The EUMETSAT
Network of
Satellite Application
Facilities

OSI SAF
Ocean and Sea Ice

Validation Report for the Atlantic High Latitude L3 Sea Surface Temperature product

OSI-203

Version 1.0

Steinar Eastwood
Norwegian Meteorological Institute

26. April 2011

This page is intentionally left blank.

Document Change Record

Version	Date	Change	Description	Responsible
1.0	26-04-2011		Validation of extended 5km product	Steinar Eastwood

Table of Contents

1 Introduction

1.1 Scope

The purpose of this report is to document the level of agreement between the new EUMETSAT Ocean and Sea Ice Satellite Application Facility (OSI SAF) L3 5km Atlantic High Latitude (AHL) Sea Surface Temperature (SST) product and in situ observations. This new L3 SST product is replacing the high latitude part of the existing 10km Merged Atlantic Product (MAP). Because these two products are very similar in nature and the 10km MAP product has been previously validated, this validation report is not very extensive.

1.2 Overview

The EUMETSAT OSI SAF is producing a range of operational air-sea interface products, namely: wind, sea ice characteristics, Sea Surface Temperatures (SST) and radiative fluxes, Surface Solar Irradiance (SSI) and Downward Longwave Irradiance (DLI). More details on the products and OSI SAF project are available at http://www.osi-saf.org.

SST, SSI and DLI products from the OSI SAF are produced using geostationary and polar orbiting satellites and are available in level 2 and level 3 formats, with different timeliness depending on the production setup.

A specific L3 High Latitude SST product is produced at METNO covering the North Atlantic High Latitudes north of 50N. This product used to be a part of the the MAP SST products, but the MAP product is superseded by a Low and Mid Latitude product based on geostationary data and a High Latitude product based on AVHRR polar orbiting data.

The HL SST products are derived from AVHRR polar orbiter data received at the local receiving station at METNO together with data received through the EUMETCast ATOVS Retransmission Service (EARS). Intermediate L2 SST products for each pass are input to the HL L3 SST product that are delivered every 12 hours. These 12-hourly products are available in HDF5 and GRIB format through the OSI SAF High Latitude FTP server (ftp://osisaf.met.no/prod). See also http://osisaf.met.no for product monitoring, validation , news messages and other information.

This report is separated in chapters describing the in situ validation data and the results obtained. The OSI SAF AHL SST product is not described, users are referred to the Product User Manual ([RD.1]).

1.3 Glossary

Acronym	Description
AVHRR	Advanced Very High Resolution Radiometer
CMS	Centre de Méteorologie Spatiale
DLI	Downward Longwave Irradiance

Acronym	Description
DMI	Danish Meteorological Institute
GRIB	Gridded Binary Format
HDF	Hierarchical Data Format
HL	High Latitudes
HIRLAM	High Resolution Limited Area Model
LML	Low and Mid Latitudes
METNO	Norwegian Meteorological Institute
MODTRAN	Moderate Resolution Transmittance model
MSG	Advanced Very High Resolution Radiometer
NOAA	National Oceanic and Atmospheric Administration
NAR	Near Atlantic Regional
NWC	Nowcasting
RMDCN	Regional Meteorological Data Communication Network
SAF	Satellite Application Facility
SMHI	Swedish Meteorological and Hydrological Institute
SSI	Surface Solar Irradiance
SST	Sea Surface Temperature
TIGR	Tovs Initial Guess Retrieval database

1.4 Applicable documents

[RD.1] Atlantic High Latitude L3 SST Product User Manual, v2.0.

[RD.2] HL SST matchup database format, v1.1.

[RD.3] OSI SAF CDOP Product Requirement Document, v1.2.

[RD.4] Validation of the High Latitude SST product for NOAA 16 from the OSI SAF, Steinar Eastwood, September 2002.

2 Validation data and strategy

2.1 Overview

The validation of the Atlantic High Latitude SST products is performed on the matchup databases (MDB) for high latitudes built in the OSI SAF as part of the production chain. There is one MDB for the intermediate L2 SST product and one for the final AHL product. The format of the MDBs are described in [RD.2] and will not be further described here.

2.2 Brief description of datasets

The main source for validation is comparison with drifting buoys, received locally at METNO through the GTS network. Drifting buoys have shown to be the most reliable source of validation. Ships are less reliable due to the variable way of observing SST, both in observing method and observation depth. Moored buoys are less common at high latitudes compared to low and mid latitudes, and only a couple of buoys are know at high latitudes.

The number of available drifting buoys at Atlantic high latitudes varies a lot. During special campaigns there can be many buoys available, while in “normal” years quite few are deployed. This is illustrated in Figure 1, which shows that very few drifting buoys were available between 2000 and 2007. Also, not many of the buoys before 2000 reached north of 65N. Fortunately, the Poleward project has been deploying around 150 buoys in the period between summer 2007 and spring 2010, all north of 60N. This has provided very good conditions for validation of the AHL SST product. More information about the Poleward project is available here:

(http://folk.uio.no/ingako/my_files/POLEWARD_WEBPAGE_MAIN.html)

The drifting buoys collected at METNO are quality controlled and then collocated with the SST product closest in time. For the L2 SST MDB additional data such as AVHRR channel values, satellite and sun geometry etc. are also collocated. In the collocation the in situ SST is matched to the closest pixel in the SST product closest in time. Only matchups that are closer than two hours are kept for the L2 SST product, and closer than 6 hours for the 12-hourly L3 SST product. There is an inconsistency between the instant observing in situ observation and the average 12-hourly AHL SST product, as they are not representing the same temperature. But when comparing longer time scale as done here (monthly and yearly), this difference is reduced.

For the L2 comparisons the duplicate observations have been removed, that is only the closest-in-time matchup is kept for each in situ platform when there are several in situ observations within the time span (+/- 2 hours) for each satellite passage. For the L3 comparisons the duplicate observations have been kept, to reduce the difference mentioned in the previous paragraph.

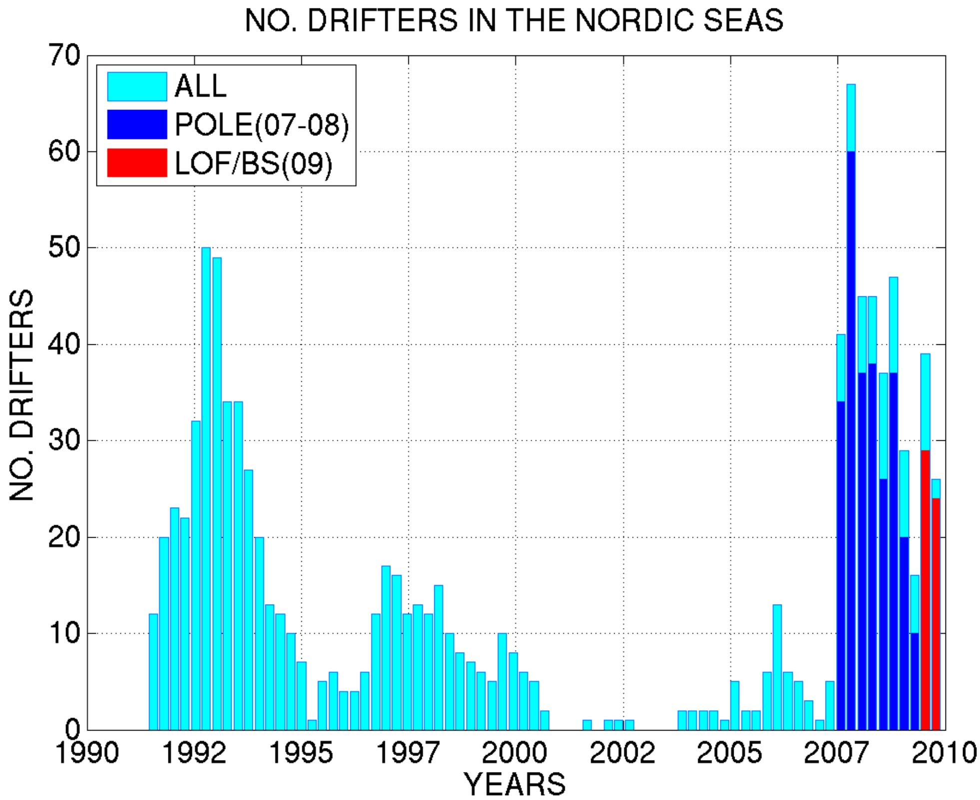


Figure 1: Number of drifting buoys available in the Nordic Seas since 1991. Dark blue and red are buoys deployed by the Poleward project. Light blue are buoys deployed by other programs. Courtesy of Inga Koszalka, University of Oslo, Norway.

During the collocation of in situ and satellite data, both the values for the matching pixel and an average of all cloud free pixels in a 15x15pixels box around the matching pixel are kept for the L2 SST MDB.

For this validation study the period of January to December 2010 have been used. This period was chosen to have a full year with all different conditions.

2.3 Target accuracy

The target accuracy is defined as the bias and standard deviation of the primary SST calculations when comparing with drifting buoy measurements, determined on a monthly basis using night time data (see [RD.3] for details):

- Bias: the biases should remain within +/-0.5°C by night on a monthly basis. By day at summer local biases of 1-2 K are possible.
- Standard deviation: the standard deviation should be lower than 0.8°C by night on a monthly basis.

The number of available drifting buoys at Atlantic high latitudes vary. The number of deployments at high latitudes have been very few for periods, the area is often cloudy and at summer there night is very short or none existent north of the polar circle for periods. For these reasons it can be difficult to determine representative night time validation results to assess the target accuracy. Therefor both night time, twilight and daytime validation results are monitored and reported.

3 Validation results

The validation results are split in two sections. First the results for the final L3 AHL SST product are presented. The main focus is on this validation. Then validations results for the intermediate L2 SST product are shown to include more detailed validation results to conclude on the algorithm performance.

3.1 L3 AHL SST validation results

The validation experiment has been run for one full year (2010.01 to 2010.12) to assess the quality of the final product. First the difference in using the ice and cloud probability masks (described in [RD.1] , chapter 3.2.6), and the results are shown in Table 1. The results are split in the 00UTC product which is dominated by night time data, and the 12UTC product which is dominated by day time data. Still, at high latitudes the length of night varies strongly with time of year. At mid summer time there are few nigh time data available, and at mid winter there are few day time data available. The results improve when including the ice and cloud probability masks, both for the 00UTC and the 12UTC product. The 12UTC product performs better than the 00UTC product, both in terms of bias and standard deviation.

	00UTC product			12UTC product		
	Bias	Std.dev	Num	Bias	Std.dev	Num
No prob masks	-0.52	0.86	64258	-0.41	0.73	76524
With prob masks	-0.39	0.75	54011	-0.29	0.63	68911

Table 1: Validation results for 2010.01-2010.12 for 12 hourly 00 and 12UTC AHL SST product including all quality levels. Validation results are shown both running the production without and with the ice and cloud probability masking. Num is number of matchup collocations used.

The rest of the validation results shown in this chapter are all including the ice and cloud probability masks. In Table 2 are shown the results when splitting on confidence level. The confidence level varies from 2 (representing the worst quality) to 5 (representing the best quality), and are further described in [RD.1] , chapter 4.3.2. In this comparison the mean of the confidence level in a 15x15 pixels box surrounding the central SST value have been used. Value 5 represent an area with only confidence level 5 pixels, and with the highest quality values. The results show that the quality degrades with decreasing confidence level, with increasing standard deviation and more negative bias. This is more pronounced for the 00UTC product than the 12UTC product.

Confidence level	00UTC product			12UTC product		
	Bias	**Std.dev**	**Num**	**Bias**	**Std.dev**	**Num**
5	-0.06	0.62	12132	-0.03	0.55	17599
4.5	-0.42	0.73	12417	-0.27	0.65	17186
4	-0.57	0.77	12217	-0.41	0.65	13515
3.5	-0.56	0.70	10993	-0.46	0.60	12545
3	-0.58	0.70	6467	-0.49	0.53	7736
2.5	-1.27	0.90	260	-0.55	0.82	655
2	-1.91	0.97	41	-0.76	0.64	35

Table 2: Validation results for 2010.01-2010.12 for 12 hourly 00 and 12UTC AHL SST products, split on confidence level. The confidence level values are the mean for the 15x15 pixels surrounding the central SST value. Value 5 represent an area with only confidence level 5 pixels.

The validation results have also been split in months to monitor the variation from month to month. The results are shown in Figure 2 for the 00UTC product and in Figure 3 for the 12 UTC product. Only the confidence level 4 and 5 have been included. This is mainly to check whether the the target accuracy stated in the Product Requirement Document ([RD.3]), which is described in 2.3 , are met for the high quality data in the product. The 12UTC product has more variation in the standard deviation compared to the 00UTC product, while less variation in bias.

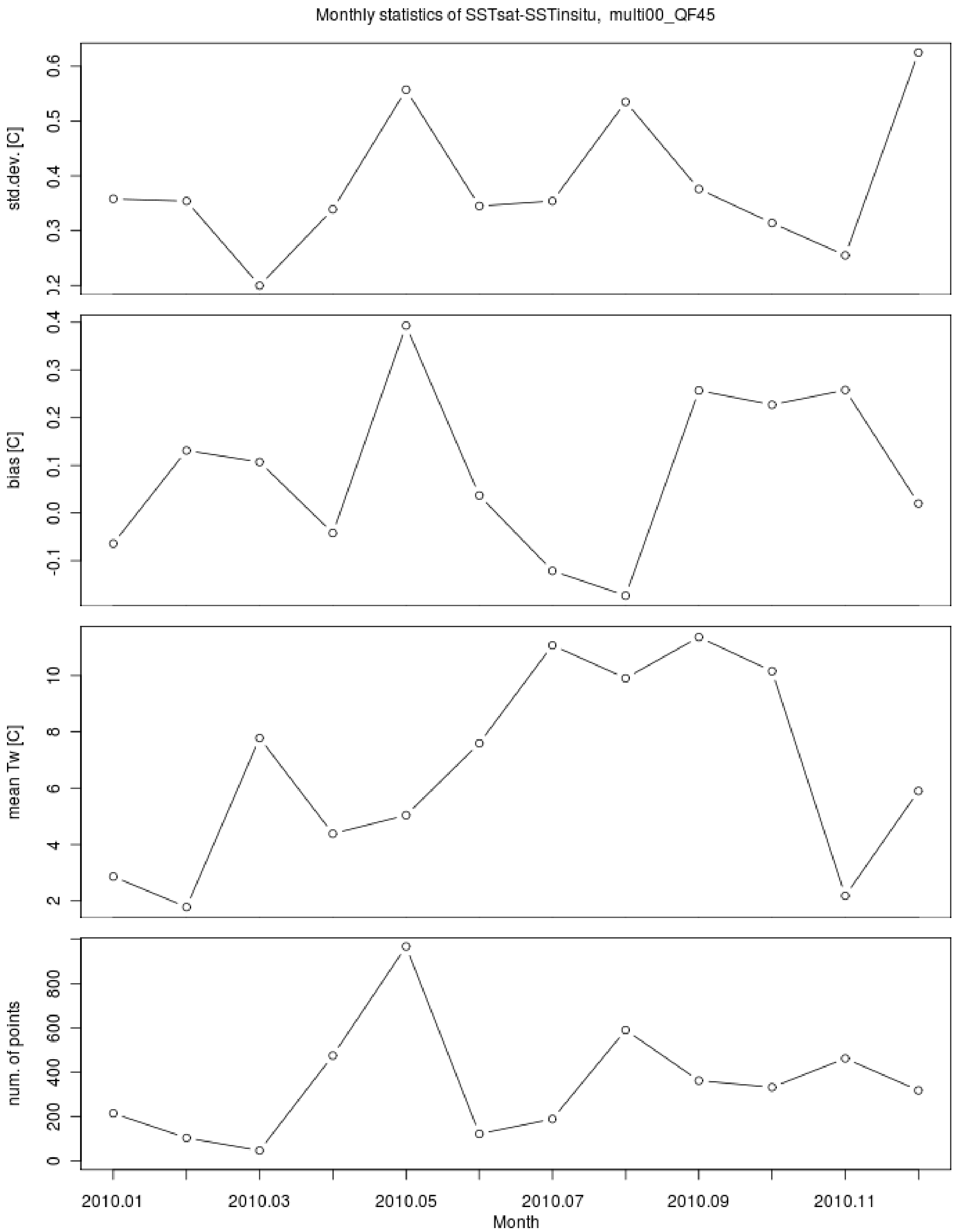


Figure 2: Monthly validation results for 2010.01-2010.12 for the 12 hourly AHL SST product, centred on 00 UTC. Only values with quality level 4 and 5 are included. Top panel shows standard deviation, second figure bias, third figure mean in situ SST and lower panel the number of validation points used.

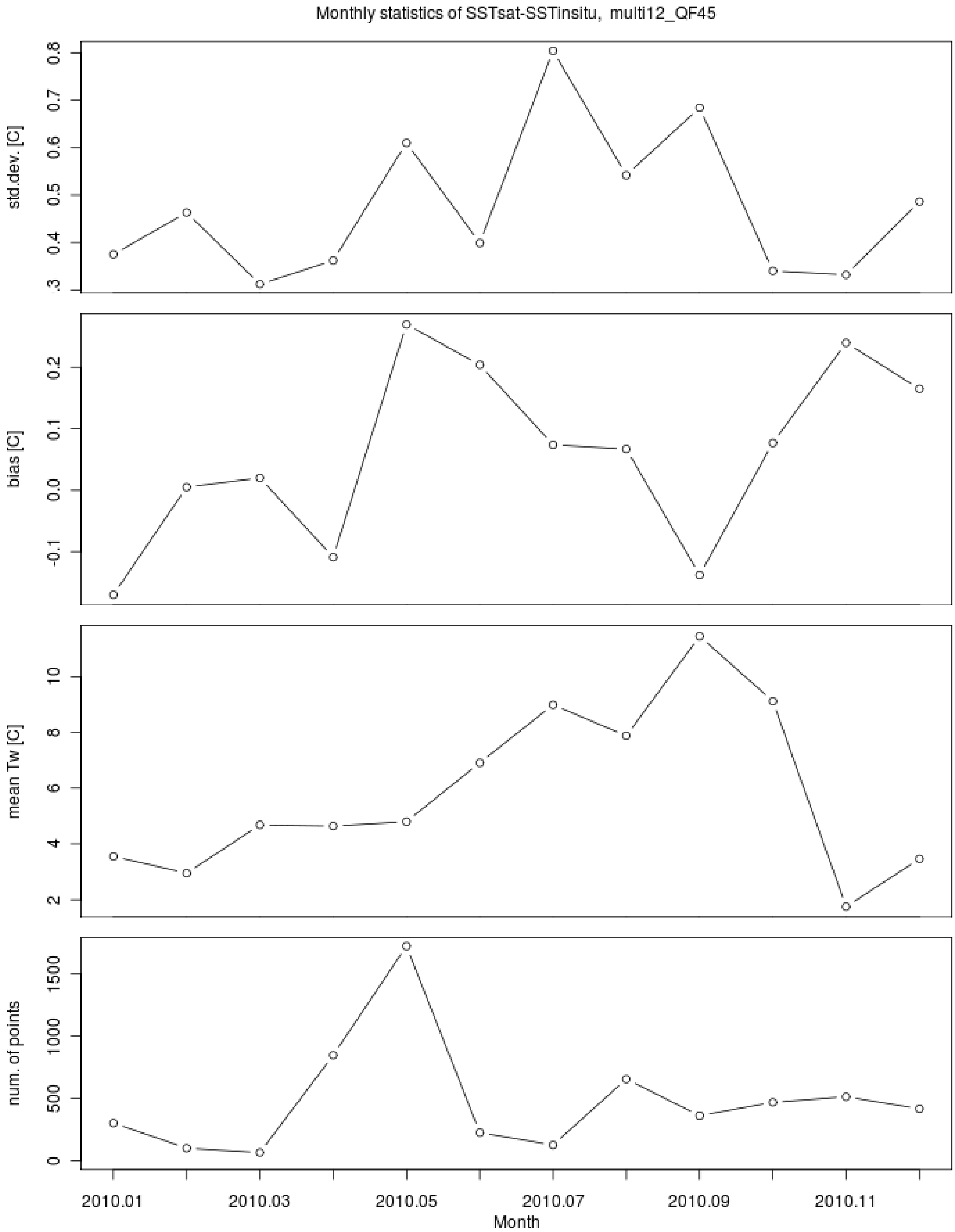


Figure 3: Monthly validation results for 2010.01-2010.12 for the 12 hourly AHL SST product, centred on 12 UTC. Only values with quality level 4 and 5 are included. Top panel shows standard deviation, second figure bias, third figure mean in situ SST and lower panel the number of validation points used.

3.2 L2 SST validation results

The intermediate L2 SST product is not distributed and the validation results presented here are only shown to document the performance of the intermediate steps of the AHL processing chain. All results shown are after the ice and cloud probability mask has been applied. All results are from comparing the average satellite SST in a 15x15 pixel box around the central pixel, except in Table 3, where both average and central pixel are given. Only matchups where the cloud cover is less than 10% within the 15x15 pixels are used. This is done to minimize the number of outliers caused by cloud masking problems, so that the results addresses the SST algorithm performance without including possible problems with cloud masking. The validation has been performed on the one year period from November 2009 to October 2010. This period is slightly different from the period used for the L3 validation, as the validation results for the last two months of 2010 were lost by a mistake.

The L2 SST validation is not done extensively since the L2 SST algorithm has been validated previously, e.g. in [RD.4] . Only results for NOAA 18 are provided in this report. Regularly validation results for all satellites included in the AHL SST product will be provided in the Quarterly Reports from the OSI SAF.

The overall results for NOAA 18 are shown in Table 3, for day time, night time and twilight conditions. Twilight is here defined at the period when the solar zenith angle is between 85 and 95 degrees. The day time standard deviations are a bit lower compared to the night time results. Twilight conditions gives the lowest standard deviations, but with much less matchups than at day time and night time to make this a significant difference. The difference between day time and night time bias is on average 0.16°C. The results for average and central pixel comparison are very similar, and therefore only average comparisons are used hereafter.

The corresponding monthly validation results for day time and night time are shown in Figure 4 and Figure 5, when only confidence levels 4 and 5 are included. Usually the standard deviation is below 0.5°C and bias within +/- 0.2°C. The night time results show a few more cases with standard deviation higher than 0.5°C compared to day time, while there is only one month with bias outside +/- 0.2°C for both day time and night time.

	Average SST			**Center pixel SST**		
	Bias	**Std.dev**	**Num**	**Bias**	**Std.dev**	**Num**
Day time	0.07	0.52	1297	0.09	0.52	1297
Night time	-0.09	0.59	798	-0.07	0.6	796
Twilight	-0.03	0.50	110	-0.04	0.56	110

Table 3: Validation results for 2009.11-2010.10 for intermediate L2 SST product from NOAA18 only, including all quality levels. Ice and cloud probability masks have been applied. Average is the results when averaging all available satellite SST in a 15x15 pixel box around the central value, while central pixel is when using only the central value. Num is the number of matchup collocations used.

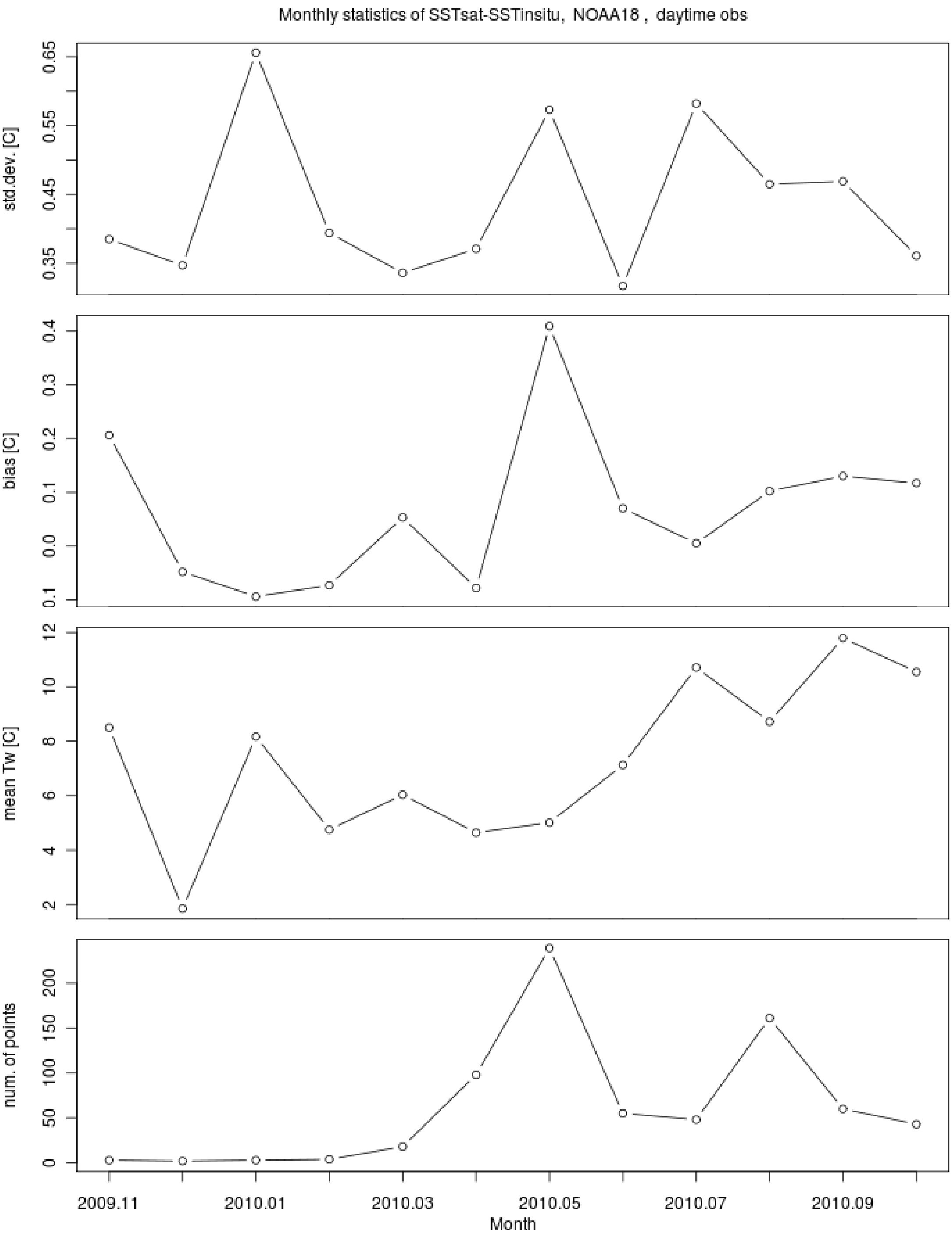


Figure 4: Monthly day time validation results from 2009.11-2010.10 for the intermediate L2 SST product, with only NOAA 18 data. Top panel shows standard deviation, second figure bias, third figure mean in situ SST and lower panel the number of validation points used. More details provided in the text.

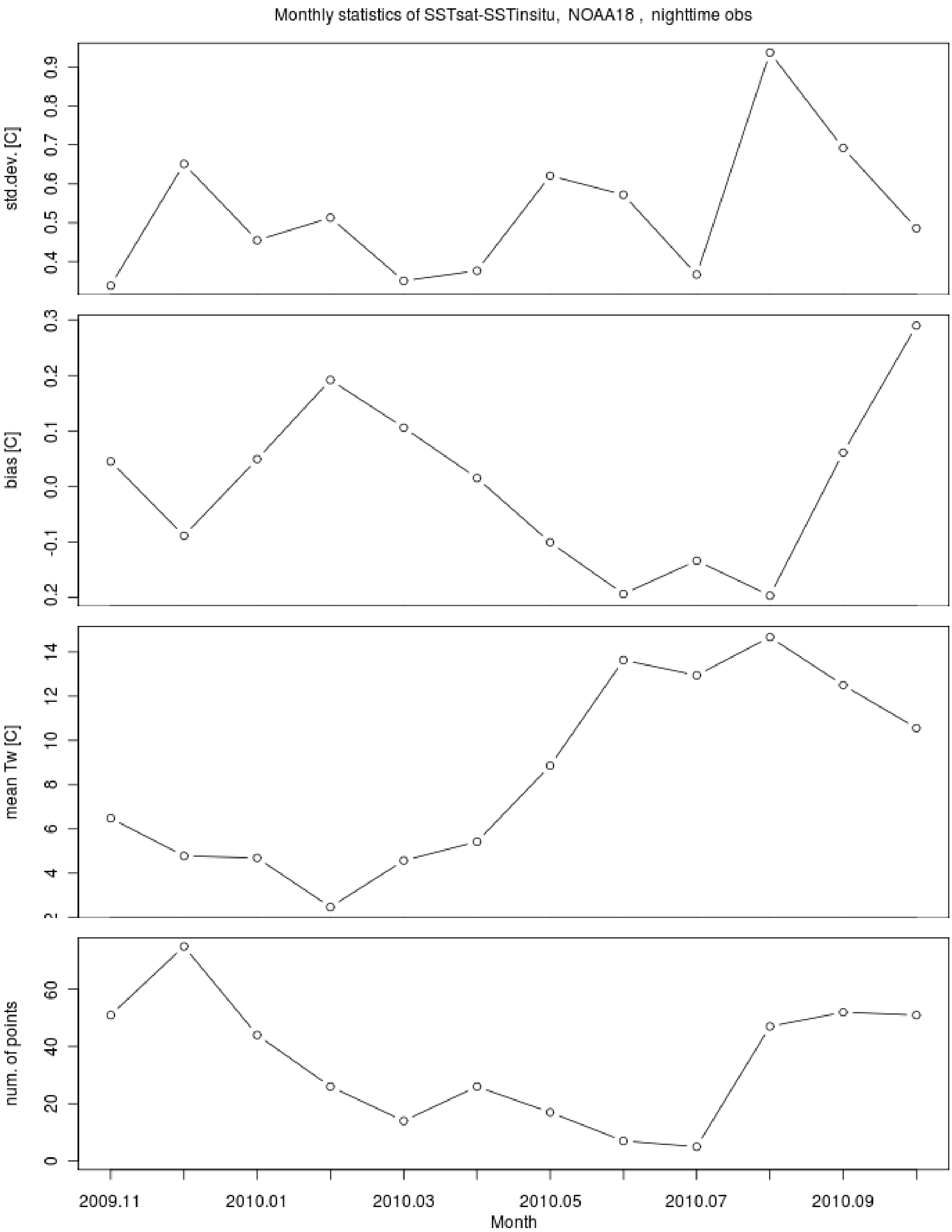


Figure 5: Monthly night time validation results from 2009.11-2010.10 for the intermediate L2 SST product, with only NOAA 18 data. Top panel shows standard deviation, second figure bias, third figure mean in situ SST and lower panel the number of validation points used. More details provided in the text.

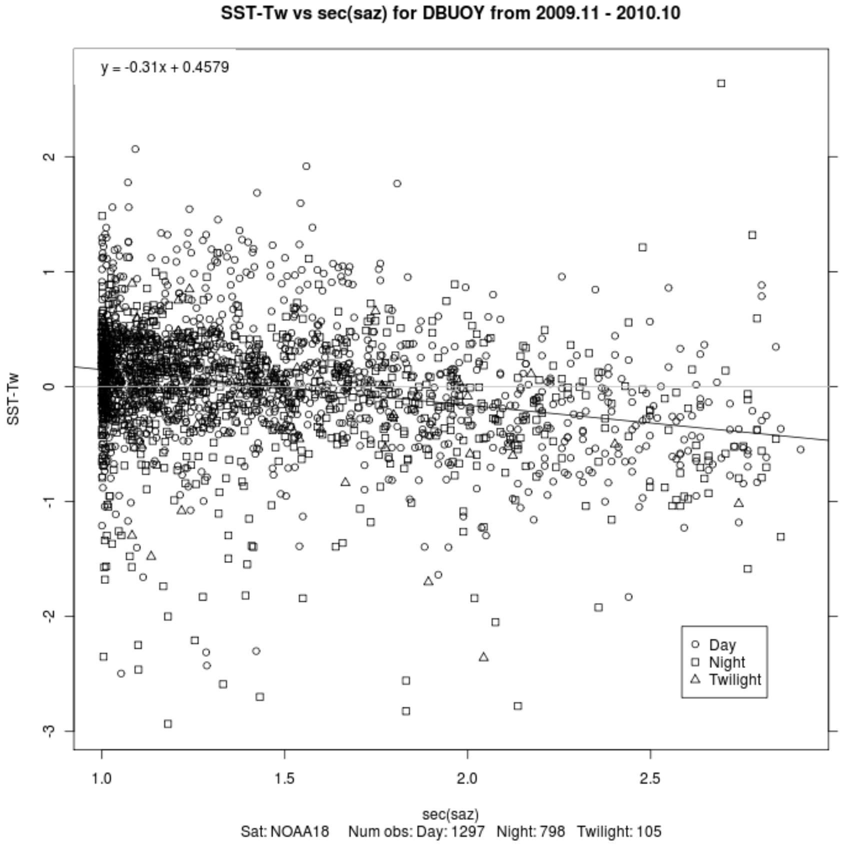
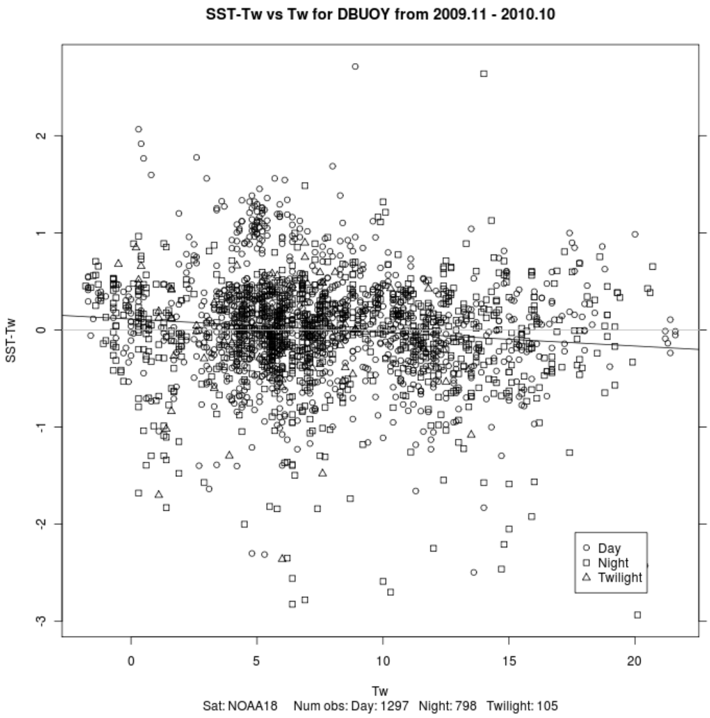


Figure 6: Scatter plot of the difference between average satellite SST and observed in situ SST (Tw) versus secant of satellite zenith angle. For NOAA 18 data in the period 2009.11 – 2010.10. Different symbols are used for day time, night time and twilight.

The scatter plots in Figure 6 and Figure 7 show how the difference between satellite SST and in situ SST varies with in satellite zenith angle, situ SST and latitude, respectively. In Figure 6 there is an increasing negative difference with increasing satellite zenith angle. Figure 7 show little variation in the temperature difference with temperature and latitude, except possibly an area with positive differences around 5°C and 73N.



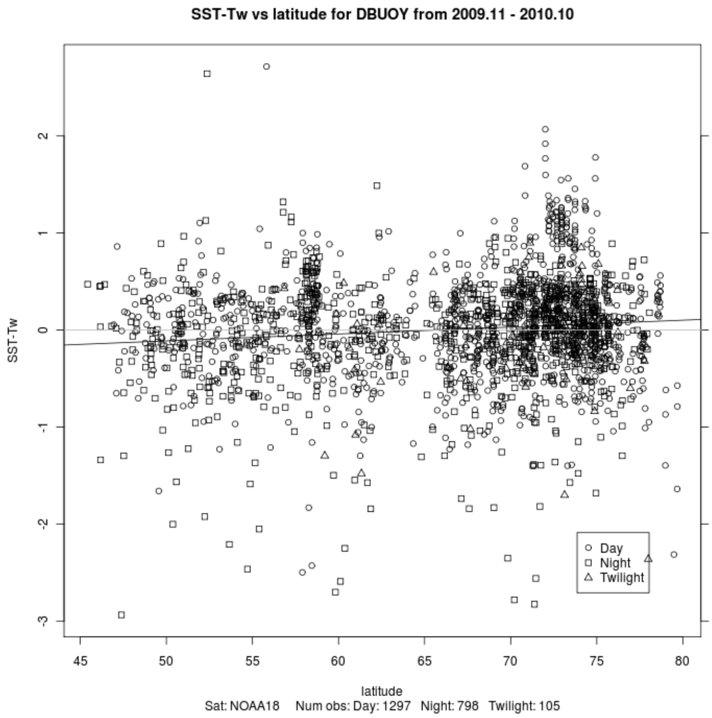


Figure 7: As Figure 6, but for in situ SST (upper) and latitude (lower).

4 Discussion

4.1 L3 AHL SST product

The validation of the L3 AHL SST product has been performed using drifting buoys and the results are shown in chapter 3.1 .

The usefulness of the additional ice and cloud probability masks has been demonstrated. As seen in Table 1, the standard deviation is reduced and the bias improved when the masks are applied, with equal impact on the 00UTC and 12UTC product. This illustrates that there is a need for additional ice and cloud masking when using the NWC SAF PPS software for SST retrieval purposes.

The SST performance as function of confidence level is shown in Table 2, and illustrates how the quality of the SST product degrades as the confidence level decreases. An increasing negative bias is observed with decreasing confidence level, most probably due to undetected clouds and ice. Similarly, the standard deviations increase. This is useful information for the users, and users must take this into account when selecting which SST estimates to use. Users have different preferences on quality, and therefore the quality level information is very useful.

The monthly validation results shown in Figure 2 and Figure 3 shows that there is no particular seasonal pattern in the performance. The number of available in situ observations do vary and is due to varying cloud conditions and varying length of day and night at high latitudes. These effects could contribute to hiding monthly patterns in the bias and standard deviation. Anyway, all the monthly values are within the target accuracy of +/-0.5°C mean difference and 0.8°C standard deviation, when using confidence levels of 4 and 5.

4.2 L2 SST intermediate product

The validation of the L2 SST intermediate product has also been performed using drifting buoys and the results are shown in chapter 3.2 . The main purpose of validating both the L2 and L3 products is to see that the results are consistent between the level 2 and level 3 processing, and in particular for the L2 validation to check the algorithm performance in detail and if there are any noticeable patterns in the difference between estimated and observed SST.

The results in Table 3 show that the quality on average is a bit better at day time compared to night time in terms of standard deviation. This could be caused by the cloud and ice masking steps are expected to perform better at day time than night time, due to the availability of visible channels at day time. It is also observed that on average the difference between day time and night time bias is 0.16°C. This is probably caused by the skin temperature cooling at night time and skin warming at day time. This difference can be much higher in particular cases, during low wind and clear sky conditions when diurnal warming can reach day-night differences of several degrees.

Figure 4 and Figure 5 shows the monthly variation in the validation results and as for the monthly L3 results, there are no particular systematic patterns, except maybe a tendency against a negative night time bias during the summer months.

Figure 6 and Figure 7 shows details in the validation results when plotting against the satellite zenith angle, in situ observed SST (Tw) and latitude. In Figure 6 the difference in estimated and observed SST becomes increasingly negative at high satellite zenith angles. The quality of the SST estimates at high satellite zenith angles are generally know to be less accurate, especially due to the reduced resolution, and also for this product. Still, further investigations will be performed to check whether the systematic difference with satellite zenith angle can be improved through algorithm correction.

In Figure 7 the majority of the observations are from temperatures below 10°C and north of 65N. This high number of validation data at very high latitudes is unusual and is due to the Poleward project deploying many buoys in the Barents Sea (see chapter 2.2). The plots show little variation in the temperature difference with temperature and latitude, except possibly an area with positive differences around 5°C and 73N.

5 Conclusion

The validation of the upgraded 5km AHL SST product shows that the performance is as expected in terns of bias and standard deviation, when comparing with previous validation results for the 10km HL SST product. The target accuracy of 0.8°C standard deviation and within +/- 0.5°C bias on a monthly basis are met when the final product estimates with high confidence level (level 4 and 5) are compared.