Objective Determination of Feature Resolution in Two Sea Surface Temperature Analyses

RICHARD W. REYNOLDS

Cooperative Institute for Climate and Satellites, North Carolina State University, and 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

JONAH ROBERTS-JONES AND MATTHEW J. MARTIN

Met Office, Exeter, United Kingdom

DIMITRIS MENEMENLIS

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

CHRISTOPHER JOHN MERCHANT

University of Edinburgh, Edinburgh, United Kingdom

(Manuscript received 5 November 2012, in final form 22 January 2013)

ABSTRACT

Considerable effort is presently being devoted to producing high-resolution sea surface temperature (SST) analyses with a goal of spatial grid resolutions as low as 1 km. Because grid resolution is not the same as feature resolution, a method is needed to objectively determine the resolution capability and accuracy of SST analysis products. Ocean model SST fields are used in this study as simulated ''true'' SST data and subsampled based on actual infrared and microwave satellite data coverage. The subsampled data are used to simulate sampling errors due to missing data. Two different SST analyses are considered and run using both the full and the subsampled model SST fields, with and without additional noise. The results are compared as a function of spatial scales of variability using wavenumber auto- and cross-spectral analysis.

The spectral variance at high wavenumbers (smallest wavelengths) is shown to be attenuated relative to the true SST because of smoothing that is inherent to both analysis procedures. Comparisons of the two analyses (both having grid sizes of roughly $^1/_{20}°$) show important differences. One analysis tends to reproduce small-scale features more accurately when the high-resolution data coverage is good but produces more spurious small-scale noise when the high-resolution data coverage is poor. Analysis procedures can thus generate small-scale features with and without data, but the small-scale features in an SST analysis may be just noise when high-resolution data are sparse. Users must therefore be skeptical of high-resolution SST products, especially in regions where high-resolution (~5 km) infrared satellite data are limited because of cloud cover.

1. Introduction

During the last decade the number of sea surface temperature (SST) analyses on a high-resolution grid has steadily increased as producers strive to satisfy the needs of the user community. This increase is partly due to the establishment of the Group for High-Resolution Sea Surface Temperature (GHRSST) (Donlon et al. 2007), which encourages the development of high-resolution data and analysis products. High-resolution analyses are facilitated by a recent increase in the number of satellite microwave (MW) and infrared (IR)

Corresponding author address: Richard W. Reynolds, NOAA/Cooperative Institute for Climate and Satellites, 151 Patton Avenue, Asheville, NC 28801.
E-mail: richard.w.reynolds@noaa.gov

DOI: 10.1175/JCLI-D-12-00787.1

SST data products available from a variety of geostationary and polar-orbiting satellites.

Assessment of the various analysis products is difficult because the analyses use differing combinations of in situ and remotely sensed data and are computed over different regions and time periods with different spatial and temporal grid resolutions, as detailed for a number of GHRSST products in Martin et al. (2012). Moreover, the true SST is not known so it is often difficult to determine whether one analysis is better than another. In any analysis procedure, irregularly spaced data are smoothed and interpolated onto a regular grid by some objective analysis procedure. As determined by the input data and smoothing parameters in the analysis, the feature resolution will in general vary in time and space. The problem for many users is not to find an SST analysis but to figure out which one is the best for their particular purpose.

Many users and developers of SST analysis products confuse analysis grid resolution with the scales of the SST features that can be resolved in the analysis. From a comparison of six analyses for a 2-yr period, Reynolds and Chelton (2010) found that the grid spacing of an SST analysis is not an indicator of the feature resolution of an analysis. Spatial grid scales and analysis feature resolution are thus not the same. Feature resolution in fact changes as actual coverage from the various satellite instruments change due to cloud cover for IR and precipitation for MW. Moreover, analyses fill in missing data and can produce small-scale noise when high-resolution data are not available.

The purpose of this study is to show how the space–time distributions of data can degrade the small-scale signal in an analysis. If high-resolution data are not available, it may seem intuitive that analyses will not show small-scale features. It is shown that this assumption is not necessarily correct and that high-resolution analyses can generate small-scale features both with and without high-resolution data. In the latter case the small-scale features may be just noise.

The procedure used in this study to demonstrate how limited data impacts analyses begins with daily complete SST fields from a high-resolution global model, which are treated as the ''true'' SST. Both low-resolution (~50 km) and high-resolution (~5 km) pseudodata were generated from the model to represent MW and IR satellite data, respectively. These full fields of low- and high-resolution SST were then subsampled based on actual times and locations of satellite SST data from one MW instrument and one IR instrument. The full and subsampled SST data fields were then analyzed using two different analysis procedures to assess the influence of missing low-resolution and high-resolution data. Differences between the full and analyzed SST fields quantify the effects of filtering and sampling errors in each of the two analysis procedures considered here. To study the impact of noise in satellite measurements of SST, realistic measurement noise was added to the full and subsampled model data. The analyses were then recomputed with the added noise.

The simulated datasets used in this study are based on sampling distributions of IR data from the Advanced Very High-Resolution Radiometer (AVHRR) and MW data from the Advanced Microwave Scanning Radiometer (AMSR). The analyses are intended primarily to illustrate a technique for assessing resolution capability and to draw attention to the fact that small-scale features in an analysis can be artifacts of noise, depending on the sparseness of the high-resolution measurements included in the analysis. This problem can be mitigated by inclusion of IR measurements from sources besides the AVHRR data considered in this study. The improved IR sampling that would be achieved from multiple satellites with IR sensors is simulated roughly in this study by utilizing AVHRR measurement locations over three successive days in each daily SST analysis (see section 2b). In regions or time periods of persistent cloud cover over multiple days, there will still be significant limitations in the resolution capability of a high-resolution SST analysis. These limitations can be quantified for any analysis procedure using the methodology presented here.

The different simulated datasets and different analyses that are considered in the sections that follow are defined as they are introduced in the text. Each is given an abbreviated name to simplify the discussion in the text and the labeling in the figures. The complete and short names are listed in Table 1.

2. Pseudodata

a. The ocean model

The true SST fields for this study were obtained from a quasi-global (78.7°S–80°N), eddy-admitting ocean simulation provided by the Estimating the Circulation and Climate of the Ocean, Phase II project (ECCO2; Menemenlis et al. 2008). This simulation is described by Hill et al. (2007). It was carried out using the Massachusetts Institute of Technology general circulation model (Marshall et al. 1997) on a grid with $1/16°$ horizontal grid spacing in the zonal direction and $1/16°$ cos (latitude) in the meridional direction. The meridionally varying grid spacing keeps each model grid cell approximately square; horizontal grid spacing ranges from 6.9 km at the equator to 1.2 km at 80°N. The model has

TABLE 1. Abbreviated and full names: All datasets use 3 days of data.

Abbreviation	Full name
Model-Hi-Full	High-resolution ocean model full dataset without noise
Model-Hi-Red	High-resolution ocean model reduced dataset without noise
Model-Lo-Full	Low-resolution ocean model full dataset without noise
Model-Lo-Red	Low-resolution ocean model reduced dataset without noise
OI-Hi-Full	High-resolution OI using the full dataset without noise
OI-Hi-Red	High-resolution OI using the reduced dataset without noise
OI-Lo-Full	Low-resolution OI using the full dataset without noise
OI-Lo-Red	Low-resolution OI using the reduced dataset without noise
OSTIA-Full	OSTIA using the full dataset without noise
OSTIA-Red	OSTIA using the reduced dataset without noise
Model-Hi-Full+Noise	High-resolution ocean model full dataset with added noise
Model-Hi-Red+Noise	High-resolution ocean model reduced dataset with added noise
Model-Lo-Full+Noise	Low-resolution ocean model full dataset with added noise
Model-Lo-Red+Noise	Low-resolution ocean model reduced dataset with added noise
OI-Hi-Full+Noise	High-resolution OI using the full dataset with added noise
OI-Hi-Red+Noise	High-resolution OI using the reduced dataset with added noise
OI-Lo-Full+Noise	Low-resolution OI using the full dataset with added noise
OI-Lo-Red+Noise	Low-resolution OI using the reduced dataset with added noise
OSTIA-Full+Noise	OSTIA using the full dataset with added noise
OSTIA-Red+Noise	OSTIA using the reduced dataset with added noise

50 vertical levels ranging in thickness from 10 m at the surface to 457 m at a maximum depth of 6135 m. Bathymetry was derived from the National Geophysical Data Center 2-minute gridded elevations/bathymetry for the world (ETOPO2), version 2001.

The simulation was initialized on 1 January 1992 from rest and from January temperature and salinity fields from the *World Ocean Atlas 2001* (*WOA01*) (Conkright et al. 2002). It was forced at the surface using fluxes (wind stress, heat, and freshwater) and penetrating shortwave radiation from the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis (Kistler et al. 2001). Surface forcing also includes time-mean river runoff from Large and Nurser (2001) and SST and salinity relaxation terms to the NCEP–NCAR reanalysis and the monthly-mean *WOA01* surface salinity, respectively, with a relaxation time scale of 60 days. Vertical mixing follows the *K*-profile parameterization of Large et al. (1994). Except for the surface temperature relaxation term, the simulation does not include a representation of sea ice. SSTs therefore occasionally fall below the freezing point in polar regions.

Daily averaged model temperatures from 1 November 1992 through 6 July 1994 at the level nearest to the surface (0–10-m depth) were used as the true SST. The SST variability for this period is typical of the variability over the last few decades (e.g., Fig. 2 of Reynolds et al. 2002). The dynamic range of the SST values is between −2° and 39°C, excluding roughly 5% of high-latitude grid cells, in which temperatures were below the freezing point of seawater because of the lack of a sea ice model. In the analyses that follow, any model SST less than −2°C was set to −2°C.

It is important to note that the model fields do not have to be completely realistic in every respect to examine the impact of sparse high-resolution data on analyses. All that is required is that the spatial variability on scales smaller than 100 km and the temporal variability on scales of days be statistically consistent with actual variability so that the challenges imposed on SST analysis procedures by missing data are realistically simulated. While the results obtained from model output from a higher-resolution numerical simulation would differ qualitatively from the results presented in this study, it is apparent from sections 4–7 that the grid resolution limitation of ECCO2 is less of an issue than the resolution limitations of the SST analysis procedures considered here.

The temporal variability of the model SST was checked by computing monthly mean anomalies along with the monthly standard deviations from the daily anomalies on a ¼° grid for 1993. These fields were compared with the monthly-mean anomalies and monthly standard deviations computed from the daily optimum interpolation (OI) analysis of Reynolds et al. (2007) for the same time period and same grid. The analysis was computed using in situ data and one or more IR satellites using an AVHRR instrument. The two anomaly means and standard deviations for July 1993 are shown in Figs. 1 and 2, respectively. The model mean anomaly field resembles strong El Niño conditions in the tropical Pacific SSTs with magnitudes up to 6°C larger than the analyzed anomalies. However, the two standard deviation fields are qualitatively similar, except that the model fields are stronger by up to 2°C in western boundary current regions such as the Gulf Stream. It is likely that the Reynolds et al. (2007) analysis underestimates the variability in these regions as

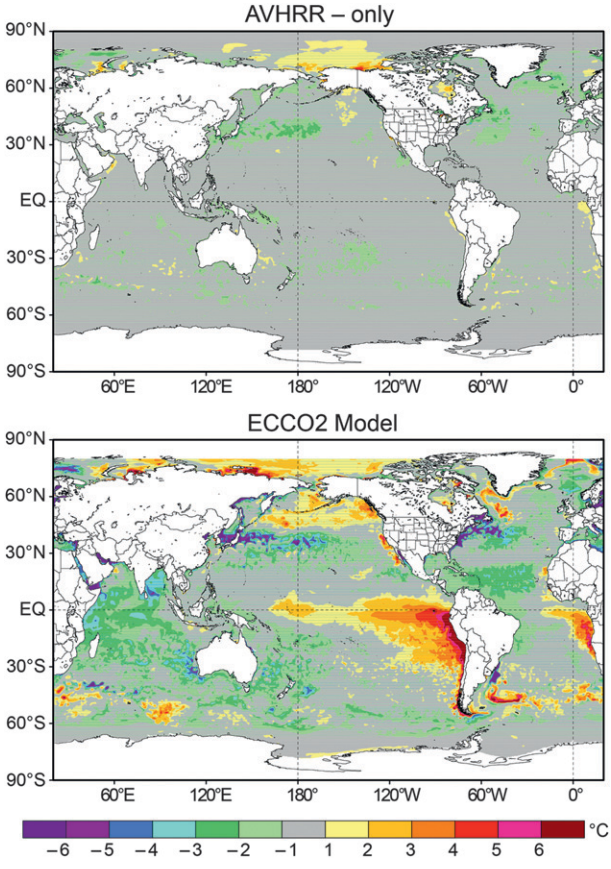

FIG. 1. Mean SST anomaly for July 1993 for (top) the AVHRR-only analysis of Reynolds et al. (2007) and (bottom) the ECCO2 model. The color scale is in degrees Celsius.

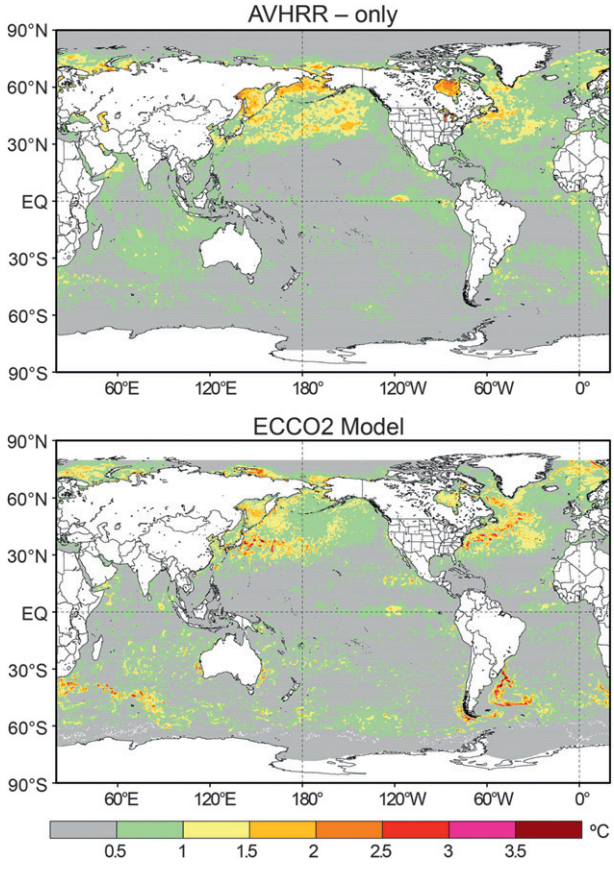

FIG. 2. As in Fig. 1, but standard deviation of the daily SST anomalies for July 1993.

it was based on only one AVHRR instrument during the July 1993 time period. In later periods, when MW data from the AMSR instrument were included in the analysis, the monthly variability increased in these western boundary current regions (Reynolds et al. 2007).

b. Data processing

To simulate satellite data coverage realistically, the MW and IR measurement times and locations were based on actual measurements from one AMSR instrument and one AVHRR instrument, respectively. Because global AMSR data were not available until June 2002, it was necessary to use a time period for the satellite data that differs from the 1993 period of the ECCO2 model run. The model dates were thus increased by 11 yr to link the model with the satellite coverage. This shift in the dates of satellite coverage with respect to the dates of the model is not critical, as only the satellite coverage, not actual satellite SSTs, is used in this study. The data coverage of both AVHRR and AMSR is typical of the instruments and quite good for this period. [Coverage for 2003 is shown in Fig. 2 of Reynolds et al. (2007).]

As described by Chelton and Wentz (2005), AMSR SST retrievals have a footprint size of about 50 km and are contaminated by precipitation and within about 75 km of land or ice. The primary advantage of AMSR data is the near-all-weather measurement capability. Version-5 AMSR data were obtained from Remote Sensing Systems (RSS) as twice daily (daytime and nighttime) SST fields on a 0.25° grid (or 28 km at the equator; from http://www.ssmi.com/).

AVHRR SST retrievals were obtained from Pathfinder version 5 (Kilpatrick et al. 2001), which is based on one satellite instrument and produced twice daily (daytime and nighttime) on a 0.043 95° grid (~1/23° or 4.9 km at the equator). AVHRR retrievals require a cloud-free view of the ocean but can be retrieved near land and sea ice.

The gridded SST fields for both AMSR and AVHRR use latitude and longitude spacing. The latitude spacing in km is therefore constant. However, the longitude spacing in kilometers decreases with increasing latitude as cos(latitude).

To obtain complete high-resolution datasets without any satellite coverage restrictions, the daily model SST fields were linearly interpolated to the 0.043 95° (~1/23°) AVHRR Pathfinder v5 grid. This interpolation adds a small amount of small-scale noise to the original model SST fields because the linear interpolation uses a series of linear segments with discontinuous slopes at segment junctions. This noise is not a significant issue for this study. The resulting SST fields on the 0.043 95° grid are defined as truth and are referred to as the high-resolution full daily model data. Complete low-resolution datasets were similarly obtained from the model SST fields by approximating the AMSR 50-km footprint as averages over ½° × ½°. To represent the oversampling of AMSR to the 0.25° grid on which the data are delivered, these averages are linearly interpolated to a ¼° grid. The resulting SST dataset on the ¼° grid is referred to as the low-resolution full daily model data.

To reduce the sampling of the high-resolution dataset based on actual measurement times and locations, the Pathfinder AVHRR daytime and nighttime data for each day were examined. If there was a daytime or nighttime value on the day corresponding to each day of high-resolution daily model data, the model data value was kept; otherwise it was set to missing. This reduction simulates IR data loss from cloud cover as well as the swath width of the AVHRR instrument. The resulting dataset is the high-resolution reduced daily model data.

Initially, the actual RSS AMSR daytime and nighttime data were processed by a similar procedure. However, as AMSR data are not available within 75 km of land, this would make any SST analysis variability based only on AMSR coverage too low near the coast. When permitted by clear-sky conditions, the Pathfinder AVHRR daytime and nighttime data were therefore averaged onto a ¼° grid and also included as IR data would normally be included in any SST analysis. The low-resolution daily reduced model data were then obtained on a ¼° grid by using both AVHRR and AMSR coverage. If either daytime or nighttime AVHRR or AMSR data were available, the low-resolution model data value was kept; otherwise, it was set to missing. Because of the superior open-ocean AMSR coverage, the addition of AVHRR had little impact on the coverage except in coastal regions and near sea ice margins.

The high-resolution reduced daily data are typically very sparse on any particular day. Furthermore, many current SST analyses often use multiple IR instruments from different satellites. To compensate for the use of only one IR instrument in this study, coverage more representative of multiple IR instruments was achieved by averaging all four of the daily datasets described above over three adjacent days. Data from the middle day, if available, were given twice the weight of the other days in the 3-day average. The SST analyses considered in this study (see the following section) are based on these 3-day averages. To simplify the names, both "3 day" and "data" are henceforth dropped. The full and reduced high-resolution datasets will be referred to as "Model-Hi-Full" and "Model-Hi-Red." The corresponding full and reduced low-resolution datasets will be referred to as "Model-Lo-Full" and "Model-Lo-Red" (see Table 1 for the complete naming convention).

The SST fields for the four datasets are shown in Fig. 3 for the region of the Gulf Stream southeast of Nova Scotia for the 3-day period centered on 1 July 1993. The Model-Lo-Full and Model-Hi-Full (Fig. 3, left) show the existence of small-scale variability that is not resolved by the low-resolution data. The Model-Lo-Red (Fig. 3, top right) has almost the same coverage as the Model-Lo-Red, thus showing the benefits of the all-weather coverage of the AMSR data. The only missing values are in the northwest region where the AMSR data are unavailable because of the presence of land. Some incomplete coverage in this region is provided by the available AVHRR data averaged over ¼° as described above. The Model-Hi-Red (Fig. 3, bottom right) based only on 3 days of AVHRR data shows poor coverage compared to Model-Hi-Full. Note in particular that the cold pattern near 39° and 58°N is only very partially resolved in Model-Hi-Red. Although sparsely sampled, these days had much better than usual reduced high-resolution coverage for this region. Other days, especially in winter, had almost no reduced high-resolution coverage, even with the 3-day averaging period. It is therefore often impossible for any analysis to be able to properly resolve all the high-resolution features both in space and in time using the reduced high-resolution model data.

3. Analyses

By definition an SST analysis produces complete fields on a regular grid and must smooth available data to fill in missing grid points. Analyses can therefore be thought of as spatial and temporal low-pass filters. The four model datasets summarized in section 2b were used in two different analysis procedures to investigate how high-resolution accuracy is lowered as the high-resolution full data are reduced by actual data coverage. The two analysis procedures were run daily for two 2-month time periods, 1 December 1992–31 January 1993 and 30 June–31 July 1993. The results are summarized in section 4 based on the second month of each period. Both analysis procedures use the previous day as a first guess. The first month in each period allows for complete spinup of the analyses.

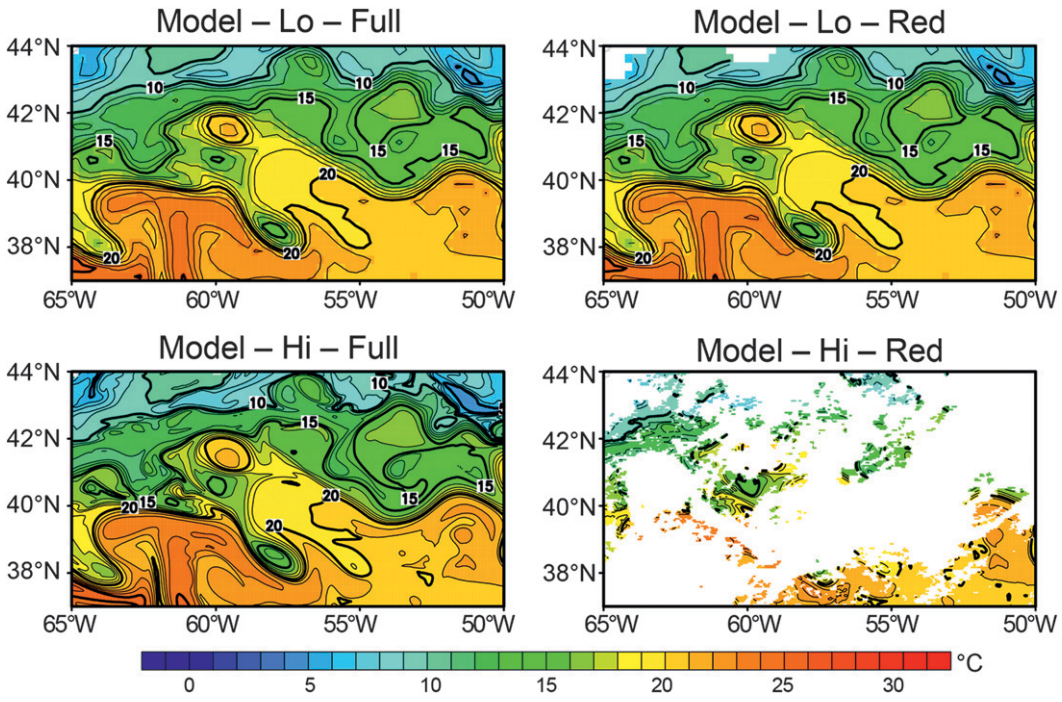

FIG. 3. (top) Low-resolution and (bottom) high-resolution ocean model SST data (3-day average) for 1 Jul 1993 (see text). (right) The data reduced by actual MW and IR data coverage and (left) the full data. The color scale is in degrees Celsius.

a. Two-stage OI

This analysis procedure is carried out in two stages using the full and reduced datasets on both the ¼° and the 0.043 95° (~1/23°) grids. The two stages both use an OI procedure. The first stage consists of a low-resolution analysis that has been described in Reynolds et al. (2007). As implemented here, this first-stage analysis does not use in situ data or include a preliminary bias correction step. The damping, noise-to-signal ratios and spatial correlation scales follow the results described in Reynolds et al. (2007). The ''OI-Lo-Full'' and ''OI-Lo-Red'' analyses use the Model-Lo-Full and Model-Lo-Red datasets, respectively.

The second stage produces a high-resolution analysis on a 0.043 95° grid using the high-resolution data and a first guess. The first guess is a combination of the low-resolution analysis at the current time step and the damped difference between the high-resolution analysis and the low-resolution analysis at the previous time step. (The damping field was computed from 1-day lagged autocorrelations of 3-day-averaged AVHRR data for 2003. The field was fairly smooth ranging from 0.4 to 0.7 with an average of 0.6. The larger damping values occurred between 40°S and 40°N.) The low-resolution analysis impacts the high-resolution analysis through the low-resolution first guess. However, the high-resolution analysis has no effect on each daily low-resolution analysis: that is, the link between the two stages is only one way. The two-stage processing is designed to have the high-resolution analysis relax to the low-resolution analysis in the absence of any high-resolution data.

The spatial error correlations and noise-to-signal variance ratios were derived from the values defined in Reynolds et al. (2007) for the low-resolution analysis. In this version the high-resolution spatial error correlations were reduced from the low-resolution values by a factor of 1/R; the high-resolution noise-to-signal variance ratios were increased from the low-resolution values by a factor of R. The value of R was determined by experimentation to be 2.8, which is half the ratio of the high-resolution to low-resolution grid sizes. The ''OI-Hi-Full'' analysis uses the Model-Hi-Full data and the OI-Lo-Full analysis; the ''OI-Hi-Red'' analysis uses the Model-Hi-Red data and the OI-Lo-Red data.

b. OSTIA

The other SST analysis procedure considered here is the Operational SST and Sea Ice Analysis (OSTIA) produced by the Met Office on a 1/20° grid that uses in-situ, AVHRR, AMSR, the Advanced Along Track Scanning Radiometer (AATSR), the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI),

and the geostationary Spinning Enhanced Visible and Infrared Imager (SEVIRI) data. Geostationary Operational Environmental Satellites (GOES) East and Infrared Atmospheric Sounding Interferometer (IASI) data were added to the OSTIA analysis in October 2011. (AMSR data are not available after the AMSR instrument failed in October 2011; AATSR data are not available after that instrument failed in April 2012.) The analysis procedure uses two background error correlation scales, 10 and 100 km, with associated error variances that vary depending on the location. All satellite SST data are adjusted for bias errors by reference to a combination of AATSR and in situ SST measurements. Operational users of the OSTIA system (e.g., numerical weather prediction systems) do not want the characteristics of the analysis to which they are sensitive (such as SST gradients) to change solely because of changes in data coverage. The OSTIA system was therefore designed to produce a temporally consistent SST analysis that is robust to changes in high-resolution IR data coverage. A full discussion on the OSTIA system design choices and further details can be found in Donlon et al. (2012) (and at http://ghrsst-pp.metoffice.com/pages/latest_analysis/ostia.html).

For this study, OSTIA was run with either the full or the reduced low- and high-resolution model datasets without any bias correction based only on the simulated AVHRR and AMSR data derived as described in section 2b. These analyses are referred to as ''OSTIA-Full'' and ''OSTIA-Red,'' respectively.

An important point to note is that exactly the same simulated data were used for both of the SST analyses considered in this study. Any differences between the SST fields from the OSTIA and two-stage OI analyses of the simulated SST data derived as described in section 2b are thus indicative of differences between the results of the SST analysis procedures rather than differences in the data input to the two procedures.

4. Results

All results that follow are shown on the high-resolution 0.043 95° (~1/23°) AVHRR Pathfinder v5 grid. The low-resolution first stage of the two-stage OI is considered only a preliminary internal processing step. However, the low-resolution first stage interpolated onto the high-resolution analysis grid is shown when needed to highlight any improvements made by the high-resolution data. The OSTIA analysis is also interpolated from the 1/20° grid to the Pathfinder grid. Both the two-stage OI and OSTIA were computed globally. Six regions were selected to be of special interest. These included three western current boundary regions (Agulhas, Kuroshio, and Gulf Stream), a region with persistent clouds (U.S. West Coast), a relatively clear region (Sargasso Sea), and a region in the tropical eastern Pacific. Comparisons for these six regions showed that only two representative regions are needed to characterize the results: the Gulf Stream region (37°–44°N, 65°–50°W) and the Sargasso Sea region (20°–30°N, 60°–30°W). The Gulf Stream region was selected because of its strong SST gradients and because it is often cloudy, especially in the winter. The satellite data available for SST analyses in this region are therefore often restricted to low-resolution MW data. The Sargasso Sea region was selected as a counter example because it is often cloud free throughout and thus often has IR high-resolution data available for SST analyses. The daily analysis results are presented for two months, January and July 1993, which were selected to capture seasonal changes in both the SST field and the sampling distribution of the simulated IR data.

An example of the daily results is presented in Fig. 4 for the Gulf Stream region for 1 July 1993, the same date shown in Fig. 3. The full high-resolution model data and analyses are shown in the left panels. Features in the full data (Model-Hi-Full) are captured by the OI-Hi-Full with some slight smoothing. OSTIA-Full is similar but clearly shows heavier smoothing of the model data; for example, note that the tight eddy gradients in the model data are reduced by the OSTIA processing. The reduced high-resolution model data and analyses are shown in the right panels. Clearly it is not possible to do a high-resolution analysis near 40°N and 55°W where there are no high-resolution model data. Note that the Model-Hi-Red is noisier than the Model-Hi-Full, even in regions where both datasets are defined. This difference is because model data averaged over three consecutive days are based on whatever data are available. Data were not always available for all three days. Because small-scale features tend to evolve relatively quickly, the undersampling of this variability in the simulated IR data manifests as noise in the Model-Hi-Red SST fields. This date was actually a time with reasonably good coverage compared with other days. The missing data are the reason that the OI-Hi-Red field is noticeably noisier than the OI-Hi-Full.

Differences between the reduced and full versions of OSTIA are much smaller. Because the OSTIA analysis procedure lowers the analysis resolution, as noted above, the input data are effectively smoothed and the impact of missing data is not as strong as it is for the OI analysis procedure.

Maps for the various SST fields for most of the Sargasso Sea region are shown in Fig. 5. Comparison of the full and reduced model datasets shows that spatial variability of the SST fields is much smoother compared to the Gulf

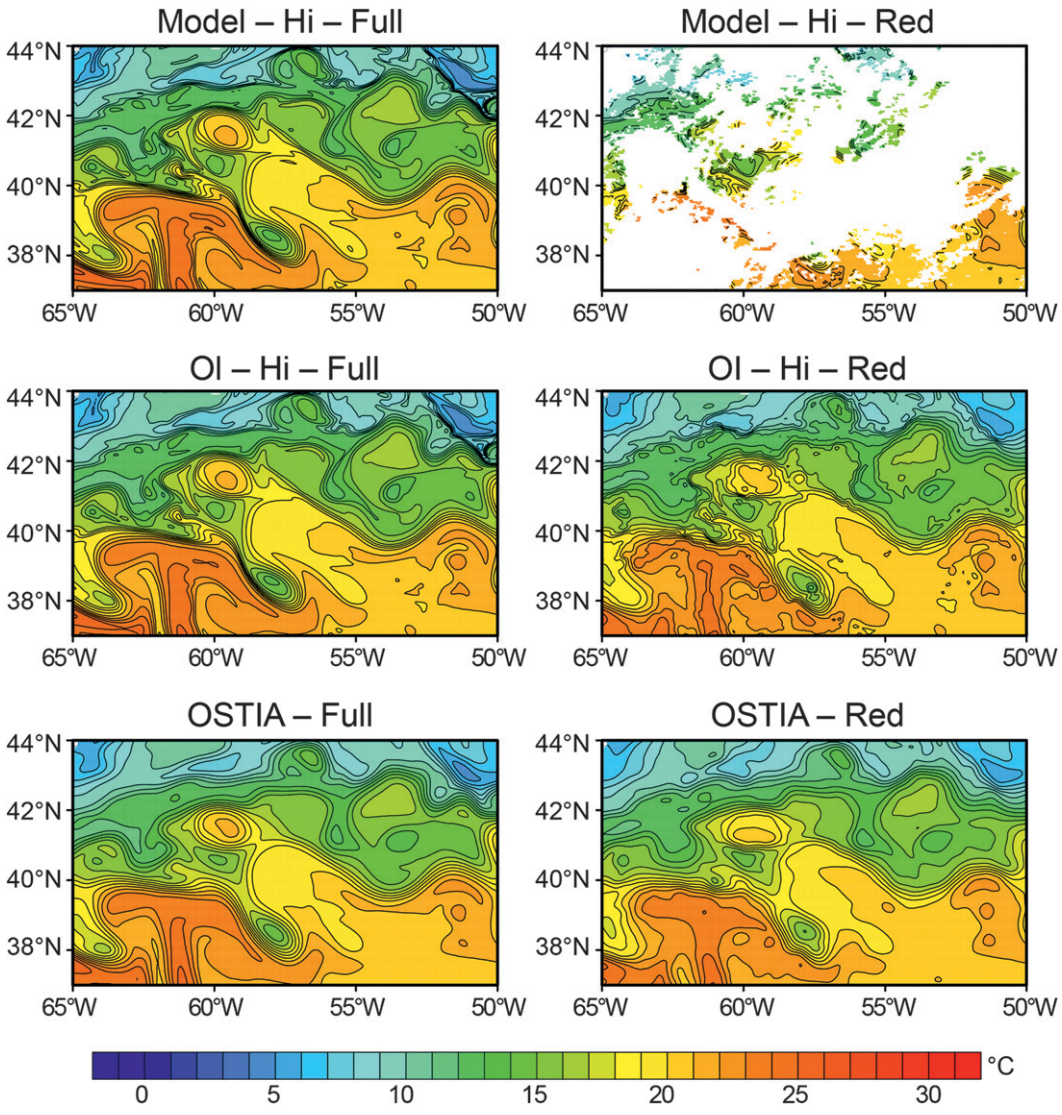

FIG. 4. SST model data and SST analyses for 1 Jul 1993 for the Gulf Stream. The (left) full and (right) reduced model data and analyses; the (top) high-resolution model data, (middle) second high-resolution stage of the two-stage OI, and (bottom) OSTIA analyses of the full model data. The color scale is in degrees Celsius. See Table 1 for further details.

Stream. Differences among the four analyses are relatively small, although it can be seen that OSTIA is slightly smoother than the OI.

a. Autospecta

All spectra were computed daily along the zonal line at the latitudinal center of each box, 40.5° and 25°N for the Gulf Stream and the Sargasso Sea regions, respectively. The daily spectra were band averaged to obtain 30 degrees of freedom and were further averaged over each month (January and July 1993).

The autospectra for July 1993 for the Gulf Stream and the Sargasso Sea regions are shown in Fig. 6. The black curve labeled Model-Hi-Full is the assumed high-resolution "truth." The peak variance at the lowest wavenumber is roughly an order of magnitude larger for the Gulf Stream region than for the Sargasso Sea. The spectra gradually drop off to roughly 10^{-4} °C cycle^{-1} km^{-1} at wavenumbers higher than about 0.08 cycles km^{-1} ($\lambda = 12.5$ km). The curves tend to flatten beyond 0.08 because of the interpolation from the original model grid to the Pathfinder grid beyond the Nyquist wavenumbers of the model grid.

Five additional spectra are also shown in each panel of Fig. 6. The lowest curve (magenta) in each panel is the OI first-stage low-resolution analysis based on the reduced data and interpolated to the high-resolution grid (labeled as OI-Lo-Red). The interpolation, which is

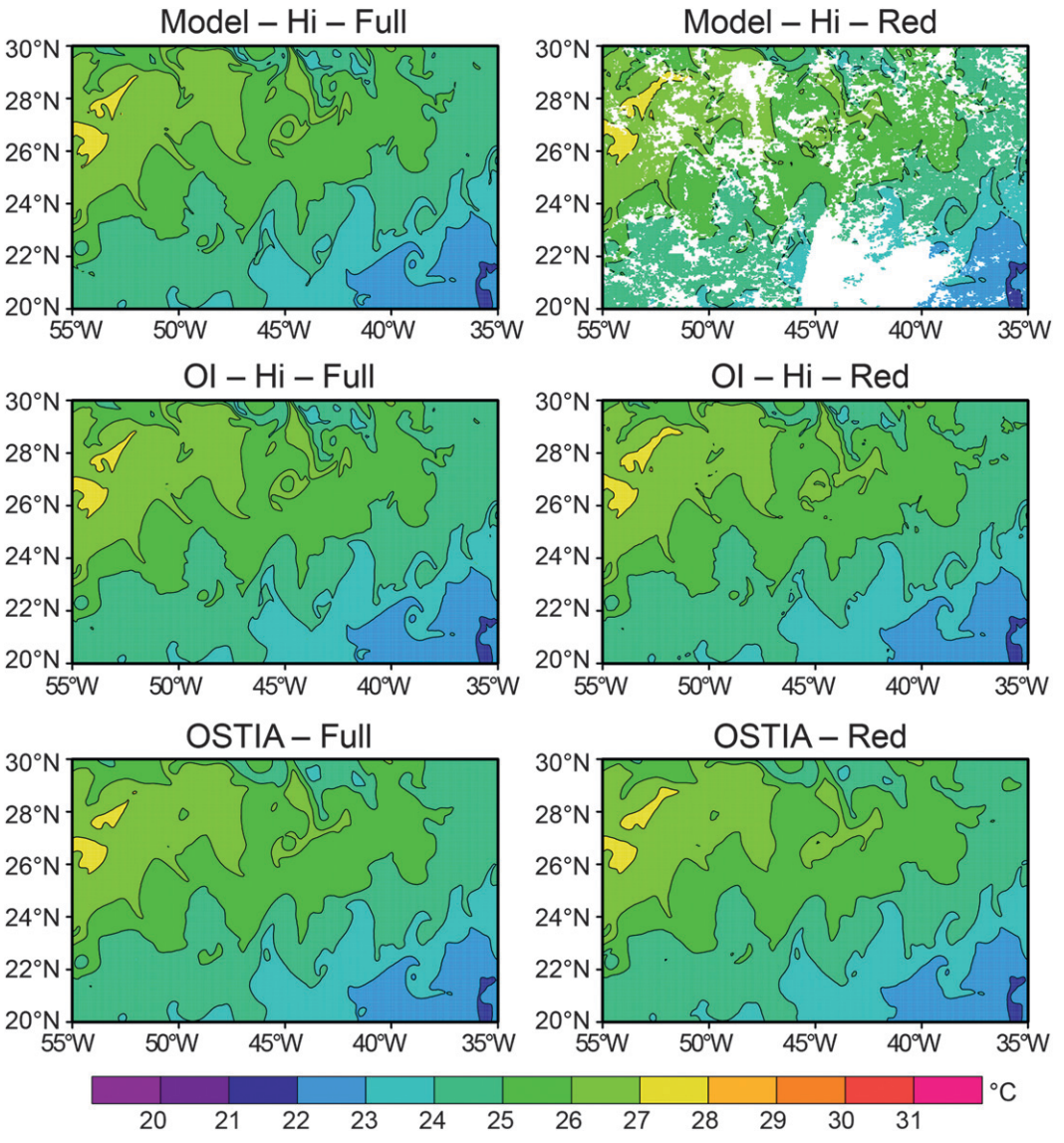

FIG. 5. As in Fig. 4, but SST model data and SST analyses for 1 Jul 1993 for the Sargasso Sea.

required to provide part of the first guess for the high-resolution analysis, uses Gaussian weights to avoid filter sidelobes at high wavenumbers that would arise if bilinear interpolation were used. Because there was no visible difference between the reduced and full low-resolution analyses, OI-Lo-Full is not shown. The OI-Lo-Red curves in both panels match the model spectra at the lowest wavenumbers but quickly drop off to amplitudes below 10^{-5} °C cycle^{-1} km^{-1} at 0.02 cycles km^{-1} (λ = 50 km) and at higher wavenumbers.

The high-resolution OI analyses using the full and reduced datasets (OI-Hi-Full and OI-Hi-Red) are shown by the solid red and green lines, respectively. In the Gulf Stream, OI-Hi-Full has lower spectral power than the model data for wavenumbers greater than 0.04 cycles km^{-1}(λ = 25 km) while OI-Hi-Red actually has higher power. The drop at high wavenumbers in OI-Hi-Full is expected from the smoothing inherent in any analysis procedure. The increase in the spectral density in OI-Hi-Red is surprising. It is an indication of noise generated by filling in missing data and may be expected from the large region of missing high-resolution data in the top-right panel of Fig. 4. In the Sargasso Sea region, OI-Hi-Full is lower than the full model dataset for wavenumbers greater than 0.04 cycles km^{-1}. However, the spectrum for OI-Hi-Red lies almost on top of that of the full model dataset.

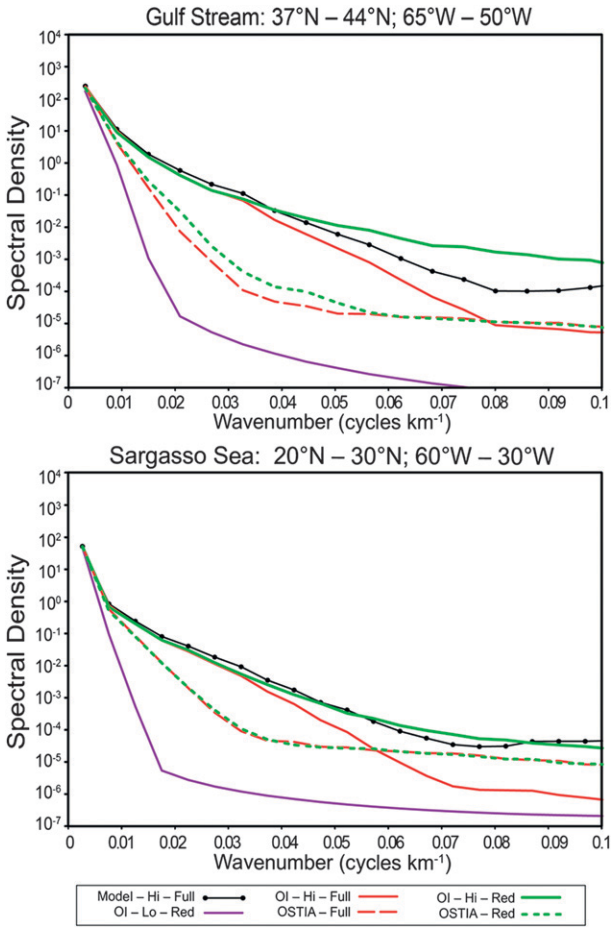

FIG. 6. Autospectra of the high-resolution full model data and the analyses on the high-resolution grid for July 1993 for (top) the Gulf Stream and (bottom) the Sargasso Sea. The original model data, the two-stage OI (high and low resolution), and OSTIA are shown using full and reduced model data. The horizontal axis shows the wavenumber in cycles per kilometer. The vertical axis shows the spectral density ($°C^2$ cycle^{-1} km^{-1}).

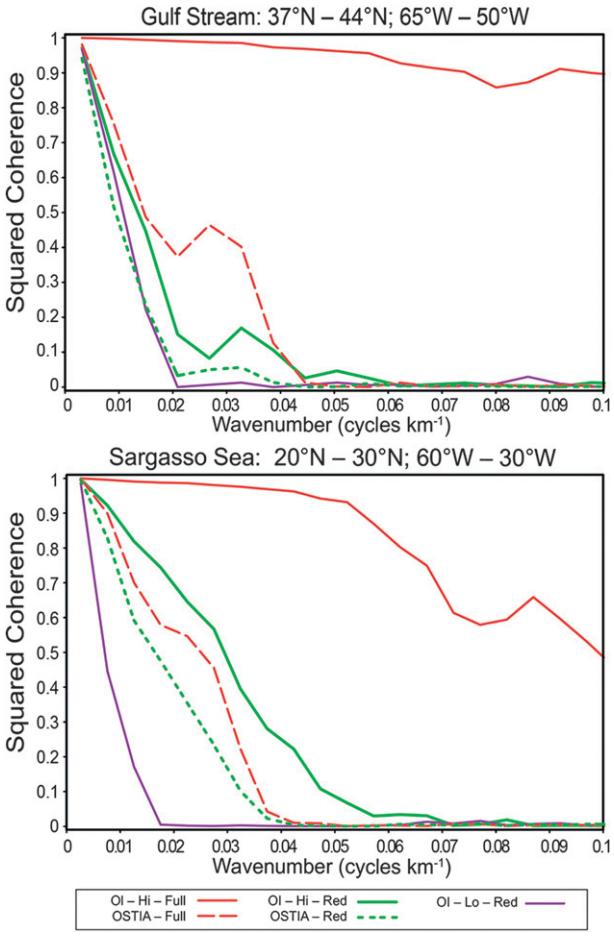

FIG. 7. Squared coherence of the two-stage OI (high and low resolution) and OSTIA analyses for July 1993 for July 1993 for (top) the Gulf Stream and (bottom) the Sargasso Sea with respect to the full model data. The horizontal axis shows the wavenumber in cycles per kilometer. The vertical axis is the squared coherence.

The OSTIA analyses derived from the full and reduced model datasets (OSTIA-Full and OSTIA-Red) are shown by the dashed red and green lines, respectively. In both regions the spectral variances of the OSTIA analyses are much lower than the model dataset, especially in the Gulf Stream. In contrast to the two analyses for the OI procedure, there is little difference between the two curves. This is due to the choices made in the design of the correlation scales and associated variances in order to make the OSTIA system robust to changes in data coverage.

b. Cross-spectra

Since the true SST is not generally known, it is usually not possible to determine how much of the spectral power is noise and how much is signal. In this study, however, the true SST is defined to be the full high-resolution model dataset, Model-Hi-Full. The squared wavenumber coherence γ^2 can then be computed between the respective analyses shown in Fig. 6 and the truth. (The coherence is effectively a spectral decomposition of the squared cross correlation as a function of wavenumber.) The results for the two regions are shown in Fig. 7 for July 1993. In both regions OI-Lo-Red drops quickly from near $\gamma^2 = 1$ at the smallest wavenumbers to roughly 0.5 at 0.01 cycles km^{-1} ($\lambda = 100$ km) and then falls to near zero at wavenumbers above 0.02 cycles km^{-1}. The feature resolution of the low-resolution OI is thus restricted to wavelengths longer than 100 km. For the Gulf Stream, γ^2 for OI-Hi-Full exceeds 0.8 at all wavenumbers. For the Sargasso Sea, γ^2 for OI-Hi-Full is above 0.5 at all wavenumbers and exceeds 0.8 for wavenumbers less than 0.05 cycles km^{-1} ($\lambda = 20$ km).

In contrast, γ^2 for OI-Hi-Red behaves very differently. In the Gulf Stream region, γ^2 for OI-Hi-Red drops very quickly from the lowest wavenumbers and is very similar to the OI-Lo-Red. In the Sargasso Sea region, OI-Hi-Red is higher than it was at the same wavenumbers in the Gulf Stream and drops gradually from 1 to 0.5 at 0.03 cycles km^{-1} ($\lambda = 33$ km). This difference indicates that it is not possible to produce an accurate SST analysis of the small-scale variability (wavelengths less than about 50 km) because of the sparseness of the high-resolution IR data in the Gulf Stream. For wavenumbers higher than about 0.02 cycles km^{-1}, the autospectral variance for OI-Hi-Red is therefore almost all noise in the Gulf Stream (Fig. 6). Small-scale variability is better produced in the Sargasso Sea because the cloud cover there is lower than it is in the Gulf Stream. Even there, however, the spectral power at wavelengths shorter than 30 km is purely noise.

For OSTIA-Full, γ^2 is much lower than the OI-Hi-Full and only slightly better than for the OI-Hi-Red in the Gulf Steam for wavenumbers between 0.02 and 0.04 cycles km^{-1}. Also γ^2 for OSTIA-Red is very similar to the low-resolution OI. In the Sargasso Sea γ^2 for OSTIA-Full is lower than for OI-Hi-Red and only slightly lower than for OSTIA-Full. The smoothing in OSTIA reduces both the signal and the noise, as expected from Fig. 6.

c. *Daily scatterplots*

The results in both Figs. 6 and 7 show that reduced high-resolution IR coverage primarily from cloud cover strongly impacts the squared coherence between OI-Hi-Red and the full model data. Furthermore, the smoothing in the OSTIA analysis procedure causes the squared coherence for both versions of OSTIA with the full model data to be similar to each other and much closer to the OI-Hi-Red than the OI-Hi-Full. These results are all based on monthly averages of daily averaged auto- and cross-spectra. To retain approximately the same number of degrees of freedom in the daily spectra as shown in Figs. 6 and 7 for the monthly averages, the spectra were computed for each day for 31 contiguous zonal lines (including the line used in the monthly version) and then averaged. The resulting daily coherences were further averaged between wavenumbers of 0.02 and 0.04 cycles km^{-1} (λ between 50 and 20 km), which is the region that is most sensitive to changing analyses as shown in Fig. 7. The resulting band-averaged coherences for each day were then compared with the daily fractional coverage computed for each day over the 31 contiguous zonal lines used for the daily spectral averages. The fractional coverage is defined as the number of ocean high-resolution grid boxes with reduced data divided by the total number of ocean high-resolution grid boxes.

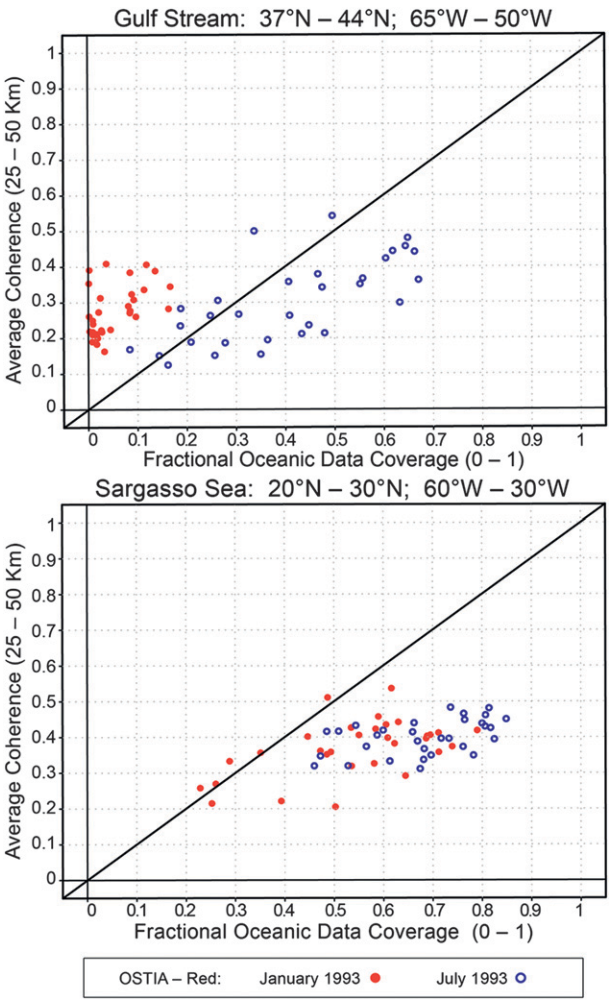

FIG. 8. Scatterplot of the reduced OSTIA daily coherence vs the fractional ocean coverage for January and July 1993 for (top) the Gulf Stream and (bottom) the Sargasso Sea. The daily coherence is computed daily for the reduced OSTIA with respect to the full high-resolution model dataset without added noise and averaged between 25 and 50 km. The fractional coverage is the number of ocean high-resolution grid boxes with reduced data divided by the total number of ocean high-resolution grid boxes.

Scatterplots of coherence versus fractional coverage computed for both months (January and July 1993) using the OSTIA-Red analysis are shown for the two regions in Fig. 8. Coherence was chosen over squared coherence to better spread the values along the coherence range of 0–1. The results for the Sargasso Sea (Fig. 8, bottom) show a rough linear relationship between coherence and fractional coverage, although with only weak sensitivity to data coverage (i.e., small slope). As expected, coherence increases with increasing factional coverage. The differences between January

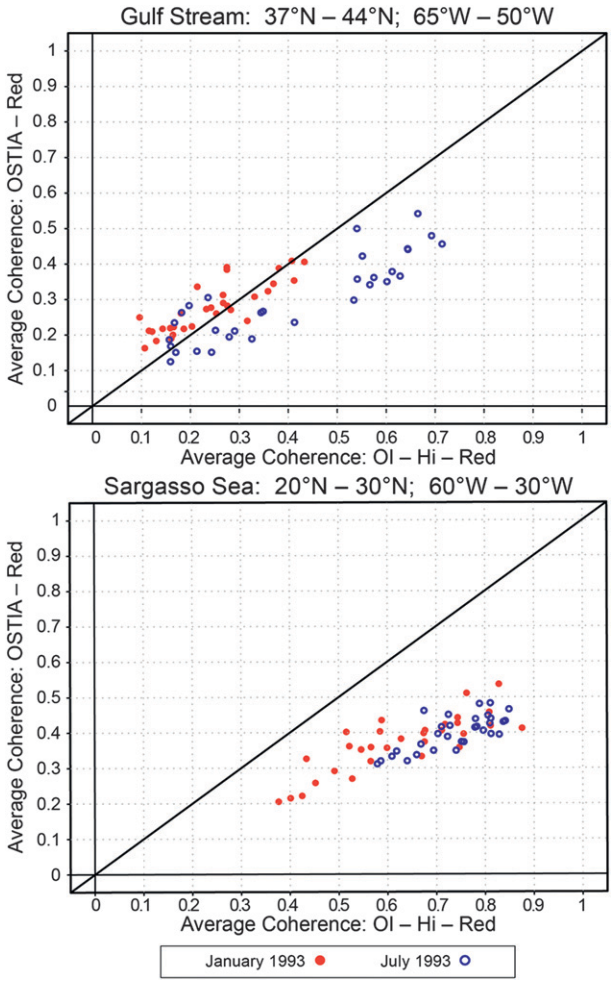

FIG. 9. Scatterplot of the average daily coherence of the high-resolution reduced OI and reduced OSTIA without noise for January and July 1993 for (top) the Gulf Stream and (bottom) the Sargasso Sea. The coherence is computed daily with respect to the full high-resolution model dataset without noise and averaged between 25 and 50 km.

and July are small, except that there were more days in January with relative low coverage and associated low coherence.

The results for the Gulf Stream are very different. The July values of coherence versus fractional coverage are roughly proportional although the magnitudes of both are mostly lower than for the Sargasso Sea. The coherences in January are uniformly low, below 0.42, with factional coverage never exceeding 0.2. The lack of coverage due to cloud cover thus results in little high-resolution signal in the Gulf Stream in January.

The daily OSTIA-Red coherences are compared with the daily OI-Hi-Red coherences in Fig. 9 for January and July 1993. For the Gulf Stream in winter, the OI and OSTIA coherences are similar and fairly low, falling close to the line with a slope of 1. In July, however, the OSTIA Gulf Stream coherences are systematically lower than the OI by approximately a factor of 2. For the Sargasso Sea, the OSTIA coherences are lower than the OI for both months. The higher coherence for the OI means it is more accurate than OSTIA regardless of the fact that it adds more noise (Fig. 6).

5. Generation of simulated measurement noise

The results in the previous section showed that generation of small-scale noise in an SST analysis occurs very easily in data-sparse regions. The calculations for the SST analyses using reduced data have thus far been based on simulated satellite measurements obtained from model SST fields without errors. The small-scale noise in Figs. 4–9 is therefore attributable entirely to sampling errors because of missing data primarily because of cloud contamination. Actual satellite measurements are subject also to measurement errors. The results presented in section 4 are thus optimistic. The influence of measurement errors on the accuracy of SST analyses is investigated in this section by adding simulated noise to the model SST values.

Realistic simulation of measurement errors requires knowledge of the mean value, variance, and spatial and temporal scales of the variability of the errors. Ideally, the measurement errors should be specified spectrally as a function of wavenumber and frequency. Quantitative information about the wavenumber–frequency characteristics of measurement errors does not appear to exist. Moreover, the error characteristics differ geographically for a given satellite instrument. Furthermore, the measurement errors and the corrections obtained from the various algorithms sometimes differ substantially. Finally, the measurement error characteristics also differ in general for each satellite instrument. It is therefore not practical to conduct a comprehensive analysis of the effects of measurement errors. The error analysis presented below in section 6, based on errors generated as described in this section, nonetheless provides useful insight into the effects of measurement errors on an SST analysis.

Qualitatively, it is known that SST measurement errors consist of three primary contributions. There are persistent large-scale errors that manifest as an overall bias for any particular region. For IR high-resolution satellite observations, these large-scale errors are primarily attributable to regional–seasonal-scale climatological structure in the atmosphere, particularly the vertical distribution of water vapor (Merchant et al. 2006). A recurrent example of regional bias of this type is SST retrieved in the equatorial Atlantic (Merchant

et al. 2009). Large-scale biases are also introduced during periods of elevated stratospheric aerosol loading following major volcanic eruptions, with the year following the 1991 eruption of Mt. Pinatubo being the most significant example (e.g., Reynolds 1993).

Large-scale biases can have significant magnitudes and sometimes persist for long periods of time. Most SST analysis procedures address these bias errors by imposing large-scale adjustments to the satellite measurements based on a standard reference such as in situ data or a standard satellite instrument. While there are surely shortcomings in the various bias removal techniques, especially at times and locations were no in situ observations exist, we have not attempted to represent this contribution to satellite measurement errors for this study.

Another contribution to the total measurement error for a satellite sensor will be referred to as ''synoptic-scale noise.'' This component is associated with atmospheric systems that have spatial scales ranging from about 100 to 1000 km and is attributable primarily to imperfect corrections because of variability in the vertical distribution of water vapor within these atmospheric systems. In some regions, variable loading of desert dust aerosol in different air masses also contributes errors on these scales. This component of error was simulated for this study by generating Gaussian-distributed random noise on both the low-resolution 1/4° and high-resolution 0.043 95° latitude–longitude data grids and then smoothing this random noise with isotropic Gaussian weighting with an *e*-folding scale of 200 km. The Gaussian weighted fields (i.e., synoptic-scale noise) were computed for two time periods: 30 November 1992–1 February 1993 and 31 May–1 August 1993. These time periods begin one day earlier and end one day later than the two analysis periods so that the 3-day averaging could be carried out without edge-effect problems.

The GHRSST Single Sensor Error Statistics (SSES) standard deviations provided by the data producers (https://www.ghrsst.org/ghrsst-science/science-team-groups/stval-wg/sses-single-sensor-error-statistics/) were used to estimate the low- and high-resolution standard deviations for AMSR and AVHRR as 0.7° and 0.4°C, respectively. For the two time periods given above, the simulated synoptic-scale noise was normalized to 0.7° and 0.4°C at each low- and high-resolution grid point, respectively. (Note that this normalization does not alter the spatial correlations because the normalization is over a 2-month period. Thus, there is one adjustment over the entire time period, not an adjustment for each day.) An example of this simulated synoptic-scale noise is shown in the top panel of Fig. 10.

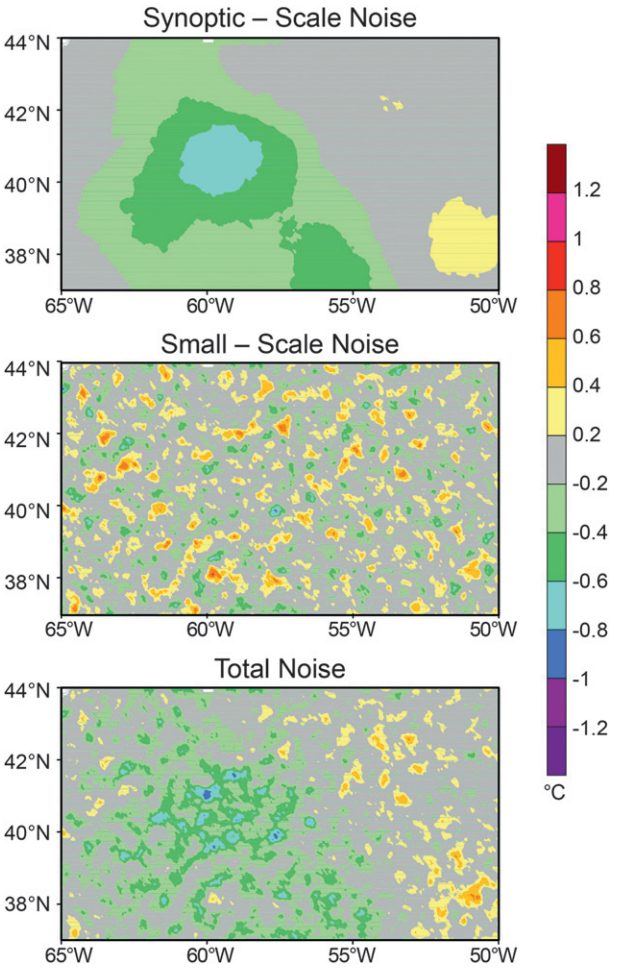

FIG. 10. Noise fields for 1 Jul 1993 for the Gulf Stream: (top) synoptic-scale noise; (middle) small-scale noise; and (bottom) total noise, which is the average of the synoptic- and small-scale noise. The color scale is in degrees Celsius.

The third component of measurement error consists of small-scale noise. This noise derives largely from errors in observed radiances from thermal noise in the detectors that propagate through the retrieval process. Depending on the instrument design, this noise can be random from pixel to pixel and well represented by a Gaussian distribution. Other errors arise occasionally at small scales from other sources, such as imperfect screening for cloud contamination, and these can affect several adjacent pixels within images. The full small-scale component of measurement error was simulated by extracting the higher wavenumbers from the synoptic-scale noise: that is, the ''imperfections'' of the filtering used to generate the synoptic-scale noise. This was done using the synoptic-scale noise field on both the low- and high-resolution grids and then spatially high-pass filtering the fields to attenuate variability with scales larger than

1.5° in latitude and longitude (corresponding to about 166 km at the equator). The high-passed fields (i.e., small-scale noise) were computed for two time periods given above and renormalized to 0.7° and 0.4°C at each low- and high-resolution grid point, respectively. An example of this simulated small-scale measurement error is shown in the middle panel of Fig. 10.

The total measurement errors (i.e., total noise) for this study were defined to be the superposition of the simulated synoptic-scale and small-scale measurement errors described above. An example is shown in the bottom panel of Fig. 10.

The wavenumber characteristics of the two components of synoptic- and small-scale noise and the total noise are shown by the three colored lines in both panels of Fig. 11 for the high-resolution grid. For reference, the black lines in the top and bottom panels correspond to the wavenumber spectrum of the model SST for the Gulf Stream and Sargasso Sea, respectively. The wavenumber at which the spectrum of the total noise (black line) crosses the spectrum of the model SST or signal (blue line) effectively defines a resolution limitation of the measurements. Features in the SST field at wavenumbers for which the spectral variance is smaller than that of the measurement errors are ''lost in the noise'' and are therefore indistinguishable from noise. This crossover point is about 0.04 cycles km^{-1} for the Gulf Stream region and about 0.01 cycles km^{-1} for the Sargasso Sea. These correspond to wavelength scales of about 25 and 100 km, respectively. These wavelength scales are analogous to Gaussian features with *e*-folding scales of about 6 and 24 km, respectively (see the analysis in appendix A.3 of Chelton et al. 2011). The ability to detect small-scale variability in the Gulf Stream occurs because of the higher power at all wavenumbers and hence higher signal-to-noise ratios compared with the Sargasso Sea over a wider range of wavenumbers.

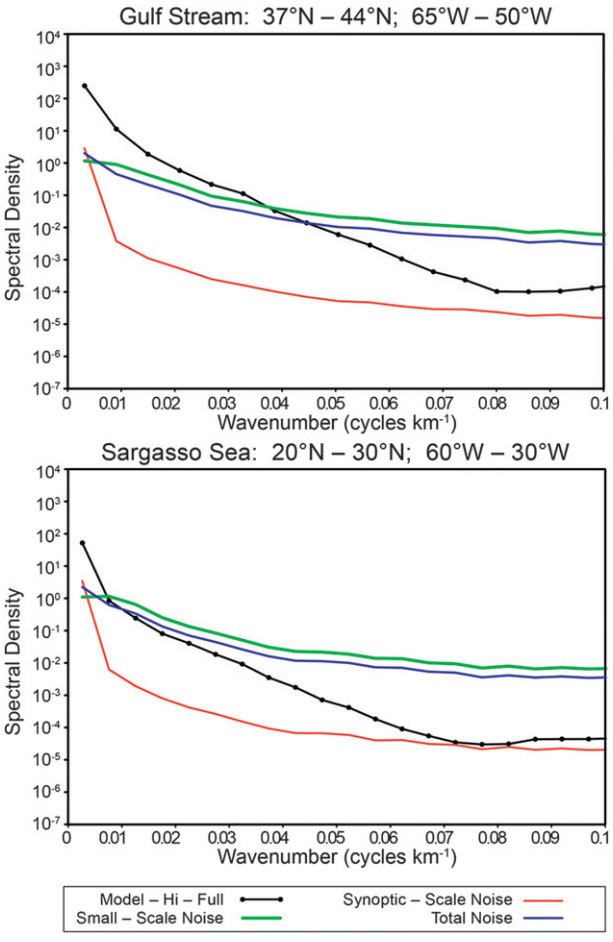

FIG. 11. Autospectra of the noise for July 1993 for (top) the Gulf Stream and (bottom) the Sargasso Sea: the original high-resolution model data, synoptic-scale noise, small scale-noise, and the total noise (average of the synoptic and small-scale noise). The variance is normalized for all noise fields (see text). The horizontal axis shows the wavenumber in cycles per kilometer. The vertical axis shows the spectral density ($°C^2$ $cycle^{-1}$ km^{-1}).

6. Results with simulated measurement noise

The simulated measurement errors were added to the low- and high-resolution model datasets using the full and reduced data coverage. The analyses were then rerun with the model+noise datasets for the two periods of interest: 1 December 1992–31 January 1993 and 1 June through 31 July 1993. The SST fields with the added noise are referred to by the same names used in the earlier sections but with the added suffix ''+Noise'' (see also Table 1). An example of the resulting daily SST fields with the added noise is shown in Fig. 12 for the Gulf Stream region for 1 July 1993. Compared with the noise-free results (Fig. 4), the added noise shows up as small-scale wiggles in the shading and contours. Only the small-scale noise is apparent because the large dynamic range of the SST signal in the Gulf Stream dominates the added noise except at the smallest scales (highest wavenumber). This result can be anticipated from the top panel of Fig. 11. As the OSTIA analysis is smoother than the OI analysis, there is little difference between the two OSTIA versions with and without noise, either in the full or reduced-data cases. The OSTIA smoothing thus attenuates most of the small-scale noise.

The SST fields with the added noise are shown in Fig. 13 for most of the Sargasso Sea region. When compared with the noise-free results, Fig. 5, the added noise is more apparent on all scales than in the Gulf Stream examples. This result is consistent with the spectra in the lower panel of Fig. 11. Note, in particular, the small cold-core eddy near 27°N and 45°W that is present in

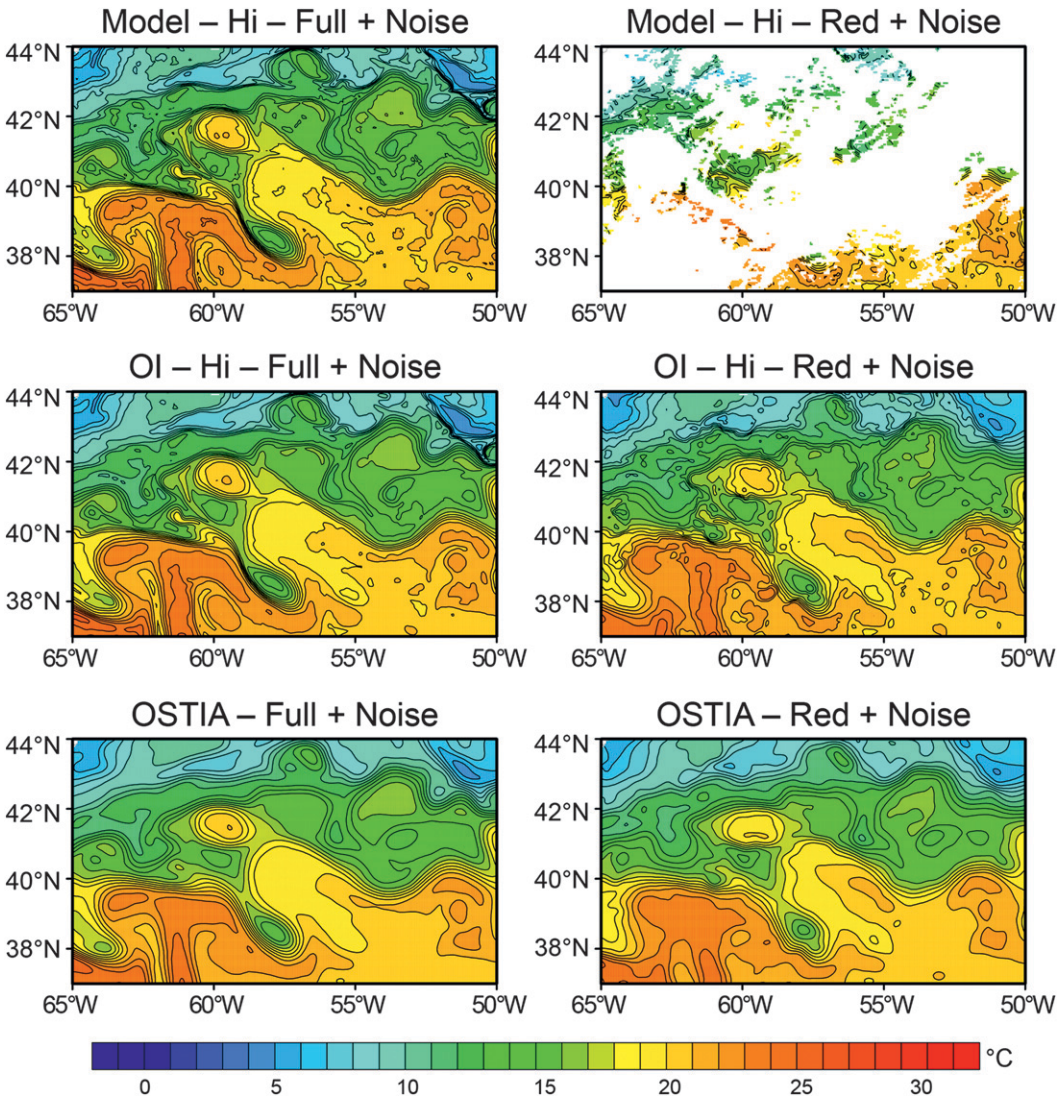

FIG. 12. As in Fig. 4, but SST model data with added noise and SST analyses with noise for 1 Jul 1993 for the Gulf Stream.

most of the noise-free maps of Fig. 5 but is not evident in Fig. 13. The added noise overwhelms this eddy signal. The noise is much less apparent in the OSTIA SST fields than in the OI fields because of the stronger OSTIA smoothing.

a. Autospectra

The autospectra for the SST analyses with simulated measurement errors for the Gulf Stream and Sargasso Sea are shown for July 1993 in Fig. 14. The spectrum from the high-resolution model without noise is the standard of comparison and is included as the black line. These results should be compared with the autospectra computed from the noise-free SST analyses in Fig. 6. The added noise increases the spectral variance of the OI-Hi-Red+Noise compared to OI-Hi-Red in both the Gulf Stream and Sargasso Sea with a larger increase in the Sargasso Sea as expected. The OI-Hi-Full+Noise spectral variance shows a strong increase over OI-Hi-Full only in the Sargasso Sea, where the signal power is smaller by about an order of magnitude compared with the Gulf Stream. The change in spectral variance is very small for OSTIA compared to the OI, again showing that the noise is suppressed (along with the signal at high wavenumbers) by the stronger smoothing in the OSTIA analysis procedure.

b. Cross-spectra

The squared coherences for the SST analyses with the added noise are shown for July 1993 in Fig. 15. The squared

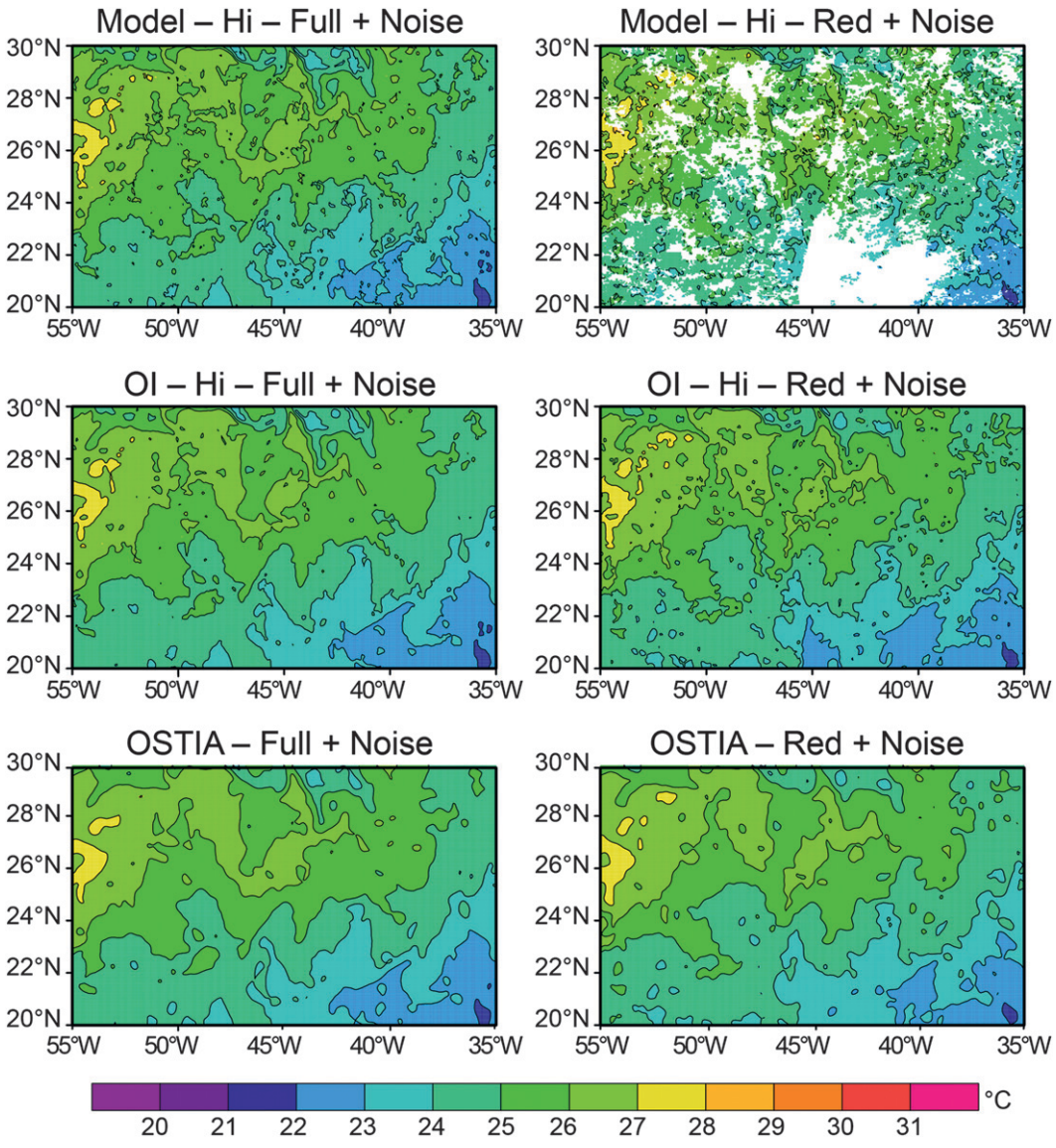

FIG. 13. As in Fig. 5, but SST model data with added noise and SST analyses with added noise for 1 Jul 1993 for the Sargasso Sea.

coherences are calculated with respect to the high-resolution full model data without noise. These results can be compared with the squared coherences for the noise-free SST fields in Fig. 7. For both the Gulf Stream and the Sargasso Sea, the OI-Hi-Full+Noise analysis shows a large decrease in the squared coherence compared with the OI-Hi-Full case. As expected from the autospectral results in Fig. 14, the decrease is greater for the Sargasso Sea than for the Gulf Stream at all but the lowest wavenumbers. The added noise has relatively little impact on the OI-Hi-Red for the Gulf Stream region because the signal spectral power is much higher at all wavenumbers in this region. The changes in the OSTIA analyses with noise from their respective noise-free cases are modest.

7. Data coverage and high-resolution signal

The results presented in sections 4 and 6 clearly show that the ability of any analysis procedure to reproduce small-scale signals accurately depends on the data coverage as well as the filtering properties of the analysis procedures. This is quantified in Figs. 8 and 9 for the noise-free measurements of SST. These figures suggest that it may be possible to assess how often accurate small-scale features can be resolved based on data coverage.

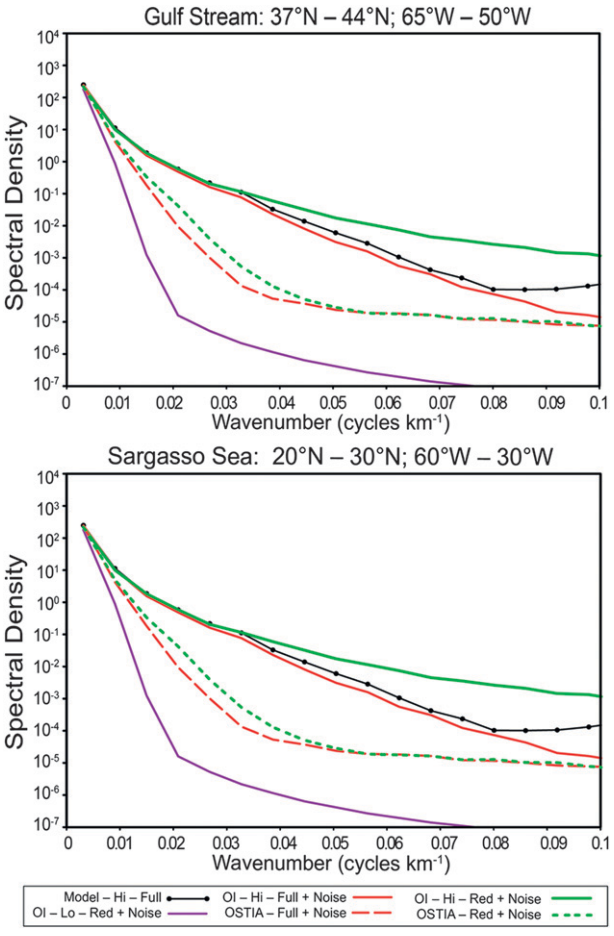

FIG. 14. As in Fig. 6, but autospectra of the high-resolution model data without noise and the two-stage OI and OSTIA analyses with added noise for July 1993 for (top) the Gulf Stream and (bottom) the Sargasso Sea.

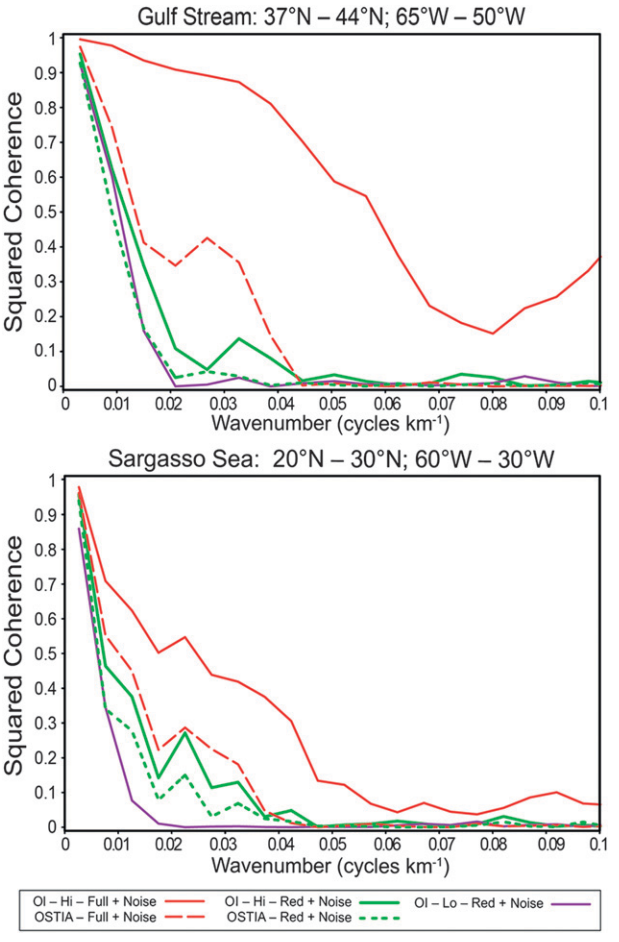

FIG. 15. As in Fig. 7, but squared coherence of the two-stage OI and OSTIA with noise for July 1993 for (top) the Gulf Stream and (bottom) the Sargasso Sea with respect to the high-resolution full model data without noise.

The presence of measurement errors, which have considerable variance at small scales (see Figs. 11, 12), complicates matters. The challenge is to distinguish signal from noise at small scales.

To gain insight into the sampling issues, the fractional coverage was computed at daily intervals. The fractional coverage is defined to be the daily ratio of the number of high-resolution ocean grid boxes with data compared to the total number of ocean grid boxes on a 1° grid. The number of days with coverage above 50% for January and July 2004 are shown in Fig. 16. (The year is labeled as 2004 as that is the year of the actual coverage used for the 1993 model data; see section 3). In the Sargasso Sea region, a high-resolution analysis is possible by this metric for roughly 18 days in January and for more than 27 days in July. For the Gulf Stream region, a high-resolution analysis is never possible in January and is only possible for roughly 15 days during July. This difference between January and July is large because of the strong seasonal cycle in cloud cover in the Gulf Stream region. The figure shows similarly large seasonal differences in other regions: for example, the northern Indian Ocean (better coverage in winter) and the Mediterranean Sea (better coverage in summer).

Figure 16 provides a qualitative assessment of where products claiming to resolve small-scale variability may actually be capable of achieving such claims. Attempts to resolve small-scale features in regions of persistent cloud cover will either produce small-scale noise (e.g., the OI analysis considered here) or will accurately reproduce only large-scale features because of the smoothing in the analysis procedure.

8. Conclusions

There has been a trend over the past decade to produce SST analyses at higher and higher grid resolutions.

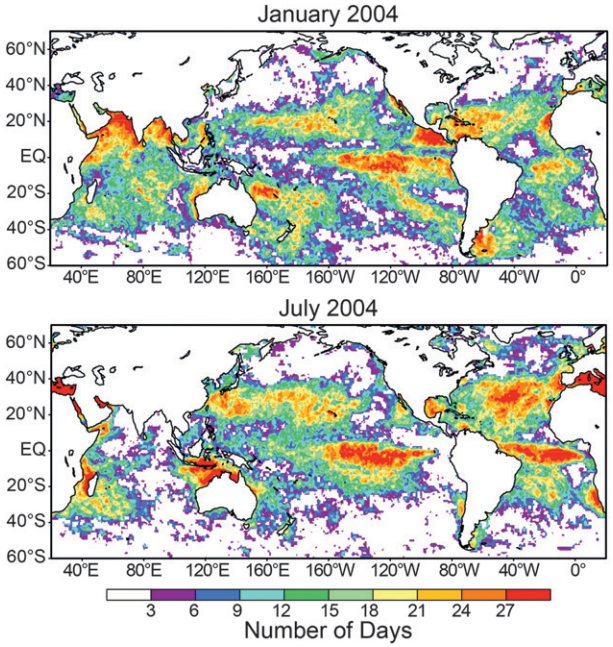

FIG. 16. Number of days in (top) January and (bottom) July 2004 where the fractional coverage of the number of ocean grid points with data to the total exceeds 50%. The fractional coverage was computed on a 1° spatial grid. The coverage for 2004 was used for the model year 1993.

As high-resolution SST IR satellite data coverage is limited primarily because of cloud cover, it is unclear where and when small-scale SST features in an analysis are accurately reproduced and where they are artifacts of the analysis procedure. In an effort to answer this question, SST fields from the ECCO2 ocean model were used here as ''true'' SST data. MW satellite data were simulated by smoothing over 50 km and interpolating to a ¼° spatial grid to mimic AMSR data. IR satellite measurements were simulated by interpolating the model SSTs to a 0.043 95° grid to mimic Pathfinder version 5 AVHRR data. Both sets of simulated observations were then subsampled to represent actual MW and IR satellite data coverage in precipitation-free and cloud-free regions, respectively.

Other IR sources of data are available besides the AVHRR data considered in this study. The availability of these other IR datasets was simulated somewhat roughly here by including AVHRR data from three consecutive days in each daily SST analysis. While inclusion of multiple IR data sources can improve the accuracy and resolution of an SST analysis, sparse sampling of high-resolution data will still occur in regions of persistent cloud cover (e.g., the Gulf Stream in winter).

To investigate seasonal dependence of the sampling, results were examined for two months (January and July 1993). Two SST analysis procedures were considered: a two-stage low-resolution (¼°) and high-resolution (~1/23°) OI analysis and the (1/20°) OSTIA analysis. Auto- and cross-spectra were examined for two regions: the Gulf Stream and Sargasso Sea. Because the true input data were known in these simulations based on model SST, it was possible to determine how the analyses based on both the full and reduced (subsampled) low- and high-resolution datasets were impacted by sampling errors. The impacts of measurement errors were also considered by adding realistic noise to the model data.

Although the ECCO2 model has a grid spacing of 1/16° or 6.9 km at the equator, the results of this study show that the resolution limitations of the two-stage OI and OSTIA analysis procedures considered here (and many other SST analysis procedures) are much coarser than the grid size limitation of ECCO2. It is important to also point out that diurnal variability is not included in the ECCO2 model simulations. In regions where the diurnal variability is significant, this is another potential source of sampling error in SST analyses. Because of the lack of diurnal forcing and the 1/16° grid resolution limitation of the ECCO2 model, the simulated true SST fields generated for this study must be interpreted as providing a lower bound on the effects of sampling errors in the accuracy and resolution limitations of SST analyses. In reality, the sampling issues are likely somewhat more severe.

The results for the noise-free analyses showed that the low-resolution first-stage of the two-stage OI is almost identical for both the full and reduced low-resolution model datasets. This is understandable because MW satellite SST coverage is only reduced by precipitation and not cloud cover since clouds are essentially transparent at MW frequencies. Differences between the full and reduced low-resolution model OI also occur in regions without MW data: for example, in regions within 75 km of the coast where MW retrievals are not possible.

The autospectra computed from the SST analyses with full data coverage showed that the spectral variance at high wavenumbers (small wavelengths) was lower than that of the input data. These results can be expected as analyses are designed to fill and smooth the input data and thus act like a low-pass filter. The filtering properties of an analysis procedure are determined by the details of the procedures and in general differ for each analysis products. For the two products considered in this study, the smoothing is much stronger in OSTIA than in the two-stage OI.

The autospectra computed from SST analyses with reduced data showed that the spectral variance at high wavenumbers is higher than the same SST analyses

with the full data and can, in fact, be larger than the original input data. In such cases, the analysis procedure adds small-scale noise when filling in the missing data. Cross-spectral analysis with respect to the input data confirmed that high-resolution analysis procedures add noise to an analysis when measurements are sparse.

Comparisons of the SST fields produced by the high-resolution OI (second stage) procedure and OSTIA showed important differences. Because of the different filtering properties of the two analysis procedures noted above, the OI tends to reproduce small-scale variability more accurately than OSTIA when the high-resolution data coverage is good. However, the OI tends to produce more small-scale noise than OSTIA when the high-resolution data coverage is poor. This result suggests that the analysis noise variance should be kept more constant as coverage of high-resolution data varies. Keeping the noise levels near constant of course means that feature resolution must be sacrificed during periods of sparse high-resolution data.

The spectral analysis was repeated with simulated measurement errors with a constant standard deviation added as noise to the SST fields. For both the Gulf Stream and Sargasso Sea, the analyses with added noise showed that the wavenumber squared coherence decreases as expected compared to the noise-free analyses. Because the Gulf Stream has an order of magnitude higher power at all wavenumbers compared with the Sargasso Sea, the effects of added noise are less apparent in the Gulf Stream. Correspondingly, the squared coherence decrease due to noise was less for the Gulf Stream than the Sargasso Sea. The decrease of coherence when adding noise is less for OSTIA than for the OI because the OSTIA analysis is inherently smoother as noted above and the coherence with respect to the true SST is already lower for OSTIA than for the OI.

Maps of the numbers of days per month with 50% or greater coverage of the high-resolution data provide qualitative guidance for where and when reliable estimates of small-scale signals may be possible in an SST analysis. Depending on the smoothing inherent in an analysis procedure, small-scale signals may be attenuated even when clear skies are present. During the 2004 observational period considered in this study, high-resolution coverage in the Gulf Stream region was very poor in January and only possible about half the time in July. The Sargasso Sea region had high-resolution data coverage during much of January and most of July because cloud coverage there is much less than in the Gulf Stream. As shown in section 6, however, the ability to detect small-scale features in the Sargasso Sea is limited by measurement errors.

This study shows that the two-stage OI procedure should be modified to increase the attenuation of small-scale features in regions and periods where high-resolution data are sparse or missing. This is likely true for other high-resolution SST products. The results presented here show that the use of the model full and reduced data as truth can be helpful in fine tuning an analysis procedure to accomplish this goal.

It is clear from the results presented here that users must be skeptical of feature resolution claims of high-resolution SST products. The ability of an SST analysis product to represent small-scale features in the SST field varies geographically as well as temporally. It would be helpful to users if producers of high-resolution analyses include monthly coverage maps of their input data and distribute them along with their SST products. Coverage maps such as those shown in Fig. 16 are helpful for assessing when and where small-scale features in an SST analysis can be considered reliable.

The best way forward may be the development of an improved analysis with variable resolution and with noise variance that is more constant in space and time. This new analysis would have high resolution of small-scale features in regions of good coverage and lower resolution in regions of poor coverage.[1] It is important that users know when and where small-scale features are detectible. All analyses must therefore include information about data coverage or some metric for small-scale resolution capability.

Acknowledgments. This work was funded in part by NOAA's Climate Data Record Program, managed by the National Climatic Data Center. We are grateful to NCDC and the NOAA/Climate Program Office, which provided partial support for this work. The graphics for all of the figures were computed using the Grid Analysis and Display System (GrADS; available online at http://grads.iges.org/grads) and finalized by the NCDC Graphics Department. Both DCB and RWR were partially supported by NASA Grant NS214A funded through Oregon State University. The research of JRJ and MJM leading to these results has received funding from the European Community's Seventh Framework Programme FP7/2007-2013 under Grant Agreement 283367 (MyOcean 2). DM performed this work at the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA.

[1] It is planned that the two-stage OI analysis will be produced as an "experimental" product.

REFERENCES

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.

——, M. G. Schlax, and R. M. Samelson, 2011: Global observations of nonlinear mesoscale eddies. *Prog. Oceanogr.,* **91,** 167–216, doi:10.1016/j.pocean.2011.01.002.

Conkright, M. E., R. A. Locarnini, H. E. Garcia, T. D. O'Brien, T. P. Boyer, C. Stephens, and J. I. Antonov, 2002: *World Ocean Atlas 2001:* Objective analyses, data statistics, and figures. National Oceanographic Data Center Internal Rep. 17, 21 pp.

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

——, M. Martin, J. D. Stark, J. Roberts-Jones, and E. Fiedler, 2012: The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system. *Remote Sens. Environ.,* **116,** 140–158, doi:10.1016/j.rse.2010.10.017.

Hill, C., D. Menemenlis, B. Ciotti, and C. Henze, 2007: Investigating solution convergence in a global ocean model using a 2048-processor cluster of distributed shared memory machines. *Sci. Program.,* **15,** 107–115.

Kilpatrick, K. A., G. P. Podestá, 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), 9179–9198.

Kistler, R., and Coauthors, 2001: The NCEP–NCAR 50-Year Reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.,* **82,** 247–267.

Large, W. G., and A. J. G. Nurser, 2001: Ocean surface water mass transformation. *Ocean Circulation and Climate,* G. Siedler, J. Church, and J. Gould, Eds., Academic Press, 317–336.

——, J. C. McWilliams, and S. Doney, 1994: Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization. *Rev. Geophys.,* **32,** 363–403.

Marshall, J., A. Adcroft, C. Hill, L. Perelman, and C. Heisey, 1997: A finite-volume, incompressible Navier–Stokes model for studies of the ocean on parallel computers. *J. Geophys. Res.,* **102** (C3), 5753–5766.

Martin, M., and Coauthors, 2012: Group for High Resolution Sea Surface Temperature (GHRSST) analysis fields inter-comparisons. Part 1: A GHRSST multi-product ensemble (GMPE). *Deep-Sea Res. II,* **77,** 21–30, doi:10.1016/j.dsr2.2012.04.013.

Menemenlis, D., J.-M. Campin, P. Heimbach, C. Hill, T. Lee, A. Nguyen, M. Schodlok, and H. Zhang, 2008: ECCO2: High resolution global ocean and sea ice data synthesis. *Mercator Ocean Quarterly Newsletter,* No. 31, Mercator Ocean, Ramonville Saint-Agne, France, 13–21.

Merchant, C. J., L. A. Horrocks, J. Eyre, and A. G. O'Carroll, 2006: Retrievals of sea surface temperature from infra-red imagery: Origin and form of systematic errors. *Quart. J. Roy. Meteor. Soc.,* **132,** 1205–1223.

——, A. R. Harris, H. Roquet, and P. Le Borgne, 2009: Retrieval characteristics of non-linear sea surface temperature from the Advanced Very High Resolution Radiometer. *Geophys. Res. Lett.,* **36,** L17604, doi:10.1029/2009GL039844.

Reynolds, R. W., 1993: Impact of Mount Pinatubo aerosols on satellite-derived sea surface temperatures. *J. Climate,* **6,** 768–774.

——, and D. B. Chelton, 2010: Comparisons of daily sea surface temperature analyses for 2007–08. *J. Climate,* **23,** 3545–3562.

——, 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.