



# ATSR V3.0

## Sea Surface Temperature Validation Report

Technical note for the ATSR QWG

Written by: Gary Corlett .....

Approved by: Hugh Kelliher .....



## TABLE OF CONTENTS

<b>EXECUTIVE SUMMARY</b>	<b>5</b>
<b>1 Introduction</b>	<b>6</b>
<b>2 Validation data</b>	<b>7</b>
<b>3 Validation Methodology</b>	<b>8</b>
<b>4 Results</b>	<b>9</b>
4.1 ATSR NR data	9
4.2 Results – L2P	11
<b>5 Summary</b>	<b>13</b>
<b>6 References</b>	<b>15</b>
<b>7 APPENDIX 1 – Detailed AATSR Validation Results</b>	<b>16</b>
7.1 Comparisons between ATS_NR SSTs and drifting buoys	17
7.2 Comparisons between ATS_NR SSTs and Argo	18
7.3 Comparisons between ATS_NR SSTs and the GTMBA	19
7.4 Comparisons between ATS_NR SSTs and radiometers	20
7.5 Statistical analysis of ATS_NR results	21
7.6 Comparisons between ATS_L2P SSTs and drifting buoys	22
7.7 Comparisons between ATS_L2P SSTs and Argo	23
7.8 Comparisons between ATS_L2P SSTs and the GTMBA	24
7.9 Comparisons between ATS_L2P SSTs and radiometers	25
7.10 Statistical analysis of ATS_L2P results	26
<b>8 APPENDIX 2 – Detailed ATSR-2 Validation Results</b>	<b>27</b>
8.1 Comparisons between AT2_NR SSTs and drifting buoys	28
8.2 Comparisons between AT2_NR SSTs and the GTMBA	29
8.3 Comparisons between AT2_NR SSTs and radiometers	30
8.4 Statistical analysis of AT2_NR results	31
8.5 Comparisons between AT2_L2P SSTs and drifting buoys	32
8.6 Comparisons between AT2_L2P SSTs and the GTMBA	33
8.7 Comparisons between AT2_L2P SSTs and radiometers	34
8.8 Statistical analysis of AT2_L2P results	35
<b>9 APPENDIX 3 – Detailed ATSR-1 Validation Results</b>	<b>36</b>
9.1 Comparisons between AT1_NR SSTs and drifting buoys	37
9.2 Comparisons between AT1_NR SSTs and GTMBA	38



---

<b>9.3</b>	<b>Statistical analysis of AT1_NR results</b>	<b>39</b>
<b>9.4</b>	<b>Comparisons between AT1_L2P SSTs and drifting buoys</b>	<b>40</b>
<b>9.5</b>	<b>Comparisons between AT1_L2P SSTs and GTMBA</b>	<b>41</b>
<b>9.6</b>	<b>Statistical analysis of AT1_L2P results</b>	<b>42</b>



**Document Control**

<b>Version Number:</b>	<b>Date of issue:</b>	<b>Comment:</b>	<b>Issued to:</b>
1A	6 <sup>th</sup> March 2015	First Issue	All
1B	7 <sup>th</sup> September 2016	Updated with post-June 2003 ATSR-2	All



## EXECUTIVE SUMMARY

This report summarises validation of SST products from the ATSR V3.0 dataset. Two different SST products are available (1) NR and (2) L2P. In general, global biases are very similar for AATSR and ATSR-2 in both the NR and L2P products, with both being slightly warm but within specification ( $< 0.3$  K). ATSR-1 L2P data exhibit a slightly cooler global bias compared to NR data but still within specification. In all cases the robust standard deviation is lower for the L2P product due to a combination of reduced regional biases and improved cloud masking.

The following conclusions are drawn from the results presented in this report:

- AATSR and ATSR-2 SST dual-view retrievals are  $< 0.15$  K for both NR and L2P products.
- Regional variations exist particularly for known retrieval issues such as water vapour affecting N2 retrievals and tropospheric aerosol affecting N2 and N3 retrievals.
- The L2P product has reduced regional biases and improved cloud clearing.
- ATSR-1 dual-view two-channel retrievals are  $< 0.25$  K for both the NR and L2P products
- ATSR-1 NR nadir-view retrievals are significantly affected by aerosol dust during the period around Mount Pinatubo (from 1991 to 1993) and should not be used for quantitative analysis.

It is recommended that:

1. The ATSR V3.0 SST dataset be released
2. That users should use the L2P product rather than the NR product



## 1 INTRODUCTION

This report presents an assessment of sea surface temperature (SST) data quality of the V3.0 Along-Track Scanning Radiometer (ATSR) dataset, covering data from the ATSR-1, ATSR-2 and Advanced ATSR (AATSR) series of sensors. The assessment is carried through comparisons to validation data taken *in situ* or from ships of opportunity and research vessels. In the V3.0 ATSR dataset there are two different SST products that require validation. These are:

- NR  
These products contain SSTs generated using latitudinally banded retrieval coefficients provided in ENVISAT format files with SADIST cloud masking.
- L2P  
These products contain SSTs generated using total column water vapour banded retrieval coefficients provided in the Group for High Resolution SST (GHRSSST) L2P format files with Bayesian cloud masking.

The ATSR series of instruments have three spectral bands that are used for SST retrieval, with nominal band centres at 3.7  $\mu\text{m}$ , 11  $\mu\text{m}$  and 12  $\mu\text{m}$ . During the day the 3.7  $\mu\text{m}$  channel is not used due to solar contamination and so, as each Earth scene is viewed in both nadir and off-nadir, there are four possible retrieved SSTs, referred to as N2 (nadir-only 11  $\mu\text{m}$  and 12  $\mu\text{m}$ ), N3 (nadir-only 3.7  $\mu\text{m}$ , 11  $\mu\text{m}$  and 12  $\mu\text{m}$ ), D2 (dual-view 11  $\mu\text{m}$  and 12  $\mu\text{m}$ ) and D3 (dual-view 3.7  $\mu\text{m}$ , 11  $\mu\text{m}$  and 12  $\mu\text{m}$ ). The N2, N3, D2 and D3 nomenclature is used through the rest of this report. Note: The L2P product only contains the best SST available and does not therefore contain nadir-view only SSTs.

The retrieval scheme itself uses linear regression algorithms with coefficients derived from radiative transfer models (see for example Závody et al., 1995, Merchant et al., 1999) that perform a linear regression of SST to simulated brightness temperatures (BTs) at the nominal band centres. Significant improvements to all ATSR retrievals have been demonstrated as part of the ATSR Reprocessing for Climate (ARC) project (Embury et al., 2012; Merchant et al., 2012). For the V3.0 dataset the NR SST retrieval coefficients were updated to use the physical basis and forward modelling used within ARC. For further details on the updated retrieval coefficients the reader is referred to Embury and Corlett (2010). For the L2P product, the full ARC processing scheme was used at the native resolution of the ATSRs, which is roughly 1km; the ARC scheme was developed to generate 0.1degree gridded data.

The two SST products use different methods for screening clouds from the BTs. The NR product uses the SADIST scheme originally developed for ATSR-1, which is based on a set of threshold and histogram tests to a set of pre-computed thresholds. The L2P product uses a method of clear-sky identification resulting from applying Bayes theorem to simulated clear-sky BTs output from a radiative transfer model.



## 2 VALIDATION DATA

The primary validation dataset for SST<sub>skin</sub> measurements provided by Infra-red (IR) satellite instruments such as the ATSRs are ship-borne radiometers, which also have the advantage of being completely traceable to agreed SI standards through national metrology institutes. However, the coverage and availability of ship-borne radiometric data is not ideal, which limits any long-term drift assessment and does not allow for a global assessment of the SST<sub>skin</sub> data quality. Other potential reference data include surface drifting buoys, Argo floats, moored buoys, and conventional ship measurements from engine room intakes or hull-mounted sensors. Although the uncertainty of these other datasets is not always traceable to an SI temperature standard, they are important datasets for satellite SST validation due to their significantly improved global coverage compared to other potential reference datasets, particularly for drifting buoys. In addition, these other datasets provide SST at depth and require adjustment to skin to be used correctly for SST validation. However, under certain conditions (wind speed > 6 ms<sup>-1</sup>) a constant offset of around -0.17 K is expected. No adjustments are made in the analysis presented in this report and so calculated differences are expected to be around zero for comparisons to radiometers and around -0.17 K for comparison to drifters, moorings and the Argo array. A summary of all validation data used in the analysis presented in this report is given in Table 2-1.

Data type	Time period	Coverage	Depth
Ship-borne IR radiometers	1998 - 2012	Various single cruise tracks plus long-term deployments in Caribbean Sea and Bay of Biscay	SST <sub>skin</sub>
Argo floats	2000 – 2012	Global	SST <sub>depth</sub>
GT MBA	1991 – 2012	Tropics	SST <sub>depth</sub>
Drifting buoys	1991-2012	Global	SST <sub>depth</sub>

**Table 2-1:** Content of validation dataset

All available ship-borne radiometer data were acquired from the instrument Principle Investigators (PI). Data was used from the Calibrated Infrared In situ Measurement System (CIRIMS), Infrared SST Autonomous Radiometer (ISAR), Scanning Infrared Sea Surface Temperature Radiometer (SISTeR) and Marine-Atmosphere Emitted Radiance Interferometer (M-AERI) radiometers. No additional quality control (QC) was applied to the radiometer data at this point. Drifting buoy and moored buoy data were provided by the Met Office Hadley Centre (MOHC) taken originally from the National Centres for Environmental Prediction (NCEP) near-real-time (NRT) marine reports (see



<http://icoads.noaa.gov/nrt.html>). Only moored buoys that form part of the Global Tropical Moored Buoy Array (GT MBA) are used. The drifter and GT MBA data are quality controlled by MOHC and no further QC was done. Argo data are extracted from the MOHC EN3 dataset. The nearest to the surface measurement passing the MOHC quality tests between 3.5 m and 5.5 m in depth was used and no further QC was done.

### 3 VALIDATION METHODOLOGY

Match-ups between ATSR and the various validation data were found to the nearest ATSR pixel within  $\pm 3$  hours of the satellite overpass, and stored within a match-up dataset (MD) for subsequent statistical and graphical analysis. The same analysis software was used to match ATSR to all validation data in order to remove the possibility of errors due to inconsistent analysis.

Although the reference data have been subject to some basic quality control (QC) by the data provider, instances of poor data quality can cause outliers in any subsequent validation analysis. Outliers can also be due to the spatial and temporal miss-match, and of course issues with ATSR data quality. To deal with a small number of expected outliers, robust statistics are used in the analysis.

Discrepancies and robust standard deviations are calculated for each type of validation data. Time series of discrepancy and robust standard deviation are provided, as well as the dependence of key instrument and retrieval parameters. Spatial variations are also presented.

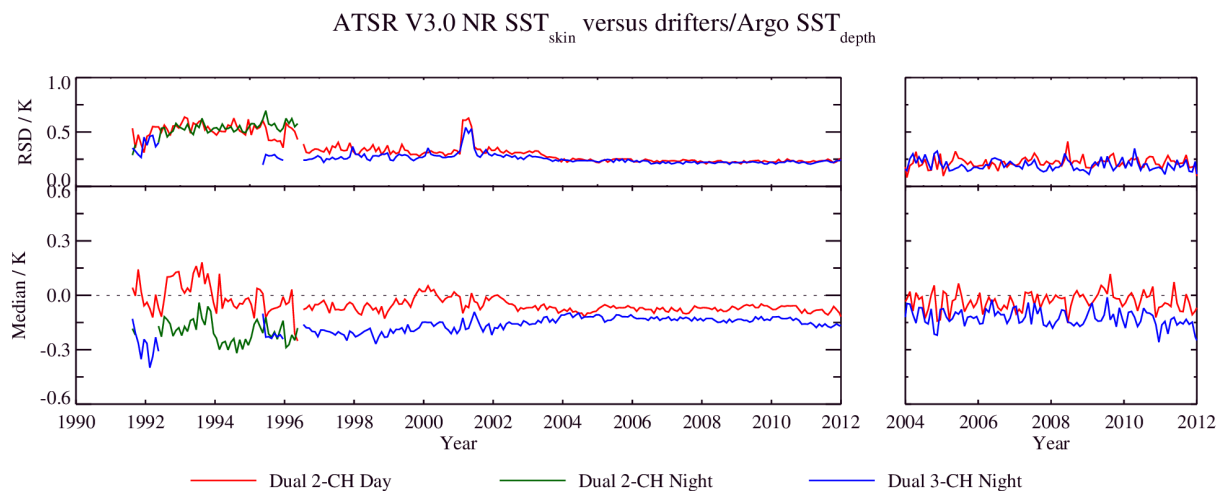
As the L2P product uses the ARC processing scheme at 1km, the data are expected to be noisier than their NR equivalent. This is due to there being no atmospheric correction smoothing in the L2P product compared to the NR product, which reduces the impact of radiometric noise from the instrument on the retrieved SSTs. Consequently, the validation presented here uses the same approach as Embury et al. (2012) whereby 5 by 5 pixel averages centred on the match-up location are used instead of just the central pixel itself.



## 4 RESULTS

### 4.1 ATSR NR data

A time series of all ATSR NR data compared to drifting buoys and Argo (for the period it is available) is shown in Figure 4-1. Results are shown for dual-view retrievals only.



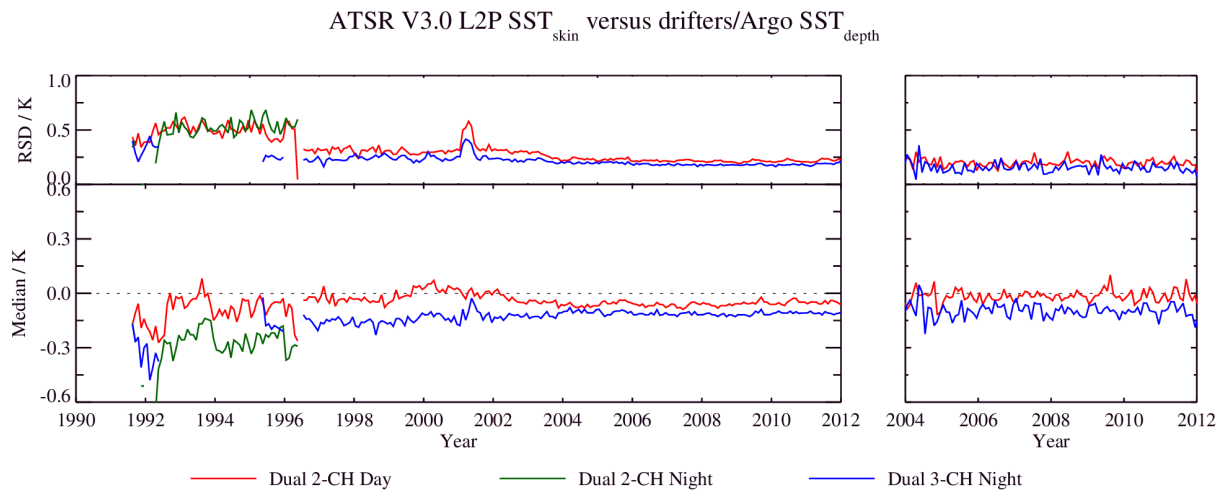
**Figure 4-1:** (Left) Time series of (lower) median discrepancy and (upper) robust standard deviation (RSD) for the ATSR V3.0 NR dataset compared to drifting buoys. Results are shown for daytime dual-view 2-channel (red), nighttime dual-view 2-channel (green) and dual-view 3-channel nighttime (blue) match-ups. Also, shown (right) is the equivalent time series for comparisons to Argo. Each point in the time series has at least 30 match-ups.

The time series in Figure 4-1 has several notable features. First, there are at least three clear regimes, 1991 to 1996, 1996 to 2005 and 2005 to 2012 in both the median and RSD. The first period corresponds to the ATSR-1 mission but the step change in 2005 does not correspond to any instrument changes and is likely driven by changes in the drifting buoy network. The RSD shows a notable increase, particularly in the early years of the ATSR-1 period and this is due to an increase in ATSR-1 retrieval noise because of the increasing 12 micron detector temperature over the mission lifetime; no equivalent pattern is seen in the median results. A notable increase and then decrease in RSD is seen in early 2001 corresponding to the loss of the final ERS-2 gyro. The reduction in noise to the pre-gyro-failure level indicates that the additional corrections used to geolocate the data were successful. Small changes in the median are also noticeable during this period. A notable day/night difference is seen in the median but not in the RSD showing the consistency of the retrievals between day and night, with the difference in median being due to diurnal warming in the daytime match-ups, which has not been reduced in this analysis.



## 4.2 Results – L2P

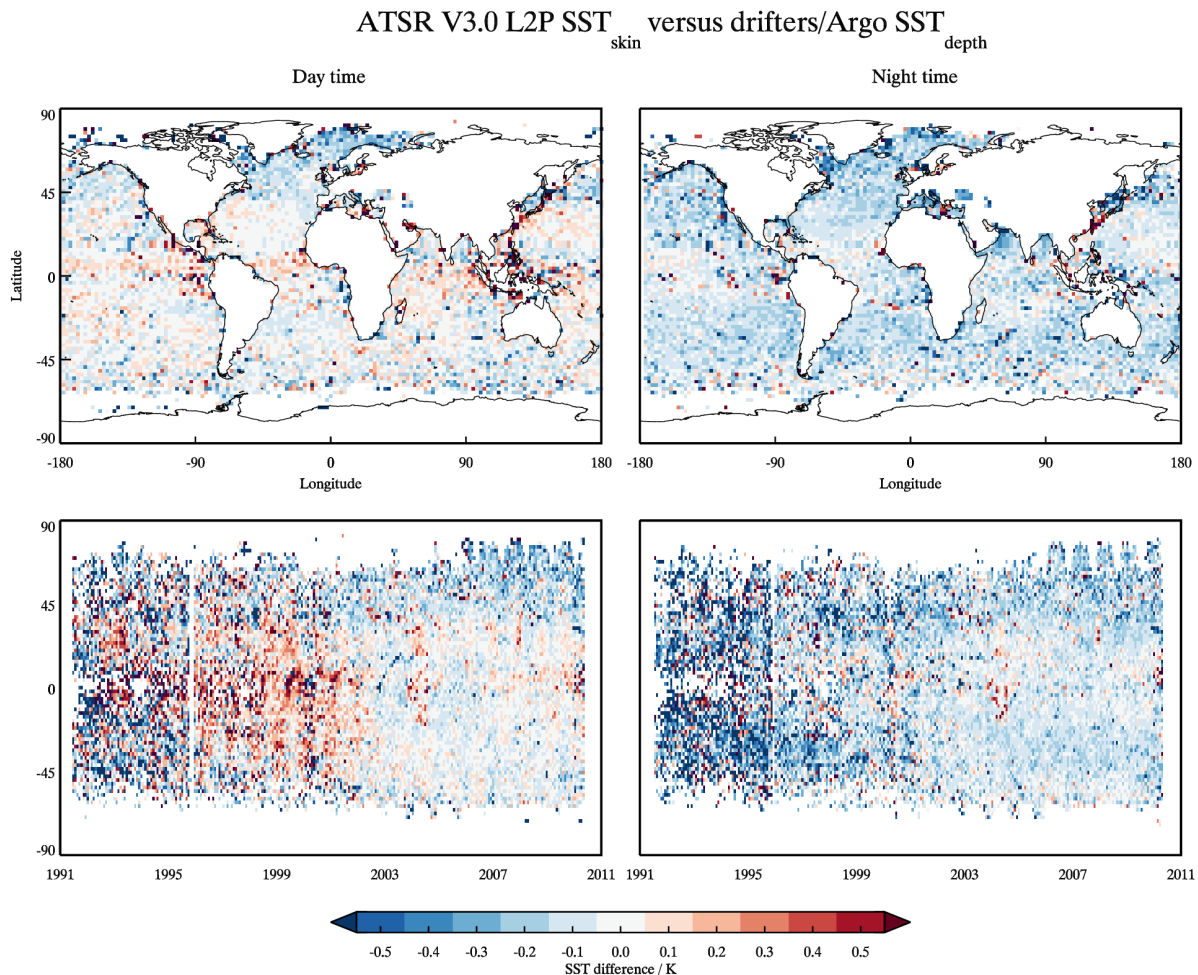
A time series of the ATSR V3.0 L2P dataset compared to the drifting buoys and Argo dataset (for the period it is available) is shown in Figure 4-3. As for the NR case, results are only shown for dual-view retrievals.



**Figure 4-3:** (Left) Time series of (lower) median discrepancy and (upper) robust standard deviation (RSD) for the ATSR V3.0 L2P dataset compared to drifting buoys. Results are shown for daytime dual-view 2-channel (red), nighttime dual-view 2-channel (green) and dual-view 3-channel nighttime (blue) match-ups. Also, shown (right) is the equivalent time series for comparisons to Argo. Each point in the time series has at least 30 match-ups.

The time series in Figure 4-3 is very similar to that shown in Figure 4-1 and only minor differences are evident. First, the median during the ATSR-1 period is more negative in sign indicating the ATSR-1 L2P data are slightly cooler than those in the NR products. Second, the median during the ATSR-1 and AATSR periods is more positive, indicating the data are slightly warmer during this time than in the NR products. These two features are consistent between day and night. Third, the RSD, particularly for ATSR-2 and AATSR, is not as consistent as it was for the NR data and is slightly noisier during daytime.

The spatial distribution of the discrepancies for the ATSR V3.0 L2P dataset is shown in Figure 4-4, which includes the latitude/longitude variation and time/latitude variation for both daytime and nighttime.



**Figure 4-4:** (Upper) Latitude/longitude variation of the median discrepancy for the AATSR V3.0 L2P dataset compared to drifting buoys for (left) daytime and (right) nighttime, and (lower) time/latitude variation of the same statistical measure. Each cell has at least 30 match-ups.

The maps in Figure 4-4 have some notable differences compared to those seen in Figure 4-2. The most obvious difference is the non-observation of the high latitude variations evident in Figure 4-2. This is due to a combination of improved cloud masking but also the introduction of total column water vapour banded retrieval coefficients that better account for the very low amounts of atmospheric water vapour in these regions in the L2P retrieval scheme. The second most obvious difference is the improved regional consistency for both day and night results, although some evidence of potential residual cloud is seen during the day in the areas of persistent cloud in the tropics (but of lower magnitude than for the NR data). As in Figure 4-2, no notable signature of Saharan dust is evident in Figure 4-4.

Although the results shown in Figure 4-3 indicate that the L2P data is slightly warmer during the AATSR-2 and AATSR periods compared to the NR data, and slightly cooler during the AATSR-1 period, the notable increase in regional consistency and reduction of high-latitude biases in Figure 4-4 compared to Figure 4-2 indicates that the user should use the L2P product as their main source of 1km AATSR SSTs and not the NR product.



## 5 SUMMARY

This report summarises the validation of SST products from the AATSR V3.0 dataset. Two products are available (1) NR and (2) L2P. A comparison of NR versus L2P match-ups to drifting buoys is given below in Table 5-1. In general, the global biases are very similar for the AATSR and ATSR-2 sensors for both the NR and L2P products, with both being slightly warm but well within specification (biases to be  $< 0.3$  K). ATSR-1 L2P data exhibit a slightly cooler global bias compared to NR data but are still within specification. In all cases the robust standard deviation is lower for the L2P product due to a combination of reduced regional biases and improved cloud masking.

Reference	Retrieval	Number	Median (K)	RSD (K)
<b>AATSR</b>	<i>NR D2 Day</i>	450733	-0.08	0.24
	<i>L2P D2 Day</i>	580885	-0.06	0.22
	<i>NR D3 Night</i>	510338	-0.14	0.22
	<i>L2P D3 Night</i>	431431	-0.11	0.19
<b>ATSR-2</b>	<i>NR D2 Day</i>	100988	-0.07	0.31
	<i>L2P D2 Day</i>	78442	-0.02	0.31
	<i>NR D3 Night</i>	93008	-0.18	0.28
	<i>L2P D3 Night</i>	57531	-0.14	0.24
<b>ATSR-1</b>	<i>NR D2 Day</i>	21872	+0.01	0.55
	<i>L2P D2 Day</i>	24782	-0.08	0.53
	<i>NR D2 Night</i>	20220	-0.21	0.54
	<i>L2P D2 Night</i>	15920	-0.25	0.52
	<i>NR D3 Night</i>	1513	-0.29	0.38
	<i>L2P D3 Night</i>	1214	-0.35	0.35

**Table 5-1:** Validation statistics from comparing AATSR V3.0 NR and L2P SSTs to those from drifting buoys. Results are reported as the total number of match-up, the median discrepancy and the robust standard deviation.

The following conclusions are drawn from the results presented in this report:

- AATSR and ATSR-2 SST dual-view retrievals are  $< 0.15$  K for both the NR and L2P products.
- Regional variations exist particularly for known retrieval issues such as water vapour affecting N2 retrievals and tropospheric aerosol affecting N2 and N3 retrievals.
- The L2P product has reduced regional biases and improved cloud clearing.
- ATSR-1 dual-view two-channel retrievals are  $< 0.25$  K for both the NR and L2P products.
- ATSR-1 NR nadir-view retrievals are significantly affected by aerosol dust during the period around Mount Pinatubo (from 1991 to 1993) and should not be used for quantitative analysis.



It is recommended that:

1. The ATSR V3.0 SST dataset be released
2. That users should in the first instance use the L2P product



## 6 REFERENCES

Embury, O., and Corlett, G. K (2010). AATSR SST Retrieval: Updated retrieval coefficients based on ARC project findings. Technical note for the ATSR QWG, UL-SST-P04, Issue 1A.

Embury, O., Merchant, C. J., and Corlett, G. K (2012). A reprocessing for climate of sea surface temperature from the along-track scanning radiometers: initial validation, accounting for skin and diurnal variability. *Remote Sensing of Environment*, 116, 62-78. doi: 10.1016/j.rse.2011.02.028

Merchant, C.J., Harris, A.R., Murray, M.J., and Závody, A.M., (1999). Toward the Elimination of Bias in Satellite Retrieval of Sea Surface Temperature 1. Theory, Modelling and Interalgorithm Comparison. *Journal of Geophysical Research*, 104, 23565-23578.

Merchant, C. J., Embury, O., Rayner, N. A., Berry, D. I., Corlett, G. K., Lean, K., and Saunders, R. (2012). A 20 year independent record of sea surface temperature for climate from along-track scanning radiometers. *Journal of Geophysical Research C: Oceans*, 117(12). doi: 10.1029/2012JC008400

Závody, A.M., Mutlow, C.T., and Llewellyn-Jones, D.T., (1995). A Radiative Transfer Model for Sea Surface Temperature Retrieval from the Along-Track Scanning Radiometer. *Journal of Geophysical research*, 100, 937-952.



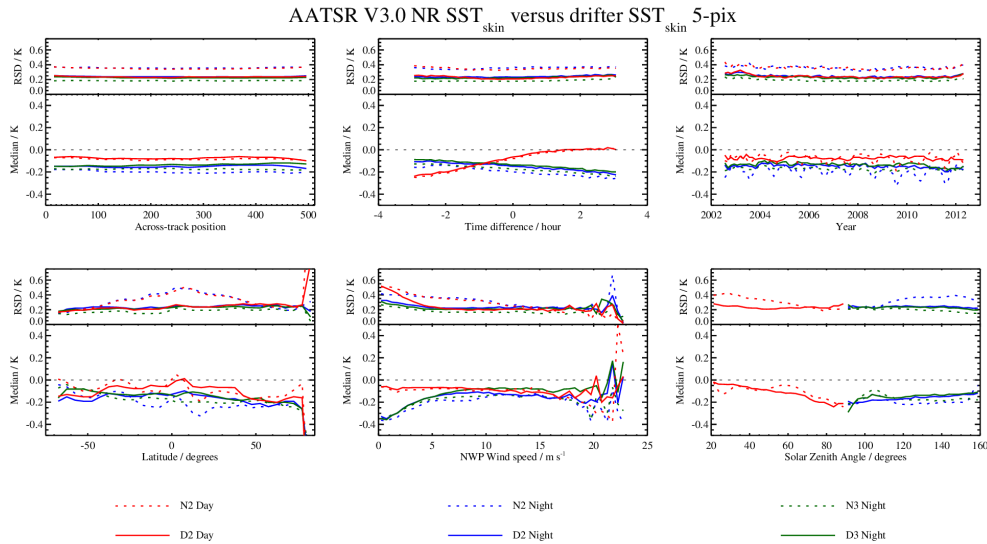
## 7 APPENDIX 1 – DETAILED AATSR VALIDATION RESULTS

The following section contains the detailed validation results for the AATSR V3.0 NR and L2P products. The results include:

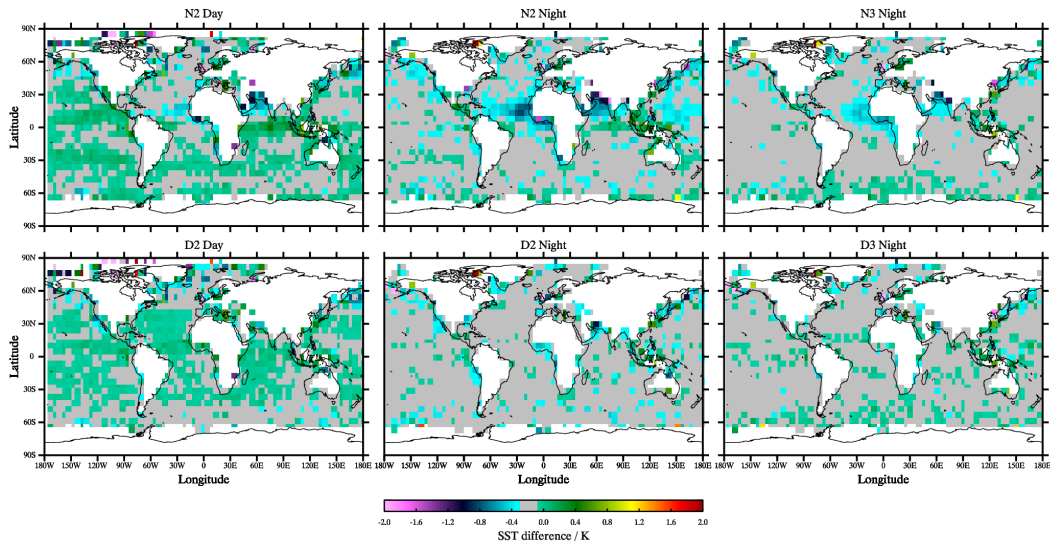
- Match-ups to four different reference datasets
  - Drifting buoys providing  $SST_{depth}$
  - Argo floats providing  $SST_{depth}$
  - GTMBA providing  $SST_{depth}$
  - Ship-borne radiometers providing  $SST_{skin}$
- Dependence plots of median and robust standard deviation of the discrepancy between the satellite  $SST_{skin}$  versus reference  $SST_{depth}$  or  $SST_{skin}$ .
  - For both NR and L2P match-ups, dependences are provided for latitude, time difference between satellite and drifter measurements, year, wind speed, solar zenith angle and across-track position.
- Spatial maps of the median discrepancy between the satellite and the reference datasets.
- Results are shown for 5-pixel averages centred on the match-up location.
- A summary table showing the median discrepancy and robust standard deviation of the complete set of match-ups to each available validation dataset.



### 7.1 Comparisons between ATS\_NR SSTs and drifting buoys

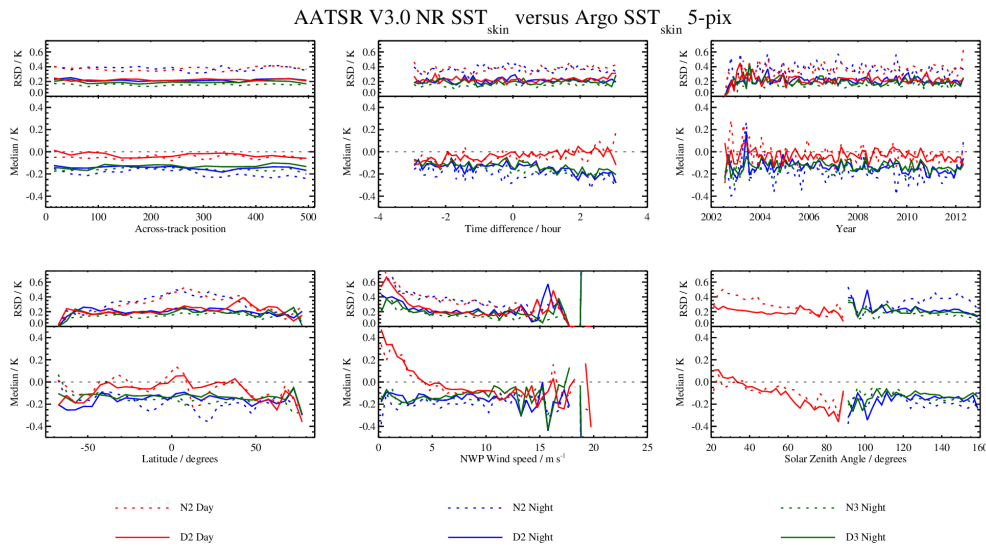


**Figure 7-1:** Dependence of the median and robust standard deviation between AATSR NR SST<sub>skin</sub> and drifter SST<sub>depth</sub> discrepancies as a function of across-track position, time difference, year, latitude, wind speed and solar zenith angle. Daytime results are shown in red, nighttime 2-channel results are shown in blue and nighttime 3-channel results are shown in green. Dual-view results are represented as solid lines and nadir-only results as dashed lines.

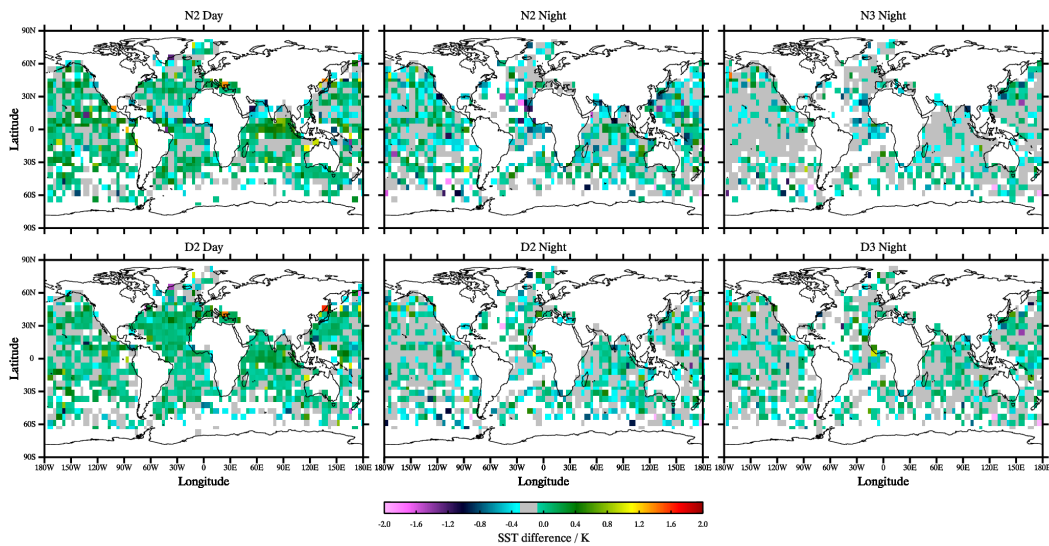


**Figure 7-2:** Spatial distribution of the median discrepancy between AATSR NR SST<sub>skin</sub> and drifter SST<sub>depth</sub>. The greyed region indicates a band of +/- 0.1 K around the expected mean skin offset of -0.17 K.

## 7.2 Comparisons between ATS\_NR SSTs and Argo

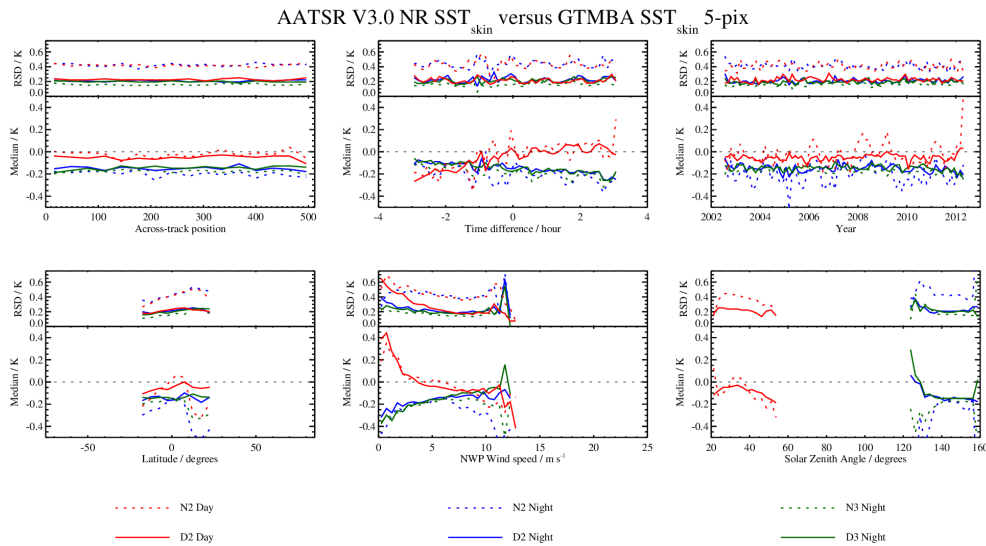


**Figure 7-3:** Dependence of the median and robust standard deviation between AATSR NR SST<sub>skin</sub> and Argo SST<sub>depth</sub> discrepancies as a function of across-track position, time difference, year, latitude, wind speed and solar zenith angle. Daytime results are shown in red, nighttime 2-channel results are shown in blue and nighttime 3-channel results are shown in green. Dual-view results are represented as solid lines and nadir-only results as dashed lines.

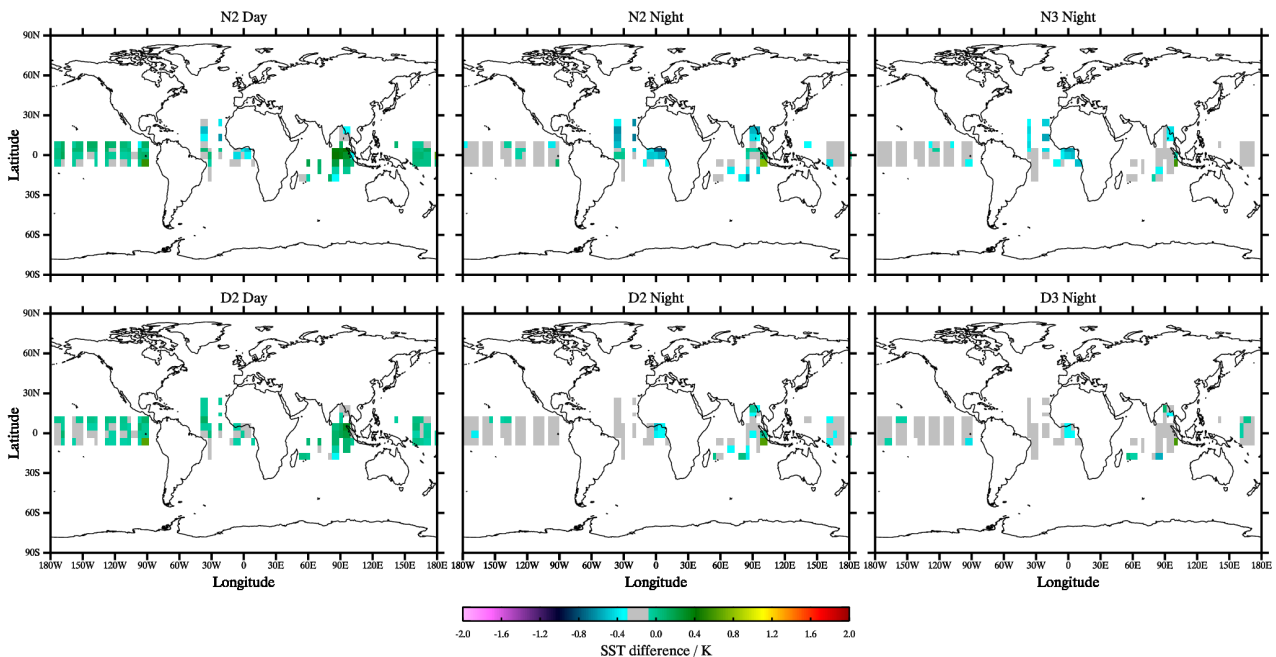


**Figure 7-4:** Spatial distribution of the median discrepancy between AATSR NR SST<sub>skin</sub> and Argo SST<sub>depth</sub>. The greyed region indicates a band of +/- 0.1 K around the expected mean skin offset of -0.17 K.

### 7.3 Comparisons between ATS\_NR SSTs and the GTMBA

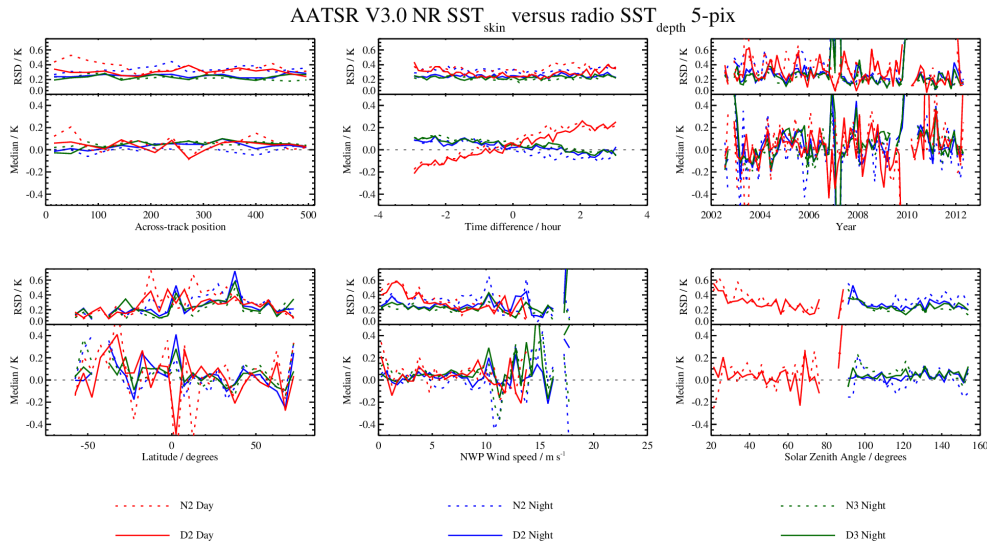


**Figure 7-5:** Dependence of the median and robust standard deviation between AATSR NR SST<sub>skin</sub> and GTMBA SST<sub>depth</sub> discrepancies as a function of across-track position, time difference, year, latitude, wind speed and solar zenith angle. Daytime results are shown in red, nighttime 2-channel results are shown in blue and nighttime 3-channel results are shown in green. Dual-view results are represented as solid lines and nadir-only results as dashed lines.

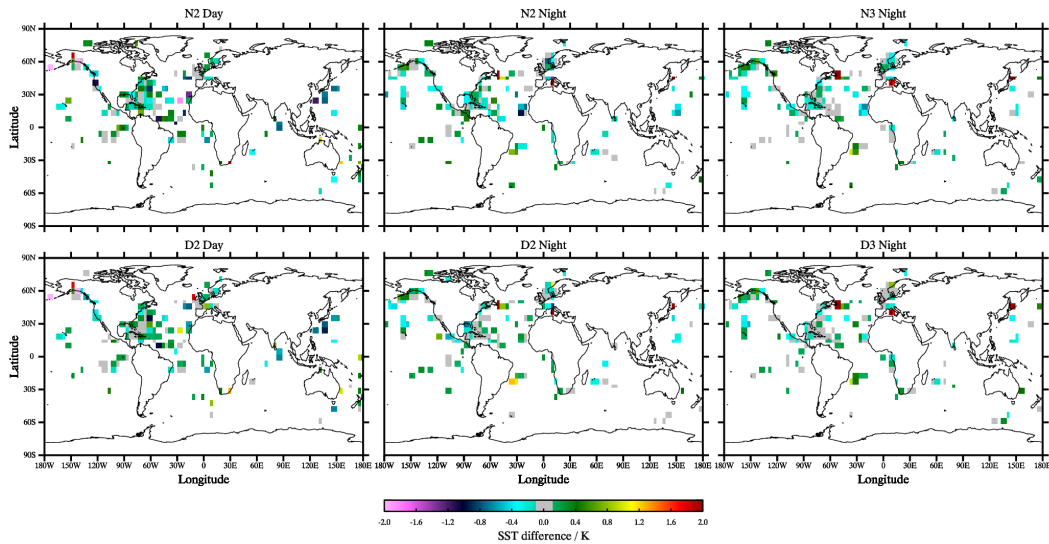


**Figure 7-6:** Spatial distribution of the median discrepancy between AATSR NR SST<sub>skin</sub> and GTMBA SST<sub>depth</sub>. The greyed region indicates a band of +/- 0.1 K around the expected mean skin offset of -0.17 K.

### 7.4 Comparisons between AATSR NR SSTs and radiometers



**Figure 7-7:** Dependence of the median and robust standard deviation between AATSR NR SST<sub>skin</sub> and radiometer SST<sub>skin</sub> discrepancies as a function of across-track position, time difference, year, latitude, wind speed and solar zenith angle. Daytime results are shown in red, nighttime 2-channel results are shown in blue and nighttime 3-channel results are shown in green. Dual-view results are represented as solid lines and nadir-only results as dashed lines.



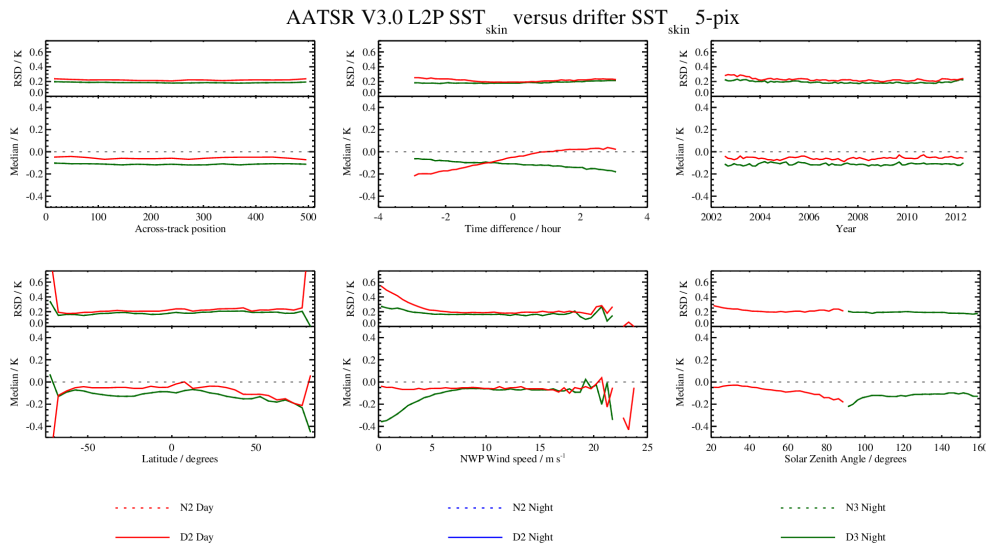
**Figure 7-8:** Spatial distribution of the median discrepancy between AATSR NR SST<sub>skin</sub> and radiometer SST<sub>skin</sub>. The greyed region indicates a band of +/- 0.1 K around zero.

**7.5 Statistical analysis of ATS\_NR results**

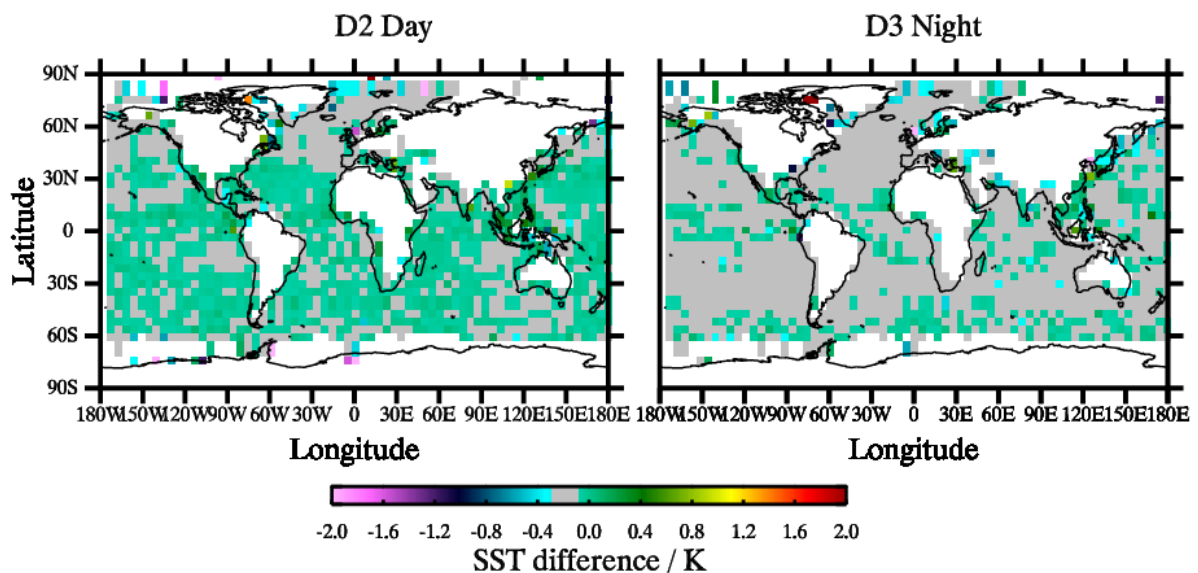
Reference	Retrieval	Number	Median (K)	RSD (K)
<b>Drifters</b>	<i>Day N2</i>	450733	-0.09	0.35
	<i>Day D2</i>	450733	-0.08	0.24
	<i>Night N2</i>	510338	-0.20	0.36
	<i>Night N3</i>	510338	-0.18	0.19
	<i>Night D2</i>	510338	-0.15	0.24
	<i>Night D3</i>	510338	-0.14	0.22
<b>GTMBA</b>	<i>Day N2</i>	17961	-0.03	0.42
	<i>Day D2</i>	17961	-0.05	0.22
	<i>Night N2</i>	19662	-0.20	0.43
	<i>Night N3</i>	19662	-0.18	0.16
	<i>Night D2</i>	19662	-0.15	0.21
	<i>Night D3</i>	19662	-0.15	0.19
<b>Argo</b>	<i>Day N2</i>	5283	-0.05	0.37
	<i>Day D2</i>	5283	-0.03	0.26
	<i>Night N2</i>	3946	-0.20	0.37
	<i>Night N3</i>	3946	-0.16	0.16
	<i>Night D2</i>	3946	-0.15	0.21
	<i>Night D3</i>	3946	-0.13	0.19
<b>Radiometers</b>	<i>Day N2</i>	9596	+0.06	0.36
	<i>Day D2</i>	9596	+0.04	0.31
	<i>Night N2</i>	14264	+0.01	0.33
	<i>Night N3</i>	14264	+0.06	0.22
	<i>Night D2</i>	14264	+0.03	0.25
	<i>Night D3</i>	14264	+0.05	0.23

**Table 7-1:** Global validation statistics from comparing AATSR NR V3.0 to the available validation datasets.

### 7.6 Comparisons between ATS\_L2P SSTs and drifting buoys

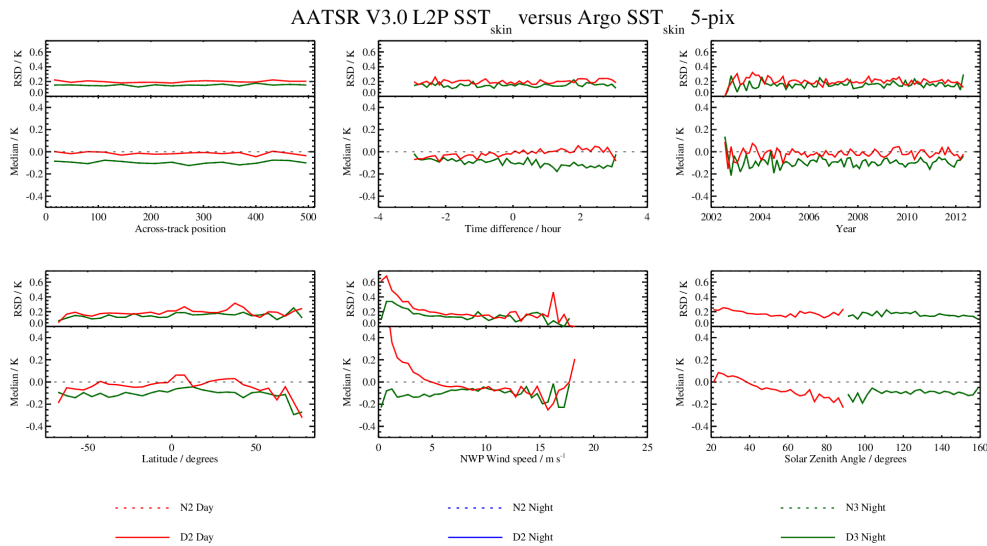


**Figure 7-9:** Dependence of the median and robust standard deviation between AATSR L2P SST<sub>skin</sub> and drifter SST<sub>depth</sub> discrepancies as a function of across-track position, time difference, year, latitude, wind speed and solar zenith angle. Daytime results are shown in red, nighttime 2-channel results are shown in blue and nighttime 3-channel results are shown in green. Dual-view results are represented as solid lines and nadir-only results as dashed lines.

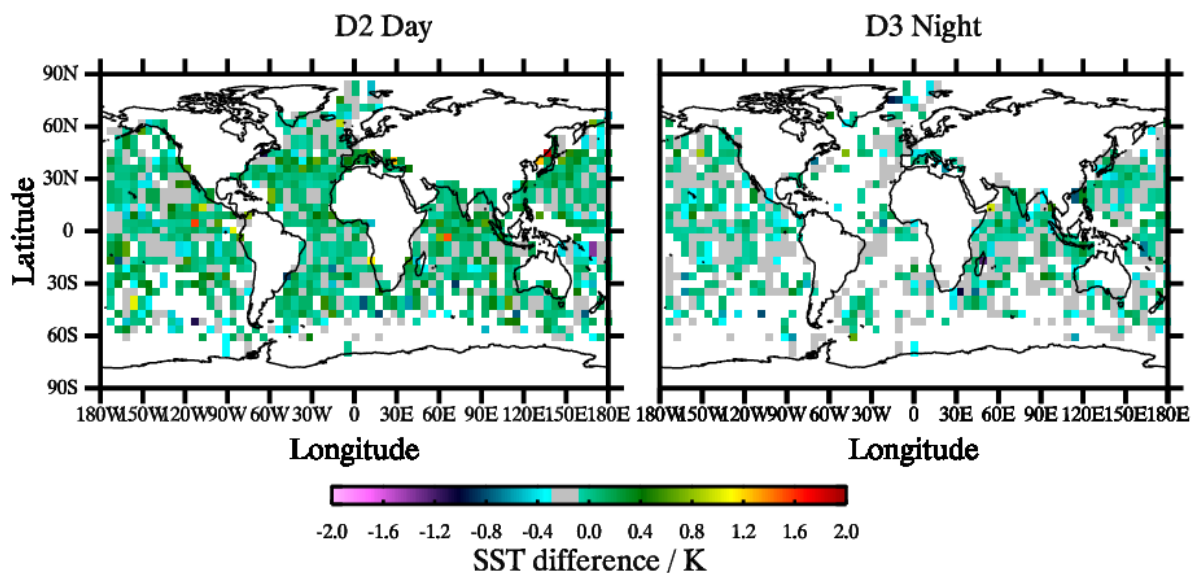


**Figure 7-10:** Spatial distribution of the median discrepancy between AATSR L2P SST<sub>skin</sub> and drifter SST<sub>depth</sub>. The greyed region indicates a band of +/- 0.1 K around the expected mean skin offset of -0.17 K.

### 7.7 Comparisons between ATS\_L2P SSTs and Argo

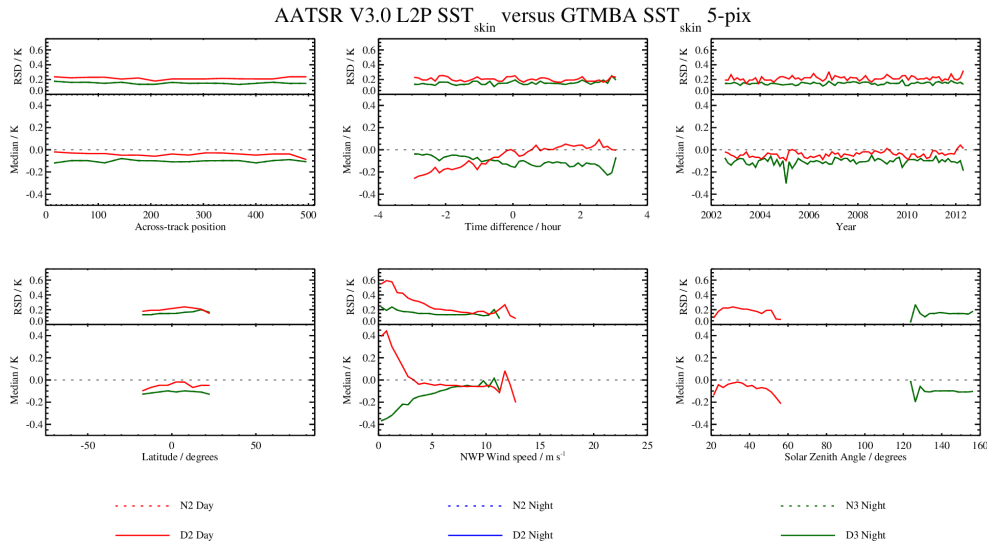


**Figure 7-11:** Dependence of the median and robust standard deviation between AATSR L2P SST<sub>skin</sub> and Argo SST<sub>depth</sub> discrepancies as a function of across-track position, time difference, year, latitude, wind speed and solar zenith angle. Daytime results are shown in red, nighttime 2-channel results are shown in blue and nighttime 3-channel results are shown in green. Dual-view results are represented as solid lines and nadir-only results as dashed lines.

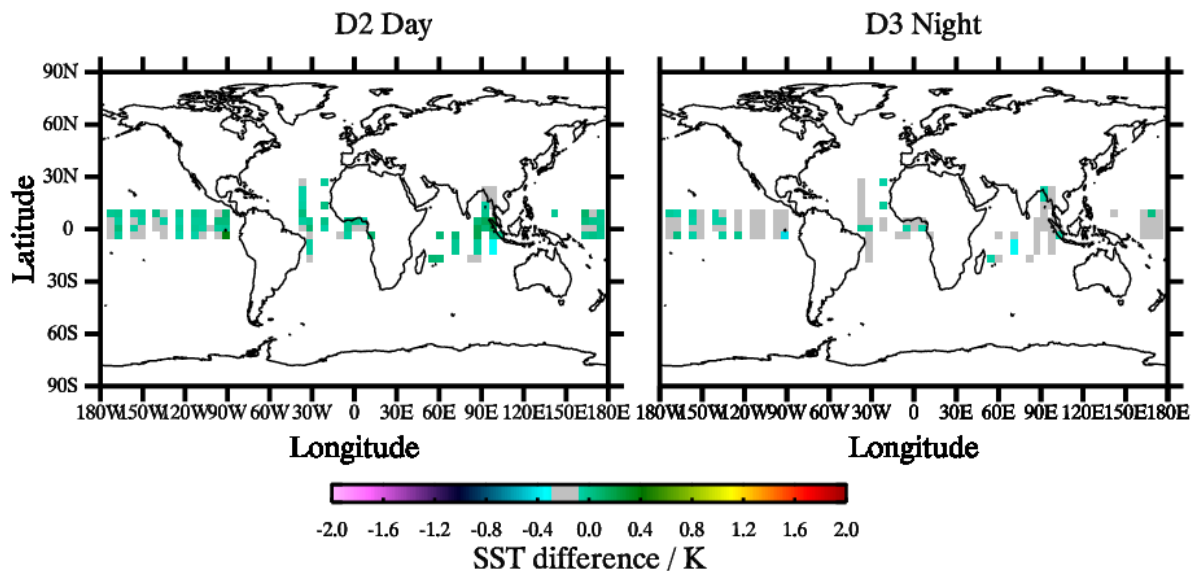


**Figure 7-12:** Spatial distribution of the median discrepancy between AATSR L2P SST<sub>skin</sub> and Argo SST<sub>depth</sub>. The greyed region indicates a band of +/- 0.1 K around the expected mean skin offset of -0.17 K.

### 7.8 Comparisons between ATS\_L2P SSTs and the GTMBA



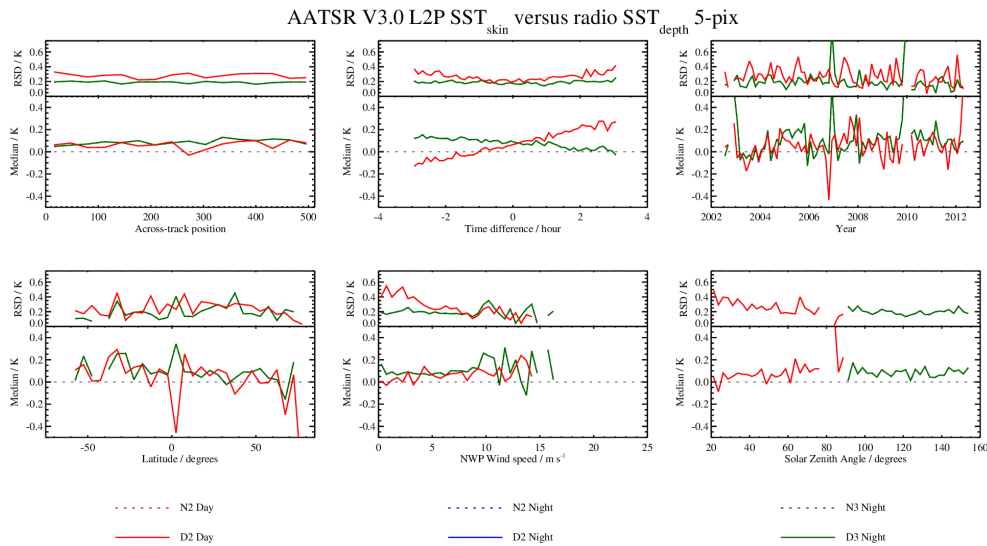
**Figure 7-13:** Dependence of the median and robust standard deviation between AATSR L2P SST<sub>skin</sub> and GTMBA SST<sub>depth</sub> discrepancies as a function of across-track position, time difference, year, latitude, wind speed and solar zenith angle. Daytime results are shown in red, nighttime 2-channel results are shown in blue and nighttime 3-channel results are shown in green. Dual-view results are represented as solid lines and nadir-only results as dashed lines.



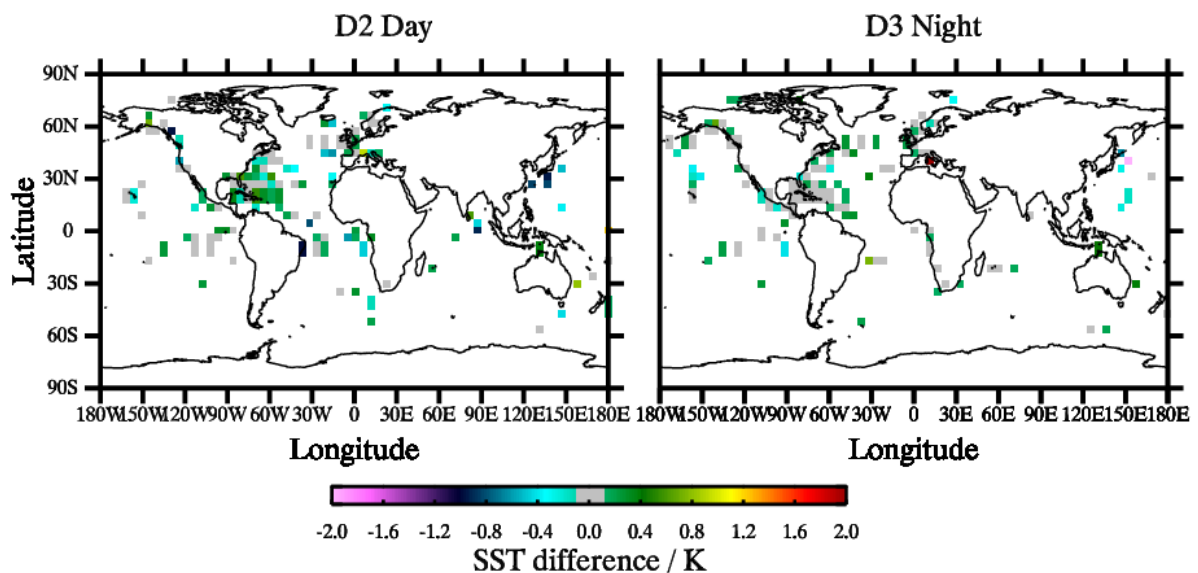
**Figure 7-14:** Spatial distribution of the median discrepancy between AATSR L2P SST<sub>skin</sub> and GTMBA SST<sub>depth</sub>. The greyed region indicates a band of +/- 0.1 K around the expected mean skin offset of -0.17 K.



### 7.9 Comparisons between ATS\_L2P SSTs and radiometers



**Figure 7-15:** Dependence of the median and robust standard deviation between AATSR L2P SST<sub>skin</sub> and radiometer SST<sub>skin</sub> discrepancies as a function of across-track position, time difference, year, latitude, wind speed and solar zenith angle. Daytime results are shown in red, nighttime 2-channel results are shown in blue and nighttime 3-channel results are shown in green. Dual-view results are represented as solid lines and nadir-only results as dashed lines.



**Figure 7-16:** Spatial distribution of the median discrepancy between AATSR L2P SST<sub>skin</sub> and radiometer SST<sub>skin</sub>. The greyed region indicates a band of +/- 0.1 K around zero.



## 7.10 Statistical analysis of AATSR\_L2P results

Reference	Retrieval	Number	Median (K)	RSD (K)
<b>Drifters</b>	<i>Day D2</i>	580885	-0.06	0.22
	<i>Night D3</i>	431431	-0.11	0.19
<b>GTMBA</b>	<i>Day D2</i>	21940	-0.04	0.22
	<i>Night D3</i>	15600	-0.10	0.15
<b>Argo</b>	<i>Day D2</i>	6525	-0.01	0.20
	<i>Night D3</i>	2213	-0.10	0.15
<b>Radiometers</b>	<i>Day D2</i>	12128	+0.06	0.48
	<i>Night D3</i>	13806	+0.09	0.19

**Table 7-2:** Global validation statistics from comparing AATSR L2P V3.0 to the available validation datasets. L2P format files contain only the best SST available and so do not contain nadir-only SSTs.



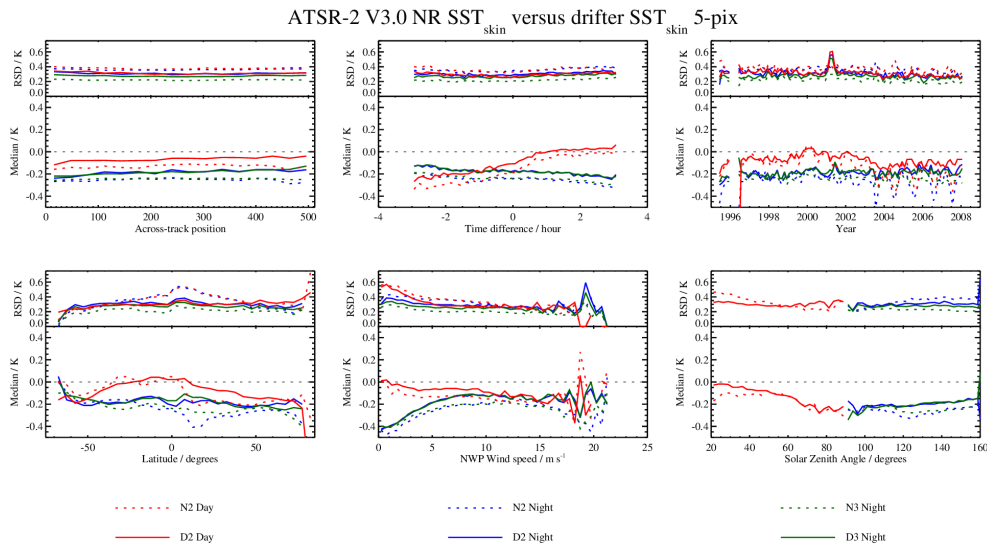
## 8 APPENDIX 2 – DETAILED ATSR-2 VALIDATION RESULTS

The following section contains the detailed validation results for the ATSR-2 V3.0 NR and L2P products. The results include:

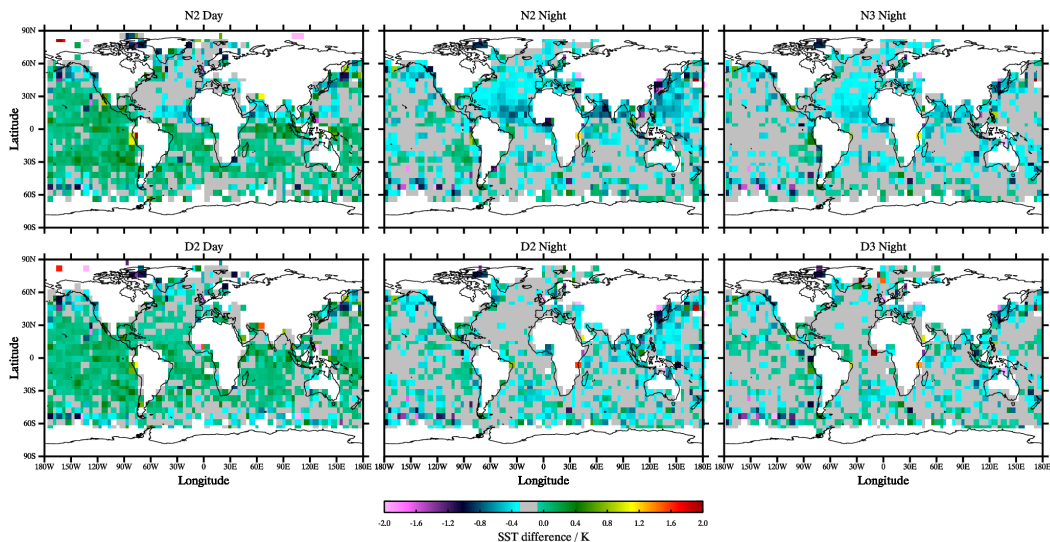
- Match-ups to three different references datasets
  - Drifting buoys providing  $SST_{depth}$
  - GTMBA providing  $SST_{depth}$
  - Ship-borne radiometers providing  $SST_{skin}$
- Dependence plots of median and robust standard deviation of the discrepancy between the satellite  $SST_{skin}$  versus reference  $SST_{depth}$  or  $SST_{skin}$ .
  - For both NR and L2P match-ups, dependences are provided for latitude, time difference between satellite and drifter measurements, year, wind speed, solar zenith angle and across-track position.
- Spatial maps of the median discrepancy between the satellite and the reference datasets.
- A summary table showing the median discrepancy and robust standard deviation of the complete set of match-ups to each available validation dataset.

Note: The NR results cover the period from June 1995 to January 2008; L2P results cover the period from June 1995 to June 2003.

### 8.1 Comparisons between AT2\_NR SSTs and drifting buoys

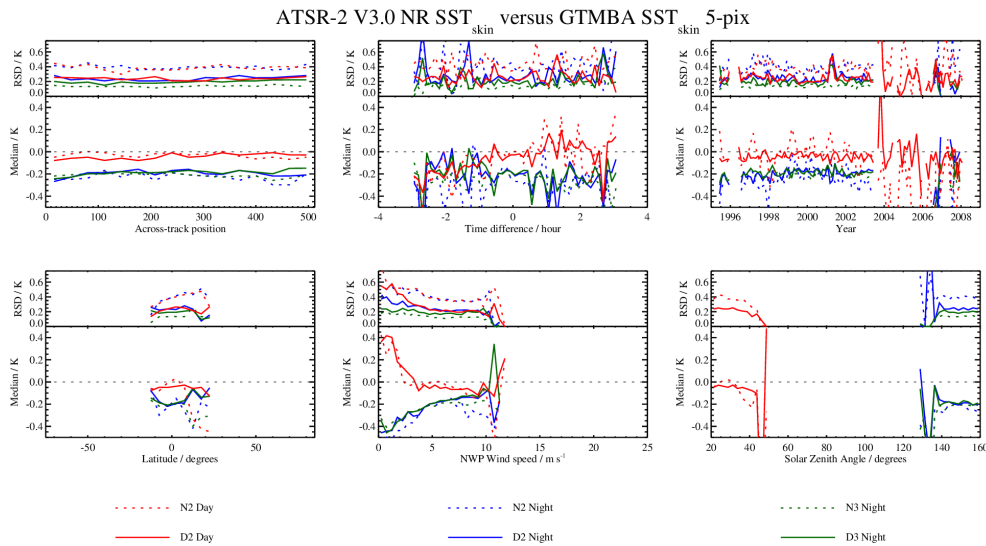


**Figure 8-1:** Dependence of the median and robust standard deviation between ATSR-2 NR SST<sub>skin</sub> and drifter SST<sub>depth</sub> discrepancies as a function of across-track position, time difference, year, latitude, wind speed and solar zenith angle. Daytime results are shown in red, nighttime 2-channel results are shown in blue and nighttime 3-channel results are shown in green. Dual-view results are represented as solid lines and nadir-only results as dashed lines.

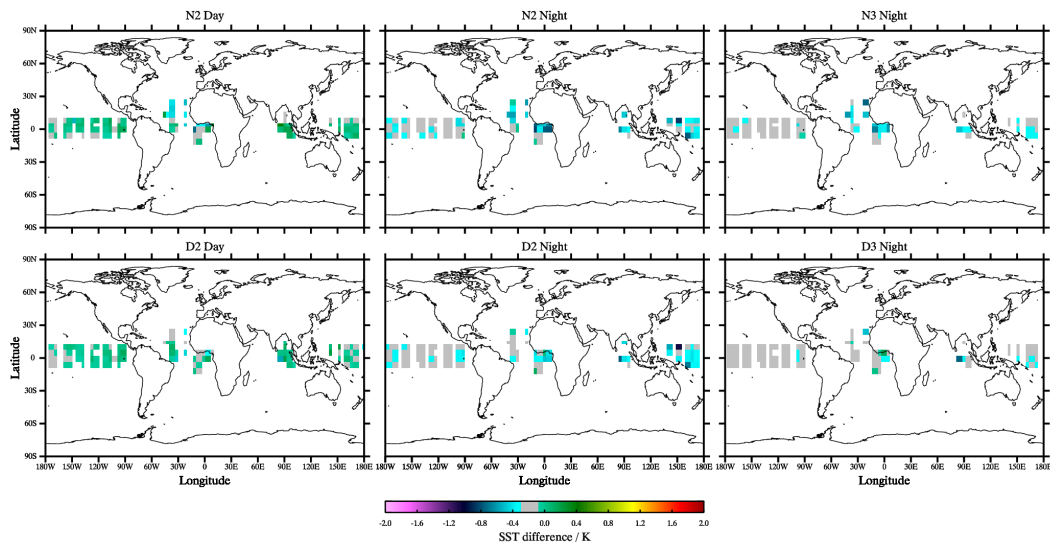


**Figure 8-2:** Spatial distribution of the median discrepancy between ATSR-2 NR SST<sub>skin</sub> and drifter SST<sub>depth</sub>. The greyed region indicates a band of +/- 0.1 K around the expected mean skin offset of -0.17 K.

## 8.2 Comparisons between AT2\_NR SSTs and the GTMBA

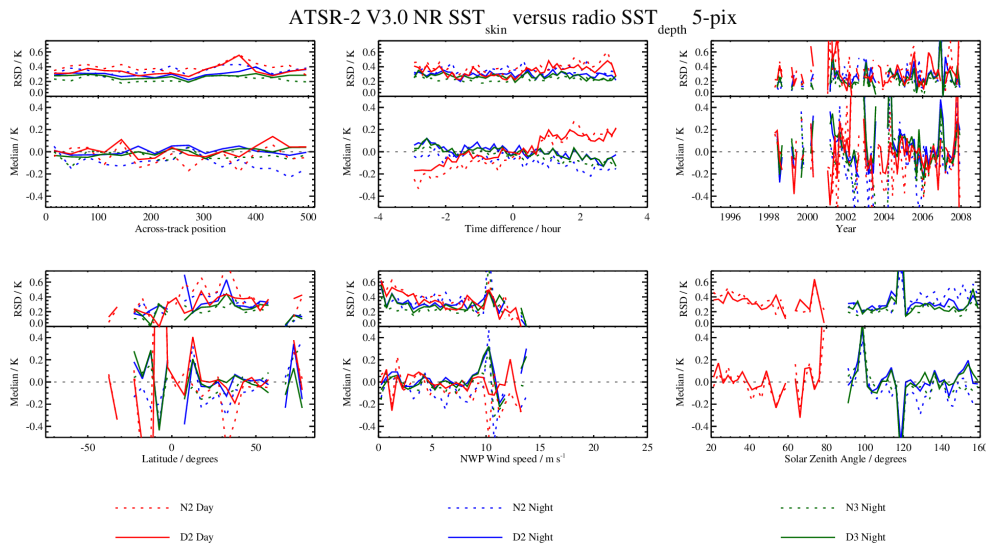


**Figure 8-3:** Dependence of the median and robust standard deviation between ATSR-2 NR SST<sub>skin</sub> and GTMBA SST<sub>depth</sub> discrepancies as a function of across-track position, time difference, year, latitude, wind speed and solar zenith angle. Daytime results are shown in red, nighttime 2-channel results are shown in blue and nighttime 3-channel results are shown in green. Dual-view results are represented as solid lines and nadir-only results as dashed lines.

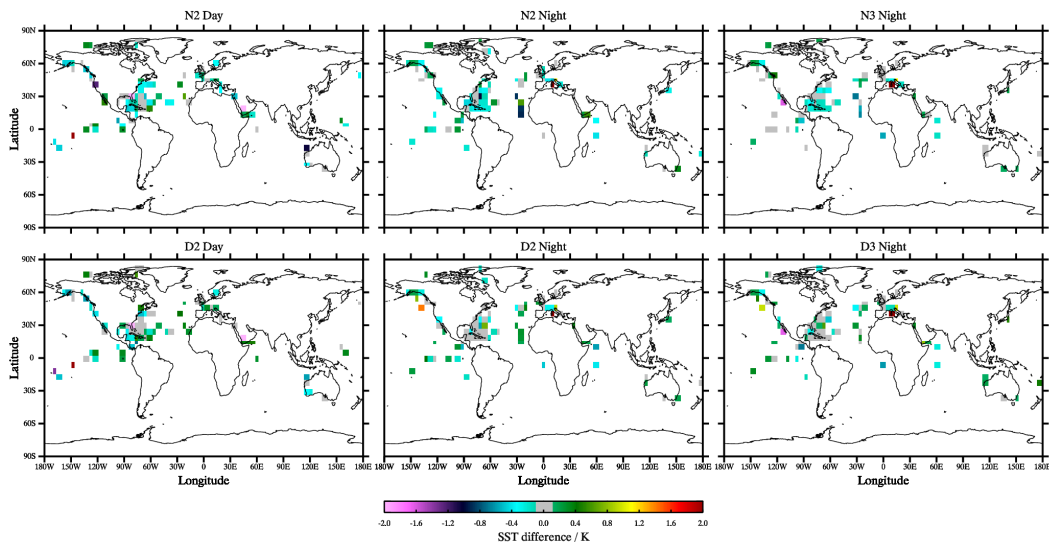


**Figure 8-4:** Spatial distribution of the median discrepancy between ATSR-2 NR SST<sub>skin</sub> and GTMBA SST<sub>depth</sub>. The greyed region indicates a band of +/- 0.1 K around the expected mean skin offset of -0.17 K.

### 8.3 Comparisons between AT2\_NR SSTs and radiometers



**Figure 8-5:** Dependence of the median and robust standard deviation between ATSR-2 NR SST<sub>skin</sub> and radiometer SST<sub>skin</sub> discrepancies as a function of across-track position, time difference, year, latitude, wind speed and solar zenith angle. Daytime results are shown in red, nighttime 2-channel results are shown in blue and nighttime 3-channel results are shown in green. Dual-view results are represented as solid lines and nadir-only results as dashed lines.



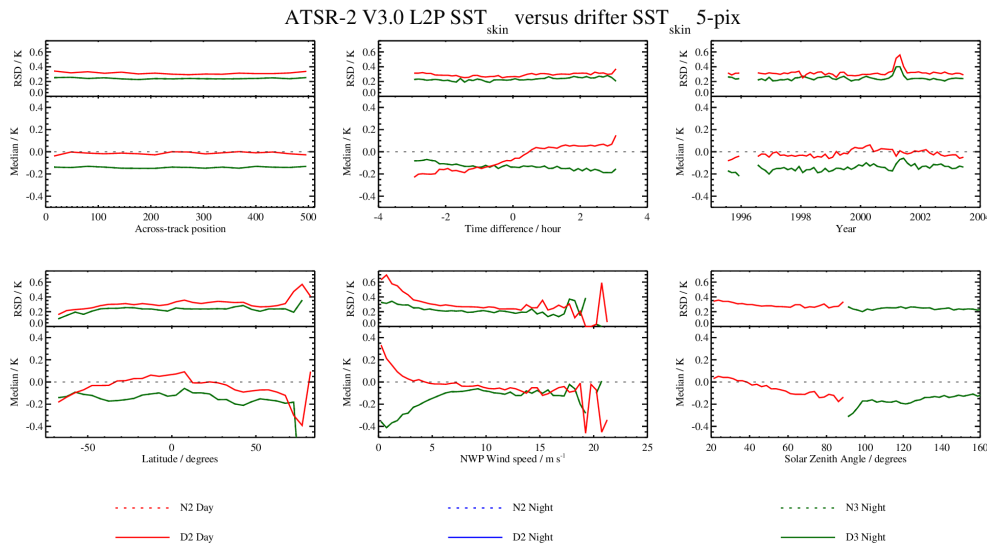
**Figure 8-6:** Spatial distribution of the median discrepancy between ATSR-2 NR SST<sub>skin</sub> and radiometer SST<sub>skin</sub>. The greyed region indicates a band of +/- 0.1 K around zero.

**8.4 Statistical analysis of AT2\_NR results**

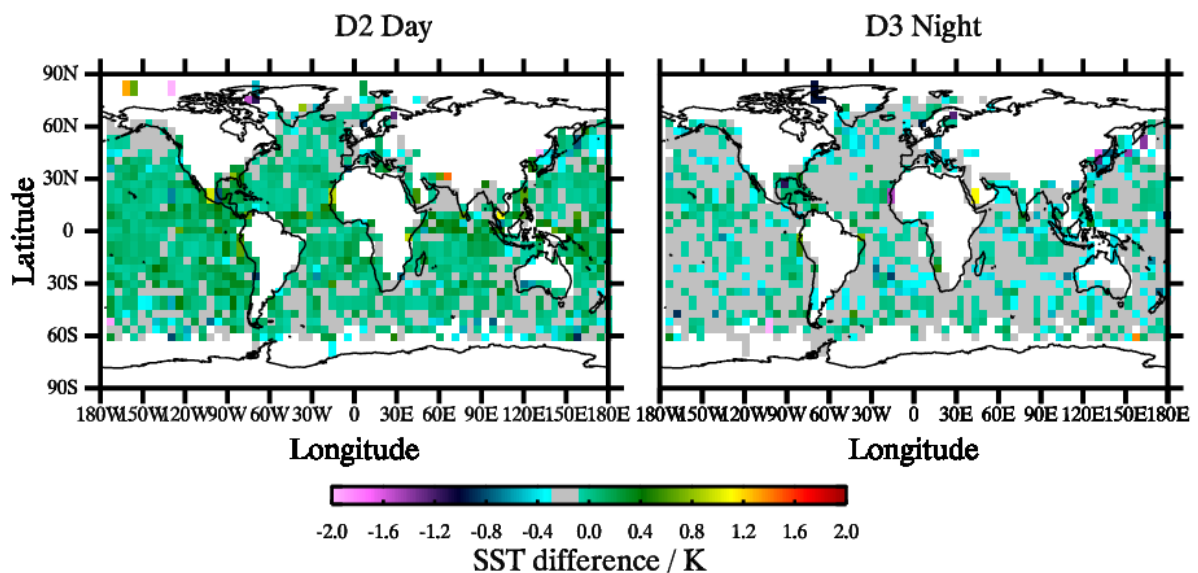
Reference	Retrieval	Number	Median (K)	RSD (K)
<b>Drifters</b>	<i>Day N2</i>	93008	-0.13	0.38
	<i>Day D2</i>	93008	-0.07	0.31
	<i>Night N2</i>	100988	-0.26	0.36
	<i>Night N3</i>	100988	-0.24	0.22
	<i>Night D2</i>	100988	-0.18	0.31
	<i>Night D3</i>	100988	-0.18	0.27
<b>GTMBA</b>	<i>Day N2</i>	7953	-0.03	0.39
	<i>Day D2</i>	7953	-0.05	0.24
	<i>Night N2</i>	7946	-0.23	0.40
	<i>Night N3</i>	7946	-0.21	0.14
	<i>Night D2</i>	7946	-0.20	0.24
	<i>Night D3</i>	7946	-0.19	0.19
<b>Radiometers</b>	<i>Day N2</i>	5701	-0.04	0.41
	<i>Day D2</i>	5701	+0.01	0.35
	<i>Night N2</i>	5951	-0.08	0.36
	<i>Night N3</i>	5951	-0.03	0.24
	<i>Night D2</i>	5951	+0.01	0.30
	<i>Night D3</i>	5951	-0.01	0.27

**Table 8-1:** Global validation statistics from comparing ATSR-2 NR V3.0 to the available validation datasets.

### 8.5 Comparisons between AT2\_L2P SSTs and drifting buoys



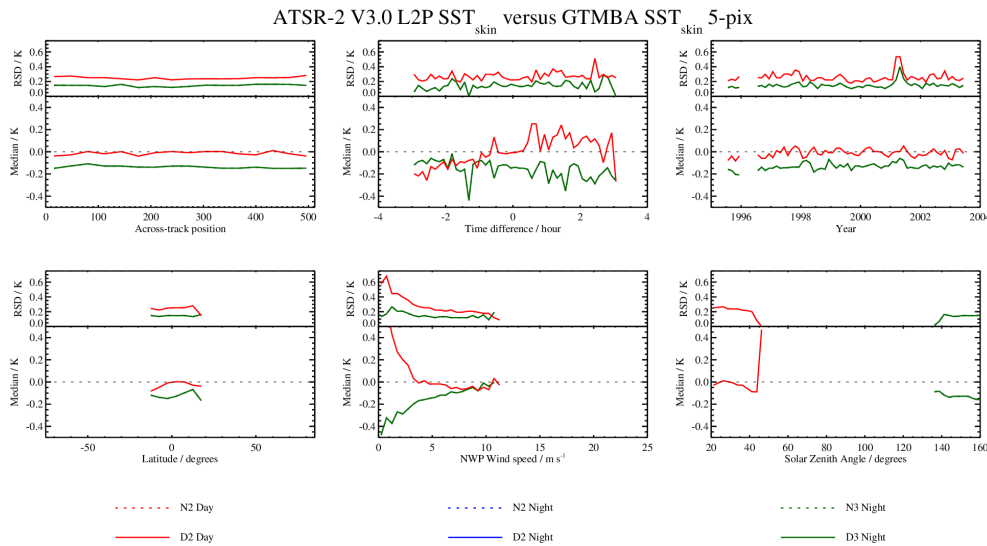
**Figure 8-7:** Dependence of the median and robust standard deviation between ATSR-2 L2P SST<sub>skin</sub> and drifter SST<sub>depth</sub> discrepancies as a function of across-track position, time difference, year, latitude, wind speed and solar zenith angle. Daytime results are shown in red, nighttime 2-channel results are shown in blue and nighttime 3-channel results are shown in green. Dual-view results are represented as solid lines and nadir-only results as dashed lines.



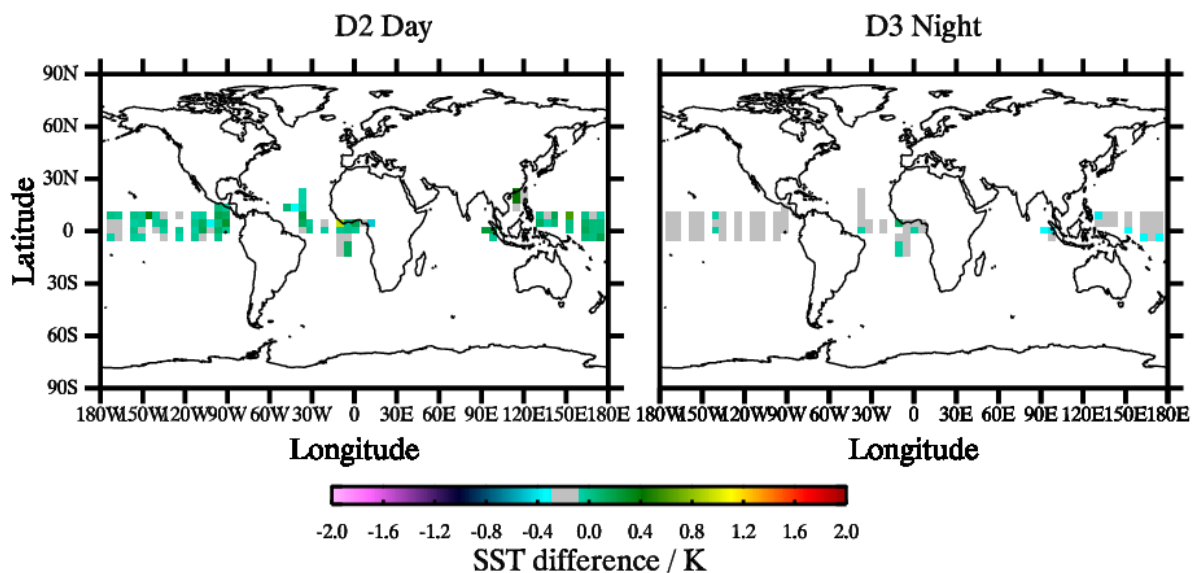
**Figure 8-8:** Spatial distribution of the median discrepancy between ATSR-2 L2P SST<sub>skin</sub> and drifter SST<sub>depth</sub>. The greyed region indicates a band of +/- 0.1 K around the expected mean skin offset of -0.17 K.



### 8.6 Comparisons between AT2\_L2P SSTs and the GTMBA

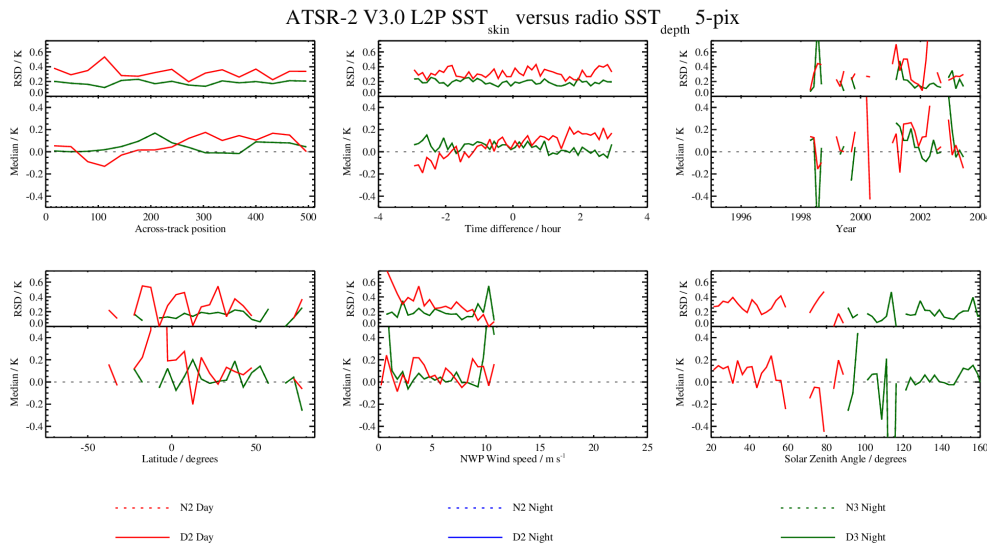


**Figure 8-9:** Dependence of the median and robust standard deviation between ATSR-2 L2P SST<sub>skin</sub> and GTMBA SST<sub>depth</sub> discrepancies as a function of across-track position, time difference, year, latitude, wind speed and solar zenith angle. Daytime results are shown in red, nighttime 2-channel results are shown in blue and nighttime 3-channel results are shown in green. Dual-view results are represented as solid lines and nadir-only results as dashed lines.

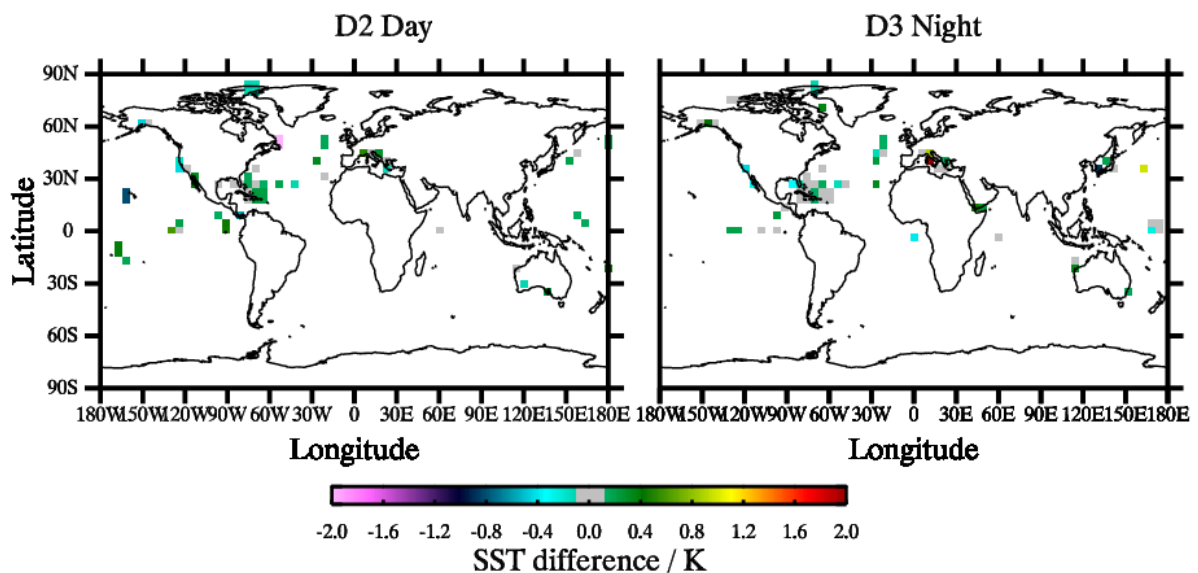


**Figure 8-10:** Spatial distribution of the median discrepancy between ATSR-2 L2P SST<sub>skin</sub> and GTMBA SST<sub>depth</sub>. The greyed region indicates a band of +/- 0.1 K around the expected mean skin offset of -0.17 K.

### 8.7 Comparisons between AT2\_L2P SSTs and radiometers



**Figure 8-11:** Dependence of the median and robust standard deviation between ATSR-2 L2P SST<sub>skin</sub> and radiometer SST<sub>skin</sub> discrepancies as a function of across-track position, time difference, year, latitude, wind speed and solar zenith angle. Daytime results are shown in red, nighttime 2-channel results are shown in blue and nighttime 3-channel results are shown in green. Dual-view results are represented as solid lines and nadir-only results as dashed lines.



**Figure 8-12:** Spatial distribution of the median discrepancy between ATSR-2 L2P SST<sub>skin</sub> and radiometer SST<sub>skin</sub>. The greyed region indicates a band of +/- 0.1 K around the expected mean skin offset of -0.17 K.



## 8.8 Statistical analysis of AT2\_L2P results

Reference	Retrieval	Number	Median (K)	RSD (K)
<b>Drifters</b>	<i>Day D2</i>	78442	-0.02	0.31
	<i>Night D3</i>	57531	-0.14	0.24
<b>GTMBAs</b>	<i>Day D2</i>	10478	-0.01	0.25
	<i>Night D3</i>	7005	-0.14	0.22
<b>Radiometers</b>	<i>Day D2</i>	1268	+0.07	0.33
	<i>Night D3</i>	1363	+0.04	0.20

**Table 8-2:** Global validation statistics from comparing ATSR-2 L2P V3.0 to the available validation datasets. L2P format files contain only the best SST available and so do not contain nadir-only SSTs.

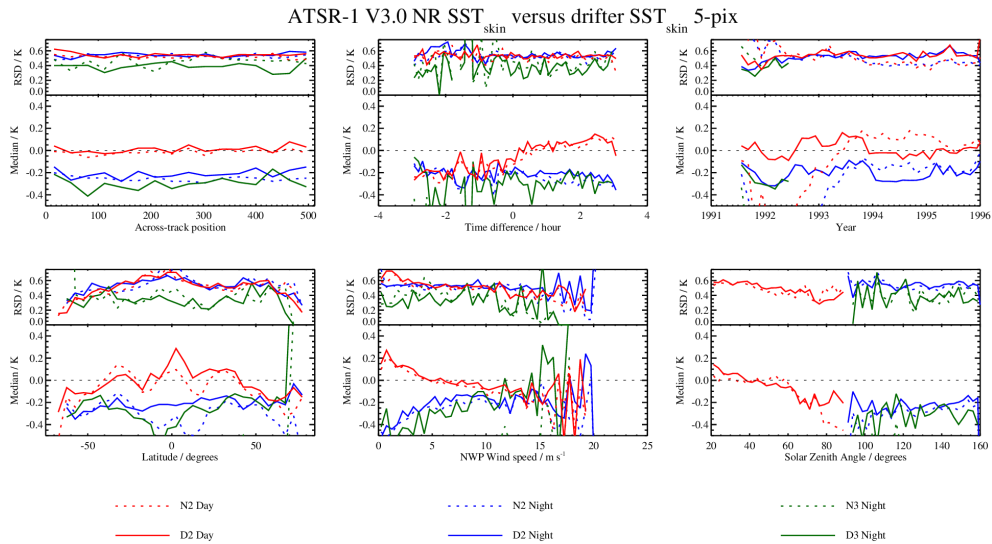


## 9 APPENDIX 3 – DETAILED ATSR-1 VALIDATION RESULTS

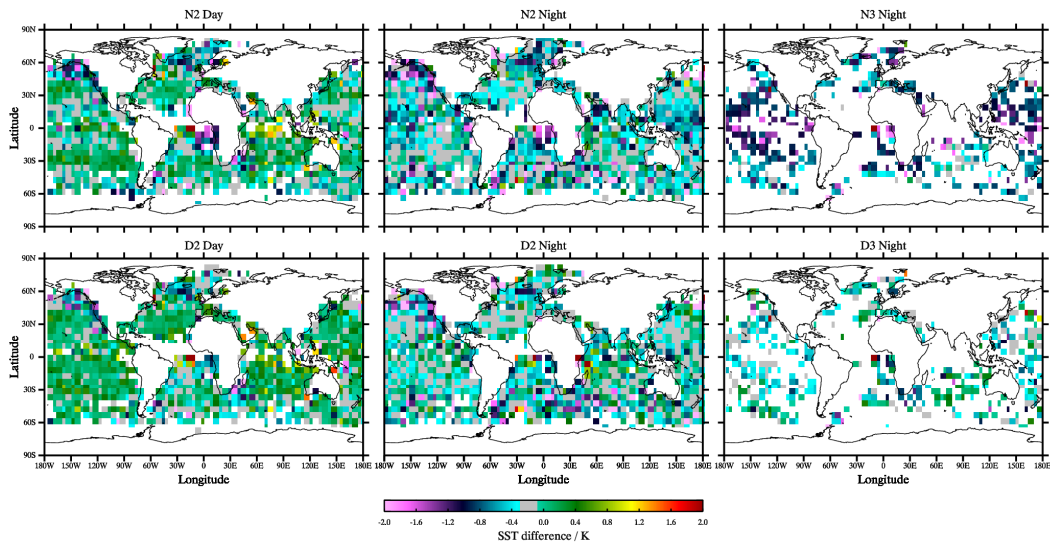
The following section contains the detailed validation results for the AATSR V3.0 NR and L2P products. The results include:

- Dependence plots of the median and robust standard deviation of the discrepancy between the satellite  $SST_{skin}$  and the drifter  $SST_{depth}$ .
  - For both NR and L2P match-ups, dependences are provided for latitude, time difference between satellite and drifter measurements, year, wind speed, solar zenith angle and across-track position.
- Spatial maps of the median discrepancy between the satellite and the drifters.
- A summary table showing the median discrepancy and robust standard deviation of the complete set of match-ups to each available validation dataset.

### 9.1 Comparisons between AT1\_NR SSTs and drifting buoys

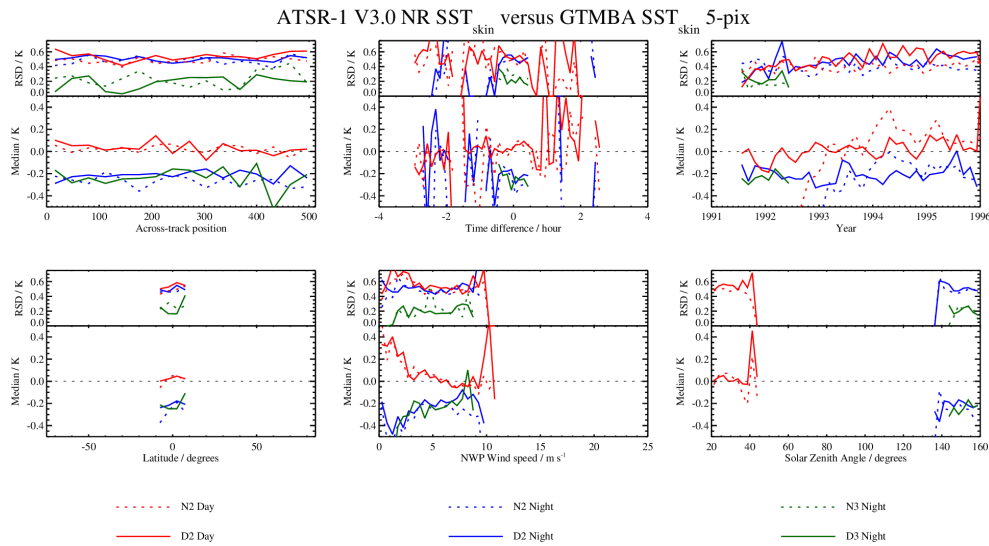


**Figure 9-1:** Dependence of the median and robust standard deviation between ATSR-1 NR SST<sub>skin</sub> and drifter SST<sub>depth</sub> discrepancies as a function of across-track position, time difference, year, latitude, wind speed and solar zenith angle. Daytime results are shown in red, nighttime 2-channel results are shown in blue and nighttime 3-channel results are shown in green. Dual-view results are represented as solid lines and nadir-only results as dashed lines.

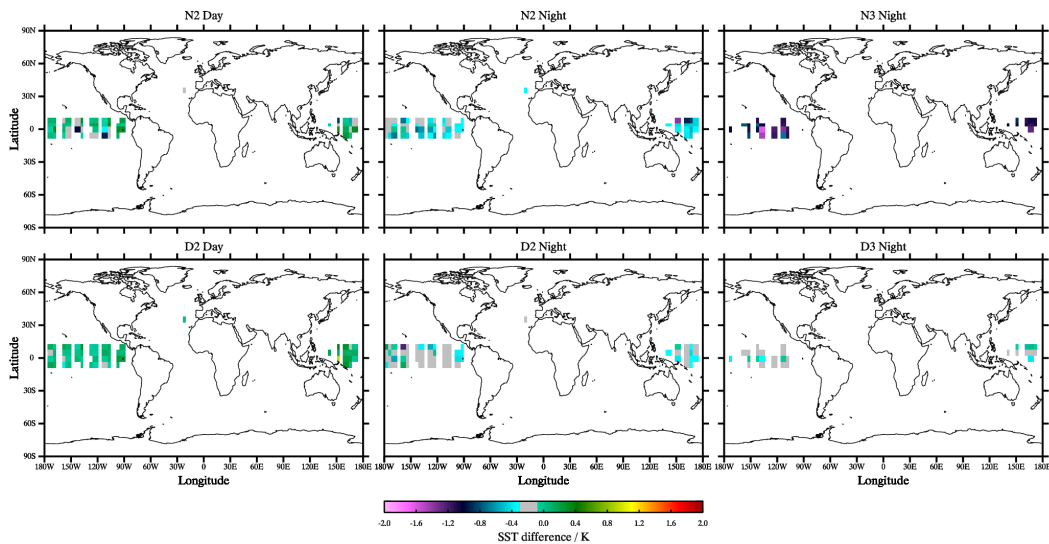


**Figure 9-2:** Spatial distribution of the median discrepancy between ATSR-1 NR SST<sub>skin</sub> and drifter SST<sub>depth</sub>. The greyed region indicates a band of +/- 0.1 K around the expected mean skin offset of -0.17 K.

## 9.2 Comparisons between AT1\_NR SSTs and GTMBA



**Figure 9-3:** Dependence of the median and robust standard deviation between ATSR-1 NR SST<sub>skin</sub> and GTMBA SST<sub>depth</sub> discrepancies as a function of across-track position, time difference, year, latitude, wind speed and solar zenith angle. Daytime results are shown in red, nighttime 2-channel results are shown in blue and nighttime 3-channel results are shown in green. Dual-view results are represented as solid lines and nadir-only results as dashed lines.



**Figure 9-4:** Spatial distribution of the median discrepancy between ATSR-1 NR SST<sub>skin</sub> and GTMBA SST<sub>depth</sub>. The greyed region indicates a band of +/- 0.1 K around the expected mean skin offset of -0.17 K.

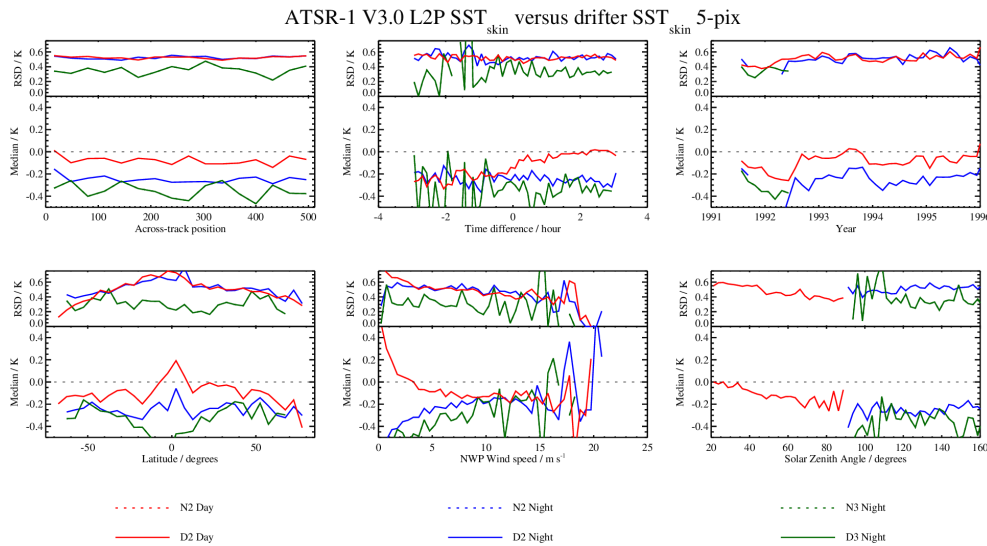


### 9.3 Statistical analysis of AT1\_NR results

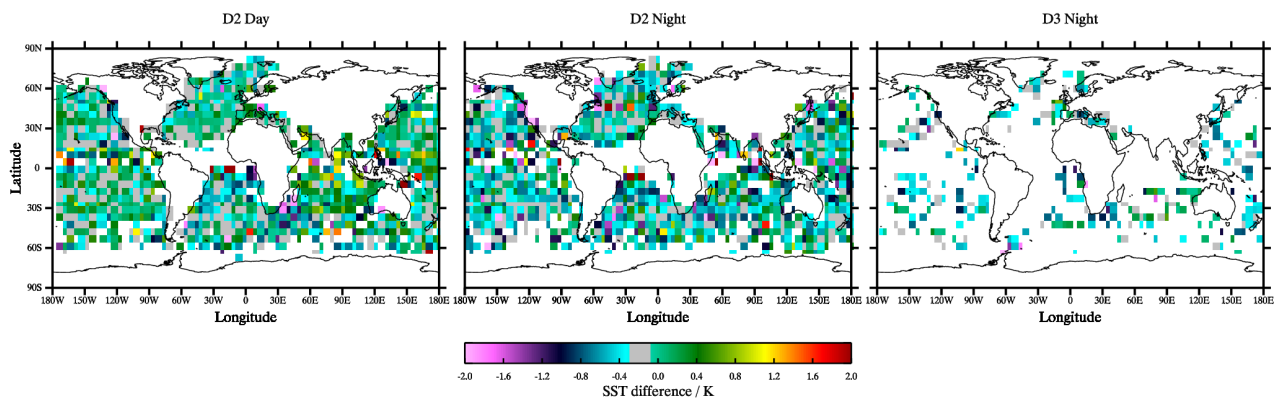
Reference	Retrieval	Number	Median (K)	RSD (K)
<b>Drifters</b>	<i>Day N2</i>	21872	-0.02	0.52
	<i>Day D2</i>	21872	+0.01	0.55
	<i>Night N2</i>	20220	-0.25	0.52
	<i>Night N3</i>	1044	-0.82	0.44
	<i>Night D2</i>	20220	-0.21	0.54
	<i>Night D3</i>	1044	-0.29	0.38
<b>GTMBA</b>	<i>Day N2</i>	3503	+0.02	0.49
	<i>Day D2</i>	3503	+0.02	0.55
	<i>Night N2</i>	3561	-0.27	0.47
	<i>Night N3</i>	182	-1.08	0.29
	<i>Night D2</i>	3561	-0.21	0.49
	<i>Night D3</i>	182	-0.24	0.20

**Table 9-1:** Global validation statistics from comparing ATSR-1 NR V3.0 to the available validation datasets.

### 9.4 Comparisons between AT1\_L2P SSTs and drifting buoys



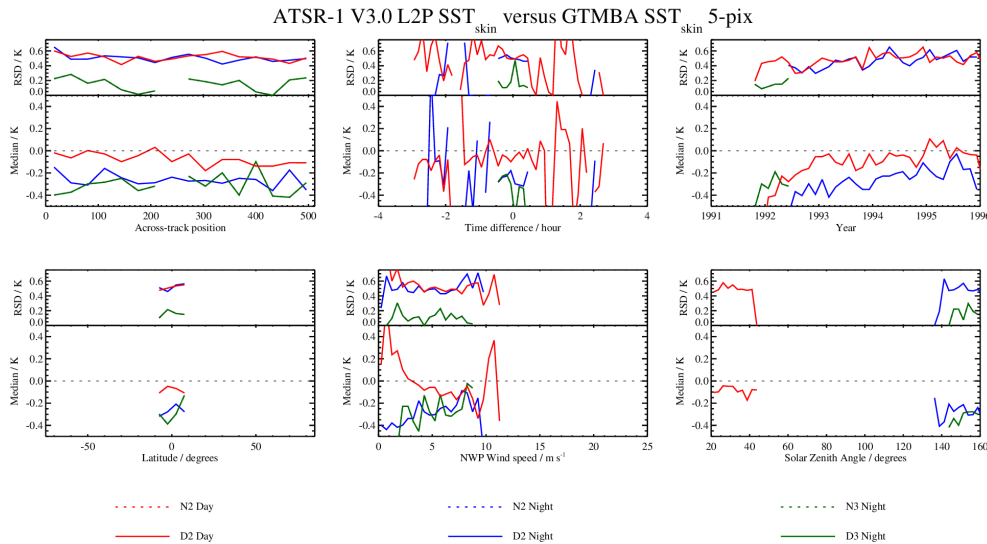
**Figure 9-5:** Dependence of the median and robust standard deviation between ATSR-1 L2P SST<sub>skin</sub> and drifter SST<sub>depth</sub> discrepancies as a function of across-track position, time difference, year, latitude, wind speed and solar zenith angle. Daytime results are shown in red, nighttime 2-channel results are shown in blue and nighttime 3-channel results are shown in green. Dual-view results are represented as solid lines and nadir-only results as dashed lines.



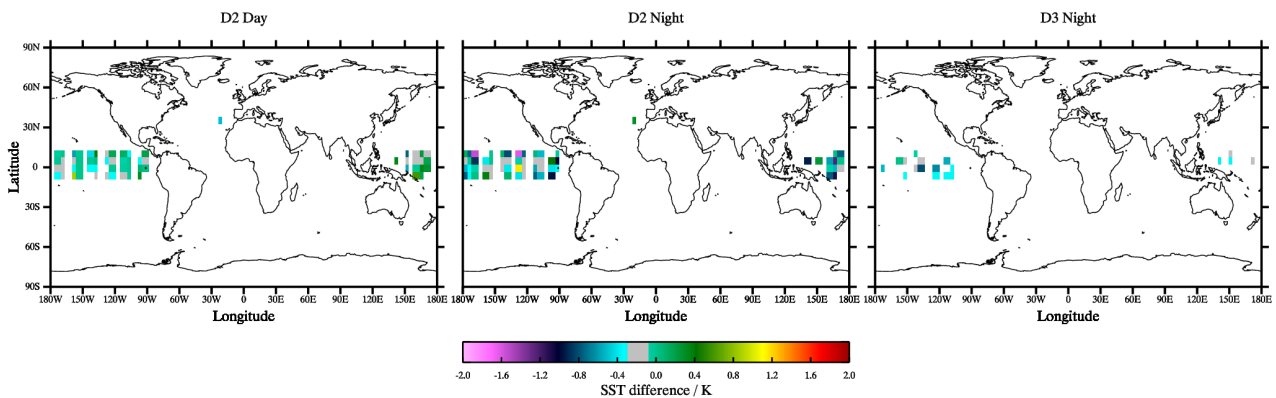
**Figure 9-6:** Spatial distribution of the median discrepancy between ATSR-1 L2P SST<sub>skin</sub> and drifter SST<sub>depth</sub>. The greyed region indicates a band of +/- 0.1 K around the expected mean skin offset of -0.17 K.



### 9.5 Comparisons between AT1\_L2P SSTs and GTMBA



**Figure 9-7:** Dependence of the median and robust standard deviation between ATSR-1 L2P SST<sub>skin</sub> and GTMBA SST<sub>depth</sub> discrepancies as a function of across-track position, time difference, year, latitude, wind speed and solar zenith angle. Daytime results are shown in red, nighttime 2-channel results are shown in blue and nighttime 3-channel results are shown in green. Dual-view results are represented as solid lines and nadir-only results as dashed lines.



**Figure 9-8:** Spatial distribution of the median discrepancy between ATSR-1 L2P SST<sub>skin</sub> and GTMBA SST<sub>depth</sub>. The greyed region indicates a band of +/- 0.1 K around the expected mean skin offset of -0.17 K.



## 9.6 Statistical analysis of AT1\_L2P results

Reference	Retrieval	Number	Median (K)	RSD (K)
<b>Drifters</b>	<i>Day D2</i>	24782	-0.08	0.53
	<i>Night D2</i>	15920	-0.25	0.52
	<i>Night D3</i>	1214	-0.35	0.35
<b>GT MBA</b>	<i>Day D2</i>	24131	-0.07	0.52
	<i>Night D2</i>	2555	-0.27	0.50
	<i>Night D3</i>	91	-0.30	0.16

**Table 9-2:** Global validation statistics from comparing ATSR-1 L2P V3.0 to the available validation datasets. L2P format files contain only the best SST available and so do not contain nadir-only SSTs.