Intercomparison and Sensitivity study of GOES-R ABI Cloud Mask and SST Quality Control prototyped with MSG/SEVIRI

N. Shabanov[*1,2], A. Ignatov[1], B. Petrenko[1,2], Y. Kihai[3] and A. Heidinger[1,4]

[1]NOAA/NESDIS, Center for Satellite Applications and Research (STAR), Camp Springs, MD
[2]IMSG Inc., Kensington, MD
[3]DELL Services Federal Government, Fairfax, VA
[4]University of Wisconsin, Madison, WI

Author for correspondence:
Nikolay V. Shabanov
NOAA/NESDIS, 5200 Auth Road, Suite 603,
Camp Springs, MD 20746-4304, USA
e-mail: Nikolay.Shabanov@noaa.gov
fax: (301)763-8572

To be submitted to JGR

Abstract:

Geostationary Operational Environmental Satellite R-Series (GOES-R) will carry Advanced Baseline Imager (ABI) onboard. This paper presents intercomparison and sensitivity analysis of ABI Cloud Mask (ABI CM) and Sea Surface Temperature Quality Control (SST QC), prototyped with data from the Spinning Enhanced Visible and IR Imager (SEVIRI) onboard Meteosat Second Generation (MSG). The ABI CM provides to all downstream products initial liberal cloud detection by analyzing spectral, spatial and temporal signatures in satellite radiances. SST QC is specifically developed for stringent SST data screening irrespective of contamination source (cloud, aerosols, etc) by matching satellite SST and corresponding brightness temperature (BT) data with clear-sky references (blended satellite & *in situ* SST and radiative transfer model simulated BTs). Sequential implementation of the two masks reduces the processing time and provides a consistency check. Such implementation is feasible as ABI CM cloud false detection rate is small (0.5%) compared to 'Clear' rate (15-20%). Analysis of SST QC mask suggests that SST product should be used up to 60° view zenith angle, beyond which threshold product performance rapidly degrades and clear-sky sample diminishes. We analyzed performance of SST QC at the level of individual filters using new concept, *filter transmission curve*, a quantitative, data-independent characteristic of a filter screening capabilities across the full range of global SST anomalies. We quantified a *performance trade-off* between robust statistics (Mean, STD) and amount of clear-sky observations, a consequence of a *modeling trade-off* between over- and under- screening by the filter *transmission curve*.

1. INTRODUCTION

Geostationary Operational Environmental Satellite R-Series (GOES-R) will carry Advanced Baseline Imager (ABI) onboard (WWW1). A suite of algorithms is currently being developed under the GOES-R Algorithm Working Group (AWG) for retrievals of environmental products [WWW2]. Sea Surface Temperature (SST) is generated with the Advanced Clear-Sky Processor for Oceans (ACSPO) system, which was initially developed for the Advanced Very High Resolution Radiometer (AVHRR) sensors flown onboard NOAA and Metop polar-orbiting platforms (e.g., Liang et al., 2009, Petrenko et al., 2009). Prior to launch, GOES-R ABI SST retrievals are being prototyped with the data from the Spinning Enhanced Visible and IR Imager (SEVIRI) onboard Meteosat Second Generation (MSG) platforms (Shabanov et al., 2009; Ignatov, 2010).

Robust product quality assessment is a challenging task due to global variability of environmental contamination conditions (clouds, aerosols), further complicated by sensor performance issues (calibration degradation, digital noise, unfavorable illumination/observation conditions). The ABI Cloud Mask (ABI CM) is a threshold-based approach for cloud screening implemented by analyzing spectral, spatial and temporal signatures in satellite radiances. The approach builds upon heritage systems employed for AVHRR and MODerate resolution Imaging Spectroradiometer (MODIS): Clouds from AVHRR Extended (CLAVR-x), Cloud and Surface Parameter Retrieval (CASPR), and MOD35 Cloud Mask (Heidinger, 2002, 2005, 2010). Prior experience suggests that a generic Cloud Mask may be suboptimal for individual downstream products (Heidinger, 2010, Martins et al., 2002, OS&I SAF SST Product User Manual). Hence the ABI CM was designed to be liberal (allows residual cloud leakage under constrain of minimal cloud false detection) and reconfigurable (allows to build a user-specific cloud mask from selected individual CM tests). Further, while clouds constitute a major contamination on a global scale, ABI CM does not provide the screening for other contamination sources, such as aerosols, which may be especially important at a regional scale. Individual products screening can be customized by combining ABI CM with ABI Aerosol Detection Mask (Kondragunta, 2010) or by developing a product-specific Quality Control (QC).

ABI SST Quality Control (SST QC) has been especially tailored for SST applications, it provides SST product screening independent from source of contamination/sensor performance. SST QC is a threshold-based approach, which has been initially implemented for AVHRR (Petrenko et al., 2010) and extended to SEVIRI in this work. The approach extensively relies on using blended satellite and *in-situ* SST reference fields (Reynolds et al, 2008) and Community Radiative Transfer Model (CRTM, WWW3) simulations of Brightness Temperatures at sensor's thermal bands, using global SST reference fields and atmospheric profiles of temperature, water vapor, ozone, wind speed from the National Center for Environmental Prediction Global Forecast System (NCEP/GFS, WWW4).

This study covers two interrelated topics in the operation implementation of quality control of GOES-R ABI SST product: (a) Analysis of relative performance of the external ABI CM and internal SST QC for robust and operationally efficient SST data screening, and (b) Sensitivity analysis of performance of SST QC up to the level of individual filters as function of key environment and modeling parameters. This paper is organized as follows. Section 2 describes

data analyzed in this work. Section 3 provides relevant details about ABI CM and SST QC algorithms. Section 4 provides results of intercomparison of performance of ABI CM and SST QC. Section 5 reports on sensitivity of SST QC and individual filters. Key results and future work are summarized in Section 6.

2. DATA

Meteosat-9 (MSG-2) SEVIRI data for June, 2008 were used in this study. Time Series of ABI CM and SST (including SST QC) 15-min Full Disc (FD) products were generated at hourly intervals. Full Disk image corresponds to a 3712x3712 pixel matrix, with 4.8 km resolution at nadir (Schmetz et al., 2002). The "SATELLITE" projection is used to store and display data with the sensor positioned at (0.0;0.0) Lat/Lon position. Algorithms primary input include Brightness Temperatures (BTs) from two Thermal Infrared Channels (Ch9 @ 10.80 μm, and Ch10 @ 12.00 μm), observation/illumination geometry (View Zenith Angle, Solar Zenith Angle, and Relative Azimuth), geolocation (Latitude and Longitude) and Land/Water (L/W) mask. SEVIRI Mid-Infrared (MIR) Channel (Ch04 @ 3.9 μm) was excluded from input for two reasons (1) data noise, temporal discontinuity, artifacts cased by Glint effect, (2) abnormally high discrepancy with CRTM simulations, caused mainly by inability of CRTM to model complex absorption processes at broad SEVIRI MIR channel. However, potentials of MIR channels will be explored when ABI data will be available.

Two major ancillary data sources include global weekly 1° and daily 0.25° fields of Reynolds SST (Dash et al., 2010, Reynolds et al., 2002, 2007) and the National Centers for Environmental Prediction Global Forecast System (NCEP/GFS) 1° 6-hour upper air data (WWW4). SST reference fields, GFS data and other ancillary information provide input to Community Radiative Transfer Model (CRTM, v2.02 in this study), which simulates BTs and its derivatives (Jacobians) at the sensor wavebands at coarse spatial resolution (Liang at al., 2009, WWW3). Reference fields, GFS data and simulated BT/Jacobians were bi-linearly interpolated from low-resolution grid to SEVIRI "SATELLITE" projection.

Unless specified otherwise, in all numerical exercises we use: (1) VZA < 60°, a limit beyond of which SEVIRI SST field is contaminated by artifacts, as is shown in this publication; (2) 6-hours GFS forecast; (3) Daily Reynolds as First Guess (for SST and SST QC) and as a Reference (for SST anomalies).

3. ALGORITHMS

3.1 ABI Cloud Mask

The objective of ABI CM mask is to provide initial cloud detection for all ABI downstream products. The ABI CM has been developed by the AWG Cloud Application Team at the University of Wisconsin and builds upon heritage approaches employed for AVHRR and MODIS (Heidinger, 2002, 2005, 2010). The algorithm is based on analysis of spectral, spatial and temporal signatures. It includes up to 30 different tests, with 10 tests being relevant to SST applications (cf. Table 1). Note that ABI CM is under development and some test will be implemented later (such as

temporal continuity, etc). Thresholds in the individual ABI CM tests have been tuned against CALIPSO Lidar measurements. Online Pressure-layer Fast Algorithm for Atmospheric Transmittances (PFAAST) RTM simulations are employed in several ABI CM tests (12, 13, 21, and 22, cf. Table 1). Target misclassification rate ('False Clear' + 'False Cloudy') is 13%. The latest tests over ocean demonstrated a misclassification rate of 8.8% wrt CALIPSO data. The ABI CM output contains results of the individual cloud tests, which are further aggregated into overall ABI CM with four states: 'Clear', 'Pr. Clear', 'Pr. Cloudy', and 'Cloudy', where "Pr." stands for "Probably" (cf. Table 2). ABI CM is set as follows: 'Clear'- if no tests were triggered, 'Pr. Clear'- if only 'Thermal Uniformity' test was triggered, 'Cloudy'- if any of the tests were triggered, except 'Thermal Uniformity', if 'Cloudy' pixel touches 'Clear' or 'Pr. Clear' rename 'Cloudy' into 'Pr. Cloudy'. For ABI SST applications we reconstruct ABI CM from individual thermal tests only to avoid temporal discontinuity caused by Optical and MIR filters (day-night discontinuity, artifacts in vicinity of glint, misclassification of aerosol contamination). In the remaining of this text when we refer to ABI CM, we imply Thermal ABI CM mask.

3.2 ABI SST

Advanced Clear-Sky Processor for Oceans (ACSPO) has been adopted for generating GOES-R ABI SST product (Ignatov, 2010, Petrenko et al., 2010). ACSPO is an end-to-end production system comprised from infrastructure and core modules. Infrastructure modules (I/O, CRTM, ancillary data processors, numeric service routines, etc.) provide the required data for product generation by core modules. Among main functions of infrastructure modules are: (1) initializing all system parameters, (2) ingesting SEVIRI dynamic and static data, (3) ingesting and spatially interpolating SST Reference field and GFS data, (4) initializing and performing CRTM simulations to calculate BTs/Jacobians at coarse GFS grid following by bi-linear spatial interpolation to the SEVIRI projection. Given inputs, core modules are executed, SST followed by SST QC. Next, control is returned to infrastructure to store output. For the GOES-R ABI project, stand-alone ACSPO system has been integrated into AWG framework [WWW2]. Only ACSPO core modules were supplied to AWG framework, which assumes role of infrastructure and also manages data input from upstream ABI CM into downstream SST. Sequential execution of ABI CM and SST QC is depicted in Fig. 1 and will be discussed in Section 4.1. Below we detail ACSPO SST and SST QC core modules.

3.2.1 SST Hybrid Algorithm

The hybrid SST algorithm has been recently developed under GOES-R project (Petrenko et al., 2011). The algorithm combines best features of two previously known approaches, regression split-window Non-Linear SST (Walton et al., 1998) and RTM inversion-based Optimal Estimation SST (Merchant et al., 2008, 2009). Due to efficient use of the *a priory* information the hybrid algorithm delivers the highest overall SST precision and the most uniform distributions of regional SST accuracy and precision, still being fast and requiring minimum assumptions. The hybrid SST algorithm is formulated as follows (Petrenko et al., 2011),

$$\begin{aligned} SST = SST_{FG} + a_0 + a_1 \Delta BT_{09} \\ + [\Delta BT_{09} - \Delta BT_{10}]\times[a_2(SST_{FG} - 273.15) + a_3(\sec(\theta) - 1)], \end{aligned} \quad (4)$$

where, SST is the retrieved SST, SST_{FG} is the First Guess SST, $\Delta BT \equiv BT_{CRTM} - BT_{SEVIRI}$ is the BT anomaly, a per-pixel discrepancy between CRTM-predicted and satellite-measured BT (all temperatures in Kelvin). The coefficients for hybrid algorithm were derived from matchups of SEVIRI observations and *in situ* drifter buoy measurements (a_0=0.743279, a_1=1.07488, a_2=0.0589083, a_3=0.734534).

3.2.2 SST Quality Control

The objective of the SST QC is screening of SST contaminated data independent from its source (clouds, water vapor, aerosols, sensor noise, etc). The approach is based on matching satellite SST and corresponding brightness temperature (BT) data with clear-sky references (blended satellite & *in situ* SST and CRTM simulated BTs). By design SST QC is conservative, it screens data to closely approximate clear-sky SST statistics. It currently utilizes 3 tests (Radiance, Adaptive SST and SST Uniformity, all described below). The initial version of the SST QC has been developed for ACSPO-AVHRR (Petrenko et al., 2010) and has been adopted in this work for SEVIRI. The SST QC output contains results of the individual QC tests, which are further aggregated into an overall SST QC with three states: 'Optimal', 'Sub-Optimal' and 'Poor' (cf. Table 2). SST QC is set as follows: 'Optimal'- if no tests were triggered, 'Sub-Optimal'- if only Uniformity test was triggered, 'Poor'- if Radiance or Adaptive SST, but not Uniformity tests were triggered.

Radiance Test

Radiance test is a per-pixel test, predictors are Brightness Temperature (BT) anomalies in SEVIRI Channels 09 and 10. Test compares discrepancies between sensor-measured and CRTM modeled BT,

$$\mathrm{Exp}[-(\Delta BT_{09} - \langle\langle\Delta BT_{09}\rangle\rangle)/STD(\Delta BT_{09})] + \mathrm{Exp}[-(\Delta BT_{10} - \langle\langle\Delta BT_{10}\rangle\rangle)/STD(\Delta BT_{10})] < \mathrm{Thresh}, \quad (1)$$

if true, pixel is classified as 'Clear', otherwise as 'Contaminated''. In Eq. (1) $\Delta BT \equiv BT_{CRTM} - BT_{SEVIRI}$, is BT anomaly, a per-pixel discrepancy between CRTM-predicted and satellite-measured BT, $\langle\langle\Delta BT\rangle\rangle$ is the systematic bias (most probable value) of ΔBT, dynamically evaluated over all ocean pixels in FD (Petrenko et al., 2010). Parameters (*Thresh* and *STD(ΔBT)*) are listed in Table 3. Several alternative versions of the filter have been tested. For instance, a quadratic form (sum of quadrates) is inferior because it: (a) allows substantial cloud leakage, (b) possess symmetry with respect to peak of histogram of SST anomalies (cf. Section 5.2).

Adaptive SST Test

Adaptive SST test is a space window test, predictor is the SST anomaly. The test is implemented as iterative technique and works as follows. All ocean pixels with valid satellite observations are divided initially into two clusters based on the condition,

$$\Delta SST < -k*STD(SST), \quad (2a)$$

if true, pixel belongs to 'Contaminated', otherwise to 'Clear' cluster. In Eq. (2a) $\Delta SST \equiv SST - Reference$, is SST anomaly, a per-pixel discrepancy between retrieved SST and Reference SST. Parameter *k* listed in Table 3 and parameter *STD(SST)* is sourced from FG SST field. This test provides an initial approximation of clusters, which can be refined by taking into account spatial aspects of the ΔSST field. Sliding window (15×15 pixels) is selected and statistics (Mean and STD) over 'Contaminated' and 'Clear' clusters are calculated. If pixel belongs to 'Contaminated' cluster, its classification is kept unchanged; otherwise, pixel is tested to which cluster it is closer,

$$[(\Delta SST - \langle\Delta SST\rangle_{CTM})/STD(SST)_{CTM}]^2 < [(\Delta SST - 0)/STD(SST)_{CLR}]^2, \quad (2b)$$

if true, pixel is reclassified as 'Contaminated', no change otherwise. Left and right sides of Eq. (2b) can be interpreted as normalized distances in the ΔSST space. Two assumptions were made for this test: (a) Mean(ΔSST) over 'Clear' pixels is close to zero, $\langle\Delta SST\rangle_{CLR}=0$; (b) distance over clear pixels should not be less than noise level, $[(\Delta SST - 0)/STD_{CLR}]^2 > 1$; (c) perform reclassifications only if there are enough 'Contaminated' pixels to generate statistics, $Num_{CLD} > f*Num_{WIN}$, where parameters Num_{WIN} and *f* are listed in Table 3. This test is repeated two times over FD.

SST Uniformity Test

The Uniformity test is a space window test, predictor is SST, enhanced with SST anomaly. Test compares spatial variability in retrieved SST with that expected around the pixel, namely,

$$Var(SST)*\exp[-(\Delta SST/STD(SST)] < Thresh, \quad (3)$$

if true, pixel is classified as 'Clear', otherwise as 'Contaminated'. Statistics (Mean, STD and Variance) are generated over a sliding window (3×3 pixels). Parameters *STD(SST)* and *Thresh* are listed in Table 3. Initially, the Uniformity test has been implemented in its standard form, to test for high Variance only, indicative of contamination. However, this formulation results in high rate of false alarms, where filter is triggered on the sensor noise over clear-sky data (speckles of mask over continuous clear-sky regions). To mitigate this problem, test was enhanced with exponential factor with SST anomaly as an argument: negative anomalies enhance filtering, while positive suppress it.

4. INTER-COMPARISON OF ABI CM AND SST QC

In comparison of the ABI CM and SST QC we pursue to investigate three main problems: (1) consistency ABI CM and SST QC, (2) possibility of sequential execution of ABI CM followed by SST QC, (3) possibility of improving clear-sky SST statistics under sequential execution of ABI CM and SST QC. To address the above problems we implemented side-by-side comparison and confusion matrix analysis.

4.1 Comparison of ABI CM and SST QC Masks

Diurnal cycle of the components of ABI CM ('Clear', 'Pr. Clear', 'Pr. Cloudy' and 'Cloudy') and SST QC ('Optimal', 'Sub-optimal', and 'Poor') for June 03, 2008 is shown in Fig. 1. The fraction of the 'Optimal' SST QC (18-23%) pixels quite closely matches that of 'Clear' ABI CM (19-25%) during whole diurnal cycle. However, the fraction of 'Pr. Clear' and 'Pr. Cloudy' ABI CM pixels (~20% each) is much higher than that of 'Sub-Optimal' SST QC pixels (~5%). These properties are consequence of masks design: ABI CM is liberal, exhibits large amount of the two 'Probably' categories, while SST QC is conservative, with thin 'Sub-Optimal' transitioning category, as majority SST applications require 'Usable' vs. 'Not Usable' binary decision. Nevertheless, for operational implementation 'Clear' and 'Pr. Clear' categories were joined to avoid false alarms triggered by Thermal Uniformity ABI CM filter. Both ABI CM and SST QC time series are smooth, there is no Day/Night discontinuity, as both masks were constructed from thermal channels only. Both masks exhibit diurnal variability with a minimum/maximum at 03:00/14:00 UTC. Currently it still remains unclear if this observed diurnal variability reflects real change in cloud cover during the course of the day, and what physical mechanism is responsible.

Sample spatial distributions (FD and zoom-in) of ABI CM and SST QC masks are shown in Fig. 2 for data acquired on June 03, 2008 at 12:00 UTC. Spatial distribution further demonstrate features observed in time series analysis: ABI CM is more liberal and spatially continuous, while SST QC is conservative and more fragmented. Across spatial scales, both masks are generally consistent, that is, cover the same major cloud formation zones. Classes in the ABI CM are highly clustered with broad transitioning zones ('Pr. Cloudy' and 'Pr. Clear'). Classes in SST QC are more fragmented, with thin transitioning zone ('Sub-Optimal') at clouds edges.

Confusion matrix analysis complements side-by-side comparison of ABI CM and SST QC, as it is important to match not only the total amount of pixels, but also quantify spatial consistency of both masks. The confusion matrix is defined in the general way between 4 components of ABI CM and 3 components of SST QC, namely as a ratio of pixels of particular component to the total amount of ocean pixels. We will analyze below the following components of the confusion matrix:

$$\begin{aligned} &\textbf{Both CM \& QC} \equiv [\text{QC=Optimal}] \cap [\text{CM=Clear} \,||\, \text{Pr. Clear}], \\ &\textbf{Only QC} \equiv [\text{QC=Optimal}] \cap [\text{CM=Pr. Cloudy} \,||\, \text{Cloudy}], \\ &\textbf{Only CM} \equiv [\text{QC=Sub-Optimal} \,||\, \text{Poor}] \cap [\text{CM=Clear} \,||\, \text{Pr. Clear}]. \end{aligned} \tag{5}$$

Figure 3 shows a snapshot of confusion matrix for 12:00 UTS and diurnal cycle of components for June 03, 2008. The logic behind selection of particular form of components is as follows. In order to implement sequentially ABI CM and SST QC, there should be only a minimum amount of *false-cloudy* ('Only QC') pixels, those, masked-out by ABI CM, but in fact being 'Optimal' according to SST QC. Consider values of the Confusion matrix elements: while fractions of ABI CM 'Clear' is very close to SST QC 'Optimal' (27.6% vs. 26.6%), still SST QC 'Optimal' overlaps only over 21.3% of ABI CM 'Clear', with the remaining spread mostly over ABI CM 'Pr. Clear' category (4.9%). Thus, if we combine 'Clear' with 'Pr. Clear' ABI CM categories,

they will nearly fully cover 'Optimal' SST QC category (21.3%+4.9%) with cloud false identification rate of 0.3+0.1=0.4%.

Compare the diurnal cycle of components and the confusion matrix. 'Optimal' SST QC category is mostly inside 'Both CM and QC' (18-27%). Therefore, the diurnal cycle of 'Both CM and QC' components comes from diurnal cycle in the 'Optimal' SST QC with peak near 14:00 UTC (cf. Fig 2). Fraction of 'Only CM' is quite large (20%), but nearly constant during the diurnal cycle. In essence, 'Only CM' corresponds to mixed/contaminated pixels which are passed through by ABI CM (due to inclusion of 'Pr. Clear'), but screened out by SST QC. Therefore, 'Only CM' includes pixels which can be further fine-screened with SST QC. Finally, 'Only QC' is stratified within 0.5% corridor during whole diurnal cycle. This indicates possibility of sequentially executing ABI CM followed by SST QC with accuracy of 0.5%. Refer to Fig. 1 which provides flow chart of sequential execution of ABI CM followed by SST QC in the AWG framework. Recall, we initially reconstruct ABI CM from individual thermal tests to ensure temporal continuity. The processing efficiency achieved by this approach is as follows: around 60% of pixels ('Cloudy' + 'Pr. Cloudy') are passed through by SST QC as 'Poor', the remaining 40% ('Clear' + 'Pr. Clear') are further analyzed by SST QC, resulting in about 15-20% of 'Optimal' for SST applications pixels during diurnal cycle, and remaining classified as 'Sub-Optimal' (Fig. 1).

4.2 Comparison of SST screened with ABI CM and SST QC Masks

Side-by-side comparison of capabilities of ABI CM and SST QC masks for screening SST anomalies is shown in Figs 4-5. According to the results of the previous section, we use 'Clear' and 'Pr. Clear' ABI CM and 'Optimal' SST QC categories to filter out pixels suitable for SST applications. Figure 4 shows spatial distribution of SST anomalies, screened with the two masks for June 02, 2008, 12:00 UTM. Note, we use gridding to show the maps, namely, we averaged any available values in the 6x6 pixel window to represent ΔSST in the grid cell displayed in the gridded image. Generally, patters are similar: (1) patterns of gaps, corresponding to clouds formations are similar, (2) on the valid areas both maps indicate "hot spots" along the African coast and in the Mediterranean Sea, which are not captured by coarse resolution Reference field (0.25° Daily Reynolds SST). However, residual clouds are much more pronounced in the ABI CM (e.g., cold anomalies in the NW part of the Atlantic Ocean and African coast, Mediterranean Sea, etc). Also, regions with high aerosols contaminations off Sahara desert and Middle-East coasts (Le Borgne et al, 2006) are passed undetected by ABI CM (='Clear'), but marked as contaminated by SST QC (='Poor'). This feature is a consequence of masks design: while ABI CM was designed to screen only clouds, SST QC is designed to screen any type of contamination. The use of ABI aerosols mask (Kondragunta, 2010) should enhance capabilities of ABI CM data screening.

Statistics for ΔSST screened with ABI CM and SST QC are presented in Fig. 5. Instantaneous histograms at 12:00 UTM and diurnal cycle of mean and STD are shown for June 03, 2008. In the case of ABI CM, cold tail of the ΔSST histogram is well pronounced, while in the case of SST QC left and right shoulders of the histogram are symmetric and values concentrate in the [-2,+2] K interval. Right shoulders of both histograms coincide, indicating consistency of both masks with respect to warm anomalies. The left shoulder of SST QC histogram is steeper and

below that for ABI CM starting from ΔSST^* (peak of the histogram). Thus, it may seem that SST QC mask over-screens "valid" pixels at ΔSST^*. In fact, those pixels correspond to ambient cloudiness, and were masked out by 'Sub-Optimal' SST QC category (cf. Section 5.2). Next, consider diurnal cycle of mean and STD of ΔSST. Mean ΔSST screened with ABI CM is biased negative by nearly 0.7 K wrt those screened with SST QC mask, due to residual cloud contamination in ABI CM. However both masks indicate diurnal cycle of Mean ΔSST with amplitude of 0.2K. Cloud contamination passed through ABI CM can be further observed in the time series of standard deviations, which show that STDs for the ABI CM (1.2K) is a factor of ~2 larger than for SST QC (0.5K). Statistically significant estimates of Mean and STD over whole month of June 2008 are reported in Section 5.1.

Statistics of ΔSST screened with the components of ABI CM / SST QC confusion matrix are shown in Fig. 6 (presentation is similar to Fig. 5). From comparison of Fig. 3 and Fig. 6 it follows: (a) majority of pixels fall into 'Both CM and QC' histogram, which is centered at ΔSST^*; (b) minor amount of pixels falls into 'Only QC' histogram, which is also centered at ΔSST^*; (c) moderate amount of pixels falls into 'Only CM' histogram, however distribution has peak outside of 'Both CM and QC'- it is shifted toward cold tails. Thus, 'Only QC' category represents minor amount of pixels over-filtered by ABI CM, but useful for SST applications. In contrast 'Only CM' represent moderate amount of ambient cloud/cloudy pixels, kept by ABI CM, but filtered out by SST QC. Diurnal cycle of Mean and STD of ΔSST indicates that statistics for 'Both CM and QC' and 'Only QC' are clustered together, but 'Only CM' captures negative ΔSST anomalies.

Statistics of ΔBT in SEVRI Channel 09 and 10 screened with ABI CM and SST QC are shown in Fig. 7 (presentation is similar to Fig. 5). Distribution of ΔBT for Channel 10 is slightly broader than that for Channel 09. Mean(ΔBT) for both Channels exhibits a systematic bias of about -0.5K when screened with SST QC, and degrade to -1.2K when screened with ABI CM. Respectively, STD(ΔBT) is about 0.6K (for SST QC) and degrades to 1.2K (for ABI CM). The systematic bias for SST QC is a consequence of limitation of CRTM to model SEVIRI observations even for clear-sky conditions. Also, compare Figs 5 and 7 and note that the amplitude of the diurnal cycle is higher for SST than for BT. Detailed analysis of the diurnal cycle is beyond the scope of this paper but possible reasons include (a) transformation BT to SST, (b) differences in the temporal resolution of the references (6-hours GFS forecast vs. daily Reynolds SST).

Overall, analysis presented in this section clearly indicates that the ABI CM alone is not sufficient to provide high accuracy SST retrievals, and SST QC is effective in improving accuracy/precision of the clear-sky SST product.

5. SENSITIVITY ANALYSIS

In this section we discuss performance of SST QC as function of environmental conditions and modeling errors. We start with reviewing several aspects of sensitivity of overall SST QC mask, and follow with in-depth analysis of individual filters.

5.1 Sensitivity of overall SST QC mask

A month of SST retrievals (June 1-30, 2008) is reported in Fig. 8. To understand impact of SST reference field on clear-sky SST (SST masked with SST QC), we constructed four possible combinations, where two SST Reference fields (Weekly and Daily Reynolds) were used in two roles: (a) to construct SST anomaly, ΔSST ≡ SST-Reference; (b) to calculate SST (First Guess SST in Eq. (4)) along with SST QC (Reference SST in Eqs. (2)-(3)). Consider variations of Mean and STD of four combinations and contrast those against inconsistencies between References itself. SST anomaly clearly exhibits diurnal cycle with amplitude of about 0.3K, Mean SST anomaly oscillate 0, and STD is within [0.5-0.75] interval, which matches to reported in literature values (Brisson et al., 2002). In addition to diurnal cycle, SST field exhibits multiple lower frequency cycles.

The following should be noted concerning Mean(ΔSST) shown in Fig. 8. The discrepancy between Weekly and Daily Reference fields varies significantly during month of June 2008: in first half of the month it rises up 0.2K, however both fields are close to each other at the second half. Weekly cycles in the Mean(ΔSST) are not in phase with discrepancies in References, suggesting, that anomalies capture satellite signal, instead of just following References variability. Mask change has the same level of impact as Reference change on Mean(ΔSST). Discrepancies in Mean(ΔSST) for different combinations of Masks and References is of the same order as discrepancies as one between References.

The following should be noted concerning STD(ΔSST) shown in Fig. 8. The lowest STD is achieved when Daily Reynolds is used both as reference for calculation of ΔSST and generation of SST/SST QC; any combination involving Weekly Reynolds increases STD(ΔSST). This value of STD is on the boundary of minimum STD of 0.5-0.6K, STD of discrepancies between reference fields. Possible reason- diurnal cycle of SST enhances STD(ΔSST). To contrast, the accuracy and precision of the on-line clear-sky RTM simulations is about 0.5K (Liang et al., 2009).

Consider another SST QC sensitivity case study, Mean and STD of ΔSST as function of View Zenith Angle, shown in Fig. 9. Combination of two factors impact the shape of those dependencies: (a) number of clear-sky observations in particular bin of VZA descretization; (b) discrepancies between SST and Reference. Analysis of Fig. 9 indicates that valid range of VZA for SST applications is $0\text{-}60^0$. On the upper boundary, STD is increasing while number of observations is decreasing sharply. The reasons for degradation of performance include (a) increase of cloudiness as function of VZA, (b) degradation of CRTM and SST Reference fields accuracy at slant angles.

5.2 Sensitivity of Individual SST QC Tests

Diurnal cycle of triggering rates of individual filters and categories of ABI CM and SST QC masks is shown in Fig.10. ABI CM individual filters (Table 1) have wide distribution of triggering rates, caused by different sensitivity of filters to environmental conditions. Triggering rates of thermal channels based filters are flat through the diurnal cycle with rate changing from

75% ('Thermal Uniformity') to below 1% ('Uniform Low Stratus'). Optical Channels based filters ('Reflectance Uniformity', 'Reflectance Gross Contrast', 'Relative Visible Contrast', and 'Uniform Low Stratus') have wide variation (0-75%) of triggering rates through the diurnal cycle, with maximum at 12:00 UTM. Mid-Infrared channel based filter ('4-micron emissivity') have moderate variability of 20-50%. As diurnal cycle in SST is small (0.2K), we removed both Optical and MIR tests from ABI CM to minimize discontinuity. 'Clear' category corresponds to cases when none of the filters were triggered; 'Pr. Clear' corresponds to those, when only 'Thermal Uniformity' filter was triggered.

SST QC mask is constructed based on different principles, and its performance is different from that of ABI CM: triggering rates of individual SST QC filters are clustered together (~75%). Triggering rates of 'Radiance' and 'Adaptive SST' tests are close, and that for 'SST Uniformity' is somewhat higher. 'Optimal' category corresponds to cases, when no filters were triggered; 'Sub-Optimal' category corresponds to those, when only 'SST Uniformity' filter was triggered. The reason behind close clustering of 'Radiance' and 'Adaptive SST' filters is that they both were build on the same idea of matching (processed) satellite data against Reference. However, while *general idea* is similar between filters, the *implementation form* is different: 'Radiance' filter is purely empirical and involves BT, while 'Adaptive SST' is physically-based and involves SST. Impact of the *form* on the performance can be traced with 'Unique Contribution' and 'Unique Omission' rates, defined as:

$$\textbf{Test(i) Unique Contribution} \equiv [\text{Test(i)=ON}] \cap_{j\neq i} [\text{Test(j)=OFF}],$$
$$\textbf{Test(i) Unique Omission} \equiv [\text{Test(i)=OFF}] \cap_{j\neq i} [\text{Test(j)=ON}]. \quad (6)$$

Diurnal cycle of 'Unique Contribution' and 'Unique Omission' rates are shown in Fig. 11. 'Adaptive SST' filter generally exhibits higher 'Unique Contribution' and 'Unique Omission' rates compared to that of the 'Radiance' filter. 'SST Uniformity' filter demonstrates highest 'Unique Contribution' rate by construction: this filter creates not only 'Poor', but also a 'Sub-Optimal' category of ambient cloudiness. In contrast, 'Unique omission' rate of the 'Uniformity' filter should be zero in ideal case so each cloudy cluster is framed by 'Sub-optimal' category. However, in practice some ambient clouds are not detected by 'SST Uniformity' filter (cf. Fig. 2, zoom-in for SST QC mask) resulting in 'Unique Omission' rate of ~2%.

Next, we analyze filtering mechanism of an individual SST QC filter. Filter design has been active research area in Remote Sensing (Ackerman et al., 1998, Heidinger 2005, 2010). How to build an "optimal" filter and what is the methodology to assess accuracy of filtering? How to minimize simultaneous over- and under-screening errors? Developers of ABI CM addressed some of the above question by validating ABI CM mask with CALIPSO Lidar measurements. Analysis presented in former sections indicates that SST QC improves ABI CM mask for SST applications. Therefore, one may expect that reference information may give further insight how to build "optimal" mask and guide sensitivity studies.

Sensitivity study of SST QC relies on the following concepts. We optimize each filter regardless of others; optimal performance of each individual filter guarantees optimal performance of the overall mask. Optimal performance is understood in terms of maximizing screening of contaminated pixels until 'False alarms' (over-screening) will appear; under-screening can be

mitigated by combining multiple filters. In our sensitivity studies we use combination of the following data: a) Reference SST; b) masks of Filter, SST QC, ABI CM, and Confusion matrix of SST QC and Filter. In terms of techniques we use both histograms and images, as combination provides additional insights. As performance of filters is similar, we will demonstrate sensitivity analysis using 'SST Uniformity'filter. Results are presented in Figs 12-13.

As a starting point, consider a fundamental characteristic of a filter, "*filter transmission curve"*, defined as follows (c.f., Fig. 12). For a particular interval of ΔSST, transmission is defined as ratio of pixels transmitted by filter to the total amount of pixels processed in this interval, with rates ranging in interval 0-100%. The distinctive feature of the transmission curve is that this is a characteristic of a filter itself, rather than characteristic of data, processed by filter. Transmission curve is applied to input data stream to get histogram of transmitted pixels. Data independence is a convenient feature to analyze and tune filter performance regardless of variations in input data stream. The transmission curve is asymmetric wrt ΔSST=0: at large negative ΔSST (contamination due to clouds and aerosols) filter transmits 0% of pixels; as ΔSST is increasing, filter incrementally transmits more, reaching 100% transmission at some positive ΔSST value. Warm anomalies are typically due to: a) real features of SST field, such as warm currents, not captured by coarse-resolution References; or b) artifacts, such as mixed Ocean-Land pixels (heating of coastal line), or underestimation of SST by a Reference field. Both cases are unrelated to atmospheric contamination, and corresponding pixels should be passed through by SST QC filter. The particular shape of the transmission curve is defined by a particular model of a filter (function and parameters), which determines 'Unique Omission' and 'Unique contribution' rates of the filter. Although this is beyond the scope of this paper, three distinct regions in the transmission curve can be identified and parameterized (Fig. 12): 1) sharp convergence to 0% transmission for cold anomalies; 2) steep increase in transmission in the intermediate region, 3) sharp convergence to 100% transmission near ΔSST=0 toward positive anomalies.

We tuned filters which generate 'Poor' category of SST QC ('Radiance' and 'Adaptive SST') to 100% transmission starting at ΔSST^* (peak of ΔSST histogram). Transmission curve for 'SST Uniformity' filter has been elevated further to mark pixels at the boundaries of cloud formations as 'Sub-Optimal' (Fig. 12-b). This effect could best be demonstrated by inspecting spatial pattern of ΔSST screened out and transmitted by the filter (Fig. 13). Figure 13-a shows ΔSST screened-out by 'SST Uniformity' filter: black represents transmitted pixels, while screened pixels are shown in color for ΔSST $\in$[-4,+4], and cold anomalies beyond this range are shown in white (i.e., clouds). It clearly can be seen that pixels with small anomalies are located at the boundaries of cloud formations, representing ambient cloudiness. In contrast, pixels, transmitted by the filter represent clusters of clear-sky domain with ΔSST mostly in the range [-1,+1] (Fig. 13-b).

Now, return back to histograms shown in Fig. 12. Consider additional details provided by screening ΔSST with components of confusion matrix (Fig 12-a). Majority of pixels belong to 'Both CM and Filter' category which selects only clear pixels from ABI CM mask, while category 'Only ABI CM' selects contaminated pixels up to extended cold tiles. Category 'Only Filter' is minor. Overall, histograms for individual filter resemble those of total SST QC mask (Figs 6-7).

Figure 12-b is the summary plot illustrating relative performance of ABI CM, SST QC and SST QC filters. Transmission curve of total mask is defined as the one separating 'Optimal' vs. 'Sub-Optimal'+'Poor' in case of SST QC, and 'Clear'+'Pr. Clear' vs. 'Pr. Cloudy'+'Cloudy' in case of ABI CM. Transmission curves of 'Radiance' and 'Adaptive SST' filters are clustered together, while transmission curves for 'Uniformity' filter is positioned above. Transmission for two former filters rises to 100% at ΔSST=0, while 'SST Uniformity' filter screens pixels even at this area. Transmission curve for overall SST QC passes above individual filters. ABI CM mask is liberal wrt clouds, rises to ~100% near ΔSST=0, however increasingly screens-out warm anomalies. This screening of large warm anomalies is an artifact of the 'Thermal Uniformity Test' of the ABI CM mask, which has quadratic form, and therefore symmetric in vicinity of ΔSST=0. Still, 'False alarm' rate of ABI CM is small (<0.5%, cf. Fig.3)

As a final exercise, consider sensitivity of SST QC filters to the selection of threshold values. For consistency with former discussion, consider 'SST Uniformity' filter and *Threshold* and *STD(SST)* parameters (Figs. 14 a and b, respectively). We assess performance of the filter based on three metrics: Mean(ΔSST), STD(ΔSST), and NOBS, number of clear-sky observations, as function of parameter values. To calculate the above metrics, we utilized full set of the developed masks: ABI CM, SST QC, SST QC filter, and Components of Confusion Matrix between ABI CM and SST QC Filter. Consider Fig. 14a, which provides dependencies for *Threshold* parameter. In general, as *Threshold* is increasing (changing from tight to lose filter), Mean(ΔSST) for all masks is decreasing (passing-through colder pixels), STD(ΔSST) for all masks is increasing (inclusion of variability of contaminated pixels), and NOBS is increasing for all but 'Only CM' mask (residual between constant ABI CM and expanding Filter). Note, most significant changes take place for 'Threshold' values $\in[0,1]$, while for values higher than 1, dependencies don't saturate, but clearly become much weaker. This effect can be better understood by referencing Fig.12. As filter becomes tighter, it starts filtering-out pixels not from the tails of histogram but from central peak, this explains steep dependence at the tight end. As we consider pixels within central peak being not contaminated (close to reference), the value of filter parameter should be selected at the region just after sharp dependencies subside to focus on filtering cold tiles of histogram. From other hand, consider effect exhibited by SST QC curves. They saturates much faster than any other. The reason for this effect is that at some point, filter become so loose that it contributes little to the union with other filters (= SST QC). Similar general dependencies are valid for the STD(SST) parameter (Fig. 14b). One major difference is that curves for 'Only QC Filter' are much steeper. This is a strong effect of exponent in the filter formulation (cf. Eq. 3), filter becomes quickly weaker, passes through cold tails as STD(SST) is increasing. Based on those observations, we selected values Threshold=3, and STD(SST)=0.8, just off the region of steep changes.

Overall, results of sensitivity of SST QC filter to threshold indicate that this, in essence, is a problem of "trade-offs" between accuracy of statistics and number of clear-sky pixels. Specifically, screening mixed pixels helps to improve statistics, but at expense of lower number of clear-sky pixels. However, one have to recognize, that reducing number of observations, not necessary leads to improving product accuracy, due to limitation of SST QC approach itself: (1) deficiencies of Reference field (over- and under-estimation of SST) translate to errors in calculating SST QC, (2) limited accuracy of modeling filter's transmission curve, (3) difficulties of discriminating clouds and cold ocean currents.

6. CONCLUSIONS AND FUTURE WORK

This paper discussed intercomparison and sensitivity study of ABI CM and SST QC under GOES-R operational set up. ABI CM provides initial 'liberal' cloud screening for all downstream products, while SST QC delivers stringent SST data screening over oceans independent from contamination source. Main results of this study are as follows:

1. Thermal-bands only based ABI CM is recommended for GOES-R SST applications to minimize product temporal discontinuity (day-night transition, glint area). SST QC substantially improves SST statistics. Sequential implementation of the two masks (ABI CM followed by SST QC over 'Clear' + 'Pr. Clear' pixels) reduces SST QC processing time by 60%. Such implementation is feasible as ABI CM cloud false detection rate is small (0.5%) compared to 'Clear' rate (15-20%) (cf. Figs 1 and 3).
2. While reliance of the SST QC on reference fields allowed improving SST statistics, it also set the approach limits: inaccuracies of reference fields may translate into inaccuracies of SST QC and retrieved SST statistics (cf. Fig. 8). This further signifies utility of ABI CM, as independent check for systematic biases and trends in SST QC. Our analysis suggests that product should be used up to 60° view zenith angle, beyond which threshold the SST performance rapidly degrades and clear-sky sample diminishes.
3. Performance of the SST QC was analyzed at the level of individual filters using a new concept, *filter transmission curve*, a quantitative, data-independent characteristic of a filter screening capabilities across a range of SST anomalies. We quantified a *performance trade-off* between robust statistics (Mean, STD) and amount of clear-sky observations, a consequence of a *modeling trade-off* between over- and under- screening by the filter *transmission curve*.

GOES-R ABI CM and SST products are currently under development and further refinement/modifications are possible. Open problems for ABI SST are:

1. Understand physical mechanism underlying several observed phenomenon (the diurnal variability of the amount of clear-sky pixels, and a systematic pattern of cold SST anomalies).
2. Further refine SST algorithm (i.e., explore alternative forms of SST equation) and SST QC filters (optimization of transmission curve with alternative forms of filters) to achieve better trade-off between robust statistics (Mean, STD) and amount of clear-sky observations
3. Efficiency of sequential implementation of ABI CM and SST QC can be improved if ABI CM 'Pr. Clear' category is minimized without increasing false cloud identification, which is currently mainly caused by the implementation of the Thermal Uniformity filter.

To facilitate performance assessment, web-based tool, SQUAM/SEVIRI Section has been set up to monitor SEVIRI SST production, validation and intercomparison online in Near Real time regime [WWW5]. Analysis of time-series will be presented in the upcoming separate publication.

Acknowledgement

The work is conducted under the Algorithm Working Group funded by GOES-R Program Office. Thanks to Prasanjit Dash, John Sapper, XingMing Liang, John Stroup of NOAA/NESDIS for helpful discussions. The views, opinions, and findings contained in this report are those of the authors and should not be construed as an official NOAA or US Government position, policy, or decision.

References

Ackerman, S., K. Strabala, W. Menzel, R. Frey, C. Moeller, and L. Gumley, 1998: Discriminating clear sky from clouds with MODIS, *J. Geophys. Res.*, **103**, 32,141–32,157.

Brisson A., P. Le Borne, and A. Marsouin, 2002: Results of One Year of Preparational Production of Sea Surface Temperatures form GOES-8, *J. Tech.*, **19**, 1638-2109

Dash, P., A. Ignatov, Y. Kihai, J. Sapper, 2010: The SST Quality Monitor (SQUAM), *J. Tech.*, **27**, 1899-1917, doi:10.1175/2010JTECHO756.1.

Heidinger, A., V. Anne, and C. Dean, 2002: Using MODIS to estimate cloud contamination of the AVHRR data record. *J. Tech.*, **19**, 586–601.

Heidinger, A., 2010: ABI Cloud Mask Product Algorithm Theoretical Basis Document, http://goesrtest2.woc.noaa.gov/products/ATBDs/baseline/Cloud_CldMask_v2.0.pdf

Ignatov, A., 2010: ABI SST Product Algorithm Theoretical Basis Document, http://goesrtest2.woc.noaa.gov/products/ATBDs/sea_surface_temperature.pdf

Kondragunta, S, 2010: ABI Aerosol Detection Product ATBD, http://goesrtest2.woc.noaa.gov/products/ATBDs/baseline/AAA_AIP_v2.0.pdf

Le Borrgne P., G. Legendre and A. Marsouin, 2006: Operational SST Retrieval from MSG/SEVIRI data, EUMETSAT Meteorological Satellite Conference, Helsinki, Finland, 12-16 June 2006. http://www.eumetsat.int/Home/Main/AboutEUMETSAT/Publications/ConferenceandWorkshopProceedings/2006/groups/cps/documents/document/pdf_conf_p48_s2c_03_leborgne_v.pdf

Liang X., A. Ignatov, and Y. Kihai, 2009: Implementation of Community Radiative Transfer Model (CRTM) in Advanced Clear-Sky Processor for Ocean (ACSPO) and validation against nighttime radiances, *J. Geophys. Res.*, **114**, D067112, doi:10.1029/2008JD010960.

Martins, J., D. Tanre, L. Remer, Y. Kaufman, S. Mattoo, and R. Levy 2002: MODIS cloud screening for remote sensing of aerosols over oceans using spatial variability, *Geophys. Res. Lett.*, 29(12), 8009, doi:10.1029/2001GL013252.

Merchant, C., A. Harris, E. Maturi, and S. MacCallum, 2005: Probabilistic physically based cloud screening of satellite infrared imagery for operational sea surface temperature retrieval. *Q.J.R. Meteorol. Soc*., **131**, 2735–2755.

Merchant, C., L. Horrocs, J. Eyre and A. O'Carrol, 2006: Retrievals of sea surface temperature from infrared imagery: origin and form of systematic errors. *Q.J.R. Meteorol. Soc*., **132**, 1205–1223.

Merchant, C.J., P. Le Borgne, A. Marsouin, and H. Roquet, 2008: Optimal estimation of sea surface temperature from split window observations. *Rem. Sens. Environ.*, **112**, 445–457.

Merchant, C., P. Le Borgne., H. Roquet., and A. Marsouin, 2009: Sea surface temperature from a geostationary satellite by optimal estimation. *Rem. Sens. Environ.*, **113**, 445–457.

OS&I SAF SST Product User Manual, v2.1, Meteo-France, Nov 2009, http://www.osi-saf.org/biblio/docs/ss1_pum_leo_sst_2_1.pdf

Petrenko, B., A. Ignatov, Y. Kihai, A. Heidinger, A., 2010: Clear-sky mask for the Advanced Clear-Sky Processor for Oceans, *J. Tech.*, **27**, 1609–1623, doi:10.1175/2010JTECHA1413.1.

Petrenko, B., A. Ignatov, N. Shabanov, Y., Kihai, 2011: Development and evaluation of Sea Surface Temperature algorithms for GOES-R ABI using MSG SEVIRI as proxy, *Rem. Sens. Environ.*, Submitted Feb 2011.

Reynolds, R., N. Rayner, T. Smith, D. Stokes, and W. Wang, 2002: An improved *in situ* and satellite SST analysis for climate. *J. Clim*., **15**, 1609–1625.

Reynolds, R., T. Smith, C. Liu, D. Chelton, K. Casey, and M. Schlax, 2007: Daily high-resolution-blended analyses for sea surface temperature, *J. Climate*, **20**, 5473–5496

Schmetz J., P. Pili, S. Tjemkes, D. Just, J. Kerkmann, S. Rota, A. Ratier, 2002: An Introduction to Meteosat Second Generation, *Bull. Amer. Soc.*, **83**, 977-992.

Schmit, T., M. Gunshor, W. Menzel, J. Gurka, J. Li, and A. Bachmeier, 2005: Introducing the next-generation Advanced Baseline Imager on GOES-R. *Bull. Amer. Soc.*, 86, 1079–1096.

Shabanov N., A. Ignatov, B. Petrenko, Y. Kihai, 2009: Prototyping SST retrievals from GOES-R ABI with MSG SEVIRI data. AMS Conf, Jan 11-15, 2009, Phoenix, AZ (poster JP1.12), http://ams.confex.com/ams/pdfpapers/143903.pdf

Thie´baux, J., E. Rogers, W. Wang, and B. Katz, 2003: A new high-resolution blended real-time global sea surface temperature analysis. *Bull. Amer. Meteor. Soc*., 84, 645–656.

Walton C., 1998: Nonlinear Multichannel Algorithms for Estimating Sea Surface Temperature with AVHRR Satellite Data. *J. Appl. Meteorol*, **27**, 115-124.

WWW links:

[WWW1] GOES-R web-site,
www.goes-r.gov

[WWW2] GOES-R Algorithm Working Group,
http://goesrtest2.woc.noaa.gov/users/awg.html

[WWW3] Han Y., P. Van Delst, Q. Liu, F. Weng, B. Yan, and J. Derber, 2005: User's guide to the JCSDA Community Radiative Transfer Model (Beta v.), NOAA/NASA/JCSDA,
www.star.nesdis.noaa.gov/smcd/spb/CRTM/crtm-code/CRTM_UserGuide-beta.pdf

[WWW4] NCEP/GFS atmospheric fields,
www.emc.ncep.noaa.gov/modelinfo

[WWW5] SST Quality Monitor, SEVIRI section,
www.star.nesdis.noaa.gov/sod/sst/squam/SEV/

Table Captions:

Table 1. ABI Cloud Mask tests relevant to Ocean applications. Shaded are Tests, applicable to SST product, but excluded from the mask to preserve temporal continuity.

Table 2. Cross-walking between ABI CM and SST QC categories.

Table 3. Parameters of the SST QC Tests and their values.

Figure Captions:

Fig. 1. Sequential implementation of ABI CM and SST QC in the GOES-R AWG operational framework. Top row shows a schematic plot of the implementation. Bottom row illustrates three stages of processing by showing diurnal cycle of: ABI CM (first), thermal-channels based ABI CM (second) and SST QC (third) categories. Each category is defined as the amount of pixels falling in the state normalized by total amount of ocean pixels. MSG-2 SEVIRI 15-min FD data on June 03, 2008.

Fig. 2. Comparison of spatial distribution of Full Disc (FD) and zoom-in of ABI CM (a) and SST QC (b) masks. Large portion of the ABI CM image is covered by Probably categories. ABI CM seems insensitive to aerosols, while SST QC mask out aerosols-covered regions (i.e., Saharan coast). MSG-2 SEVIRI 15-min FD data at 12:00 UTC on June 03, 2008.

Fig. 3. Confusion matrix between ABI CM and SST QC masks. Components of the matrix are calculated as the amount of pixels falling in each category normalized by the total amount of ocean pixels. Color-coded are the three components of interest of the confusion matrix: 'Both CM and QC' (green), 'Only QC' (blue), 'Only CM' (red). Snapshot of the confusion matrix for 12:00 UTM is shown on the left, while diurnal cycle of components is on the right. MSG-2 SEVIRI 15-min FD data on June 03, 2008.

Fig. 4. Spatial distribution of SST anomalies (Hybrid SST – Daily Reynolds SST) screened with ABI CM (a) and SST QC (b) masks. White areas correspond to SST anomalies < -2.0[C]. MSG-2 SEVIRI 15-min Full Disk (FD) data at 12:00 UTC on June 03, 2008.

Fig. 5. Statistics of the SST anomalies (Hybrid SST – Daily Reynolds SST) screened with ABI CM (in red) and SST QC (in blue) masks. MSG-2 SEVIRI 15-min FD data on June 03, 2008.

Fig.6. Statistics of the SST anomalies (Hybrid SST – Daily Reynolds SST) screened with components of the ABI CM and SST QC confusion matrix 'Both CM and QC' (green), 'Only QC' (blue), 'Only CM' (red). MSG-2 SEVIRI 15-min FD data on June 03, 2008.

Fig.7. Statistics of the Ch09 and Ch10 BT anomalies (BT SEVIR - CRTM) screened with ABI CM (in red), SST QC (in blue). MSG-2 SEVIRI 15-min FD data on June 03, 2008.

Fig. 8. Impact of First Guess (FG) and Reference Fields on statistics (Mean and STD of SST anomalies) of the retrieved SST anomalies screened with the SST QC mask. MSG-2 SEVIRI 15-min (at hourly interval) FD data for June 2008.

Fig. 9. Dependence on View Zenith Angle of statics (Mean and STD of SST anomalies and Number of observations). Cut-off is 60° for SST applications. MSG-2 SEVIRI 15-min Full Disk (FD) data at 12:00 UTC on June 03, 2008.

Fig. 10. Panel (a) Diurnal cycle of data screening rate by individual ABI CM filters. Also shown ABI CM mask- area between dotted lines corresponds 'Pr. Clear', a transition between 'Cloudy+ Pr. Cloudy' (bottom part) and 'Clear' (upper part). MSG-2 SEVIRI 15-min FD data on June 03, 2008. Panel (b) Diurnal cycle of data screening rate by individual SST QC tests. Also shown SST QC mask- area between dotted lines corresponds 'Sub-Optimal', a transition between 'Poor' (bottom part) and 'Optimal' (upper part).The same date as in Fig 10-a.

Fig. 11. Time series of Unique Contribution (left) and Unique Omission (right) of SST QC filters. MSG-2 SEVIRI 15-min FD data on June 03, 2008.

Fig. 12. Filters Transmission analysis: (a) Transmission curve and other histograms for 'SST Uniformity' filter, (b) Cross-comparison of transmission curves for SST QC individual filters, SST QC mask and ABI CM mask. Note, Transmission curve values are shown in percentage on the left-hand side scale, while histograms are presented in pixels numbers at the right-hand side scale.

Fig. 13. Images for SST screened with 'SST Uniformity' filter of SST QC (a)Test is triggered ON, (b) Test is triggered OFF. MSG-2 SEVIRI 15-min Full Disk (FD) data at 12:00 UTC on June 03, 2008.

Fig. 14. Sensitivity of Mean, STD and NOBS to selection of threshold (a) and STD (b) of the Uniformity filter. Selected values for both parameters are marked with dashed line. MSG-2 SEVIRI 15-min Full Disk (FD) data at 12:00 UTC on June 03, 2008.

Table 1. *ABI Cloud Mask tests relevant to Ocean applications. Shaded are Tests, applicable to SST product, but excluded from the mask to preserve temporal continuity.*

Test ID	Test Name	Description
09	RUT- Reflectance Uniformity Test	STD of the observed 0.6 μm reflectance within a 3x3 box surrounding each pixel checked against a globally fixed threshold (Reflectance analog of TUT).
10	TUT- Thermal Uniformity Test	STD of the observed 11 μm BT within a 3x3 box surrounding each pixel checked against a globally fixed threshold (Thermal analog of RUT).
11	RTCT- Relative Thermal Contrast Test	BT difference @ the 11 μm (Pixel minus Nearest Warm Center in 5x5 box) checked against a globally fixed threshold.
12	ENTROP- Emissivity at Tropopause.	Effective emissivity of a pixel is compared against a fixed threshold. For cloud at the tropopause, emissivity is elevated, while for clear sky it approaches 0.
13	PFMFT- Positive 4-5 Test	Split window test for semi-transparent cloud ΔBT=BT(11 μm)-BT(12 μm) is checked against the pre-calculated ΔBT as a function of BT(11 μm).
16	RFMFT - Relative 4-5 Test	Split-window test. Significant deviations of pixel's ΔBT in 5x5 box from that at the NWC (positive or negative) are indicative of cloud.
17	RGCT - Reflectance Gross Contrast Test	Clouds exhibit large values of the visible reflectance compared to clear sky.
18	RVCT - Relative VIS Contrast Test	Relative VIS Contrast Test - over small region (3x3 box, cloud edge), cloudy pixels have largest contrast in VIS reflectance. Unlike RGCT, the RVCT test dynamically calculates its thresholds.
21	EMS4 – 3.9 μm Emissivity Test	3.9 μm emissivity for clouds is augmented, and near zero for clear sky.
22	ULST - Uniform Low Stratus Test	Low uniform stratus clouds are more reflective (less emissive) than the surface in the 3.9 μm. Test compares pixel emissivity with clear sky prediction @ Night.

Table 3. *Cross-walking between ABI CM and SST QC categories.*

ABI CM		SST QC	
Clear	3	0	Optimal
Probably Clear	2	0	Optimal
Probably Cloudy	1	1	Sub-Optimal
Cloudy	0	2	Poor
----- / --------		3	Unprocessed

Table 4. *Parameters of the SST QC Tests and their values.*

Test	Parameter	Description
Radiance	STD(BT09) = 1.0	STD of BT Ch09, a normalization parameter in Eq. (1)
	STD(BT10) = 1.0	STD of BT Ch10, a normalization parameter in Eq. (1)
	Threshold = 4.0	Threshold value for the Test, Eq. (1)
Adaptive SST	Wnd = 15	Window size for the Test
	f = 80%	Min percentage of poor quality (SST QC= 'Poor') pixels required to calculate statistics over cluster of poor pixels in the window. If insufficient this test will be skipped for this pixel.
	k = 8.0	Multiplication factor in Eq. (2a) for initial screening of pixels in poor quality cluster (SST QC= 'Poor').
SST Uniformity	Wnd = 3	Window size for the Test
	STD(SST) = 0.8	STD of SST, a normalization parameter in Eq. (3)
	Threshold= 3.0	Threshold value for the test, Eq. (3)

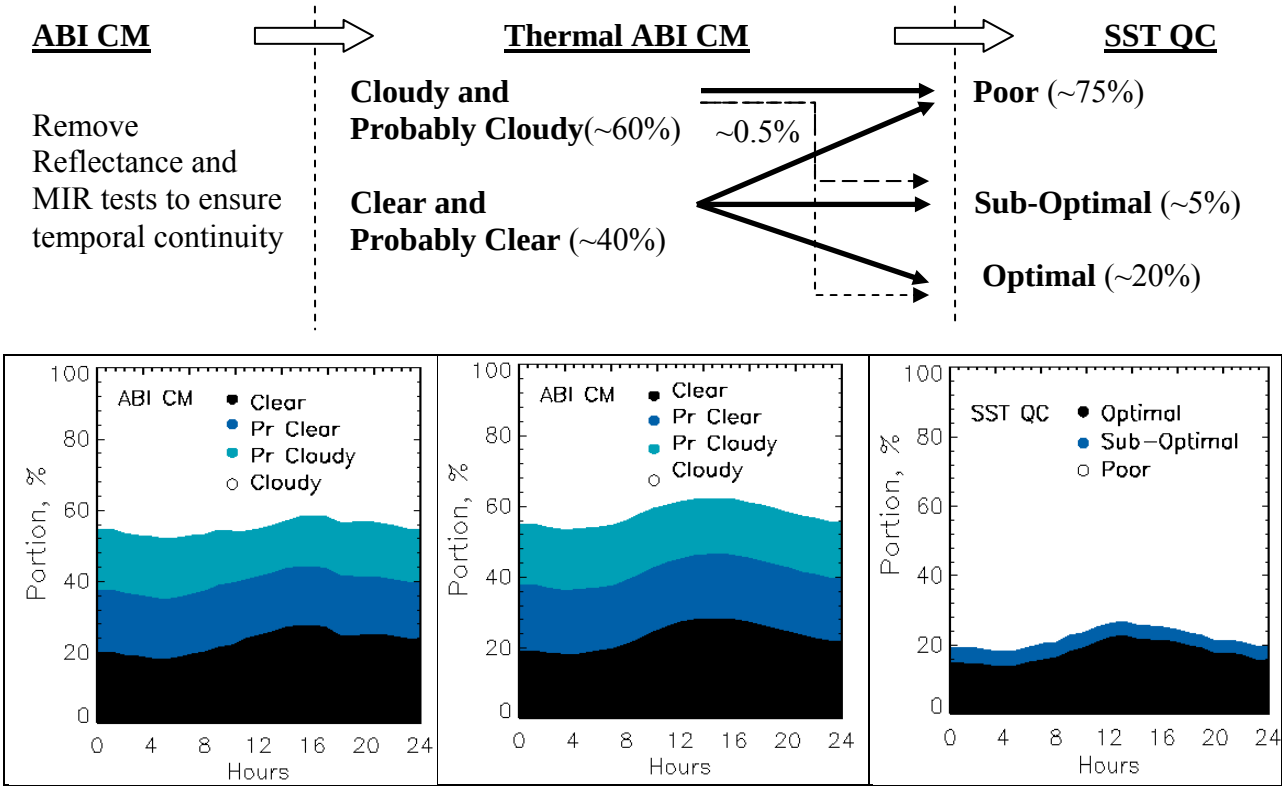

Fig. 1. *Sequential implementation of ABI CM and SST QC in the GOES-R AWG operational framework. Top row shows a schematic plot of the implementation. Bottom row illustrates three stages of processing by showing diurnal cycle of: ABI CM (first), thermal-channels based ABI CM (second) and SST QC (third) categories. Each category is defined as the amount of pixels falling in the state normalized by total amount of ocean pixels. MSG-2 SEVIRI 15-min FD data on June 03, 2008.*

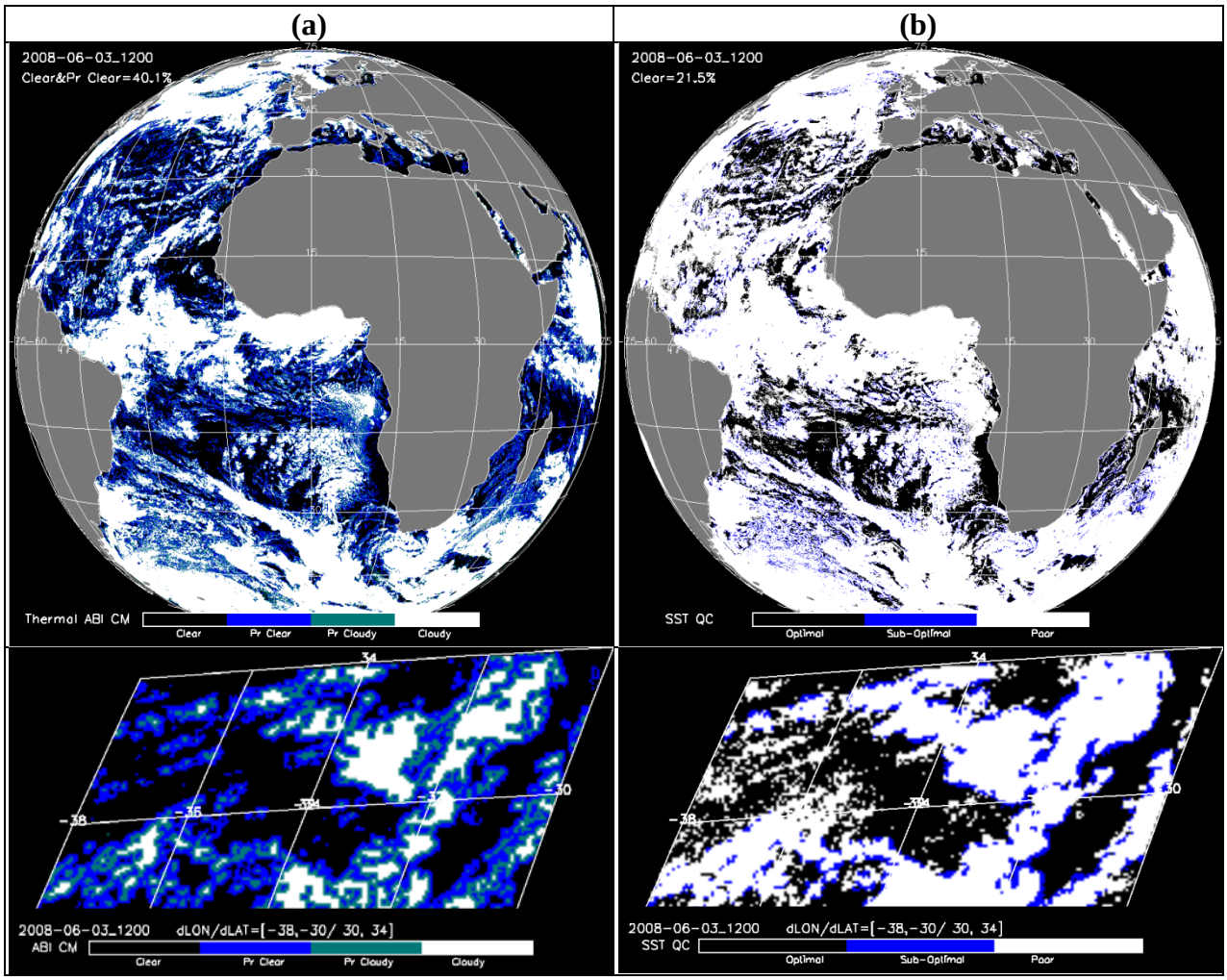

Fig. 2. *Comparison of spatial distribution of Full Disc (FD) and zoom-in of ABI CM (a) and SST QC (b) masks. Large portion of the ABI CM image is covered by Probably categories. ABI CM seems insensitive to aerosols, while SST QC mask out aerosols-covered regions (i.e., Saharan coast). MSG-2 SEVIRI 15-min FD data at 12:00 UTC on June 03, 2008.*

June 03, 2008 12:00 UTC	SST QC	Optimal	Sub-Optimal	Poor
ABI CM	100% =	26.6% +	4.7% +	68.6%
Clear	27.6% +	21.3%	0.7%	5.6%
Probably Clear	12.8%+	4.9%	3.4%	9.4%
Probably Cloudy	15.5% +	0.3%	0.5%	14.7%
Cloudy	39.1%	0.1%	0.12%	38.9%

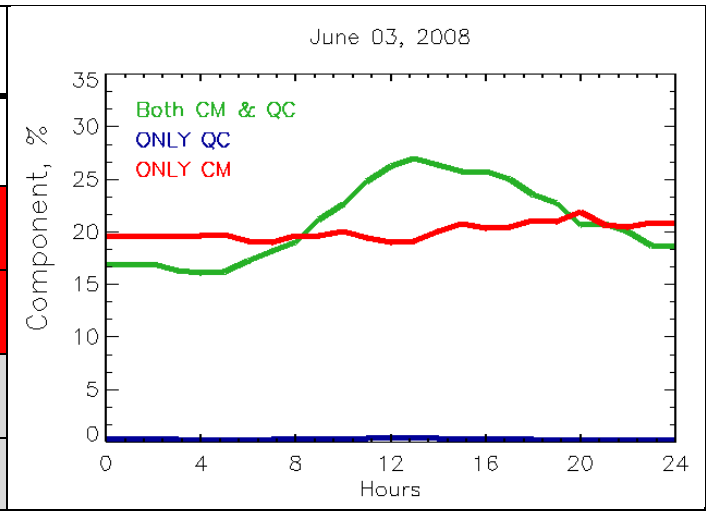

Fig. 3. *Confusion matrix between ABI CM and SST QC masks. Components of the matrix are calculated as the amount of pixels falling in each category normalized by the total amount of ocean pixels. Color-coded are the three components of interest of the confusion matrix: 'Both CM and QC' (green), 'Only QC' (blue), 'Only CM' (red). Snapshot of the confusion matrix for 12:00 UTM is shown on the left, while diurnal cycle of components is on the right. MSG-2 SEVIRI 15-min FD data on June 03, 2008.*

(a)	(b)

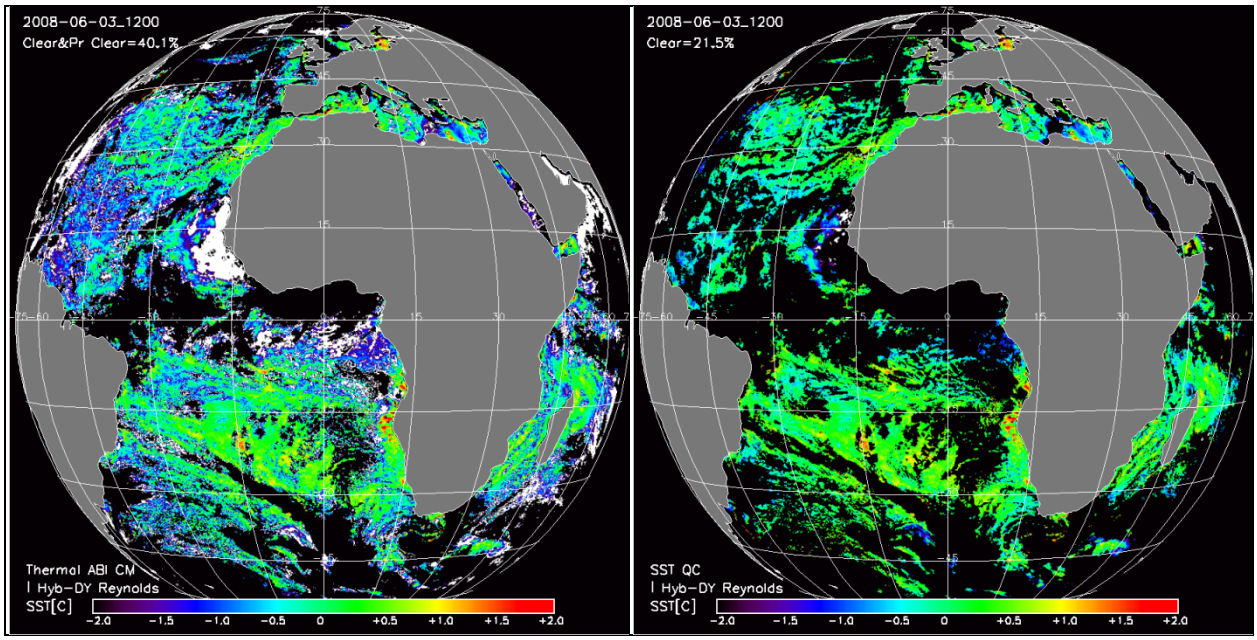

Fig. 4. *Spatial distribution of SST anomalies (Hybrid SST – Daily Reynolds SST) screened with ABI CM (a) and SST QC (b) masks. White areas correspond to SST anomalies < -2.0[C]. MSG-2 SEVIRI 15-min Full Disk (FD) data at 12:00 UTC on June 03, 2008.*

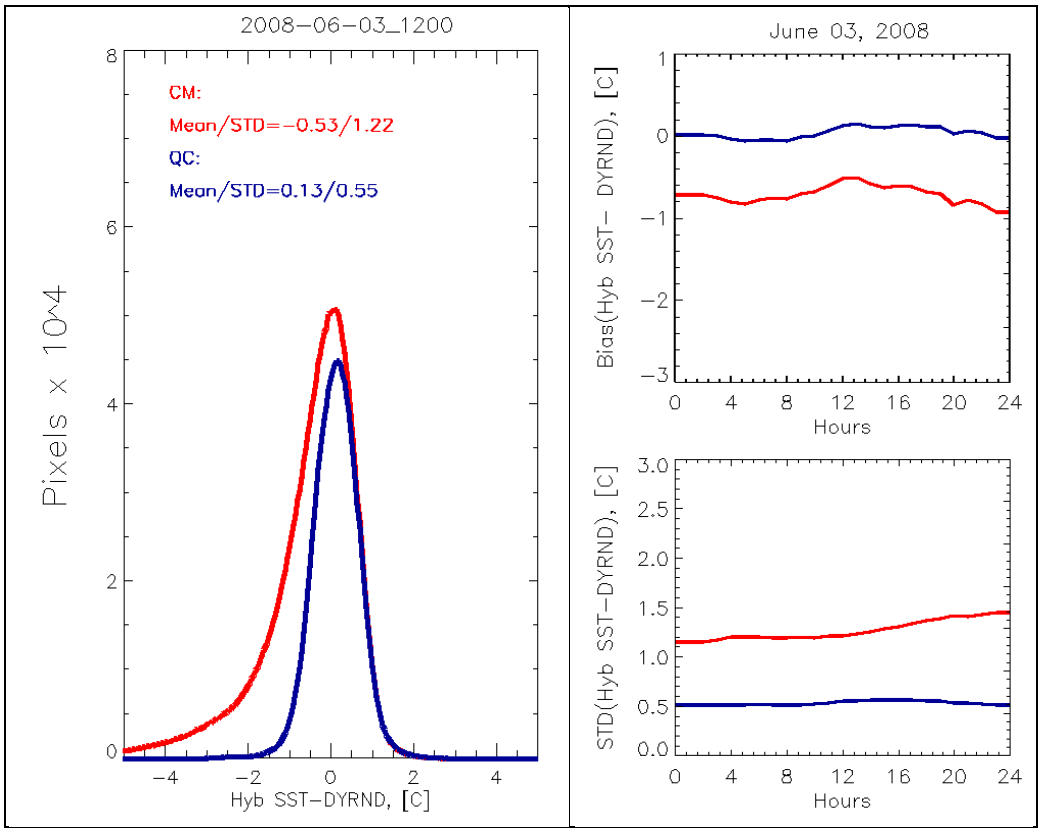

Fig. 5. *Statistics of the SST anomalies (Hybrid SST – Daily Reynolds SST) screened with ABI CM (in red) and SST QC (in blue) masks. MSG-2 SEVIRI 15-min FD data on June 03, 2008.*

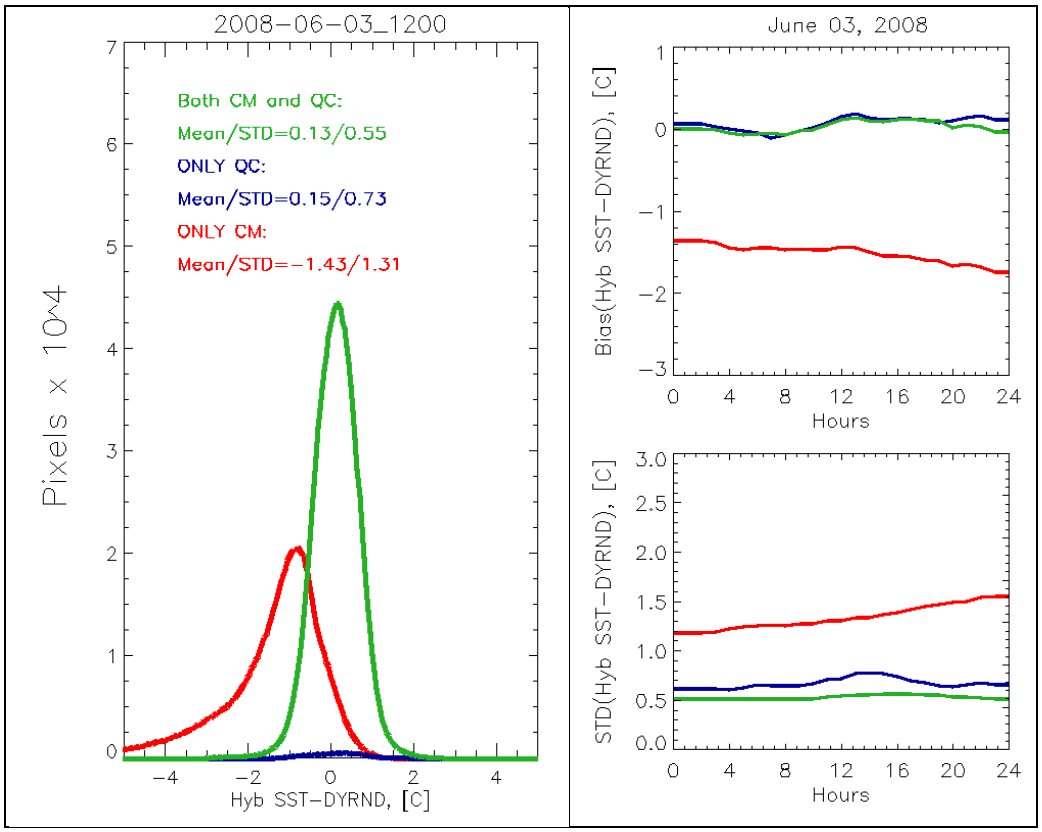

Fig.6. *Statistics of the SST anomalies (Hybrid SST – Daily Reynolds SST) screened with components of the ABI CM and SST QC confusion matrix 'Both CM and QC' (green), 'Only QC' (blue), 'Only CM' (red). MSG-2 SEVIRI 15-min FD data on June 03, 2008.*

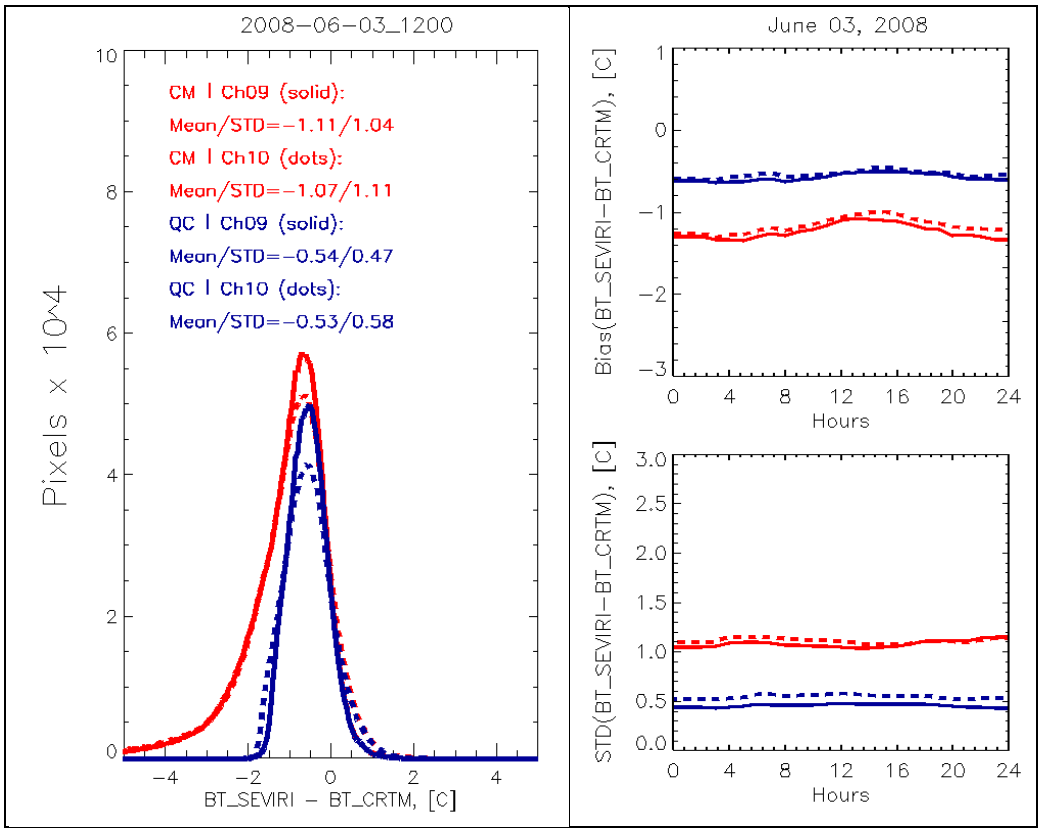

Fig.7. *Statistics of the Ch09 and Ch10 BT anomalies (BT SEVIR - CRTM) screened with ABI CM (in red), SST QC (in blue). MSG-2 SEVIRI 15-min FD data on June 03, 2008.*

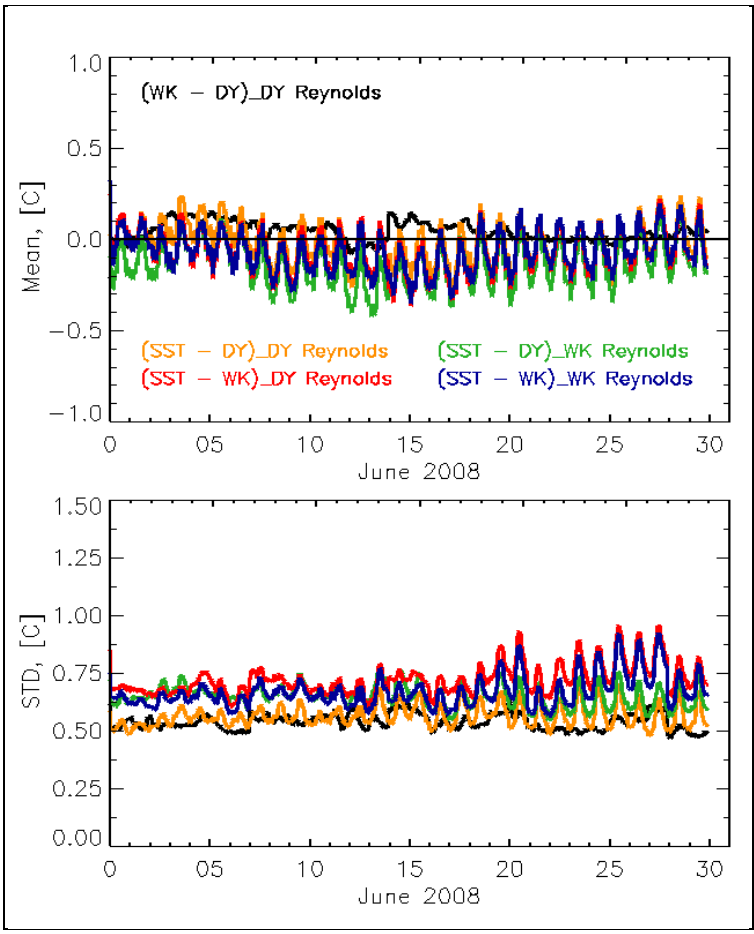

Fig. 8. *Impact of First Guess (FG) and Reference Fields on statistics (Mean and STD of SST anomalies) of the retrieved SST anomalies screened with the SST QC mask. MSG-2 SEVIRI 15-min (at hourly interval) FD data for June 2008.*

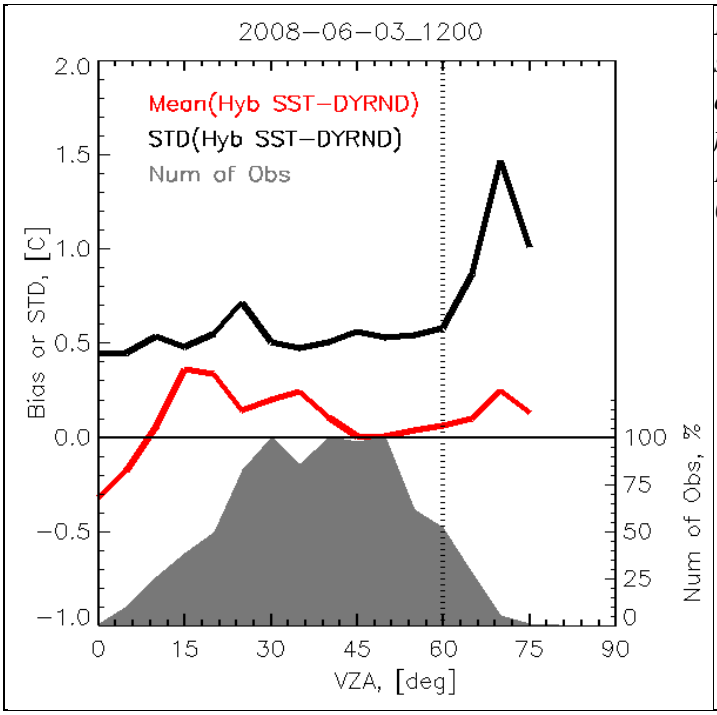

Fig. 9. *Dependence on View Zenith Angle of statics (Mean and STD of SST anomalies and Number of observations). Cut-off is 60° for SST applications. MSG-2 SEVIRI 15-min Full Disk (FD) data at 12:00 UTC on June 03, 2008.*

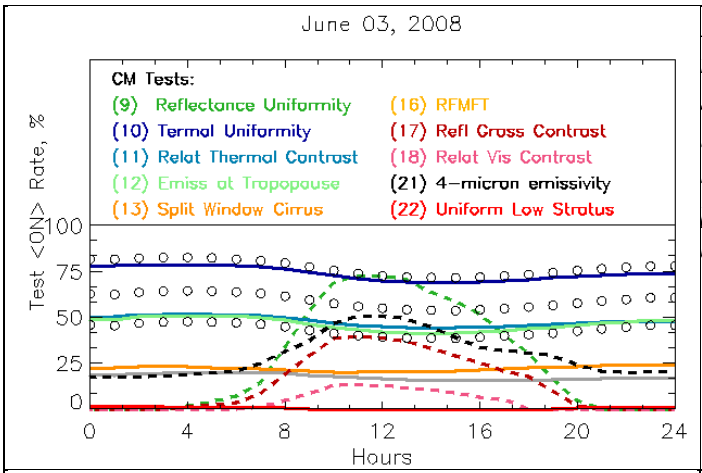

Fig. 10-a. *Diurnal cycle of data screening rate by individual ABI CM filters. Also shown ABI CM mask- area between dotted lines corresponds 'Pr. Clear', a transition between 'Cloudy+ Pr. Cloudy' (bottom part) and 'Clear' (upper part). MSG-2 SEVIRI 15-min FD data on June 03, 2008.*

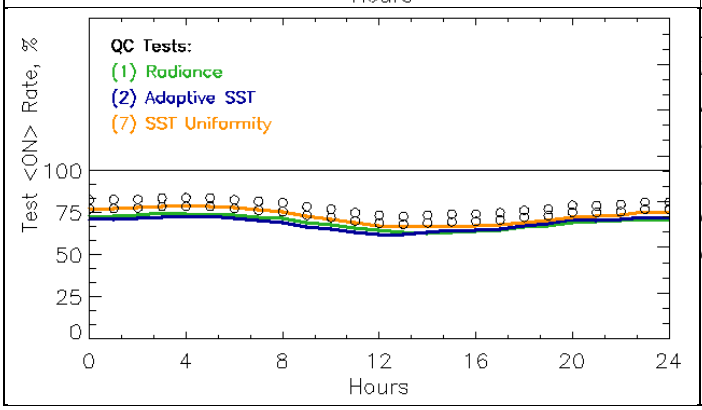

Fig. 10-b. *Diurnal cycle of data screening rate by individual SST QC tests. Also shown SST QC mask- area between dotted lines corresponds 'Sub-Optimal', a transition between 'Poor' (bottom part) and 'Optimal' (upper part).The same date as in Fig 10-a.*

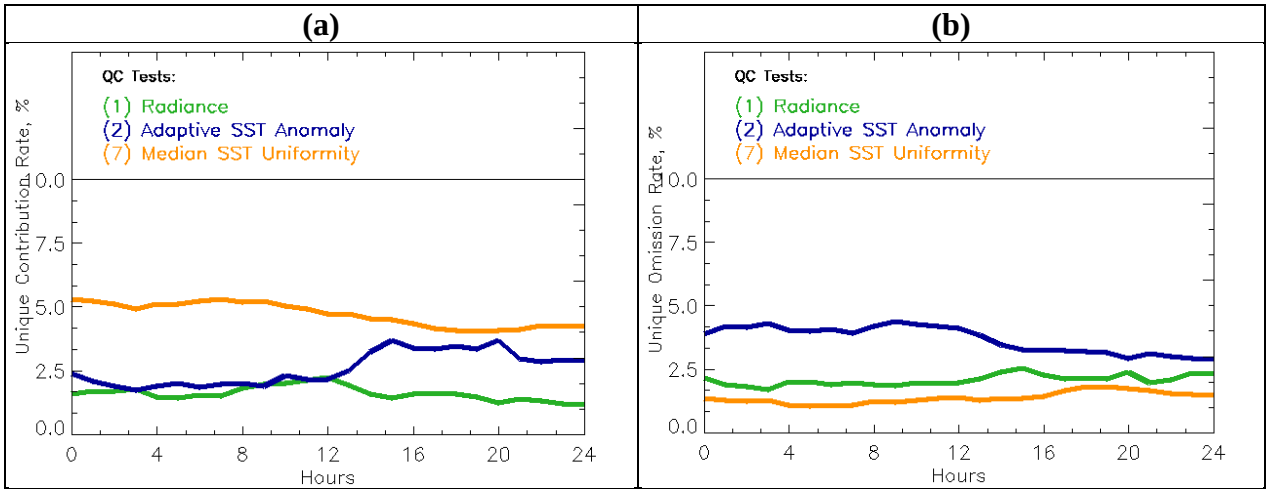

Fig. 11. *Time series of Unique Contribution (left) and Unique Omission (right) of SST QC filters. MSG-2 SEVIRI 15-min FD data on June 03, 2008.*

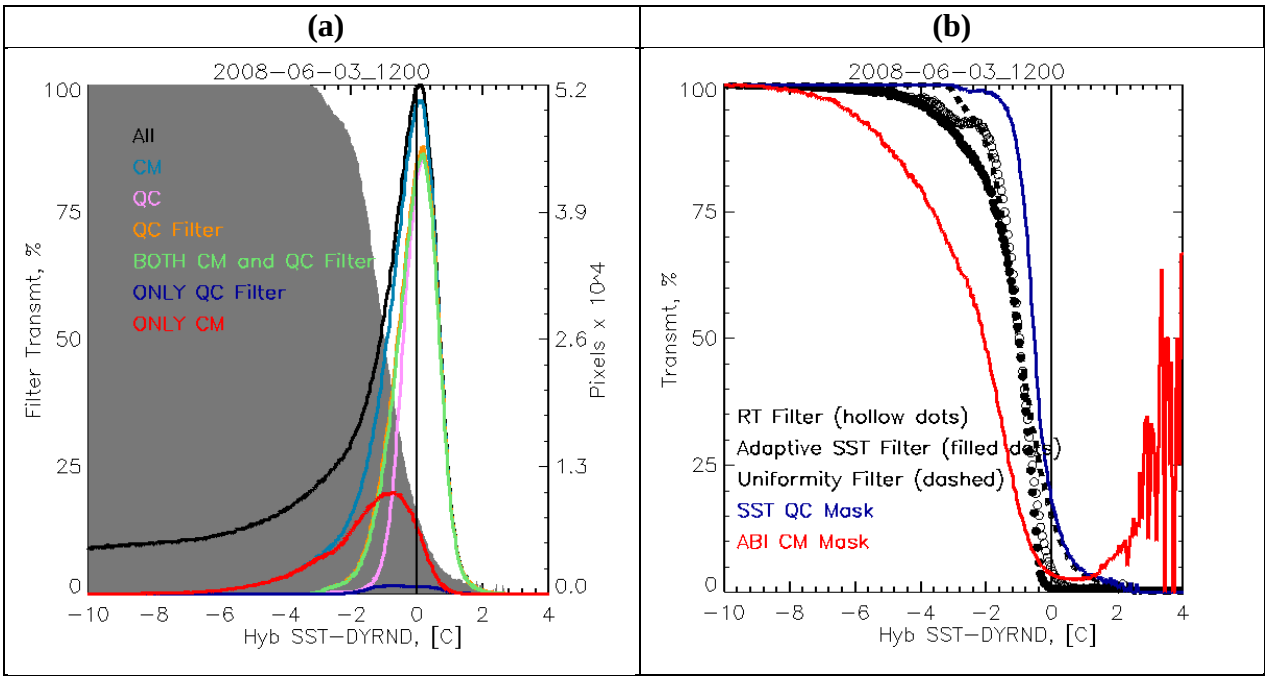

Fig. 12. *Filters Transmission analysis: (a) Transmission curve and other histograms for 'SST Uniformity' filter, (b) Cross-comparison of transmission curves for SST QC individual filters, SST QC mask and ABI CM mask. Note, Transmission curve values are shown in percentage on the left-hand side scale, while histograms are presented in pixels numbers at the right-hand side scale.*

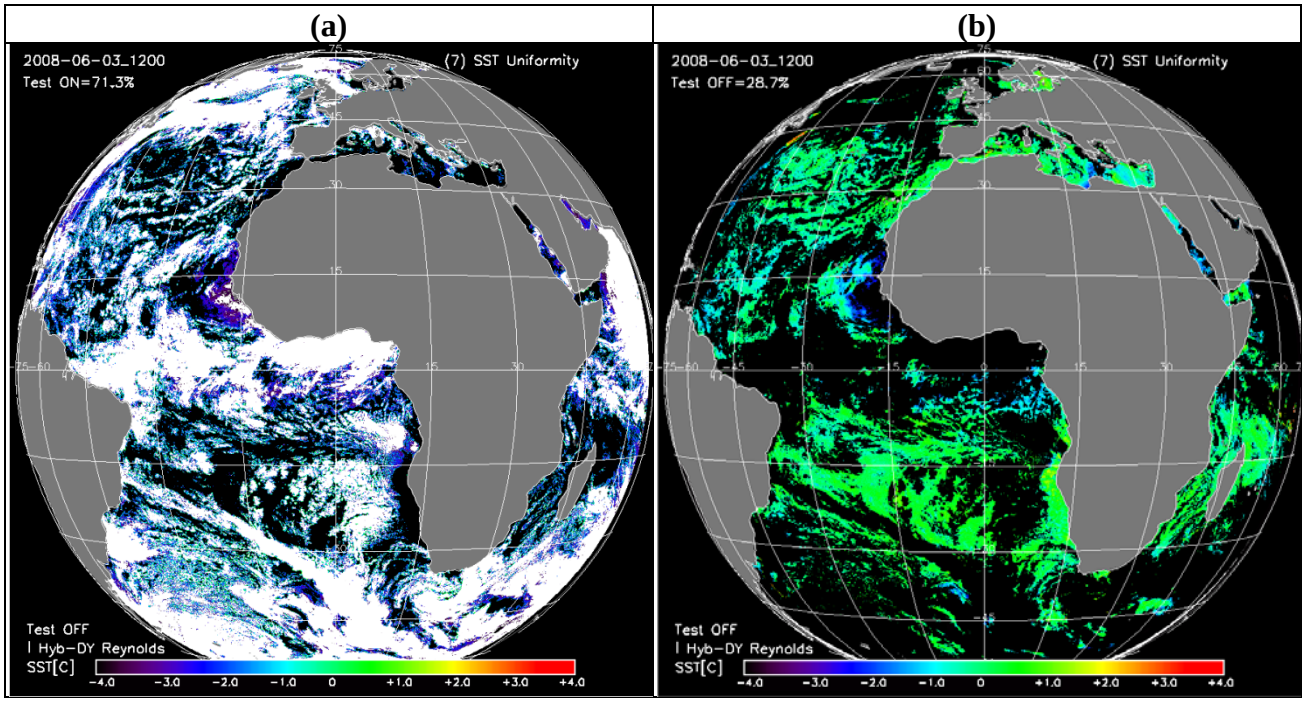

Fig. 13. *Images for SST screened with 'SST Uniformity' filter of SST QC (a)Test is triggered ON, (b) Test is triggered OFF. MSG-2 SEVIRI 15-min Full Disk (FD) data at 12:00 UTC on June 03, 2008.*

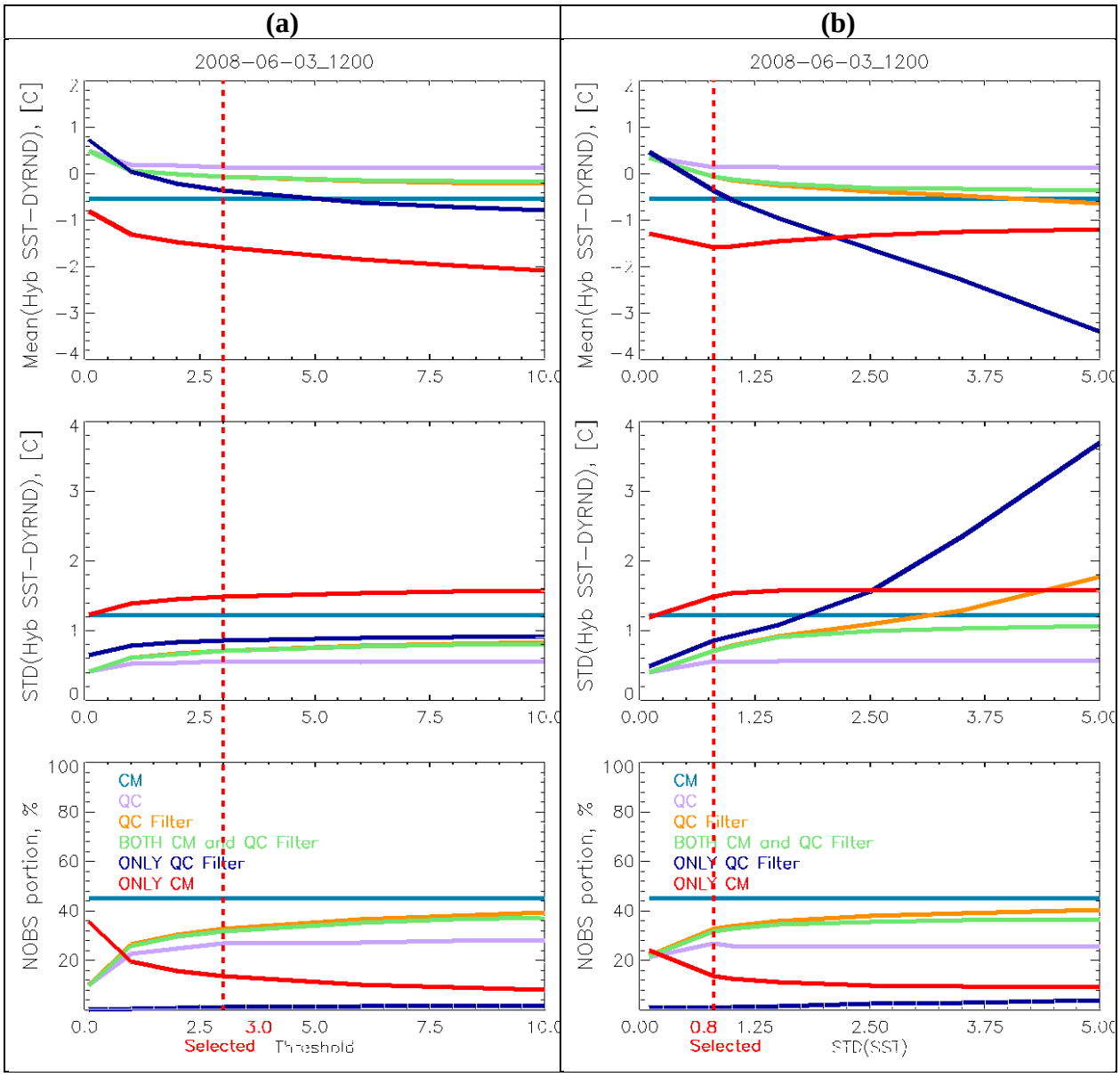

Fig. 14. *Sensitivity of Mean, STD and NOBS to selection of threshold (a) and STD (b) of the Uniformity filter. Selected values for both parameters are marked with dashed line. MSG-2 SEVIRI 15-min Full Disk (FD) data at 12:00 UTC on June 03, 2008.*