JP1.13

TOWARDS CONTINUOUS ERROR CHARACTERIZATION OF SEA SURFACE TEMPERATURE IN THE ADVANCED CLEAR-SKY PROCESSOR FOR OCEANS

Feng Xu[*1,2], Alexander Ignatov[1] and XingMing Liang[1,2]

1. NOAA / NESDIS / STAR, Camp Springs, MD; 2. CSU / CIRA, Fort Collins, CO

ABSTRACT

A new approach to error characterization in Sea Surface Temperature (SST) is proposed and tested with the SST data of NOAA-17, NOAA-18, and MetOp-A produced by the Advanced Clear Sky Processor for Oceans (ACSPO v1.0) newly developed in NESDIS. The approach is based on using continuous analytical functions rather than conventional look-up tables. Also, SST anomalies are analyzed with respect to the 0.25° daily Optimum Interpolation (OI v2.0) Reynolds and 0.05° OSTIA SST analyses rather than with respect to *in situ* SSTs. Bias dependencies are represented by continuous analytical functions, whose parameters are estimated through curve/surface fitting after binning along the factors. These functions have a simple form and have been constructed with consideration of the major physical factors affecting errors in the retrieval SST. Preliminary results of fitting bias vs. number of ambient clear-sky pixel (NAC) are presented. The fit parameters are largely consistent across different platforms as well as relatively stable as a function of time. The analyses in this study are only performed for nighttime ACSPO products and the focus is on the mean error (bias). Future work will include analyses of daytime data and of standard deviation (stdv) of error. Also, the same methodology will be employed to evaluate the model minus observation bias in ACSPO (Liang et al., 2009).

KEYWORDS – Error characterization, Sea Surface Temperature, ACSPO

1. INTRODUCTION

Sea surface temperatures (SST) have been operationally retrieved from satellites since the 1970s from infrared (IR) top-of-atmosphere brightness temperatures (BT) measured by sensors, such as the Advanced Very High Resolution Radiometer (AVHRR). The 3.7 and 10-12 μm IR spectral windows are usually employed and SST is estimated as a linear combination of BTs in dual or triple windows (e.g., McClain et al., 1985; Walton, 1988). The National Oceanic and Atmospheric Administration (NOAA) maintains a series of operational satellites, which carry AVHRR instruments and from which SST products of high resolution (4 km) and accuracy (approaching 0.4K rms) (Kearns et al., 2000) have been generated. The European satellite MetOp-A launched in 2006 also carries an AVHRR sensor, and its data are analyzed in this study as well.

Customarily, accuracy and precision of satellite SST are evaluated against *in situ* SST and characterized by one global mean bias and one standard deviation (stdv) of retrieved minus *in situ* SST (e.g., Walton et al., 1998; Kilpatrick et al., 2001). For the purpose of quality control, the retrieved SST in each pixel may further contain quality flags or confidence levels. The Group for High-Resolution SST (GHRSST) has become increasingly interested in supplying error estimates in each individual pixel, the so-called Single-Sensor Error Statistics (SSES) (Donlon et al, 2007). Such information is critically important for blending different SST products. Also, these estimates may be useful for SST developers to evaluate the performance of their SST product, identify potential areas for improvement, and guide its future optimization. For SST users, the SSES can provide comprehensive quality information about the SST product and facilitate decisions on the domains of data suitable for their particular applications. Although the current ACSPO v1.0 does not comply with this GHRSST requirement yet, work is underway to incorporate the SSES in the future version of ACSPO. This study reports our intermediate progress towards this goal.

It has been shown that the currently operational regression SSTs have large observational-condition dependent biases (e.g., Merchant et al., 2008). They may arise from various sources, including error in specifying retrieval coefficients (e.g., Merchant and Le Borgne, 2004), residual cloud (e.g., Merchant et al., 2005), and aerosols (e.g., Walton et al., 1998; Vasquez-Cuervo et al., 2004), just to name a few. The linear regression SST may also be subject to two other systematic biases due to prior error and error arising from the non-linearity of the converting radiances to BTs at some extreme conditions (Merchant et al., 2008). Improving regression (e.g., Emery et al., 1994; Walton et al., 1998) and physical (i.e., RTM-based inversion) SST retrievals (e.g. Merchant et al., 2008) has been subject of ongoing research. Also, cloud screening and quality control vary largely from one SST product to another. It is thus imperative to develop an objective methodology to quantify SST errors as a function of the full retrieval space, for different SST products.

Currently, the SSESs are routinely provided by different SST producers to the GHRSST and used for global analyses such as OSTIA (Stark et al., 2007).

[*] Correspondence to Feng.Xu@noaa.gov

However, there have been few publications on the specifics of how the error characterization of satellite SSTs have been performed. For instance, Minnett and Evans (2006) introduced the 7-dimensional SSES hypercube of bias and stdv of MODIS SST as a function of season, day/night, latitude, view angle, surface temperature, water vapor proxy, and retrieval quality level. The hypercube technique has been used to provide SSES information to the GHRSST Pilot Project based on satellite-*in situ* match-up databases. Castro et al. (2008) studied SST error of infrared and microwave satellite retrievals vs. *in situ* observation. Bias/stdv dependencies on retrieval conditions, such as wind speed, water vapor, view angle, and SST, were represented using look-up tables and 1-to-5% stdv reduction was achieved after bias correction. Vazquez-Cuervo et al. (2004) analyzed correlations between the Pathfinder SST and the Along-Track Scanning Radiometer (ATSR-2) versus aerosol/cloud information derived from the Total Ozone Measuring Spectrometer (TOMS). They have shown that consistency between these two SST products can be improved after flagging aerosol/cloud-contaminated pixels using regression analyses of SST error vs. aerosol/cloud.

This study examines the SSES in the new SST product available from the Advanced Clear-Sky Processor for Oceans (ACSPO) recently developed at NOAA/NESDIS. Analyses were conducted with several months of global AVHRR data from NOAA-17, -18, and MetOp-A. Instead of *in situ* SST, the retrieved SSTs in this study were referenced to the global analysis SSTs. Two reference global daily analyses were used, the Optimum Interpolation (OI) daily 0.25° (Reynolds et al., 2007) SST (hereafter Reynolds SST), and the Operational SST and Sea Ice Analysis (OSTIA) daily 0.05° (Stark et al., 2007). This allows us to substantially increase the match-up statistics while largely preserving their quality because both reference fields used here have been anchored to *in situ* SST. SST anomaly was analyzed as a function of ambient clear-sky information available from ACSPO data itself, and atmospheric and surface information saved in ACSPO granules from the NCEP Global Forecast System (GFS) data. Pronounced dependencies of SST anomaly on the number of ambient clear-sky pixels (NAC), column water vapor, and view angle were observed. Our study initially focuses on the NAC dependence and its characterization, which was found to be one of the strongest factors affecting the ACSPO SST bias.

In contrast to the previous studies, which used the LUT approach (Minnett and Evans, 2006; Castro et al., 2008), this study employs continuous analytical functions to represent the error statistics vs. multiple environmental factors. An exponential fit function was tested to approximate the bias dependence vs. number of ambient clear-sky pixel (NAC) (cf., Liang et al., 2009). Note that NAC characterizes a fraction of ambient cloud, land, and ice, but it may also be an indicator of residual subpixel contamination by those factors (Dash and Ignatov, 2008). To minimize the effect of outliers, a robust least-square method was adopted in fitting. Cross-platform consistency was examined and long-term stability monitored over time for the fit coefficients of bias vs. NAC.

The ACSPO SST data and other reference/ancillary data used in this study are described in Section 2, where the definitions of retrieval conditions are also introduced. Section 3 presents the statistical model of error characterization and the corresponding methods for measurement and fitting. Analyses and discussion comprise Section 4.

2. DATA

The ACSPO has been developed at NOAA/NESDIS and its version 1.0 became operational in May 2008. Its further improvements are underway, with the objective to ultimately replace the heritage clear-sky radiances (CSR), SST, and aerosol products currently produced by the heritage Main Unit Task (MUT) system, which has been in operational use since the early 1980s (McClain et al., 1985; Walton et al., 1998). Similarly to the MUT, ACSPO generates AVHRR clear-sky radiances over oceans, from which SST and aerosol products are derived. A major improvement in ACSPO over the MUT is the full integration of global clear-sky AVHRR radiances with the National Centers for Environmental Prediction Global Forecast System (NCEP/GFS) atmospheric and Reynolds et al. (2002, 2007) SST fields. These fields are used as input to the Community Radiative Transfer Model (CRTM) to predict the top-of-atmosphere BTs under clear-sky conditions.

The SST data used in this study is the ACSPO nighttime clear-sky satellite regression SST. The algorithm used is the triple-window MCSST

$$\begin{aligned} T_S &= a_1 T_4 + a_2 T_3 + a_3 T_5 + a_4 (T_3 - T_5)(\sec\theta - 1) \\ &\quad + a_5 (\sec\theta - 1) + a_6 \end{aligned} \tag{1}$$

where T_3, T_4, T_5 are the measured AVHRR brightness temperatures in channels 3B, 4, and 5, and θ is the view zenith angle. In the first version of ACSPO, the form of the regression equation and the coefficients $a_1 \sim a_6$ have been adopted from MUT to quickly evaluate the consistency of the two SST systems. The cloud mask used in ACSPO v1.0 is documented in Petrenko et al. (2008).

Reynolds OI v2. daily 0.25° (AVHRR only) (Reynolds et al, 2007) and daily OSTIA 0.05° (Stark et al., 2007) SSTs were chosen as reference SSTs for anomaly analyses in this study. Reynolds SST (AVHRR only) is a bulk SST obtained by blending AVHRR satellite SST from NAVOCEANO SEATEMP system (May et al., 1998) and *in situ* observations. OSTIA is a foundation SST obtained by blending multiple infrared and microwave satellite retrievals with *in situ* observations. Using different references can help to quantify and minimize the reference-SST specific errors on the SSES.

Atmospheric data available in ACSPO granules are derived from the NCEP GFS data. The embedded reference SST data in ACSPO v1.0 come from the weekly Reynolds (Reynolds et al., 2002). Daily

Reynolds and OSTIA SST were appended to ACSPO granules offline for this study. Note that in ACSPO v1.1, which is currently being tested, weekly Reynolds was replaced by daily AVHRR-based SST product from Reynolds et al. (2007).

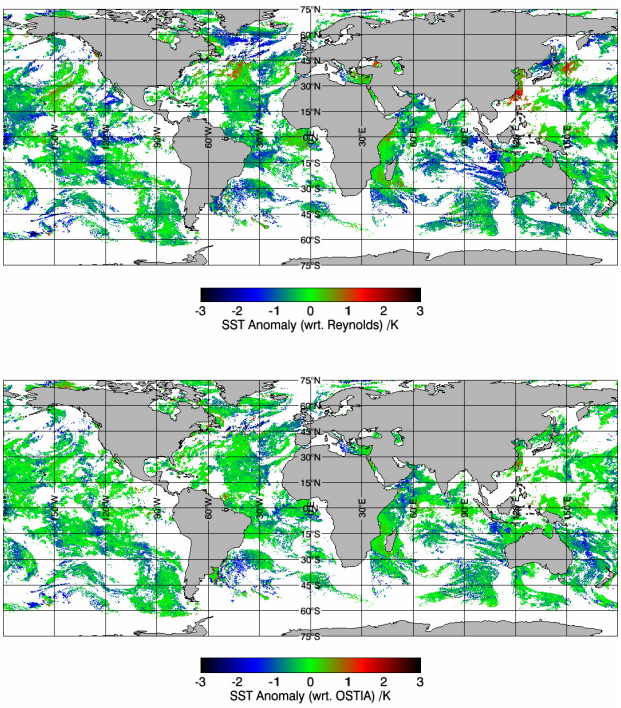


Figure 1. Global nighttime maps of SST anomalies from MetOp-A with respect to Reynolds and OSTIA for 17 September 2008.

Fig. 1 shows maps of the global SST anomalies "ACSPO-Reynolds" and "ACSPO-OSTIA" from nighttime MetOp-A for 17 September 2008. Cold anomalies are found in most areas. In particular, anomalies are more pronounced in dynamic areas (e.g., Gulf Stream) and near land areas (coastlines in the Northern Hemisphere). As observed from anomalies maps for other days (not shown here), anomalies not only vary in space but also in time. Interestingly, ACSPO SST appears closer to OSTIA than to Reynolds SST.

The geographical distribution of the bias in Fig.1 likely reflects distribution of the parameters that affect the SST retrieval. In what follows, a conglomerate of these parameters will be referred to as the "retrieval space," which can be loosely defined by three aspects: observation geometry, atmospheric status, and air-sea boundary conditions. Atmospheric state may include ambient clear-sky conditions, water vapor, and aerosol. In this study, the ambient clear-sky conditions are defined by the number of clear-sky ocean pixels in a predefined proximity of each pixel, NAC. It represents how many non-clear-sky ocean pixels (cloud, ice, land) are found in the vicinity of a pixel. Column water vapor is used here to represent the full water vapor profile. (Note that aerosol information has not been included, as well as the air-sea boundary conditions including SST itself, air-sea temperature difference and wind speed, and time history of those parameters that affect the surface fluxes.)

These factors span the multidimensional "retrieval space" and can be represented as a vector of factors

$$\vec{\alpha} = [N, \theta, W, V, T_s, T_a, \tau] \quad (2)$$

which include NAC, N; view angle, θ; water vapor, W; wind speed, V; aerosol optical depth, τ; SST, T_s; and air temperature, T_a. In this study, only NAC is considered, and view angle and column water vapor are preliminarily analyzed. Column water vapor (calculated by the integration of NCEP GFS water vapor profiles) and view angle are reported in the ACSPO granules. The NAC was calculated offline, using the ACSPO cloud, ice, and land mask information (Petrenko et al., 2008).

Note also that Minnett and Evans (2006) use previously available quality levels as one of the governing factors (additional dimension of the SSES hypercube). Here, we have adopted a different approach when full retrieval domain is considered in the SSES analyses, with the expectation that the error characterization itself should provide us such quality information.

NAC was counted within a circle of 12-pixel radius surrounding each clear-sky pixel. After performing tests using different radii, 12-pixel radius was found to be the optimal choice, which adequately accounts for the effect from ambient pixels. Search radii based on the distance (e.g., 50 km) were also tested but not adopted for routine processing due to the significant increment in computation time, which had only minor effects on the result. Note that at the edge of each scan, NAC can only be counted in a "partial circle." This limitation was compensated by scaling NAC to an equivalent full circle. The minimum NAC is zero and the maximum is ~470.

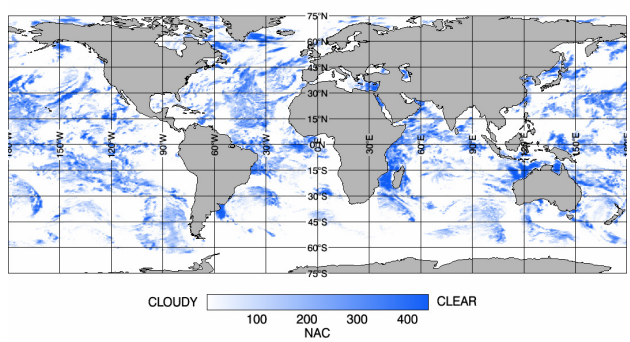


Figure 2. Global map of number of ambient clear-sky ocean pixels (NAC) corresponding to Fig.1.

Fig. 2 shows an example of an NAC global map derived from the same data as Fig. 1. Over most open ocean, the NAC indicates the probability (or the volume) of ambient (and potentially residual subpixel) cloud. Larger NAC means current pixel is situated in the

middle of mostly clear-sky area, which is also likely associated with a lower probability of subpixel residual cloud, and vice versa. Indirectly, NAC also serves as a measure of the distance from the current pixel to cloud, ice, or land. The smaller the NAC, the closer the distance to the edge of the SST valid retrieval domain.

3. METHODOLOGY

The basic assumption of error characterization is that the SST residual, in some statistical sense, continuously depends on the retrieval conditions. Our analyses have shown that probability density functions (PDF) of the biases are close to a Gaussian shape, which can be described by its mean and variance. Therefore, a normal distribution of the retrieval error can be represented as

$$\Delta T \sim N(b,\sigma) \tag{3}$$

where b and σ denote mean bias and stdv, respectively.

Since the error distribution continuously varies in the retrieval space spanned by environmental factors, its mean bias and stdv can be expressed as continuous functions of the retrieval space as follows

$$\begin{aligned} &b = f(\vec{\alpha}), \quad \sigma = g(\vec{\alpha}) \\ &\vec{\alpha} = [\alpha_i], \quad \alpha_i = N, \theta, W, ... \end{aligned} \tag{4}$$

where $\vec{\alpha}$ denotes the multidimensional vector of retrieval space. In this study, its elements α_i are NAC, N; view zenith angle, θ; and column water vapor, W. Note that higher orders of statistics of the distribution, such as skewness and kurtosis, should be taken into account in the same way as mean and stdv, if the real PDF deviates from a normal distribution.

The objective of error characterization is to estimate the dependence of b and σ on the retrieval space. These analytical functions can then be used to predict SST accuracy and precision in each retrieval point, and to perform SST bias corrections (e.g., Castro et al., 2008). In this study we concentrate on the mean bias, b.

For each sample of observations, there is an SST error and a set of corresponding measured and environmental factors. To estimate statistical information, retrieval space has to be first stratified into bins along each dimension. The b and σ are then estimated within each bin as

$$\begin{aligned} &\hat{b} = \sum_n \Delta\hat{T} \Big/ n \\ &\hat{\sigma} = \sum_n (\Delta\hat{T} - \hat{b})^2 \Big/ (n-1) \end{aligned} \tag{5}$$

Here, $\Delta\hat{T}$ is the observed bias in a pixel; $\hat{b}$, $\hat{\sigma}$ is the estimated mean bias and stdv; and n is the sample size within each bin. The estimates of $\hat{b}$ and $\hat{\sigma}$ can then be plotted in the retrieval space as data points as a function of bin position, from which continuous hyper-surfaces will be fitted. These hyper-surfaces are geometric realizations of the functions $f(\cdot)$ and $g(\cdot)$ in Eq.(4). Now the problem becomes a curve-fitting problem, i.e.,

$$\hat{b} = f_{\vec{\beta}}(\vec{\alpha}) + \varepsilon \tag{6}$$

where ε denotes the error between the binned estimate of bias, $\hat{b}$, and the fitted curve. $\vec{\beta}$ denotes the parameters of function $f(\cdot)$.

One possible approach is to empirically select the analytical form of the $f(\cdot)$ function and then fit its parameters via mathematical approaches, such as non-linear least square methods like Levenberg-Marquardt algorithm (e.g. Nocedal and Wright, 1999). In this case, the curve-fitting problem is in fact a minimization of errors between the data points and fit functions,

$$\vec{\beta}^* = \arg\min_{\vec{\beta}} \sum_m \left| \hat{b} - f_{\vec{\beta}}(\vec{\alpha}) \right|^2 \tag{7}$$

where m denotes the number of binned observations.

Assuming the form of the fit function is correctly selected, the error of observation comes from several sources. One major error source is the measurement error in the x (factors) and y (SST) values. Another major error may come from other potential factors that have not been included in the analysis due to a lack of measurement or forecast information about them (e.g., aerosol in this study). Errors may also result from the statistical estimation of bias/stdv within each bin, but these errors can be minimized by reducing the size of the bin while increasing statistics in it (such as by increasing time interval for data collection; this approach is explored below in this study).

Some errors may severely affect the fitting process or may even lead to its failure to converge. Hence, a high robustness of fitting is required to minimize the effect of these random and uncontrollable noises on the result. A robust fit method, based on an iterative reweighting least square (Holland and Welsch, 1977) was adopted in this study. It iteratively performs the weighted non-linear least square. The weights are calculated based on the distance of each data point's deviation from the fitted function in the last iteration. The bi-square like weight (DuMouchel and O'Brien, 1989) is used

$$w_i = \begin{cases} \left(1 - d_i^2 / D^2\right)^2 & d_i < D \\ 0 & else \end{cases}, \quad D = 6 \times \mathrm{Median}(d_i) \tag{8}$$

where d_i denotes the deviation of the i-th data point. This approach reduces the weight of each data point in inverse proportion to the fitting error, i.e. the further the data point deviates from the fit function, the less

significant its contribution will be to the fitting in the next iteration.

4. ANALYSIS AND CHARACTERIZATION of NAC DEPENDENCE

In the initial analyses reported here, the scope was limited to a more simple, as well as stable and uniform, subsample of the ACSPO SST product. This decision was made to minimize the uncertainty from the reference data, uncontrollable factors, and strong SST anomalies, which may affect the match-up error. Therefore, only nighttime ACSPO data were initially considered. The statistical analyses were performed in one-day and one-week increments to assure a statistically significant sample pool and to check sensitivity to the sample size.

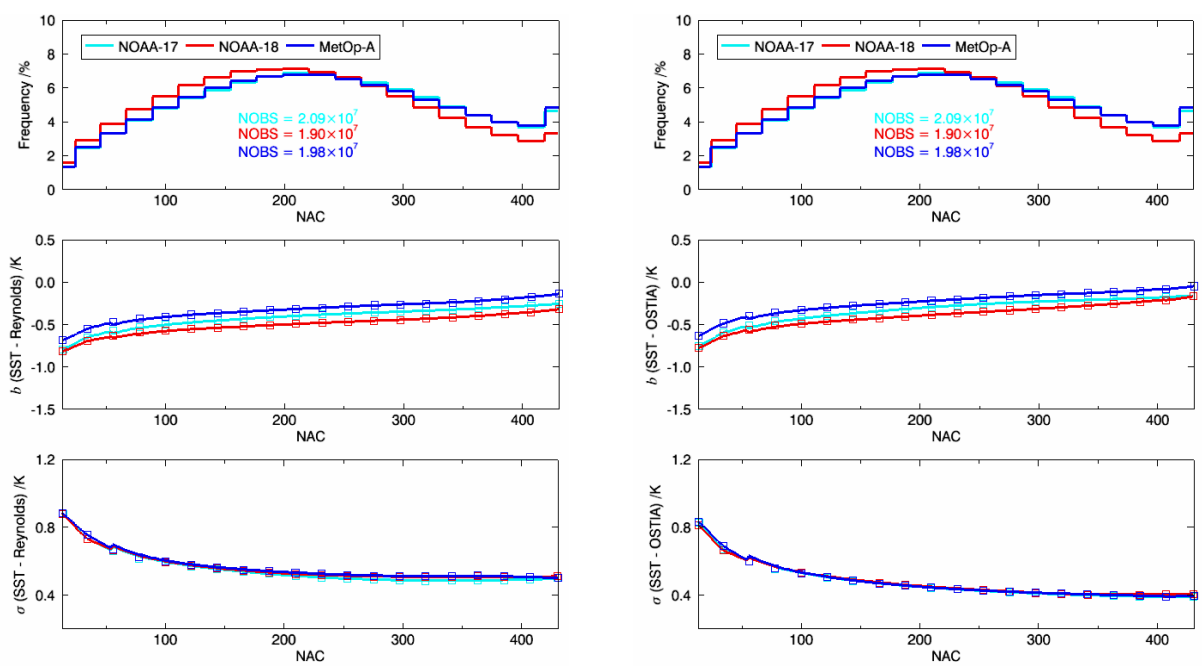
To investigate the bias dependencies, samples are binned equidistantly along each factor. Special analyses have shown that the θ and W dependencies are coupled and should be analyzed together, whereas the NAC dependence is well decoupled from both θ and W. The number of bins was selected to be 20 for a single independent factor (NAC) and 14 for two interrelated factors (θ and W). Subsequent analyses have shown that the bin size has negligible impact on the result, when the number of bins is varied from 10 to 50.

Fig. 3 shows b and σ as a function of NAC, along with the normalized histogram of the NAC. Statistics are calculated based on one week of global nighttime ACSPO data from 14 to 20 September 2008. Both mean bias and stdv show pronounced trends for both reference SSTs (Reynolds and OSTIA) and for all three platforms. As expected, both mean bias and stdv gradually approach an asymptotic regime as NAC increases, suggesting that the influences of ambient non-clear-sky ocean conditions progressively diminish as a pixel moves away from the boundaries of the clear-sky ocean domain. As an example, if an NAC>100 threshold is used to define the "confidently cloud-free" zone, this corresponds to 85% of total clear-sky area in the current ACSPO product. Note that the NAC dependencies are fairly consistent for three platforms (NOAA-17, NOAA-18, and MetOp-A) except that NOAA-18 has a little different histogram shape. This may be due to its different local overpass time around 2:00 PM, whereas NOAA-17 and MetOp-A take their observations around 9:30 AM local time.

An exponential fit function was selected to describe bias vs. NAC dependence in Fig. 3, such as

$$b \equiv f_N(N) = a_0 + a_1 \exp(-a_2 N) \tag{9}$$

where parameter a_0 is the asymptotic limit for a mean bias over clear-sky ocean (i.e., when NAC is approaching infinity); a_1 is the amplitude of the mean bias caused by the NAC; and a_2 is its decay rate.



(a) ΔSST (ACSPO-Reynolds) vs. NAC (b) ΔSST (ACSPO-OSTIA) vs. NAC

Figure 3. Histograms (with total number of observations) of NAC, and bias/stdv of SST anomalies as a function of NAC, from 14-20 September 2008.

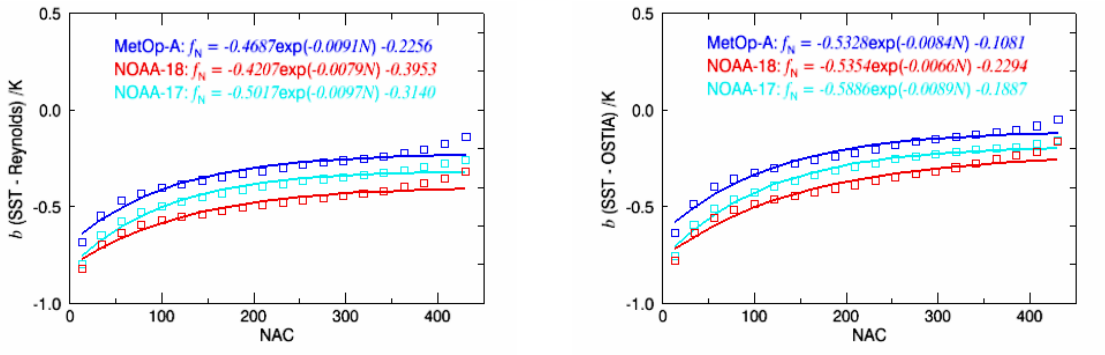
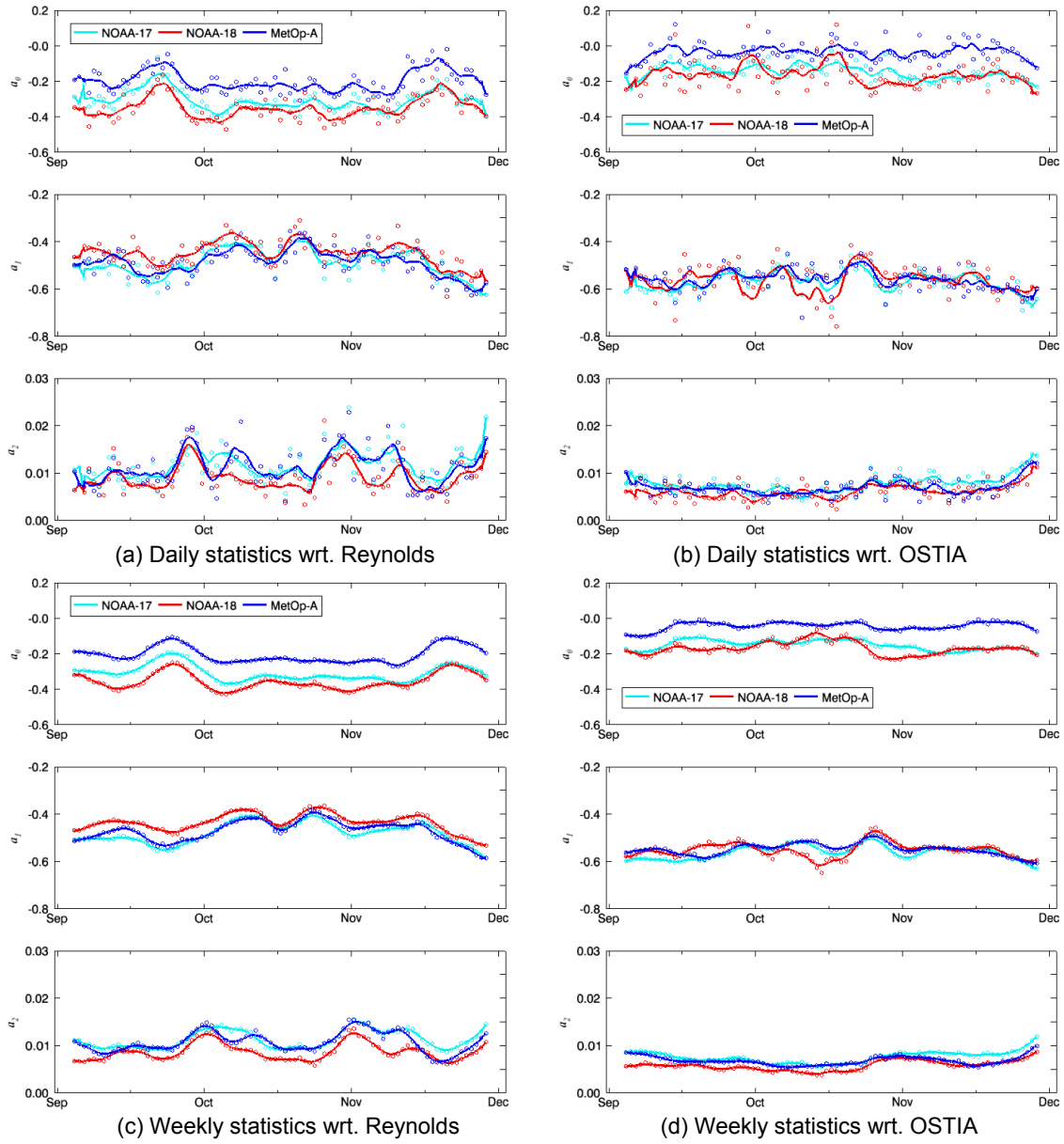


Figure 4. Fitting bias vs. NAC dependencies shown in Fig.3.



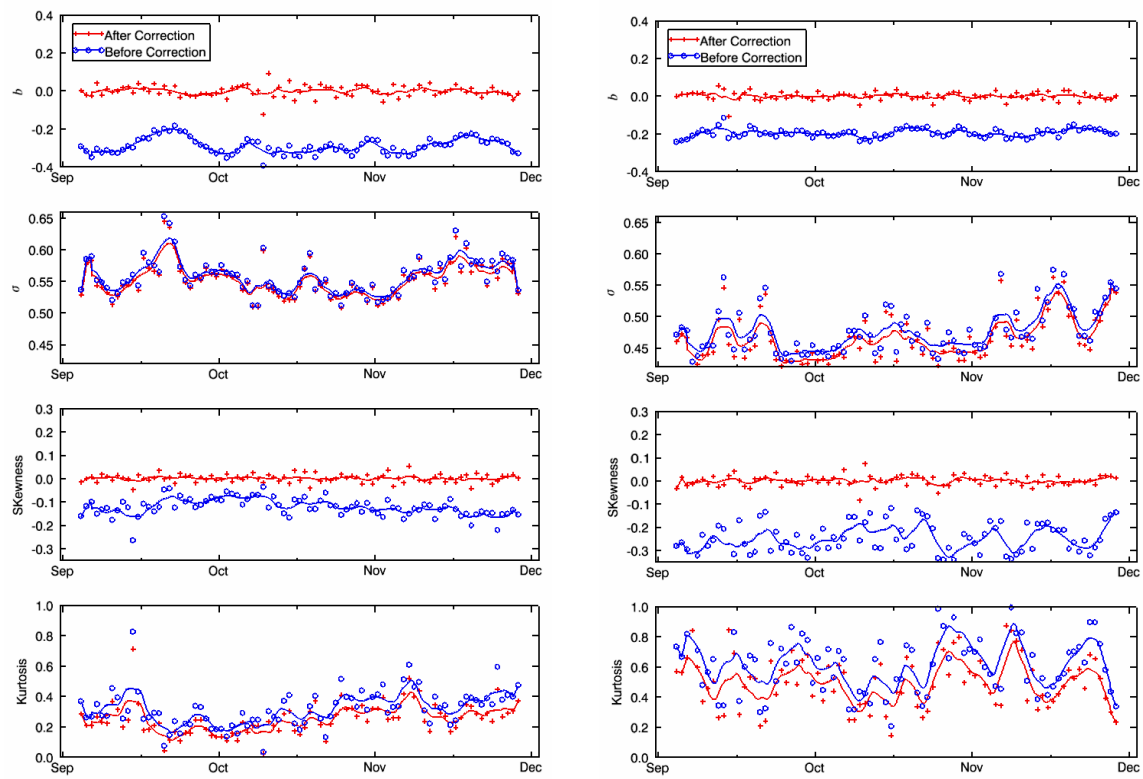
(a) Daily statistics wrt. Reynolds (b) Daily statistics wrt. OSTIA

(c) Weekly statistics wrt. Reynolds (d) Weekly statistics wrt. OSTIA

Figure 5. Time series of fitted parameters of $b = f_N(N)$ with respect to Reynolds and OSTIA reference SSTs, for daily and weekly statistics.

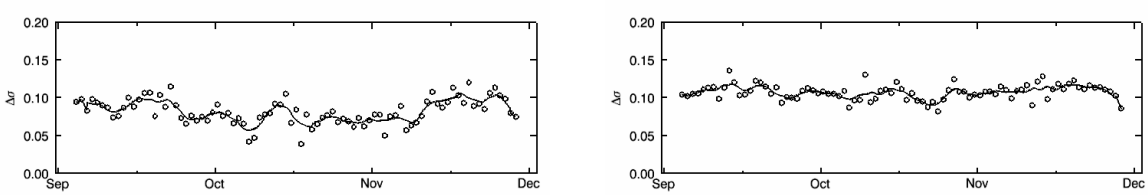
Fig. 4 gives the results of fitting the curves in Fig. 3 using Eq.(9). A good consistency is observed across platforms. Recall that the definition of NAC includes several different factors. For example, a low NAC could result from scattered cloud around the pixel or from the pixel lying too close to a large continuous cloud area. These two situations may affect the SST bias differently, but this difference is not considered in the current approach. Besides, influence from non-ocean neighbor pixels (land, ice) would also contribute to this NAC count. Overall, NAC appears to be a good integral indicator of a wide variety of these different effects. On the other hand, this wide range of factors may explain the abnormal fluctuating patterns in the bias dependencies, which may affect the long-term stability in the time series of the fitted parameters.

The stability of fitting is further investigated by trending the fit parameters in time in Fig. 5. Three-month data (from September to November 2008) have been processed. Each data point stands for the fitting of a one-day (daily statistics) or one-week moving window (weekly statistics) of ACSPO data. (Note that two successive points have only one-day shift, and the fitted parameters from the previous point are used as an initial guess for the non-linear fitting in the next point.)



(a) Wrt. Reynolds (b) Wrt. OSTIA

Figure 6. Time series of first four moments of statistics before and after correction of bias vs. NAC (corresponding to MetOp-A of Fig.6, coefficients are derived from previous day data and then applied to correction of current day data).



(a) Wrt. Reynolds (b) Wrt. OSTIA

Figure 7. Time series of root-mean-squared improvement in stdv after correction of bias vs. NAC (corresponding to Fig.7).

The parameter a_0 shows that, on average, the ACSPO v1.0 product is biased low with respect to Reynolds daily SST by 0.1-0.4K. The bias is much smaller with respect to OSTIA from 0.0-0.3K. This is a known problem in ACSPO v1.0, and work is underway to improve this bias in ACSPO v1.1, which is currently under testing. The amplitude of mean bias, a_1, ranges from 0.4 to 0.6 K for Reynolds and 0.5 –to 0.7 K for OSTIA and is fairly stable for both SST anomalies. The decay rate, a_2, appears smaller and more stable in the OSTIA case than Reynolds, which suggests sensitivity to the reference field. Weekly statistics show much more stable performance than the daily results. Analyses are underway to better understand and independently verify these results using *in situ* SST reference.

To evaluate the effectiveness of error characterization, bias correction was applied to independent data sample. Time series of the first four orders of statistics—mean, stdv, skewness, and kurtosis—of daily ACSPO data before and after bias correction are shown in Fig. 6. Since results of different platforms are very comparable, only MetOp-A is shown here. The data are the same as in Fig. 5, except that for each day, the bias correction was done using the fit coefficients derived from previous day. The global mean bias and skewness are both centered close to zero after bias correction. The significant improvement in skewness indicates that the major reason for the negative asymmetry of the original SST anomaly distribution was contamination by residual cloud. The stdv and kurtosis also improve in the bias-corrected sample. The corresponding reduction in stdv, $\Delta\sigma$, can be estimated from $\Delta\sigma^2 = \sigma^2_{\text{before}} - \sigma^2_{\text{after}}$ and is plotted in Fig. 7. Note that a 0.1 K root-mean-squared reduction in stdv is significant. Our preliminary analyses indicate that the entire stdv root-mean-squared reduction after three-factor bias correction could be up to ~0.2 K.

5. DISCUSSION AND FUTURE WORK

Further analyses are underway, including fitting bias and stdv dependencies as a function of other multidimensional factors. Fig. 8 gives an example of SST anomaly (NAC-detrended using Eq.(9)) vs. water vapor (W) dependencies for a number of different view angles. It is clearly seen that water vapor dependence is significantly affected by view angle. Since this coupling is nonlinear, multifactor dependence analysis and fitting must be used to characterize SST error behaviors. These analyses are currently underway and their results will be presented elsewhere.

Comparing to the look-up table approaches adopted by Minnett and Evans (2006) and Castro et al. (2008), the continuous error characterization analyzed in this study has at least two advantages. First, it is simple and easy to document and reuse, as it only requires knowledge of a few parameters in a fixed functional form. The fit parameters, however, change in time. Second, it is more stable (less sensitive to noise) because it only permits a limited number of degrees of freedom, sufficient to describe the intrinsic factors affecting the SST error. As such, it may offer a natural way of separating "signal" from "noise" in SSES. However, effective use of this approach requires successfully resolving the difficulties in selecting the appropriate form of the fitting functions and overcoming multidimensional curve fitting challenges.

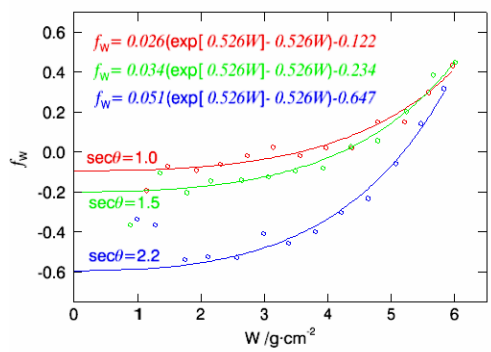


Figure 8. SST bias vs. water vapor dependencies under different view angles.

Analyses of ACSPO SST anomalies with respect to Reynolds and OSTIA reference SSTs show clear and stable dependencies of biases as a function of NAC. Using analytical functions to continuously describe error characteristics in the retrieval space thus appears feasible. For different platforms, error characteristics may vary greatly, due to differences between sensors, local overpass times, and SST regression equations. Preliminary analyses in this study show that the exponential representation of the NAC dependence can achieve reasonable stability and cross-platform consistency, at least in the three-month time period and for NOAA-17, -18, and MetOp-A data considered here.

Several factors may affect the accuracy of this study, such as other potential predictors of SST error not included in this analysis (e.g., aerosol) and errors in calculation (or measurement) of the dependent and independent factors used in the statistical analyses (including errors in the reference SST). In future studies, more factors will be included in the analyses. Stdv dependencies will be analyzed and fitted in the same manner as the bias considered in this study. Different reference SSTs, including *in situ* SSTs, will be used to minimize the effect of possible errors in the reference SST fields on the error characteristics. This method will be extended to more general regimes, such as daytime and more seasons. Newer versions of ACSPO will be consistently tested to measure the improvement. We also plan to apply this methodology to the analysis of the Model (CRTM) minus Observation (AVHRR) bias, which is critically important for physical SST retrievals.

One last issue to be addressed is the independence between reference SST and object SST. Most global SST analyses such as OSTIA use satellite SST as input and bias-correct it and blend it with other SST products using the SSES provided in the data. Independence between reference SST and satellite SST

must be ensured in error characterization to avoid a possible vicious circle where positive feedbacks could deteriorate the robustness of error characterization. A potential solution is using *in situ* SST in validation of error characterization results to restore independence and robustness of error characterization.

ACKNOWLEDGEMENTS

This work was supported by the GOES-R Algorithm Working Group (Manager Dr. Mitch Goldberg), funded by the GOES-R Program Office. F. Xu also acknowledges the CIRA visiting scientist fellowship. We thank our colleagues John Sapper (OSDPD), Prasanjit Dash (STAR/CIRA), Boris Petrenko (STAR/IMSG), Nikolay Shabanov (STAR/IMSG), Yury Kihai (STAR/PSGS) and Denise Frey 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

Castro S. L., G. A. Wick, D. L. Jackson, and W. J. Emery. 2008. Error characterization of infrared and microwave satellite sea surface temperature products for merging and analysis, *J. Geophys*. Res., 113, C03010, doi:10.1029/2006JC003829.

Dash, P., and A. Ignatov. 2008. Validation of clear-sky radiances over oceans simulated with MODTRAN4.2 and global NCEP GDAS fields against nighttime NOAA15-18 and MetOp-A AVHRRs data. *Remote Sens. Environ*., 112(6), 3012–30.

Donlon et al. 2007. The Global Ocean Data Assimilation Experiment High-Resolution Sea Surface Temperature Pilot Project, *Bull. Am. Meteorol. Soc*., 88(8), 1197–1213.

DuMouchel, W. and F. O'Brien. 1989. "Integrating a Robust Option into a Multiple Regression Computing Environment," in *Computing Science and Statistics: Proceedings of the 21st Symposium on the Interface*, (K. Berk and L. Malone, eds.), American Statistical Association, Alexandria, VA, pp. 297–301.

Emery W. J., Y. Yu, G. A. Wick, P. Schluessel, and R. W. Reynolds. 1994. Correcting infrared satellite estimates of sea surface temperature for atmospheric water vapor attenuation, *J. Geophys. Res*., 99(C3), 5219–5236.

Holland, P. W. and R. E. Welsch. 1977. Robust regression using iteratively reweighted least-squares, *Communications in Statistics - Theory and Methods*, Volume 6, Issue 9, pages 813–827.

Kearns, E. J., J. A. Hannifin, R. H. Evans, P. J. Minnett, and O. B. Brown. 2000. An independent assessment of Pathfinder AVHRR sea surface temperature accuracy using the Marine Atmosphere Emitted Radiance Interferometer (MAERI), *Bull. Am. Meteorol. Soc*., 79, 397–407.

Kilpatrick, K. A., G. P. Podesta and R. Evans, 2001, Overview of the NOAA/NASA advanced very high resolution radiometer Pathfinder algorithm for sea surface temperature and associated matchup database, J. Geophys. Res., 106, C5, 9179-9197

Liang, X., A. Ignatov, and Y. Kihai. 2009. Implementation of the Community Radiative Transfer Model (CRTM) in AVHRR Clear-Sky Processor for Oceans (ACSPO) and validation against nighttime radiances, *J. Geophys. Res*., submitted.

May, D. A., M. M. Parmeter, D. S. Olszewski, and B. D. McKenzie. 1998. Operational processing of *satellite sea surface temperature retrievals at the Naval Oceanographic Office. Bull. Amer. Meteor. Soc*., 79, 397–407.

McClain, E. P., W. G. Pichel, and C. C. Walton. 1985. Comparative performance of AVHRR-based multichannel sea surface temperatures. *J. Geophys. Res*., 90, 11 587–11 601.

Merchant, C. J., P. Le Borgne, A. Marsouin, and H. Roquet. 2008. Optimal estimation of sea surface temperature from split-window observations, *Remote Sensing of Environment* Volume 112, Issue 5, and *Earth Observations for Terrestrial Biodiversity and Ecosystems* Special Issue, 15 May 2008, Pages 2469–2484.

Merchant, C. J. and P. Le Borgne. 2004. Retrieval of sea surface temperature from space based on modeling of infrared radiative transfer: Capabilities and limitations, *Journal of Atmospheric and Oceanic Technology* 22 (11), pp. 1734–1746.

Merchant C. J., A. R. 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.

Minnett P. J. and R. H. Evans. 2006. Validation of sea surface temperatures from MODIS, presented at the MODIS Science Team Meeting, Adelphi, MD, Oct. 31 to Nov.2 2006. (Available at http://modis.gsfc.nasa.gov/sci_team/meetings/200610/presentations/cal/minnett.pdf)

Nocedal, J. and S. J. Wright. 1999. *Numerical Optimization*, Springer, New York.

Petrenko, B., A. Heidinger, A. Ignatov, and Y. Kihai. 2008. Clear-Sky Mask for the AVHRR Clear-Sky Processor for Oceans. AGU Ocean Sciences Meeting, Orlando, FL, 2-7 March 2008.

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

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

Stark, J. D., C. J. Donlon, M. J. Martin, and M. E. McCulloch. 2007. OSTIA: An operational, high resolution, real time, global sea surface temperature analysis system, Oceans '07 IEEE Aberdeen, conference proceedings. Marine challenges: coastline to deep sea. Aberdeen, Scotland.

Vazquez-Cuervo, J., E. M. Armstrong, and A. Harris. 2004. The effect of aerosols and clouds on the retrieval of infrared sea surface temperature, *Journal of Climate* 17 (20), pp. 3921–3933.

Walton, C. C. 1988. Nonlinear multichannel algorithms for estimating sea surface temperature with AVHRR satellite data, *J. Appl. Meteorol., 27*, 115.

Walton, C. C., W. G. Pichel, J. F. Sapper, and D. A. May. 1998. The development and operational application of nonlinear algorithms for the measurement of sea surface temperatures with the NOAA polar-orbiting environmental satellites, *Journal of Geophysical Research* 103, pp. 27,999–28,012.