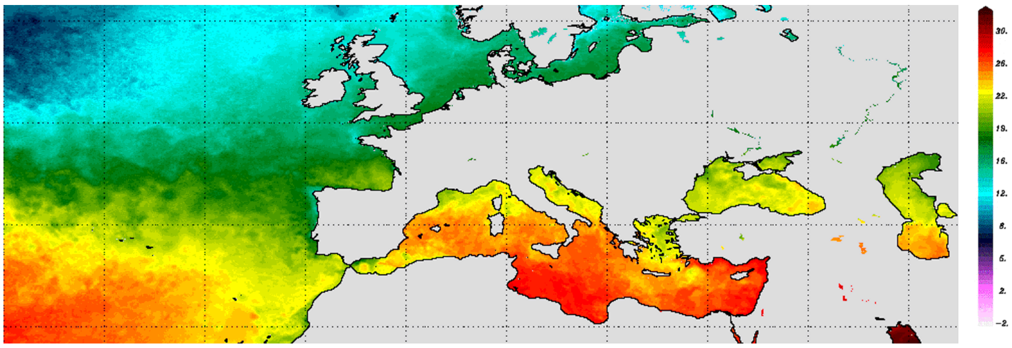
ESA CCI Phase 3 Sea Surface Temperature (SST)



End-To-End ECV Uncertainty Budget D2.2 v3.1

Issue Date: **03 November 2023**

Document Ref: SST_CCI_D2.2_E3UB_v3.1

Contract: ESA/AO/1-9322/18/I-NB

SIGNATURES AND COPYRIGHT

Title : **ESA CCI Phase 3 Sea Surface Temperature (SST)**

Volume : **End-To-End ECV Uncertainty Budget D2.2 v3.1**

Issued : **03 November 2023**

Authored :

Digitally signed by Sarah Connors
Date: 2023.12.12 13:35:11 Z

Dr Owen Embury (UoR)

Authored : Prof Chris Merchant (UoR)

C J Merchant

Address : University of Reading,
Whiteknights,
Reading,
Berkshire,
RG6 6AH,
United Kingdom

Copyright : © University of Reading 2023. The Copyright of this document is the property of the University of Reading. It is supplied on the express terms that it be treated as confidential, and may not be copied, or disclosed, to any third party, except as defined in the contract, or unless authorised by the University of Reading in writing.

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Purpose and Scope

In any climate data record (CDR), the uncertainty of data should be stated, and this should be done per datum in the record if uncertainty is variable [Reference Document 1, RD 1]. In ESA's climate change initiative project for Sea Surface Temperature (SST), this is done using methods addressed in this and previous reports on "End-to-end ECV Uncertainty Budget".

This document is the final (third) SST_CCI End-to-End Error and Uncertainty Budget (E3UB) report relevant to the SST CCI Version 3 Reprocessing available from:

- The CCI Open Data Portal: https://climate.esa.int/en/odp
- https://catalogue.ceda.ac.uk/uuid/1dc189bbf94209b48ed446c0e9a078af

This version covers the uncertainty validation of Level 3 and 4 products of the version 3 CDR. To put in place the context of this E3UB report, note the following:

- E3UB v1 [RD 2] described the new uncertainty model for optimal estimation retrieval of sea surface temperature used in v3.
- E3UB v2 [RD 3] described the method of estimation, within a framework of "bias aware optimal estimation", of error covariance matrices for AVHRR sensors required within the uncertainty model for optimal estimation retrieval.

This report reports the validation of uncertainty, thereby assessing the developments reported in E3UB v1 and v2.

This report should be read in conjunction with the following Reference Documents:

RD 1	Merchant, C. J., Paul, F., Popp, T., Ablain, M., Bontemps, S., Defourny, P., Hollmann, R., Lavergne, T., Laeng, A., de Leeuw, G., Mittaz, J., Poulsen, C., Povey, A. C., Reuter, M., Sathyendranath, S., Sandven, S., Sofieva, V. F., and Wagner, W. (**2017**) Uncertainty information in climate data records from Earth observation. Earth System Science Data, 9, 511-527. doi:10.5194/essd-9-511-2017
RD 2	Merchant, C. J. (**2020**) Uncertainty Characterisation (E3UB) D2.2 v1. ESA SST CCI technical report. https://climate.esa.int/media/documents/E3UB_D2.2_v1_signed.pdf
RD 3	Merchant, C. J. (**2021**) Uncertainty Characterisation (E3UB) D2.2 v2. ESA SST CCI technical report. https://climate.esa.int/media/documents/E3UB_D2.2_v2_signed.pdf
RD 4	Poli, P., Lucas, M., O'Carroll, A., Le Menn, M., David, A., Corlett, G.K., Blouch, P., Meldrum, D., Merchant, C.J., Belbeoch, M., Herklotz, K. (**2019**) The Copernicus Surface Velocity Platform drifter with Barometer and Reference Sensor for Temperature (SVP-BRST): genesis, design, and initial results. Ocean Sci. 15, 199–214. doi:10.5194/os-15-199-2019

1.2 Executive Summary

This document shows the end-to-end uncertainty validation results for sea surface temperature (SST) level 3 (gridded) and level 4 (gap-filled) products in the ESA climate change initiative project for SST.

Uncertainty validation uses the differences between satellite and matched reference data together with knowledge of in situ uncertainties to assess whether stated uncertainties in SST products (which are provided per datum) are consistent with the uncertainties seen in practice.

The document addresses in turn the uncertainty validation for sensor series used in SST CCI:

- dual-view sensors (ATSRs, SLSTRs)
- single-view sensors (AVHRRs)

The document also addresses the uncertainty validation for the gap-filled level 4 analysis product of SST.

The key results are:

- While ATSR-1 uncertainties are underestimated overall by ~30%, the uncertainties for ATSR-2 and AATSR are largely unbiased and capture well the variations in uncertainty for these sensors (§2.2).
- SLSTR A & B uncertainties are well estimated for the vast majority of data, although the largest uncertainties are somewhat underestimated.
- For AVHRRs, some retrieval uncertainties are very well estimated, and in other circumstances, the uncertainty model seems not to predict variations in uncertainty with significant skill (§2.3).
    - Retrievals at night using either 2 channels (3.7 and 11 µm, early AVHRR/1 sensors) or 3 channels (3.7, 11 and 12 µm, all later AVHRR sensors) have well estimated uncertainties in SST.
    - Retrievals in daytime using split-window channels in optimal estimation are not skilful in discriminating more and less uncertain SSTs. In general, uncertainties are over-estimated (the results are pessimistic).
- For the level 4 (gap-filled analysis) product, more and less certain data are well distinguished. The majority of data have quantitatively well estimated uncertainties, although in general these are pessimistic by about 20% to 30% (§2.4).

The results are overall satisfactory, other than for daytime single-view uncertainty estimates. The reason for the problem with these split-window retrievals is not known, and would be an area for further investigation in future studies.

1.3 Acronyms

The following acronyms and abbreviations have been used in this report with the meanings shown:

Acronym	Definition
AMSR	Advanced Microwave Scanning Radiometers
AVHRR	Advanced Very High Resolution Radiometers
ATSR	Along-Track Scanning Radiometers
C3S	Copernicus Climate Change Service
CCI	Climate Change Initiative
CDR	Climate Data Record
CMEMS	Copernicus Marine Environment Monitoring Service
DV	Diurnal Variability
ECMWF	European Centre for Medium-Range Weather Forecasts
EPS	EUMETSAT Polar System
ERS	European Remote Sensing
ESA	European Space Agency
GAC	Global Area Coverage
GTMBA	Global Tropical Moored Buoy Array
GTS	Global Telecommunication System
HadIOD	Hadley Centre Integrated Ocean Dataset
ICOADS	International Comprehensive Ocean-Atmosphere Dataset
L2	Level 2
L3	Level 3
L3C	Level 3 Collated
L3U	Level 3 Uncollated
L4	Level 4
MD	Matchup Dataset
MMS	Multi-sensor Matchup System
NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical Weather Prediction
POES	Polar Operational Environmental Satellites
RSD	Robust Standard Deviation
SIRDS	SST CCI Independent Reference Data Set
SLSTR	Sea and Land Surface Temperature Radiometers
SST	Sea Surface Temperature
SST-CCI	ESA Climate Change Initiative on SST
VOS	Voluntary Observing Ship
WMO	World Meteorological Organization

2. VALIDATION OF UNCERTAINTIES

2.1 Methodology

As the uncertainties for infra-red SST products have been estimated independently from in-situ data, we can use the in situ data to validate the uncertainty estimates. The approach used is to compare the validation statistics (robust standard deviation of difference between the satellite retrieval and reference data) against the expected uncertainty – including both the estimated uncertainty in the SST retrieval and the uncertainty in the reference in situ data. Statistics were generated for different levels of uncertainty ascribed to the retrievals, in order to determine if the uncertainties were valid across the full range of possible uncertainties.

Figure 1 shows a typical uncertainty validation plot. Along the x-axis is the estimated uncertainty which is calculated as part of the retrieval (i.e., it is estimated without the use of any in situ data). The y-axis represents the discrepancy between the satellite retrieval and in situ measurement, which will be affected by errors in both the satellite and in situ observations. Given an estimate of the uncertainty in the in situ data (in this example we assume 0.2 K for drifter observations [RD 4]), we can estimate the expected spread in the satellite to in situ comparison as $\sigma = \sqrt{\sigma_{ins} + \sigma_{sat}}$ which is shown in the plot with solid blue lines.

The shaded grey area shows the Robust Standard Deviation (RSD) difference as a function of the estimated uncertainty (x-axis). If the uncertainties have been correctly estimated, then the shaded area should match the expected RSD envelope shown in the solid blue lines. In the example shown in Figure 1 (left panel) the uncertainties are over estimated (grey area fall short of the envelope), while in Figure 1 (right panel) the uncertainties are well estimated (grey area matches the envelope). Additionally, we show the bias as a function of the estimated uncertainty with the orange error bars, and the density distribution of estimated uncertainties in the green violin plot. In the example plot we see that daytime retrievals have expected uncertainty 0.15 K and 0.95 K, though the majority of data are found between 0.25 K and 0.45 K; while the night-time data range between 0.15 K and 0.65 K, with the majority under 0.25 K.

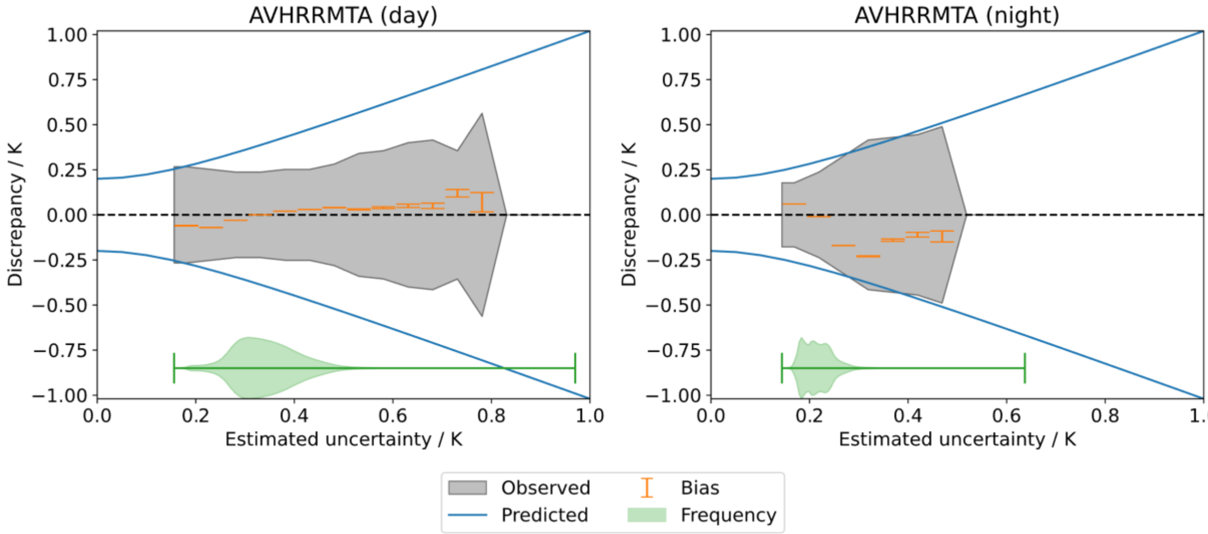


Figure 1: Example of uncertainty validation plot against in situ drifters for daytime (left) and night-time (right) data. Shaded grey area shows RSD discrepancy as a function of estimated uncertainty. Solid blue line shows expected relationship based on assumed in situ uncertainty of 0.2 K. Orange error bars show median discrepancy in each bin. Green violin plot shows the density distribution of the data.

2.2 Dual-View ATSR/SLSTR

Comparison of estimated ATSR uncertainties against drifters is shown in Figure 2 and tropical moorings in Figure 3. The ATSR2 and AATSR uncertainties are well estimated (shaded grey area is a good match to the blue, predicted, envelope), with the majority of data expected to be highly accurate with expected uncertainty $\lesssim 0.2$ (day) and $\lesssim 0.1$ (night). The discrepancies against in situ are then consistent with the assumed uncertainties in the in situ (0.2 K for drifters, and 0.1 K for GTMBA).

However, in the case of ATSR-1 the uncertainties appear to be underestimated, with the observed discrepancies appearing wider than the predicted envelope. This indicates that either the retrieval uncertainties or in situ uncertainties (or both) were underestimated. The drifter uncertainty is assumed to be fixed at 0.2 K; however, in practice the accuracy of drifter SSTs have improved over the years and 0.2 K may be too low for the early 1990s [RD 4]. The ATSR-1 sensor was also affected by the loss of the 3.7 micron channel shortly after launch (this is why the ATSR-1 results show the “D2” retrieval for night as well as day) and problems with the onboard cooler meant the detectors were operated at higher temperatures than intended resulting in higher levels of radiometric noise. This is reflected in the ATSR-1 estimated uncertainties which are larger than the latter two sensors – the majority of ATSR-1 uncertainties are estimated to be in the range 0.2-0.4 K with some uncertainties estimated at over 1 K.

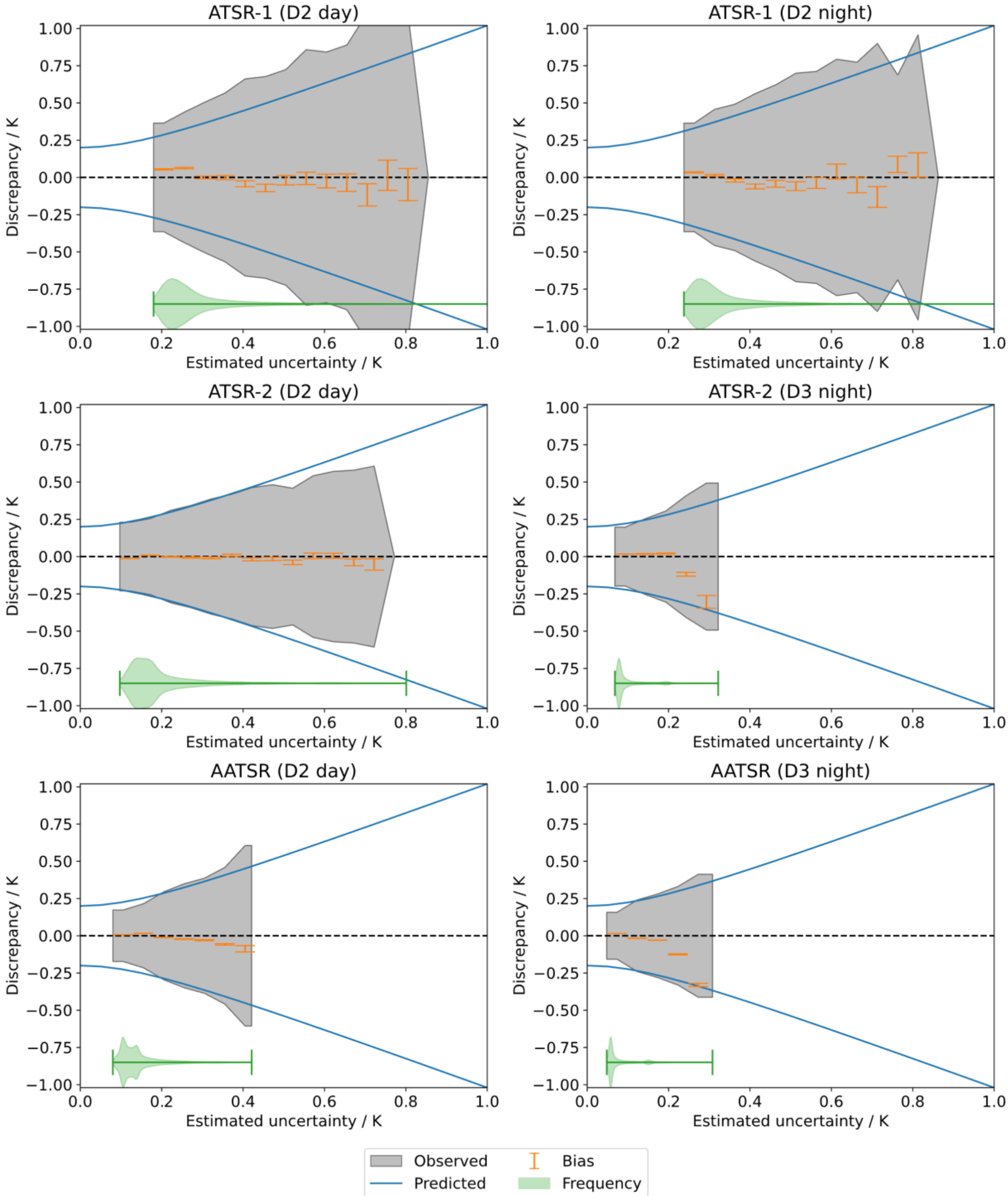


Figure 2: ATSR uncertainty validation against drifter in situ. See Figure 1 for explanation of plots.

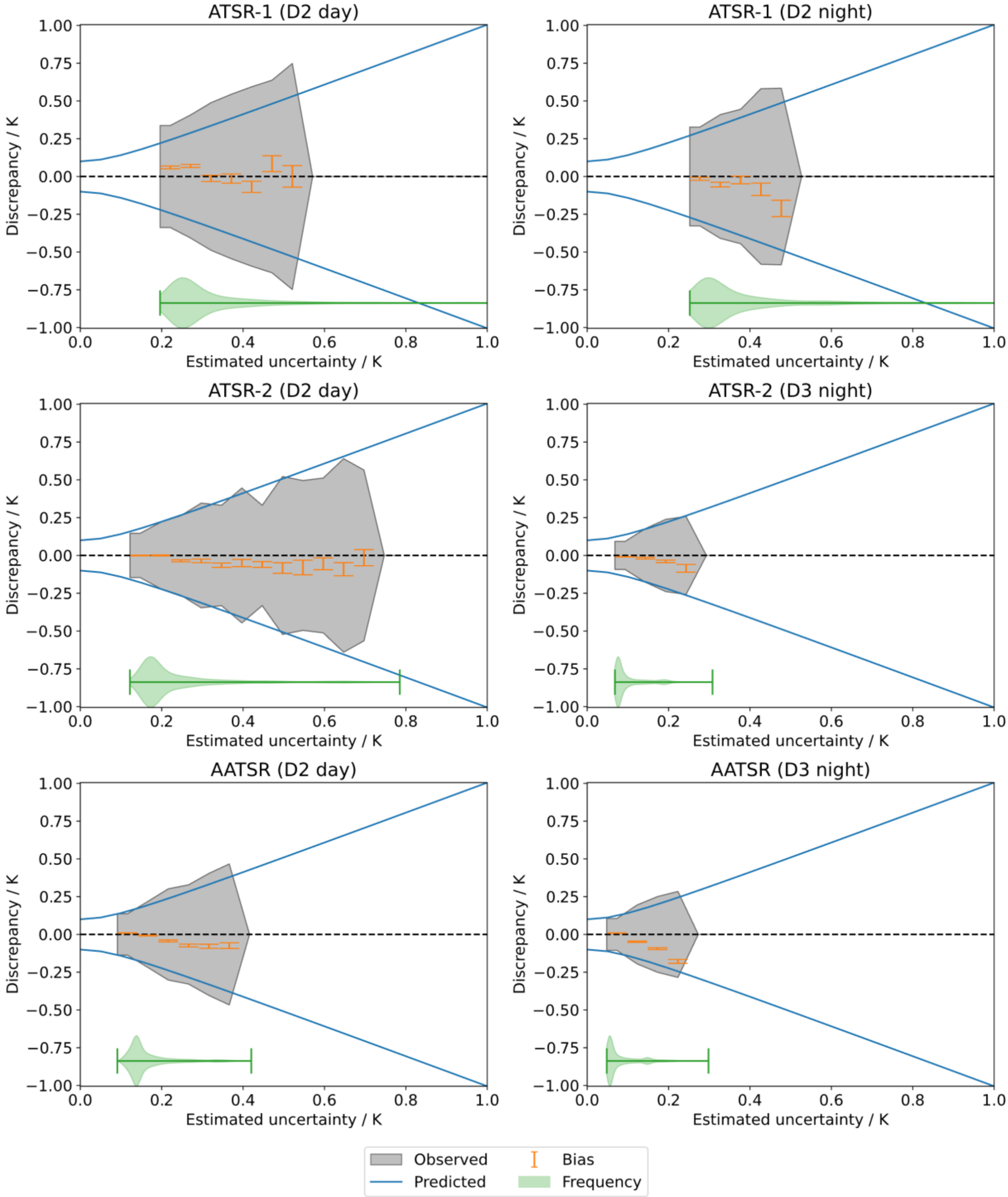


Figure 3: ATSR uncertainty validation against GTMBA in situ. See Figure 1 for explanation of plots.

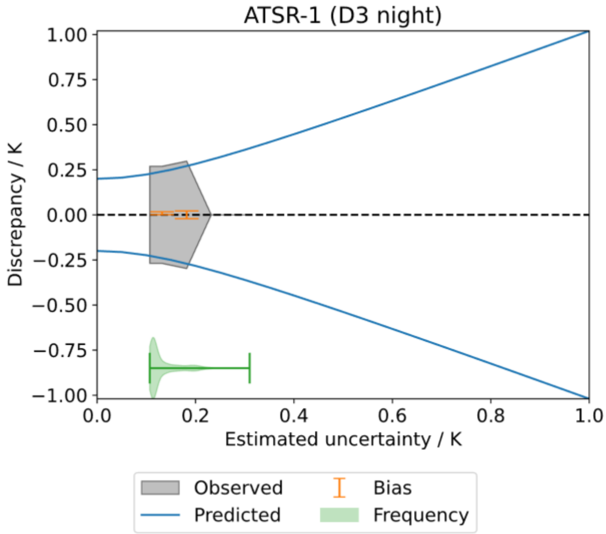


Figure 4: ATSR-1 D3 night uncertainty validation against drifter in situ. See Figure 1 for explanation of plots.

Figure 4 shows the comparison of the ATSR-1 "D3" retrieval against drifters while the 3.7 micron channel was still operating (to May 1992). These results are consistent with the later sensors, indicating that the ATSR-1 retrieval uncertainties are underestimated later in the operation of the sensor.

Comparison of estimated SLSTR uncertainties against drifters is shown in Figure 5 and tropical moorings in Figure 6. The majority of data have low, well-estimated uncertainties $\lesssim 0.2$ (day) and $\lesssim 0.1$ (night) similar to the ATSR sensors. However, there are a small minority of data with higher estimated uncertainties which are under-estimated.

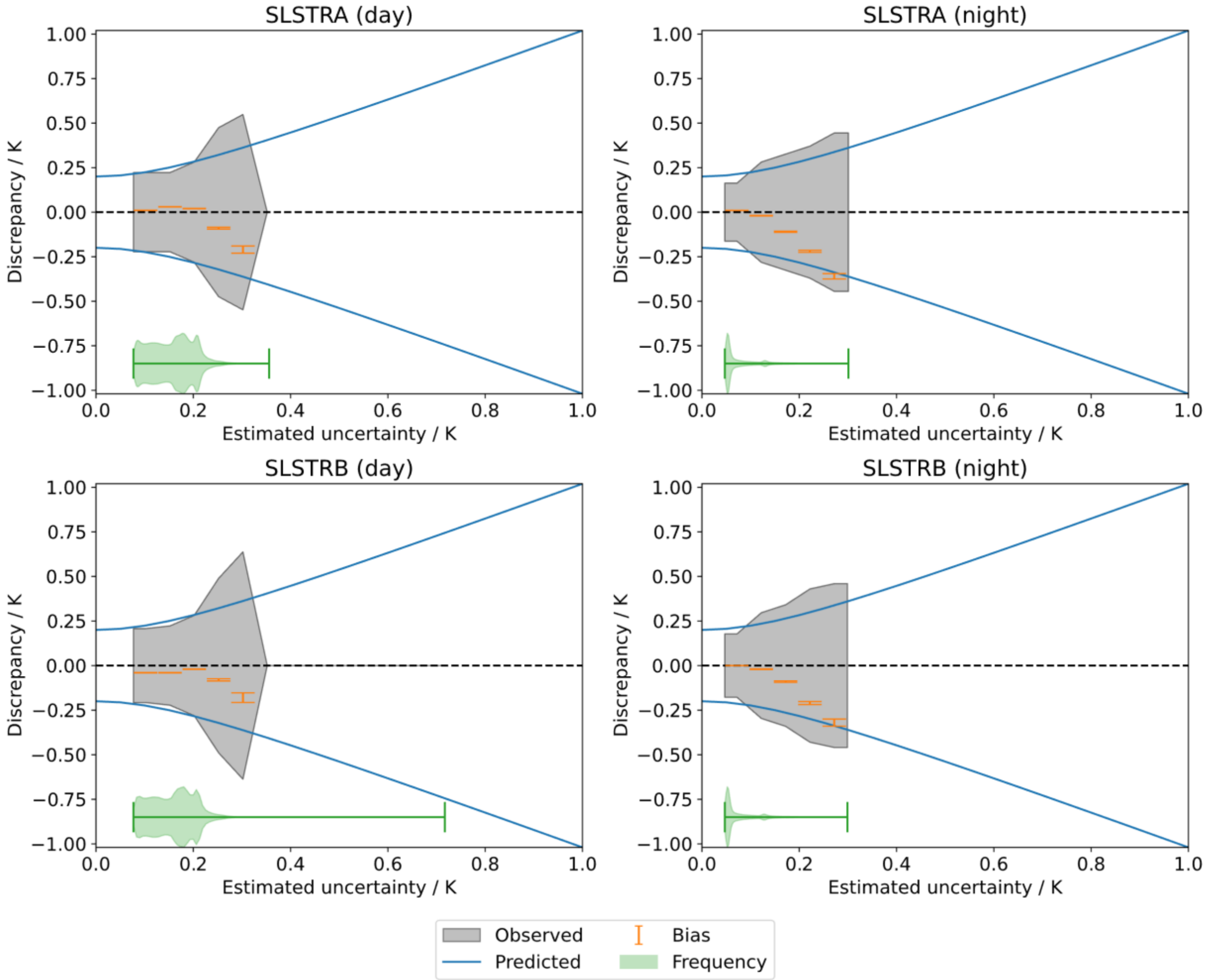


Figure 5: SLSTR uncertainty validation against drifter in situ. See Figure 1 for explanation of plots.

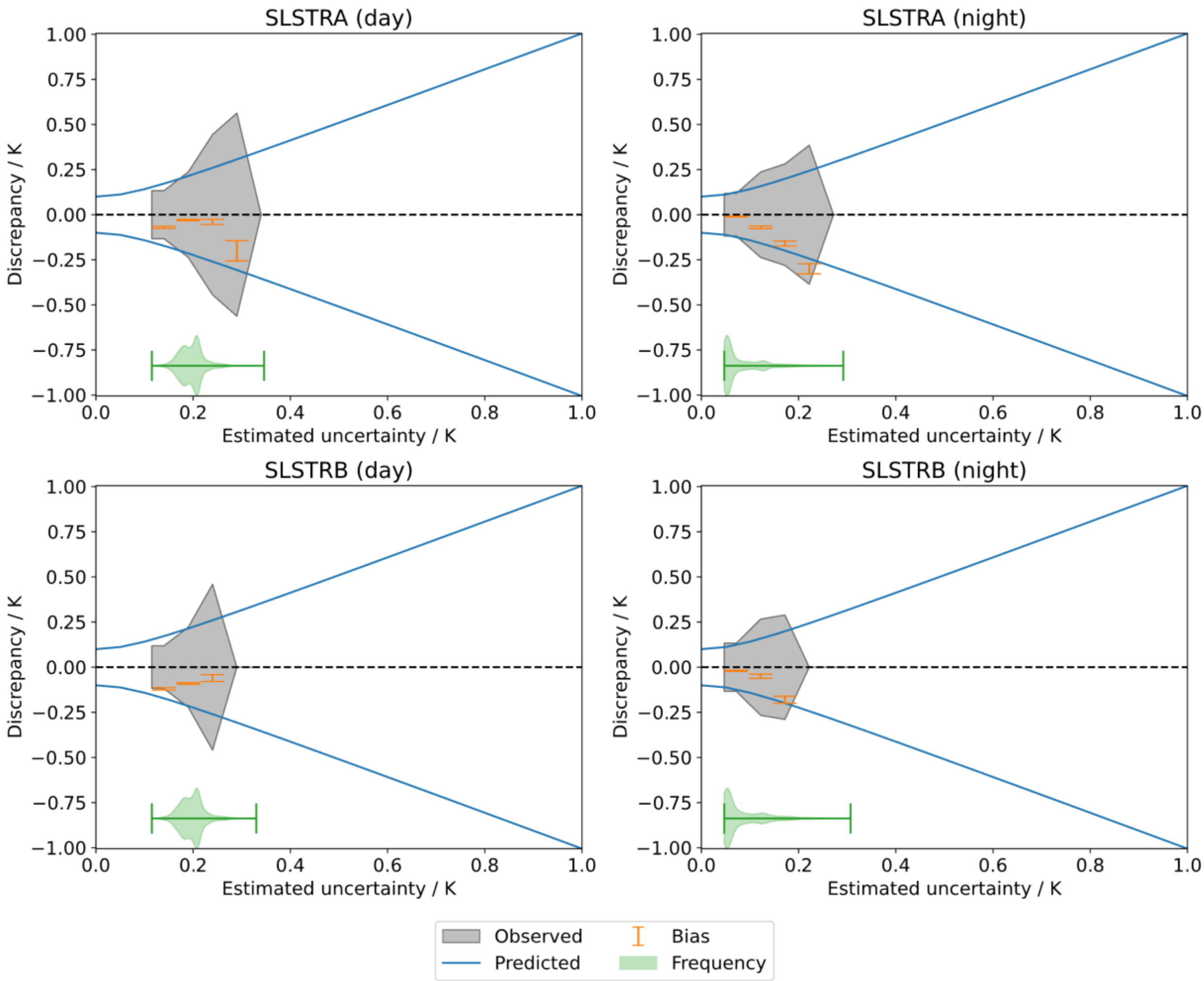


Figure 6: SLSTR uncertainty validation against GTMBA in situ. See Figure 1 for explanation of plots.

2.3 Single-View AVHRR

Comparisons of estimated AVHRR uncertainty against reference and ship in situ for the AVHRR/1 sensors are shown in Figure 7. The original AVHRR/1 design of the sensor, used in the 1980s, lacked a 12 micron channel so they are only used to generate SSTs from night-time observations. Due to the limited availability of in situ SST observations in the 1980s the comparisons are split into two: reference in situ that includes all measurements which are not categorised as ships (i.e., drifters, moorings, CTDs, MBTs, XBTs etc), and ship measurements which form the majority of the 1980s in situ record. While the discrepancies between the satellite and in situ measurements are generally large the uncertainties are well-estimated.

Comparisons for the AVHRR/2 instruments operating in the 1980s are shown in Figure 8 for both day and night comparisons against the reference in situ. Ship-based comparisons are now shown as the uncertainty in the ship measurements is significantly higher than the satellite uncertainties. The night-time uncertainties are well estimated; however, the daytime uncertainties are being overestimated – especially for the later AVHRR-11 instrument.

Further comparisons for the later AVHRR instruments against drifter in situ are shown in Figure 9, Figure 10, and Figure 11. These continue the pattern seen in the early AVHRRs, with night-time uncertainties being well-estimated while the daytime uncertainties are all over estimated. Figure 12 shows the comparison against the more accurate GTMBA in situ for a selection of AVHRRS (12, 15, and 17), where the over estimation of daytime uncertainty is even clearer.

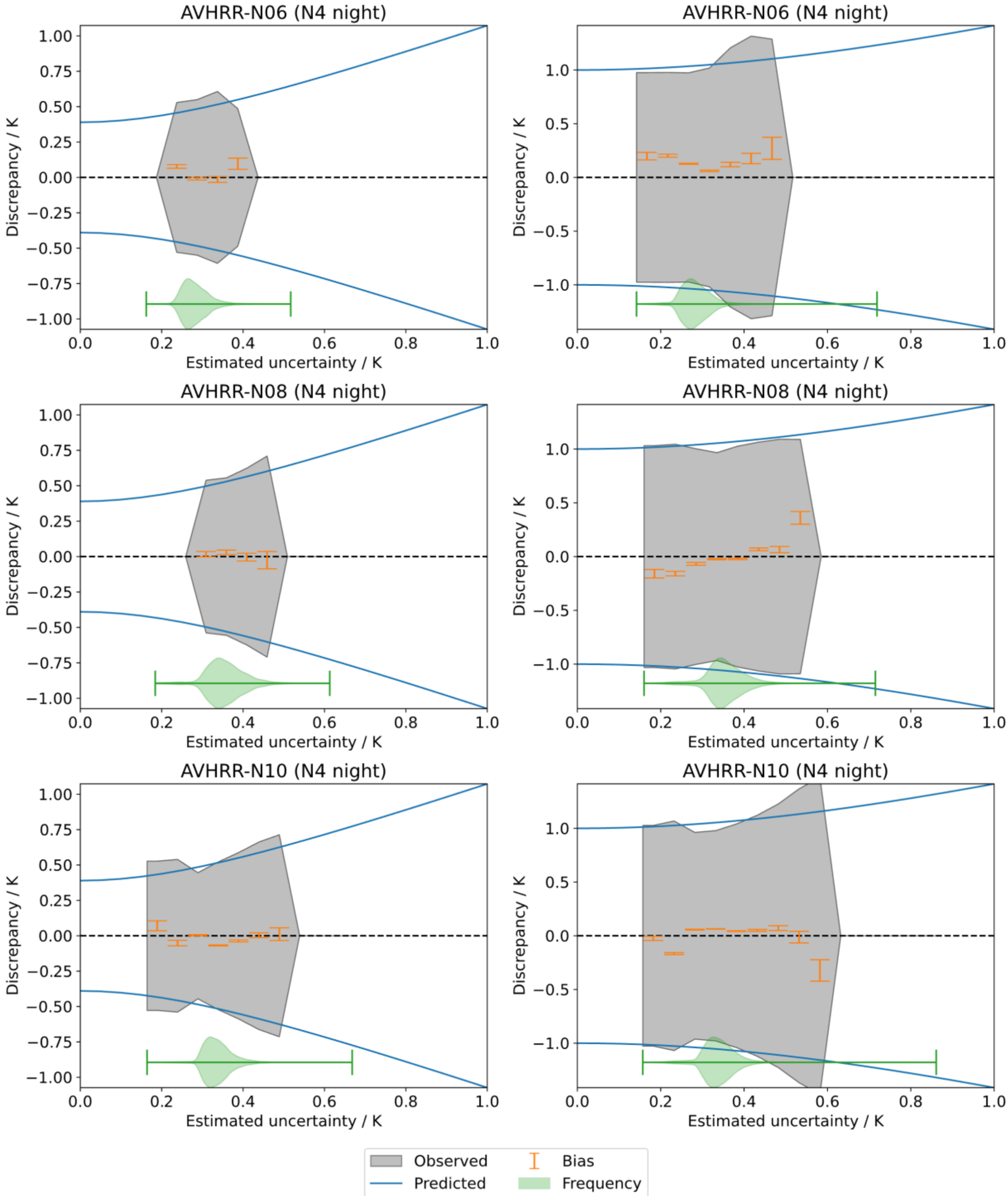


Figure 7: AVHRR/1 uncertainty validation against reference in situ (left) and ship in situ (right). No daytime data are shown as AVHRR/1 sensors only produce night-time SST. See Figure 1 for explanation of plots.

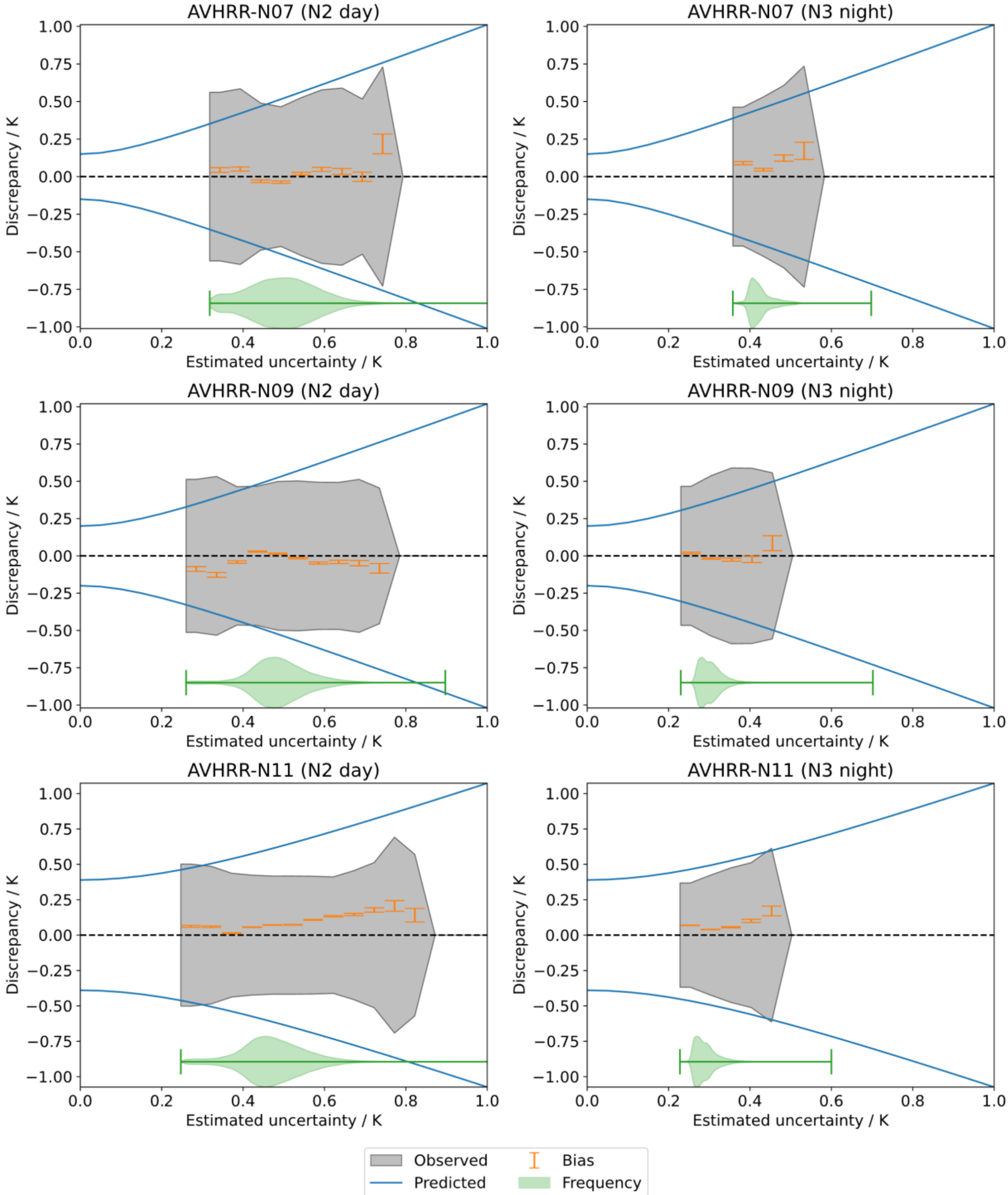


Figure 8: AVHRR uncertainty validation against reference in situ for day (left) and night (right) data. See Figure 1 for explanation of plots.

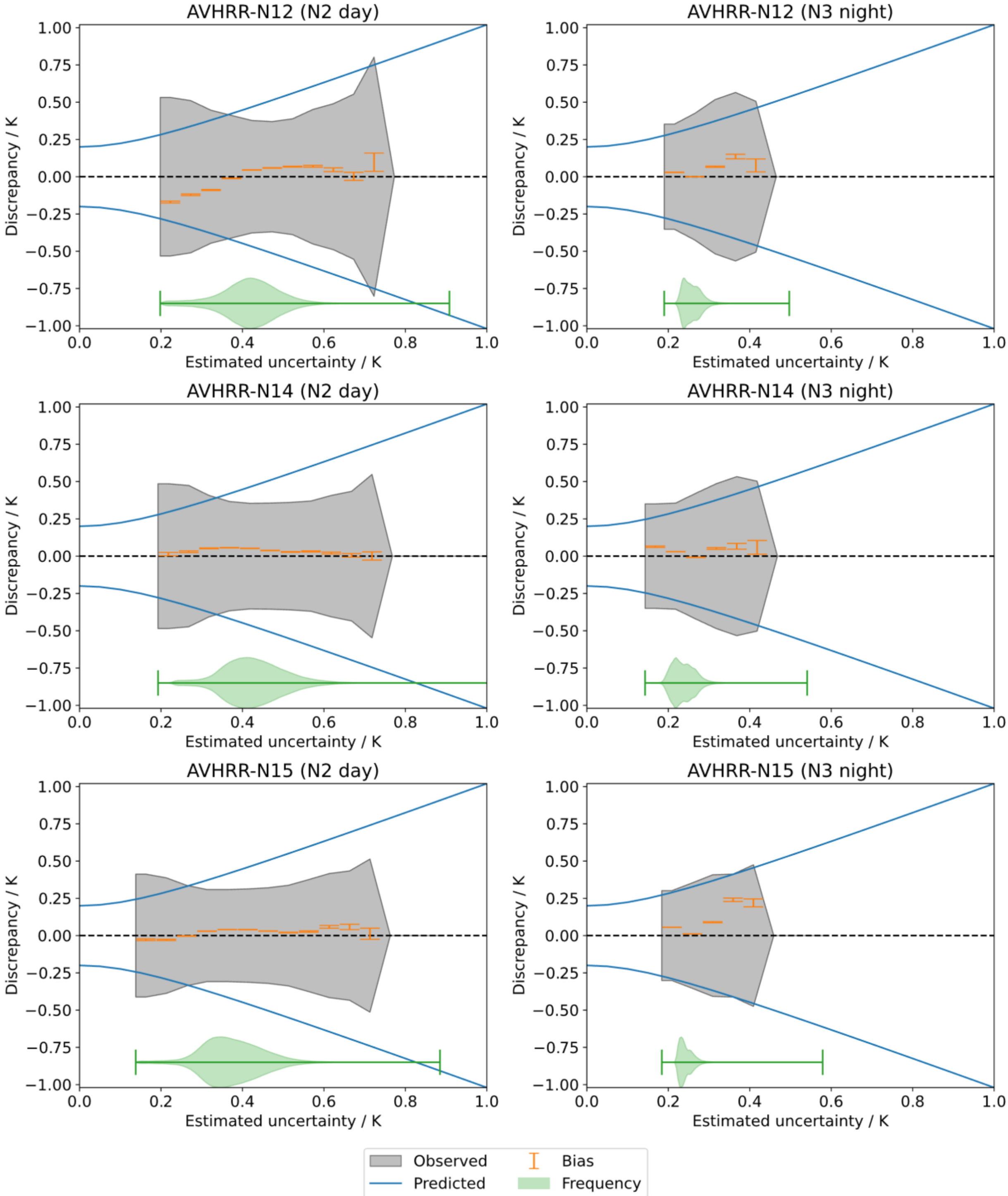


Figure 9: AVHRR uncertainty validation against drifter in situ for day (left) and night (right) data. See Figure 1 for explanation of plots.

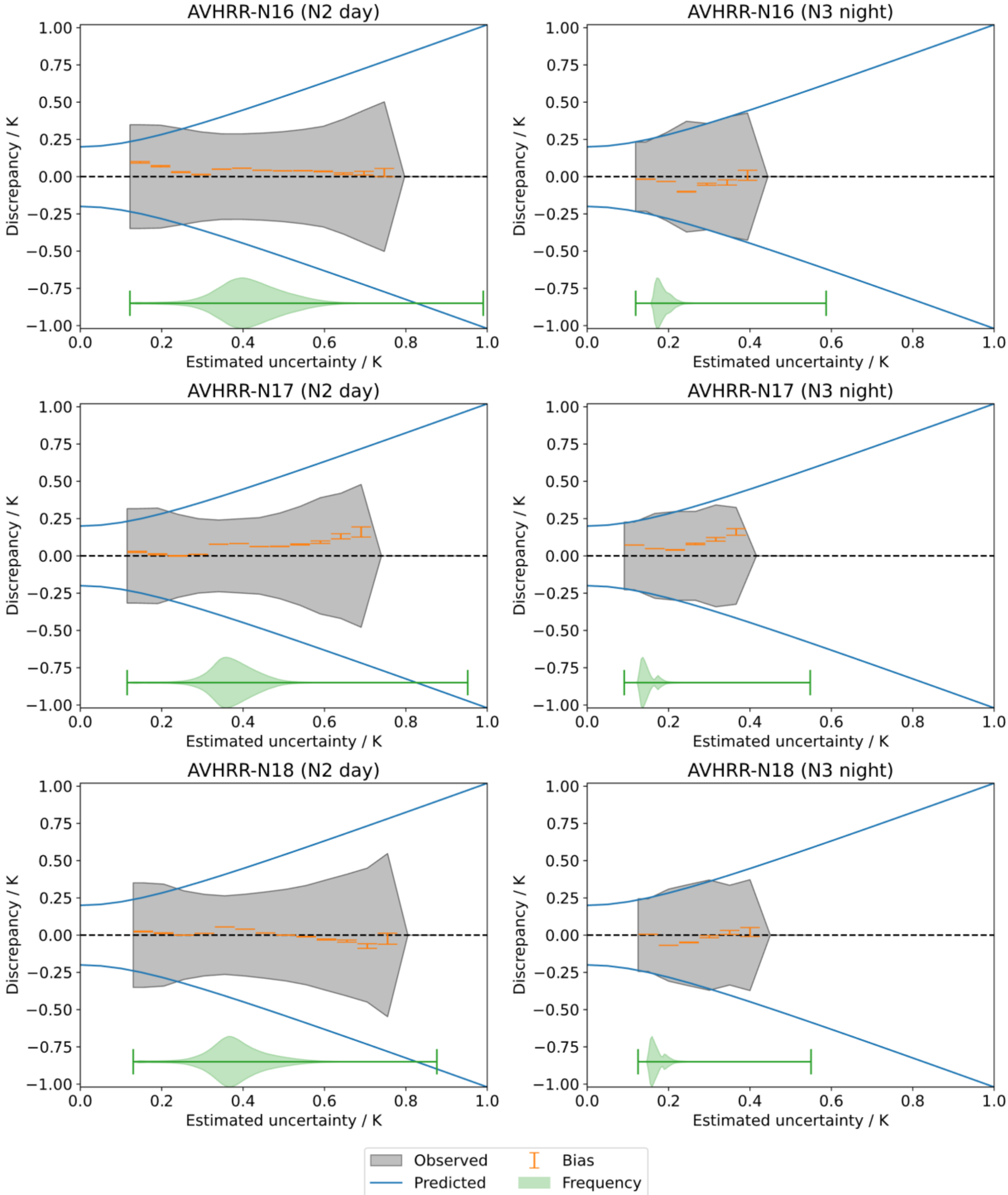


Figure 10: AVHRR uncertainty validation against drifter in situ for day (left) and night (right) data. See Figure 1 for explanation of plots.

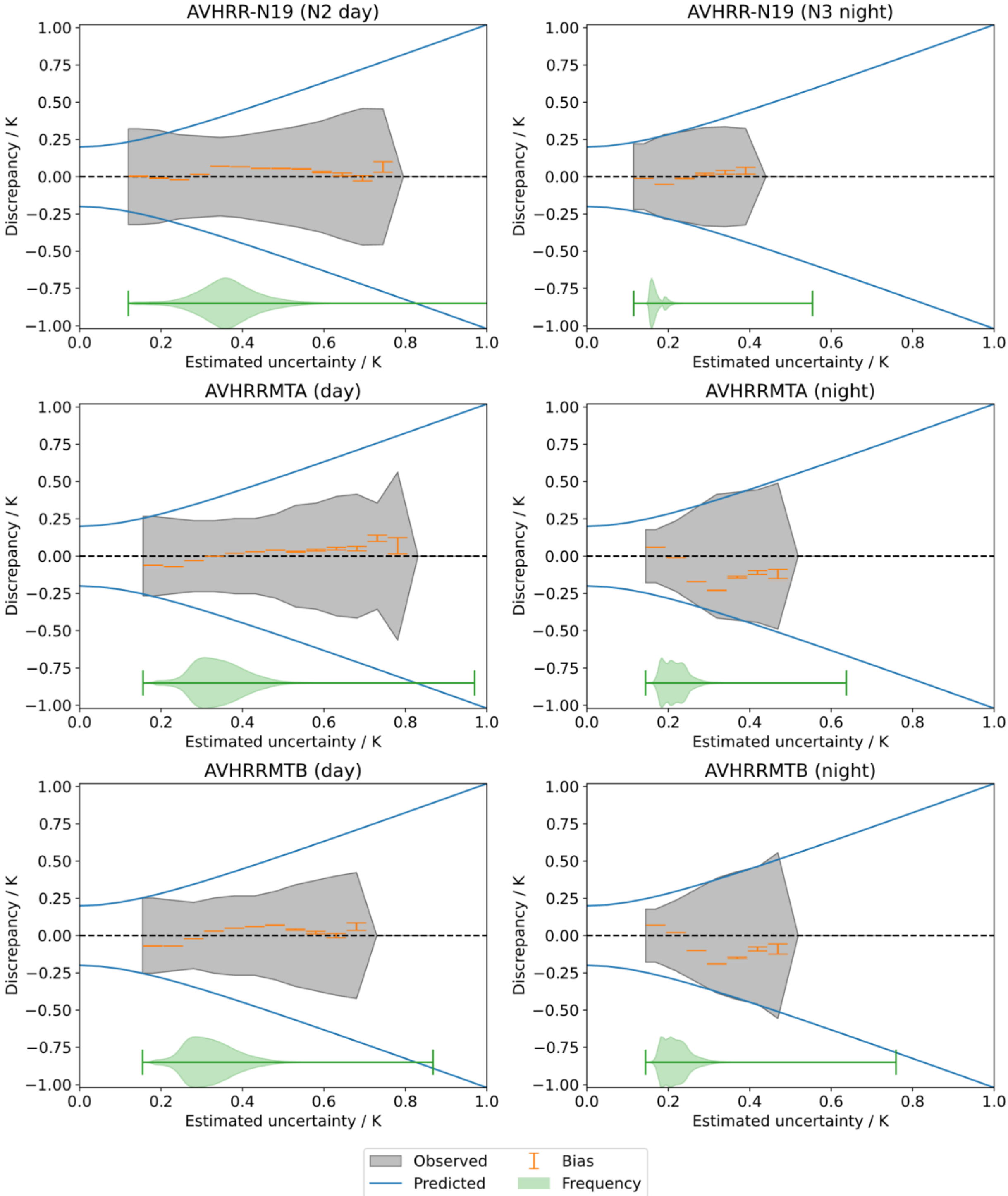


Figure 11: AVHRR uncertainty validation against drifter in situ for day (left) and night (right) data. See Figure 1 for explanation of plots.

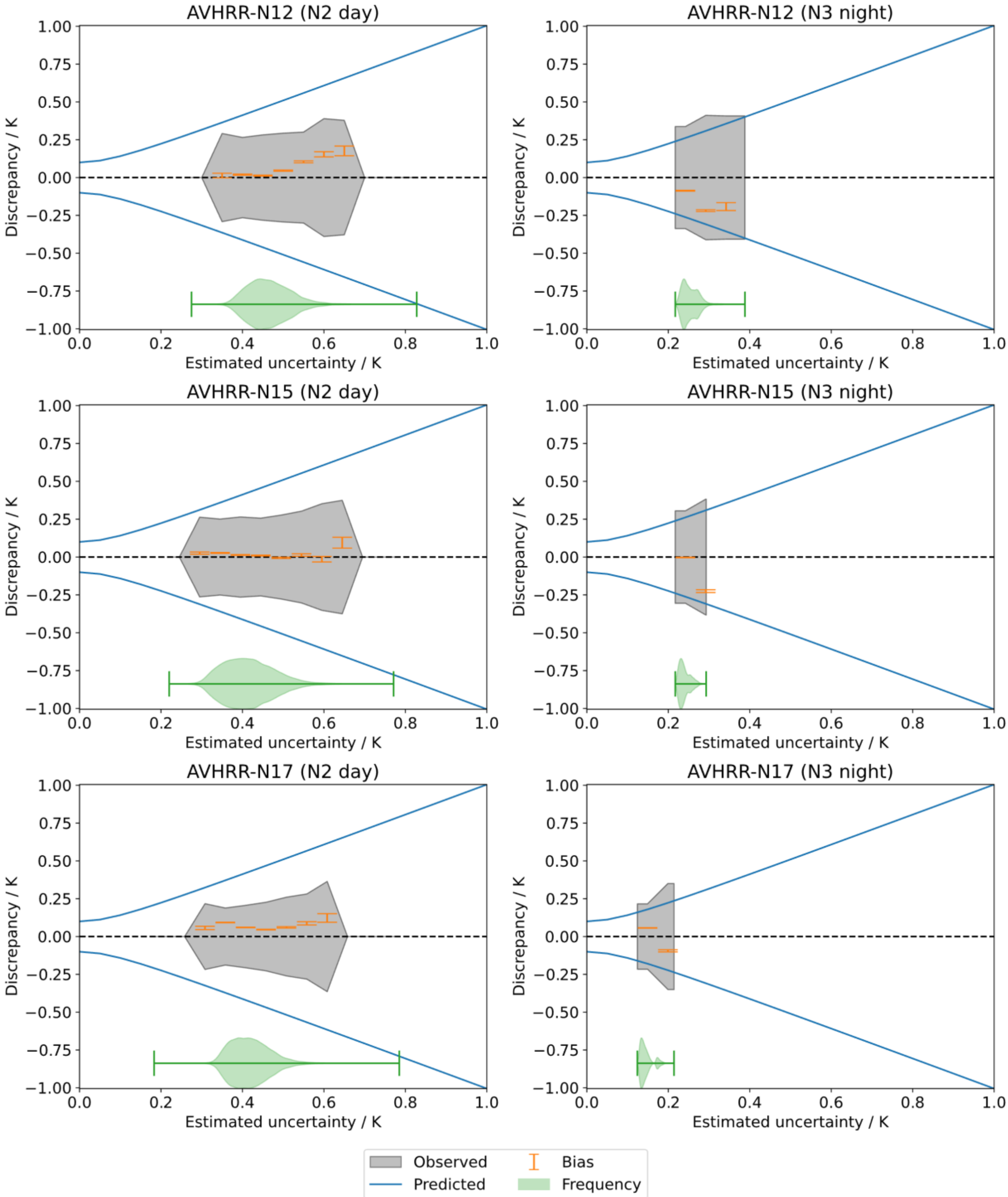


Figure 12: AVHRR uncertainty validation against GTMBA in situ for day (left) and night (right) data. See Figure 1 for explanation of plots.

2.4 Level 4 Analysis

Uncertainty validation plots for the global and coastal Level 4 data against reference in situ and ship are show in Figure 13. In all cases the spread of RSDs is slightly narrower than the expected envelope – indicating that either the estimated Level 4 uncertainties are too high, or the assumed in situ uncertainty is too high. In this case we believe the assumed in situ uncertainty is too high: as the reference data includes a mixture of in situ types we use the median reported uncertainty from the SIRDS dataset which at 0.39 K is significantly higher than 0.2 K assumed for drifter-only datasets.

In the comparison of ocean data to reference in situ (Figure 13 top-left) we see that distribution of estimated uncertainty peaks between 0.2 and 0.3 K with the majority of all data under 0.5 K. With coastal data (Figure 13 lower-left) the estimated uncertainties are noticeably higher, with a peak in the distribution between 0.3 and 0.5 K and a noticeable fraction up to 1.0 – 1.5 K. In both regions, extreme cases with estimated uncertainties over 3 K do exist (mostly in the earlier record), there are insufficient to estimate the RSD or median biases above about 2 – 2.5 K. In the comparison against ships (Figure 13 right column) the distributions are shifted towards slightly higher estimated uncertainties due to the spatial and temporal sampling.

Figure 14 shows the uncertainty validation against reference in situ and ships for each of the four complete decades: 1980s, 1990s, 2000s, 2010s.

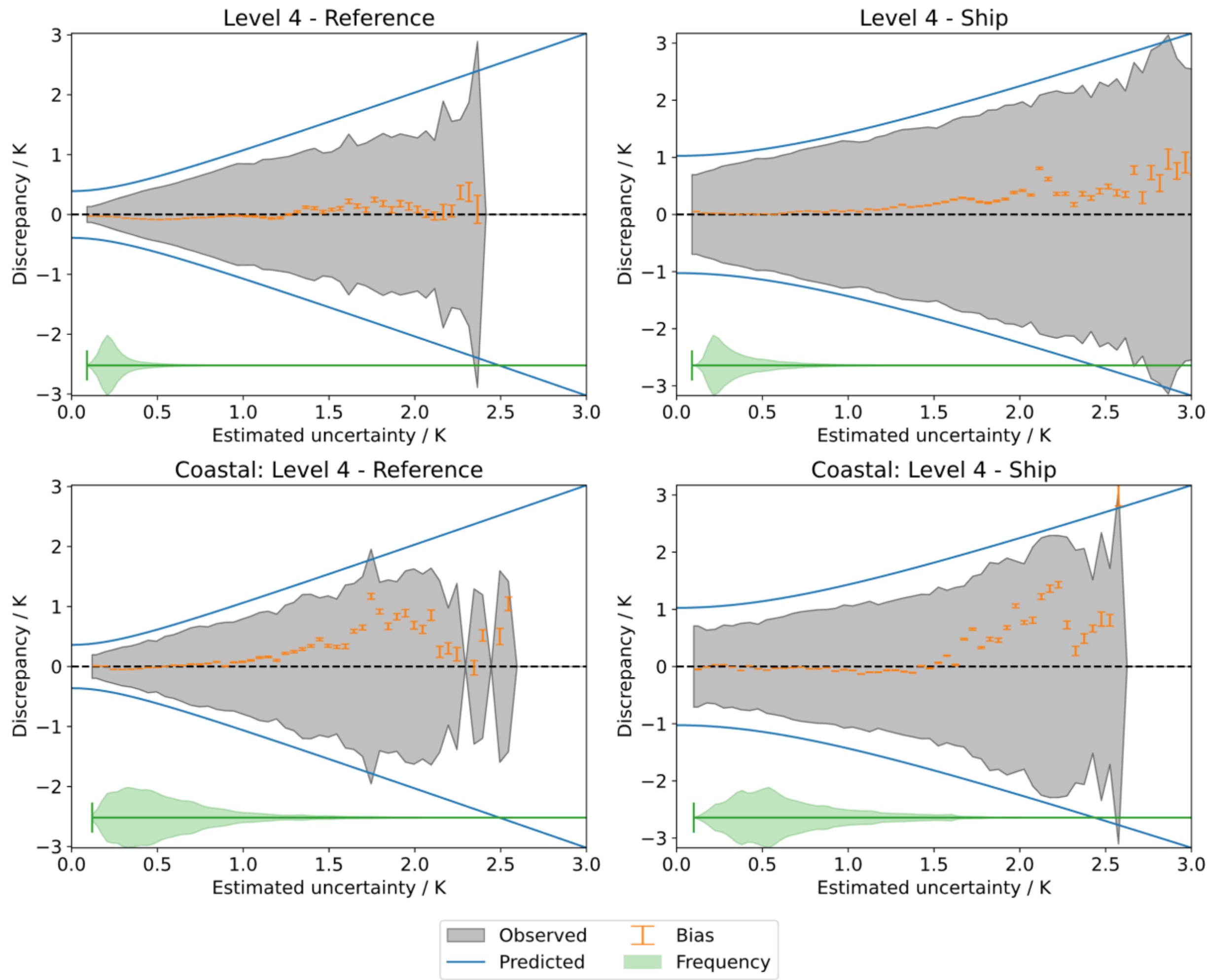


Figure 13: Dependence of the median (error bars) and robust standard deviation (grey shaded area) between L4 SST and in situ SST discrepancies as a function of estimate uncertainty. See Figure 1 for explanation of plots.

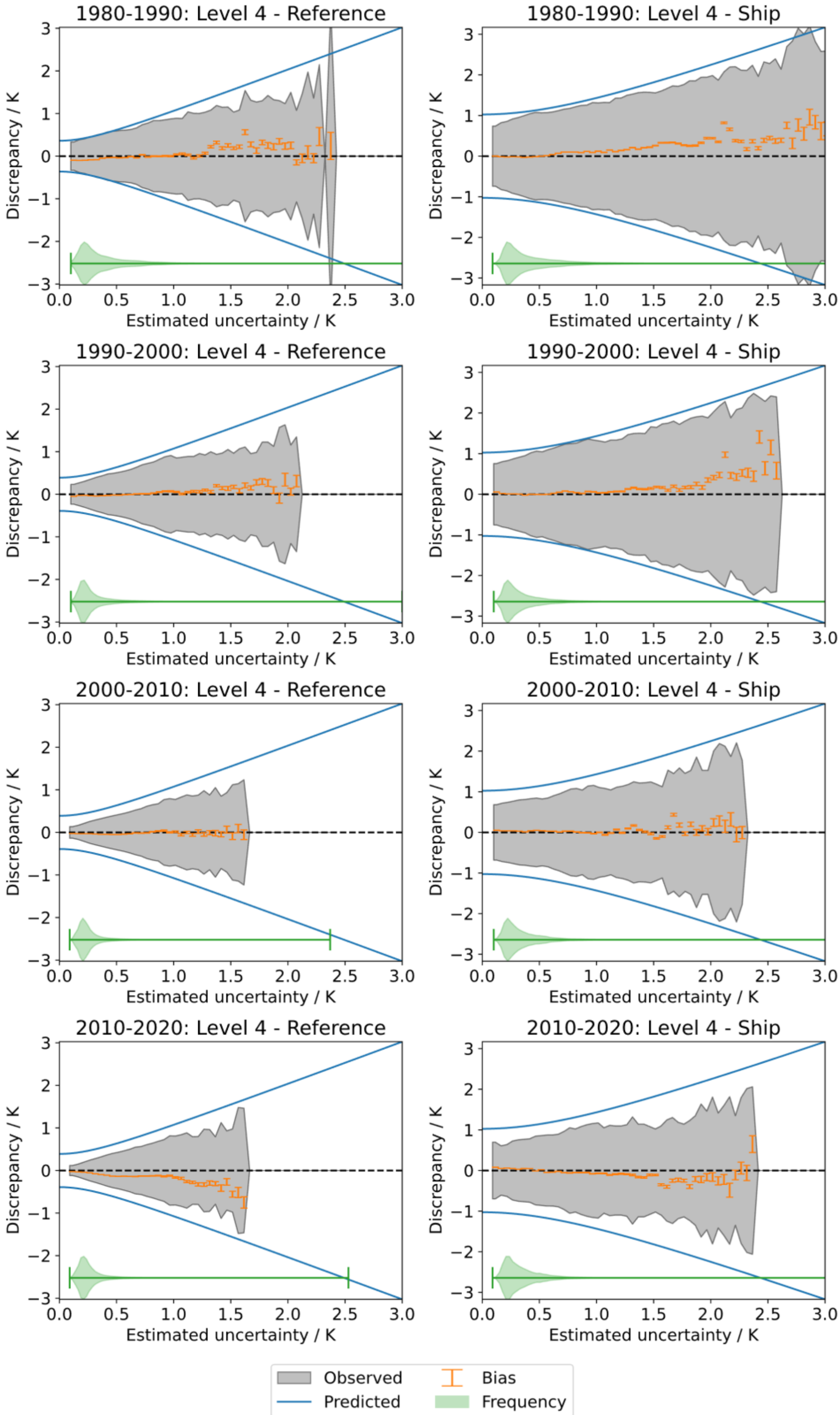


Figure 14: Dependence of the median (error bars) and robust standard deviation (grey shaded area) between L4 SST and in situ SST discrepancies as a function of estimate uncertainty by decade. See Figure 1 for explanation of plots.

3. SUMMARY AND CONCLUSIONS

The key results shown in previous sections are:

- While ATSR-1 uncertainties are underestimated overall by ~30%, the uncertainties for ATSR-2 and AATSR are largely unbiased and capture well the variations in uncertainty for these sensors (§2.2).
- SLSTR A & B uncertainties are well estimated for the vast majority of data, although the largest uncertainties are somewhat underestimated.
- For AVHRRs, some retrieval uncertainties are very well estimated, and in other circumstances, the uncertainty model seems not to predict variations in uncertainty with significant skill (§2.3).
    - Retrievals at night using either 2 channels (3.7 and 11 µm, early AVHRR/1 sensors) or 3 channels (3.7, 11 and 12 µm, all later AVHRR sensors) have well estimated uncertainties in SST.
    - Retrievals in daytime using split-window channels in optimal estimation are not skilful in discriminating more and less uncertain SSTs. In general, uncertainties are over-estimated (the results are pessimistic).
- For the level 4 (gap-filled analysis) product, more and less certain data are well distinguished. The majority of data have quantitatively well estimated uncertainties, although in general these are pessimistic by about 20% to 30% (§2.4).

The uncertainty results are satisfactory overall.

The exception is for SSTs obtained by single-view split-window retrieval by OE (used for daytime retrieval). The developments in uncertainty estimation for OE have not provided good results for this category of retrieval, despite satisfactory results for night-time observations. This is difficult to understand, as the same theory and code-based are applied day and night. Clearly, further investigation of this point is required.