Comparisons of Daily Sea Surface Temperature Analyses for 2007–08

RICHARD W. REYNOLDS

NOAA/National Climatic Data Center, Asheville, North Carolina

DUDLEY B. CHELTON

*College of Oceanic and Atmospheric Sciences, and Cooperative Institute for Oceanographic Satellite Studies,
Oregon State University, Corvallis, Oregon*

(Manuscript received 16 June 2009, in final form 18 February 2010)

ABSTRACT

Six different SST analyses are compared with each other and with buoy data for the period 2007–08. All analyses used different combinations of satellite data [for example, infrared Advanced Very High Resolution Radiometer (AVHRR) and microwave Advanced Microwave Scanning Radiometer (AMSR) instruments] with different algorithms, spatial resolution, etc. The analyses considered are the National Climatic Data Center (NCDC) AVHRR-only and AMSR+AVHRR, the Navy Coupled Ocean Data Assimilation (NCODA), the Remote Sensing Systems (RSS), the Real-Time Global High-Resolution (RTG-HR), and the Operational SST and Sea Ice Analysis (OSTIA); the spatial grid sizes were $1/4°$, $1/4°$, $1/9°$, $1/11°$, $1/12°$, and $1/20°$, respectively. In addition, all analyses except RSS used in situ data. Most analysis procedures and weighting functions differed. Thus, differences among analyses could be large in high-gradient and data-sparse regions. An example off the coast of South Carolina showed winter SST differences that exceeded 5°C.

To help quantify SST analysis differences, wavenumber spectra were computed at several locations. These results suggested that the RSS is much noisier and that the RTG-HR analysis is much smoother than the other analyses. Further comparisons made using collocated buoys showed that RSS was especially noisy in the tropics and that RTG-HR had winter biases near the Aleutians region during January and February 2007. The correlation results show that NCODA and, to a somewhat lesser extent, OSTIA are strongly tuned locally to buoy data. The results also show that grid spacing does not always correlate with analysis resolution.

The AVHRR-only analysis is useful for climate studies because it is the only daily SST analysis that extends back to September 1981. Furthermore, comparisons of the AVHRR-only analysis and the AMSR+AVHRR analysis show that AMSR data can degrade the combined AMSR and AVHRR resolution in cloud-free regions while AMSR otherwise improves the resolution. These results indicate that changes in satellite instruments over time can impact SST analysis resolution.

1. Introduction

Sea surface temperature (SST) analyses are useful for many purposes including hurricane forecasting, fisheries operations, air–sea flux studies as the ocean surface boundary condition for atmospheric models, and for studies of climate change and prediction. In recent years the number of satellite instruments has increased. This has helped facilitate an increase in the number of additional data and analysis products that are operational or under development. Many of these products are produced by the Group for High-Resolution Sea Surface Temperature (GHRSST) (see Donlon et al. 2007; http://www.ghrsst-pp.org/, see in particular "Data Access"). GHRSST products are used for many purposes (see publication list online at http://www.ghrsst.org/Peer-reviewed-articles.html). The analyses use in situ and remotely sensed data from a variety of geostationary and polar-orbiting satellites and are computed over different regions and time periods with different spatial and temporal resolutions. Users now have a choice of analyses that was never possible before GHRSST was established. Most of these products tend to cover roughly the last five years when satellite instruments such as the microwave (MW) Advanced Microwave Scanning Radiometer (AMSR) and

Corresponding author address: Richard W. Reynolds, NOAA/ National Climatic Data Center, 151 Patton Avenue, Asheville, NC, 28801-5001.
E-mail: richard.w.reynolds@noaa.gov

TABLE 1. Summary of the six analyses sorted by increasing equatorial grid spacing. All analyses except the RTG-HR product are available online at ftp://podaac.jpl.nasa.gov/pub/GHRSST/data/L4/GLOB/. In situ observations are SSTs from ships and buoys except for NCODA, which also includes also includes in situ hydrographic temperature and salinity profiles.

Analysis name	Input data	Equatorial resolution	Temporal resolution	Start date
1) AVHRR-only	AVHRR, in situ	28 km ($^1/_4°$)	Daily	September 1981
2) AMSR+AVHRR	AMSR, AVHRR, in situ	28 km ($^1/_4°$)	Daily	June 2002
3) NCODA	AVHRR, in situ, altimeters, atmospheric forcing	12 km ($^1/_9°$)	6 h	October 2005
4) RSS	AMSR, TMI, MODIS	10 km ($^1/_{11}°$)	Daily	August 2005
5) RTG-HR	AVHRR, in situ	9 km ($^1/_{12}°$)	Daily	September 2005
6) OSTIA	AVHRR, AMSR, TMI, AATSR, SEVIRI, in situ	6 km ($^1/_{20}°$)	Daily	March 2006

the infrared (IR) Moderate Resolution Imaging Spectroradiometer (MODIS), joined two longer time series of IR instruments as sources of global SST observations: the Advanced Very High Resolution Radiometer (AVHRR), since November 1981, and the Along Track Scanning Radiometer (ATSR), since August 1991.

The problem for users is not to obtain an SST analysis but to choose from the many that are available the one that is best suited to their particular purpose. Every analysis is designed to produce a regularly gridded product from irregularly spaced data. Most SST analyses use statistical techniques to produce gridded products without dynamical considerations. Analyses can be produced at any temporal interval for any spatial grid. As the resolution increases, the signal may increase depending on the resolution of the data. However, as the resolution increases, the noise also increases. The bottom line is that no analysis works when there are no observations nearby in space and/or time. Of course, as more high-resolution satellite data become available, the analysis resolution can be increased. With presently available data, differences among daily analyses are larger than differences among monthly analyses, and the daily differences can exceed 5°C, as will be discussed below.

The details of the design of each analysis vary. One of the first choices is which SST data should be used in the analysis procedure. Then choices have to be made on the spatial grid spacing and the update frequency of the analysis. Next, bias corrections need to be considered along with analysis parameters such as error correlation scales and the signal-to-noise ratios. These and other choices that must be considered in the design of an analysis procedure may lead to very different results. For example, SST observations that exceed some predetermined threshold are often discarded. Thus, observations near this threshold are either kept or discarded based on small differences.

In this paper, six different analyses are compared for the 2007–08. The results produced here may have to be reevaluated in the future as the accuracy and resolution of analyses evolve as satellite instruments change and as improvements in analysis procedures are implemented. The most useful result of this comparison is to identify problems in the various analyses, which will hopefully lead to improvements.

2. Overviews of the six SST analysis products

Analyses were selected that were global with at least daily resolution and available for the two year period, 2007–08. The minimum requirement of 2 yr eliminates several newer analyses but allows comparisons over a substantial time period. Five analyses meeting the requirements were selected from the GHRSST Global Data Assembly Center (GDAC) at the Jet Propulsion Laboratory (JPL) Physical Oceanography Distributed Active Archive Center (PO.DAAC) (see table online at http://ghrsst.jpl.nasa.gov/GHRSST_product_table.html). One additional analysis from the National Oceanic and Atmospheric Administration's (NOAA) National Centers for Environmental Prediction (NCEP) was added because it is presently used operationally in two forecast models: the regional North American Mesoscale (NAM) model at NCEP (Thiébaux et al. 2003) and the global forecast model at the European Centre for Medium-Range Weather Forecasts (ECMWF) (Chelton and Wentz 2005). All analyses can be obtained from ftp://podaac.jpl.nasa.gov/pub/GHRSST/data/L4/GLOB/ unless specifically noted below. The analyses are discussed in order of increasing grid spacing and summarized in Table 1. GHRSST (Donlon et al. 2007) recommends that analyses be produced at the foundation temperature, defined as the temperature at the shallowest depth below any diurnal variations.

There are several sources of in situ and AVHRR data. For the period considered here, the in situ SST data are from ships and buoys and collected over the real-time Global Telecommunication System (GTS); the AVHRR

data are from the U.S. Navy (May et al. 1998). This AVHRR algorithm separately corrects daytime and nighttime satellite data with collocated buoy data.

a. AVHRR-only and AMSR+AVHRR analyses

Two of the analyses are produced daily on a $\frac{1}{4}°$ grid at NOAA's National Climatic Data Center (NCDC) as described by Reynolds et al. (2007). One analysis (AVHRR-only) uses in situ and AVHRR data. The second analysis (AMSR+AVHRR) adds AMSR data. Both analysis procedures are the same. In situ data from ships and buoys are used to provide a large-scale bias correction of the satellite data. All data are used for a given day, where the day is defined by coordinated universal time (UTC). However, the satellite data are separated into daytime and nighttime bins and corrected separately using all in situ data. Then the in situ and corrected satellite data are analyzed using an optimum interpolation (OI) procedure. These two analyses were computed to determine the impact of an analysis variance jump when AMSR became available in June 2002. The error correlation scales range from 50 to 200 km with smaller scales at higher latitudes (especially in western boundary current regions) and larger scales in the tropics.

Version 2 of the OI procedure, as described by Reynolds et al. (2007), is used here. The changes from version 1 are relatively small and primarily consist of additional temporal smoothing. The temporal smoothing includes using 3 consecutive days of data, with the middle day weighted higher than the other two days. The date of the analysis is defined to be the middle day of each 3-day analysis period. The temporal smoothing also includes additional smoothing of the bias corrections, which tend to be noisy because of limited in situ observations. In addition, ship SSTs are corrected relative to the buoy SSTs by subtracting 0.14°C from all ship observations before they are used to bias correct the satellite data. Thus, all observations are bias corrected with respect to buoy SSTs and there are no corrections to foundation temperature. The changes in version 2 are described in more detailed in http://www.ncdc.noaa.gov/oa/climate/research/sst/papers/whats-new-v2.pdf. Additional up-to-date information is available at http://www.ncdc.noaa.gov/oa/climate/research/sst/oi-daily.php.

b. NCODA analysis

The U.S. Navy Coupled Ocean Data Assimilation (NCODA) analysis (Cummings 2005) is computed operationally using in situ data and AVHRR data. In addition to the ship and buoy SST in situ data used by other analyses, the NCODA analysis also includes in situ hydrographic temperature and salinity profiles.

The analysis is performed on a $\frac{1}{9}°$ grid on the equator with gradual reductions in latitudinal intervals to keep the size of the grid boxes nearly equal area between 80°S and 80°N. The NCODA analysis is done every 6 h using data within ±3 h of the date of the analysis. However, NCODA has a floating "look back time" for every instrument to ensure that all data get into the analysis (J. A. Cummings 2009, personal communication). Because of this potential delay, incoming data may not be centered on the analysis time window. No bias correction of satellite data is performed. However, diurnal warming events are flagged and some flagged observations may be eliminated in light winds. The error correlation scales are determined by the Rossby radius of deformation obtained from Chelton et al. (1998) and range from ~10 km near the pole to ~220 km at the equator. As described in Cummings (2005), NCODA is implemented as the data assimilation component of the Hybrid Coordinate Ocean Model (for details see http://www.hycom.org/dataserver/glb-analysis/expt-90pt8). In this assimilation system altimeter data and atmospheric forcing fields are also used. NCODA is the only analysis considered here that is linked to a dynamical model. All the other analyses are statistical analyses only. NCODA therefore has a strong potential advantage if the ocean dynamics from the model are realistic. Analyses from July 2005 to present are available online at http://usgodae2.fnmoc.navy.mil/ftp/outgoing/fnmoc/models/ghrsst/; the JPL Web site has analyses beginning in April 2008.

c. RSS analysis

The Remote Sensing Systems (RSS) analysis is computed on a ~$\frac{1}{11}°$ grid using AMSR, Tropical Rainfall Measuring Mission Microwave Imager (TMI) and MODIS data. The analysis uses 3 days of consecutive data with the analysis date referenced by the middle day. The analysis has a constant error correlation scale of 1.5° (~165 km on the equator) and is the only analysis that does not use in situ data directly. There is no correction for a foundation temperature. However, AMSR and TMI retrievals are calibrated and validated by buoys. Furthermore, large-scale MODIS biases are removed by adjustment to AMSR. All observations are adjusted to remove any diurnal signal based on the local time of day and wind speed, following Gentemann et al. (2003). This analysis is not yet published (C. L. Gentemann 2009, personal communication), although some details are available online at http://www.ssmi.com/sst/microwave_oi_sst_browse.html.

d. NCEP RTG-HR analysis

The NCEP Real-Time Global High-Resolution (RTG-HR) is operationally computed daily using in situ and

AVHRR data on a $1/12°$ grid. The analysis is computed at 0000 UTC using observations within $±12$ h. The analysis has no diurnal correction or depth adjustment to foundation temperature. The error correlation scales range from 50 to 450 km based on the climatological SST gradients with smaller (larger) scales related to larger (smaller) gradients (see Thiébaux et al. 2003). The satellite data are corrected relative to the in situ data following the Poisson scheme of Reynolds and Smith (1994). This version is described by the unpublished manuscript Gemmill et al. (2007), which is available online at http://polar.ncep.noaa.gov/sst/ along with other analysis details. Only the most recent year of analyses is available for download at ftp://polar.ncep.noaa.gov/pub/history/sst/ophi/. The RTG-HR analysis is not part of the GHRSST program and so is not available from the GDAC. Thiébaux et al. (2003) describe the earlier version of this analysis on a coarser grid (see also Chelton and Wentz 2005).

e. Met Office OSTIA analysis

The Met Office produces an Operational SST and Sea Ice Analysis (OSTIA) on a $1/20°$ grid that uses in situ, AVHRR, AMSR, TMI, Advanced ATSR (AATSR), and geostationary Spinning Enhanced Visible and Infrared Imager (SEVIRI) data. The analysis is run daily at 0600 UTC using data from a 36-h period ending at 0600 using two error correlation scales, 10 and 100 km, which vary depending on the region and the input data. To produce a foundation temperature, the input data are then filtered to remove daytime observations with wind speed <6 m s^{-1}, which may contain diurnal surface warming. All satellite SST data are adjusted for bias errors by reference to a combination of AATSR and in situ SST measurements from drifting buoys. Further details can be found in Donlon et al. (2009, manuscript submitted to *Remote Sens. Environ.*) and online at http://ghrsst-pp.metoffice.com/pages/latest_analysis/ostia.html.

f. Some general comments

It is important to point out that all the analyses except the NCODA have some type of satellite bias correction using either in situ data or one type of satellite data (or both in the case of OSTIA) as a reference. If this is not done, small biases may lead to jumps as data are contributed by different satellite instruments. Reynolds et al. (2010) show that this bias occurs (see in particular Figs. 4 and 5) even for two different nighttime AVHRR instruments using the same algorithms.

Only the RSS analysis specifically corrects the data to remove the diurnal cycle. This removal may lead to errors because accurate information on the surface heat fluxes is needed for accurate estimates of the diurnal cycle. However, such information is not available for use

in the RSS analysis procedure. In OSTIA and NCODA, there are adjustments to remove some diurnal signal by deleting daytime observations in low winds. The other analyses bias correct satellite data to match spatially and temporally averaged in situ data in an effort to reduce the impact of diurnal warming.

3. Qualitative differences

Qualitative comparisons of SST maps from the various products show areas with large differences (exceeding several degrees Celsius) during the 2-yr time period 2007–08 considered here. The regions with the largest differences tend to occur near the coast, in strong gradient regions such as western boundary currents, and at high latitudes where SST measurements (both in situ and satellite) tend to be sparse. In addition, SSTs are simulated in some of the analyses based on sea ice concentrations (see Reynolds et al. 2007 for details), which may contribute to the differences among analyses at high latitudes.

Figure 1 shows regional SST fields off the Carolina coast for 1 January 2007. This region was selected because the warm Gulf Stream is found off shore in winter while colder shelf water is present between the Gulf Stream and the coast. The shelf water cannot be detected by MW retrievals because it is too near the coast where MW observations are contaminated by land in the antenna sidelobes. Furthermore, winter clouds can limit IR retrievals. The detection of the shelf water is especially difficult off the South Carolina coast because there are no moored buoys there. Thus, it will sometimes be difficult to detect the shelf water in statistical-only analyses. In Fig. 1, the colder shelf water is evident in the AVHRR-only, the AMSR+AVHRR, the NCODA, and the RTG-HR analyses, but it is partly missing in the RSS and the OSTIA analyses. In addition, the small-scale variability is highest in the RSS analysis and lowest in RTG-HR and OSTIA.

Figure 2 shows the SST gradient magnitudes for the same region and time as in Fig. 1. The SST gradients associated with the Gulf Stream are clearest off the Carolinas in NCODA and weakest in RSS and OSTIA. The gradients in the AVHRR-only and AMSR+AVHRR analyses are not as clear adjacent to the coast because of the relatively coarse $1/4°$ grid compared with the grid spacing of the other analyses; the gradients were computed with centered differences only if all east–west and north–south nearest grid values were ocean values. North of about 35°N, the gradients in the AVHRR-only and the AMSR+AVHRR sharpen and become very sharp in the RSS.

It is noteworthy that the gradients north of 35°N are actually a little sharper in the AVHRR-only analysis
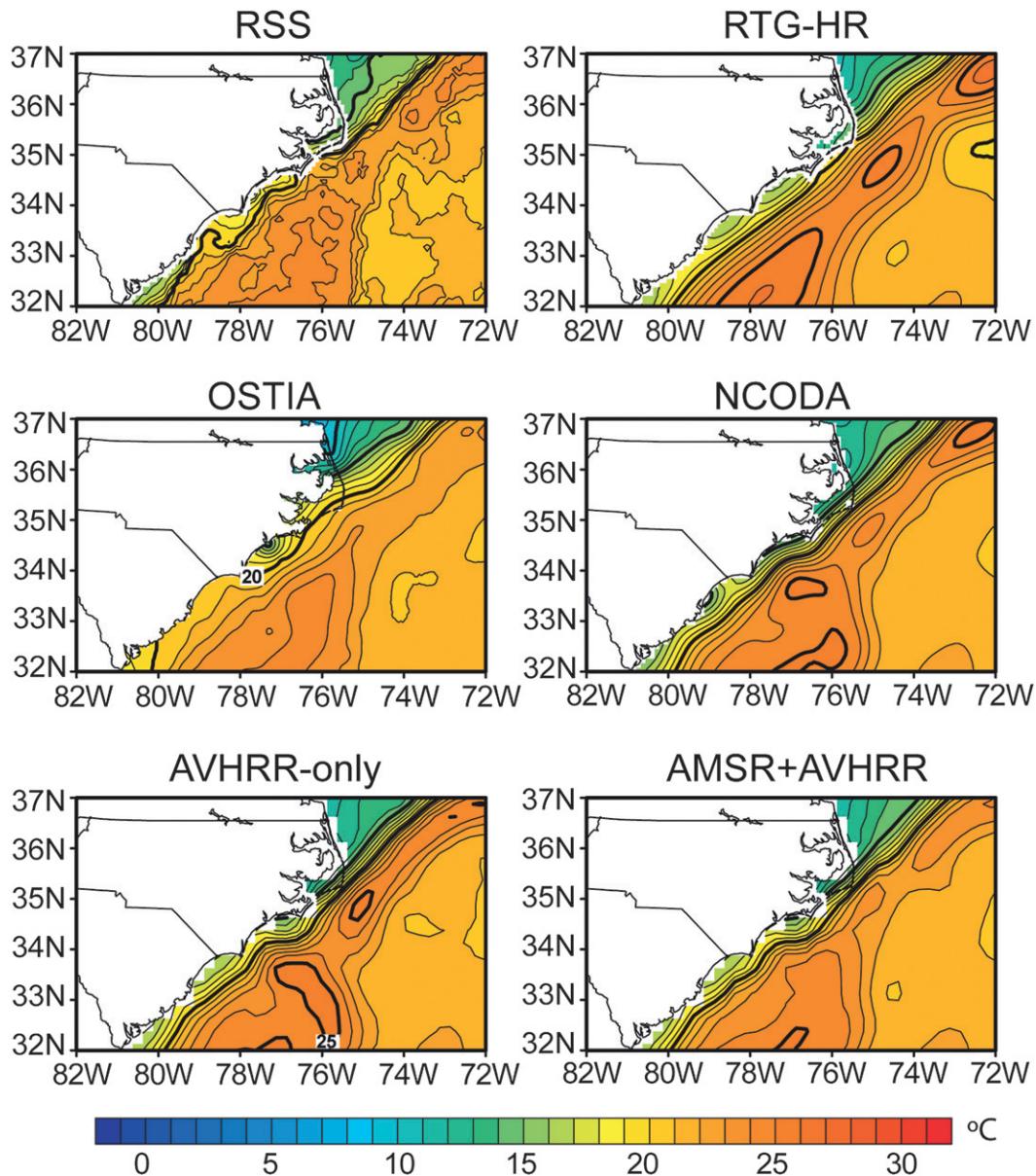
FIG. 1. Daily SST fields for 1 Jan 2007 from the six analyses considered in this study. Note in particular the differences near the South Carolina coast (roughly 33°N, 80°W). The contour interval is 1°C.

than in the AMSR+AVHRR analysis. The AVHRR and AMSR data were gridded on the same ¼° grid in the AVHRR-only and AMSR+AVHRR analyses. However, the AMSR data resolution is only about 50 km while the AVHRR data resolution is 4–9 km. Thus, AMSR data slightly reduced the gradients in this case. However, in cloud-covered regions farther off the coast, AVHRR observations become sparse and the AMSR can improve the analysis resolution relative to AVHRR-only (see Reynolds et al. 2007, Fig. 9).

Figure 3 shows SST fields for a region in the western tropical Pacific for 1 January 2007. SST variability and gradients are small in this region. In addition, SST temperatures between 10°S and 10°N are near 30°C, as shown in the figure, with abundant rainfall year-round due to strong convective activity. Thus, IR satellite retrievals are limited by cloud cover, and to a lesser extent MW satellite retrievals are limited by precipitation. The results show that the features are smoothest in the RTG-HR analysis. The RSS analysis has considerable small-scale structure that apparently arises from inclusion of 1-km data from MODIS. However, because MODIS data are limited by swath width and clouds, they are not available every day for the region shown in the figure. Thus, some
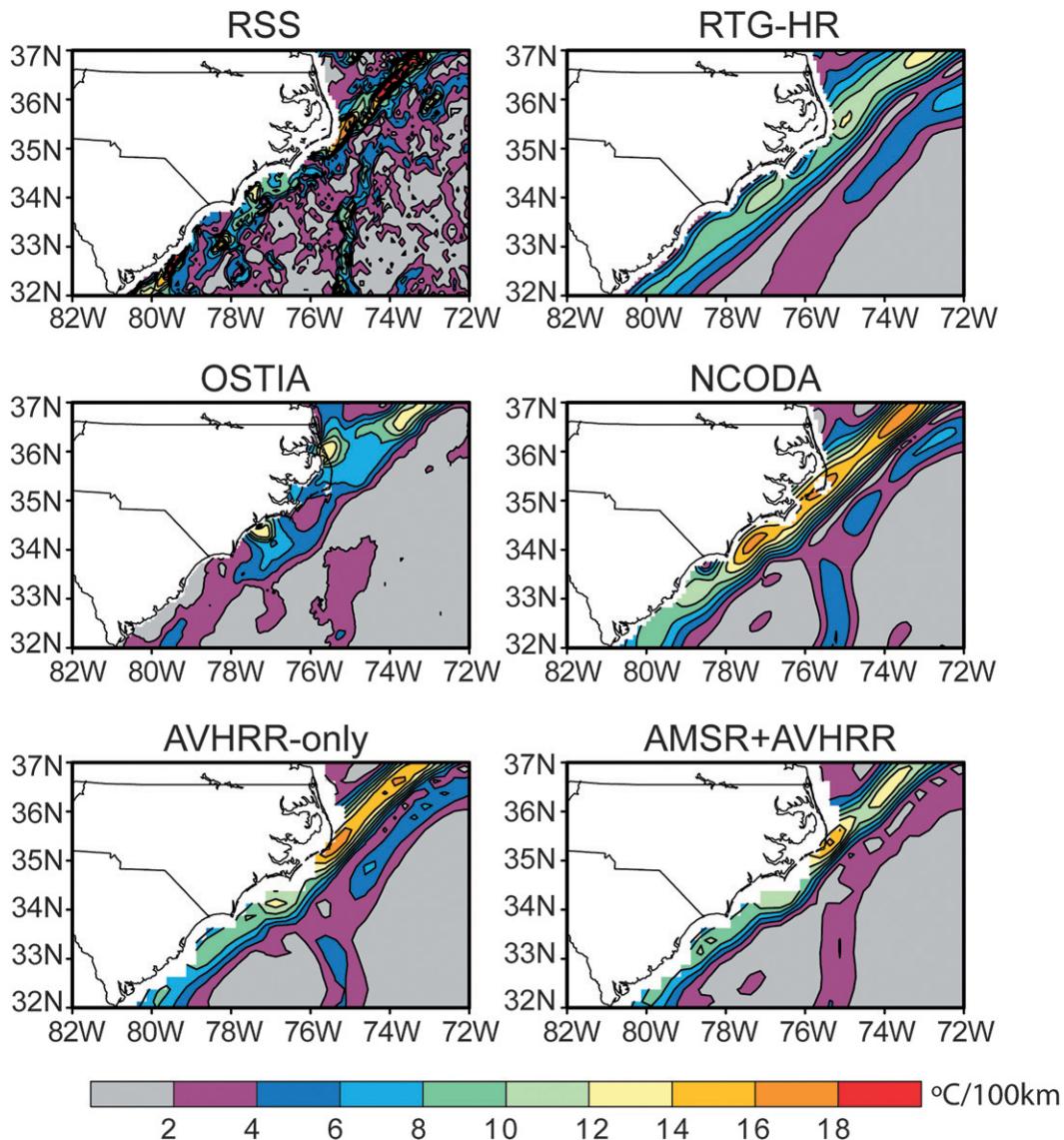
FIG. 2. Daily SST gradient magnitudes for 1 Jan 2007 from the six SST analyses. Gradients are computed as centered differences only if all four of the nearest east–west and north–south neighbors are present. The contour interval is $2°C (100 km)^{-1}$.

of the RSS small-scale details may arise from conditions several days earlier and persistence in the RSS analysis procedure.

Figure 4 shows the SST gradient magnitude off the West Coast of the United States for 1 September 2008, a time of year when coastal upwelling is generally strong. Upwelling is evident along the coasts of Oregon, Washington, and northern California from the narrow band of strong SST gradients adjacent to the coast. The strong and meandering SST gradients farther offshore are frontal regions associated with eddies and meanders of the California Current (Castelao et al. 2006). The SST gradients are strongest in the NCODA analysis. The

coastal gradients are somewhat weaker in OSTIA and even weaker in the AMSR+AVHRR and AVHRR-only analyses. The coastal gradients are the weakest and SST gradients are almost nonexistent in the open-ocean regions offshore. RSS shows very strong and finescale SST gradients nearshore. However, the SST gradients are again very noisy everywhere in the RSS analysis.

4. Wavenumber spectra

The feature resolutions of the six SST analysis products considered in this study are evident from the zonal wavenumber spectra in Fig. 5. Spectra are shown for
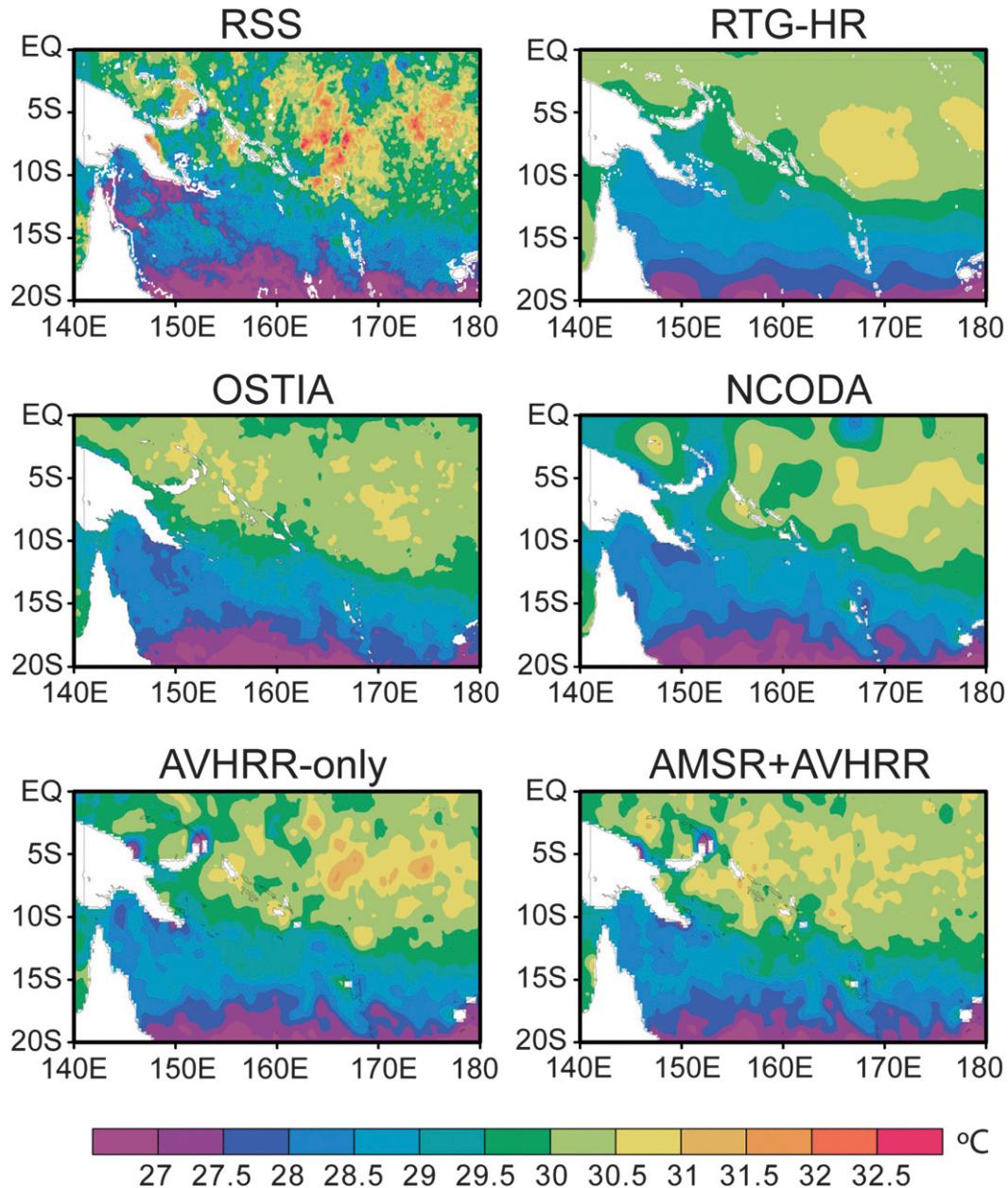
FIG. 3. Daily SST fields for 1 Jan 2007 from the six SST analyses. Small-scale features are most evident in RSS and RTG-HR is the smoothest. The shading interval is 0.5°C.

a Southern Hemisphere midlatitude region [the Agulhas Return Current (ARC) in the south Indian Ocean], a Northern Hemisphere midlatitude region [the Kuroshio Extension (KE) in the North Pacific Ocean], and the western tropical Pacific (WTP). Each spectrum is the ensemble average of 31 individual zonal wavenumber spectra computed from daily SST fields over the months of January 2007 and July 2007.

A consistent feature in all six panels of Fig. 5 is that the RSS SST fields have much higher spectral energy than any of the other SST products at wavelengths shorter than about 300 km, with the RSS spectra rolloff with zonal wavenumber k as approximately k^{-2} in all cases. In comparison, the wavenumber dependence of the OSTIA spectra ranges from k^{-4} to k^{-5}. The NCODA, AVHRR-only, and AMSR+AVHRR spectra are somewhat steeper. The spectral energy in the RSS fields is more than two orders of magnitude higher than any of the other SST analyses at the highest wavenumbers. The existence of a k^{-2} spectrum is not in general an
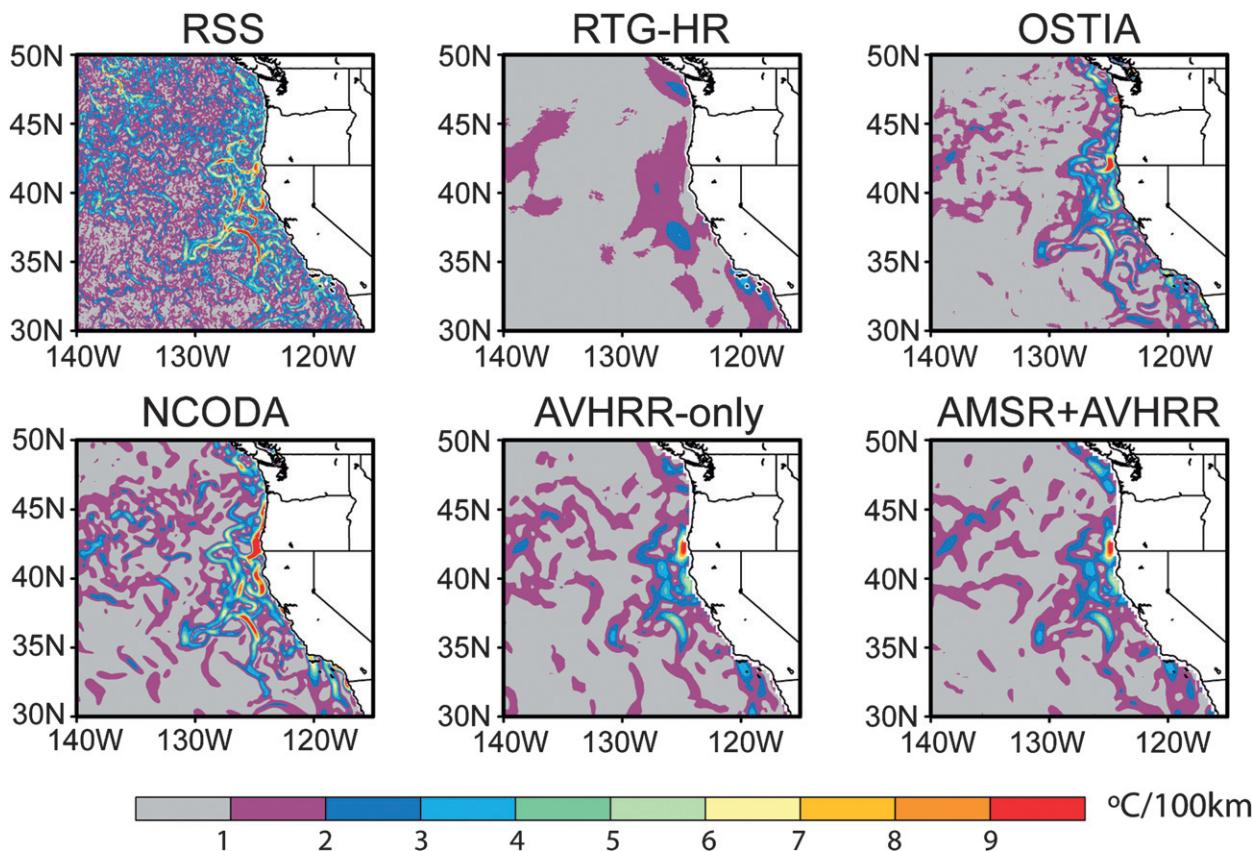
FIG. 4. SST gradient magnitudes for 1 Sep 2008 from the six SST analyses. Note the coastal gradients due to upwelling. Gradients are computed as in Fig. 2. The shading interval is $1°C (100\ km)^{-1}$.

indication of noise in the data. However, the fundamentally different spectral behavior of the RSS SST fields compared to the other SST analyses and the noisy, "speckled," appearance of the RSS SST fields in Figs. 2–4 is a clear indication that the RSS fields are dominated by small-scale noise. It should also be noted that the k^{-2} falloff is also evident at the shortest wavelengths in the ARC region in all analyses except the OSTIA. The k^{-2} falloff at these short scales may be interpreted as an indication of "red" noise in all the analyses except OSTIA rather than an abrupt change in the spectral characteristics of SST at short wavelengths.

Another consistent feature in all six panels of Fig. 5 is that the RTG-HR SST fields have significantly lower feature resolution than any of the other SST products, as evident from the steep rolloff of the spectra at wavelengths shorter than about 250 km in the ARC and KE regions. This rolloff indicates that the analysis procedure evidently attenuates the variability with scales shorter than about 250 km. In the two midlatitude regions, the RTG-HR spectra are very similar to the other spectra at wavelengths longer than about 250 km. The RTG-HR SST fields are characterized by a red noise

with roughly k^{-3} spectral rolloff that extends to quite long wavelengths of about 150 km in the ARC and KE and more than 500 km in the WTP. This is presumably due to spatial smoothing of noisy SST fields as part of the RTG-HR analysis procedure.

In the WTP region, the spectral rolloff of the NCODA SST fields is approximately k^{-4} at wavelengths shorter than about 750 km, indicating that NCODA attenuates signals with these scales. As a result, the NCODA variance at wavelengths shorter than about 250 km in January and about 150 km in July is smaller than the $\sim k^{-3}$ noise variance in the RTG-HR fields. As a result, RTG-HR has the illusion of having higher resolution than NCODA. It would be risky to interpret these shorter-scale features as real in the RTG SST fields, given the approximate k^{-3} red spectral characteristics of the noise that evidently exists in the RTG fields.

The wavenumber spectral analysis thus identifies important problems in the RSS and RTG-HR SST products in all three of the regions considered here and the NCODA product in the WTP region. The differences between the spectra of the other SST products are more subtle, but some generalizations can nonetheless be
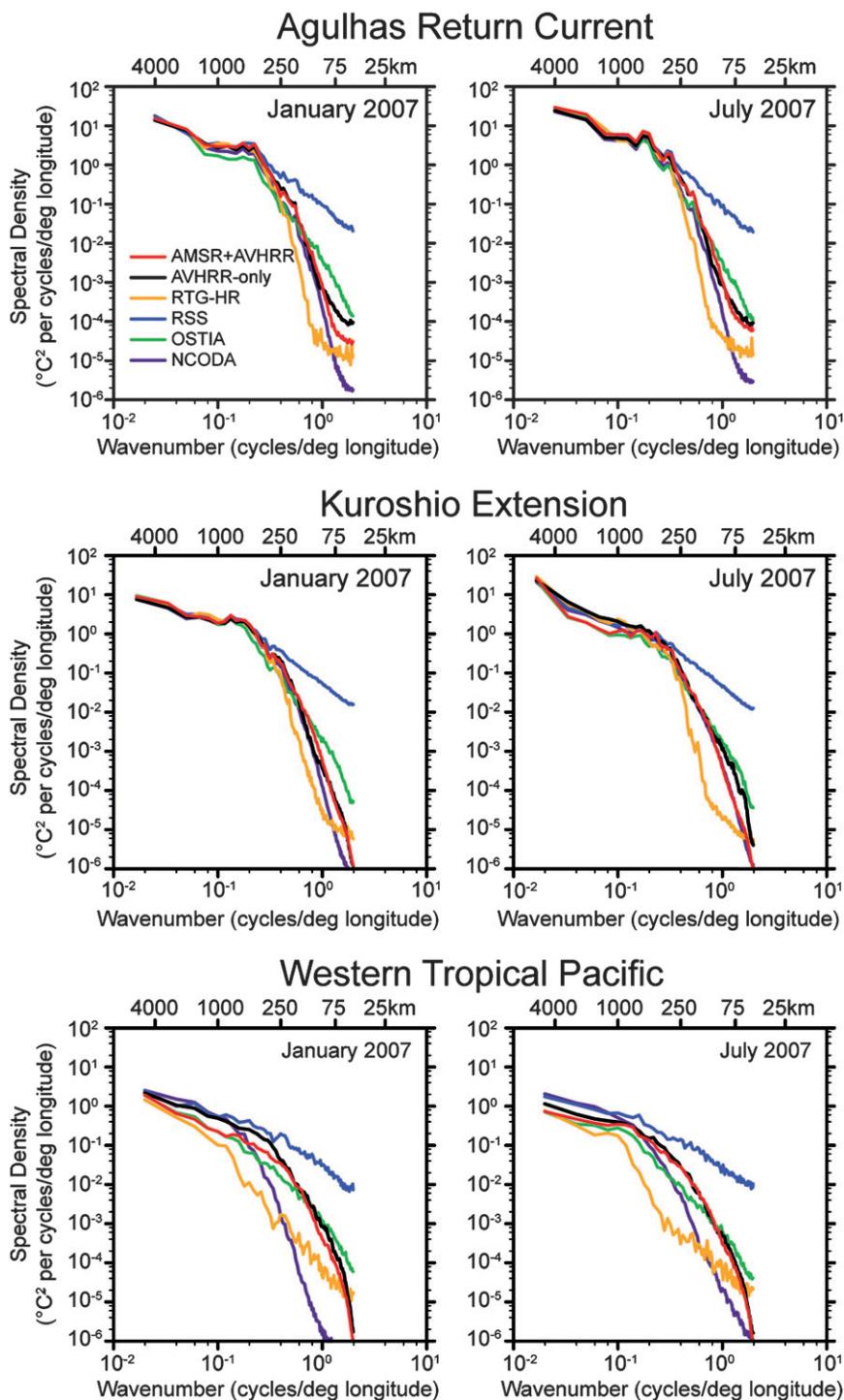
FIG. 5. Zonal wavenumber spectra for January 2007 (left) and July 2007 (right) for three regions of the World Ocean: the Agulhas Return Current (47°–38°S, 45°–85°E), the Kuroshio Extension (30°–45°N, 150°E–150°W), and the western tropical Pacific (5°S–10°N, 150°E–160°W). For each region, wavenumber spectra were computed from daily averaged SST fields in each month along each latitude of grid points within the specified domain. These individual spectra were then ensemble averaged over the latitudes and the 31 days of each month. (top left) The color key for the six SST products is shown.

made. In all six panels of Fig. 5, the OSTIA SST analyses have consistently more energy than AVHRR-only, AMSR+AVHRR, and NCODA at wavelengths shorter than about 200 km; the differences approach an order of magnitude in some cases. This higher energy may be an indication that the OSTIA analysis procedure utilizes the information content of the IR measurements more effectively than the other analyses.

In the two midlatitude regions, the spectra of the NCODA SST analyses are very similar to the other SST products for wavelengths longer than about 150–200 km but have a steeper spectral rolloff of $\sim k^{-4}$ at shorter wavelengths. As noted above, the underestimation of the spectral energy at short scales in NCODA is much worse in the WTP, where the discrepancy is more than an order of magnitude in July 2007 and nearly three orders of magnitude in January 2007. The steepness and linearity of the rolloffs of the NCODA spectra, and the visually smooth appearance of the NCODA SST fields in Fig. 3, strongly suggest that the NCODA SST analyses are overly smooth in all three of the open-ocean regions considered in these wavenumber spectral analyses. This contrasts with the high resolution of the NCODA analysis that is visually apparent in the California Current region (Fig. 4). The NCODA analysis procedure for open-ocean regions thus apparently utilizes the IR measurements less effectively than the other analyses except for RTG-HR.

The differences between the AVHRR-only and AMSR+AVHRR SST analyses in Fig. 5 are surprising. The AMSR+AVHRR spectral energy at the higher wavenumbers is significantly less in the midlatitude regions during the summertime (January in the ARC and July in the KE) and slightly less in the WTP during January. Although the summertime differences are not great, it may be counterintuitive that the AVHRR-only fields have more small-scale energy than the AMSR+AVHRR fields. This difference is likely attributable to the greater prevalence of AVHRR data in the summertime in the ARC and KE and during January in the WTP because of reduced cloudiness, which allows accurate mapping of the SST field without inclusion of AMSR data. Inclusion of the much coarser ~50-km resolution AMSR data at these times of prevalent AVHRR data evidently results in a smoothing of the SST fields that would otherwise be obtained from the ~25-km resolution AVHRR data alone. (The AVHRR data are degraded to the grid spacing, 1/4° or ~25 km, in the AVHRR-only analysis.)

In the three regions considered here, the differences between the AVHRR-only and AMSR+AVHRR thus suggest that the AMSR observations do not contribute much additional information beyond what can be obtained from AVHRR data alone and in fact can be detrimental to

the resolution of SST analyses during times of good AVHRR coverage. This result provides a very encouraging assessment of the accuracy of the long record of AVHRR-only fields that date back to September 1981, long before the availability of all-weather microwave observations of SST from TMI (since December 1997) and AMSR (since June 2002). However, on a daily basis AMSR provides the only SST retrievals in persistently cloud-covered regions. This is clearly shown in the better resolution of AMSR+AVHRR over AVHRR-only in the Gulf Stream in Fig. 9 from Reynolds et al. (2007).

5. Comparisons with buoy data

To further quantify these results, comparisons were carried out using buoy data that were gridded onto a 1/4° grid and quality controlled to eliminate extreme SST values. As noted above, buoys are used in all analyses except the RSS analysis. Thus, the buoys are not independent data. The extent to which the buoy SSTs were replicated in any particular SST product depends on how heavily they were weighted compared to the satellite and other in situ data used in the analysis procedure.

SST observations from drifting and moored buoys are typically made by a thermistor or hull contact sensor and usually are obtained in real time by satellites. Although the accuracy of the buoy SST observations varies, the random error is usually smaller than 0.5°C. Further information on buoy data is available from McPhaden et al. (1998) and Bourlès et al. (2008).

a. Buoy data

Over the 2007–08 period, the gridded buoy data were screened for locations with at least one daily observation for 90% of the days (658 days out of 731). The resulting 92 locations are shown in Fig. 6 where the data are grouped into the 10 regions shown. (If a 95% threshold were used, the number of buoy locations dropped to 81. The 95% threshold especially affected the eastern tropical Pacific region where locations east of 120°W were missing altogether). Both moored and drifting buoy data were used. However, because of the temporal coverage restrictions, almost all of the data were from moored buoys. A typical distribution of moored and drifting buoy data can be seen in Fig. 1 of Reynolds et al. (2002) and the most recent weekly distribution is available online at http://www.emc.ncep.noaa.gov/research/cmb/sst_analysis/images/inscol.png.

It should be noted that these buoy data are real-time GTS observations. The number of moored tropical Pacific (http://www.pmel.noaa.gov/tao/) and tropical Atlantic (http://www.pmel.noaa.gov/pirata/) buoy observations are denser when delayed observations are included. Figure 6
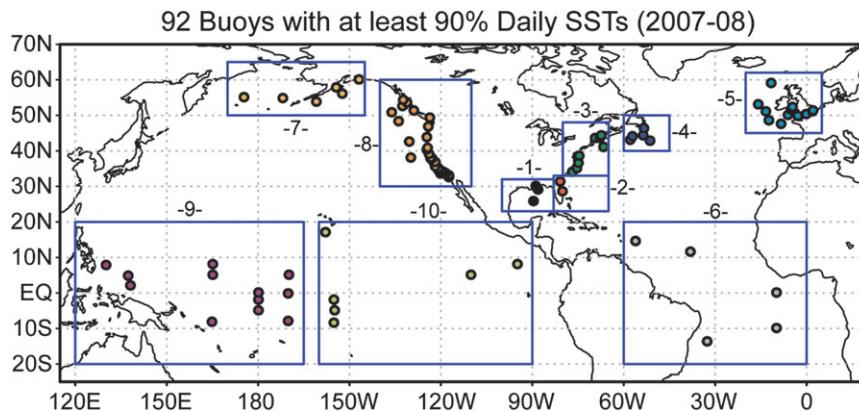
FIG. 6. Location of ¼° grid points with daily buoy SST observations for at least 90% of the time period: 1 Jan 2007–31 Dec 2008. The 10 regions are indicated in numerical order: Gulf of Mexico (GM), Gulf Stream Florida (GSF), Gulf Stream East Coast (GSEC), Gulf Stream Extension (GSE), United Kingdom/Ireland (UKI), tropical Atlantic (TA), Aleutians (AL), North America West Coast (NAWC), tropical west Pacific (TWP), and tropical east Pacific (TEP). (The TWP region shown here and used in the figures that follow is not the same as the WTP region used in section 4.)

indicates the location of 0.25° grid boxes within which the daily return from the buoys was at least 90%. Mooring locations may vary with time for several reasons. Depending on their design they may move within a watch circle with a radius of a few kilometers, and they may also move farther when pulled on by fishermen. In addition, the location may change by several kilometers or more when buoys are replaced with new moorings. This variation in location resulted in data from some mooring sites being spread between two or more 0.25° grid cells over the 2-yr period, thus reducing the amount of data for our comparisons.

Time series from each of the six SST analyses were constructed by spatial linear interpolation to the buoy locations. Missing data in all time series (both data and analysis) were filled by temporal linear interpolation. In the results that follow, statistics were computed at each buoy location and then averaged over each region.

b. Regional biases

To examine the large-scale regional biases, differences between monthly averages of the buoys and each of the analyses were computed at each buoy location (Fig. 6). The monthly differences were averaged over the region to produce a regional monthly bias. The final step was to compute the RMS of the regional monthly biases for the 24 months in 2007–08. RMS values were used rather than a simple average to avoid any possible impact of alternating signs in lowering the final bias estimates.

RMS values of the monthly regional biases are shown in Fig. 7. In almost all regions, RSS shows the largest biases, and RTG-HR has the second largest biases, while NCODA has the lowest bias. The biases are generally smaller in the tropical regions than the higher-latitude regions. The climatological SST standard deviation from the International Comprehensive Ocean–Atmosphere Dataset (ICOADS; e.g., Worley et al. 2005) increases from the tropics (roughly 1°C) to the Gulf Stream (roughly 3°C). Thus, the latitudinal variation of the bias in Fig. 7 should be expected. Four areas stand out as having the largest overall biases: the three Gulf Stream regions and the Aleutians region. In the three Gulf Stream regions, the RSS clearly has the largest biases, especially in the region off Florida. The Gulf Stream Extension region shows the largest bias difference between the AVHRR-only and the AMSR+AVHRR where the AMSR+AVHRR bias is larger than AVHRR-only. This region usually has the most cloud cover of the three Gulf Stream regions and hence shows the strongest impact from using or not using AMSR data.

The fact that AMSR+AVHRR analysis has a larger bias than AVHRR-only in the Gulf Stream Extension region may be due to bad in situ data. In winter, any bad in situ data would be more likely to be offset by an analysis that used AMSR. An example of a bad buoy just outside the Gulf Stream Extension region was found near 38°N and 63°W in January 2007 (see Fig. 10 from Reynolds et al. 2010). That figure shows a bull's-eye in an AATSR-only analysis but not in the AMSR+AVHRR analysis. The problem of bad data can of course occur in both satellite and in situ data. The comparison of different SST analyses is a good way to detect data problems because data are handled differently in each analysis product.

In the Aleutians region, the RTG-HR analysis has its strongest bias. Individual maps showed the RTG-HR had negative biases in the Aleutians and near Iceland in
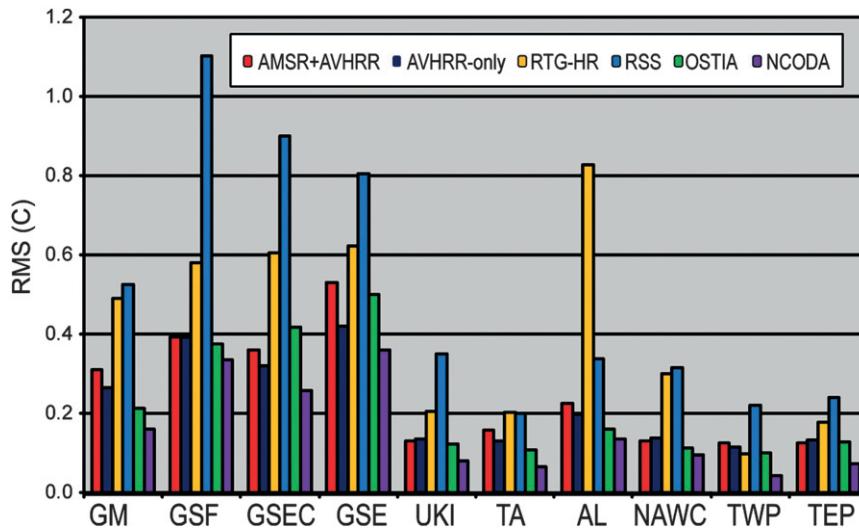
FIG. 7. RMS of the bias of the monthly average buoy − analysis difference for the six analyses and 10 regions (see Fig. 6). The largest biases occur in the Gulf Stream regions for all analyses. The buoy − NCODA RMS differences are the lowest for each region; the RSS and RTG-HR differences are generally larger than the others.

early 2007. In the RTG-HR (and AVHRR-only and AMSR+AVHRR), sea ice coverage is converted to SSTs by climate regression (Reynolds et al. 2007). If the sea ice coverage were too large, the generated SSTs would be too cold. Thus, negative biases in the Aleutians may be related to sea ice coverage.

To investigate the biases in detail, daily regional average analysis minus buoy differences were computed and temporally smoothed by a 13-day triangular filter. The smoothed differences are plotted in Fig. 8 for the Gulf Stream East Coast and Aleutians regions. The differences for the Gulf Stream region show large positive biases for the RSS analysis and large negative biases for the RTG-HR analyses. Furthermore, the OSTIA analysis shows positive biases in the first few months of 2008. In the Aleutians region, the RTG-HR analysis is more than 2°C colder than the buoys in the winter of 2007. The RTG-HR cold difference also occurred in December 2006 (not shown). The simulated SSTs derived from sea ice and used in the AVHRR-only and AMSR+AVHRR analyses did not lead to negative differences.

c. Frequency spectra

Frequency spectra were computed at each buoy location from the buoy and analysis time series. The individual spectra were averaged over each region. Figure 9 shows the frequency spectra for 3 regions. All spectra shown here are red with the spectral variance decreasing by more than three orders of magnitude from the lowest frequency (0.007 cpd) to the highest frequency (0.5 cpd). In Fig. 9 (top), the Gulf Stream East Coast region, all

spectra agree with each other within the confidence limits at low frequencies (<0.02 cpd). At middle and high frequencies, the buoy spectrum has the largest variance followed by the RSS and NCODA spectra and then by the other analyses. The RTG-HR has the lowest variance between 0.08 and 0.2 cpd. At the highest frequencies (>0.3 cpd) the AVHRR-only analysis has the lowest variance. At frequencies >0.1 cpd the buoy spectrum is often above the upper confidence limits for the RSS or NCODA spectra and is always above the confidence limits for the other analyses. These differences between buoys and analyses suggest that all of the analyses underestimate daily SST variability in high-gradient regions such as this region of the Gulf Stream.

Figure 9 (middle) shows the frequency spectra in the Aleutians region. At low frequencies (<0.02 cpd) and at the highest frequencies, the RTG-HR significantly differs from the other analyses and the buoys at the 95% confidence limit. This difference is attributable to the winter bias offset shown in the bottom panel of Fig. 8. At middle frequencies (between 0.02 and 0.2), the RSS and RTG-HR analyses have significantly greater variability than the buoys and the other analyses. In addition, the RTG-HR has highest variability at the highest frequencies while the AVHRR-only has the lowest. These high-frequency differences are significant and indicate that the RTG-HR is too noisy and that the AVHRR-only is too smooth.

The RSS differences with respect to other analyses are not due to the use of TMI in the RSS analysis. Reynolds et al. (2010) shows that the use of TMI provides only
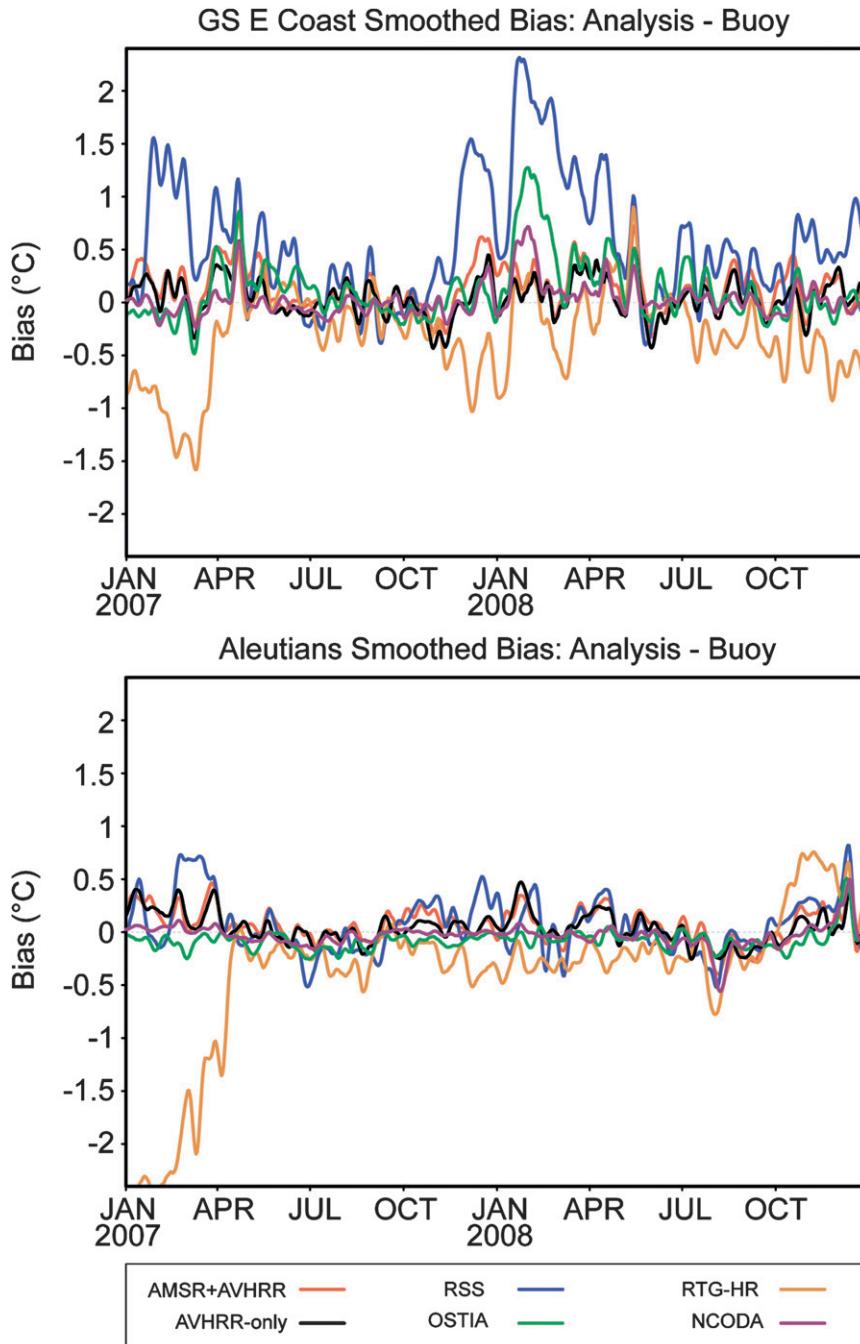
FIG. 8. Daily biases of the buoy − analysis difference for the six analyses for the Gulf Stream East Coast Extension region (top) and the Aleutians region (bottom). A 13-day triangular weighted running average was applied to the time series.

modest analysis improvements because TMI coverage is limited to between 38°S and 38°N. However, problems with small-scale noise are evident in the RSS SST fields at all latitudes. The differences are evidently due primarily to the use of the very high-resolution MODIS data in the RSS analysis.

Figure 9 (bottom) shows the frequency spectra in the western tropical Pacific region. All spectra agree with each other within the confidence limits at low frequencies (<0.02 cpd). However, the RSS analysis has significantly larger variance than the other spectra at middle and high frequencies; the RSS SST variance over these frequencies
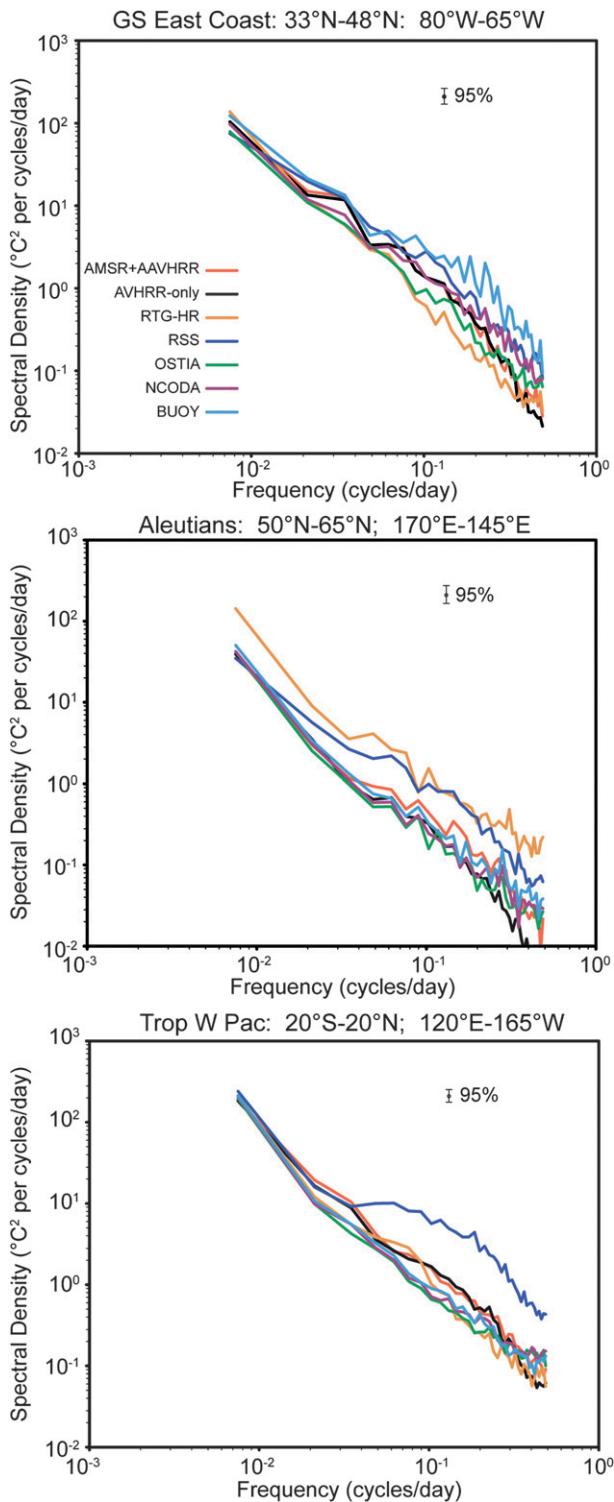
FIG. 9. Average frequency spectra for the buoys and the six analyses at the buoy locations shown in Fig. 6 for the (top) Gulf Stream East Coast region, (middle) Aleutians region, and (bottom) tropical west Pacific region. The spectra are computed at each buoy location and averaged over the region. Note the large RTG-HR difference from the other spectra in the Aleutians, and the high values for the RSS analysis for frequencies >0.02 cpd.

is roughly half an order of magnitude larger than any of the other SST spectra. A similar RSS difference occurs in other tropical regions (tropical Atlantic and eastern tropical Pacific) and is the most dramatic difference among the analyses. The RSS difference is consistent with the wavenumber spectral analysis in section 4 and the maps in section 3, which clearly indicate that the RSS analysis needs more spatial smoothing to remove small-scale noise.

d. Cross correlations

Because the spectra are red as shown in Fig. 9, the correlations between buoys and analyses will be dominated by the lowest frequencies. The correlations presented here were therefore computed from filtered time series. The filtering consisted of a half-power cutoff frequency of 0.2 cpd (a 5-day period) to separate the low- and high-frequency bands. Cross correlations were computed between the buoys and analysis time series for each region. The average time series was formed by concatenating (in the same order) the individual buoy and analysis time series at each buoy location.

The buoy-to-analysis correlations for each region are shown in Fig. 10 for the two frequency bands. The low frequencies (Fig. 10, upper panel; <0.2 cpd) all show correlations above 0.57. The two analyses with the lowest correlations are the RSS (lowest for the Gulf Stream East Coast region) and RTG-HR (lowest for the Aleutians region). If the RSS and RTG-HR correlations are excluded, all other correlations exceed 0.80.

The high-frequency correlations (Fig. 10, lower panel; >0.2 cpd) are much lower and show more differences among analyses. The NCODA correlations are the highest (between 0.70 and 0.86); the analyses with the next highest correlations vary with region among the AVHRR-only, AMSR+AVHRR and OSTIA, with OSTIA more often better correlated with the buoy observations. The correlations for RTG-HR and RSS are almost always lower than for the other products except in the Gulf Stream Extension region, where the RTG-HR has a slightly higher correlation than the AMSR+AVHRR analysis. The AVHRR-only and AMSR+AVHRR correlations are very similar (within 0.10) except in the Gulf Stream Extension region where the AVHRR-only is higher than the AMSR+AVHRR (0.61 and 0.37, respectively). The difference may be related to the cloud cover in this region. If the AVHRR data are restricted because of clouds, the buoy data will be relatively more important in the AVHRR-only analysis than the AMSR+AVHRR analysis. Thus, seasonal changes in cloud cover may impact the correlations. Note, however, that the quality of the AVHRR-only SST analysis can be expected to degrade away from the immediate vicinity of the buoys during periods of persistent cloud cover.

Buoy-Analysis: Correlations at Low Frequencies

Buoy-Analysis: Correlations at High Frequencies
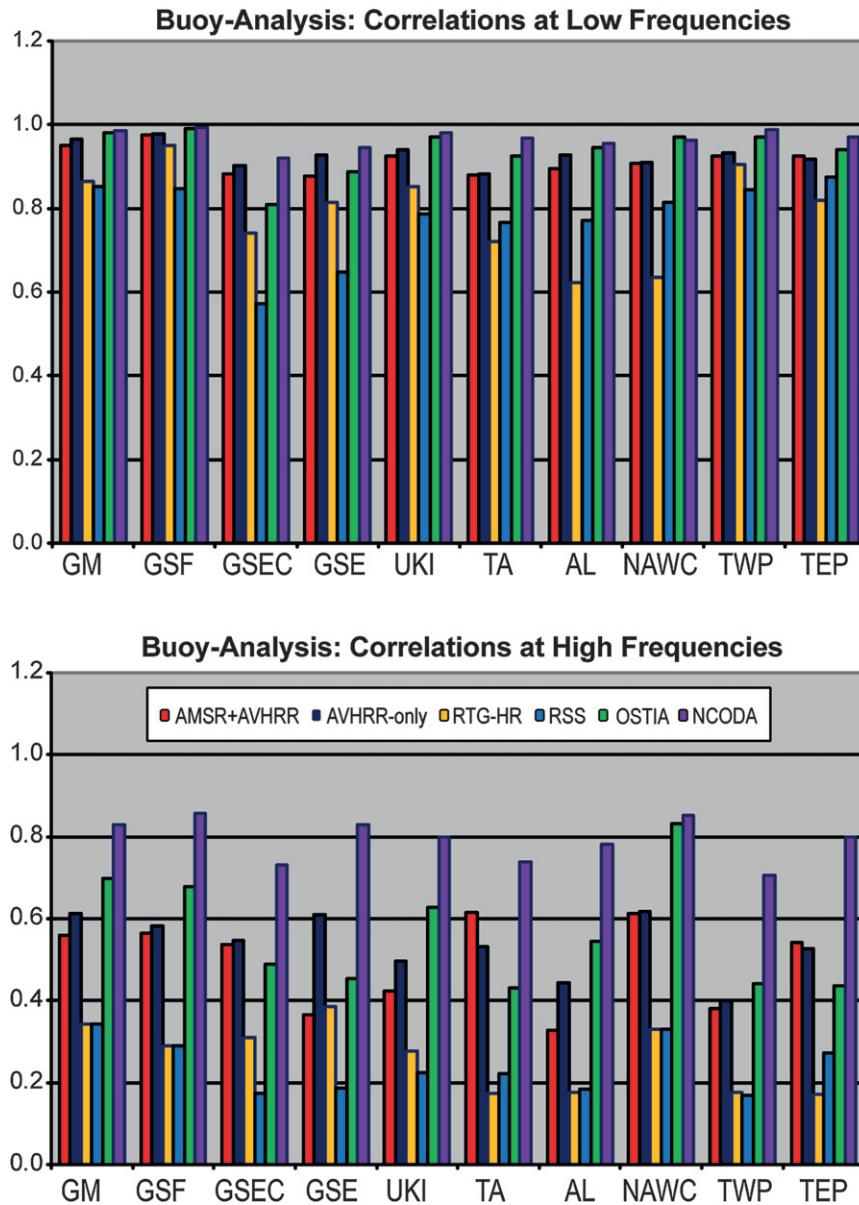

FIG. 10. Correlations for (top) low frequencies (<0.2 cpd) and (bottom) high frequencies (>0.2 cpd) computed over the 2007–08 time period for the six analyses and 10 regions (see Fig. 6). Note that the NCODA − buoy high frequency correlations are the largest.

To examine the seasonal dependence of the high-frequency correlations, the correlations were recomputed for the months of December, January, and February (DJF) and for June, July, and August (JJA), as shown in Fig. 11. Outside of the two tropical Pacific regions, correlations were generally higher for JJA than for DJF. The differences in the JJA correlations compared with the DJF correlations ranged from −0.15 to 0.40. The regions with correlation changes >0.2 were the Gulf Stream East Coast, Gulf Stream Extension, United Kingdom/Ireland, Aleutians, and North America West Coast regions. These regions are all midlatitude regions where winter seasonal cloud cover restricts IR retrievals. Furthermore, SST gradients are reduced in summer in most regions. For example, the colder shelf water during winter in the Gulf Stream region (Fig. 1) is usually not present in summer. Thus, it is easier to "fill in" any missing data in summer analyses and that process enhances the agreement between analyses and buoys. The large DJF correlation difference between AMSR+AVHRR and AVHRR-only in the Gulf Stream Extension became almost the zero in JJA. This difference again suggests that clouds
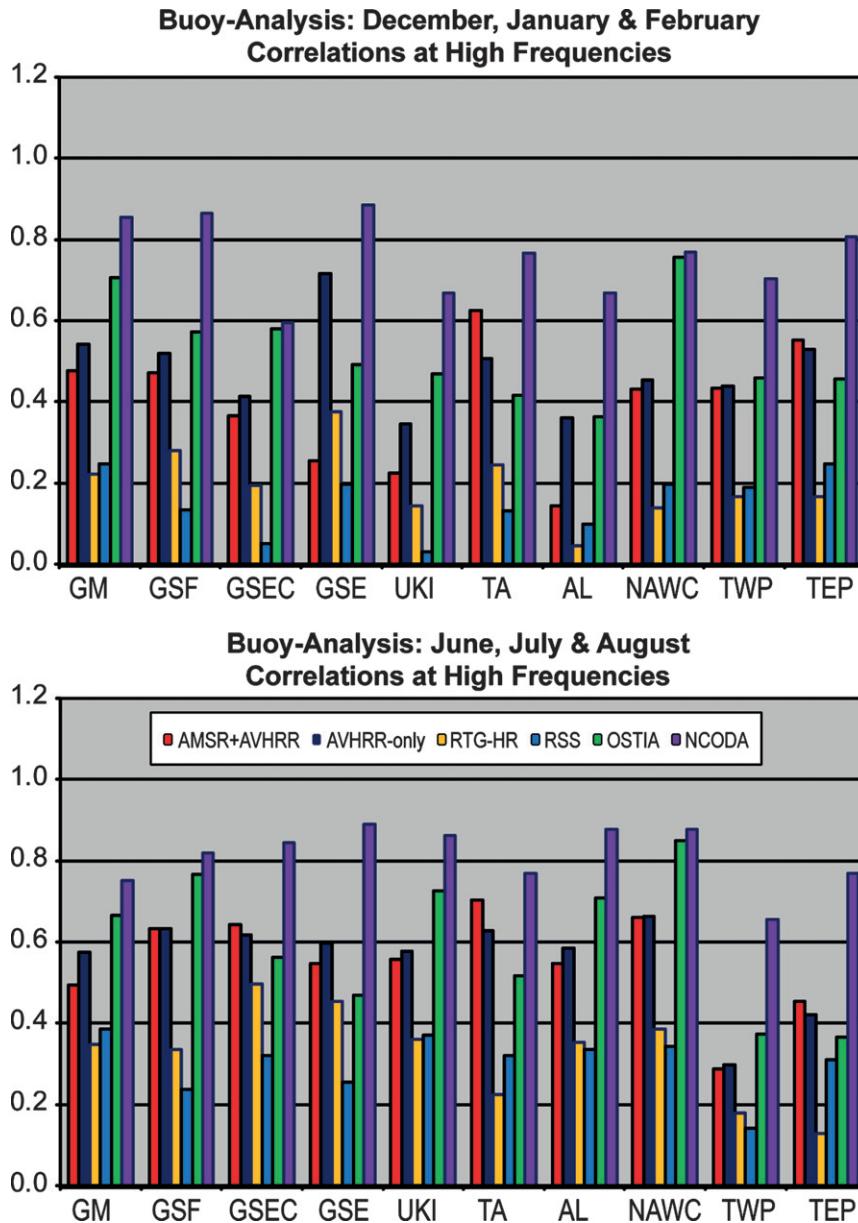
FIG. 11. Correlations for high frequencies (>0.2 cpd) for (top) DJF and (bottom) JJA computed over the 2007–08 time period for the six analyses and 10 regions (see Fig. 6).

may introduce seasonal variations of the resolution of SST in the AVHRR-only analysis.

6. Conclusions

Six different SST analyses have been compared with each other and with buoy data for the 2-yr period 2007–08. As summarized in section 2 and Table 1, the analyses are AVHRR-only, AMSR+AVHRR, NCODA, RSS, RTG- HR, and OSTIA with spatial grid sizes of ¼°, ¼°, ⅑°, 1/11°, 1/12°, and 1/20°, respectively. All analyses used differing sets of satellite data. In addition, all analyses except RSS used in situ data. Except for the AVHRR-only and AMSR+AVHRR pair of analyses, the analysis procedures and weighting functions differ for all of the products. Thus, it is not surprising that the analyzed SSTs also differ. An example off the coast of South Carolina showed SST differences that can exceed 5°C in winter. In this example, the OSTIA and RSS analyses do not clearly show the expected cold coastal water. In other regions and time periods, there are large differences among the various

analyses and the best analysis is more difficult to determine.

To assess the feature resolution capability of the various SST products, wavenumber spectra were computed at three different locations. These results clearly indicate that the RSS analysis is much noisier and the RTG-HR analysis is much smoother than the other analyses. It was shown that the feature resolution is always coarser than the grid spacing. A perhaps surprising result of the wavenumber spectra is that the spatial resolution for each SST product in general varies geographically and temporally (summer versus winter). It is therefore not possible to assign a single spatial resolution to any of the SST products considered here.

Further comparisons were made using collocated buoys for 10 regions from time series using frequency spectral analysis and cross correlations of low- and high-pass-filtered time series from the buoy data and the six SST analyses. These comparisons also showed that RSS is too noisy in the tropics. They also revealed that RTG-HR had winter biases in the Aleutians region during January and February 2007. The correlation results showed that buoy-to-analysis correlations at high frequencies (>0.2 cpd) were best for the NCODA and OSTIA analyses and worst for the RTG-HR and RSS analyses. The high correlations suggest that NCODA and, to a somewhat lesser extent, OSTIA were strongly tuned locally to buoy data, where they exist.

It is important to distinguish between the grid spacing and the feature resolution of an SST analysis. The grid spacing defines the smallest possible features that could be resolved in an analysis. A grid scale that is unnecessarily smaller than the features that are actually resolved by the analysis is computationally inefficient. Moreover, such mismatch between grid spacing and feature resolution conveys a false impression of the resolution capability to the casual user who often misinterprets grid resolution as feature resolution. The feature resolution is determined by analysis parameters such as error correlation scales and signal-to-noise ratio and can be limited by the resolution of the input observations (e.g., the footprint sizes of about 50 km for MW, 4–9 km for AVHRR, and 1 km for MODIS) and by the sampling density over the temporal period of the analysis. In the examples shown here, the OSTIA analysis has the smallest grid size and yet the feature resolution is not significantly smaller than the other analyses. If the analysis procedure does not incorporate sufficient smoothing, as in the case of the RSS analysis, the analysis will contain many small-scale features that are artifacts of small-scale measurement noise rather than signal.

Another important factor is the impact of seasonal variations in cloud cover on an SST analysis. Consider, for example, a region with 1-km IR data and 50-km MW data. During cloudy periods the IR data will be limited while the MW data will not be impacted (except in the case of rain). Thus, any analysis that attempts to obtain the highest resolution possible based on IR data unavoidably degrades this resolution when the IR data are missing or the coverage is reduced. This change in IR coverage can result in apparent temporal inhomogeneity in the small-scale variance that could wrongly be interpreted as real and may be problematic for some applications.

An additional issue is that there is generally a strong correlation between time and space scales; small features are less persistent than large features. If the analysis procedure attempts to resolve very small features in the SST field (e.g., the RSS analysis), there may be insufficient high-resolution data during cloudy periods, thus resulting in noise in the SST analysis. If the smoothing in the analysis procedure is too large (e.g., the RTG-HR analysis), the SST fields will be unnecessarily smooth. One possible solution to this dilemma is to produce an analysis with two analysis stages. In the first stage, a coarse-resolution SST analysis is produced based on combined MW and IR data similar to that described by Reynolds et al. (2007). In the second stage, a high-resolution analysis is produced with a finer grid spacing using only the available IR data; MW and in situ data are not be used directly in the high-resolution analysis. This procedure could improve the low-resolution product when sufficient IR data are available. In regions of sparse IR data, the high-resolution product would damp toward the low-resolution product. An important feature of such a 2-stage analysis is that it must include information about the coverage of the IR data so that users can assess when and where small-scale features can be adequately resolved in the second-stage analysis.

The AVHRR-only analysis is uniquely useful for climate studies because it is the only daily SST analysis that extends back to September 1981. It was shown in section 3 that inclusion of AMSR data can actually reduce the resolution of the combined AMSR and AVHRR analysis compared with the AVHRR-only analysis during cloud-free time periods. When there is extensive cloud cover, however, as often occurs at midlatitudes in the wintertime, AMSR improves the AMSR+AVHRR analysis over AVHRR-only. These results indicate that spatial and temporal variations in the satellite data that are incorporated in an SST analysis (e.g., AMSR versus AVHRR or MODIS) can impact the resolution of the resulting SST fields and result in artificial spatial and temporal variability of the apparent scales of features in the SST field. This should be kept in mind when using the SST analyses to investigate climate variability.

REFERENCES

Bourlès, B., and Coauthors, 2008: The PIRATA Program: History, accomplishments, and future directions. *Bull. Amer. Meteor. Soc.,* **89,** 1111–1125.

Castelao, R. M., T. P. Mavor, J. A. Barth, and L. C. Breaker, 2006: Sea-surface temperature fronts in the California Current System from geostationary satellite observations. *J. Geophys. Res.,* **111,** C09026, doi:10.1029/2006JC003541.

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. Meteor. Soc.,* **86,** 1097–1115.

——, R. A. de Szoeke, M. G. Schlax, K. El Naggar, and N. Siwertz, 1998: Geographical variability of the first baroclinic Rossby radius of deformation. *J. Phys. Oceanogr.,* **28,** 433–460.

Cummings, J. A., 2005: Operational multivariate ocean data assimilation. *Quart. J. Roy. Meteor. Soc.,* **131,** 3583–3604.

Donlon, C., and Coauthors, 2007: The Global Ocean Data Assimilation Experiment High-resolution Sea Surface Temperature Pilot. *Bull. Amer. Meteor. Soc.,* **88,** 1197–1213.

Gemmill, W., B. Katz, and X. Li, 2007: Daily real-time global sea surface temperature—High resolution analysis at NOAA/NCEP. NOAA/NWS/NCEP/MMAB Office Note 260, 39 pp. [Available online at http://polar.ncep.noaa.gov/sst/.]

Gentemann, C. L., C. J. Donlon, A. Stuart-Menteth, and F. J. Wentz, 2003: Diurnal signals in satellite sea surface temperature measurements. *Geophys. Res. Lett.,* **30,** 1140, doi:10.1029/2002GL016291.

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.

McPhaden, M. J., and Coauthors, 1998: The Tropical Ocean–Global Atmosphere observing system: A decade of progress. *J. Geophys. Res.,* **103** (C7), 14 169–14 240.

Reynolds, R. W., and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. *J. Climate,* **7,** 929–948.

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

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

——, C. L. Gentemann, and G. K. Corlett, 2010: Evaluation of AATSR and TMI satellite SST data. *J. Climate,* **23,** 152–165.

Thiébaux, 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.

Worley, S. J., S. D. Woodruff, R. W. Reynolds, S. J. Lubker, and N. Lott, 2005: ICOADS release 2.1 data and products. *Int. J. Climatol.,* **25,** 823–842.