

Assimilation of Altimeter Wave Measurements into Wavewatch III

Paul A. Wittmann
Fleet Numerical Meteorology and Oceanography Center
7 Grace Hopper Ave, Stop 1, Monterey, CA 94943-5501

James A. Cummings
Oceanography Division
Naval Research Laboratory
7 Grace Hopper Ave, Stop 2, Monterey CA 93943-5502

Abstract. Assimilation of altimeter measured significant wave heights (SWH) into a global implementation of the Wavewatch III model was performed for March 2004, using SWH data obtained from ENVISAT and JASON satellites. The wave model is forced by 3-hourly Navy Operational Global Atmospheric Prediction System (NOGAPS) marine surface winds. A 6-hour time window about the synoptic time is used to select the altimeter SWH data for the assimilation. The satellite measurements are quality controlled and bias corrected before being used in the analysis. An Optimum Interpolation (OI) scheme is used to compute the SWH increment field from the altimeter SWH innovations. The “first guess” 6-hour model forecast directional wave spectra are then corrected by the ratio of the analysis wave height over the first guess wave height. This correction is distributed uniformly over the wave model spectra. Prior to the March 2004 assimilation run, a six-month analysis-only run (no forecast model update) was performed. The SWH innovations from the analysis-only run are used to compute the statistical parameters required in the OI; observation errors, Wavewatch III prediction errors at the 6-hour forecast period, and spatial covariance functions. Observation errors are found to vary with satellite, prediction errors are found to vary with position, and a second-order autoregressive function is found to be an adequate fit to the bin-averaged spatial autocorrelation estimates. Initial testing of the assimilation system shows a decrease in wave model SWH forecast mean and root mean square errors when compared to selected deep-water wave buoys and yet-to-be-assimilated altimeter SWH observations. Spatial correlation analysis of the analysis residuals shows that the analysis is effectively extracting all of the information in the altimeter SWH measurements.

Introduction. The assimilation of radar altimeter wave heights into numerical wave models has progressed over the last 15 years with the deployment of altimeters on a number of satellites orbiting the earth. The significant wave height (SWH) is estimated from the backscatter of the altimeter pulse. The narrow footprint gives high resolution along track, but sparse data coverage between tracks. Two main issues need to be considered: 1) the method of interpolation of the wave height corrections, and 2) the method used to modify the first guess directional wave spectra of the model based on the wave height analysis.

The first attempts to assimilate altimeter measured wave heights in numerical wave models were made by Esteva (1998) and Lionello et al. (1992), using SEASAT and GEOSAT data. Both of these studies used standard optimum interpolation (OI) techniques to create wave height analysis. Esteva scaled the wave model spectra by the ratio of the first guess SWH to the analyzed SWH, while Lionello et al. used a more sophisticated method using the local wind velocity to modify the sea and swell

spectral components. Since that time, SEASAT and GEOSAT have failed, but other altimeter satellites have been launched. Currently, JASON-1, GFO and ENVISAT satellite altimeters provide wave height measurements to a number of operational weather centers (Bidlot and Holt, 1999). Greenslade (2001) looked at the effect of the spectral adjustment method and the error correlation length. She found that the results were more sensitive to the length scale than the choice of spectral adjustment method.

More recent studies have focused on the sensitivity of the wave model to the simultaneous assimilation of data from several altimeters (Skandrani et al., 2003), and the choice of the spatial autocorrelation functions used in the OI method (Greenslade, 2004). Unlike NWP models, wave models are strongly forced by surface winds, so the impact of the assimilation is often diminished over forecast time, particularly in the wind sea portion of the directional wave spectra. However, it has been shown that corrections to the low frequency portion of the spectra retain the corrections for a longer time (Bender and Glowacki, 1996). In

general, these studies have found that assimilation of altimeter data into the operational wave models has a positive effect on the wave model bias in the short term (0-36 hour) forecast.

Wave Model Configuration. The Wavewatch III version 2.22 (Tolman, 1990) configuration used for the assimilation test is identical to that of the Fleet Numerical Meteorology and Oceanography Center (FNMOC) operational global model. The model is run on a 0.5-degree resolution spherical grid, using an ice analysis to mask points under the ice. The model is initialized by the 6-hour forecast, or first guess, spectra from the previous run. The wind forcing time step is 3 hours. The spectral resolution of the wave model is 24 directions (15 deg angular resolution) and 25 frequencies, ranging from 0.42 to 0.04 hertz (Wittmann, 2002).

Assimilation Method. The wave model data assimilation is performed by the Naval Research Laboratory (NRL) Coupled Ocean Data Assimilation (NCODA) system. NCODA is a fully three-dimensional multivariate optimum interpolation system developed as part of the Office of Naval Research (ONR) sponsored Navy coupled modeling project (Cummings, 2003). In this study, NCODA is executed in two-dimensional mode to provide updated SWH fields for the Wave Watch III wave forecast model using a sequential incremental update cycle. The analysis background field, or first guess, is generated from a short-term wave model forecast. In the wave model data assimilation runs described here a six-hour update cycle is used. NCODA computes corrections to the first-guess SWH field using all of the altimeter SWH observations that have become available since the last analysis was made. The forecast model with the new initial conditions is then run forward in time to produce the next forecast.

The optimum interpolation problem is formulated in NCODA as,

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{P}_b \mathbf{H}^T (\mathbf{H} \mathbf{P}_b \mathbf{H}^T + \mathbf{R})^{-1} [\mathbf{y} - \mathbf{H}(\mathbf{x}_b)]$$

where \mathbf{x}_a is the analysis, \mathbf{x}_b is the background, \mathbf{P}_b is the background error covariance, \mathbf{H} is the forward operator, \mathbf{R} is the observation error covariance, and \mathbf{y} is the observation vector. The observation vector contains all of the synoptic SWH observations within the geographic and

time domains of the Wavewatch forecast model grid and update cycle. The forward operator in NCODA is simply a spatial interpolation of the forecast model grid to the observation location performed in two dimensions. Thus, $\mathbf{H} \mathbf{P}_b \mathbf{H}^T$ is approximated directly by the background error covariance between observation locations, and $\mathbf{P}_b \mathbf{H}^T$ directly by the error covariance between observation and grid locations. For the purposes of discussion, the quantity $[\mathbf{y} - \mathbf{H}(\mathbf{x}_b)]$ is referred to as the innovation vector, $[\mathbf{y} - \mathbf{H}(\mathbf{x}_a)]$ is the residual vector, and $\mathbf{x}_a - \mathbf{x}_b$ is the increment (or correction) vector.

Specification of the background and observation error covariances in the analysis is very important. The NCODA background error covariances are separated into a background error variance and a correlation. In two-dimensional mode only the horizontal correlation component needs to be specified. The horizontal correlation is modeled as a second order autoregressive (SOAR) function of the form,

$$C_h = (1 + s_h) \exp(-s_h)$$

where s_h is the horizontal distance between two locations (observations or observation and a grid point). The distance is normalized by the geometric mean of the horizontal correlation length scales prescribed *a priori* at the two locations. NCODA allows the correlation length scales to vary with location, but in the assimilation experiment reported here the SWH error correlation length scale is set to a constant value (223 km). This value was computed using the innovation correlation method (Hollingsworth and Lonnberg, 1986) from a non-assimilative JASON-1 altimeter SWH innovation time series created in a six-month run of the analysis from June through December 2003. Statistical analysis of the innovations is the most common, and the most accurate, technique for estimating observation and forecast error covariances. Fig. 1 shows the bin-averaged autocorrelation estimates as a function of distance, and a non-linear least squares fit of the SOAR model. As can be seen in Fig. 1, a SOAR function accurately models the long positive tail of the estimated correlations. In comparison, the spatial autocorrelation analyses and SOAR models fit to the JASON-1 and ENVISAT altimeter SWH innovations from the March 2004 assimilation run are shown in Fig. 2. The correlation length scale derived from the non-

assimilative SOAR model is almost twice as large as the length scale computed from the assimilation run. However, the functional form of the SOAR models is very similar between the two innovation time series. A longer innovation time series from an assimilation run is needed to determine if these estimated differences in correlation length scales are real.

The background error variances in NCODA (\mathbf{E}_b^2) vary with location and evolve with time. The error variances are computed from a time history of the analyzed increment fields and updated at the end of each update cycle. A climate error growth rate parameterization is used to account for the inherent sampling limitations of the altimeters. In the long-term absence of altimeter SWH observations, the background error variances are slowly restored to climate variability values using a climate decorrelation time scale of ~ 96 hours. The climate decorrelation time scale is calculated from observations and assumes a zero mean SWH climate field. In practice, the background error variances reflect the long-term average prediction error variances of the model forecast at the analysis update time. To initialize the assimilation run the background error variances are computed from the time history of the non-assimilative analyzed increments (Fig. 3). Note that because of the assumption of a zero mean SWH climate field, the background error variances in Fig. 3 computed using the climate error growth scheme are likely to be inflated.

The observation errors and the background errors are assumed to be uncorrelated, and errors associated with observations made at different locations and at different times are also assumed to be uncorrelated. As a result of these assumptions, the observation error covariance matrix \mathbf{R} is set equal to $\mathbf{1} + \mathbf{e}_o^2$ along the diagonal and zero elsewhere. Note that \mathbf{e}_o^2 represents observation error variances that have been normalized by the background error variances interpolated to the observation location ($\mathbf{e}_o^2 = \mathbf{E}_o^2 / \mathbf{E}_b^2$). Observation error variances are computed from the non-assimilative innovation time series using the innovation correlation method. The SOAR correlation function that is fit to the bin-averaged observed covariances is extrapolated to zero distance and the background error variance is computed. The difference between this value and the innovation variance is the observation error variance. The method assumes horizontally uncorrelated observation

errors, and it is only possible to obtain a horizontally homogeneous (domain-averaged) estimate of the background error variance using this method. Observation and background errors for JASON-1 and ENVISAT computed using the innovation correlation method are shown in Table 1 for both the assimilation and non-assimilation control runs of the wave model.

Quality Control and Observation Preprocessing. All altimeter SWH observations are subject to quality control (QC) procedures prior to assimilation. The primary purpose of the QC system is to identify observations that are obviously in error, as well as the more difficult process of identifying measurements that fall within valid and reasonable ranges, but nevertheless are erroneous. The need for quality control is fundamental to any data assimilation system. Accepting erroneous data can cause an incorrect analysis, while rejecting extreme, but valid, data can miss important events. The SWH QC procedures include land/sea boundary checks, shallow water retrieval checks, and background field checks against Wavewatch III model forecast fields using 6-hour prediction error variances. Cross validation checks are also performed between the altimeter SWH observations and sea ice concentration to check for impossible SWH retrievals. Sea ice analyses are performed at the same time as the SWH analysis to provide the QC procedure with a contemporaneous sea ice concentration field. SSM/I sea ice retrievals from the DMSP series of satellites are used in the sea ice analysis. The QC processes result in the assignment of a probability of gross error to each altimeter SWH retrieval. The magnitude of an acceptable gross error probability is a user-defined parameter in NCODA, and thus an integral component of the space/time queries performed on the QC data files when gathering SWH observations for assimilation.

A "super observation" algorithm is used to thin the data prior to the analysis. Thinning of the relatively high volume altimeter SWH observations is a necessary step in the analysis in order to remove redundancies in the data and minimize horizontal correlations among observations. NCODA uses an adaptive algorithm to compute super-observations by averaging SWH retrieval innovations into bins dependent on grid resolution and observation data type (satellite). The algorithm is adaptive in that as the model grid resolution increases the

actual number of innovations averaged into a super-observation decrease until, eventually, the original data are directly assimilated. The resolution of the altimeter SWH retrievals is ~ 7 km along track, and the analysis is performed on a global 0.5-degree spherical grid. This discrepancy in resolution between the observations and the model grid results in SWH super-observations being formed, typically, from ~ 7 altimeter SWH retrievals.

The altimeter SWH bias corrections of Cotton (2002) for GFO, ERS2, and Topex are applied to the SWH retrievals prior to assimilation. Bias corrections do not exist for JASON-1 and ENVISAT at the time of the wave model data assimilation runs, so these satellite data are not bias corrected. Bias corrections are applied prior to the QC and prior to the data thinning procedures.

Validation and Verification. Simple bulk measures of root-mean-square (RMS) error and mean bias of the innovations are computed every update cycle. These statistics are used to assess the quality of the analysis. Spatial autocorrelation analysis of the SWH analysis residual vectors $[\mathbf{y} - \mathbf{H}(\mathbf{x}_a)]$ is used to determine the fit of the analysis to the altimeter SWH observations. In theory, the analysis residuals should be uncorrelated at all spatial lags greater than one. Any spatial correlation remaining in the residuals represents information that has not been extracted by the analysis (Hollingsworth and Lonnberg 1989). Fig. 4 shows the residual autocorrelation analyses of JASON-1 and ENVISAT altimeter SWH observations from the assimilation run. As expected, autocorrelations at all spatial lags greater than one are close to zero, which indicates an effective analysis.

In the Wavewatch analysis update cycle, innovations of the ENVISAT and JASON-1 altimeter tracks synoptic about the analysis time are computed and processed through the NCODA analysis scheme to produce the analyzed increments. The analyzed increment field is added to the Wavewatch 6-hour SWH forecast (\mathbf{H}^f) valid at the analysis time, to produce the corrected SWH analysis field (\mathbf{H}^a). The analyzed wave model spectrum (\mathbf{F}^a) as a function of frequency (\mathbf{f}) and direction (Θ) is then obtained from the forecast spectrum (\mathbf{F}^f) using a simple scaling strategy,

$$a = (H^a / H^f)^2$$

$$F^a(f, \Theta) = aF^f(f, \Theta)$$

The assimilation run and a non-assimilative control run are compared to independent buoy and yet-to-be assimilated altimeter SWH measurements. The 18 moored buoy locations are shown in Fig. 5. The buoy SWH measurements are plotted against collocated model forecast SWH fields from the assimilation and non-assimilation control runs of the Wavewatch model (Fig. 6). Fig. 6 shows a 32% reduction in bias and a 15% reduction in root mean square error for the assimilation run at the 6-hour forecast period. Further impacts of the assimilation can be seen from individual buoy time series. For example, National Data Buoy Center (NDBC) buoy 44004 is located 200 nautical miles east of Cape May, New Jersey, in 3124 meters of water. The time series of the buoy 44004 SWH measurements show a 8.5 m wave event on day 70, under predicted by almost 2 m in the control run, that is closely predicted in the 6-hour forecast from the assimilation run (Fig. 7). Fig. 8 shows similar, improved agreement of the 6-hour wave model SWH forecast from the assimilation with the buoy SWH trace for NDBC buoy 46059, located in the North Pacific, as compared to the non-assimilative control run of the wave model. Fig. 9 shows the time series of altimeter SWH innovations and residuals at each update cycle. The 6-hour Wavewatch SWH forecasts at the altimeter observation locations from the operational free run of the model are also shown in Fig. 9. The stability and the effect of the assimilation system is seen in the unbiased residuals and in the consistent reduction in error of the innovations from the control run. The average 6-hour forecast RMS error over the 30-day period is 0.61 m in the control run, and 0.46 m in the assimilation run.

Discussion. The experiment described here is a first attempt to assimilate altimeter SWH into the FNMOC global Wavewatch III model. Future work will include testing the sensitivity of the spectral modification method and the effect of the assimilation on the wave model forecast at forecast periods longer than the 6-hour update cycle. Also, work is underway to look at spatial dependence of the horizontal correlation length scales used in the assimilation. A real-time operational test of the FNMOC wave model assimilation system is planned for the 2004-2005

northern hemisphere winter. Once the assimilation method is verified it will be included in the FNMOG operational wave model run.

References

Bender L.C. and T Glowacki (1996). The assimilation of altimeter data into the Australian wave model. *Aust. Met. Mag.* 45, 41-48.

Bidlot, J.R. and M.W. Holt (1999). Numerical wave modeling at operational weather centres, *Coastal Engineering* 37, 209-429.

Cotton, P.D. (2002). *Satellite Observing Systems*, Ltd, UK

Cummings, J.A. (2003). Ocean Data Assimilation, In *COAMPS: version 3 model description*, pp 21-28. NRL Publication NRL/PU/7500-03-448.

Esteva, D.C. (1988). Evaluation of preliminary experiments assimilating Seasat significant wave heights into a spectral wave model. *J. Geophys. Res.* 93, 14099-14106.

Greenslade D.J.M (2001). The assimilation of ERS-2 significant wave height data in the Australian region. *Journal of Marine Systems*, 28, 141-160.

Greenslade D.J.M and I.R. Young (2004). The Impact of Altimeter Sampling Patterns on Estimates of

Background Errors in a Global Wave Model. *Journal of Atmospheric and Oceanic Technology*, Submitted.

Hollingsworth, A. and P. Lonnerberg (1986). The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus* 38A:111-136.

Hollingsworth, A. and P. Lonnerberg (1989). The verification of objective analyses: Diagnostics of analysis system performance. *Meteor. Atmos. Phys.* 40:3-27.

Lionello, P., H. Gunther, and P. Janssen, (1992). Assimilation of altimeter data in a global third generation wave model. ECMWF Tech. Report No. 67.

Skandrani C., J.M. Lefevre, L.Aouf, P. Queffeuilou (2003). Impact of multi-sources of altimeter data (ERS2, ENVISAT, JASON) on wave forecasts. CD-ROM proceedings of EGS-AGU-EUG Joint Assembly, Vol. 5, NICE, France, 06-11 Apr. 2003.

Tolman, H.L. (1990). A third-generation model for wind waves on slowly varying, unsteady and inhomogeneous depths and currents. *Journal of Physical Oceanography*, 21, 782-787.

Wittmann, P.A. (2002). Implementation of Wavewatch III at Fleet Numerical Meteorology and Oceanography Center. Conf. Proceedings: MTS/IEEE: Conference and Exposition. Nov 5-8, 2001 Honolulu, HI (sponsored by the Marine Technology Society and IEEE), 1474-1479.

Table 1. Altimeter SWH observation and Wavewatch III SWH prediction errors (m) estimated from the spatial autocorrelation functions computed from the non-assimilation (June-December 2003) and assimilation (March 2004) innovations time series.

Satellite	Non-assimilative Control Run		Assimilation Run	
	Observation	Prediction	Observation	Prediction
GFO	0.30	0.45		
ERS2	0.43	0.48		
ENVISAT			0.30	0.37
JASON-1	0.40	0.60	0.43	0.44

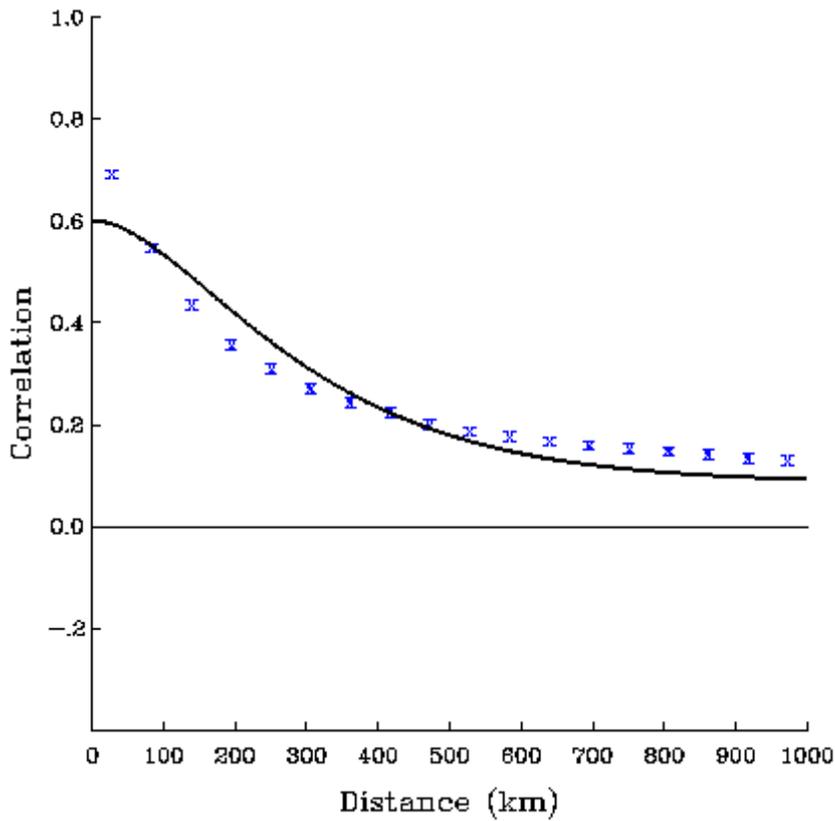


Figure 1. Bin averaged correlations (χ) for JASON-1 altimeter SWH observations estimated from a non-assimilative run of the analysis system from June-December 2003. The solid line is a least squares fit of a SOAR function to the bin averaged correlation estimates. A total of 775,500 altimeter innovations are used in the calculations. The correlation length scale is estimated to be ~223 km.

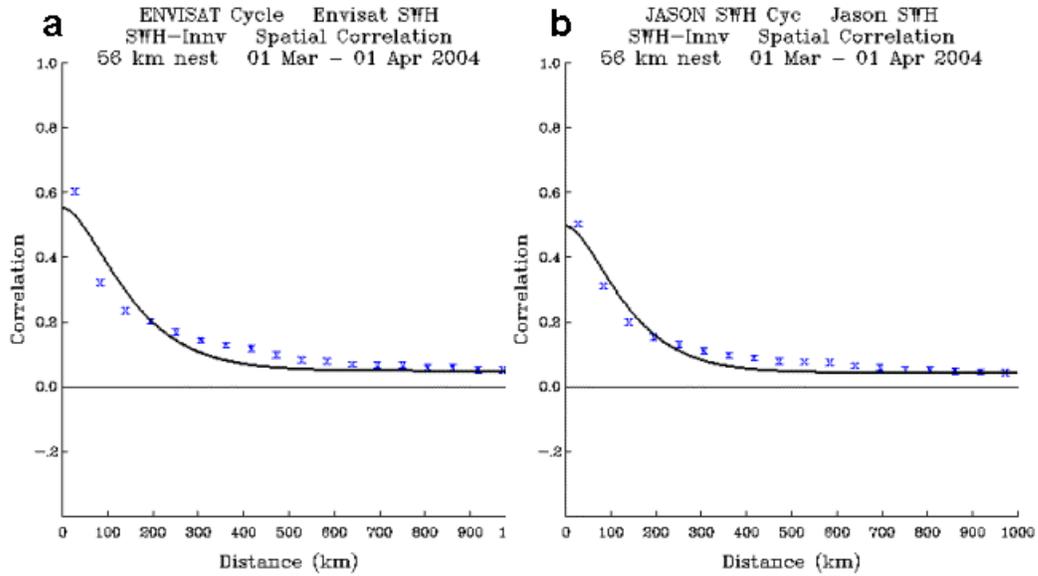


Figure 2. ENVISAT (a) and JASON-1 (b) SWH autocorrelation functions computed from the March, 2004 assimilation innovation time series. The bin-averaged correlation estimates are marked with an x , and the non-linear least squares fit of a SOAR function to the correlation estimates is shown as solid curves. A total of 287,072 JASON-1 and 188,898 ENVISAT innovations are used in the calculations. The correlation length scales are estimated to be 110 km for JASON-1 and 114 km for ENVISAT.

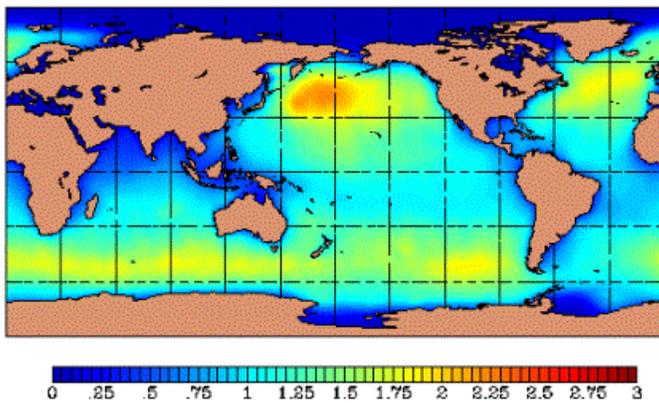


Figure 3. Wave Watch III significant wave height 6-hour prediction error variances (m^2) computed from the June–December, 2003 non-assimilative innovation time series. See text for details on how the background error variances are computed.

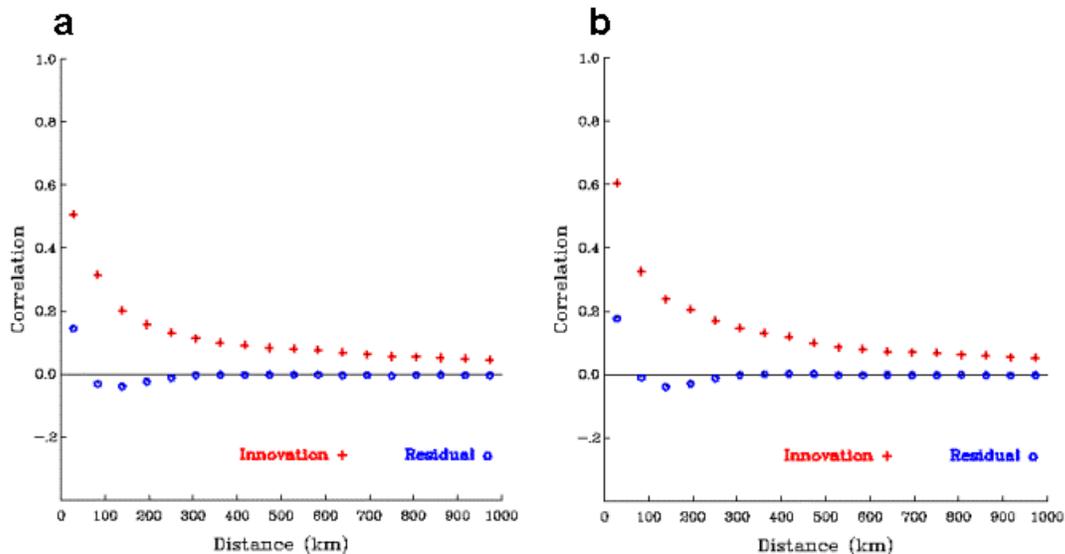


Figure 4. Residual autocorrelation analyses of ENVISAT (a) and JASON-1 (b) altimeter SWH innovation time series from March 2004 assimilation run. The residual autocorrelation estimates are marked with an *o*, and for reference purposes the innovation autocorrelation estimates are shown marked with a +. The analysis residuals are essentially uncorrelated after one spatial lag.

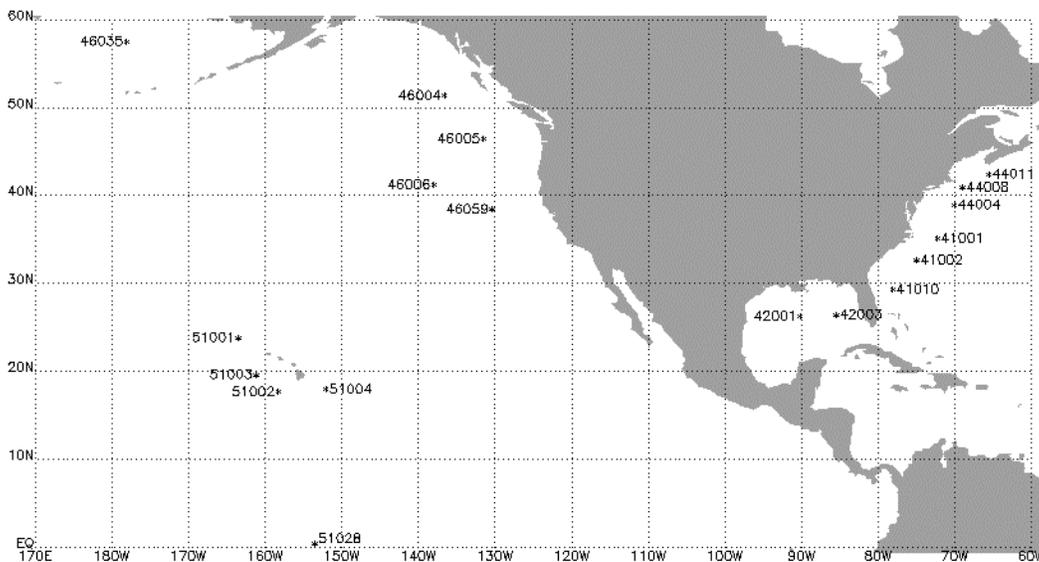


Figure 5. NDBC Buoy locations of the 18 buoys used to verify the control and assimilation runs.

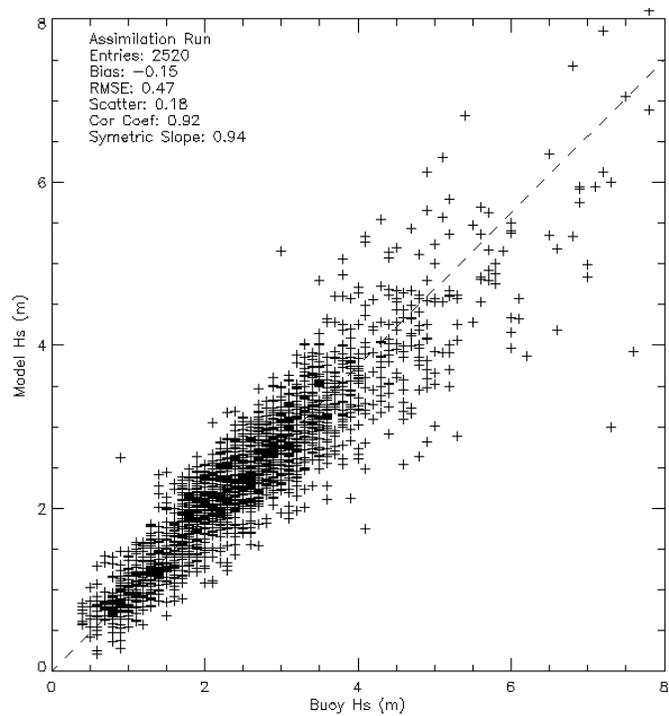


Figure 6a. Wave height (Hs) measurements from 18 NDBC deep-water wave buoys plotted against the WW3 assimilation run, for March, 2002. Forecast time is 0.

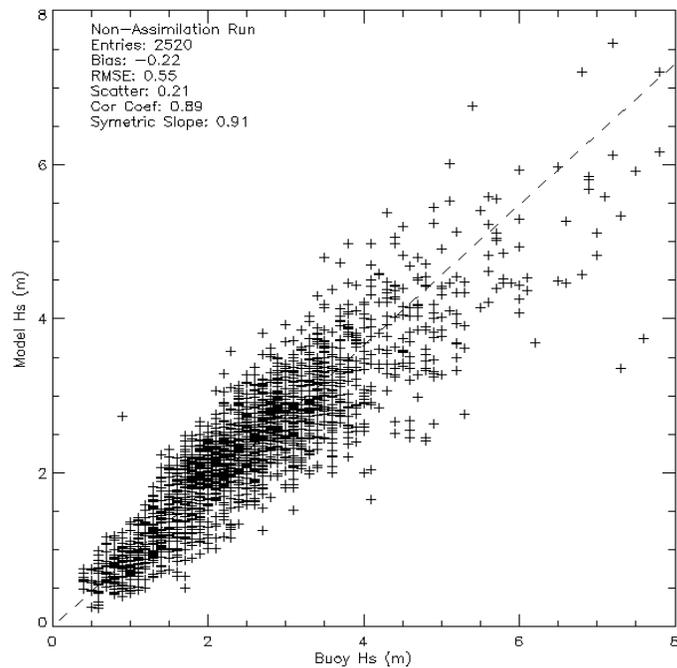


Figure 6b. Same as 6a, except model values are from the control run.

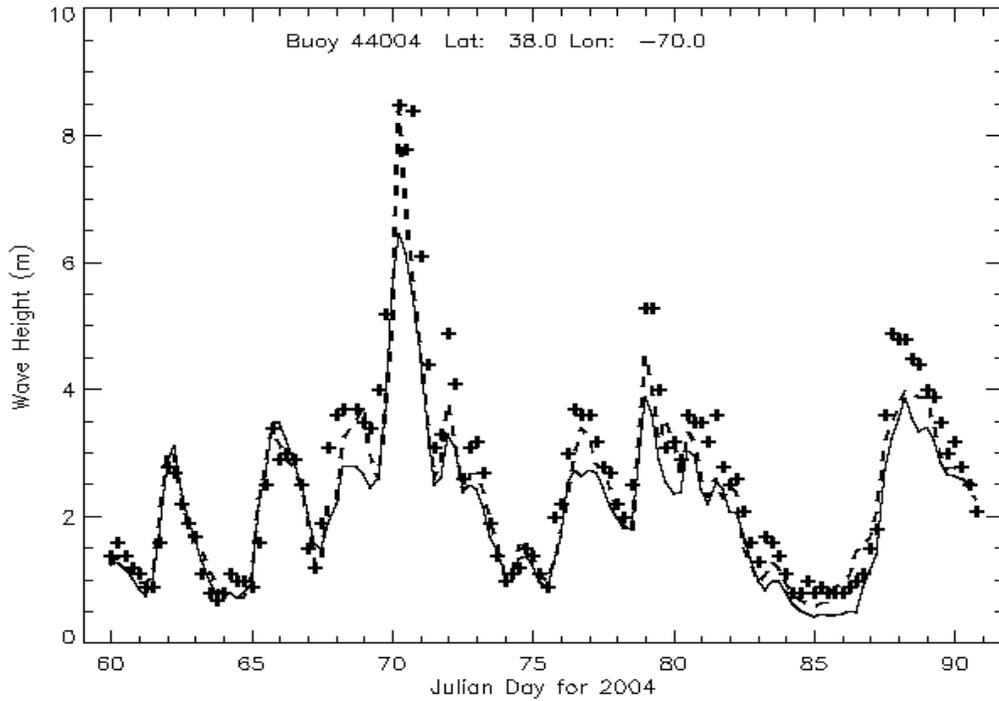


Figure 7. Time series plot of control run (solid line), assimilation run (dashed line) and wave height measurements (crosses) from NDBC buoy 44004, located in the northwest Atlantic.

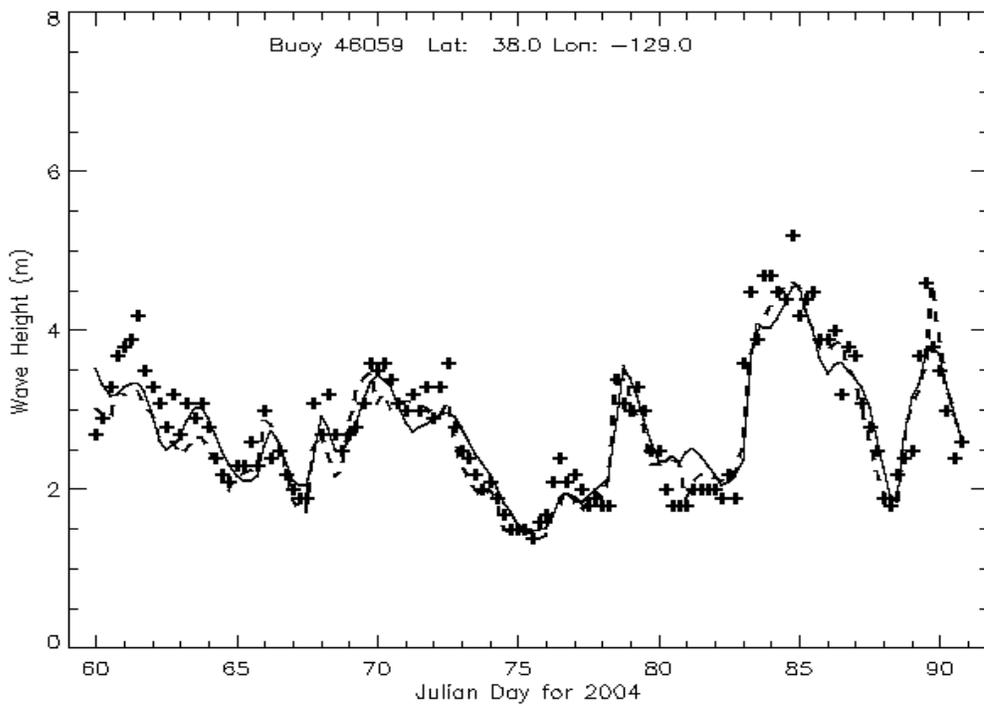


Figure 8. Same as Figure 7, except for NDBC buoy 46059, located in the northeast Pacific.

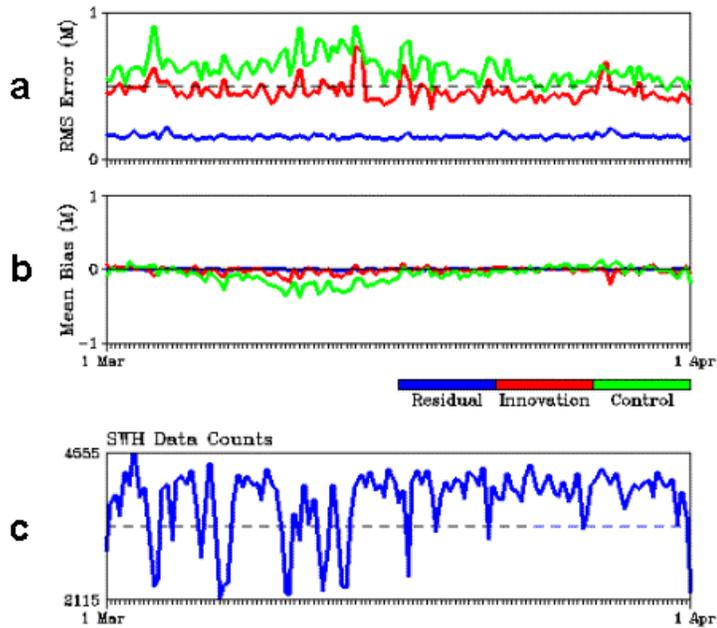


Figure 9. Verification of March 2004 assimilation and control runs using altimeter SWH observations. In the top two frames, 6-hour forecast errors of the free run of the model (control) are shown in green and the errors of the assimilation run (innovation) are shown in red; analysis residuals are shown in blue. (a) RMS error, (b) mean bias error, (c) data counts of ENVISAT and JASON-1 SWH super-observations used in each assimilation update. Each tick mark along the time axis represents an assimilation update cycle.