applied sciences

MDPI

Article

Towards Improved Quality Control of In Situ Sea Surface Temperatures from Drifting and Moored Buoys in the NOAA *i*Quam System

Boris Petrenko [1,2,*], Alexander Ignatov [1,*], Victor Pryamitsyn [1,2] and Olafur Jonasson [1,2]

[1] STAR, NOAA Center For Weather and Climate Prediction (NCWCP), College Park, MD 20740, USA; victor.pryamitsyn@noaa.gov (V.P.); olafur.jonasson@noaa.gov (O.J.)
[2] Global Science and Technology, Inc., Greenbelt, MD 20770, USA
* Correspondence: boris.petrenko@noaa.gov (B.P.); alex.ignatov@noaa.gov (A.I.)

Abstract: The NOAA in situ Sea Surface Temperature (SST) Quality Monitor (*i*Quam) online system collects in situ SSTs from various sources, performs quality control (QC), and provides QC'ed data to users. Like many other in situ QCs, the *i*Quam QC employs comparisons with Level 4 SST analysis. However, the current daily L4 analyses do not capture the diurnal cycle, nor do they resolve the fine structure of SST in dynamic areas. As a result, high-quality in situ SSTs significantly deviating from the L4 SST may be rejected. This paper discusses the new Diurnal Reference Check (DRC), which addresses overscreening for buoys whose sampling frequency is sufficient for resolving the diurnal cycle. The DRC separates records from individual buoys into 24-h segments and characterizes each segment with the median nighttime (MNT) SST and the amplitude of the diurnal signal (ADS). The segment is rejected if the ADS is unrealistically large or if the difference between the MNT and L4 SST exceeds a geographically dependent threshold. The outliers are further screened out by comparison of individual in situ SSTs with the MNT. All thresholds are determined from the analysis of matchups with reprocessed NOAA SSTs from multiple low-orbiting satellites. The satellite matchups are also used to validate the QC results. The DRC minimizes the overscreening, increases the number of high-quality in situ data by ~5%, and reduces the QC reliance on the L4 analysis. In addition, a new retrospective satellite-based quality check is introduced to identify matchups, which are most useful for training SST algorithms and validation of reprocessed satellite data.

Keywords: in situ SST; *i*Quam; quality control; reference SST; satellite SST; L4 analysis; matchups

Citation: Petrenko, B.; Ignatov, A.; Pryamitsyn, V.; Jonasson, O. Towards Improved Quality Control of In Situ Sea Surface Temperatures from Drifting and Moored Buoys in the NOAA *i*Quam System. *Appl. Sci.* **2023**, *13*, 10205. https://doi.org/10.3390/app131810205

Academic Editor: Gabriele Pieri

Received: 25 July 2023
Revised: 31 August 2023
Accepted: 4 September 2023
Published: 11 September 2023

1. Introduction

In situ sea surface temperatures (SST; see Abbreviations) measured by drifting and tropical moored (DTM) buoys are widely used in satellite remote sensing for training SST retrieval algorithms, validation of satellite retrievals, and their de-biasing before assimilation into higher-level models and analyses. In order to facilitate the use of in situ SSTs, T_{IS}, NOAA has established the in situ SST Quality Monitor (*i*Quam) system [1–3]. The *i*Quam collects in situ data from various external sources, performs quality control (QC), monitors QC'ed data online, and provides them for NOAA applications and all interested external users. Since its establishment in 2009, the *i*Quam has been in demand as a reliable source of quality in situ data. The *i*Quam QC identifies degraded SST measurements (e.g., obtained from malfunctioning sensors, corrupted during transmission to satellite and back to the ground, erroneously time stamped or geo-positioned, etc.). A common way to identify low-quality T_{IS}'s is to compare them with SSTs obtained from alternative sources, such as L4 analyses 'or climatologies' [2,4–6] or satellite retrievals [3,5,7,8]. In this study, we explore enhancements to the current *i*Quam QC aimed at improved utilization of the available reference SST information.

All current L4 analyses are daily gap-free products. They provide continuous coverage of the world ocean with a 'reference SST', T_{L4}, which allows examining deviations $\Delta T_{L4} = T_{IS} - T_{L4}$ at any location. The Reference Check (RC) is a critical part of the current *i*Quam QC, which identifies most low-quality T_{IS}'s [2,9]. The current RC in the latest *i*Quam v2.10 employs a Bayesian approach [10] to transform the ΔT_{L4} into posterior probability of 'gross error' (PGE), under certain assumptions on prior distributions of ΔT_{L4} and noise in T_{IS}. The quality level (QL) is downgraded for a specific T_{IS} if its PGE exceeds the predefined threshold [2]. It was recognized, however, that the efficiency of the current RC is limited due to intrinsic features of the available L4 SST analyses [5,9], which often represent the so-called 'foundation' SST (i.e., corresponding to depths of ~10 m, at which the diurnal warming cycle is absent [11]). Such analyses do not capture the diurnal warming at the depths of drifting (~0.2 m) and moored (~1 m) buoys, which causes overscreening in the presence of significant diurnal signal (DS) in T_{IS}. Intensive diurnal SST variations were observed under conditions of prolonged insolation and suppressed mixing in the ocean upper layer (typically associated with clear skies and low winds) in both satellite [12–15] and in situ [15–17] SSTs in various oceanic regions, from the Tropics all the way to high latitudes. Furthermore, the gridded daily L4 analyses often fail to track fine-scale spatial and rapid temporal SST variations [9], also causing overscreening, mainly in the dynamic oceanic areas, such as the Gulf Stream, Kuroshio and Agulhas currents, upwellings, downwellings, etc. In this study, we explore the new Diurnal Reference Check (DRC), which largely mitigates the overscreening problems related to both diurnal warming in T_{IS} and inaccurate T_{L4} in the dynamic areas.

We also explore using the satellite SSTs, T_{SAT}, as an additional source of reference information. The advantage of satellite SSTs is that they are timelier and often more accurate than T_{L4}. However, the comparisons with T_{SAT} are only possible for T_{IS} measured at times and locations of satellite overpasses and under clear-sky conditions (for T_{SAT} obtained from IR radiometer measurements). Ideally, the QC should take advantage of comparisons with both T_{L4} and T_{SAT}. In [3], the *i*Quam data were compared with satellite SSTs from NOAA-17 AVHRR and ENVISAT AATSR. However, the current *i*Quam QC does not use satellite SSTs routinely. During recent years, the full missions of multiple low Earth orbiting (LEO) satellite IR radiometers have been reprocessed with the NOAA Advanced Clear Sky Processor for Ocean (ACSPO) enterprise SST system [18–20], and multi-year data sets of satellite SSTs (ACSPO RANs) and the corresponding matchup data sets (MDS) have been accumulated for a period from September 1981–present. In this study, we explore their potential use for adjusting the DRC thresholds, validation of the QC results, and retrospective evaluation of the *i*Quam data.

2. Data

2.1. In Situ Data

The *i*Quam collects in situ SSTs from multiple types of platforms, including conventional and scientific ships, drifting buoys, tropical and coastal moored buoys, and Argo floats available from various data sources, such as ICOADS, FNMOC, CMEMS, AOML, GDACs, and IMOS. Detailed descriptions of types of platforms and sources of in situ information can be found in [2,9]. The full set of in situ data from September 1981 –present, with *i*Quam QLs appended, is available at [1]. The *i*Quam QC employs five binary checks, including the Duplicate Removal (DR), the Plausibility/Geolocation check (GC), the platform Track Check (TC), the SST Spike Check (SC), and the platform ID Check (IC) [2,9]. Two other checks, the RC and the Buddy Check (BC), employ the Bayesian method [10]. The highest *i*Quam QL = 5 is set when all seven checks pass. If the T_{IS} fails at least one of these checks, then QL = 3 or 4 are assigned. Note that the *i*Quam does not exclude low-quality T_{IS}'s from records but rather assigns those lower QLs, so users have access to these data and may explore them. This study focuses on QC'ing T_{IS}'s from drifters and tropical moored (DTM) buoys, which sample T_{IS} with frequencies sufficient for resolving the diurnal signal. Figure 1 shows that the number of DTM measurements from September

1981–December 2021 had increased by more than two orders of magnitude and, today, represents the major source of in situ data for satellite Cal/Val.

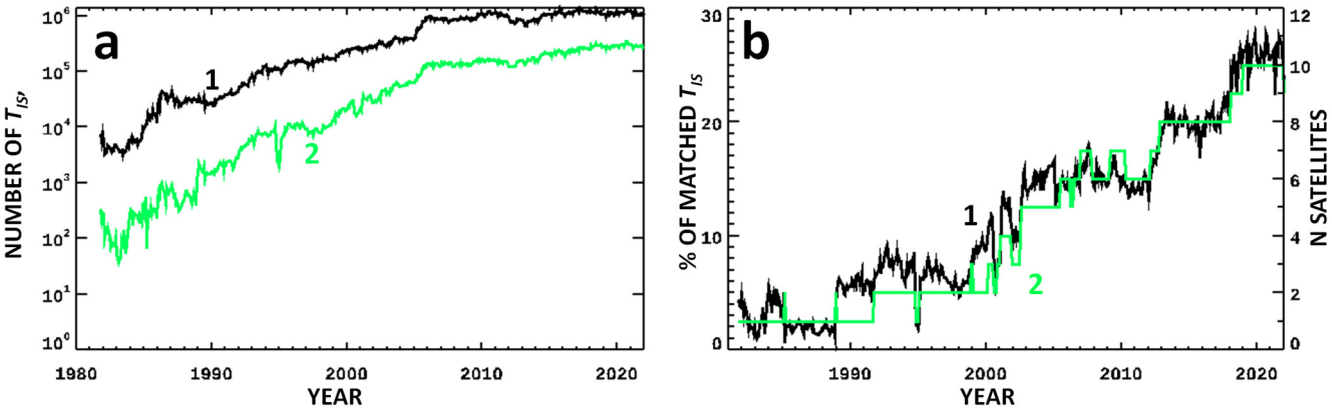

Figure 1. (**a**) Monthly number of (1) DTM T_{IS}'s and (2) DTM T_{IS}'s matched with satellite T_{SAT}'s. (**b**) (1) Monthly fraction of T_{IS}'s matched with T_{SAT}'s (%) and (2) Number of satellites whose MDS are available during each month.

2.2. Matchups with Satellite SSTs

The matchups of T_{IS} (with all *i*Quam QLs) with clear-sky T_{SAT} have been accumulated during ACSPO reprocessings (RANs) of multiple LEO satellites [18–20]. Table 1 shows the periods covered by each MDS and the types of the satellites' orbits. Altogether, these MDSs cover 42 yr+ period since September 1981. The orbits of the heritage NOAA satellites from NOAA-07 to NOAA-19 have been drifting during their missions, and their LEXT varied within wide ranges [21].

Table 1. Satellites and instruments represented in each MDS and local equator crossing times.

Satellite, Instrument	Covered Period (DD.MM.YYYY)	Orbit
NOAA-07 AVHRR/2 (GAC)	09.01.1981–02.02.1985	Variable, p.m.
NOAA-09 AVHRR/2 (GAC)	01.31.1985–11.07.1988	Variable, p.m.
NOAA-11 AVHRR/2 (GAC)	11.08.1988–09.13.1994	Variable, p.m.
NOAA-12 AVHRR/2 (GAC)	09.16.1991–12.14.1998	Variable, p.m.
NOAA-14 AVHRR/2 (GAC)	01.19.1995–10.19.2001	Variable, a.m.
NOAA-15 AVHRR/3 (GAC)	11.01.1998–present	Variable, a.m./p.m.
NOAA-16 AVHRR/3 (GAC)	10.26.2000–09.17.2007	Variable, p.m.
NOAA-17 AVHRR/3 (GAC)	07.10.2002–03.08.2010	Variable, a.m.
NOAA-18 AVHRR/3 (GAC)	06.06.2005–present	Variable, p.m./a.m.
NOAA-19 AVHRR/3 (GAC)	02.22.2009–present	Variable, p.m./a.m.
Terra MODIS	02.25.2000–present	Stable, 10:30 a.m.
Aqua MODIS	07.04.2002–present	Stable, 1:30 p.m.
Metop-A AVHRR/3 (FRAC)	12.01.2006–11.25.2021	Stable, 9:30 a.m.
Metop-B AVHRR/3 (FRAC)	10.19.2012–present	Stable, 9:30 a.m.
Metop-C AVHRR/3 (FRAC)	12.04.2018–present	Stable, 9:30 a.m.
S-NPP VIIRS	12.01.2012–present	Stable, 1:30 p.m.
NOAA-20 VIIRS	01.05.2018–present	Stable, 1:30 p.m.

The ACSPO data files and corresponding matchups report two satellite SSTs: 'Subskin', produced with a global regression and highly sensitive to the temperature of the upper ~10 μm 'skin' layer of the ocean, and 'Depth', produced with a piecewise regression and representing a closer proxy of the T_{IS} [22]. In this study, we use matchups with the 'Subskin' SST, due to its higher sensitivity to the diurnal signal. (Note, however, that one should not expect full consistency between the 'Subskin' T_{SAT} and 'bulk' T_{IS}.) As discussed in Section 4, the availability of multiple MDSs, especially in the two recent decades, greatly

facilitates the validation of in situ QC algorithms. Moreover, the data of multiple satellites with different LEXTs allow validation of the diurnal signal in the T_{IS}'s.

Note that the ACSPO MDSs report 'one-to-many' matchups, i.e., each T_{IS} is matched with all clear-sky T_{SAT}'s found in its neighborhood, individually. In order to reduce the number of matchups to a manageable size, we created a single matchup for each T_{IS} by averaging all clear-sky L2P 'Subskin' SSTs within ±0.5 h and ±10 km of each T_{IS}. Figure 1a shows the time series of monthly numbers of DTM T_{IS}'s (with all QLs) and monthly numbers of T_{IS}'s matched with T_{SAT}'s in at least one MDS. Figure 1b demonstrates a close consistency between the fraction of matched T_{IS}'s in the total monthly numbers of T_{IS} measurements (curve 1) and the number of satellites' MDSs available for each month (curve 2). The fraction of matched T_{IS}'s increased from ~3% in the 1990s to ~27% in the 2020s due to the increased number of available satellites (from 1 to ~10) and their corresponding MDSs.

2.3. L4 Analyses

The current *i*Quam QC employs two T_{L4}'s. From September 1981–present, the NOAA Optimal Interpolation SST (OISST) [23,24] is used. Since 1 Sep 1991, the second T_{L4} is obtained from the Canadian Meteorological Center (CMC) SST (the 0.2° version [25] before 1 January 2016 and the 0.1° version [26] after this date). The *i*Quam RC evaluates T_{IS}'s using a single OISST PGE before 1 September 1991 and has the option to use OISST or CMC PGEs, or their combination, after this date. In this study, we use the CMC SST as a reference since 1 September 1991. Our prior analyses suggest that it is the optimal choice for this period (note that the CMC is also employed as the first guess in ACSPO during this period [18–20]). Additional analyses were performed in this study to select the reference L4 SST before September 1991 from three global products currently available: OISST [23,24], Ocean SST and Sea Ice Reprocessed (OSTIA-RAN) [27,28] and the Climate Change Initiative (CCI) [29,30].

Figure 2 shows time series of monthly medians of ΔT_{L4}, $M(\Delta T_{L4})$, and corresponding robust standard deviations (RSD), $S(\Delta T_{L4}) = 1.4826 \times M(|\Delta T_{L4} - M(\Delta T_{L4})|)$, for the T_{L4}'s obtained from the three analyses.

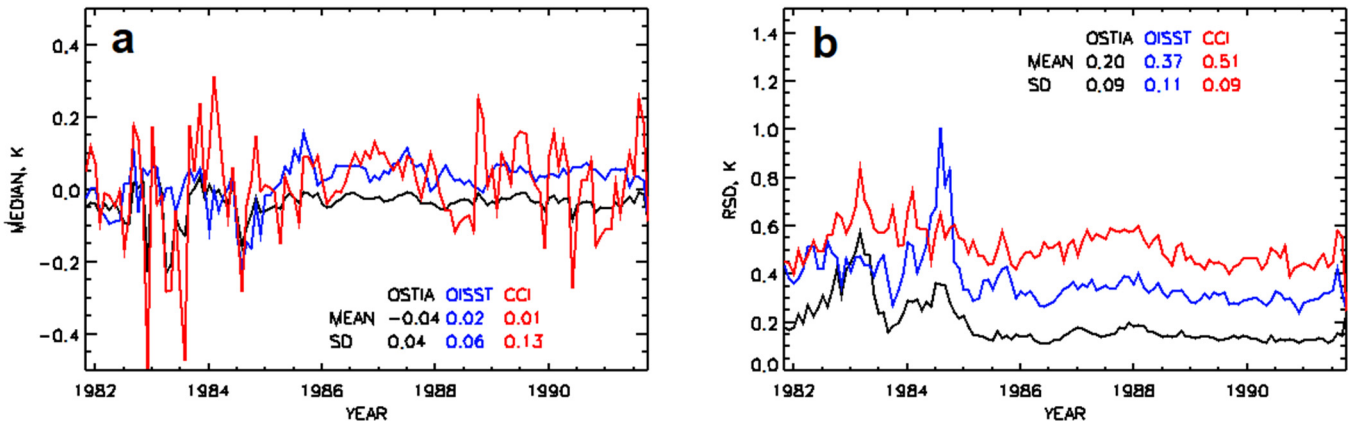

Figure 2. Time series of monthly (**a**) medians and (**b**) RSDs of nighttime DTM SST deltas from three L4 SSTs for a period from September 1981–1991: (black) OSTIA, (blue) OISST, and (red) CCI. The numbers in the plots show the temporal means and SDs calculated over the whole 10-year period.

The statistics are for the nighttime T_{IS}'s measured between 11 p.m. and 6 a.m. local solar time, without applying any QC. The deviations from CCI exhibit the smallest on average, but most variable in time medians, and the largest RSDs. The medians with respect to OSTIA are larger but more stable in time, with the smallest and least variable RSDs. Note that during this early period, comparisons of DTM T_{IS}'s with T_{SAT}'s are based on a small number of matchups, which is oftentimes insufficient for reliable training of the satellite SST algorithms. More efficient training requires adding less accurate but much

more numerous matchups with ship SSTs (at least, for the NOAA-07/09) [19]. Figure 3 re-plots time series from Figure 2 but using T_{IS}'s from both ships and DTMs. As in Figure 2, the deviations from OSTIA appear lowest in terms of both average median and average RSD, and most stable and consistent in time.

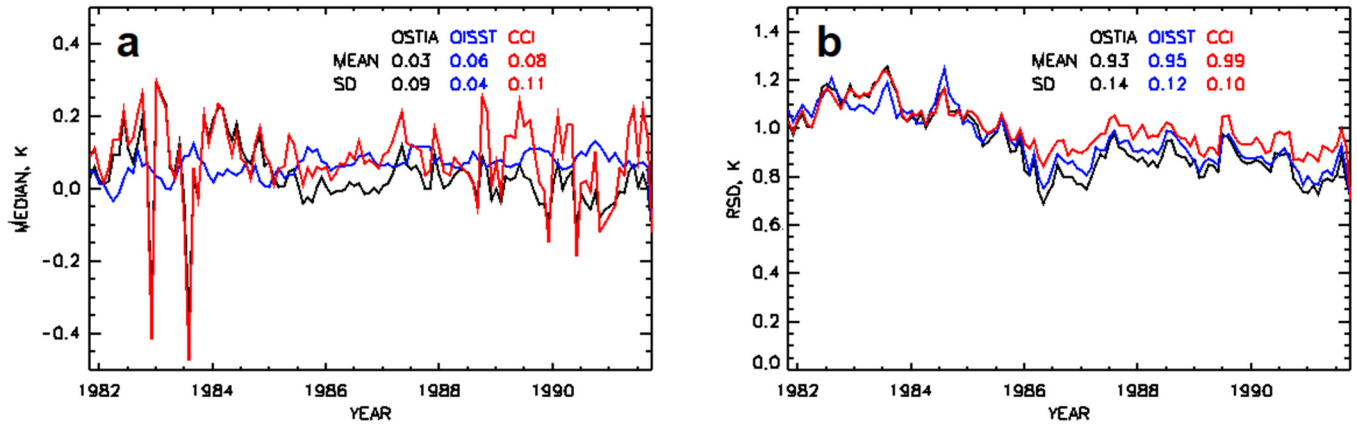

Figure 3. Same as in Figure 2 but for combined ship and DTM ΔT_{L4}'s.

Another factor in selecting L4 for use in *i*Quam QC is users' interest in in situ SSTs over lakes and in coastal areas. The availability of the OISST, OSTIA, and CCI analyses in these two domains is illustrated in Figure 4 using a representative example of May 1985.

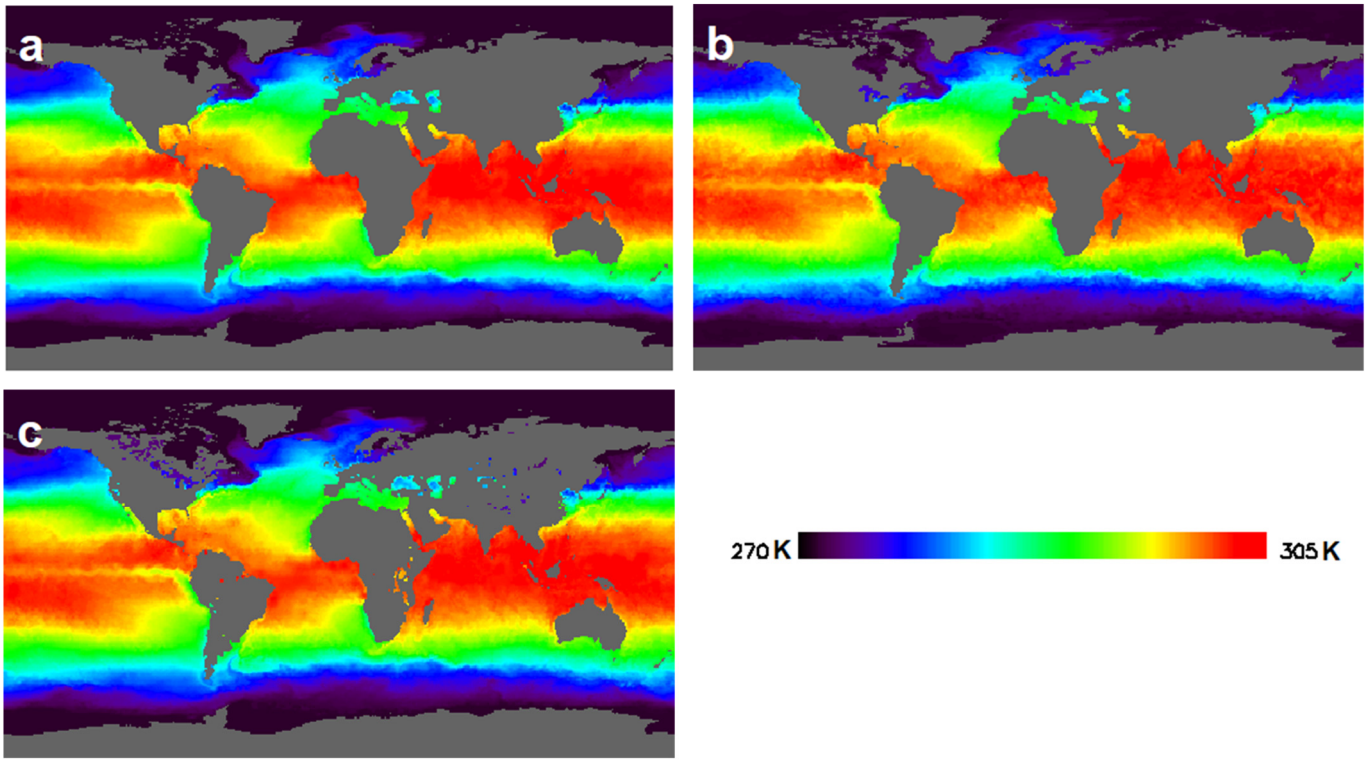

Figure 4. Monthly composite SST maps for May 1985 for (**a**) CCI, (**b**) OISST, and (**c**) OSTIA.

The CCI does not report temperatures over lakes. The OISST reports only Great Lakes and its coastal SSTs exhibit significantly fewer details compared with OSTIA (likely due to the 0.25° OISST vs. 0.05° OSTIA resolution).

Summarizing analyses in this section, OSTIA-RAN SST was selected as a reference for the period from September 1981–August 1991.

3. The Diurnal Reference Check (DRC)

3.1. Methodology

We formulate the DRC and adjust its thresholds using matchups with T_{SAT} from 2012–2018 data. Table 2 lists eight MDSs available for this period, along with the number of accumulated matchups in each MDS and LEXTs for each satellite. The orbits of the satellites listed in Table 2 are different. The Metop-A/B and Terra fly mid-morning 'am' orbits, whereas the S-NPP, NOAA-20, and Aqua are in the afternoon 'pm' orbits and observe the ocean close to the peak of the SST diurnal cycle, usually occurring around 3 p.m. LST [12,17]. These five satellites are maintained in stable orbits. The orbits of the three NOAA satellites (15/18/19) drift. In 2012–2018, the NOAA-19 was close to the peak of the DC, whereas the NOAA-15/18 overpassed at times when the SST was already decreasing.

Table 2. Total number of matchups of DTM in situ with satellite 'Subskin' SSTs and LEXT for eight satellites in 2012–2018. The LEXT bounds for the NOAA-15/18/19 were determined from the NOAA Sensor Stability for SST (3S) system [21].

Satellite Instrument	Number of Matchups	Local Equator Crossing Time
Metop-A AVHRR FRAC	2.47×10^6	9:30 a.m.
Metop-B AVHRR FRAC	2.47×10^6	9:30 a.m.
Terra MODIS	2.46×10^6	10:30 a.m.
Aqua MODIS	2.46×10^6	1:30 p.m.
S-NPP VIIRS	2.48×10^6	1:30 p.m.
NOAA-15 AVHRR GAC	2.45×10^6	4:30–6:50 p.m.
NOAA-18 AVHRR GAC	2.45×10^6	2:30–8:20 p.m.
NOAA-19 AVHRR GAC	2.45×10^6	1:30–4:30 p.m.

The goal of this study is to revisit the *i*Quam QC and improve its performance in the presence of significant DS in T_{IS}'s. This requires quantitative characterization of the DS in the T_{IS} records from individual buoys. We introduce the corresponding metrics as follows. The T_{IS} record from each buoy is subdivided into 24-h segments in terms of LST. Each segment is further subdivided into nighttime subsegment, which includes T_{IS}'s sampled between 11 p.m. and 7 a.m. LST, and daytime subsegment, with T_{IS}'s sampled between 8 a.m. and 8 p.m. LST. The segments in which either nighttime or daytime subsegment includes less than three T_{IS} samples are excluded from the analysis. In order to minimize the potential distortion of the metrics by outliers, each subset is filtered by excluding T_{IS}'s satisfying the condition $|T_{IS} - M(T_{IS})| > 4 \times S(T_{IS})$, where $M(T_{IS})$ is median, and $S(T_{IS}) = 1.4826 \times M(|T_{IS} - M(T_{IS})|)$ is the corresponding RSD, both calculated over a given subset. After that, the following metrics are calculated:

- Maximum SST in the daytime subsegment, T_{MAX}.
- Minimum SST in the nighttime subsegment, T_{MIN}.
- Median SST in the nighttime subsegment, T_{NIGHT}.
- Amplitude of the DS (ADS), $D = T_{MAX} - T_{MIN}$ if $T_{MAX} > T_{MIN}$; otherwise, $D = 0$.

The above metrics characterize each segment as a whole rather than individual T_{IS} counts. This allows for implementing separate checks for filtering the entire low-quality 24-h segments and individual T_{IS} outliers within the segments.

3.2. Setting Maximum Amplitude of the Diurnal Signal

Figure 5a shows the time series of the total monthly number of 24-h segments in the T_{IS} records, and the total number of those segments in which the DS was identified as described above. The typical monthly fraction of segments with identifiable diurnal signal is (90 ± 3)%. Figure 5b shows the fraction of segments with ADS exceeding a given threshold, D. This fraction quickly decreases with increasing D. In rare extreme cases, the ADS may reach 40 K. In addition to the sensor's malfunctioning, another possible source of unrealistically large ADS is direct solar heating of buoys' hulls [31,32]. We establish the

threshold for 'realistic' ADS, beyond which the DS is likely caused by extraneous factors other than variations in the water temperature, from the analysis of statistics of T_{SAT} vs. T_{IS} within the MDSs listed in Table 2.

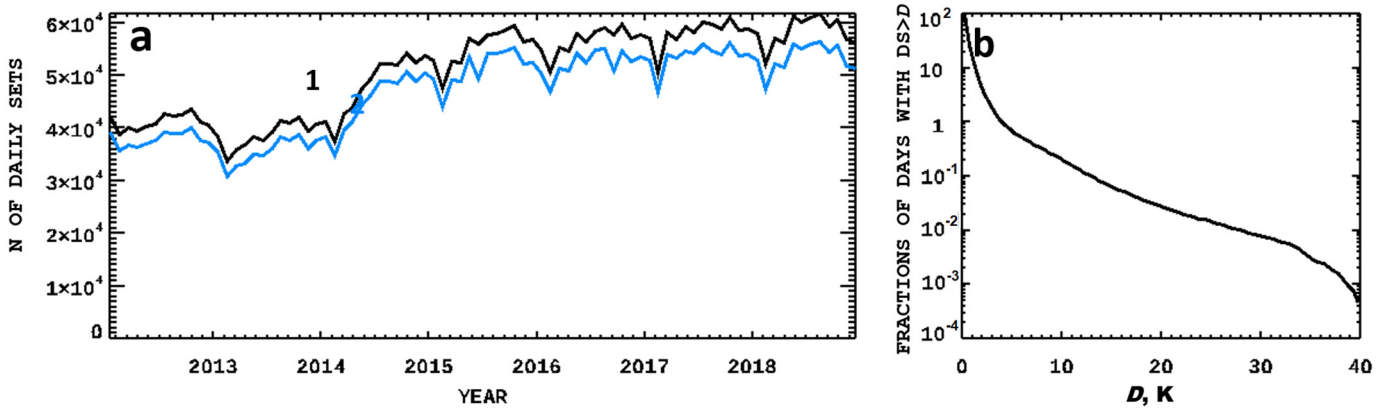

Figure 5. (**a**) Time series of the total monthly number of (1) daily segments in DTM and (2) daily segments with identifiable DS. (**b**) Fraction of daily segments with identifiable DS, in which the ADS exceeds D, as a function of D.

Figure 6a shows the median $M(T_{SAT} - T_{IS})$ computed over segments with ADS corresponding to the $D \pm 0.5$ K bins as a function of D. For all satellites, the $M(T_{SAT} - T_{IS})$ is close to 0 K in the first D bin centered at 0.5 K, because the regression coefficients of the ACSPO SST equations were trained against T_{IS} with *i*Quam QL = 5, which rejects those T_{IS} counts that are significantly affected by the diurnal warming. Within a range of 0.5 K < D < 5.5 K, the medians vary from −0.05 K to 0.15 K, and then at D > 5.5 K, they drop to essentially negative values due to increased fractions of segments with the ADSs caused by the extraneous factors.

Figure 6b shows the RSDs, $S(T_{SAT} - T_{IS})$, as a function of D. The RSDs are consistent across all satellites for Ds < ~3.5 K, gradually increasing to ~0.6 K. At D > 3.5 K, the RSDs grow at a faster rate. The largest RSDs correspond to the 'pm' satellites (S-NPP, Aqua, and NOAA-19), which observe the ocean at times close to the peak of the SST diurnal cycle, whereas the 'am' satellites (Metop-A/B) exhibit the lowest RSDs.

Figure 6c shows the correlation coefficients C_{NIGHT} of $T_{SAT} - T_{NIGHT}$ versus $T_{IS} - T_{NIGHT}$, calculated within the same D bins. The C_{NIGHT} for Aqua, NPP, and NOAA-15/18/19 satellites show pronounced maxima at D = 3.5 K, followed by a sharp decline at D = 5.5 K. For the 'am' satellites Metop-A/B and Terra, the C_{NIGHT}s are relatively small and do not show distinct maxima. This is because significant 'realistic' ADSs, with D < 5.5 K, most often occur near the peak of the diurnal cycle, whereas in the morning or evening, such ADSs are mostly caused by factors other than variations in the water temperature.

Considering that at D > ~5.5 K, the medians $M(T_{SAT} - T_{IS})$ in Figure 6a start decreasing, and the correlations C_{NIGHT} in Figure 6c converge for all satellites, we adopt 5 K as the upper threshold for realistic ADSs for the water temperature at the buoys' depth. Figure 5b shows that this threshold rejects approximately 0.7% of all segments with identifiable DS.

Figure 6d shows the correlations C_{L4} of $T_{SAT} - T_{L4}$ and $T_{IS} - T_{L4}$. For S-NPP, Aqua, and NOAA-15/18/19 at D > 2.5 K, they are higher than the corresponding C_{NIGHT}s in Figure 6c and remain relatively high at D > 5.5 K because the deviations of $T_{SAT} - T_{L4}$ and $T_{IS} - T_{L4}$ are both affected by the same variations in T_{L4}. This raises a question of the quality of the reference SST in the presence of significant DS in the buoys' records. Remember that the 'foundation' L4 analyses are not intended to reproduce the diurnal warming cycle. In particular, the CMC does not assimilate SSTs sampled during the day under wind speeds < 6 m/s between 25° S and 25° N or elsewhere within 45 days of the summer solstice [25,26]. It is instructive, therefore, to explore the quality of the $T_{IS} - T_{L4}$ statistics in the presence of diurnal SST variations.

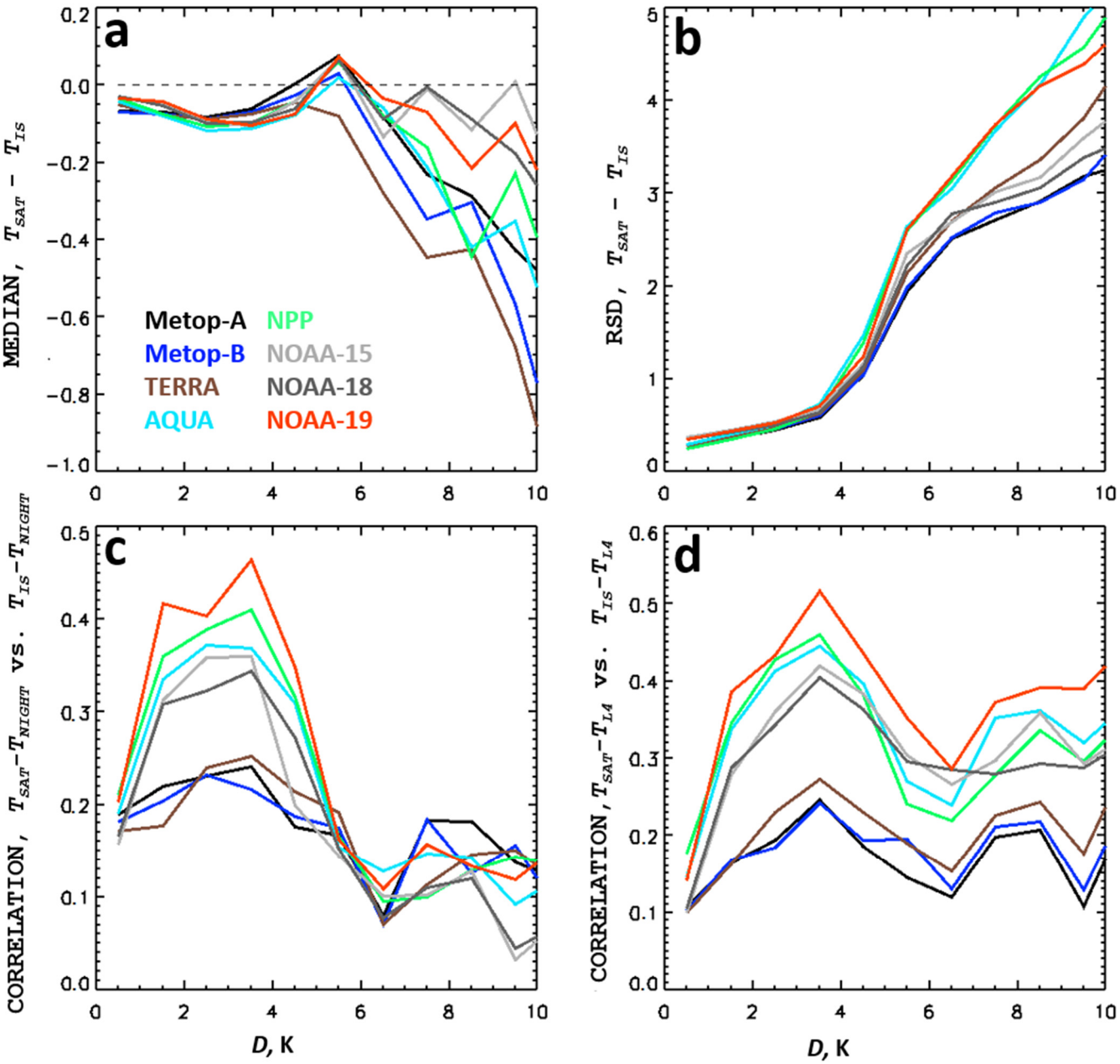

Figure 6. (**a**) Medians and (**b**) RSDs of $T_{SAT} - T_{IS}$; correlations of (**c**) $T_{SAT} - T_{NIGHT}$ versus $T_{IS} - T_{NIGHT}$ and (**d**) $T_{SAT} - T_{L4}$ versus $T_{IS} - T_{L4}$, as a function of D, for eight satellites listed in Table 2.

3.3. Quality of the L4 SST in the Presence of Diurnal SST Variations

Figure 7 shows the statistics of $T_{NIGHT} - T_{L4}$, $T_{MAX} - T_{L4}$, and $T_{MIN} - T_{L4}$, averaged over the daily segments with ADS falling within $D \pm 0.25$ K bins in 2012–2018, as a function of D. Figure 7b,c show the corresponding statistics for $T_{MAX} - T_{NIGHT}$ and $T_{MIN} - T_{NIGHT}$.

In Figure 7a, the median, $M(T_{NIGHT} - T_{L4})$, is as small as 0.01 K, and the RSD, $S(T_{NIGHT} - T_{L4}) = 0.2$ K within the range 0 K $\leq D \leq$ 0.5 K. This indicates a close consistency between the T_{NIGHT} and T_{L4}, when the diurnal warming is limited. However, when D grows, the $M(T_{NIGHT} - T_{L4})$ decreases and the $S(T_{NIGHT} - T_{L4})$ increases. One concludes that in the presence of the DS, the CMC SST becomes less consistent with the nighttime SSTs and less stable, likely due to residual diurnal variations in the assimilated daytime SSTs.

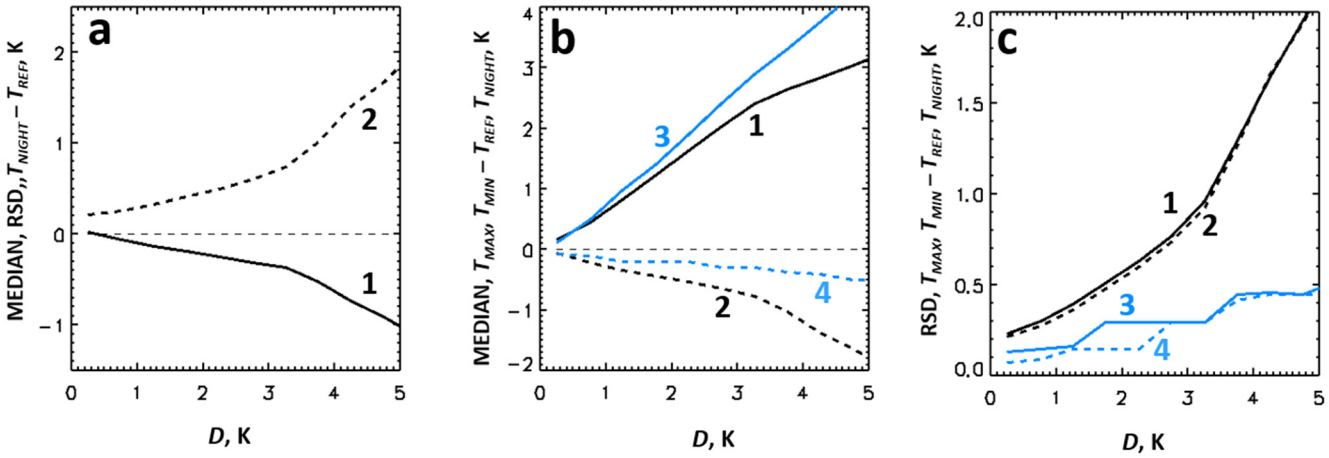

Figure 7. (**a**) (1) Median and (2) RSDs of $T_{NIGHT} - T_{L4}$; (**b**) Medians and (**c**) RSDs of (1) $T_{MAX} - T_{L4}$, (2) $T_{MIN} - T_{L4}$, (3) $T_{MAX} - T_{NIGHT}$ and (4) $T_{MIN} - T_{NIGHT}$, calculated within DS intervals $D \pm 0.5$ K, as functions of D.

In Figure 7b, the two medians, $M(T_{MIN} - T_{NIGHT})$ and, more so, $M(T_{MIN} - T_{L4})$, are both negative and decrease when D increases. This suggests a direct link between larger ADS and more intensive nighttime cooling of the ocean at the buoys' depth. The deviations of T_{MAX} and T_{MIN} from T_{L4} are colder than the corresponding deviations from T_{NIGHT}, consistent with curve 1 in Figure 7a. In Figure 7c, all four RSDs, $S(T_{MAX} - T_{L4})$, $S(T_{MIN} - T_{L4})$, $S(T_{MAX} - T_{NIGHT})$, and $S(T_{MIN} - T_{NIGHT})$, increase with D. However, the $S(T_{MAX} - T_{NIGHT})$ and $S(T_{MIN} - T_{NIGHT})$ are much smaller than the $S(T_{MAX} - T_{L4})$ and $S(T_{MIN} - T_{L4})$. Considering that variations of $T_{IS} - T_{NIGHT}$ are less dependent on the ADS compared to $T_{IS} - T_{L4}$, we conclude that the comparison of T_{IS} with T_{NIGHT} provides a more stable screening of individual T_{IS} outliers than the comparison of T_{IS} with T_{L4}.

3.4. Checking Daily Segments for Systematic Errors

To check daily segments for systematic errors, we compare T_{NIGHT} with T_{L4}. This is a more adequate use of the 'foundation' T_{L4} than comparing each individual T_{IS} with T_{L4} (as it is done e.g. in the current *i*Quam QC). The threshold for this comparison accounts for the geographical variability of the T_{L4} uncertainty. Specifically, this threshold is set to be proportional to the spatial standard deviation (SD) of the analysis SST within the $5^\circ \times 5^\circ$ neighborhood of each grid node over each month. The estimates of the T_{L4} uncertainty, S_{L4}, are obtained for any location by interpolating the gridded monthly SDs.

The condition for comparing T_{NIGHT} versus T_{L4} takes the following form:

$$\gamma_1 S_{L4} < T_{NIGHT} - d(D) - T_{L4} < \gamma_2 S_{L4} \quad (1)$$

The function $d(D)$ accounts for the dependency of $M(T_{NIGHT} - T_{L4})$ on D, corresponding to curve 1 in Figure 7, and the S_{L4} is a function of geographical coordinates. The coefficients γ_1 and γ_2 are selected based on the analysis of $T_{SAT} - T_{IS}$ statistics over 2012–2018 and all matchups, both nighttime and daytime, from the MDS listed in Table 2. Figure 8 shows these statistics as a function of γ:

$$\gamma = (T_{NIGHT} - d(D) - T_{L4})/S_{L4} \quad (2)$$

In Figure 8a, the medians $M(T_{SAT} - T_{IS})$ depend on γ, and their shapes are similar for all satellites. This confirms a direct link between the γ and systematic errors in T_{IS}. Based on the dependencies in Figure 8a, we select $\gamma_1 = -1.7$ and $\gamma_2 = 1.5$, which corresponds to the ± 1 K limits for $M(T_{SAT} - T_{IS})$. Figure 8b shows the RSDs, $S(T_{SAT} - T_{IS})$ as functions of γ. The selected values of γ_1 and γ_2 correspond to the RSDs of ~ 1 K at $\gamma_1 = -1.7$ and ~0.6 K at $\gamma_2 = 1.5$. Figure 8c shows the fractions of T_{IS}'s in all daily segments as a function

of γ. The above selection of γ_1 and γ_2 results in rejecting ~0.25% of T_{IS}'s belonging to the segments with $\gamma < \gamma_1$ and ~0.2% of T_{IS}'s in the segments with $\gamma > \gamma_2$.

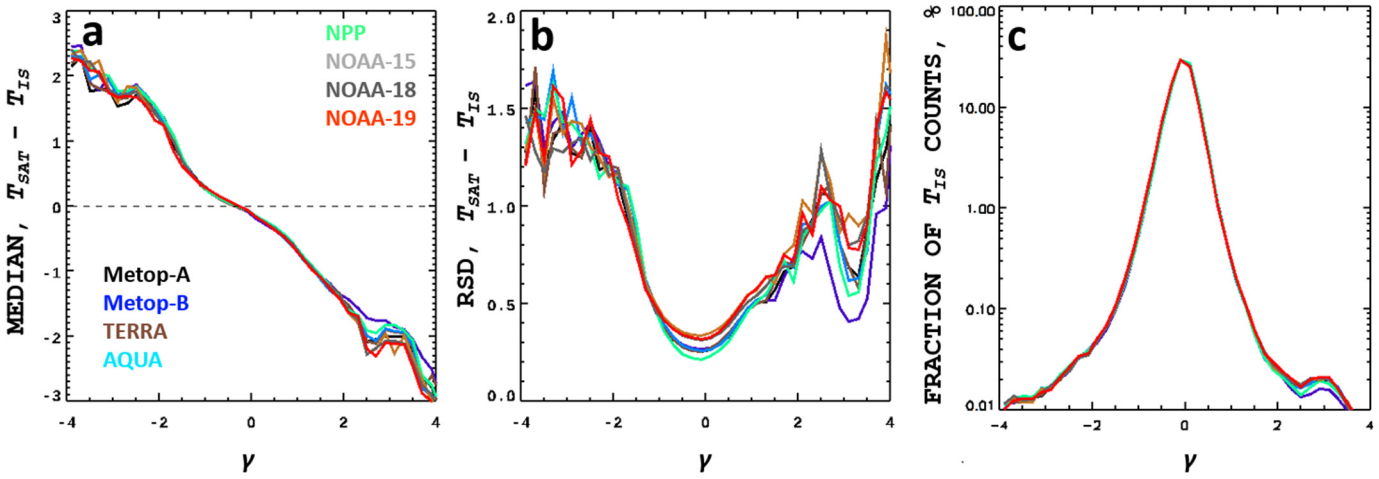

Figure 8. (**a**) Medians and (**b**) RSDs of $T_{SAT} - T_{IS}$ and (**c**) fractions of T_{IS}'s, as a function of γ for satellites listed in Table 2.

3.5. Formulation of the DRC

In summary, the Diurnal Reference Check is formulated as follows.
Each daily segment is checked with the following conditions:

$$D < 5 \text{ K} \tag{3}$$

$$T_{L4} - 1.7S_{L4} < T_{NIGHT} - d(D) < T_{L4} + 1.5S_{L4} \tag{4}$$

If at least one of the two conditions (3) or (4) is not met, the whole segment is rejected. If conditions (3) and (4) are both satisfied, then individual T_{IS}'s within a given segment are checked for outliers with the following condition:

$$a(D) - \alpha b(D) < T_{IS} - T_{NIGHT} < f(D) + \alpha c(D) \tag{5}$$

In (5), $a(D)$ and $f(D)$ are functions of D corresponding to curves 4 and 3 in Figure 7b; $b(D)$ and $c(D)$ are functions of D corresponding to curves 4 and 3 in Figure 7c; and $\alpha = 2$ is an empirical coefficient. The T_{IS} is rejected if the condition (5) is not met.

4. Results

This section validates the newly proposed DRC check.

4.1. Examples of Filtering Data of Individual Buoys

We reprocessed *i*Quam data from DTM buoys for 1981–2021 with the experimental set of QC checks (EXP QC). In addition to the DRC, the EXP QC retained the five current binary *i*Quam checks (DR, GC, TC, SC, and IC). The Buddy Check was not included in the EXP QC because its effect was found to be insignificant. The highest QL = 5 was assigned in EXP QC to T_{IS}'s that pass all six checks. Figures 9–12 compare the results of processing selected drifters' records with the current *i*Quam v2.10 QC and the EXP QC. The matching T_{SAT} were selected from the MDS listed in Table 2. If matchups for the specific T_{IS} are found in more than one MDS, then the median of all matching T_{SAT} is taken as a single matchup.

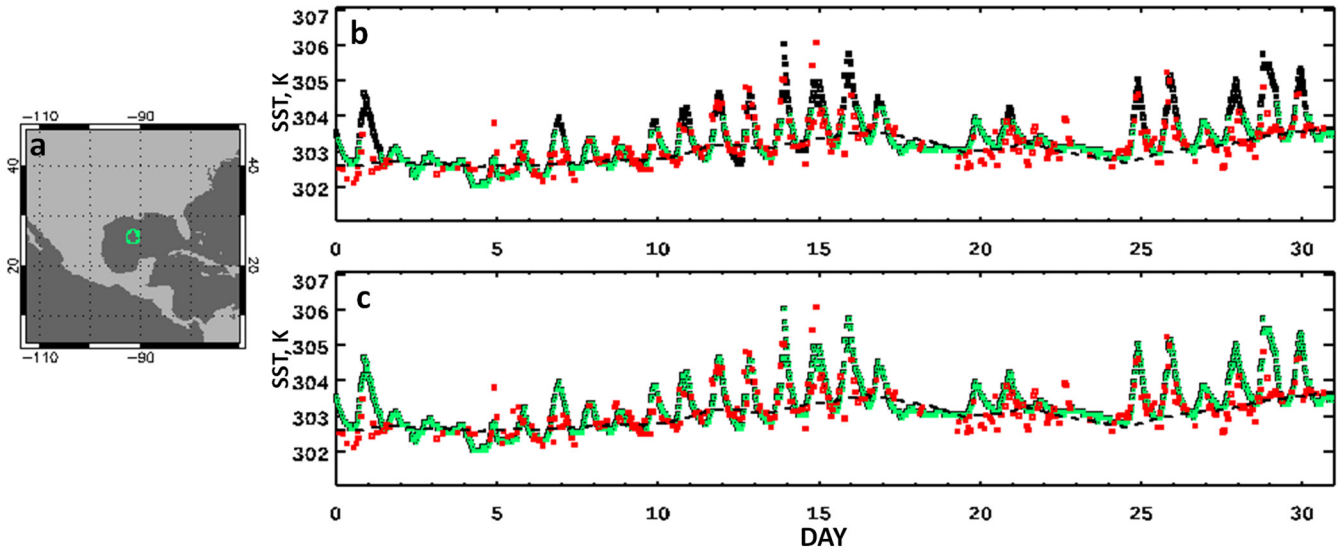

Figure 9. Filtering the T_{IS} record for the drifter 4201591 in July 2018. (**a**) Drifter's track in the Gulf of Mexico; Results of processing with (**b**) current *i*Quam v2.10 QC and (**c**) EXP QC. The points show (black) T_{IS} with QL < 5, (green) T_{IS} with QL = 5, and (red) matched T_{SAT}. T_{L4} is shown with a black dashed line.

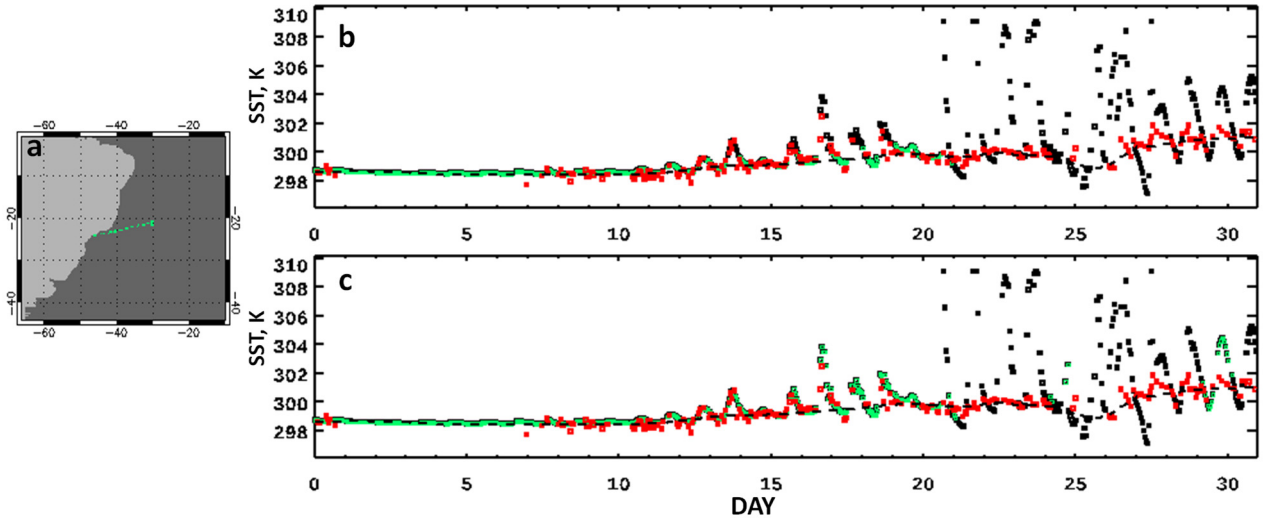

Figure 10. Same as in Figure 9 but for buoy 1500607 drifting to the Brazilian coast in December 2018.

Figure 9 shows the results of QC'ed T_{IS} for the drifter 4201591 in the Gulf of Mexico in July 2018. Significant ADSs up to 5 K were observed over 19 days, and consistent diurnal variations are seen in T_{SAT}. The current QC cuts off the peak values of daytime T_{IS}'s (cf. Figure 9b), whereas the EXP QC preserves the full DS (Figure 9c).

Figure 10 shows the SST record for the drifter 1500607 in the Atlantic Ocean near Brazil in December 2018. During the second decade of the month, the T_{IS} showed ADSs from 1–5 K and consistent diurnal variations in T_{SAT}. In the third decade, the ADSs increased to ~10 K, but the T_{SAT} did not show such diurnal variations. This suggests that these large ADSs in T_{IS} were caused by extraneous factors rather than variations in the water temperature. In Figure 10b, the current *i*Quam RC rejects all significant deviations from T_{L4} after December 10. In Figure 10c, the EXP QC preserves the diurnal T_{IS} variations from 10–20 December, consistent with T_{SAT}, but rejects the daily segments with unrealistic ADSs from 20–30 December. Note, however, that the EXP QC also preserves the DS on 29 December, when its amplitude is less than 5 K, despite its inconsistency with T_{SAT}. This

may suggest the need for a more conservative threshold for the ADS in the coastal zones. Such refinements will be explored in the future.

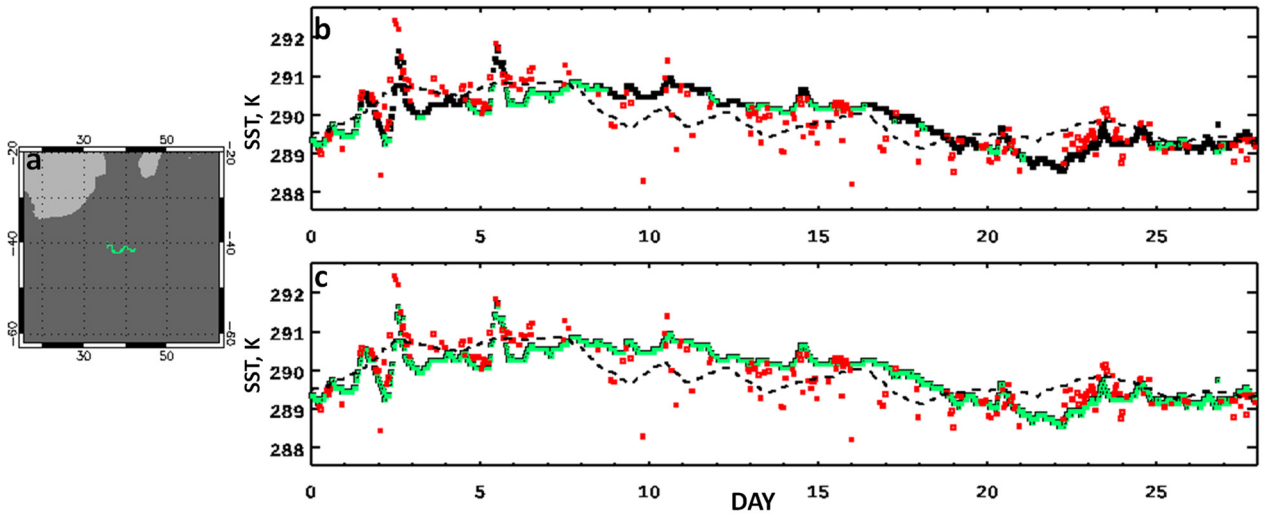

Figure 11. Same as in Figure 9 but for drifter 3300688 in February 2017.

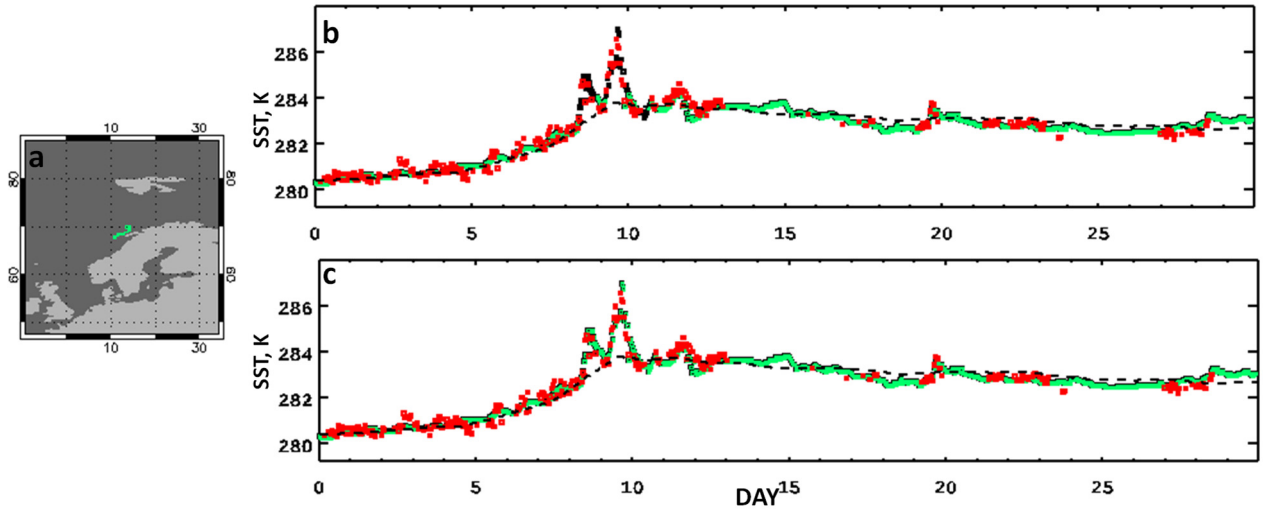

Figure 12. Same as in Figure 9 but for drifter 6500599 near the Norwegian coast in June 2017.

Figure 11 shows T_{IS} from buoy 3300688 in the Agulhas current in February 2017. The current QC rejects a significant number of T_{IS}'s (which note are reasonably consistent with T_{SAT}; Figure 11b). Significant deviations of T_{IS} from T_{L4} are due to the diurnal warming (on days 2, 3, 6, 14, 20, and 24) and inaccurate T_{L4} (typical for a dynamic zone). In contrast, the EXP QC preserves the whole record (cf. Figure 11c), thanks to the ADS-dependent thresholds for $T_{IS} - T_{NIGHT}$ and adequately liberal thresholds for $T_{NIGHT} - T_{L4}$.

Lastly, Figure 12 shows the T_{IS} for buoy 6500599 near the Norwegian coast in June 2017. Significant diurnal variations in T_{IS}, consistent with T_{SAT}, occurred on 8–9 June. The possibility of large diurnal SST variations in the high latitudes was discussed in [14,15]. The current RC rejects daytime T_{IS} due to significant deviations from T_{L4} (Figure 12b), whereas the EXP QC fully preserves the DS in T_{IS} (Figure 12c).

4.2. The Impact of the QC on Temporal and Spatial Variability of In Situ SST

Figure 13 illustrates the impact of the QC on spatial and temporal variability of high-quality ΔT_{L4}. Figure 13a,b show the composite maps of bias and SD of ΔT_{L4} produced from DTM T_{IS} counts sampled between 8 a.m. and 8 p.m. LST and averaged over 2012–

2018 without preliminary QC. Figure 13c,d show the same statistics produced from T_{IS} of QL = 5 by the *i*Quam QC.

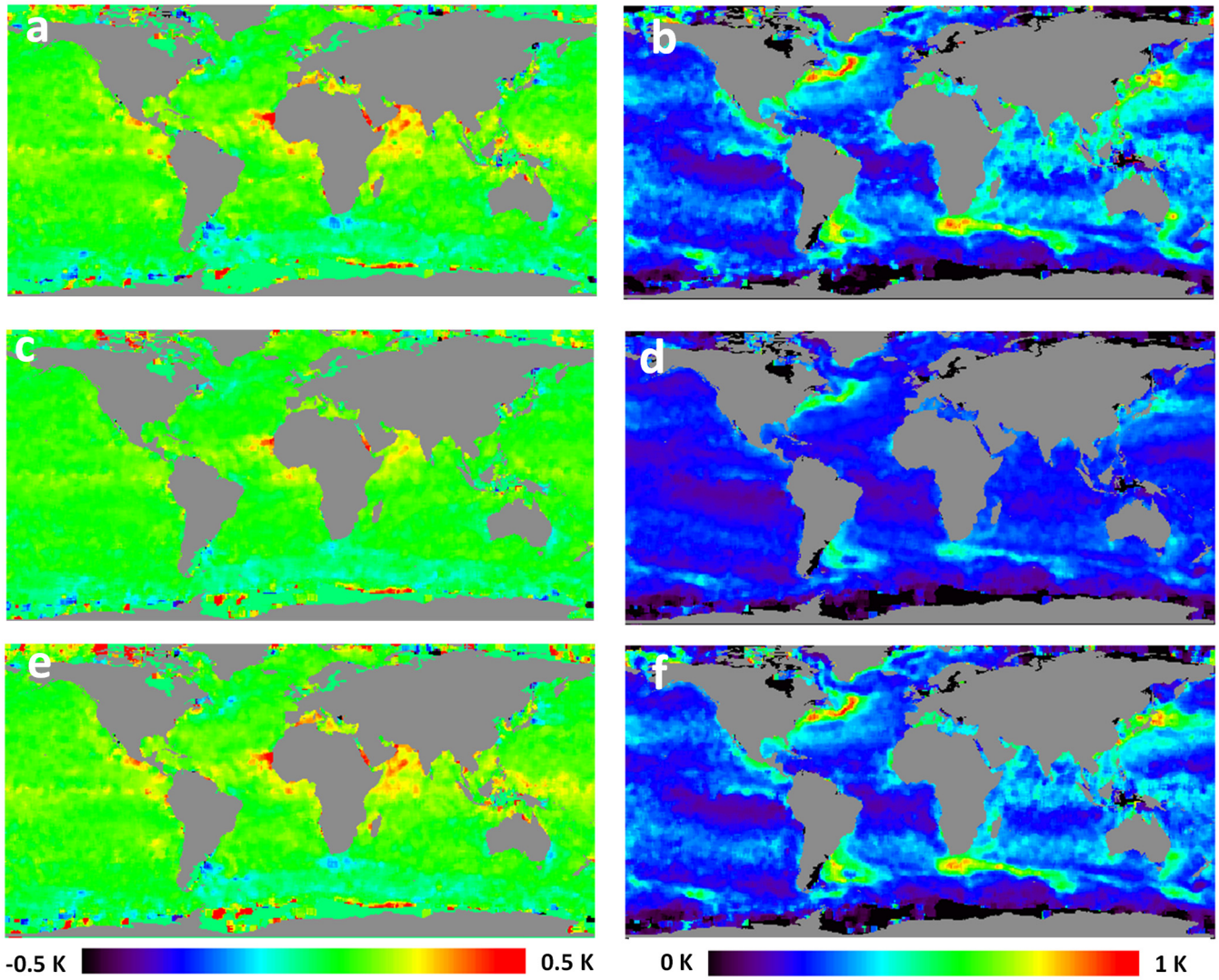

Figure 13. (**a**,**c**,**e**) Biases and (**b**,**d**,**f**) SDs of daytime ΔT_{L4}, sampled between 8 a.m. and 8 p.m. local solar time (LST) by drifting and tropical moored buoys averaged over 2012–2018, produced from T_{IS} with (**a**,**b**) current *i*Quam QL = 3–5, (**c**,**d**) current *i*Quam QL = 5 only; and (**e**,**f**) EXP QL = 5.

The current *i*Quam QC suppresses the original spatial and temporal variability of ΔT_{L4}. Figure 13e,f show the maps of similar statistics for ΔT_{IS} of QL = 5 processed with the QC EXP. These maps are much closer to those shown in Figure 13a,b. We conclude that the QC EXP better preserves the original variability of ΔT_{L4}.

As shown in Figure 7a, the negative difference $T_{NIGHT} - T_{L4}$ increases and becomes less stable with increased ADS. Recall the CMC SST is largely produced from nighttime SSTs, with a small fraction of those daytime data deemed least affected by the diurnal warming. Given that the criteria for assimilating in situ and satellite SSTs are different in different L4 analyses, it is instructive to estimate the global diurnal cycle in ΔT_{L4}'s produced by subtracting different analyses from T_{IS}'s. Figure 14a shows such global mean diurnal cycles produced from T_{IS} with QL = 5 determined by the current *i*Quam QC for eleven L4 analyses listed in Table 3. The averaging periods for each ΔT_{L4} are listed in the right column. Note that all curves are normalized at zero (i.e., L4-specific global biases subtracted from all curves to center them all at 0 K). The mean ADS in ΔT_{L4} is ~0.17 K, with maxima and minima varying within ~0.03 K between different L4s. Figure 14b re-plots

the same diurnal cycle but now produced from T_{IS} with the EXP QL = 5. The EXP QC increases global mean ADSs to ~0.23 K, ~35% larger than with the current *i*Quam QC, and mitigates the differences between various L4 analyses. One concludes that the EXP QC better preserves the diurnal variability in the T_{IS}, mitigates the excessive forcing of QL = 5 T_{IS}'s to the specific L4 analysis, and reconciles the DSs produced with different L4 analyses.

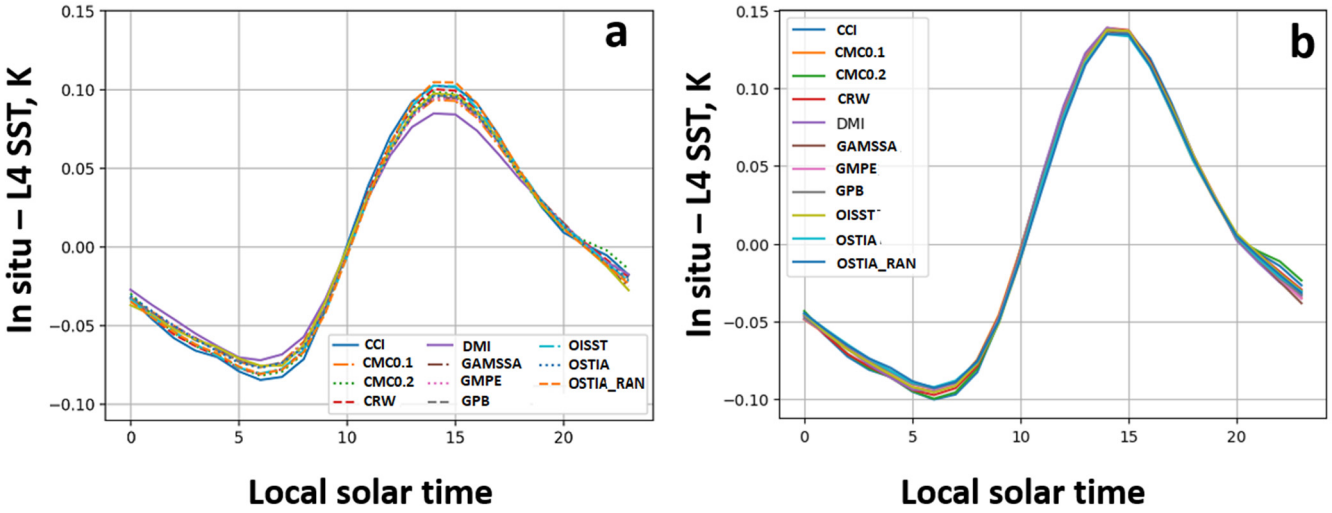

Figure 14. The average global diurnal cycle in $\Delta T_{L4} = T_{IS} - T_{L4}$ with T_{L4}'s obtained from eleven different L4 analyses listed in figures, averaged over the periods shown in Table 3. T_{IS} with QL = 5 by (**a**) current *i*Quam and (**b**) EXP QC.

Table 3. The L4 SST products used in comparison of diurnal signals and periods of their averaging. All links accessed on 22 August 2023.

L4 Analysis	Access	Period of Averaging (DD.MM.YYYY)
CCI	https://catalogue.ceda.ac.uk/uuid/aced40d7cb964f23a0fd3e85772f2d48	01.09.1981–31.12.2016
CMC0.2	https://podaac.jpl.nasa.gov/dataset/CMC0.2deg$-$CMC-L4-GLOB-v2.0	01.09.1991–31.12.2015
CMC0.1	https://podaac.jpl.nasa.gov/dataset/CMC0.1deg-CMC-L4-GLOB-v3.0	01.01.2016–31.12.2021
CRW	https://coralreefwatch.noaa.gov/product/5km/index.php#data_access	01.01.1985–31.12.2021
DMI	https://podaac.jpl.nasa.gov/dataset/DMI_OI-DMI-L4-GLOB-v1.0	30.04.2013–31.12.2021
GAMSSA	https://doi.org/10.5067/GHGAM-4FA1A	23.07.2008–31.12.2021
GMPE	http://ghrsst-pp.metoffice.gov.uk/ostia-website/gmpe-monitoring.html	17.09.2009–31.12.2021
GPB	https://coastwatch.noaa.gov/cwn/products/noaa-geo-polar-blended-global-sea-surface-temperature-analysis-level-4.html	02.06.2014–31.12.2021
OISST	https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.highres.html	01.09.1981–31.12.2021
OSTIA	https://podaac.jpl.nasa.gov/dataset/OSTIA-UKMO-L4-GLOB-v2.0	31.12.2006–31.12.2021
OSTIA-RAN	https://doi.org/10.48670/moi-00168	01.10.1981–31.12.2021

4.3. Time Series of Statistics of T_{IS} vs. T_{SAT} and Satellite-Based Retrospective QC

This section compares the statistics of T_{IS} with QL = 5 vs. T_{SAT} for 1981–2021 produced with the current *i*Quam and the EXP QCs. The T_{SAT}'s are obtained from all MDSs available for each period. As mentioned in Section 3.1, if matchups for a given T_{IS} are found in several MDSs, then the median of all matching T_{SAT} is taken as a single matchup. We also explore the potential of introducing an additional satellite-based QL (SAT QC), which is assigned to T_{IS} counts of QL = 5 by EXP QC, matching T_{SAT} and satisfying the following condition:

$$|(T_{IS} - T_{SAT} - \mu)| < 3\sigma \quad (6)$$

Here, μ and σ are the mean and SD of $T_{IS} - T_{SAT}$ averaged over a full monthly set of matchups with QL = 5 as determined by the EXP QC. The fractions of matchups satisfying

condition (6) in the monthly sets of matchups range from 98.7% to 99%. The time series of monthly statistics of T_{IS} vs. T_{SAT} for 1981–2021 are shown in Figure 15.

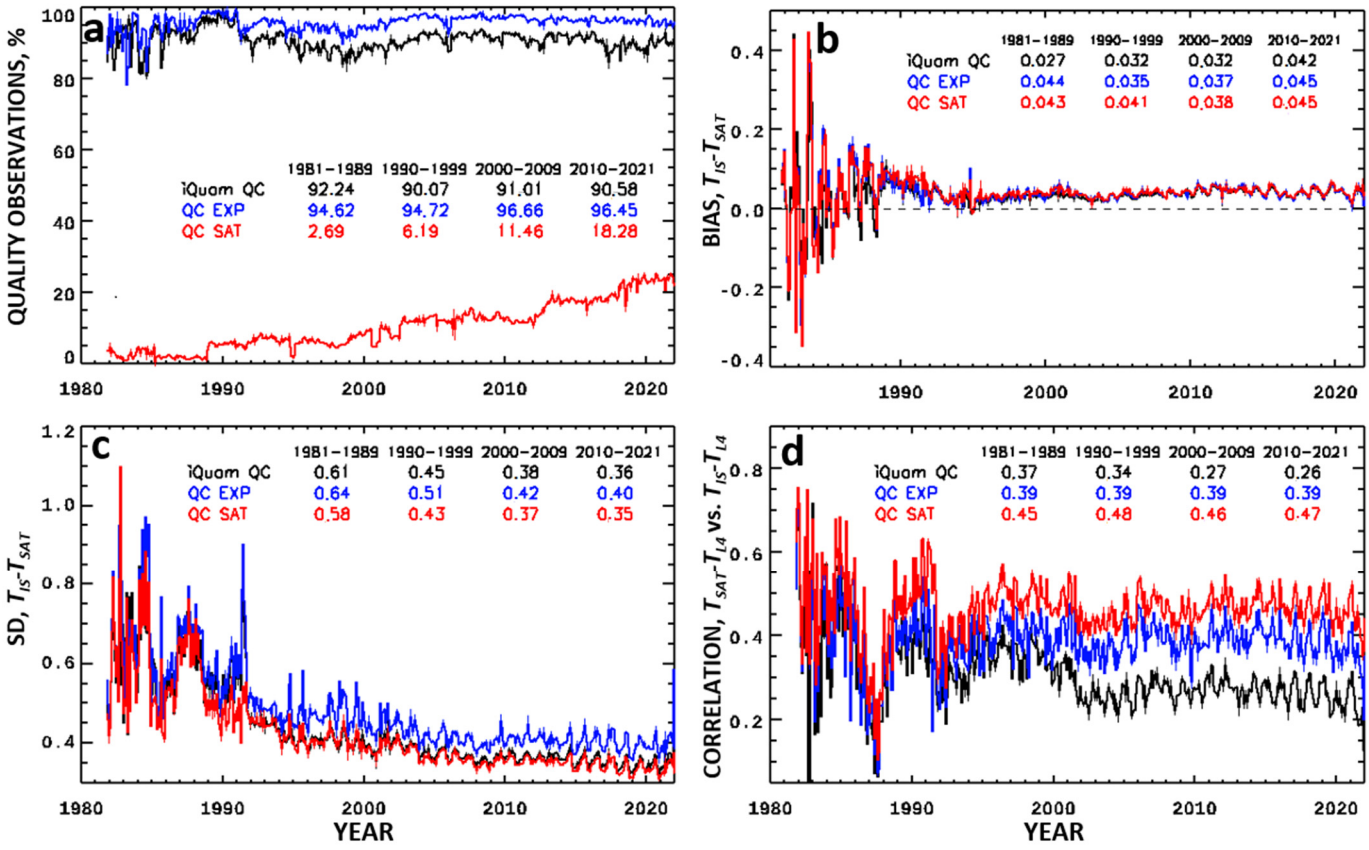

Figure 15. Time series of monthly (**a**) fractions of quality T_{IS} observations, (**b**) biases of $T_{IS} - T_{SAT}$, (**c**) SDs of $T_{IS} - T_{SAT}$, and (**d**) correlation coefficients of $T_{IS} - T_{L4}$ vs. $T_{SAT} - T_{L4}$, for (black) current *i*Quam, (blue) EXP and (red) SAT QCs. The overlaid numbers show mean values of the statistics for four periods.

Figure 15a shows a time series of fractions of QL = 5 T_{IS} identified by the current *i*Quam QC, EXP QC, and SAT QC. The EXP QC performs more liberal screening than the current QC, producing 2.5% more quality T_{IS}'s counts in 1981–1989 and 6–7% more counts in the later years. The fraction of matchups passing the SAT QC increases over the years consistently with the numbers of available satellite MDSs (cf. time series of monthly fractions of matched T_{IS} before the QC and the numbers of available MDSs in Figure 1b), reaching ~18–20% in 2010–2021 and ~25% in 2020–2021.

In Figure 15b, the time series of $\Delta T_{SAT} = T_{IS} - T_{SAT}$ produced by current, EXP, and SAT QCs are very similar. However, the biases for the EXP QC are somewhat warmer, due to better preserving the daytime T_{IS}'s affected by the diurnal warming.

Figure 15c shows the time series of SDs of ΔT_{SAT} for the three QCs. The EXP QC produces larger SDs than the current QC. The SAT QC reduces SDs below the current QC levels by rejecting a small number of outliers in ΔT_{SAT}. Note that the removal of outliers by SAT QC does not suppress the DS in the T_{IS}, as attested by the fact that in 1990–2021, the SAT QC produced biases comparable with or warmer than those produced by the EXP QC (cf. Figure 15b). Note that the SAT QC eliminated the artificial peak in SDs in 1991 due to inaccurate T_{IS}'s that survive both current *i*Quam and EXP QCs.

Large variations in biases and SDs in the 1980s are mainly due to a relatively small number of matchups (several hundred per month, according to Figure 1a). Degraded quality T_{SAT} caused by instrumental problems of the AVHRR/2 radiometers on the early

NOAA-07/09/11 satellites also contributes, as well as training SST equations for NOAA-07/09 against the combinations of DTM and ship SSTs [19].

Figure 15d shows time series of monthly correlations between $T_{IS} - T_{L4}$ and $T_{SAT} - T_{L4}$ with T_{L4} produced from the OISST [23,24]. The EXP QC increases correlation coefficients compared to the current *i*Quam QC due to improved preservation of the daytime T_{IS}'s affected by the diurnal warming and consistent with T_{SAT}. Rejecting additional outliers by the SAT QC increases the correlation even further.

Although the numbers of T_{IS}'s identified by the SAT QC are relatively small, these are the very matchups used in the Cal/Val of the satellite products listed in Table 1. Filtering outliers in the MDS with the condition (5) will facilitate the Cal/Val during subsequent reprocessings of the satellite data.

5. Conclusions

The *i*Quam QC, as well as other in situ QC systems, relies on comparing individual in situ SST measurements with reference SST obtained from L4 analyses. The challenge with such comparisons is that true significant deviations from reference SST, caused by diurnal warming or degraded feature resolution of the reference SST in the dynamic zones of the World Ocean, can be screened out. The Diurnal Reference Check (DRC), proposed in this study, mitigates the overscreening problem for drifting and tropical moored buoys, whose sampling frequency is sufficient for identifying and estimating the diurnal signal.

To optimize the DRC performance in the presence of a significant diurnal signal in in situ SSTs, the SST records from individual buoys were subdivided into 24-h daily segments, and each segment was characterized by the median nighttime SST and the amplitude of a diurnal signal. This allowed separate optimization of filtering individual SST outliers from filtering daily segments affected by significant systematic errors or extraneous factors, such as the overheating of the buoys' hulls.

The main DRC features and thresholds have been determined from the analyses of the statistics of deviations of in situ SSTs from reference (L4) SST, median nighttime SSTs, and satellite 'Subskin' SSTs averaged over the period 2012–2018. The matchups with satellite 'Subskin' SST, used in this analysis, came from long-term full-mission historical reprocessings (RANs) of the data of satellite radiometers AVHRR GAC onboard NOAA-15/18/19, AVHRR FRAC onboard Metop-A/B/C, VIIRS onboard S-NPP/NOAA-20, and MODIS onboard the Terra/Aqua satellites with the NOAA ACSPO system.

Based on the analysis of the statistics of in situ SST versus 'Subskin' SST, the upper threshold for the amplitude of the diurnal signal in the water temperature was set at 5 K. The diurnal segments, in which the amplitudes of the diurnal signals exceed the threshold, are rejected, assuming that these unrealistic amplitudes are mainly caused by extraneous factors.

It was found that in the presence of significant diurnal signals, the statistics of deviations of in situ SST from reference SST are less stable than the corresponding statistics of deviations from median nighttime SST. Therefore, the individual in situ SSTs are checked for outliers by comparison with the median nighttime SSTs rather than reference SSTs. This stabilizes the results of screening the individual outliers.

The reference L4 SST is used in the DRC only to identify the daily sets affected by significant systematic errors (calibration trends, 'gross' errors, etc.) by comparison with median nighttime SST. The thresholds used in this comparison are geographically dependent, being set proportional to the local spatial/temporal variability of the reference SST estimated on a monthly basis. The coefficient of proportionality was derived from the matchups with 'Subskin' SST such that to limit the maximum systematic error at ± 1 K.

The previous studies [5,9] suggested that the overscreening caused by comparing in situ SST with reference SST during diurnal warming events can be mitigated using models or reference SST data sets that capture the diurnal warming cycle. The DRC does not require such information because the reference SST is used only for comparison with the median nighttime SST, which, in turn, is used as a reference in filtering individual in situ

SST outliers. The DRC would rather benefit from the availability of the L4 analysis derived solely from nighttime data because assimilation of daytime SSTs, even after excluding those most affected by the diurnal warming, introduces additional 'noise' in the deviations between the median nighttime and reference SSTs.

We compared the results of the experimental QC, which includes the DRC along with a number of the heritage *i*Quam QC checks, with the ones of the current *i*Quam QC. We further validated both QCs against satellite SSTs from available data sets of matchups for September 1981–December 2021 from several ACSPO RANs, including AVHRR GAC RAN2 from NOAA-07/09/11/12/14/15/16/17/18/19; AVHRR FRAC RAN1 from Metop-A/B/C; VIIRS RAN3 from S-NPP/NOAA-20; and MODIS RAN1 from Terra/Aqua. The simultaneous use of matchups from satellites flying in different orbits allows for validating the diurnal signals observed in in situ SSTs, as well as discriminating the 'real' diurnal variations in the water temperature from 'false' diurnal signals caused by the extraneous factors.

Compared with the current *i*Quam QC, the DRC better preserves true diurnal variations in in situ SST, as well as its deviations from reference SST in the dynamic zones. At the same time, it efficiently rejects unrealistically large deviations from the reference SST. As a result, the DRC minimizes suppression of spatial and temporal variability in in situ SST and reduces the dependency of the QC results on the specific L4 analysis. In particular, the DRC reconciles the estimates of the mean global diurnal cycle derived from deviations of in situ SST from different L4 analyses, and increases their amplitudes. Overall, the DRC increases the monthly numbers of quality in situ SST measurements by 4–6% and improves the correlation between deviations of satellite and in situ SSTs from reference SST at the expense of increased SDs of deviations of in situ from satellite SST.

We also explored the potential of introducing the additional, satellite-based quality level for in situ SSTs based on matching them with satellite SSTs. The exclusion of ~1% of outliers from the matchups reduces the standard deviations of in situ SSTs minus satellite SSTs below the levels typical for the current *i*Quam QC and significantly increases the correlation between $T_{IS} - T_{L4}$ and $T_{SAT} - T_{L4}$. The fraction of in situ measurements to which this new QL can be assigned, is relatively small in the earlier years, ranging from ~2.6% in the 1980s and increasing to 18–25% in the 2010s. However, these are the very same matchups used in Cal/Val of satellite retrievals. It is expected, therefore, that the elimination of these outliers will improve the satellite Cal/Val.

Work is underway to implement the DRC in the official future releases of the *i*Quam. Future work will also aim at the development of advanced QC procedures for the platforms with a relatively low temporal sampling rate (ships, Argo floats), evaluation of the effects of the new QC on the training and validation of satellite SST retrievals, and further refining the DRC parameters as needed.

Author Contributions: Conceptualization, B.P. and A.I.; methodology, B.P.; software, B.P.; validation, B.P., V.P. and O.J.; formal analysis, B.P.; investigation, B.P., A.I. and V.P.; resources, A.I.; data curation, V.P. and O.J.; writing—original draft preparation, B.P.; writing—review and editing, A.I. and BP.; visualization, B.P. and V.P.; supervision, A.I.; project administration, A.I.; funding acquisition, A.I. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the following NOAA Programs: JPSS (Lihang Zhou, JPSS Manager; Satya Kalluri, JPSS Program Scientist; Ingrid Guch, JSTAR Manager), GOES-R (Jaime Daniels, Chair, Algorithm Working Group), ORS (Paul DiGiacomo and Marilyn Yuen-Murphy, Managers), and NESDIS Innovation (Mitch Goldberg, NESDIS Chief Scientist and Program Manager).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: In situ data processed with the current *i*Quam v.2.10 are available at https://www.star.nesdis.noaa.gov/socd/sst/iquam/ (accessed on 3 September 2023). In situ data reprocessed with the modified *i*Quam will be available after updating the official *i*Quam version.

Acknowledgments: We thank NOAA JPSS, GOES-R, ORS, and NESDIS Innovation Programs for sustained support of *i*Quam development. The views, opinions, and findings in this report are those of the authors and should not be construed as an official NOAA or U.S. Government position or policy.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

Acronym	Definition
3S	Sensor Stability for SST system
AATSR	Advanced Along-Track Scanning Radiometer
ACSPO	Advanced Clear Sky Processor for Ocean
ADS	Amplitude of Diurnal Signal
AOML	Atlantic Oceanographic and Meteorological Laboratory
AVHRR	Advanced Very High Resolution Radiometer
BC	Buddy Check
CCI	Climate Change Initiative
CMC	Canadian Meteorological Center
CMEMS	Copernicus Marine Environment Monitoring Service
CRW	Coral Reef Watch
DMI	Danish Meteorological Institute
DTM	Drifters and Tropical Moored buoys
DRC	Diurnal Reference Check
DS	Diurnal Signal
ENVISAT	Environmental Satellite
FNMOC	Fleet Numerical Meteorology and Oceanography Center
FRAC	1 km Full Resolution Area Coverage mode
GAC	4 km Global Area Coverage mode
GAMSSA	Global Australian Multi-Sensor SST Analysis
GC	Plausibility/Geolocation Check
GDAC	Global Data Assembly Centers
GMPE	SST Global Multi-Product Ensemble
GPB	Geo Polar Blended SST
IC	Platform ID Check
ICOADS	International Comprehensive Ocean-Atmosphere Data Set
IMOS	Integrated Marine Observing System (Australia)
*i*Quam	In Situ SST Quality Monitor
IR	Infrared
LEO	Low-Earth Orbiting
LEXT	Local Equator Crossing Time
MDS	Data Set of Matchups
MODIS	Moderate Resolution Imaging Spectroradiometer
OISST	Optimal Interpolation SST
OSTIA	Operational Sea Surface Temperature and Sea Ice Analysis
OSTIA-RAN	OSTIA Reanalysis
PGE	Probability of Gross Error
RSD	Robust Standard Deviation
SD	Standard Deviation
SST	Sea Surface Temperature
QC	Quality Control
QC EXP	Experimental QC
QC SAT	Satellite-based QC
QL	Quality Level
RC	Reference Check
SC	SST Spike Check
SQUAM	SST Quality Monitor
TC	Platform Track Check
VIIRS	Visible Infrared Imaging Radiometer Suite

References

1. In Situ SST Quality Monitor. Available online: https://www.star.nesdis.noaa.gov/socd/sst/iquam/ (accessed on 22 August 2023).
2. Xu, F.; Ignatov, A. In situ SST quality monitor (*i*Quam). *J. Atmos. Ocean. Technol.* **2014**, *31*, 164–180. [CrossRef]
3. Xu, F.; Ignatov, A. Error characterization in *i*Quam SSTs using triple collocations with satellite measurements. *Geophys. Res. Lett.* **2016**, *43*, 10826–10834. [CrossRef]
4. Brisson, A.; Le Borgne, P.; Marsouin, A. Results of one year of preoperational production of sea surface temperatures from GOES-8. *J. Atmos. Ocean. Technol.* **2002**, *19*, 1638–1652. [CrossRef]
5. Atkinson, C.P.; Rayner, N.A.; Roberts-Jones, J.; Smith, R.O. Assessing the quality of sea surface temperature observations from drifting buoys and ships on a platform-by-platform basis. *J. Geophys. Res. Oceans* **2013**, *118*, 3507–3529. [CrossRef]
6. Kilpatrick, K.A.; Podestá, G.P.; Evans, R. Overview of the NOAA/NASA Advanced Very High Resolution Radiometer Pathfinder algorithm for sea surface temperature and associated matchup database. *J. Geophys. Res.* **2001**, *106*, 9179–9197. [CrossRef]
7. Castro, S.L.; Wick, G.A.; Emery, W.J. Evaluation of the relative performance of sea surface temperature measurements from different types of drifting and moored buoys using satellite-derived reference products. *J. Geophys. Res.* **2012**, *117*, C02029. [CrossRef]
8. Marsouin, A.; Le Borgne, P.; Legendre, G.; Péré, S.; Roquet, H. Six years of OSI-SAF MetOp-A AVHRR sea surface temperature. *Remote Sens. Environ.* **2015**, *159*, 288–306. [CrossRef]
9. Zhang, H.; Ignatov, A.; Hinshaw, D. Evaluation of the in situ Sea Surface Temperature Quality Control in the NOAA in situ SST Quality Monitor (*i*Quam) System. *J. Atmos. Ocean. Technol.* **2021**, *38*, 1249–1263. [CrossRef]
10. Lorenc, A.C.; Hammon, O. Objective quality control of observations using Bayesian methods. Theory, and a practical implementation. *Q. J. R. Meteorol. Soc.* **1988**, *114*, 515–543. [CrossRef]
11. Donlon, C.; Robinson, I.; Casey, K.S.; Vazquez-Cuervo, J.; Armstrong, E.; Arino, O.; Gentemann, C.; May, D.; LeBorgne, P.; Piollé, J.; et al. The Global Ocean Data Assimilation Experiment High-resolution Sea Surface Temperature Pilot Project. *Bull. Am. Meteorol. Soc.* **2007**, *88*, 1197–1214. [CrossRef]
12. Gentemann, C.L.; Donlon, C.J.; Stuart-Menteth, A.; Wentz, F.J. Diurnal signals in satellite sea surface temperature measurements. *Geophys. Res. Lett.* **2003**, *30*, 1140. [CrossRef]
13. Gentemann, C.L.; Minnett, P.J.; Le Borgne, P.; Merchant, C.J. Multi-satellite measurements of large diurnal warming events. *Geophys. Res. Lett.* **2008**, *35*, L22602. [CrossRef]
14. Eastwood, S.; Le Borgne, P.; Péré, S.; Poulter, D. Diurnal variability in sea surface temperature in the Arctic. *Remote Sens. Environ.* **2011**, *115*, 2594–2602. [CrossRef]
15. Kawai, Y.; Wada, A. Diurnal Sea Surface Temperature Variation and Its Impact on the Atmosphere and Ocean: A Review. *J. Oceanogr.* **2007**, *63*, 721–744. [CrossRef]
16. Kennedy, J.J.; Brohan, P.; Tett, S.F.B. A global climatology of the diurnal variations in sea-surface temperature and implications for MSU temperature trends. *Geophys. Res. Lett.* **2007**, *34*, L05712. [CrossRef]
17. Kawamura, H.; Qin, H.; Ando, K. In-situ diurnal sea surface temperature variations and near-surface thermal structure in the tropical hot event of the Indo-Pacific warm pool. *J. Oceanogr.* **2008**, *64*, 847–857. [CrossRef]
18. Pryamitsyn, V.; Petrenko, B.; Ignatov, A.; Kihai, Y. Metop First Generation AVHRR FRAC SST Reanalysis Version 1. *Remote Sens.* **2021**, *13*, 4046. [CrossRef]
19. Petrenko, B.; Pryamitsyn, V.; Ignatov, A.; Jonasson, O.; Kihai, Y. AVHRR GAC Sea Surface Temperature Reanalysis Version 2. *Remote Sens.* **2022**, *14*, 3165. [CrossRef]
20. Jonasson, O.; Ignatov, A.; Pryamitsyn, V.; Petrenko, B.; Kihai, Y. JPSS VIIRS SST Reanalysis Version 3. *Remote Sens.* **2022**, *14*, 3476. [CrossRef]
21. Sensor Stability for SST (3S). Available online: https://www.star.nesdis.noaa.gov/socd/sst/3s/ (accessed on 22 August 2023).
22. Petrenko, B.; Ignatov, A.; Kihai, Y.; Dash, P. Sensor-Specific Error Statistics for SST in the Advanced Clear-Sky Processor for Oceans. *J. Atmos. Ocean. Technol.* **2016**, *33*, 345–359. [CrossRef]
23. Reynolds, R.W.; Smith, T.M.; Liu, C.; Chelton, D.B.; Casey, K.S.; Schlax, M.G. Daily high-resolution-blended analyses for sea surface temperature. *J. Clim.* **2007**, *20*, 5473–5496. [CrossRef]
24. NOAA Optimal Interpolation (OI) SST v.2. Available online: https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.highres.html (accessed on 22 August 2023).
25. GHRSST Level 4 CMC0.2deg Global Foundation Sea Surface Temperature Analysis (GDS Version 2). Available online: https://podaac.jpl.nasa.gov/dataset/CMC0.2deg-CMC-L4-GLOB-v2.0 (accessed on 22 August 2023).
26. GHRSST Level 4 CMC0.1deg Global Foundation Sea Surface Temperature Analysis (GDS Version 2). Available online: https://podaac.jpl.nasa.gov/dataset/CMC0.1deg-CMC-L4-GLOB-v3.0 (accessed on 22 August 2023).
27. Good, S.; Fiedler, E.; Mao, C.; Martin, M.J.; Maycock, A.; Reid, R.; Roberts-Jones, J.; Searle, T.; Waters, J.; While, J.; et al. The Current Configuration of the OSTIA System for Operational Production of Foundation Sea Surface Temperature and Ice Concentration Analyses. *Remote Sens.* **2020**, *12*, 720. [CrossRef]
28. Global Ocean OSTIA Sea Surface Temperature and Sea Ice Reprocessed. Available online: https://doi.org/10.48670/moi-00168 (accessed on 22 August 2023).
29. Merchant, C.J.; Embury, O.; Bulgin, C.E.; Block, T.; Corlett, G.K.; Fiedler, E.; Good, S.A.; Mittaz, J.; Rayner, N.A.; Berry, D.; et al. Satellite-based time-series of sea-surface temperature since 1981 for climate applications. *Sci. Data* **2019**, *6*, 223. [CrossRef]

30. ESA Sea Surface Temperature Climate Initiative Level 4 Analysis Climate Data Record, Versions 2. Available online: https://catalogue.ceda.ac.uk/uuid/aced40d7cb964f23a0fd3e85772f2d48 (accessed on 22 August 2023).
31. Bitterman, D.S.; Hansen, D.V. Evaluation of sea surface temperature measurements from drifting buoys. *J. Atmos. Ocean. Technol.* **1993**, *10*, 88–96. [CrossRef]
32. Kawai, Y.; Kawamura, H.; Tanba, S.; Ando, K.; Yoneyama, K.; Nagahama, N. Validity of Sea Surface Temperature Observed with the TRITON Buoy under Diurnal Heating Conditions. *J. Oceanogr.* **2006**, *62*, 825–838. [CrossRef]