6.12 MULTI-PLATFORM REAL-TIME SEA SURFACE TEMPERATURE ANALYSIS FOR THE INITIALIZATION OF SHORT-TERM OPERATIONAL FORECASTS

Steven M. Lazarus*, Corey G. Calvert, Michael Splitt
Florida Institute of Technology, Melbourne, Florida

Pablo Santos
NOAA/National Weather Service, Miami, Florida

David W. Sharp, Peter Blottman, and Scott Spratt
NOAA/National Weather Service, Melbourne, Florida

1. INTRODUCTION

Forecasting within the coastal zone (i.e., within 200 km of a coastline) continues to be a NOAA priority and includes wide ranging marine-related weather issues. As the need for better integrated system models and observations grows, so does the challenge of assimilating the myriad of data streams. Despite advances in short-term forecasting, issues remain. For example, relatively significant gradients in the near-shore sea surface temperatures (SSTs) can impact regional weather including boundary layer evolution, mesoscale features such as the sea breeze, stratocumulus development (e.g., Young and Sikora, 2003), coastal showers, etc. In addition to these issues, recent work examining model output winds from European Centre for Medium Range Weather Forecasts (ECMWF) has shown that higher resolution SSTs can significantly impact the surface wind stress field (Chelton and Wentz 2005; O'Neill et al. 2005, Chelton 2005).

In the absence of clouds (and sun glint), IR-based satellite sensors provide reliable radiances from which bulk (i.e., upper meter) SST estimates are derived. Satellite derived SSTs (from multiple platforms) are a significant source of high resolution data in operational analyses (e.g., He et al., 2003; Thiebaux et. al. 2003) Because of coverage and/or data latency issues, operational systems require a first-guess field which may be a composite (e.g., weekly mean) or a previous analysis. While multi-platform SST analyses are fairly common, typical real-time data assimilation challenges such as bias, error and length scale specification, cloud masking, and data latency problems remain. While a combination of both IR and microwave estimates of SSTs can mitigate the impact of clouds on an SST analysis, microwave-derived SSTs are of relatively coarse resolution (on the order of 50 km). In contrast, GOES SSTs can provide temporal resolution unavailable from polar orbiters, at spatial resolutions of 4 km, and have been shown to have comparable errors to that of the higher resolution AVHRR SSTs (Walker et al. 2003).

* *Corresponding author address:* Dr. Steven M. Lazarus, Florida Institute of Technology, Dept. of Marine and Environmental Systems, Melbourne, FL 32901-6975; email: slazarus@fit.edu.

As part of a COMET funded project, the Florida Institute of Technology is building a near real-time operational SST analysis system. This system is designed to provide high resolution SST analyses in lieu of the relatively coarse National Centers of Environmental Prediction/Marine Modeling and Analysis Branch (NCEP/MMAB) Real-Time Global Sea Surface Temperature (RTG-SST) analysis which is currently assimilated into the Advanced Regional Prediction System (ARPS) by the National Weather Service (NWS) in Melbourne, Florida. The NWS in Melbourne, Florida is cycling the ARPS four times per day over an approximate 700 km^2 domain that includes the eastern Gulf of Mexico and the northwest Bahamas (e.g., Fig. 1). The FIT analysis system combines high resolution satellite SST data obtained from both the GOES-12 and the Moderate Resolution Imaging Spectroradiometer (MODIS) on board the Terra and Aqua satellites.

Preliminary SST analyses are running operationally at the Florida Institute of Technology (FIT) with current efforts here focused on: 1.) latency-related diurnal SST adjustments, and 2.) determining spatially varying error covariances, and decorellation length scales. Evaluation of the impact of high-resolution SSTs on short-term model forecasts (i.e., on the order of a day), and an intercomparison project involving FIT generated GOES composites, SPoRT MODIS composites, and the RTG-SST analysis are discussed in related papers (LaCasse et al., 2006; Haines et al., 2006).

2. DATA

i. GOES-12

The GOES (bulk) SST data is provided by NOAA National Environmental Satellite, Data, and Information Service (NESDIS). The product is derived from the satellite radiances using two of the five available channels (3.9 and 11μm). 30 minute data are combined to produce hourly SST files. Removal of both cloud-contaminated radiances (via a cloud mask) and radiances that are affected by sun glint at 3.9 μm, precede application of a regression-based SST retrieval algorithm (Maturi et al. 2004). Area Man computer Interactive Data Access System (McIDAS) files are sub sampled (Maturi, personal communication) to produce the 6 km horizontal resolution lat/lon grids. The data are available hourly in near-real time (i.e., 4-hour lag).

ii. MODIS

The MODIS (bulk) SSTs are available twice daily (within the ARPS domain during the following intervals: 3-4 UTC, 7-8 UTC, 15-16 UTC, and 18-19 UTC) from both the AQUA and TERRA platforms, are of high spatial resolution (1 km), and can be accessed in near real-time via direct broadcast from the University of Wisconsin. The SST algorithm is based on IR retrieval methodology and uses both mid and far bands which are corrected for atmospheric absorption and cloud screened (Brown et al. 1999). The live broadcast Wisconsin SST retrieval algorithm differs from the "official" algorithm developed by the MODIS Science Data Support/Ocean Science Teams and archived by the Goddard Space Flight Center DAAC. The modified Wisconsin algorithm is part of the International MODIS AIRS Processing Package (IMAPP). In limited evaluation, the IMAPP algorithm produced SST estimates that were within 0.5^{o}C (of the Science Team SSTs) for the daytime product with smaller differences for the nighttime SST product. The SSEC IMAPP algorithm is fast and can deliver SST products via direct broadcast in near real-time (see: ftp://ftp.ssec.wisc.edu/pub/IMAPP/MODIS/Level-2/v1.5/SST_DOC.pdf).

iii. Buoy

Although somewhat sparse, buoy day is obtained from 9 active sites (within the ARPS domain, Fig. 1) via the National Data Buoy Center online archive at http://www.ndbc.noaa.gov/. These data are have been used in separate validation efforts discussed by Haines et al. (2006) and LaCasse et al. (2006). SST data from these buoys will also be used for near real-time evaluation of the Florida Tech SST analysis system that is currently under development.

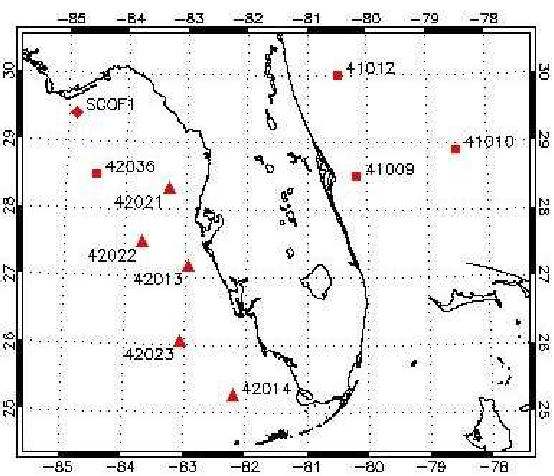

Figure 1: National Data Buoy Center (NDBC) buoy locations within the NWS/ARPS domain. NOAA buoys (red squares), University of South Florida Coastal Ocean Monitoring and Prediction System (red triangles), and C-MAN station (red diamond).

3. ANALYSIS METHOD

Although we intend to transition to a two-dimensional variational approach, the first generation FIT SST analysis employs the Bratseth method (1986) which is essentially a hybrid analysis – a successive correction technique that converges to optimum interpolation. Because of the latter, the method does not require the inversion of large matrices which can be both computationally expensive and/or impractical given current computer resources (Sashegyi et al., 1993). Although the Bratseth method was originally constructed as an iterative two-step process whereby the grid point and observation point analyses alternate (e.g., see Lazarus et al. 2002), we apply an equivalent but alternative approach that iterates a correction vector only and relegates the analysis to a single "one-shot" process (e.g., Kalnay 2003). Because we focus here on the data latency issue and the determination of analysis parameters, rather than the analyses themselves the reader is referred to Zavodsky et al. (2006) more complete description of the analysis scheme.

4. RESULTS

i. Latency

We are currently examining the quality of the GOES composites for a period (May 2004) that corresponds to concurrent/cooperative projects with SPoRT (Haines et al. 2006, LaCasse et al. 2006). Because it is our intention to use, in some fashion, the GOES composites as a first-guess field for the analyses, we compare the statistics of clear versus cloudy scenes. Relevant questions include:

Does the composite smooth gradients or create spurious gradients?

What is the impact of diurnal SST variations on the composites?

What is the contribution of the latency to the total bias?

In an attempt to address these questions we compare temperature gradient statistics for clear vs. cloudy scenes. Figure 2 depicts several different estimates of the mean temperature gradient (^{o}C/km) as a function of grid separation (Delta X) for May 2004. The "benchmark" (i.e., best case scenario) is given by the red curve which represents cases with zero latency (no compositing) and a clear sky threshold (i.e., cases where the domain is significantly clear). If we remove the clear sky criteria, the mean temperature gradient increases for all length scales shown (black curve). The "upper bound" on the mean temperature gradient is given by the composites in which all data are allowed – regardless of the

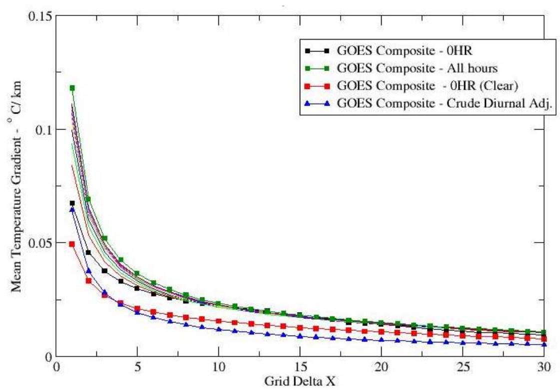

Figure 2: Mean GOES sea surface temperature gradient (°C/km) versus grid spacing (delta x) for May 2004 (see text for more).

latency (green curve). Other curves shown, but not labeled, depict various degrees of latency in which we systematically allow 'older' data to enter the composites. Note that the blue curve, a somewhat "crude" diurnal SST adjustment that employs a simple 'time of day correction' to the latent SSTs, compares relatively favorably to the GOES 'CLEAR' – suggesting that the latency issue is indeed relevant.

Is it possible to 'adjust' for the diurnal SST effect using a more physically rigorous approach? Here, we apply a parameterization developed by Kawai and Kawamura (KK, 2002) that was designed to use daily mean wind speed and maximum solar radiation as inputs and returns the difference between the maximum and minimum SST for the day. Rather than use daily inputs of wind and solar radiation however, we attempt to downscale the KK parameterization (i.e., apply it to hourly data). The approach replaces the latent SST in the composite with a combination of the most recent value of the diurnally adjusted (i.e., diurnal signal removed) SST (T_{base}) and an SST increment calculated using the current (Eta) solar/wind via the KK parameterization. Although the KK algorithm was developed for both skin and bulk SST adjustments – we choose the latter as the 1m depth would be the most appropriate for 'adjusting' the GOES SSTs as the GOES IR radiances are regressed against buoy temperatures. Additionally, because we are attempting to apply the KK parameterization in a manner that it was not necessarily intended, we test the method using inputs (wind/solar) from Eta analyses (at 00, 06, 12, and 18 UTC) for May 2004. On time scales greater than a day, T_{base} is intended to be a quasi-conserved SST value (i.e., slowly changing in time). We attempt to show that T_{base} is also approximately invariant at time scales < 1 day, and thus can be used to "update" our composites in regions of latent data.

Both monthly and daily average estimates of SST and T_{base} for May 2004 over the entire NWS/ADAS domain are shown in Figure 3. Albeit reduced, It is clear from this figure that we have not removed the entire diurnal signal in the T_{base} estimate (red dots). At this

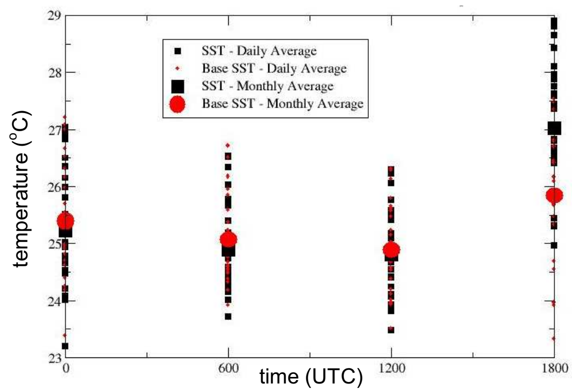

Figure 3: Daily mean (small black squares) and monthly mean (large black squares) GOES SSTs (°C), and KK estimates of the daily mean (small red circles) and monthly mean T_{base} for the NWS ARPS Florida domain (see text for details).

point is not clear why we are not able to remove more of the diurnal signal, however there are likely several factors including sun glint contamination, parameterization scalability, and the quality of the Eta solar and wind inputs. In terms of the former, we have uncovered sun glint-contaminated (GOES-12) SSTs in the shallow shelf waters within the ADAS domain (see Haines et al., 2006) that may, in part, be responsible for producing a spuriously large diurnal SST signal over portions of the region. We are currently reprocessing the GOES composites using an additional sun glint mask to remove the warm SSTs.

ii. Length Scale Determination

There is no reason to expect the analysis length scales and error covariances to be homogeneous in space or time. We have begun to determine the analysis length scales using an approach similar to that of Bormann et al. (2003). The Bormann algorithm calculates the correlations of distance-binned difference pairs (between GOES cloud track and upper air winds) using the correlation formula

$$COR(X,Y) = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}}, \qquad (5)$$

where *(X_i, Y_i)* and $(\bar{X}, \bar{Y})$ denote the innovation pairs and their bin mean respectively, and *i* is the number of pairs in each distance bin. Innovations are used so as to avoid recorrelation with distance (note however that this does not guarantee that the data will not recorrelate). Because this calculation involves an enormous number of innovation pairs, we focus on pre-selected 'representative' regions within the NWS ARPS domain (see crosses on Fig. 4). In order to ensure a statistically robust estimate of the distance-dependent correlations, we use an entire month of innovations. Additionally, because we seek spatially dependent length scale estimates, all correlation estimates are local -- using a subregion within a radius of 80 km for a chosen grid point. The 80 km radius was selected: 1.) to allow for a sufficient number of innovations within the subregion (over a month), 2.) to avoid spreading unrepresentative innovations across the Florida peninsula, and 3.) to maintain a somewhat reasonable number of calculations.

Figure 5 shows an example plot (for February 2005) of the local spatial correlation values plotted versus distance (i.e., mid-bin value for 20 km bins). These values correspond to the region directly off the east-central Florida coast (red circle in Fig. 4). The spatial lag-correlation function plotted against separation distance is substantially less than 1.0 because of the presence of observation error variance (in addition to background field error) in the denominator of Eq. (5) which is not present in the numerator as the lag approaches zero separation (e.g., Thiébaux et al. 1986). Because the best fit of the correlations may not be Gaussian (which is the assumed error correlation structure for our analysis system), we show two different regressions: Gaussian and autoregressive. For this region, the correlations between the two models are comparable (this is not the case in all regions of our domain however). For this case, the best fit returns a length scale of approximately 77 km.

In addition to accounting for spatial variability in the error covariance length scale, we are also in the process of determining the temporal variability (monthly). The degree of stationarity (or lack thereof) will determine the eventual operational analysis configuration with respect to how often we update the analysis parameters (e.g., monthly versus seasonal). A Barnes analysis will then spread the length scale information throughout the grid in order to ensure a smoothly varying length scale.

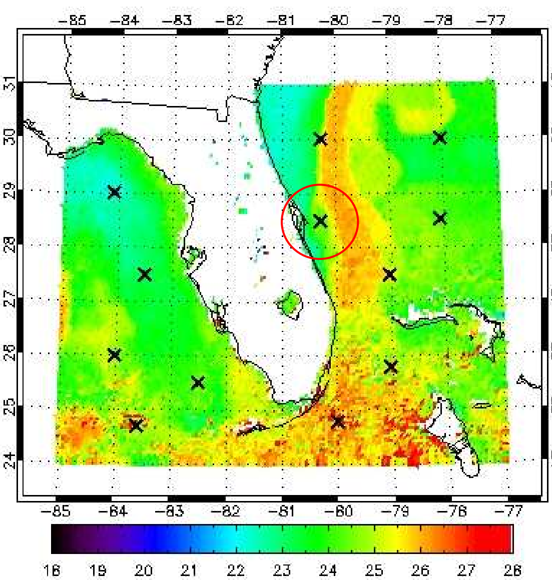

Figure 4: GOES SST (^{o}C) composite valid 07 UTC 8 May, 2004. The 'X's indicate regions selected for estimating the error covariance length scale.

Table 1: Mean SSTs (^{o}C) by buoy location for May 2004.

	#41009	#41010	#41012	#42036
buoy/day	24.72	25.29	24.15	24.19
buoy/night	24.40	25.10	23.76	23.84
GOES/night	24.71	25.05	24.00	24.17
RTG	25.19	24.83	24.21	24.32
climatology	25.50	24.57	24.56	25.07

iii. Bias and Error Variance Determination

We are currently calculating bias and error variances (GOES minus MODIS) for the ADAS domain. Because the truth is not known, we are using the innovations as a surrogate to produce both spatially and temporally dependent error and bias estimates. Since we intend to use buoy data as an independent evaluation metric, we can use buoy/GOES differences and MODIS/GOES bias estimates to tune the analysis system. Table 1 lists the mean SSTs at four of the nine buoy locations (see Fig. 1) within the ADAS domain during May 2004. Daytime (nighttime) estimates are averages over a 3 h window between 16-19 UTC (4-7 UTC) that correspond to the MODIS overpass times for the domain. Also shown are the corresponding GOES (night), RTG, and 15 yr climatological (obtained from the Jet Propulsion Laboratory AVHRR Pathfinder 9 km pentad data archived at the Goddard DAAC) SSTs at the buoy locations. All but the GOES data, which are nearest neighbor, are interpolated to the buoy location. There are no GOES SSTs corresponding to the daytime window of the MODIS overpass (17-19 UTC) due to sun glint. Except for buoy #41010, the GOES, RTG and climatological SSTs are each warmer than the buoy for May 2004 (the RTG analysis assimilates a 24 h average buoy SST). We are in the process of determining monthly GOES bias estimates at all 9 buoy locations within the domain.

In Figure 6 we show a spatial map of the SST error variance estimate (using both day/night MODIS data) for May 2004. These preliminary variance estimates are based on MODIS-GOES differences where the latent GOES SSTs have been adjusted using the KK parameterization. It is important to point out however, that these

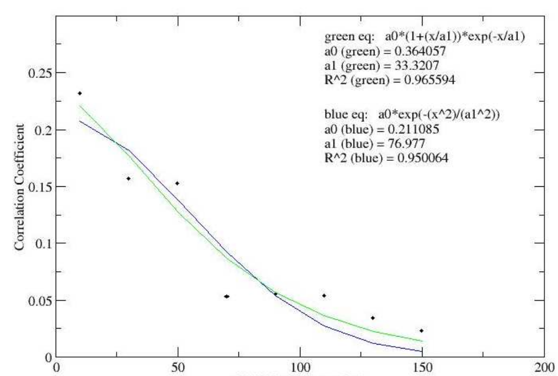

Figure 5: Local spatial lag correlation versus innovation distance separation for February 2005.

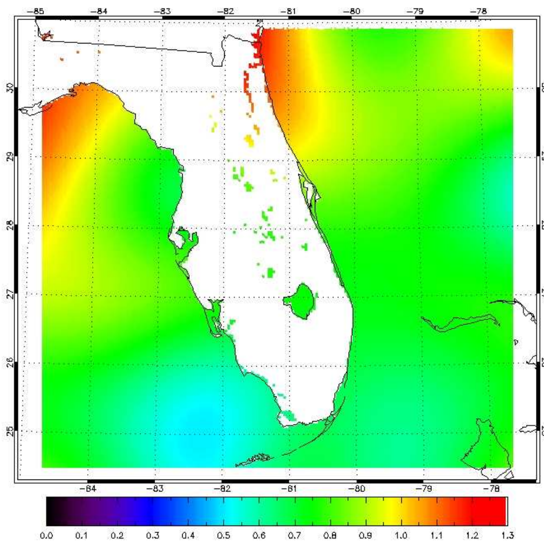

Figure 6: SST error variance estimate ($^{o}C^{2}$) from MODIS/GOES innovations for May 2004.

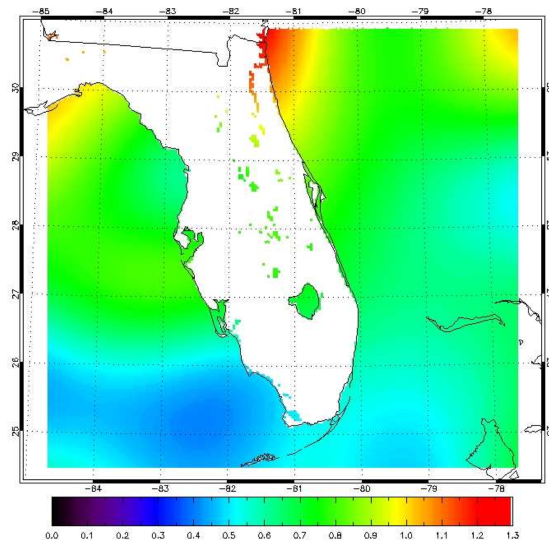

Figure 7: As in Fig. 6 but for non-composite GOES data.

variance estimates are impacted by a 4 h lag between data sets. Questions remain with how to properly address issues related to data set latencies in an operational setting (i.e., the GOES SST data is approximately 4 h older than the MODIS). Nevertheless, the variance shown is maximum (on the order of 1.3 $^{o}C^{2}$) in the northwest portion of the ADAS domain and adjacent to the northeast Florida coast with lower (and relatively uniform) variability elsewhere. Figure 7 depicts a clear-sky (i.e., non-composite) variance estimate for the same period. As in Fig. 6 there is an approximate 4 h difference between the GOES and MODIS data. The error variance estimates for the non-composite data are lower – especially in the south and southwest part of the domain (compare Fig. 6 and Fig. 7). Ultimately, we intend to use zero (time) lag and distinguish between day/night in our estimates of the error variance and thus anticipate that the error estimates will be less than those shown in Figs. 6 and 7. Also, as previously discussed, an evaluation of the month-to-month variability will determine how often we update the operational error variances within our analysis domain.

5. SUMMARY

An SST analysis system has been constructed at FIT as part of a COMET funded project to produce near real-time SST analyses for the ARPS/ADAS assimilation and forecast cycle at the Melbourne National Weather Service Forecast Office. The analysis system is fully configured except for some outstanding data quality issues – in particular with respect to the data latency problem. Systematic estimates of the analysis parameters (which include both the error correlation length scales and error covariance magnitudes) will be stratified by month and region in an effort to tune the analysis system and to determine their temporal evolution within the NWS ADAS/ARPS domain. The latter will be used to create a time-varying error specification. A relatively clear testbed period was chosen (May 2004) to evaluate the SST analyses which will be compared against both buoy data and MODIS SST composites (Haines et al., 2006). In a companion study (LaCasse et al., 2006), work is currently underway to evaluate the impact of high-resolution SSTs (derived from the MODIS composites) on short-term model forecasts.

6. REFERENCES

Borman, N., S. Saarinen, G. Kelly, and J.-N. Thepaut, 2003: The spatial structure of observation errors in atmospheric motion vectors from geostationary satellite data. *Mon. Wea. Rev.*, **131**, 706-718.

Braseth Bratseth, A.M., 1986: Statistical interpolation by means of successive corrections. *Tellus*, **38A**, 439-447.

Brown, O. B., and P. J. Minnett, 1999: *MODIS Infrared Sea Surface Temperature Algorithm*. Tech Report ATBD25, University of Miami, Miami, FL.

Chelton, D. B. and F. J. Wentz, 2005: Global Microwave Satellite Observations of Sea Surface Temperature for Numerical Weather Prediction and Climate Research. *Bull. Amer. Met. Soc.*, **86**, 1097-1115.

Chelton, D. B., 2005: The Impact of SST Specification on ECMWF Surface Wind Stress Fields in the Eastern Tropical Pacific. *J. Climate*, **18**, 530-549.

O'Neil, L. W., D. B. Chelton, S. K. Esbensen, and F. J. Wentz, 2005: High-Resolution Satellite Measurements of the Atmospheric Boundary Layer Response to SST Variations along the Agulhas Return Current. *J. Climate*, **18**, 2706-2722.

Haines S. L., G. J. Jedlovec, , S. M. Lazarus, and C. G. Calvert, 2006: An Aqua MODIS sea surface temperature composite product. Preprints, *14th Conference on Satellite Meteorology and Oceanography*, Atlanta GA, Amer. Met. Soc., January 30-February 2, 2006.

Sashegyi, K. D., D. E. Harms, R. V. Madala, and S. Raman, 1993: Application of the Bratseth Scheme for the Analysis of GALE Data Using a Mesoscale Model. *Mon. Wea. Rev.*, 121, 2331-2350,

He, Rouying, R. H. Weisberg, H. Zhang, F. E. Muller_Karger, R. W. Helber, 2003: A Cloud Free, Satellite-Derived, Sea Surface Temperature Analysis for the West Florida Shelf. Geophys. Res. Lett., **15**, 1811, doi:10.1029/2003GL017673.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 341 pp.

Kawai and Kawamura, 2002: Evaluation of the Diurnal Warming of Sea Surface Temperature Using Satellite-Derived Marine Meteorological Data. J. Oceanography, **58**, 805-814.

LaCasse, K., M., W. M. Lapenta, S. M. Lazarus, and M. E. Splitt, 2006: The Impact of MODIS SST Composites On Short-Term Regional Forecasts. Preprints, *10th Symposium on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface*, Atlanta GA, Amer. Met. Soc., January 30-February 2, 2006.

Lazarus, S. M., Steven M. Lazarus, Carol M. Ciliberti, John D. Horel and Keith A. Brewster. 2002: Near-Real-Time Applications of a Mesoscale Analysis System to Complex Terrain. *Wea. Forecasting*, **17**, 971–1000.

Maturi, E., A. Harris, N. Nalli, C. Merchant, S. McCallum, R. Meiggs, and R. Potash, 2004: NOAA's Operational Geostationary Sea Surface Temperature Products. Preprints, *20th International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology*, Seattle WA, Amer. Met. Soc., January 12-15 2004.

Thiébaux, J., E., H. L. Mitchell, and D. W. Shantz, 1986: Horizontal Structure of Hemispheric Forecast Error Correlations for Geopotential and Temperature. *Mon. Wea. Rev.*, 114, 1048-1066.

Thiébaux, J., E. Rogers, W. Wang, and B. Katz, 2003: A New High-Resolution Blended Real-Time Global Sea Surface Temperature Analysis. *Bull. Amer. Met. Soc.*, **84**, 645–656.

Walker, N., S. Mayint, A. Babin, and A. Haag, 2003: Advances in satellite radiometry for the surveillance of surface temperatures, ocean eddies and upwelling processes in the Gulf of Mexico using GOES-8 measurements during summer. *Geophys. Res. Lett.*, 30, 1854, doi:10.1029/2003GL017555.

Young, G. S. and T. D. Sikora, 2003: Mesoscale Stratocumulus Bands Caused by Gulf Stream Meanders. *Mon. Wea. Rev.*, **131**, 2177-2191.

Zavodsky, B. T., S. M. Lazarus, R. Ramachadran, and Xiang Li, 2006: Evaluation of an Innovation Variation Methodology for Real-Time Data Reduction of Satellite Data Streams. *18th Conference on Probability and Statistics*, Atlanta GA, Amer. Met. Soc., January 30-February 2, 2006.