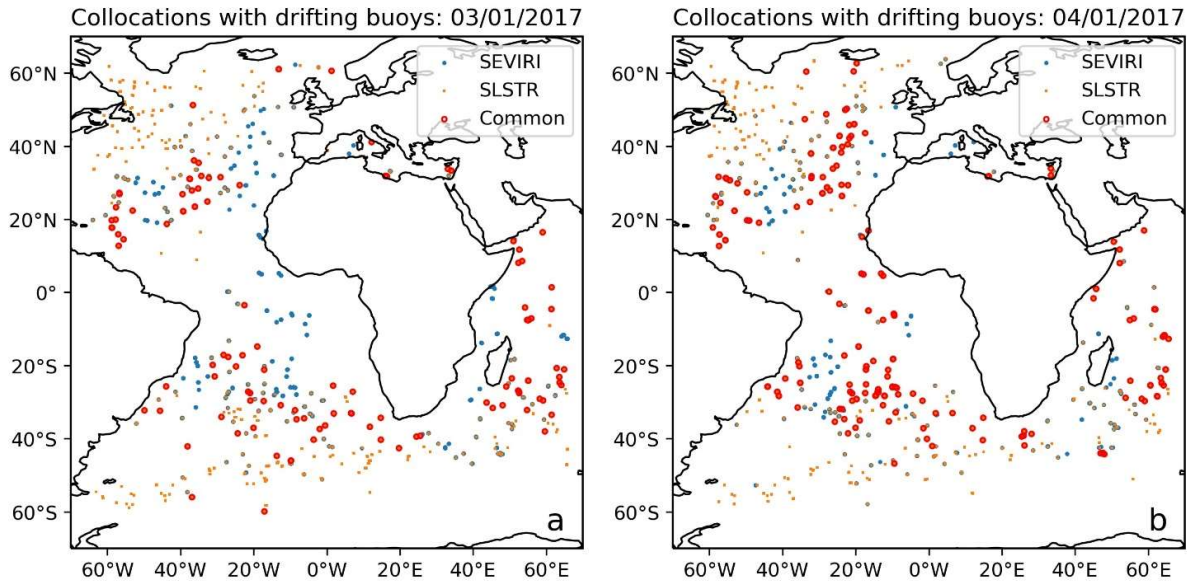


Figure 2: Examples of collocations of SEVIRI (blue dot) and SLSTR (orange square) with drifting buoys for two days of January 2017 the 3rd (left) and the 4th (right). The red circle indicates the common collocations of SEVIRI, SLSTR and drifting buoys.

4. Results

For the analysis of the results, day and night conditions will be examined separately. The separation between day and night will be performed adopting the same criteria as for SEVIRI and using its solar zenith angle information:

- Day: SEVIRI solar zenith angle less or equal to 90°.
- Night: SEVIRI solar zenith angle greater or equal to 110°.

This means that collocations with SEVIRI solar zenith angle larger than 90° and lower than 110° are not considered in the following analysis.

As it is mentioned above, in this study the focus will be on SSTs from SEVIRI with quality level 3 to 5 and from SLSTR with quality level 5. However, firstly a very basic assessment will be performed of all the quality levels for both satellites. In general, no single-sensor error statistics (SSES) are applied to the SST measured by the satellites, although the impact of applying the SSES will be briefly discussed later for SLSTR.

4.1. Breakdown by quality level

Table 1, Table 2, Table 3 and Table 4 present few basic statistic measures of the comparison between SEVIRI or SLSTR and drifting buoys for day and night conditions stratified according the different quality levels of SST from the satellite instruments. For this section only the reference quality level (QL) for SEVIRI is 2 to 5 and for SLSTR 3 to 5. The results cover the 3-year period (1/8/2016-31/7/2019) of this study.

SEVIRI QL	μ (K)	σ (K)	m (K)	s (K)	N
2 to 5	-0.03	0.67	0.01	0.49	47729
2	-0.13	0.84	-0.08	0.65	19090
3	0.01	0.63	0.02	0.55	6444
4	0.03	0.52	0.04	0.47	9001
5	0.05	0.41	0.07	0.34	13194

Table 1: Mean (μ), standard deviation (σ), median (m), robust standard deviation (s) and number of collocations (N) of the SST differences between SEVIRI and drifting buoys for different quality levels (QL) of SEVIRI and SLSTR QL 3 to 5 under daytime conditions.

SLSTR QL	μ (K)	σ (K)	m (K)	s (K)	N
3 to 5	-0.12	0.54	-0.10	0.30	47729
3	-0.35	0.68	-0.28	0.50	2374
4	-0.21	0.82	-0.21	0.49	11261
5	-0.07	0.37	-0.07	0.25	34094

Table 2: Mean (μ), standard deviation (σ), median (m), robust standard deviation (s) and number of collocations (N) of the SST differences between SLSTR and drifting buoys for different quality levels (QL) of SLSTR and SEVIRI QL 2 to 5 under daytime conditions.

For the reference QL SEVIRI has a mean (median) difference against drifting buoys of -0.03 (0.01) K and standard deviation (robust standard deviation) of 0.67 (0.49) K during daytime (Table 1). The respective values for SLSTR are -0.12 (-0.10) K and 0.54 (0.30) K (Table 2). As expected, the mean/median difference against drifting buoys is close to 0 K for SEVIRI during the day, as it is adjusted to the SST of drifting buoys. On the other hand, SLSTR has smaller standard deviations in comparison to SEVIRI, a clear indication of its better performance when comparing with drifting buoys, although for the same QL the standard deviations are quite similar between SEVIRI and SLSTR. It can be seen from Table 1 that for SEVIRI QL equal to 2 the mean/median difference against drifting buoys becomes colder at about -0.1 K. For the remaining SEVIRI QL it is positive (SEVIRI warmer than drifting buoys), while it is gradually increasing with increasing SEVIRI QL. On the other hand, the standard deviations decrease with increasing SEVIRI QL, with the standard deviation reducing from 0.84 (QL=2) to 0.41 (QL=5). This indicates the expected behaviour and confirms that the QL stratification for SEVIRI works sufficiently well. The number of available collocations is larger for SEVIRI QL=2, being 19090, followed by QL=5 (i.e. 13194).

The respective statistics for SLSTR and its distinctive quality levels keeping all SEVIRI QLs (i.e. 2 to 5) during daytime are shown in Table 2. The mean/median difference is significantly colder for QL=3 at about -0.3 K, but reduces with increasing QL reaching -0.07 K for QL=5. Surprisingly, the standard deviation for QL=4 is larger (0.82 K) than the respective one for QL=3 (0.68 K), however this behaviour is not reflected in the robust standard deviation where its value decreases with increasing QL, although the robust standard deviation for QL equal to 3 and 4 are almost indistinguishable. The vast majority of SLSTR collocations with drifting buoys are of QL=5 (34094), followed by QL=4 (11261).

The results for night time are presented in Table 3 and Table 4. For SEVIRI QL 2 to 5, it can be seen that the SEVIRI mean/median differences are colder by at least 0.1 K in comparison to the respective values during day (Table 1), while the night time standard deviations are slightly larger than the respective ones during

day. This is probably due to more challenging cloud detection during night because of the absence of visible channels, despite the fact that during day additional variability is expected in the comparisons due to the diurnal thermocline. SLSTR mean/median values for the reference QL are also colder by about 0.08 K in comparison to the respective values during the day (Table 2). However, the standard deviations are significantly reduced in comparison to daytime values, in accordance with expectations as three channels are used in the SST retrievals. SLSTR performs better than SEVIRI in term of standard deviations for QL=5, but the opposite is true for QL=4, certainly indicating an issue with SLSTR QL=4. It is not surprising that the mean/median difference of SLSTR is colder than the respective of SEVIRI, as the former provides a skin SST, while the latter is adjusted to depth SST. It is interesting to note that the number of night time collocations (53802) is larger than the respective for daytime.

By changing the SEVIRI QL but keeping the reference SLSTR QL, the same conclusions as for daytime conditions hold, but during night the mean/median differences are always negative. In addition, now the mean (median) difference for QL=2 is significantly colder being -0.37 (-0.29) K (Table 3), which is indicative of cloud contamination.

Regarding the stratification of SLSTR QL it can be observed that the vast majority of collocations (50850) are of QL=5, with only 2952 being of QL=4 and none of QL=3 (Table 4). The latter point is not surprising and it is an outcome of selection criteria of SLSTR SST (probability of cloud less than 0.1) and the definition of night time conditions (solar zenith angle larger than 110°). As expected, the statistics of SLSTR against drifting buoys improves substantially with increasing SLSTR QL, and this is also the case for SEVIRI (Table 3).

SEVIRI QL	μ (K)	σ (K)	m (K)	s (K)	N
2 to 5	-0.19	0.71	-0.11	0.55	53802
2	-0.37	0.90	-0.29	0.73	21069
3	-0.11	0.67	-0.08	0.61	7068
4	-0.09	0.54	-0.06	0.49	10333
5	-0.05	0.43	-0.02	0.37	15332

Table 3 : Mean (μ), standard deviation (σ), median (m), robust standard deviation (s) and number of collocations (N) of the SST differences between SEVIRI and drifting buoys for different quality levels (QL) of SEVIRI and SLSTR QL 3 to 5 under night time conditions.

SLSTR QL	μ (K)	σ (K)	m (K)	s (K)	N
3 to 5	-0.21	0.39	-0.17	0.25	53802
3	--	--	--	--	0
4	-0.40	0.71	-0.31	0.53	2952
5	-0.20	0.36	-0.17	0.24	50850

Table 4: Mean (μ), standard deviation (σ), median (m), robust standard deviation (s) and number of collocations (N) of the SST differences between SLSTR and drifting buoys for different quality levels (QL) of SLSTR and SEVIRI QL 2 to 5 under night time conditions.

Given that the big majority of SLSTR collocations with drifting buoys (and SEVIRI), especially during night, are with QL equal to 5 and EUMETSAT (Gary Corlett, personal communication 2019) recommends to use

mainly them, in the rest of this study only SLSTR data with QL=5 will be used. The same holds for SEVIRI, as OSI SAF recommends using SEVIRI SSTs with QL from 3 to 5 and most of the collocations belong to these QLs.

4.2. Statistics and histograms

Before moving on to the presentation of the results, it is useful to examine the time difference and the distance between the collocations of the three instruments. The results cover the 3-year period, so this will not be repeated again in this section. Figure 3 shows the time difference for the collocations of the three instruments. Outside, the different number of collocations (ordinate axis), there is no practical difference between day and night. It can be observed that both for SLSTR (orange) and SEVIRI (green) the vast majority of collocations with drifting buoys occurs within ± 30 min. For SLSTR the distribution is quite uniform in this interval, but for SEVIRI there are big spikes every about 7.5 min. On the other, the distribution of the collocations between SLSTR and SEVIRI (blue) looks more like a Gaussian with the big majority occurring within ± 60 min. This is not surprising given that the merging of the two collocation databases (i.e. SEVIRI-Drifting buoys and SLSTR-Drifting buoys) has been based on common drifting buoys in both databases.

Figure 4 shows the histograms of the distance for the collocations between pairs of the three instruments. Again there is no significant difference between night and day, outside the number of collocations which is larger during night. It is interesting to note that the distance for the majority of collocations between SLSTR and drifting buoys (orange) is smaller than 1 km, i.e. within the pixel of SLSTR. The distance distributions between SLSTR-SEVIRI (blue) and SEVIRI-Drifting buoys (green) are very similar, with most of them being within 3 km.

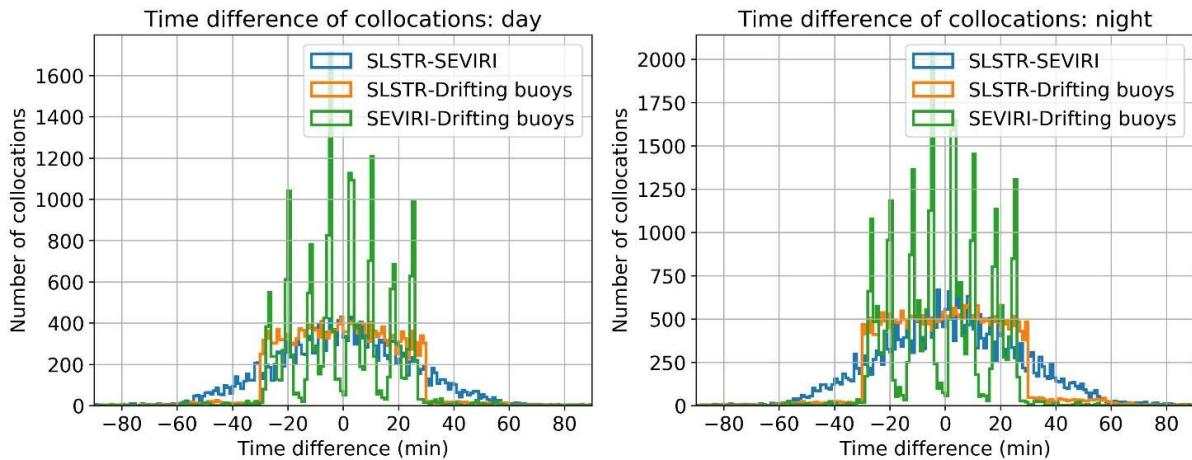


Figure 3: Histograms of time difference for the collocations between pairs of the three instruments. The left image is for day conditions, while the right is for night conditions.

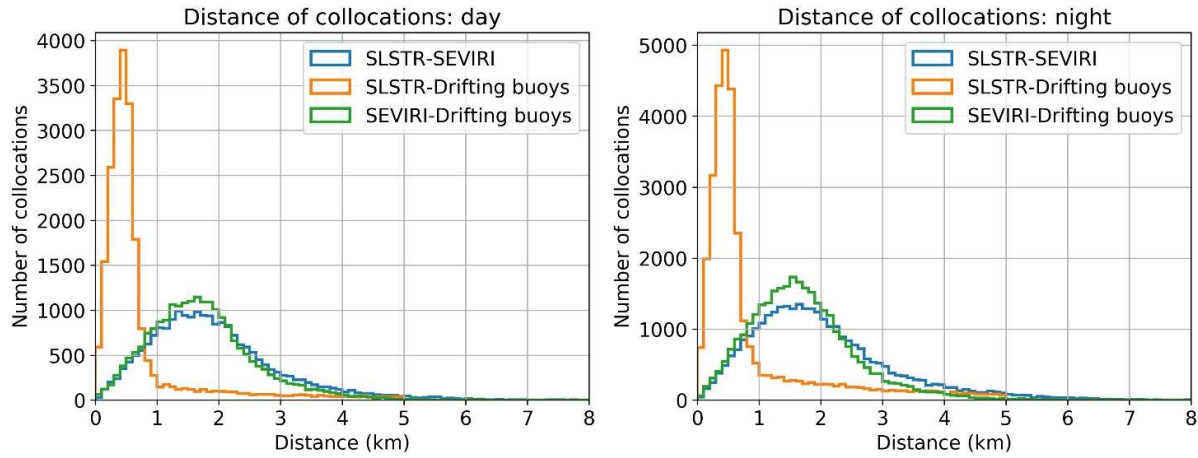


Figure 4: Histograms of distance for the collocations between pairs of the three instruments. The left image is for day conditions, while the right is for night conditions.

After establishing that the collocations both over time and space are as close as possible, a fact that minimizes the collocation uncertainties, Figure 5 shows the histograms of the SST differences against drifting buoys. The SLSTR distribution (blue) is narrower than the SEVIRI one (orange), both during night and day. Also, in both conditions SLSTR is colder than drifting buoys, while SEVIRI is slightly warmer than drifting buoys only during daytime. During night, the OSTIA distribution (green) against drifting buoys is rather close to SLSTR. This behaviour is reflected on the statistics of Table 5 (for day conditions) and Table 6 (for night conditions), which includes not only the results against the drifting buoys but also for the other pairs of datasets. Indeed, during day mean (median) difference for the pair SLSTR-Drifting buoys is -0.05 (-0.06) K with standard deviation (robust standard deviation) 0.34 (0.24) K. For the pair SEVIRI-Drifting buoys, the mean (median) difference is 0.05 (0.06) K with standard deviation (robust standard deviation) 0.49 (0.42) K. Interestingly, the standard deviation (robust standard deviation) for the pair SLSTR-SEVIRI is 0.54 (0.46) K, with both of them being larger than the respective deviations for the pairs SLSTR-Drifting buoys or SEVIRI-Drifting buoys, while the mean (median) is the sum of the other two pairs.

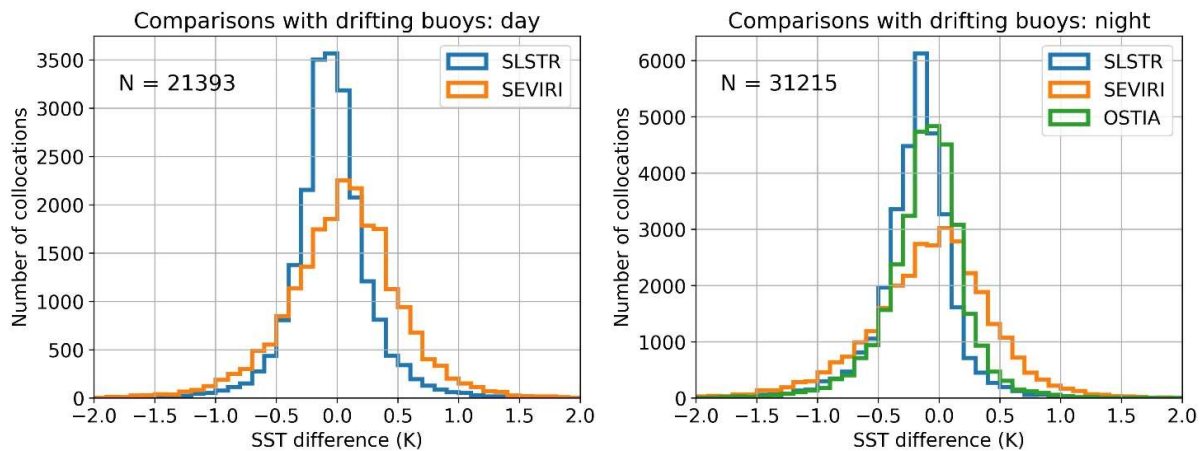


Figure 5: Histograms of the SST difference of SLSTR (blue), SEVIRI (orange) or OSTIA (green - during night only) with drifting buoys for day (left) and night (right) conditions.

During night there are more collocations than during day, 31215 vs. 21393, although it is not clear why is the case (Table 6). The mean (median) difference of SLSTR-Drifting buoys is -0.19 (-0.17) K, which is colder than during day as expected (e.g. Donlon et al., 2007), while the standard deviation (robust standard deviation) is 0.34 (0.22) K, very similar to daytime. The pair SEVIRI-Drifting buoys has mean (median) difference of -0.07(-0.04) K, with standard deviation (robust standard deviation) 0.52 (0.44) K. The respective values for OSTIA with drifting buoys are -0.11 (-0.09) K and 0.35 (0.25) K. Especially, the standard deviations for OSTIA-Drifting buoys are really close to the values of SLSTR-Drifting buoys, although these for SLSTR are slightly better. Probably, this is also because the collocation uncertainty for the pair SLSTR-Drifting buoys is minimal given the distribution of the distance for SLSTR (Figure 4) and the horizontal resolution for SLSTR (1 km) and OSTIA (0.05°), as the drifting buoys SSTs are considered point measurements. This becomes more clear when comparing directly SLSTR with OSTIA for which the standard deviation (robust standard deviation) is 0.42 (0.30) K, higher than the respective ones with drifting buoys. The larger standard deviations are observed for the pairs with SEVIRI (in the range 0.40-0.55 K), with the largest one being for SLSTR and SEVIRI.

Datasets	μ (K)	σ (K)	m (K)	s (K)
SLSTR-SEVIRI	-0.10	0.54	-0.12	0.46
SLSTR-Drifting buoys	-0.05	0.34	-0.06	0.24
SEVIRI-Drifting buoys	0.05	0.49	0.06	0.42

Table 5: Mean (μ), standard deviation (σ), median (m) and robust standard deviation (s) of the SST differences between two datasets for SEVIRI QL 3 to 5 and SLSTR QL 5 under daytime conditions. The number of collocations is common between any of two datasets being 21393.

Datasets	μ (K)	σ (K)	m (K)	s (K)
SLSTR-SEVIRI	-0.12	0.55	-0.15	0.47
SLSTR-Drifting buoys	-0.19	0.34	-0.17	0.22
SLSTR-OSTIA	-0.08	0.42	-0.10	0.30
SEVIRI-Drifting buoys	-0.07	0.52	-0.04	0.44
SEVIRI-OSTIA	0.04	0.47	0.06	0.40
OSTIA-Drifting buoys	-0.11	0.35	-0.09	0.25

Table 6: Mean (μ), standard deviation (σ), median (m) and robust standard deviation (s) of the SST differences between two datasets for SEVIRI QL 3 to 5 and SLSTR QL 5 under night time conditions. The number of collocations is common between any of two datasets being 31215.

It is worthwhile to note that application of SLSTR SSES bias on the SLSTR SSTs does not modify significantly the statistics (not shown). The results are the same with or without the SSES bias at two decimal places for all SLSTR statistics of Table 5 and Table 6, although differences appear at three decimal places.

As SLSTR has four algorithms for the SST retrieval it is useful to investigate the performance of each one of them. Note that there is also a fifth algorithm for SLSTR using the nadir view and 3 channels, which is

robust to stratospheric aerosols, but it is not activated in the normal SLSTR SST product and thus not considered in this study. Table 7 presents the same statistics as previously, with N2 and D2 retrievals applied during daytime conditions and N3, D3 only during night conditions. N2 and D2 results (mean, standard deviation and median) are practically the same as in Table 5. However, the robust standard deviation is higher for N2, being 0.28 K, while for D2 is 0.21 K, when for all daytime conditions is 0.24 K. In addition, the number of collocations of D2 is about twice the number of N2. On the other hand, there is a clear stratification between N3 and D3, with N3 having colder mean (median) difference being -0.25 (-0.19) K than D3 being -0.16 (-0.15) K and slightly larger robust standard deviation, 0.24 vs. 0.21 K. The results of Table 6 for the pair SLSTR-Drifting buoys are within the respective N3 and D3.

Algorithm	μ (K)	σ (K)	m (K)	s (K)	N
N2	-0.06	0.35	-0.06	0.28	7149
N3	-0.25	0.34	-0.19	0.24	12743
D2	-0.05	0.34	-0.05	0.21	14237
D3	-0.16	0.34	-0.15	0.21	18472

Table 7: Mean (μ), standard deviation (σ), median (m), robust standard deviation (s) and number of collocations (N) of the SST differences between SLSTR QL 5 and drifting buoys for the different SLSTR retrieval algorithms. The results of N2 and D2 are for daytime conditions, while the N3 and D3 results for night time.

In this section overall statistics for the collocations have been presented, however it is important to investigate how the statistics vary with time, geographically, satellite zenith angle and other parameters, which are known to affect the quality of SST retrievals. All these factors will be examined in the following sections.

4.3. Time series

Figure 6 depicts the monthly time series of the median difference and the robust standard deviation against drifting buoys for day and night conditions. SLSTR is shown in blue, SEVIRI in orange, OSTIA in green (only during night), while the number of common collocations each month is shown by the black dashed line. It can be observed that during day the number of collocations increases gradually from just above 400 in the second half of 2016 reaching over 1000 in summer 2019, but in January and February 2019 the number of collocations is very low even reaching 0 (see also about missing days in Sections 3.1 and 3.2). The same is true during night for January and February 2019, while in general the number of collocations is between 800 and 1000, with a higher number sometimes over 1200 during the period of the OSI SAF SLSTR database (Section 3.2) and in summer 2019.

During day, the SLSTR median difference from drifting buoys is about -0.05 K always negative, while for SEVIRI is about 0.1 K. However, both time series appear rather noisy, more for SEVIRI, despite the fact that the median is quite robust to fluctuations (more than the mean) and the number of collocations is over 400. On the other hand, during night the time-series are less noisy, with SLSTR having a median difference between -0.15 and -0.2 K, SEVIRI being about -0.05 K and OSTIA -0.1 K. These values are in agreement with the results of Table 5 and Table 6. The results for the mean values (not shown) are rather close to the median values during day, while during night the mean time-series are about ~0.2 K colder than the median ones.

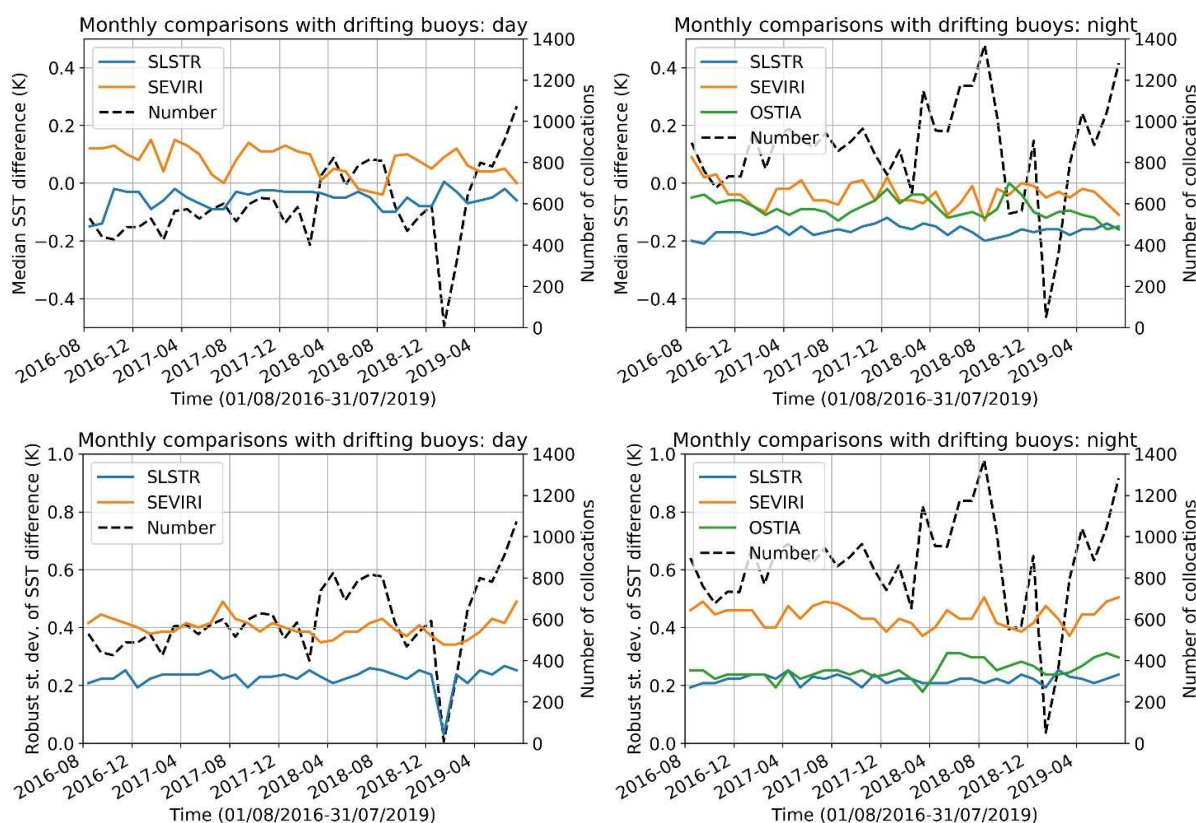


Figure 6: Monthly time series of the median SST differences against drifting buoys (top) and the robust standard deviation (bottom) during day (left) and night (right). The results are for all the common collocations with SEVIRI QL 3 to 5 and SLSTR QL 5. SLSTR is shown in blue, SEVIRI in orange and OSTIA in green (only during night). The monthly number of collocations is shown with the dashed black line with the values displayed on the right ordinate.

The robust standard deviations are less noisy than the median values. For SLSTR against drifting buoys the robust standard deviation is just above 0.2 K rather stable over time with similar values during night. The almost 0 K value during day in January 2019 is obviously due to the limited number of collocations (black dashed line) and it should not be interpreted as a real feature. The SEVIRI robust standard deviation (orange line) against drifting buoys is about 0.4 K during day and somehow larger during night. The OSTIA robust standard deviation (green line – during night) is practically the same as SLSTR until March 2018, then it increases by about 0.05 K and stays roughly stable until the end of July 2019. This increase coincides with the change from the reprocessed SLSTR database to the OSI SAF (Section 3.2), but there is no change in how OSTIA parameter is stored/treated in all three SLSTR databases (Gary Corlett, personal communication, 2019). The change of OSTIA assimilation system from multi-scale optimal interpolation scheme solved using the analysis correction method (Donlon et al., 2012a) to NEMOVAR occurs earlier in February 2018 (Section 3.3) than the change observed in Figure 6. However, based on the results from the SST Quality Monitoring (SQUAM) webpage (<https://www.star.nesdis.noaa.gov/sod/sst/squam/index.php>), the introduction of NEMOVAR improves the performance of OSTIA against drifting buoys. This is not observed in Figure 6 and instead there is the step increase. Thus, there is no explanation currently regarding this change of the OSTIA-Drifting buoys robust standard deviation. Also, it is important to note

that the standard deviations (not shown) are significantly larger than the respective robust standard deviations by 0.1-0.15 K, while the step change for OSTIA is less obvious, although present.

4.4. Geographical distribution

The geographical distribution of the median and the robust standard deviation of SST differences against drifting buoys are shown in Figure 7 for daytime conditions for SEVIRI and SLSTR. Figure 9 presents the respective results for night time conditions, with the addition of the comparisons for OSTIA. Figure 8 displays the number of collocations for day and night. The number of collocations is larger during night, as more grid boxes have yellowish colour than during day. The majority of collocations are found in the latitude zones 20°-40° both in the northern and the southern hemisphere. The number of collocations is small in the northern tropics, especially during day when many gridboxes have less than 20 collocations and appear hatched in Figure 7, Figure 8 and Figure 9. Although, results are presented for all gridboxes with at least four collocations, those that are hatched should be interpreted with caution.

Figure 7 indicates that the median SST difference of SEVIRI against drifting buoys is rather homogeneous geographically being about 0.1 K with somehow warmer gridboxes being about 0.3 K around South and Eastern Africa. On the other hand, the robust standard deviation of SEVIRI has an obvious pattern with lower values (~ 0.3 K) close to the sub-satellite point at 0° and increasing with satellite zenith angle reaching 0.6-0.7 K at the edge of SEVIRI disk. The SLSTR results during day are more homogeneous geographically in comparison to SEVIRI for both the median (being about -0.1 K or closer to 0 K) and the robust standard deviation (being ~ 0.2 K). The same results hold also for the mean and the standard deviation (not shown), although the respective maps are somehow more noisy than the robust statistics measures.

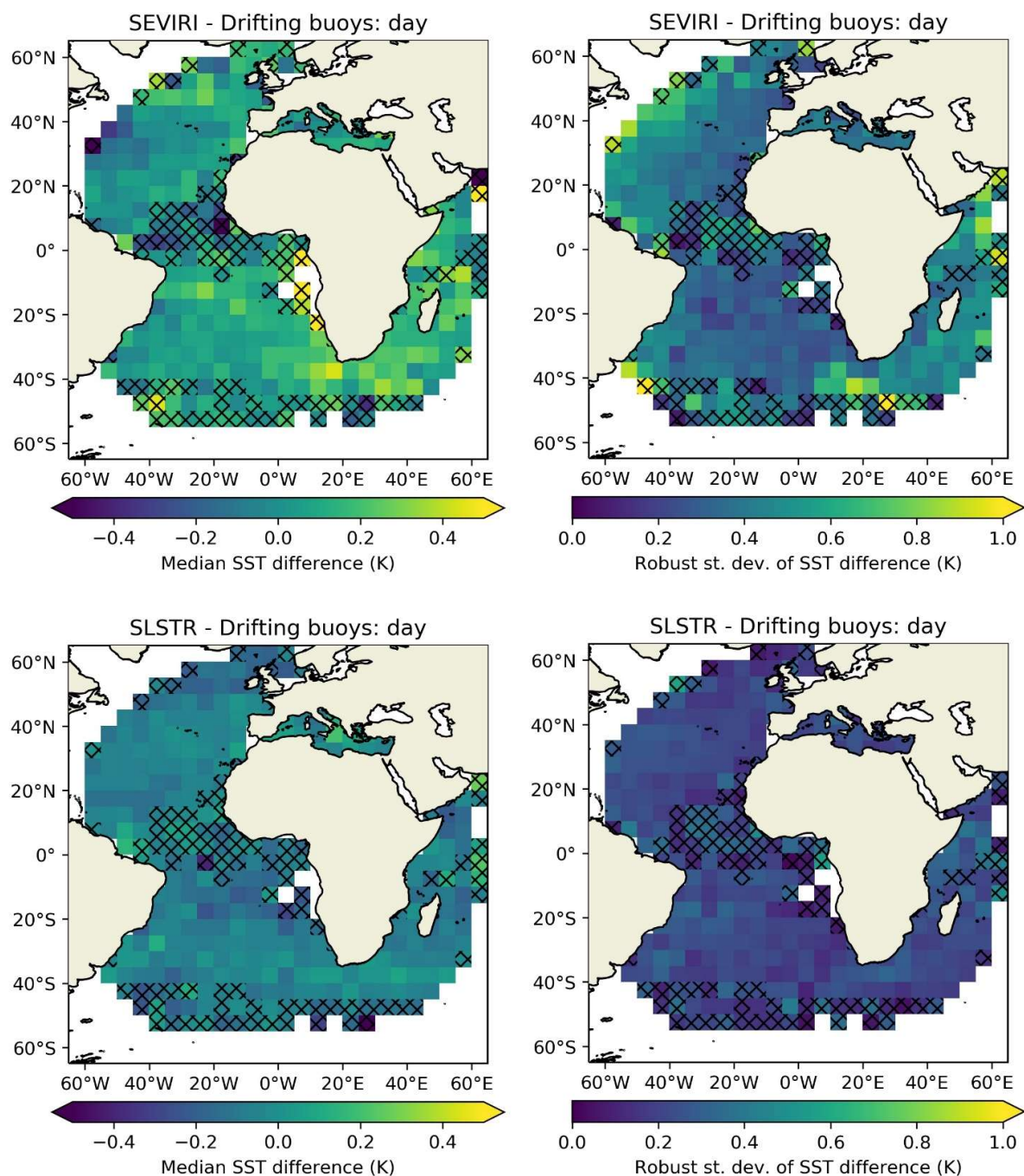


Figure 7: Geographical distribution of the median SST differences against drifting buoys (left) and the robust standard deviation (right) during daytime conditions. The upper panel shows the results for SEVIRI and the lower for SLSTR. The grid box size is 5°, while results are plotted only for grid boxes with at least 4 collocations. Grid boxes with less than 20 collocations are hatched.

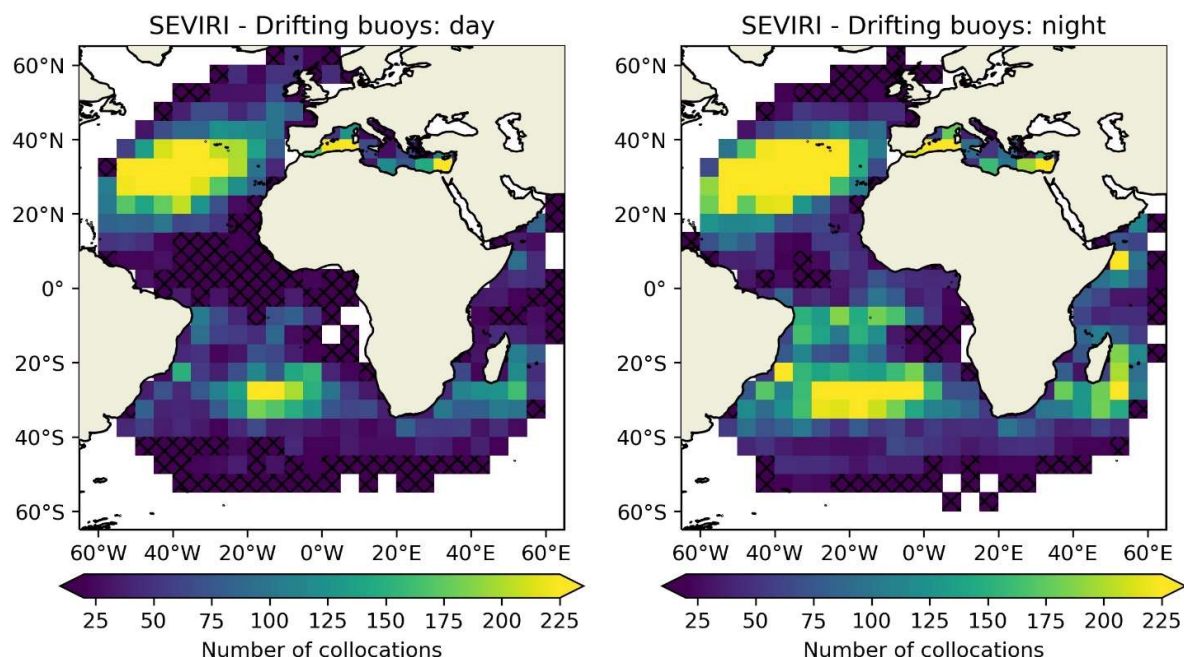
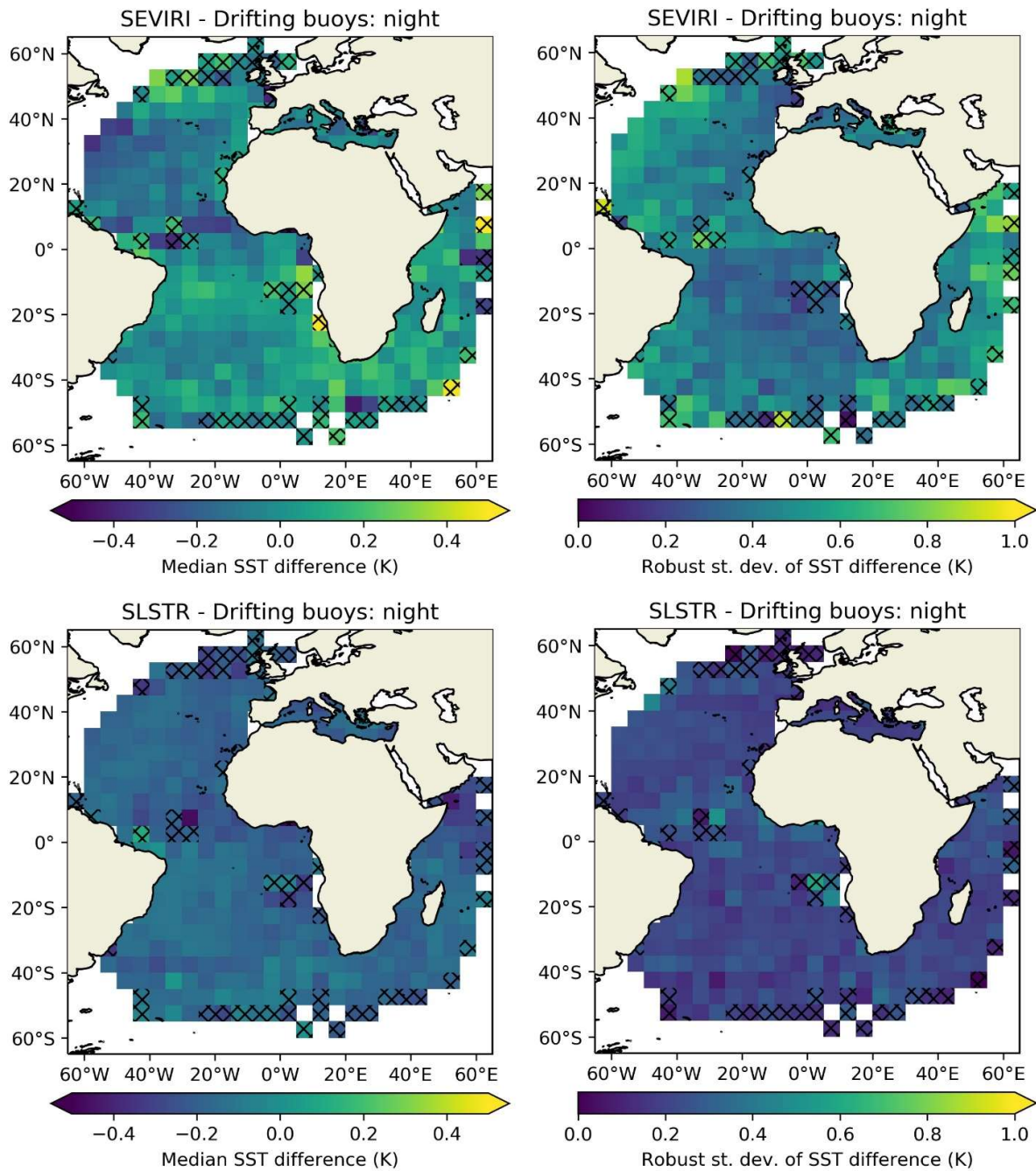


Figure 8: Geographical distribution of the number of collocations for day (left) and night (right) conditions. Although, the specific maps are for SEVIRI and drifting buoys, as the collocations are common, the same distribution applies to the remaining datasets. The grid box size is 5°, while results are plotted only for grid boxes with at least 4 collocations. Grid boxes with less than 20 collocations are hatched.

The maps during night, now including also OSTIA, are shown in Figure 9. The median difference of SEVIRI against drifting buoys has slightly colder values over the North Atlantic (about -0.1 K) than the rest of the SEVIRI disk (about 0 K). This is also reflected with higher values for the robust standard deviation over the North Atlantic, while like for the daytime robust standard deviation there is an increase with satellite zenith angle, although probably less significant than during daytime. Both the median difference and the robust standard deviation are homogeneous geographically for SLSTR, with values about -0.2 K and 0.2 K, respectively. Some colder values of the median difference at about -0.3 K are observed over the Arabian Sea, which could be due to dust aerosols contamination on the SST retrieval in this region (e.g. Peyridieu et al., 2013). In addition, the median difference of OSTIA against drifting buoys is rather homogeneous over the SEVIRI disk with value close to -0.1 K, with the exception over the northern tropical Atlantic, where the Saharan Air Layer is located (e.g. Tsamalis et al., 2013), and a little bit less obvious over the Arabian Sea. This is not surprising given that OSTIA uses either AVHRR or VIIRS as reference for its bias correction scheme (Section 3.3), which are single view instruments and as such susceptible to dust aerosols contamination. On the other hand, the OSTIA robust standard deviation displays some geographical patterns with enhanced values (0.5-0.6 K) off West and South Africa and low values (0.1-0.2 K) over the southern tropical Atlantic. The mean differences maps (not shown) are similar to the median, but the

standard deviation maps during night (not shown) are noisier, and thus somehow more difficult to extract conclusions from them.

The abovementioned results are in agreement with them of the previous sections.



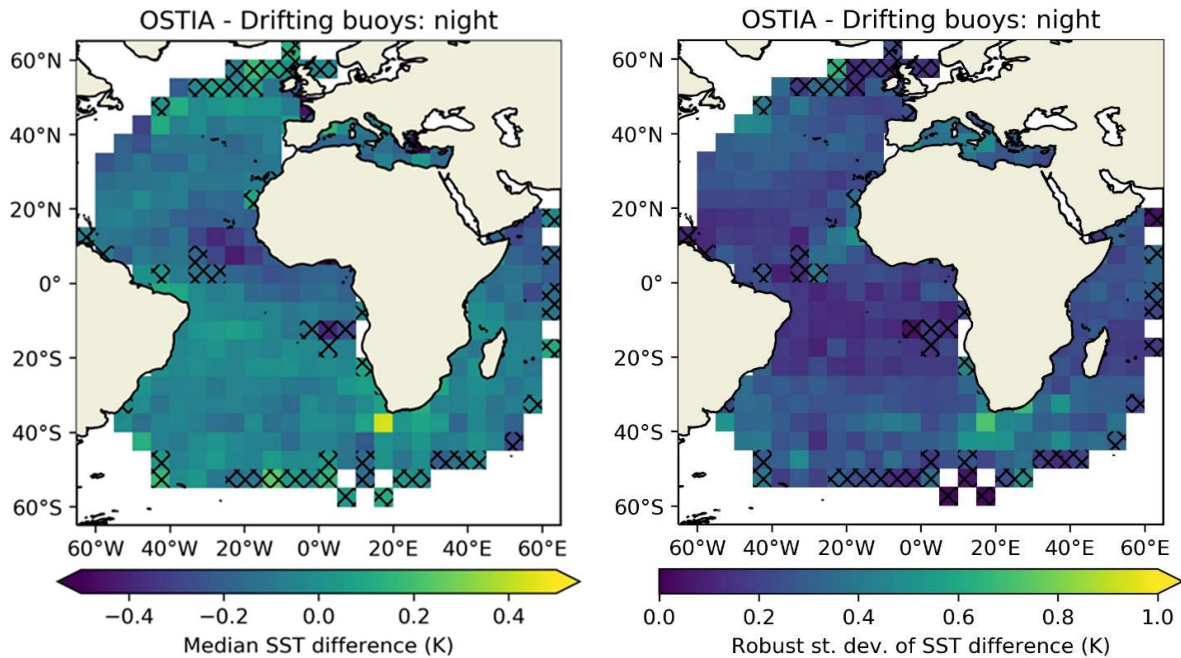


Figure 9: Geographical distribution of the median SST differences against drifting buoys (left) and the robust standard deviation (right) during night conditions. The upper panel shows the results for SEVIRI, the middle for SLSTR and the lower for OSTIA. The grid box size is 5°, while results are plotted only for grid boxes with at least 4 collocations. Grid boxes with less than 20 collocations are hatched.

4.5. Variation with satellite zenith angle

It is interesting to investigate how the statistical measures vary with satellite zenith angle. At larger zenith angles the optical path increases and the assumptions made for the retrievals or even the radiative transfer codes used for the retrievals become less accurate. An indication of that was the geographical variability of SEVIRI-Drifting buoys robust standard deviation discussed in the previous section, especially during day.

Figure 10 and Figure 11 present the median differences and robust standard deviations against the SEVIRI zenith angle and SLSTR zenith angle, respectively. Obviously, it is expected that any variation of SEVIRI zenith angle would not affect the SLSTR results, and should have a minor impact (if any) on OSTIA results, given that SEVIRI SST is assimilated into OSTIA. Similarly, the variability of SLSTR zenith angle should not impact either SEVIRI or OSTIA, as these are independent from SLSTR (OSTIA assimilates SLSTR but only for the last 5 months out of 3 years considered here, so any potential impact would be minimal). However, in order to verify this independence, results are presented for all datasets.

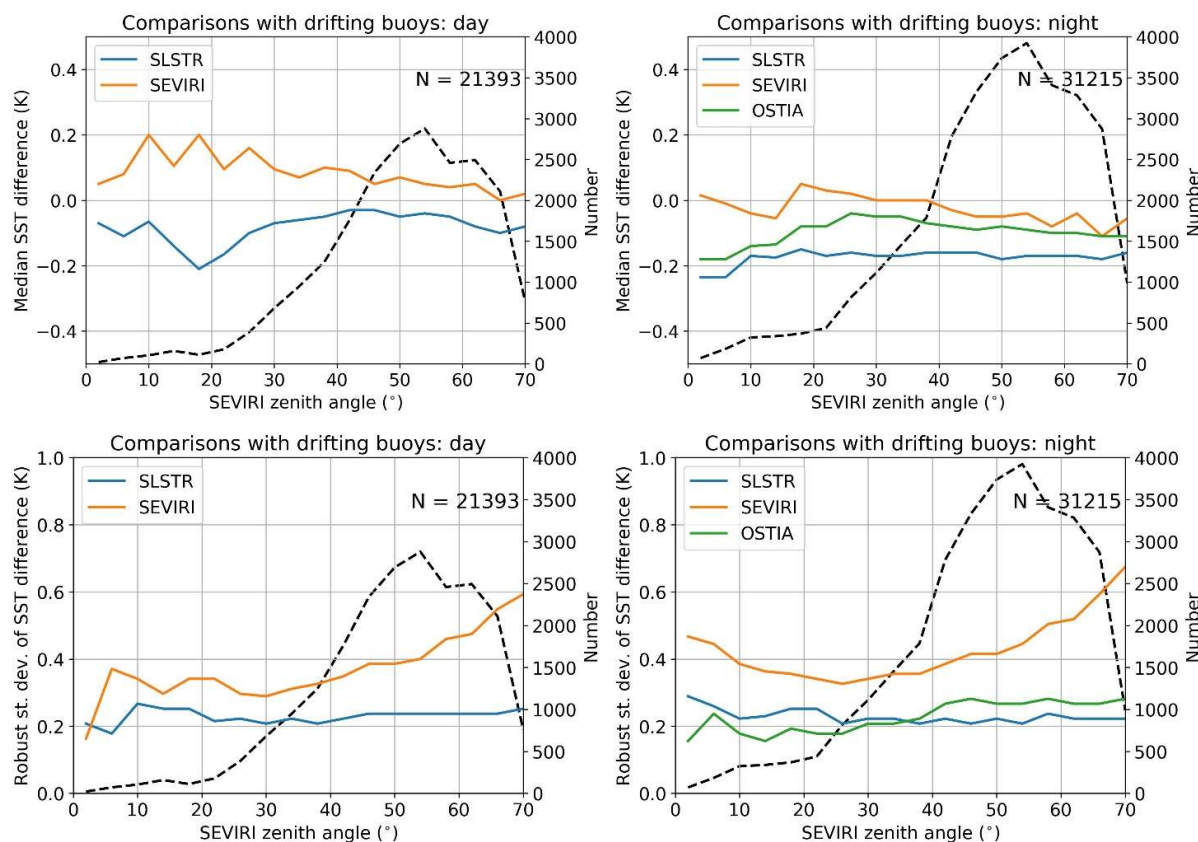


Figure 10: Median SST differences against drifting buoys (top) and the robust standard deviation (bottom) during day (left) and night (right) versus the SEVIRI zenith angle. The results are for all the common collocations with SEVIRI QL 3 to 5 and SLSTR QL 5. SLSTR is shown in blue, SEVIRI in orange and OSTIA in green (only during night). The distribution of the number of collocations is shown with the dashed black line with the values displayed on the right ordinate. The zenith angle step is 4°.

SEVIRI median difference seems to decrease slightly with SEVIRI zenith angle both during day and night. For zenith angles less than 20° there are few collocations, so this decrease is more noisy (Figure 10). During day, for SEVIRI zenith angles between 10° and 30° (which geographically are located in the longitude-latitude box: 30°W-30°E, 30°S-30°N), SLSTR-drifting buoys median difference becomes colder reaching -0.2 K, an indication of cloud or dust contamination in SLSTR N2 and/or D2 retrievals. This can be observed also in Figure 7, although it is less obvious. On the other hand, during night there is not such indication for SLSTR with the median difference being almost a straight line. OSTIA median difference against drifting buoys initially increases from about -0.2 K to -0.05 K until about 25° and then slightly decreases. This is indicative of a latitude dependence for OSTIA.

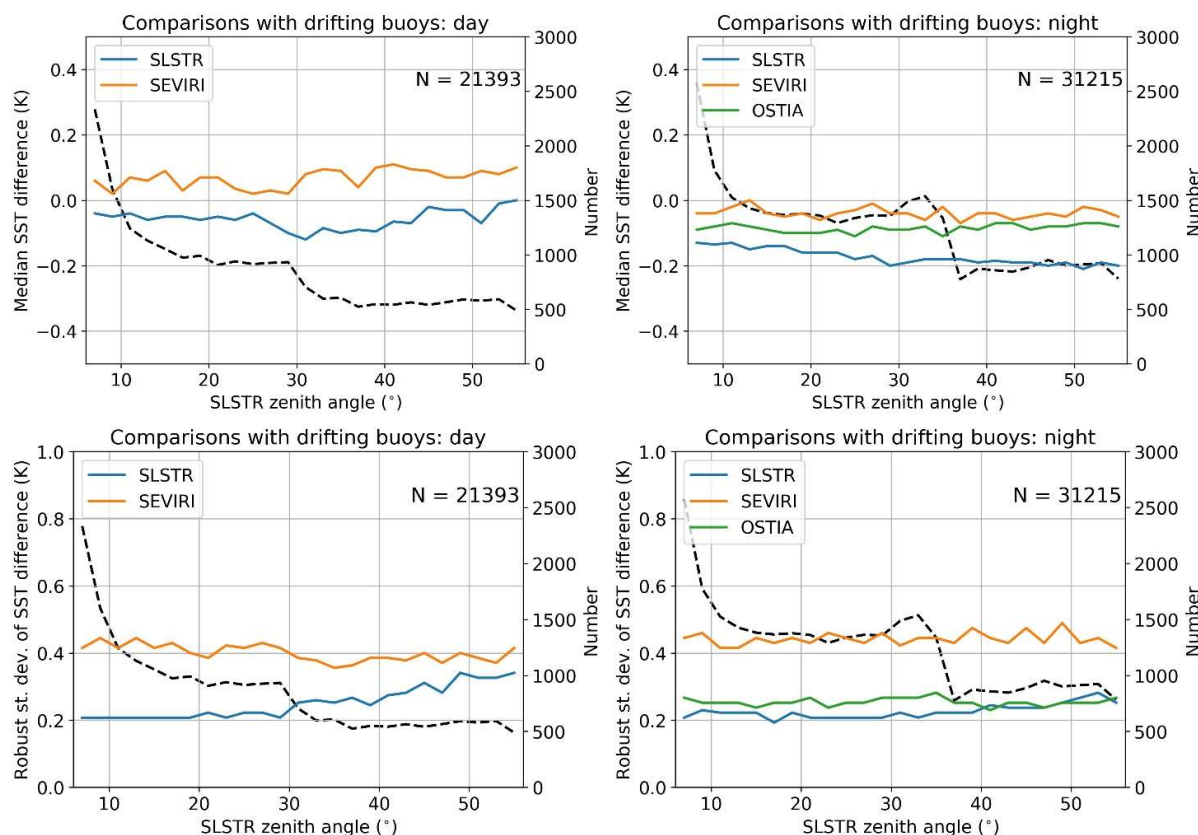


Figure 11: Similar to Figure 10, but now the zenith angle is for SLSTR. The zenith angle step is 2°.

The robust standard deviation for SLSTR against drifting buoys is independent from SEVIRI zenith angle both during day and night. As already seen in previous section, the SEVIRI robust standard deviation increases with its zenith angle, a clear indication that the applied SST retrieval and/or cloud mask is not that efficient at high optical paths. However, the change of the robust standard deviation with zenith angle is not linear for SEVIRI, but rather parabolic with the minimum occurring at about 25°. This is more obvious for night time conditions and it could be related to cloud contamination. OSTIA robust standard deviation indicates a slight increase with SEVIRI zenith angle, which as for the median difference indicates a latitude dependence.

Figure 11 presents the respective results, but for the SLSTR zenith angle. Not surprisingly, SEVIRI and OSTIA do not show obvious dependence on SLSTR zenith angle neither for the median difference nor for the robust standard deviation against drifting buoys. It is useful to note that the SLSTR zenith angle starts from 6°, as it can be seen by the abscissa, given that the instrument by design never sees the true nadir. The instrument design explains also the shape of the number of collocations versus the zenith angle. A significant number of collocations occurs for zenith angles less than 10°-15°, while there is a step-like decrease at about 30°-35° where the retrieval SST changes from dual view (D2 or D3) to nadir view (N2 or N3). The median difference of SLSTR-drifting buoys during day is about stable, while during night there is a slight decrease with increasing zenith angle from about -0.15 K to -0.2 K. During day, the robust standard deviation is stable for the D2 retrievals at 0.2 K, but for N2 retrievals is increasing from about 0.25 K to 0.35 K. Similarly during night, it is stable at 0.2 K for the D3 retrieval and incrementally increasing for the N3 retrievals, reaching about 0.25 K at the edge of the swath.

4.6. Dependence on meteorological variables

Wind speed is known that affects the stratification of the SST with depth (e.g. Donlon et al., 2002) and thus the intercomparison between different datasets; given that different instruments/analyses provide SST at different depths. Furthermore, water vapour is the main atmospheric constituent under cloud-free conditions that defines the quality of SST retrievals from satellite observations (e.g. Merchant et al., 2006a). Thus, it is useful to examine how the comparison results with drifting buoys vary with both wind speed and water vapour, for the latter in terms of total column water vapour.

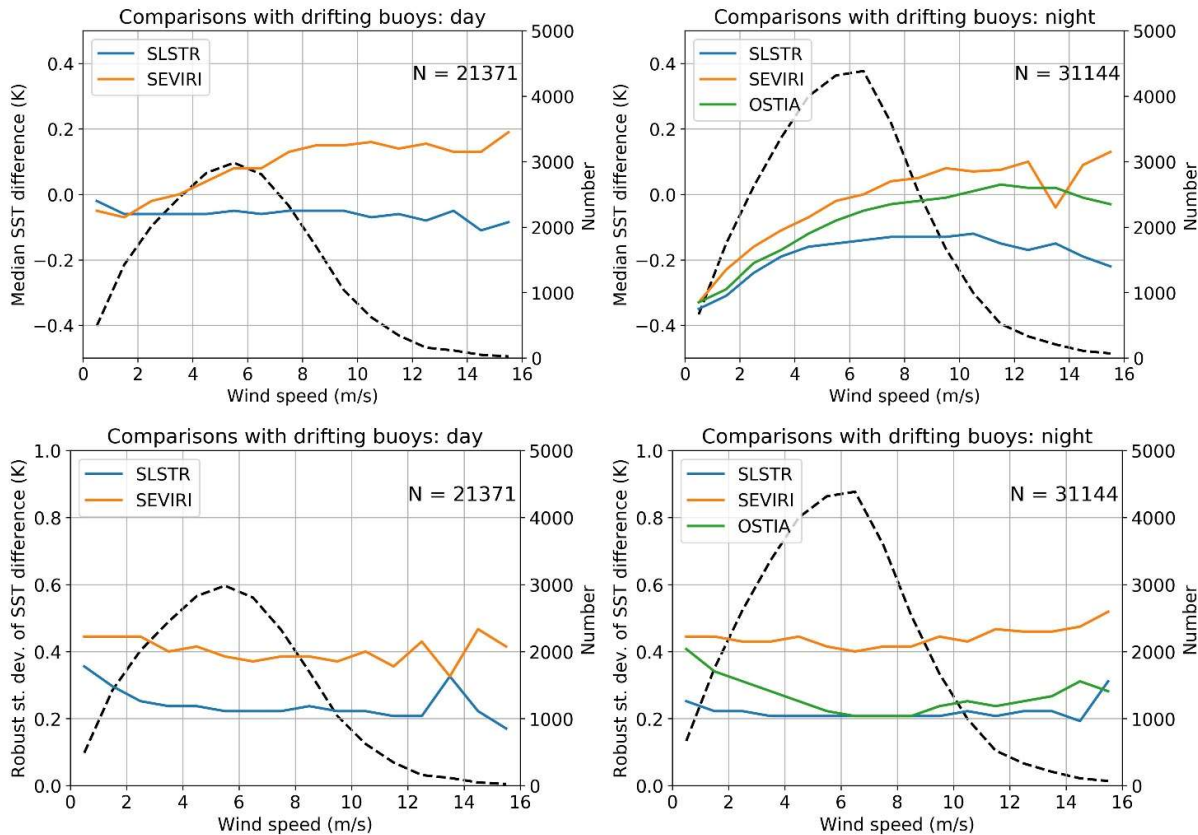


Figure 12: Median SST differences against drifting buoys (top) and the robust standard deviation (bottom) during day (left) and night (right) versus ECMWF wind speed. The results are for all the common collocations with SEVIRI QL 3 to 5 and SLSTR QL 5. SLSTR is shown in blue, SEVIRI in orange and OSTIA in green (only during night). The distribution of the number of collocations is shown with the dashed black line with the values displayed on the right ordinate. The wind speed step is 1 m/s.

Figure 12 presents the robust statistics measures against the wind speed. During the day the median difference of SLSTR-Drifting buoys is practically stable, while for SEVIRI there is a speed dependence up to 8 m/s. This is a surprising result for SLSTR, given that measures skin SST a wind dependence of the SLSTR-Drifting buoys SST difference is expected. Note also, that for SEVIRI the median difference is increasing with

wind speed and then stabilising, while the opposite is expected based on previous studies (e.g. Donlon et al., 2002). For night time conditions, SLSTR has the expected behaviour, i.e. initially large median difference decreasing with wind speed until about 6 m/s and then more or less a stable value of about -0.15 K. SEVIRI has similar form to SLSTR, although it does not seem to stabilise with increasing wind speed. Astonishingly, the same behaviour to SEVIRI holds for OSTIA, taking into account that during night OSTIA should measure the same SST as drifting buoys.

The robust standard deviation of SEVIRI-Drifting buoys does not depend on wind speed having a value of about 0.4 K (Figure 12). Idem, for SLSTR during night with a value of 0.2 K. However, during day the SLSTR robust standard deviation decreases from ~0.35 K initially to 0.2 K at 6 m/s and then stays at this level. OSTIA robust standard deviation has a more complicating form, like a parabola, with higher value of 0.4 K at calm conditions decreasing to 0.2 K for the wind speed interval 6 to 9 m/s and then increasing again reaching 0.3 K at 15 m/s.

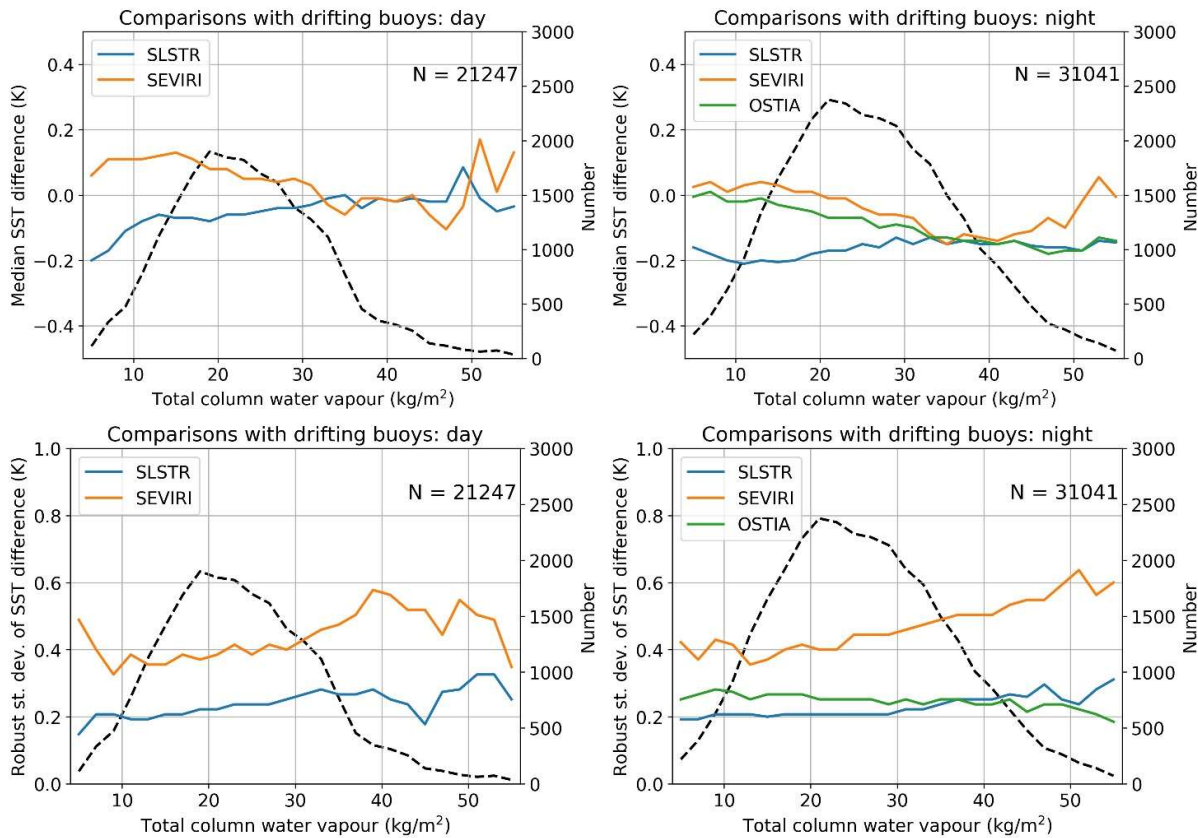


Figure 13: Similar to Figure 12, but now for total column water vapour from ECMWF. The total column water vapour step is 2 kg/m².

Both SEVIRI and SLSTR show a dependence with total column water vapour (TCWV) during day (Figure 13), with the SLSTR median difference reaching 0 K (from -0.2 K initially) above 35 kg/m². SEVIRI median difference with drifting buoys also reaches about 0 K (from 0.1 K initially) at the same level of TCWV. The positive slope of SLSTR with TCWV, probably is due to the dual view capability and SST retrievals. Similar behaviour, for both SEVIRI and SLSTR is observed during night time, although the variation is very slight for

SLSTR. OSTIA median difference becomes colder in a linear fashion as the TCWV increases. On the other hand, the robust standard deviation of OSTIA is independent from TCWV at about 0.25 K. The same is almost true for SLSTR during night with a minor increase from 0.2 K to 0.25 K when the TCWV becomes above 35 kg/m². However, the robust standard deviation of both SEVIRI and SLSTR (during day only) increases as the TCWV increases, more so for SEVIRI during night.

4.7. Impact of dust aerosols on the SST retrieval

It is known that volcanic and dust aerosols affect the quality of SST retrievals from IR instruments (e.g. Merchant et al., 1999; Le Borgne et al., 2013). During the 3-year period of this study, there were not major volcanic eruptions, which could emit aerosols reaching the stratosphere, where they stay for a long period, thus affecting SST retrievals. However, dust outbreaks over the ocean from the big deserts are quite common, especially from Sahara and the Arabian Peninsula, both of them being inside the SEVIRI disk. In order to detect the existence of dust aerosols and mitigate their impact on SEVIRI SST retrievals, Merchant et al. (2006b) created the Saharan Dust Index (SDI) for night time conditions, which uses the 3.9, 8.7, 10.8 and 12 μm channels. During day, OSI SAF developed a modified version of SDI, which uses the 8.7, 10.8, 12 and 13.4 μm channels, in order to mitigate the contamination of the 3.9 μm channel by solar radiation (Saux-Picart, 2018). Good et al. (2012) building on the SDI work have created another two indexes of dust aerosols applicable to ATSRs. There is one formulation applicable to D2 and D3 retrievals (using the 11 and 12 μm channels and both views) and a second one applicable to N3 retrievals (using the 3.7, 11 and 12 μm channels and the nadir view only). In the case of SLSTR, these two indexes are also used although referred to as Aerosol Dynamic Indicator (ADI). The second formulation of the ADI is applied also to N2 retrievals, despite the fact that the SST retrieval itself uses only the 11 and 12 μm channels (Gary Corlett, personal communication, 2019). Both SDI and ADI do not have units.

Figure 14 presents the median difference and the robust standard deviation versus SDI. It is useful to note that the number of collocations is 16910 during day and 25156 during night, significantly lower than the number of collocations in Table 5 and Table 6, being 21393 and 31215, respectively. This is because the SDI is calculated only for the latitudes between 30°S and 50°N, where dust aerosols are expected, thus not covering the whole SEVIRI disk. Note that SDI depends both on the dust aerosols load and their altitude in a complex way (Merchant et al., 2006b), but in general it is mostly the positive values of SDI that indicate significant presence of dust aerosols.

The SLSTR-Drifting buoys median difference during day is only weakly dependent on SDI, very slightly decreasing until SDI becomes about 0.1 and then very slightly increasing. During night there is no obvious dependence for SLSTR. SEVIRI median difference has the same form both during day and night (the same as SLSTR during the day), decreasing linearly with SDI until SDI becomes about 0, where it starts to increase also linearly but now with steeper slope (especially during day). On the other hand, OSTIA median difference decreases linearly with SDI being about 0 K when SDI equals -1 and reaching about -0.2 K when SDI is 0.4. Both for SEVIRI and OSTIA, there is a clear indication of dust aerosols impacting the quality of their SSTs.

For both SLSTR and OSTIA there is no dependence of the robust standard deviation on the SDI, being constant just above 0.2 K. For SEVIRI, the same is true when the SDI is negative with the robust standard deviation being 0.4 K, but increases to more than 0.6 K when the SDI becomes 0.4.

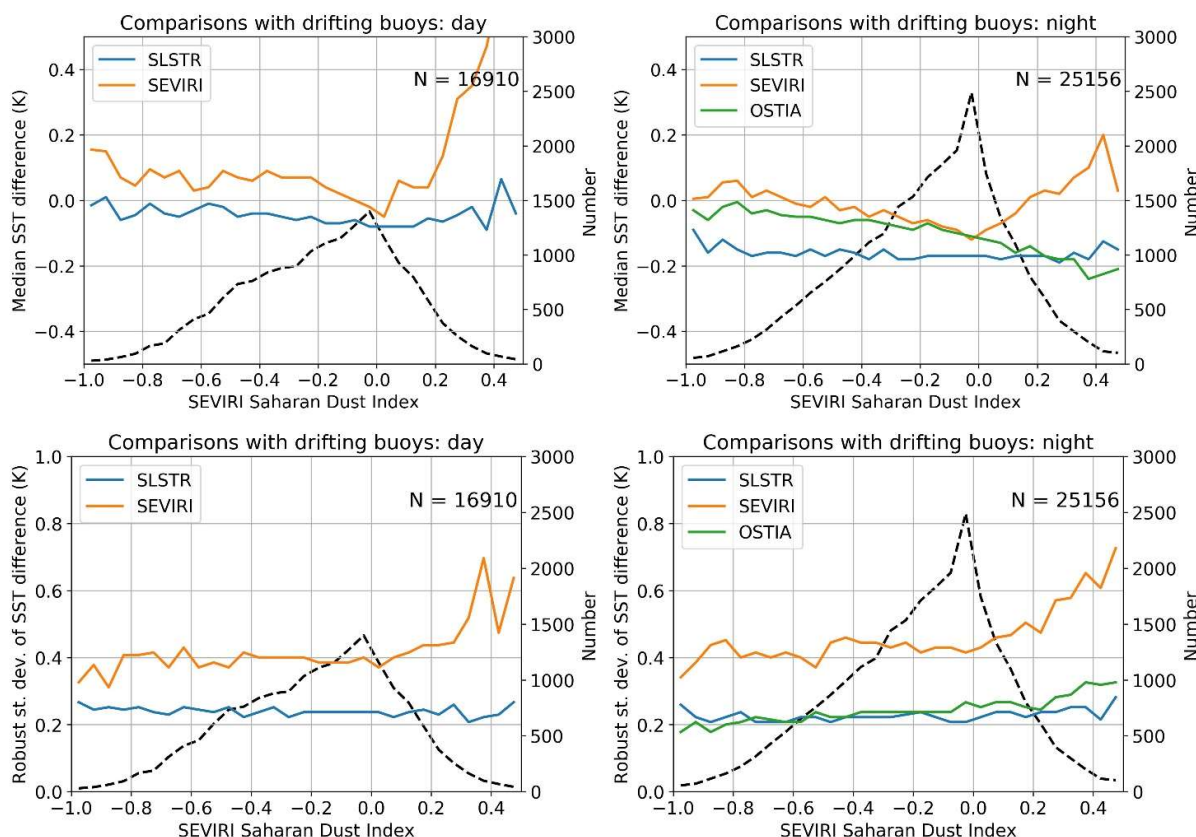


Figure 14: Median SST differences against drifting buoys (top) and the robust standard deviation (bottom) during day (left) and night (right) versus SEVIRI Saharan Dust Index (SDI). The results are for all the common collocations with SEVIRI QL 3 to 5 and SLSTR QL 5. SLSTR is shown in blue, SEVIRI in orange and OSTIA in green (only during night). The distribution of the number of collocations is shown with the dashed black line with the values displayed on the right ordinate. The SDI step is 0.05.

Examining now the dependence against the SLSTR ADI (Figure 15), the results are not entirely consistent as with SEVIRI SDI discussed above. For example, the SEVIRI-Drifting buoys median difference during day is almost constant at about 0.05 K (except from ADI values around -0.3, which is about 0.15 K), where there was an obvious dependence with SEVIRI SDI. On the other hand, the SLSTR-Drifting buoys median difference varies with ADI being positive when ADI is less than -0.2, slightly negative and constant and then further reducing as the ADI becomes larger than 0.1, again inconsistent with the almost constant results during day in Figure 14. During night, the median difference of OSTIA decreases as the ADI increases, something consistent with the results of SDI. For SEVIRI and SLSTR the results are more complex, but still denoting a dependence with ADI. For SLSTR in general the median difference becomes less negative as the ADI increases. SEVIRI median difference initially becomes slightly less negative with ADI approaching 0 K, when ADI is about 0 and then starts to become colder reaching -0.2 K when ADI is about 0.2. The robust standard deviation of SEVIRI during day is about constant at 0.4 K, while this of SLSTR has a particular shape (parabola-like), being 0.2 K when the ADI is between -0.1 and 0.1, which is a significant percentage of the collocations (see black dashed line). As with SDI, OSTIA robust standard deviation does not vary with ADI, being about 0.25 K. Both SEVIRI and SLSTR robust standard deviations with ADI have very similar form, with

this of SEVIRI being about 0.2 K larger of the SLSTR one, taking higher values when the ADI is larger than 0.15.

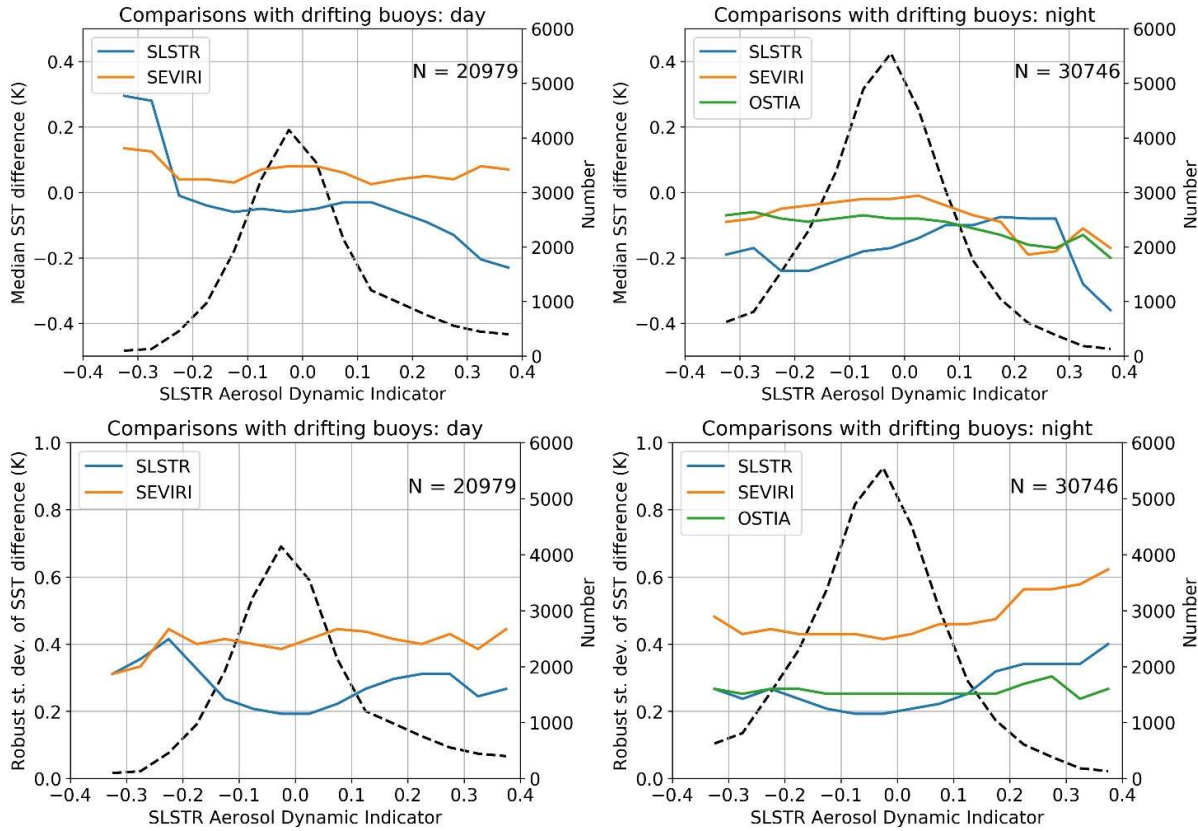


Figure 15: Similar to Figure 14, but now for SLSTR Aerosol Dynamic Indicator (ADI).

4.8. Triple collocation approach

An obvious advantage of having common collocations of SEVIRI, SLSTR and drifting buoys is that the triple collocation approach can be applied in order to deduce the random uncertainties of each individual instrument. Of course, there are some assumptions behind this approach, mainly that the observations of the instruments are independent from each other and that the collocation uncertainty is negligible. As OSTIA assimilates both drifting buoys and the satellite instruments, mainly SEVIRI, while also the drifting buoys are used for the bias correction (Section 3.3) there is no independence from them, so OSTIA will not be considered in this section. However, the three instruments are independent from each other, despite the fact that both satellites measure the radiation at similar channels, both the SST retrievals and the instrument characteristics differ between them. Regarding the collocation uncertainty, it is much smaller than in similar studies using MW imagers or IR sounders (e.g. Tsamalis and Saunders, 2018; O'Carroll et al., 2012) for the triple collocation, as most of the collocations are within 5 km and 1 hour (Figure 3 and Figure 4). Of course, this does not mean necessarily that the collocation uncertainty is negligible, but it is here as small as possible, although it is neglected in the triple collocation approach. Very often, the triple

collocation approach is applied during night conditions in order to avoid the complications arising from the thermocline during the day and the same is done here.

Figure 16 presents the monthly time series of the random uncertainties using night time collocations. The random uncertainties of the drifting buoys is pretty much stable at 0.20 K. The same is true also for SLSTR with the random uncertainty being about 0.27 K. SEVIRI random uncertainty seems to improve slightly with time, but its mean value is around 0.46 K. Despite the significant variation of the number of collocations with time, being usually between 600 and 1200, this does not impact the random uncertainty of any instrument (at least in an obvious way).

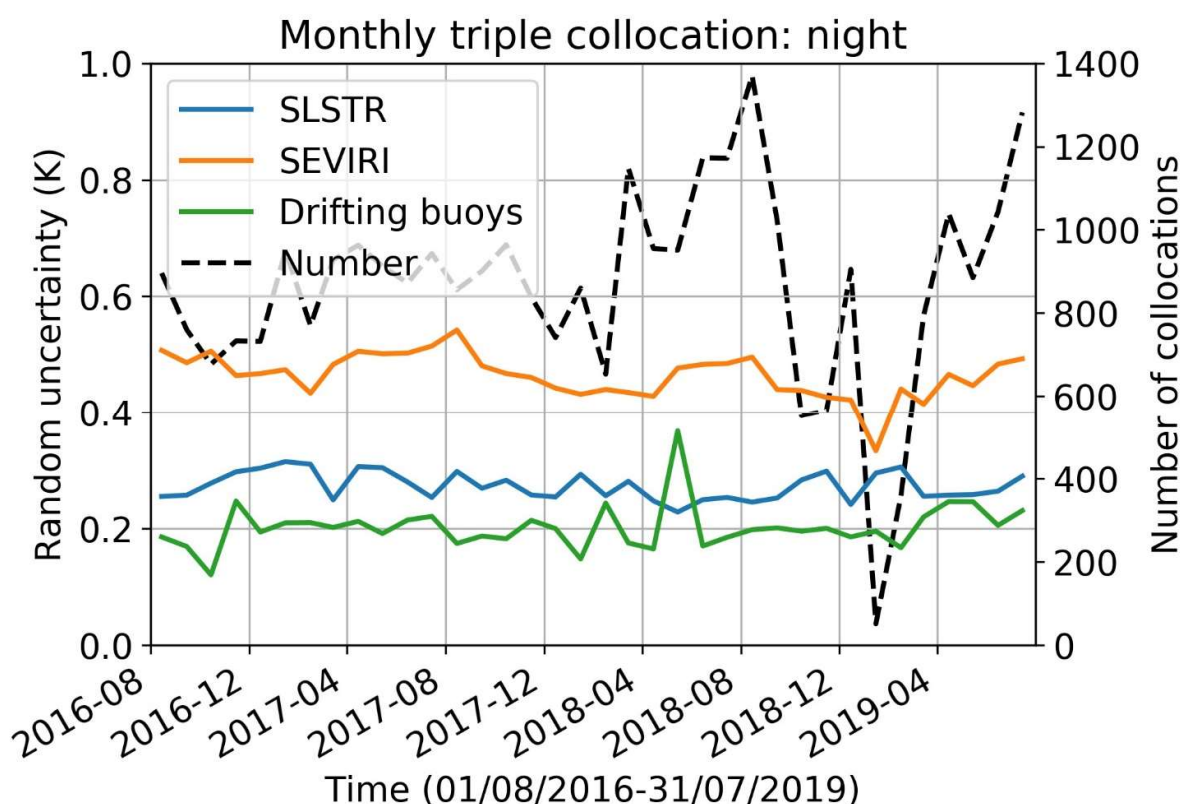


Figure 16: Monthly random uncertainty as estimated from the triple collocation approach. The results are for all the common collocations with SEVIRI QL 3 to 5 and SLSTR QL 5 during night. SLSTR is shown in blue, SEVIRI in orange and drifting buoys in green. The monthly number of collocations is shown with the dashed black line with the values displayed on the right ordinate.

5. Conclusions

In this study, the SST from SEVIRI, SLSTR, drifting buoys and OSTIA (only during night) have been intercompared over common collocations covering a 3-year period (1/8/2016-31/7/2019). This way, the number of collocations is reduced significantly, as both satellite instruments have to provide cloud-free SSTs during the overpass time of SLSTR (being the satellite instrument on a polar orbit). For this reason,

mainly results from robust statistics have been provided, which are more resilient to outliers and to noisy data. The reference SST dataset has been chosen to be the drifting buoys (as in situ instruments), which measure depth SST. In the interpretation of the results, one needs to take into account that SLSTR measures skin SST, thus a difference with drifting buoys observations is expected, but both SEVIRI and OSTIA (during night) should agree rather closely with the drifting buoys on average. In the analysis of the results, daytime and night time conditions have been examined separately. The software for the creation of the common collocation database, containing daily netCDF files, is written in Python. Both the software and the database have been delivered to OSI SAF.

In the reprocessed SLSTR database missing coverage within some parts of the SEVIRI disk has been found. Initially, SEVIRI quality levels 2 to 5 and SLSTR quality levels 3 to 5 have been investigated. The stratification between the quality levels is meaningful and fit-for-purpose for both satellite instruments, as indicated by the generally decreasing (robust) standard deviation with increasing quality level. Also, at the lowest quality level for both SEVIRI and SLSTR the mean/median difference to drifting buoys is always negative and much colder than the rest of the differences, an indication of cloud contamination.

The main part of the analysis performed with SEVIRI quality levels 3 to 5 and SLSTR quality level 5, following the advice from OSI SAF and EUMETSAT, respectively. It turns out that the common collocations are as close as possible in space and time in general within ± 3 km and ± 60 min, while in the case of SLSTR with drifting buoys the distance is even tighter being in general within ± 1 km. This translates to minimal collocation uncertainties. The histograms of the SST difference against drifting buoys have a Gaussian shape with median \pm robust standard deviation for SLSTR -0.06 ± 0.24 K and for SEVIRI 0.06 ± 0.42 K during daytime. During night, the respective results are for SLSTR -0.17 ± 0.22 K, for SEVIRI -0.04 ± 0.44 K and for OSTIA -0.09 ± 0.25 K. Both OSTIA and SLSTR have very similar performance against drifting buoys, but one has to bear in mind that the collocation uncertainty is larger in the case of OSTIA due to the larger spatial resolution than SLSTR, and that OSTIA system assimilates drifting buoys measurements. SEVIRI results indicate a poorer performance against drifting buoys in comparison to SLSTR or OSTIA, but these are in agreement with the OSI SAF validation results for it (Saux-Picart and Marsouin, 2018). As expected, the robust standard deviation of SLSTR dual view retrievals (D2 and D3) are better (0.21 K) than for nadir view only with N3 being 0.24 K and N2 being 0.28 K, although their standard deviations are almost the same (0.34-0.35 K).

The time series confirm that the median differences and robust standard deviations are about stable over the 3-year period, although quite noisy, especially for the median difference. The geographical distribution is very homogeneous for SLSTR during both day and night for both the median difference and the robust standard deviation. SEVIRI median difference is rather homogeneous during day, but during night the difference with drifting buoys is colder over the North Atlantic. On the other hand, the SEVIRI robust standard deviation is lower at the sub-satellite point increasing towards the edge of the swath. OSTIA statistics are also homogeneous geographically, with the exception of North Tropical Atlantic and the Arabian Sea. There the median difference is colder indicative of dust contamination, probably inherited by VIIRS, which is used as reference in OSTIA. The geographical dependence of SEVIRI robust standard deviation can be easily depicted also when it is plotted against the SEVIRI zenith angle with higher values (~ 0.6 K) at the edge of SEVIRI swath and lower (0.3-0.4 K) towards 20° - 30° . It is worthwhile to note that SLSTR during day has a colder median difference (down to -0.2 K) when the SEVIRI zenith angle is about 20° , indicative of a cloud/dust contamination around the 20° latitudes. The SLSTR robust standard deviation is stable with zenith angle at about 0.2 K, with the exception to N2 retrieval which increases reaching ~ 0.35 K at the edge of SLSTR swath.

During night there is dependence of OSTIA median difference with wind speed, a surprising fact given it should provide a SST similar to that observed by drifting buoys. In addition, during day SLSTR median difference does not show dependence on wind speed, when the expected result would be the inverse.

OSTIA robust standard deviation has also a dependence on wind speed. Under daytime conditions, both the median difference and the robust standard deviation indicate a variability with total column water vapour for both SEVIRI and SLSTR. The same is true for SEVIRI during night, but SLSTR has only very weak dependence, something that indicates the added value in the SST retrieval by the inclusion of the 3.7 μm channel. The median difference of OSTIA also has a clear dependence on total column water vapour, up to 40 kg/m^2 following this of SEVIRI, again probably inherited from VIIRS.

Using SEVIRI Saharan Dust Index (SDI) as a measure of dust load in the atmosphere, there is clear dependence of SEVIRI statistics with positive SDI values. On the other hand, SLSTR statistics have no or minor dependence on SDI. Nevertheless, OSTIA's median difference shows linear dependence on SDI, a fact that indicates impact of dust aerosols on OSTIA SST accuracy and confirms the results of the geographical distribution. However, when using SLSTR Aerosol Dynamic Indicator (ADI) as a measure of dust load, the results are quite the opposite with SLSTR showing in general more dependence on ADI rather than SEVIRI, certainly during day. This inconsistency between SDI and ADI needs further examination, but it is outside the scope of this study. Again, OSTIA's median difference varies roughly linearly with ADI.

Finally, the triple collocation approach has been applied in order to estimate the random uncertainties of SST from drifting buoys, SLSTR and SEVIRI. The random uncertainties were rather stable over time with value of 0.20 K for drifting buoys, 0.27 K for SLSTR and 0.46 K for SEVIRI. For SEVIRI it is the first time that the random uncertainty is estimated using the triple collocation approach. The value of 0.20 K for drifting buoys is in agreement with previous estimations (e.g. O'Carroll et al., 2012; Tsamalis and Saunders, 2018). However, the value for SLSTR is larger than this found for the precursor instrument AATSR, which is similar to drifting buoys or even better (e.g. O'Carroll et al., 2008; Tsamalis and Saunders, 2018). Taking into account that the collocation uncertainties are minimal, this indicates that there is possibly room for improvement in the case of SLSTR, despite the significant improvement in comparison to previous releases (Dybkjaer et al., 2018).

Obviously, further development and improvement of all SST datasets is desirable and it would be useful for the scientific community as a whole. The multi-collocation of SST datasets offers validation that is more rigorous by providing more robust conclusions than the widely used pairwise validation, although in the expense of reduced number of collocations. The approach taken here could be easily extend to assess the cloud mask of the satellite instruments. Another possibility is the assessment of the in situ measurements quality, which even if here have been treated as the reference it is known that some of them are subject to non-trivial errors due to their deployment in the harsh environment of the sea.

6. Acknowledgements

Gary Corlett is acknowledged for helping with the description of the SLSTR collocation databases, the interpretation of their variables and for providing the last couple of months of the NRT SLSTR database, which permitted to this study to cover exactly 3 years. I would like to thank Météo-France staff for the warm reception and hospitality.

7. References

- Blackmore, T., O'Carroll, A., Fennig, K., and Saunders, R.: Correction of AVHRR Pathfinder SST data for volcanic aerosol effects using ATSR SSTs and TOMS aerosol optical depth, *Remote Sensing of Environment*, 116, 107–117, doi:10.1016/j.rse.2011.04.040, 2012.
- Donlon, C. J., Minnett, P. J., Gentemann, C., Nightingale, T. J., Barton, I. J., Ward, B., and Murray, M. J.: Toward improved validation of satellite sea surface skin temperature measurements for climate research, *Journal of Climate*, 15, 353–369, doi:10.1175/1520-0442(2002)015<0353:TIVOSS>2.0.CO;2, 2002.
- Donlon, C., Rayner, N., Robinson, I., Poulter, D. J. S., Casey, K. S., Vazquez-Cuervo, J., Armstrong, E., Bingham, A., Arino, O., Gentemann, C., Le Borgne, D. M. P., Piolle, J., Barton, I., Beggs, H., Merchant, C. J., Heinz, S., Harris, A., Wick, G., Emery, B., Minnett, P., Evans, R., Llewellyn-Jones, D., Mutlow, C., Reynolds, R. W., and Kawamura, H.: The Global Ocean Data Assimilation Experiment High-resolution Sea Surface Temperature Pilot Project, *Bulletin of the American Meteorological Society*, 88, 1197–1213, doi:10.1175/BAMS-88-8-1197, 2007.
- Donlon, C. J., Martin M., Stark J., Roberts-Jones J., Fiedler E., and Wimmer W.: The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system, *Remote Sensing of Environment*, 116, 140–158, doi: 10.1016/j.rse.2010.10.017, 2012a.
- Donlon, C. J., Berruti, B., Buongiorno, A., Ferreira, M.-H., Femenias, P., Frerick, J., Goryl, P., Klein, U., Laur, H., Mavrocordatos, C., Nieke, J., Rebhan, H., Seitz, B., Stroede, J., and Sciarra, R.: The Global Monitoring for Environment and Security (GMES) Sentinel-3 mission, *Remote Sensing of Environment*, 120, 37–57, doi: 10.1016/j.rse.2011.07.024, 2012b.
- Dybkaer, G., Marsouin, A., Eastwood, S., Piolle, J.-F., Hoyer, J. L., Roquet, H., Saux-Picart, S., O'Carroll, A., and Tomazic, I.: Sentinel 3 SLSTR SST Validation Report, A Match-up Data Base for S3A/SLSTR SST products validation, OSI_CAF_FA 13_02, 1-26, , 2018.
- Good, E. J., Kong, X., Embury, O., Merchant, C. J., and Remedios, J.: An infrared desert dust index for the Along-Track Scanning Radiometers, *Remote Sensing of Environment*, 116, 159–176, doi:10.1016/j.rse.2010.06.016, 2012.
- Le Borgne P., Roquet H. and Merchant C.J. Estimation of Sea Surface Temperature from the Spinning Enhanced Visible and Infrared Imager, improved using numerical weather prediction. *Remote Sens. Environ.*, 115, 55-65, 2011.
- Le Borgne, P., Marsouin, A., Orain, F., and Roquet, H.: Operational sea surface temperature bias adjustment using AATSR data, *Remote Sensing of Environment*, 116, 93–106, doi:10.1016/j.rse.2010.02.023, 2012.
- Le Borgne, P., Pere, S., and Roquet, H.: Night time detection of Saharan dust using infrared window channels: Application to NPP/VIRS, *Remote Sensing of Environment*, 137, 264–273, doi:10.1016/j.rse.2013.06.001, 2013.
- Marsouin, A., Le Borgne, P., Legendre, G., Pere, S. and Roquet, H. Six years of OSI-SAF METOP-A AVHRR sea surface temperature. *Remote Sens. Environ.*, 159, 288-306, 2015.
- Merchant, C. J. Sea surface temperature (SLSTR) Algorithm Theoretical Basis document. SLSTRATBD-L2SST, http://www.eumetsat.int/website/wcm/idc/idcplg?IdcService=GET_FILE&dDocName=PDF_S3_L2_ATBD_SLSTR_SST&RevisionSelectionMethod=LatestReleased&Rendition=Web, 2012.
- Merchant, C. J., Harris, A. R., Murray, M. J., and Zavody A. M.: Toward the elimination of bias in satellite retrievals of sea surface temperature. 1. Theory, modeling and interalgorithm comparison, *Journal of Geophysical Research*, 104, C10, 23565-23578, 1999.
- Merchant, C. J., Harris, A. R., Maturi, E., and MacCallum, S.: Probabilistic physically based cloud screening of satellite infrared imagery for operational sea surface temperature retrieval, *Quarterly Journal of the Royal Meteorological Society*, 131, 2765–2755, doi:10.1256/qj.05.15, 2005.
- Merchant, C. J., Horrocks, L. A., Eyre, J. R., and O'Carroll, A. G.: Retrievals of sea surface temperature from infrared imagery: origin and form of systematic errors, *Quarterly Journal of the Royal Meteorological Society*, 132, 1205–1223, doi: 10.1256/qj.05.143, 2006a.

- Merchant, C. J., Embury, O., Le Borgne, P., and Bellec, B. Saharan dust in nighttime thermal imagery: Detection and reduction of related biases in retrieved sea surface temperature. *Remote Sens. Environ.*, 104, 15–30, 2006b.
- O’Carroll, A. G., Eyre, J. R., and Saunders, R. W. Three-way error analysis between AATSR, AMSR-E, and in situ sea surface temperature observations. *J. Atmos. Ocean. Technol.*, 25, 1197–1207, 2008.
- O’Carroll, A. G., August, T., Le Borgne, P., and Marsouin, A.: The accuracy of SST retrievals from Metop-A IASI and AVHRR using the EUMETSAT OSI-SAF matchup dataset, *Remote Sensing of Environment*, 126, 184–194, doi:10.1016/j.rse.2012.08.006, 2012.
- Peyridieu, S., Chedin, A., Capelle, V., Tsamalis, C., Pierangelo, C., Armante, R., Crevoisier, C., Crepeau, L., Simeon, M., Ducos, F., and Scott, N. A.: Characterisation of dust aerosols in the infrared from IASI and comparison with PARASOL, MODIS, MISR, CALIOP, and AERONET observations, *Atmospheric Chemistry and Physics*, 13, 6065–6082, doi:10.5194/acp-13-6065-2013, URL <http://www.atmos-chem-phys.net/13/6065/2013/>, 2013.
- Reynolds, R. W., Gentemann, C. L., and Corlett, G. K.: Evaluation of AATSR and TMI satellite SST data, *Journal of Climate*, 23, 152–165, doi:10.1175/2009JCLI3252.1, 2010.
- Roberts-Jones, J., Fiedler, E. K., and Martin, M. J.: Daily, global, high-resolution SST and sea ice reanalysis for 1985–2007 using the OSTIA system., *Journal of Climate*, 25, 6215–6232, doi:10.1175/JCLI-D-11-00648.1, 2012.
- Saux-Picart, S.: Algorithm Theoretical Basis Document for MSG/SEVIRI Sea Surface Temperature data record, doi: 10.15770/EUM_SAF_OSI_0004, URL , 2018.
- Saux-Picart, S., and Marsouin, A.: Geostationary Satellite Sea Surface Temperature Scientific Validation Report, URL , 2018.
- Schmetz, J., Pili P., Tjemkes S., Just D., Kerkmann J., Rota S., and Ratier A.: An introduction to Meteosat Second Generation (MSG), *Bulletin of American Meteorological Society*, 83, 977–992, , 2002.
- Tsamalis C. and Saunders R.: Quality Assessment of Sea Surface Temperature from ATSRs of the Climate Change Initiative (Phase 1), *Remote Sens.*, 10, 497, 2018.
- Tsamalis C., Chedin A., Pelon J. and Capelle V.: The seasonal vertical distribution of the Saharan Air Layer and its modulation by the wind. *Atmos. Chem. Phys.*, 13, 11235–11257, 2013.
- Zwieback, S., Scipal, K., Dorigo, W., and Wagner, W.: Structural and statistical properties of the collocation technique for error characterization, *Non-linear Processes in Geophysics*, 19, 69–80, doi:10.5194/npg-19-69-2012, URL <http://www.nonlin-processes-geophys.net/19/69/2012/>, 2012.