SST VALIDATION OVER THE BALTIC SEA

Claus Sølvsteen, Farvandsvæsenet
Overgaden oven Vandet 62 B, 1023 Copenhagen K, Denmark
Phone: +45 32 68 96 25
E-mail: cls@frv.dk

ABSTRACT

The O&SI SAF SST product from the MNOR area (including the Baltic Sea) has been validated against temperature measurements from ships, buoys and moored stations. Insitu observations cover primarily the southern part of the MNOR area. The validation period is from June 24th 2001 to August 2nd 2003, and only the NOAA-16 SST product has been evaluated. This product has been assigned central acquisition times of 2 UTC (night) and 12 UTC (day). It is found that in the majority of the months, the OSISAF SSTs are within the expected accuracy, and that in situ temperatures measured at depths between 0 and 4-6 meters fit equally well to the OSISAF SSTs. There do not seem to be geographical variations in the accuracy of the daytime OSISAF SSTs, while there may be some minor variations for nighttime SSTs. There are seasonal variations in the accuracy of the OSISAF SSTs, especially for daytime SSTs where the OSISAF SSTs are slightly too low during the autumn and winter, and slightly too high during spring and summer. Validation results depend largely on, which of the SST quality flags between 2 and 5 that are considered. An overall mean bias (bias = "in situ temperature minus OSISAF SST") of -0.19 K (day) and 0.06 K (night) is obtained, if only quality 5 data are used, and a mean bias of 0.00 K (day) and 0.24 K (night) is obtained, if all SST quality flags are used.

1. DATASET

373,131 in situ observations (from buoys, moored stations and ships) were collected for the validation period June 24th 2001 to August 2nd 2003. The observations covered the Baltic Sea, the inner Danish waters and the North Sea. 38% of the observations were without indication of depth. They were, however, described as "SST measurements" by the data deliverer, and it is therefore assumed that they have been measured at depths between 0 and 5 meters.

The observation set has been reduced according to the following acceptance criteria:

- Only observations measured between 22 UTC and 6 UTC (night) or between 10 UTC and 14 UTC (day) are used for validation
- Observations must correspond in time and location to a cloud free grid point in the SST product
- If several measurements lie within the time window for a given SST grid point, then only the measurement closest in time to the acquisition time of the grid point is used for validation
- In situ measurements are not allowed to deviate more than 4 K from the SST value (elimination of gross errors)

This procedure left 11435 accepted observations (6609 nighttime and 4826 daytime observations) – a 97% reduction of the data material.

The locations of the accepted daytime measurements are as indicated at figure 1. The location of nighttime measurements is similar.

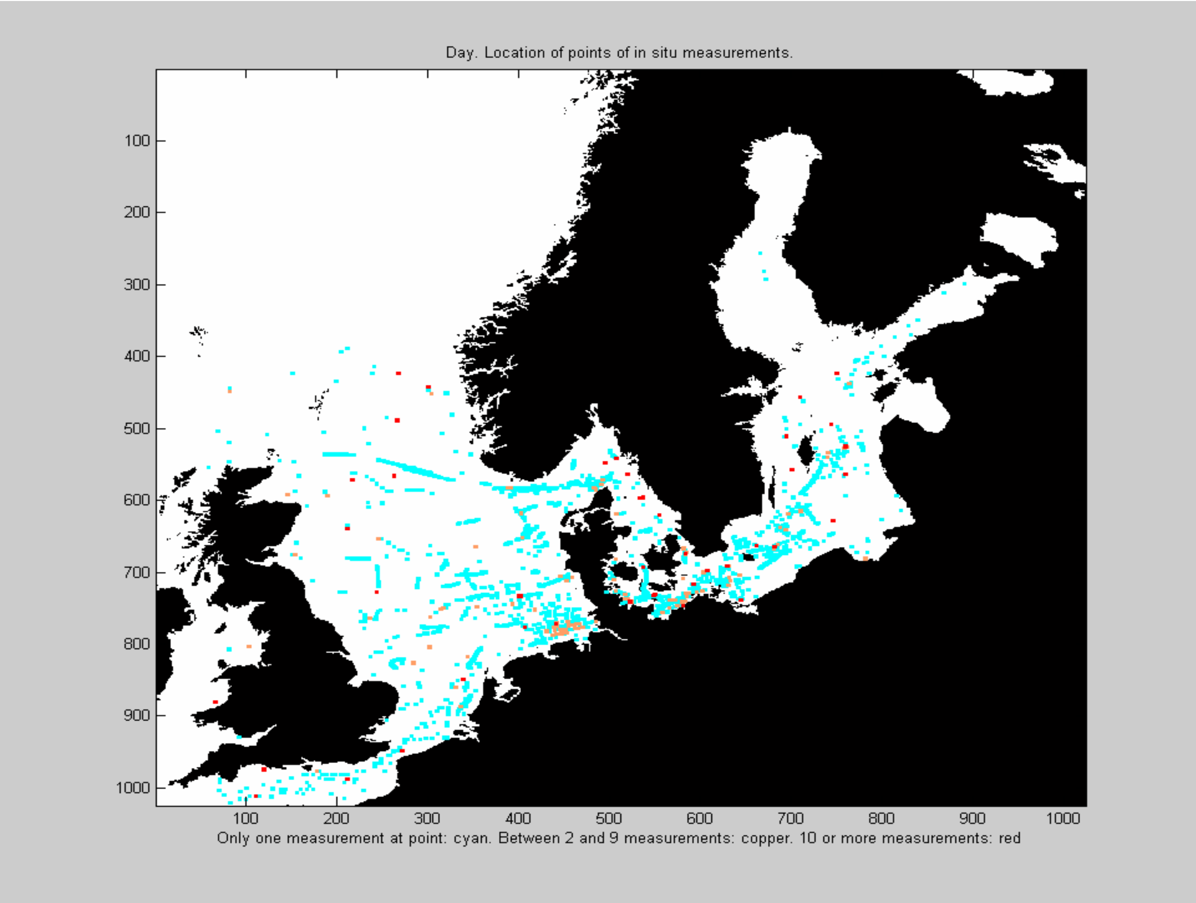

Figure 1. Location of accepted daytime measurements within the validation period. Color code: cyan: 1 measurement at the point. Copper: 2-9 measurements at the point. Red: 10 or more measurements at the point.

2. RESULTS FOR THE WHOLE DATASET

A comparison of OSISAF SSTs with in situ measurements at different depths has been performed. For each depth a mean bias and a standard deviation have been calculated for day and night observations. Bias is understood as "in situ temperature minus OSISAF temperature". Results of this analysis is shown in tables 1 and 2. It is seen that measurements taken at depths between 0 and 4-6 meters or at unknown depth fit equally well to the OSISAF SSTs (measurements at 5 meters depth fit worse than measurements at 6 meters depth, but at the same time there are only few measurements at 5 meters depth, so firm conclusions cannot be made for this depth).

Figure 2 shows an example of the distribution of biasses as a function of the time distance between in situ measurement and OSISAF SST. It is seen that the absolute biasses do not increase with increasing time distance, indicating that the selected time windows (2 hours at day and 4 hours at night) are not too large.

Depth (m)	Mean value of ‘in situ – SST’ (K)	Std. of ‘in situ – SST’ (K)	Number of observations
0	-0.07	0.74	612
1	0.08	0.70	134
2	0.14	0.86	293
3	-0.05	0.69	304
4	-0.09	0.76	1025
5	-0.37	0.93	60
6	-0.13	0.92	177
7	-1.25	1.08	17
8	-0.46	0.95	67
> 8 m	-0.53	1.57	417
Unknown	0.12	0.81	1720
Total	**-0.06**	**0.91**	**4826**

Table 1. Daytime data. Results for different depths.

Depth (m)	Mean value of ‘in situ – SST’ (K)	Std. of ‘in situ – SST’ (K)	Number of observations
0	0.26	0.81	832
1	0.18	0.36	120
2	0.17	0.57	329
3	0.20	0.47	329
4	0.18	0.53	1900
5	0.06	0.27	28
6	0.19	0.65	189
7	-0.41	0.76	16
8	0.05	0.68	72
> 8 m	-0.17	1.22	218
Unknown	0.29	0.69	2576
Total	**0.22**	**0.68**	**6609**

Table 2. Nighttime data. Results for different depths.

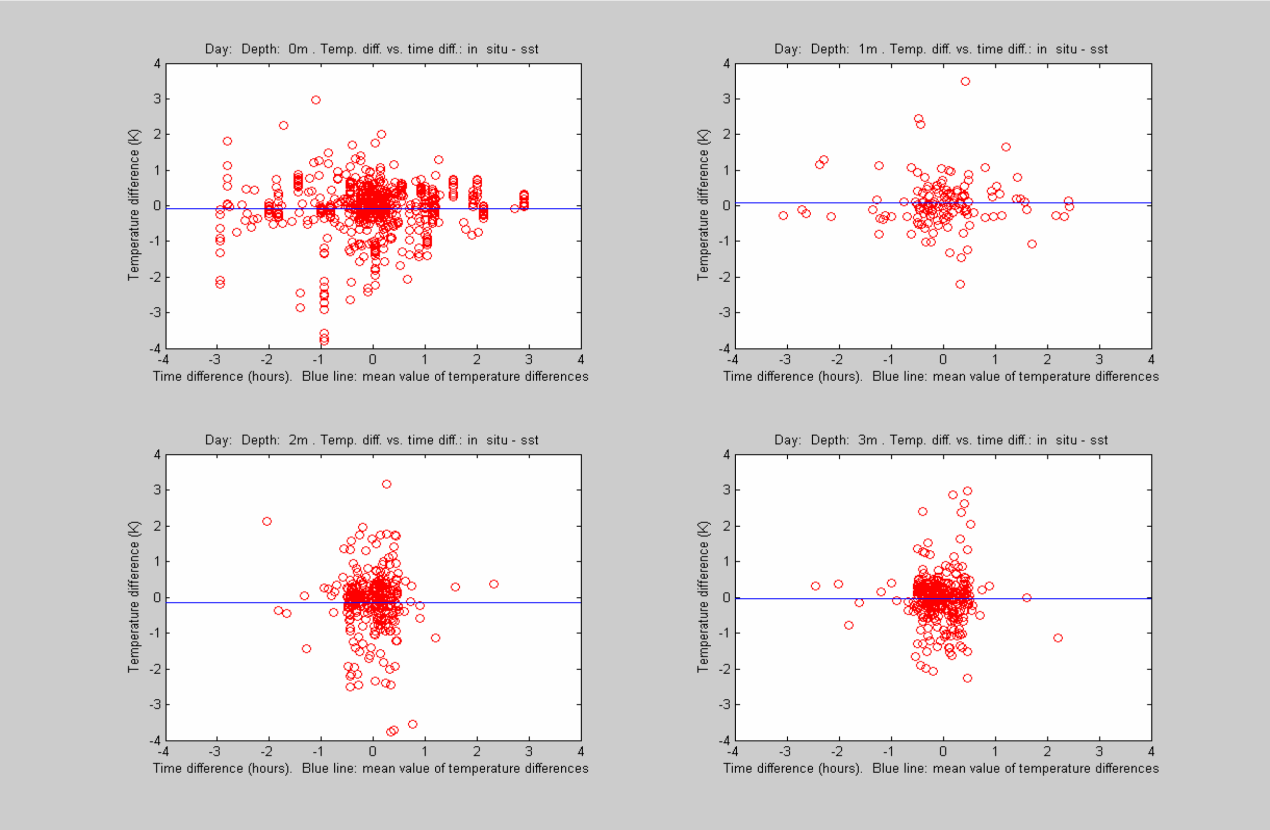

Figure 2. Daytime results. Temperature difference as a function of time distance between in situ measurement and OSISAF SST. Results for different depths: 0m, 1m, 2m and 3m. Results for larger depths and for nighttime data are similar.

Based on these results we have chosen to accept the chosen time windows and to look only at in situ measurements taken between 0 and 5 meters depth or taken at an unknown depth. 4148 daytime and 6114 nighttime measurements fullfill this criterion. For the rest of this paper we only consider these depths.

Mean biasses and standard deviations for these depths are:

Mean:	0.00 K (day)	0.24 K (night)
Std:	0.79 K (day)	0.64 K (night)

Since it is expected that biasses are season dependent due to variation of diurnal effects during the year, the mean bias and the standard deviation has been calculated for each month (figures 3 and 4). According to the product manual for the NAR SST product (http://www.osi-saf.org/biblio/docs/ss1_pmnarsst_1_5.pdf) the expected accuracy for nighttime mean biasses is ± 0.4 K for each month, and the standard deviations are expected to be lower than 0.8 K for each month (both day and night). These limits are marked with blue lines in the figures (a 0.4 K line is also drawn in the daytime bias-graph, though there is no expected accuracy for this product). It is seen that with a few exceptions the expected accuracies are fullfilled.

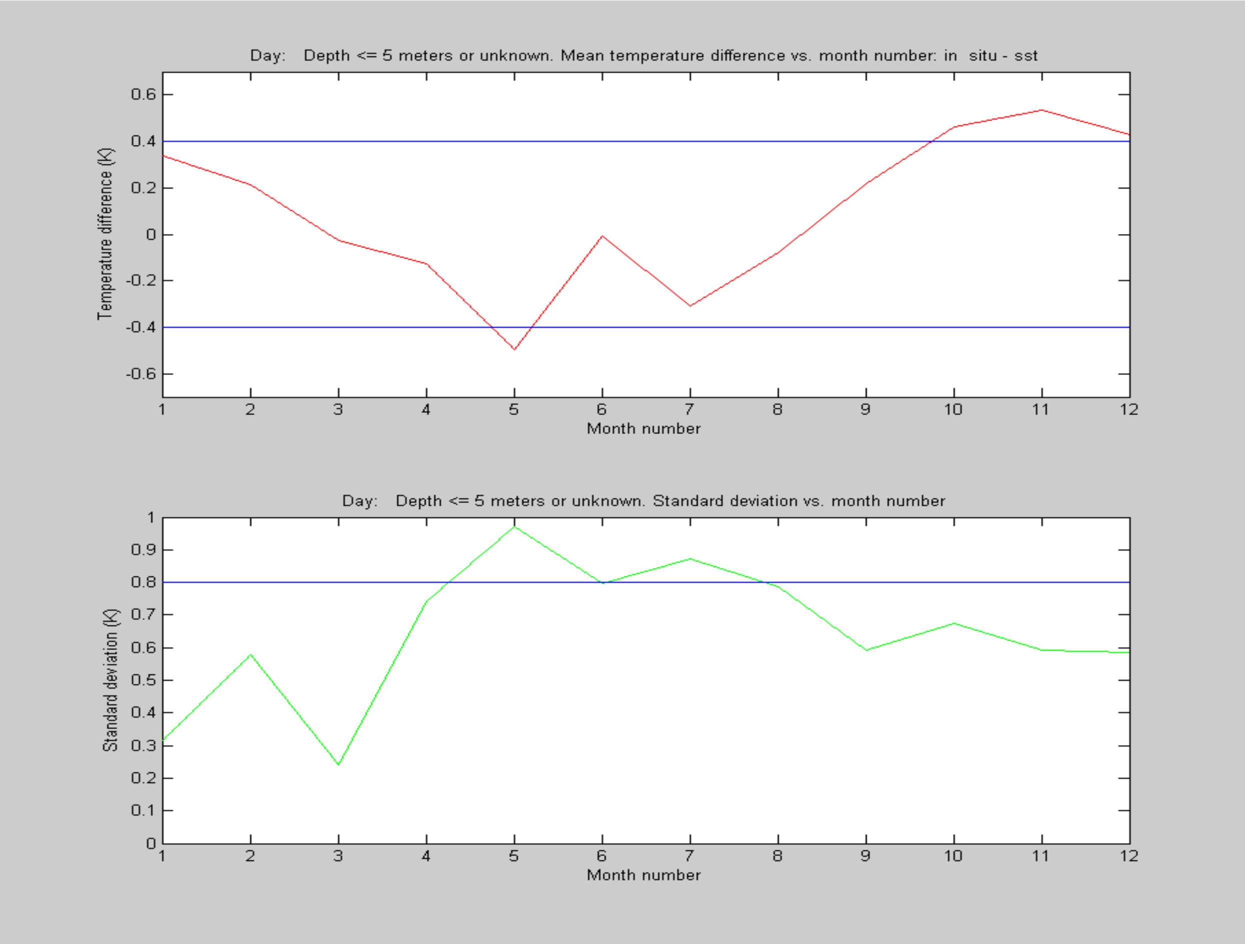

Figure 3. Daytime results as function of month. Mean bias (upper graph) and standard deviation (lower graph).

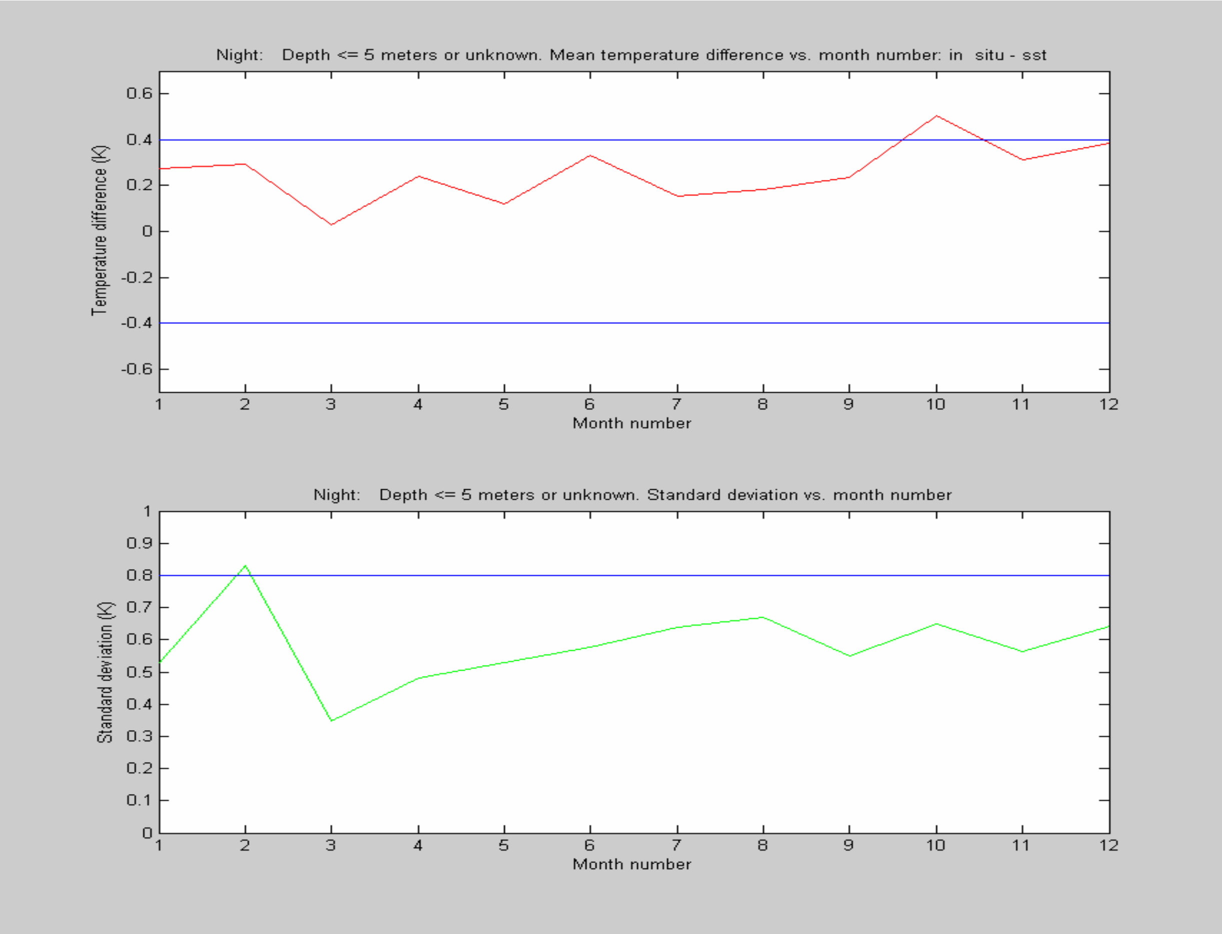

Figure 4. Nighttime results as function of month. Mean bias (upper graph) and standard deviation (lower graph).

Since there could be geographical variations in the accuracy of the OSISAF product, this topic is examined too. The results are not given here, but they show that there are no significant geographical differencies in the mean biasses and standard deviations for daytime data, but that there may be some minor differencies for nighttime observations (the Western part of the MNOR area may be slightly worse than the Eastern part). If a geographical nighttime problem should be confirmed, more data are needed.

3. RESULTS FOR DIFFERENT SST QUALITY STAMPS

During the processing of the OSISAF SSTs each grid point in the MNOR area is assigned a quality value between 0 and 5. The significance of the quality values is:

5: “Excellent”	➔	No problem
4: “Good”	➔	Close to the minimum climatologic value
3: “Acceptable”	➔	Close to a cloudy pixel
2: “Bad”	➔	Close to the minimum climatologic value and close to a cloud
1: “Erroneous”	➔	No cloud classification or failure in SSI calculations
0: “Unprocessed”	➔	Out of area, land, no satellite data

Note that SSTs with quality stamps 2 or 4 are known to be low, while SSTs with quality stamp 3 are only geographical close to a cloud and therefore not necessarily cold.

Results from a comparison between quality stamp and mean bias/standard deviation are shown in Table 3. It is seen that there are large differencies dependent on quality stamp, especially for the daytime data. Mean bias is lowest for quality 5 data (day and night), implying that the OSISAF temperatures are higher (compared to in situ SSTs) for quality 5 SSTs than for lower quality SSTs. Quality 3 and 5 data are more frequent than quality 2 and 4 data – especially for daytime SSTs. The large quality dependent variations show, that it is of great importance which SSTs are accepted in a final SST product. If only quality 5 data are used, then the results will be very different from the result if all quality flags are used – on the other hand, the number of valid SSTs will be reduced by more than 55%.

Quality	Mean temp. diff.		Std. of temp. diff.		Number of obs.		Percentage of obs.	
stamp	Day	Night	Day	Night	Day	Night	Day	Night
2	0.58 K	0.58 K	0.73 K	0.66 K	233	595	5.6 %	9.7 %
3	0.08 K	0.32 K	0.81 K	0.67 K	2035	2662	49.1 %	43.5 %
4	0.40 K	0.29 K	0.51 K	0.59 K	83	233	2.0 %	3.8 %
5	-0.19 K	0.06 K	0.73 K	0.56 K	1797	2624	43.3 %	42.9 %
Total	0.00 K	0.24 K	0.79 K	0.65 K	4148	6114	100.0 %	100.1 %

Table 3. Data analysis distributed on quality values.

The results shown in Table 3 turn up to be seasonal dependent. Table 4 shows the seasonal variation for quality 5 data only. The mean bias is lowest during spring and summer (day and night), and the daytime standard deviations are highest during spring and summer. For Winter and Autumn the mean biasses are nearly identical for day and night, indicating that a mean bias of approximately 0.10 – 0.15 K may be considered as a sort of "diurnal free" and cloudfree SST value.

Season	Mean temp. diff.		Std. of temp. diff.		Number of obs.	
	Day	Night	Day	Night	Day	Night
Winter	0.10 K	0.15 K	0.40 K	0.76 K	143	455
Spring	-0.21 K	-0.03 K	0.66 K	0.31 K	143	179
Summer	-0.29 K	0.03 K	0.78 K	0.54 K	1253	1603
Autumn	0.15 K	0.13 K	0.44 K	0.47 K	258	387
Total	-0.19 K	0.06 K	0.73 K	0.56 K	1797	2624

Table 4. Data analysis for each season, quality 5 SSTs only.

Finally, the overall results for quality 5 data alone compared to all valid SSTs (quality 2-5) are shown in Table 5. The mean bias is about 0.20 K lower for quality 5 SSTs than for all SSTs.

Quality	Mean temp. diff.		Std. of temp. diff.	
flag	Day	Night	Day	Night
5	-0.19 K	0.06 K	0.73 K	0.56 K
All (2-5)	0.00 K	0.24 K	0.79 K	0.65 K

Table 5. Summarized mean biasses and standard deviations for the whole validation period. Results shown for quality 5 SSTs alone, and for all valid SSTs.

4. CONCLUSIONS

Based on a comparison between the NAR MNOR SST product and in situ measurements from (primarily) the southern part of the MNOR area, the following conclusions can be drawn:

- In the vast majority of the months, the OSISAF SSTs are within the expected accuracy
- In situ temperatures measured at depths between 0 and 4-6 meters fit equally well to the OSISAF SSTs
- There do not seem to be geographical variations in the accuracy of the OSISAF SSTs – at least not for daytime data
- There are seasonal variations in the accuracy of the OSISAF SSTs
- A bias between 0.10 K and 0.15 K may be considered as a bias without seasonal and diurnal effects
- Results depend largely on the accepted quality flags.

Generally, the OSISAF SST product seems to be a very accurate and reliable product, which is recommended for as well operational as research purposes. The significant seasonal variations, however, indicate that care must be taken in relation to the diurnal variation problem. Furthermore, the varying results for different quality SSTs imply that a public presented SST product or analysis must inform clearly, which quality stamps that have been included in the production of the SSTs. Though quality 5 data are assumed to represent the “best” SSTs, an SST product based solely on qualilty 5 SSTs will consist of approximately 55% less valid data, than an SST product where quality stamps 2-5 are used.

5. REFERENCES

MÉTÈO-FRANCE, 2005, North Atlantic Region Sea Surface Temperature Product Manual, Version 1.5

Sølvsteen, C., 2005, Comparison between in situ temperature measurements and NOAA/AVHRR based SSTs from O&SI SAF, Version 3