Ocean Dynamics
DOI 10.1007/s10236-014-0700-2

Optimizing spectral wave estimates with adjoint-based sensitivity maps

Mark Orzech · Jay Veeramony · Stylianos Flampouris

Received: 1 October 2013 / Accepted: 25 January 2014
© Springer-Verlag Berlin Heidelberg (outside the USA) 2014

Abstract A discrete numerical adjoint has recently been developed for the stochastic wave model SWAN. In the present study, this adjoint code is used to construct spectral sensitivity maps for two nearshore domains. The maps display the correlations of spectral energy levels throughout the domain with the observed energy levels at a selected location or region of interest (LOI/ROI), providing a full spectrum of values at all locations in the domain. We investigate the effectiveness of sensitivity maps based on significant wave height (H_s) in determining alternate offshore instrument deployment sites when a chosen nearshore location or region is inaccessible. Wave and bathymetry datasets are employed from one shallower, small-scale domain (Duck, NC) and one deeper, larger-scale domain (San Diego, CA). The effects of seasonal changes in wave climate, errors in bathymetry, and multiple assimilation points on sensitivity map shapes and model performance are investigated. Model accuracy is evaluated by comparing spectral statistics as well as with an RMS skill score, which estimates a mean model–data error across all spectral bins. Results indicate that data assimilation from identified high-sensitivity alternate locations consistently improves model performance at nearshore LOIs, while assimilation from low-sensitivity locations results in lesser or no improvement. Use of sub-sampled or alongshore-averaged bathymetry has a domain-specific effect on model performance when assimilating from a high-sensitivity alternate location. When multiple alternate assimilation locations are used from areas of lower sensitivity, model performance may be worse than with a single, high-sensitivity assimilation point.

Keywords Sensitivity map · Data assimilation · Numerical adjoint · SWAN · SWANFAR · Wave spectra

Responsible Editor: John Wilkin

M. Orzech (✉) · J. Veeramony
Naval Research Laboratory, Code 7322,
Stennis Space Center, MS 39529, USA
e-mail: mark.orzech@nrlssc.navy.mil

S. Flampouris
University of Southern Mississippi, Hattiesburg, MS 39406, USA

1 Introduction

Accurate forecasts of nearshore wave conditions are important to a diverse constituency, including vacation destinations such as Miami Beach or San Diego, coastal residents, environmental scientists tracking urban runoff, municipalities with harbor facilities, and many others. Model-based wave forecasts may be improved by assimilating data from sites within the modeled domain, preferably as close as possible to the location/region of interest (LOI/ROI) where accuracy is most important. Occasionally, however, conditions make it difficult or impossible to deploy an instrument at or near the desired location. Wave breaking may be too intense, bathymetric irregularities may make deployments infeasible, or environmental sensitivities may prohibit such activities.

In such cases, an adjoint-based sensitivity map for waves can be extremely useful for identifying alternate locations whose spectral energy levels are most correlated with those at the original site. Sensitivity maps are widely used in oceanography to investigate and characterize patterns of variability, in quantities ranging from sea-surface temperature (e.g., Moore et al. 2009) to circulation (e.g., Ngodock and Carrier 2013; Zhang et al. 2008; Bugnion et al. 2006) to carbon sequestration in the ocean (e.g., Hill et al. 2004). Little or no work has been done with sensitivity maps of surface wave spectra.

Sensitivity maps are generally constructed for a selected system indicator (e.g., vorticity) by computing the differential of that indicator at all domain locations in response to a unit impulse applied to a specific system parameter at a chosen

Published online: 18 February 2014

LOI or ROI. The magnitudes of sensitivity values indicate the relative correlation of variations in the indicator throughout the domain with its variation at the impulse location(s). The highest sensitivity values mark the locations of greatest correlation. When the indicator values for a given domain are computed by a model system, the adjoint to the model may be used to construct a sensitivity map, essentially by reversing the forward model's direction and propagating the variation of the selected indicator "backward" in space and/or time.

The present study evaluates the effectiveness of spectral wave sensitivity maps in determining the best offshore alternate instrument sites for data assimilation when a given nearshore LOI or ROI is inaccessible. It makes use of an adjoint to the spectral wave model SWAN (Orzech et al. 2013), as well as a variational data assimilation system built around this adjoint (Flampouris et al. 2013). The analysis employs wave and bathymetry datasets from Duck, NC, and San Diego, CA (Field Research Facility 2013; National Data Buoy Center 2013). Sensitivity maps used here are based on a significant wave height calculation of overall sensitivity at each location:

$$H_s = 4.\sqrt{\sum_{i=1}^{I}\sum_{j=1}^{J} S\left(f_i,\theta_j\right)\Delta f_i \Delta\theta_j} \tag{1}$$

where $S(f_i, \theta_j)$ is the adjoint-estimated spectral sensitivity at discrete frequencies f_i and directions θ_j and Δf_i and $\Delta\theta_j$ are spectral bin sizes. The effects of changes in wave climate and errors in bathymetry on the shapes of H_s-based sensitivity maps are investigated. Key questions to be addressed include: Can sensitivity-guided assimilation improve spectral estimates at denied locations? How does sensitivity map effectiveness change with wave climate, LOI vs ROI, or bathymetry resolution? How is the distribution of sensitivity affected by wave climate or bathymetry resolution?

The following section briefly describes the SWAN wave model and the recently constructed numerical adjoint to SWAN, as well as the techniques for creating and interpreting sensitivity maps from adjoint output. Section 3 describes five separate sets of simulations conducted to identify optimal alternate instrument sites for LOIs and ROIs under varied wave conditions and bathymetry, as well as techniques used to evaluate model performance. Results from each analysis are presented and interpreted in Section 4, followed by further discussion and overall conclusions in Section 5.

2 Model and adjoint

2.1 SWAN wave model

SWAN (Booij et al. 1999) is a stochastic wave model that is widely used in the coastal modeling community. It produces nearshore wave forecasts and climatologies by solving the spectral action balance Eq. 2, generally initialized at the offshore boundary with spectral wave and other outputs from regional models such as WAVEWATCH III® (Tolman 2009). SWAN's principal spectral action balance equation is generally written as

$$\frac{\partial N}{\partial t} + \vec{\nabla}\cdot\left(\vec{C}N\right) = \frac{S_{\mathrm{tot}}}{\sigma} \tag{2}$$

where N is spectral action density (spectral energy density divided by an intrinsic representative wave frequency, σ). The first term of (2) expresses the time rate of change of the action density, while the second term represents the propagation of action density in both physical and spectral space as the dot product of vector gradient $\vec{\nabla} = (\partial/\partial x, \partial/\partial y, \partial/\partial\sigma, \partial/\partial\theta)$ with vector velocity $\vec{C} = \left(C_x, C_y, C_\sigma, C_\theta\right)$ and action density N. S_{tot} is a combined source/sink term, including contributions from bottom friction, nonlinear wave–wave interactions, wave breaking dissipation, white-capping, and wind-wave forcing (Booij et al 1999). SWAN must be provided with bathymetry, wave data, and source term configuration, with which it will compute the numerical solution of (2) to predict the evolution of the wave spectrum throughout the model domain.

2.2 Numerical adjoint to SWAN

The numerical adjoint to SWAN is composed of a collection of adjoint subroutines, each individually constructed from a corresponding subroutine in the original "forward" SWAN model (Orzech et al. 2013). Every subroutine in the forward model that acts on the wave action N is matched by a complementary subroutine in the adjoint model. Individual adjoint SWAN subroutines were built with the Parametric Fortran Compiler or "PFC" utility (Erwig et al. 2007), which (somewhat crudely) linearizes and/or transposes each forward subroutine by applying standard rules of automatic differentiation together with user-specified active and dependent variables.

The current version of the SWAN adjoint model, utilized in this study, includes adjoints to all source and sink subroutines in the forward model, as well as adjoint subroutines for the spatial and spectral propagation of wave action. It is configured to handle nonstationary conditions (although all simulations in the present analysis are stationary). All subroutines as well as the complete adjoint code have been fully tested for consistency using standard matrix analysis techniques (see Orzech et al. 2013). The adjoint model did not incorporate wave–current interaction at the time of this study, so currents are neglected here.

2.3 SWANFAR assimilation system

The SWAN adjoint model forms the core of a newly developed data assimilation system for SWAN, called SWANFAR

(Flampouris et al. 2013; Veeramony et al. 2012; Orzech et al. 2012). The system consists of individual forward, adjoint, and representer-perturbation ("RP" forward) versions of SWAN, which are integrated into a Fortran-based wrapper code (Fig. 1). The system initially runs forward SWAN to create "first guess" background spectra at all locations on the specified bathymetry grid. Spectra at the assimilation location(s) are then compared to observations by the wrapper, which computes a cost function to estimate the mismatch between model and observation. A conjugate gradient technique is then applied to iteratively minimize this cost function (see Appendix A of Orzech et al. 2013). The adjoint is initialized with the innovation (model–data error for each spectral bin) at the assimilation location(s) and propagates the perturbations "backward" through the model grid. Adjoint output is used to initialize the RP SWAN, which computes a new estimate for the spectral correction at the assimilation location(s). The cost function is recalculated, and the process continues in this manner until it has been minimized. The optimal spectral corrections are then added to the initial first-guess spectra to produce the revised spectral estimates throughout the domain.

Localized error covariances are not included in computations of the cost function for this study (i.e., covariance values are assumed to be uniformly equal to one). In practice, this limits the influence of assimilated spectra to smaller regions surrounding each assimilation point. A five-dimensional (x, y, f, θ, t) covariance function is presently being developed for SWANFAR but has not yet been fully validated. Test results indicate that covariance is most strongly affected by bathymetry, but only in depths where waves "feel the bottom". In deeper water where refraction and shoaling are minimal (and assuming wind gusts are light), spectral shapes and magnitudes remain relatively constant. The assumption of constant covariance is thus expected to be reasonably accurate under these more limited conditions.

The SWANFAR system is used for all data assimilation work in this study.

2.4 Sensitivity maps for wave spectra

For any type of adjoint, sensitivity maps may be constructed from adjoint output to track the response of system properties to variations in a given parameter. In general, for a given system state vector Φ,

$$\Phi(t) = [\vec{x},\ \vec{s},\ \vec{u},\ \vec{\tau}, \mathrm{SST}, S, t, \ldots]$$

an indicator J can be created to track some aspect of the system:

$$J = G[\Phi]$$

When a specific parameter A is varied at one or more locations, the differential $\delta J/\delta A$ tells how the indicator changes throughout the grid; i.e., J 's sensitivity to variations in A.

For the spectral wave model SWAN, the system state based on the wave action N is described by (2) above. In this case the indicator J is a spectrally based functional of N such as significant wave height H_s, mean period $\overline{T}$, and mean direction $\overline{\theta}$ (each a linear functional), or directional spread σ_θ (a nonlinear functional):

$$J = G[N(\vec{x},\ \vec{s}, t)] = H_s,\ \ \overline{T},\ \ \overline{\theta},\ \ \sigma_\theta,\ \ \text{etc.}$$

Following initialization with an energy impulse or a spectrum at a nearshore location, the SWAN adjoint produces sensitivity contours and spectra that are shaped by the combined effects of bathymetry, local winds and currents, and configuration of the input spectrum. As illustrated in Fig. 2,

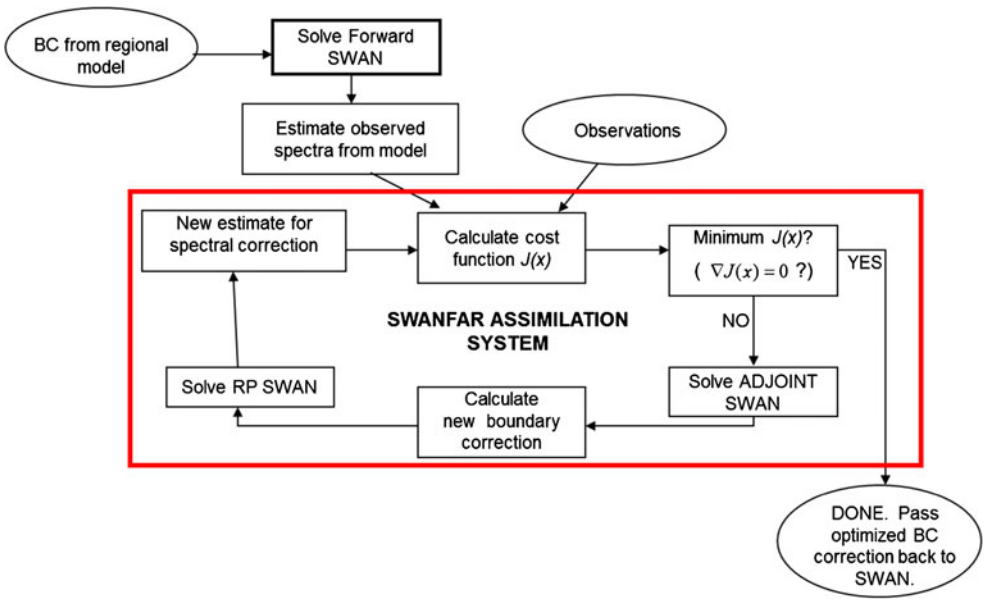

Fig. 1 SWANFAR assimilation system

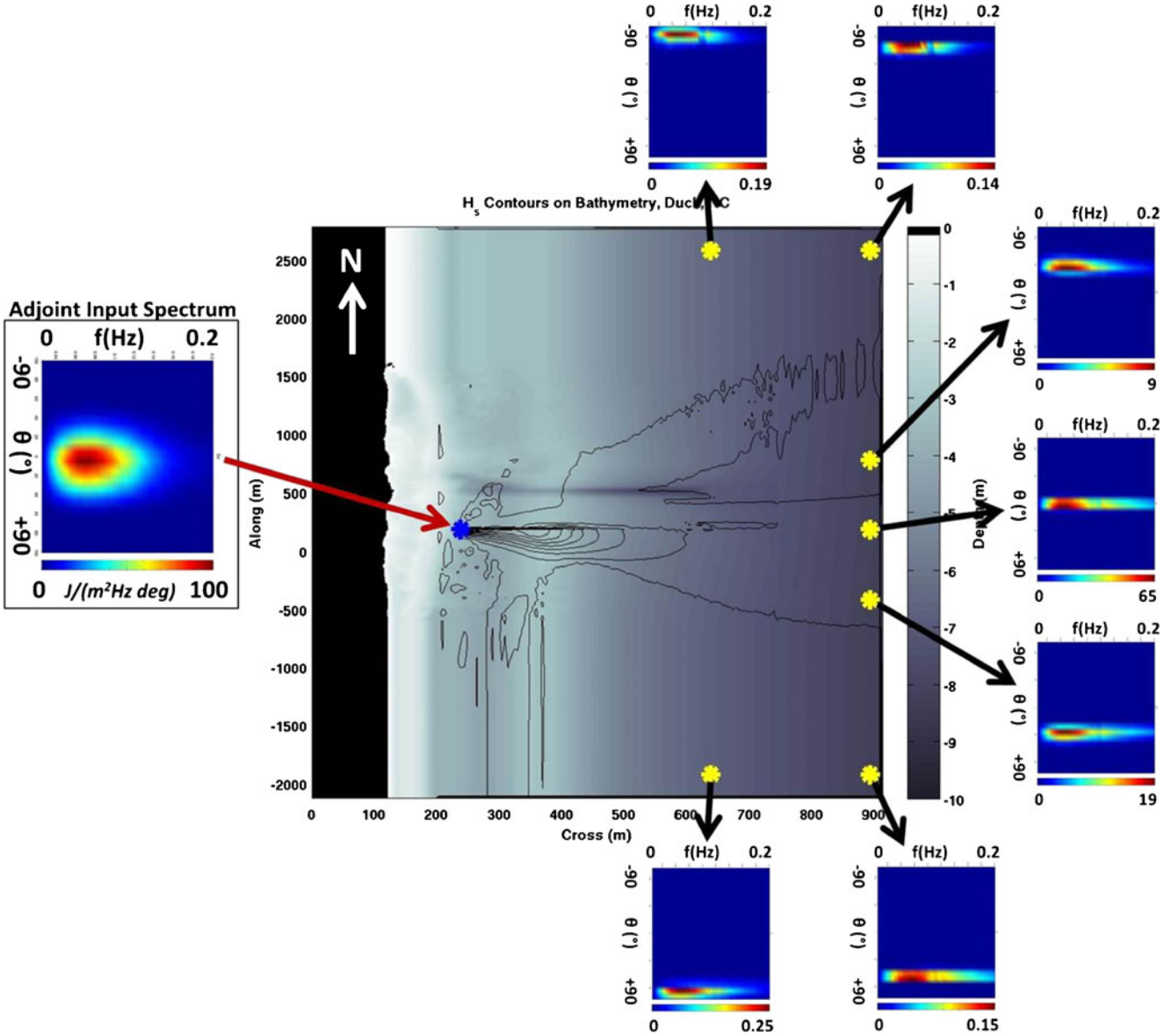

Fig. 2 Sample output from adjoint-generated spectral sensitivity map at Duck, NC. Adjoint is initialized with an artificial wave spectrum (*left panel*) at the nearshore location of interest indicated with a *blue asterisk. Black contours plotted on top of grayscale Duck bathymetry in large center panel* show variations in adjoint-estimated significant wave height; high-valued contours mark areas whose energy variations are most highly correlated with those at the LOI. Adjoint-generated frequency directional spectra are displayed for each location marked by a *yellow asterisk. Color bars* along bottom of each spectral panel indicate energy ranges in J/(m^{2}Hz deg)

adjoint output includes a complete frequency-directional spectrum at every grid location. Mean period $\overline{T}$ remains relatively constant in the output spectra, resulting in essentially "flat" sensitivity maps (although these gradients can be greater under strong winds). Mean direction $\overline{\theta}$ shifts in accordance with position relative to the initialization location, leading to sensitivity maps with contours that extend radially from the LOI. Directional spread σ_θ becomes increasingly narrow with distance from the initialization point, producing sensitivity maps with concentric contours centered on the LOI.

Total spectral energy and H_s values vary along the boundary, as dictated by directional variations in the LOI spectrum in combination with bathymetric and other environmental effects. In Fig. 2, because spectral peak direction in the initializing spectrum at the LOI is from slightly south of shore-normal, the highest sensitivity contours (with greatest spectral energy levels) also extend in that direction. In addition, some adjoint-propagated wave energy is blocked (refracted) by the Duck pier channel to the north of the LOI ($y \approx 500m$ alongshore), so that northern boundary energy levels are somewhat smaller than those in the south. Of the above-mentioned parameters, only sensitivity maps showing the variation of H_s (e.g., $\delta H_s(x_m, y_n)/\delta S(f_a, \theta_b, x_{\mathrm{LOI}}, y_{\mathrm{LOI}})$) are effective for identifying high correlation alternate assimilation sites under

these sorts of conditions. As this is the primary goal of the present study, only H_s-based sensitivity maps are utilized in the analyses that follow.

3 Methods

The goal of these simulations is to evaluate the effectiveness of the H_s-based sensitivity map in identifying optimal alternate locations for instrument deployment, when the desired LOI or ROI is inaccessible. In a realistic scenario, in situ measured wave data will not be available from the LOI, so the adjoint will need to be initialized with a "first guess" background spectrum from the forward model. The effectiveness and correctness of the sensitivity map in such cases will thus clearly be dependent on the initial accuracy of the forward model, as well as the accuracy of the bathymetry data. The following tests are designed with these limitations in mind.

Sensitivity maps are generated for two distinct regions: one very shallow-water, small-scale domain (Duck, NC) and one deeper-water, larger-scale domain (San Diego, CA). A set of 12 representative "observed" wave climates is constructed at each site by initializing the forward model with mean monthly offshore wave conditions for Jan–Dec, averaged for each month over three separate years (2009–2011). For each case, the adjoint system is initialized with first guess spectra from forward SWAN at the inaccessible LOI or ROI. The adjoint-generated map is used to identify maximum and minimum sensitivity locations within a hypothetical "accessible region" farther offshore. The SWANFAR assimilation system is then configured to assimilate the "observed" spectrum from either the maximum- or minimum-sensitivity alternate location for each case. Improvements in the LOI spectral estimate when assimilating from the optimal or "best" (high sensitivity) alternate location are compared to the corresponding results obtained when assimilating from the "worst" (low sensitivity) location. Results are compared to assimilation at the LOI/ROI itself and/or the control case with no data assimilation.

Because the locations of high- and low-sensitivity assimilation points are not known in advance in these simulations, it is not possible to use actual measured wave spectra as observations. This would require instruments at all grid locations in the accessible region, a network of ground-based radar systems, or very high-resolution satellite data, which would not be feasible for most sites. Instead, forward SWAN is first initialized with realistic boundary conditions (i.e., JONSWAP spectra based on monthly means of observed H_s, $\overline{T}$, and $\overline{\theta}$), and model output is stored as quasi-observations throughout the entire domain. For simulations at Duck, NC, observed boundary wave data come from records of the FRF's 8 m array. Observed monthly mean wave statistics at the 8-m array have the following ranges for 2009–2011: H_s=0.53–1.24 m, $\overline{T}$ =8.6–10.0 s, and $\overline{\theta}$ =0.4–20.4° (relative to shore-normal). Three-year-averaged boundary spectra have H_s=0.86 m, $\overline{T}$ = 9.3 s, and $\overline{\theta}$ =10.9°. For simulations at San Diego, CA, observed boundary data come from NDBC buoy 46231. Observed monthly mean wave heights and periods at the buoy are larger and wave directions somewhat more variable than those at Duck: H_s=0.98–1.39 m, $\overline{T}$ =12.1–14.0 s, and $\overline{\theta}$ = 280–330° (relative to true N). Three-year-averaged boundary spectra have H_s=1.17 m, $\overline{T}$ =13.3 s, and $\overline{\theta}$ =299°.

Following the generation of quasi-observations, forward SWAN is re-initialized with the following modified boundary conditions: a "low" scenario (using measured H_s −10 %, $\overline{T}$ −10 %, and $\overline{\theta}$ −10°), a "high" scenario (using measured H_s + 10 %, $\overline{T}$ +10 %, and $\overline{\theta}$ +10°), and an "average" scenario (using 3-year averages of observed H_s, $\overline{T}$, and $\overline{\theta}$ values). SWAN output spectra under each boundary scenario become the model estimates "with error". The initializing error ranges are intended to represent typical errors seen in the output of larger-scale models such as WAVEWATCH III® (Tolman 2009), which are often used to initialize SWAN for domains where buoy data are unavailable. They are not intended to comprehensively describe all possible conditions seen at either site. Although the directional error, ±10°, is equivalent to the directional resolution of the SWAN spectra used in this study, its effects are clearly distinguishable in model output.

The accuracy of spectral estimates is evaluated by calculating correlations of H_s, $\overline{T}$, $\overline{\theta}$, and directional spread σ_θ, and also by computing the RMS spectral skill score at the LOI/ROI for each assimilation, as follows:

$$\text{skill} = \max(\xi\ ,\ 0), \quad \text{where} \quad \xi = 1 - \frac{\sqrt{\sum_{i,j}\left(S_{\text{mod}}(f_i,\theta_j) - S_{\text{obs}}(f_i,\theta_j)\right)^2}}{\sqrt{\sum_{i,j}\left(S_{\text{obs}}(f_i,\theta_j)\right)^2}} \tag{3}$$

Here, $S_{\text{mod}}(f_i, \theta_j)$ and $S_{\text{obs}}(f_i, \theta_j)$ represent the spectral energy density in frequency-directional bin f_i, θ_j of the model-estimated and observed spectra, respectively. All bins of the two spectra are compared, with ξ ranging from negative

values (large difference between model and observations and/or very low observed wave energy) to one (perfect match). The RMS skill value itself is limited to the range 0–1, in order to focus on cases with better model performance. Equation 3 is most informative and useful when S_{mod} is relatively similar to S_{obs}. As will be seen, the RMS skill score provides a concise and often helpful alternate perspective on model accuracy to supplement the more traditional examination of spectral statistics.

3.1 Test set #1: Duck LOI

The first set of simulations examines the effectiveness of spectral sensitivity maps in a very shallow, mildly sloping, surf-zone environment—the CRAB-surveyed region seaward of the Field Research Facility (FRF) in Duck, NC (Fig. 3, left panel). An "inaccessible" location of interest is selected in 2.5 m water depth in the surf zone, north of the pier (blue asterisk in figure). An alternate "accessible" region roughly 200 m by 1,000 m is arbitrarily defined about 600 m farther offshore (green dashed lines). For each of 12 months (Jan–Dec), forward SWAN is employed four times to generate "observed", "low", "high", and "average" spectral datasets throughout the domain. The SWAN adjoint is then separately initialized with each estimated spectrum at the LOI, producing an H_s-based sensitivity map for each of the three (low/high/avg) cases in each month. "Best" and "worst" (i.e., maximum and minimum sensitivity) alternate locations are identified within the accessible region.

Next, the SWANFAR assimilation system is used to conduct an observing system simulation experiment. Estimated (erroneous) boundary conditions are utilized while the "observed" spectrum is assimilated from either the best or the worst alternate location. The post-assimilation spectral estimate at the inaccessible LOI is compared to the original observed spectrum there to determine how much (if any) the model estimate was improved by using observed data from the alternate locations. These assimilation results are all compared against two control cases: one in which the observed spectrum is assimilated directly from the LOI, and a second in which no assimilation is used (uncorrected estimates versus observations).

Questions to be answered by these simulations include: Does assimilation from the "best" alternate location consistently outperform assimilation from the "worst" location? How do results from the two alternative locations compare to assimilation from the LOI itself? Are there conditions under which neither alternate assimilation type is effective? Do low, high, or averaged boundary conditions tend to result in more accurate assimilation results?

3.2 Test set #2: San Diego LOI

The second set of assimilations is structured similarly to the first, except that the model domain, including Scripps and La Jolla Canyons offshore of San Diego, CA, is significantly larger, with a greater variation in water depth (Fig. 3, right panel). An inaccessible LOI is again defined, this time in roughly 25 m depth, shoreward of the two canyons. A larger accessible region, approximately 5 by 2 km, is delineated about 5 km offshore, in 30–500 m water depth. Observed and low/high/avg estimated spectra are again generated for the domain, and assimilations are conducted using both "best" and "worst" alternate locations within the accessible region, as identified by the adjoint-generated sensitivity maps. Model performance in these tests is again compared to control cases with no assimilation and with assimilation directly from the LOI.

Questions to be answered by these simulations include those of the previous section as well as: Is the performance of the model and assimilation system in the deeper, larger-scale San Diego environment significantly different from its performance in the shallower Duck domain?

3.3 Test set #3: low-resolution bathymetry

The third set of simulations examines the effects of an incomplete or approximate bathymetry on assimilation performance

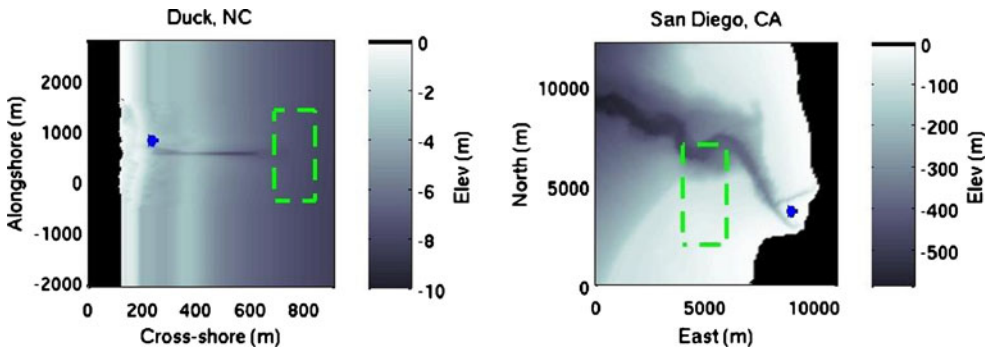

Fig. 3 Configuration for test sets #1 and #2. Model domains at Duck, NC (*left panel*) and San Diego, CA (*right panel*). *Blue asterisk in each panel* marks the inaccessible location of interest. Rectangular areas bounded by *green dashed line* are accessible regions of alternate assimilation locations. LOI depth is approximately 2.5 m at Duck and 25 m at San Diego

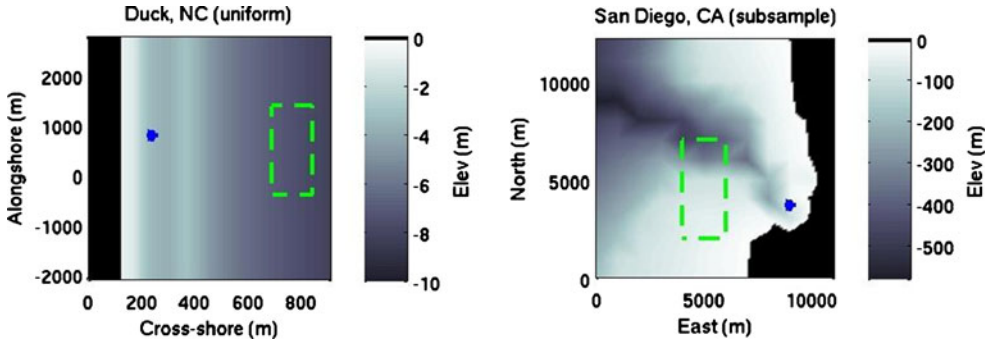

Fig. 4 Configuration for test set #3. Same configuration as in Fig. 3, but with low-resolution bathymetry at Duck, NC (*left panel*) and San Diego, CA (*right panel*). Duck data are alongshore uniform, based on a mean cross-shore profile, while San Diego data have been subsampled from original bathymetry (every tenth point in x and y). Estimated LOI depth is approximately 3.5 m at Duck and 54 m at San Diego

at the two sites (Fig. 4). For Duck, it is assumed that only a mean cross-shore depth profile is available, so that model bathymetry is alongshore uniform for all simulations. At San Diego, original bathymetry is subsampled, taking every 10th point in both x- and y-directions, and the new bathymetry is linearly interpolated from the subsampled depths to the same grid spacing as the original. In this test set, mean monthly wave spectra are again used at the boundary to create "observed spectra" for each of 12 months at each site, but only the 3-year-averaged wave conditions are employed as boundary conditions for generating model estimates. All observed spectra are produced on the original high-resolution bathymetry, while all model estimates are created using the more limited low-resolution bathymetry. Sensitivity map shapes, predicted optimal assimilation locations, and LOI skill scores are compared for the two bathymetry types at each site.

Questions to be answered by these tests include: How do low-resolution and alongshore uniform bathymetry affect sensitivity map shapes and optimal alternate assimilation locations? Is model skill consistently lower on sub-sampled bathymetry?

3.4 Test set #4: region of interest

In many coastal field projects, spectral wave estimates are required for a nearshore *region* rather than just at a single point. With this in mind, the final two test sets of this study work with a broader region of interest (ROI) rather than a single location (LOI) at each site. For both Duck and San Diego the LOI from earlier tests is now replaced by a rectangular box (Fig. 5). All estimated spectra from within each box are provided to the SWAN adjoint to compute sensitivity maps and identify alternate assimilation locations. For test set #4, a single best alternate assimilation location is chosen from the accessible region. Here, the original, accurate bathymetries are again used rather than the low-resolution approximations of the preceding section, and model estimates are again limited to 3-year-averaged wave conditions.

To depict model performance in the ROI in a concise manner, the RMS skill score is used instead of the more traditional correlations of spectrally derived wave height, period, and direction. Assimilation results are compared to observed values for the entire ROI via skill score maps for each month at each site. A final skill score colormap is created by averaging over all 12 months for each location. The skill map is used to examine variations in model performance throughout the selected region.

Questions to be answered include: Is the skill of the assimilation-aided model in the ROI comparable to its skill for a single LOI or consistently better/worse? Does model skill vary significantly across either site's ROI? If so, why?

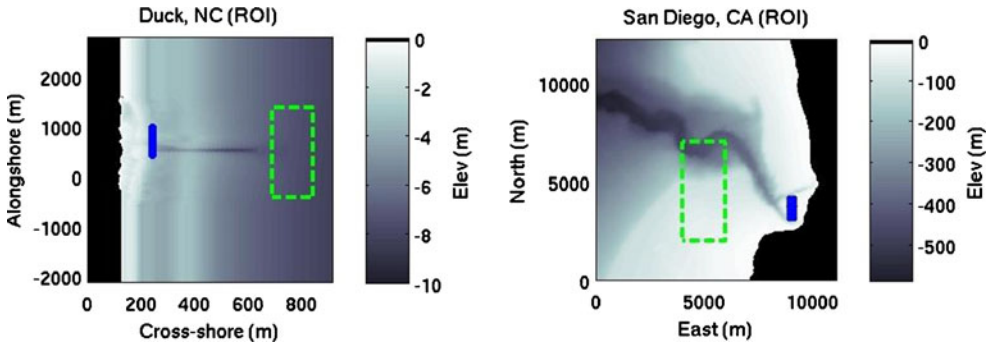

Fig. 5 Configuration for test set #4. Regions of interest (*ROIs*) are marked as *rectangular blue strips* on original bathymetry (i.e., same as Fig. 3). The Duck ROI is 15×600 m, while the San Diego ROI is 300×1,200 m. Depth ranges in the ROIs are 2.4–4.5 m at Duck and 20–90 m at San Diego. Accessible regions bounded by *green dashed lines* are the same as those used for the preceding tests

3.5 Test set #5: region of interest with two-point assimilation

In many cases, more than one instrument is available for deployment, and there are often multiple accessible regions to choose from. As a first step toward addressing this more complicated scenario, we expand on the ROI simulations of test set #4 by replacing each offshore accessible region at the Duck and San Diego sites with two separate smaller regions (Fig. 6). Sensitivity map results from test set #4 are re-accessed here to determine two optimal alternate instrument locations for each site (one within each accessible region). Assimilations with SWANFAR are rerun using two assimilated spectra instead of one, and overall model skill scores are again plotted as a colormap for the entire ROI at each site.

Questions to be addressed include: Does the use of two alternate assimilation locations consistently improve model skill in the ROI? Are results significantly different between the two sites?

4 Results

4.1 Test set #1: Duck LOI

Correlations and biases of spectral statistics at the Duck LOI based on assimilation at each alternate location and at the LOI are summarized in Table 1, with a separate value for each of three model estimate types (low/high/avg), assimilation location (best, worst, LOI, and non-assimilation), and statistic (H_s, $\overline{T}$, $\overline{\theta}$, and σ_θ). Results from all 12 months are combined in each entry. In each case, the correlation coefficient is calculated as

$$\rho_{m,o} = \frac{E[(X_m - \mu_m)(X_o - \mu_o)]}{\sigma_m \sigma_o} \tag{4}$$

and the bias of the model estimates is calculated as

$$\mathrm{Bias}(X_m) = E[(X_m - X_o)] \tag{5}$$

where E is expected value, μ is the mean, and σ is the standard deviation of individual data values represented by X, and subscripts m and o indicate model estimates and observations, respectively. Using monthly mean values with error (i.e., Table 1, rows 1–8), model–data correlations for H_s, $\overline{T}$, and $\overline{\theta}$ at the shallow water LOI are very high for both the best and worst alternate assimilation locations, as well as for the non-assimilation case. Correlations are generally lower for model estimates based on 3-year-averaged values (rows 9–12), varying considerably for estimates of $\overline{T}$, $\overline{\theta}$, and σ_θ. When no spectra are assimilated, the estimated spectrum at the LOI is the same for each month, so that $X_m = \mu_m$ and $\sigma_m = 0$ in (4), leading to a correlation coefficient of 0/0 (row 12). Correlations are significantly lower for directional spread in all cases. In contrast to correlation results, the biases for 3-year-averaged estimates are smaller than those for the monthly estimates in nearly all cases.

Somewhat unexpectedly, results are generally worse when the observed spectrum is assimilated from the LOI itself (rows 3, 7, and 11). Further investigation reveals that this anomaly is a consequence of limitations in the SWANFAR assimilation system when working with data in the shallow surf zone. If provided only with spectra from waves that have already begun breaking, the SWAN adjoint is unable to correctly estimate the offshore wave energy level (Fig. 7). This limitation arises from the wave-breaking parameterization in SWAN itself, which arbitrarily reduces the energy of spectra in the surf zone when their wave height H_s is greater than a user-specified proportion of the water depth, γh. Although it is essential to prevent SWAN-estimated wave energy from growing exponentially in shallow water, this arbitrary nonlinear correction does not have an adjoint. In order to overcome this limitation, spectra from deeper locations, seaward of the surf zone, must also be assimilated into SWANFAR. This problem does not exist for LOIs outside the surf zone that are not subject to the energy tapering, as will be seen in the assimilation results from San Diego, CA, below.

Skill scores (Eq. 3) at the Duck LOI based on assimilation at the best (i.e., highest sensitivity) and worst (i.e., lowest sensitivity) alternate locations are plotted as time series for months Jan–Dec in Fig. 8. Skill scores from model estimates without assimilation are also included for comparison. Skill values are relatively close for all three estimate types at this

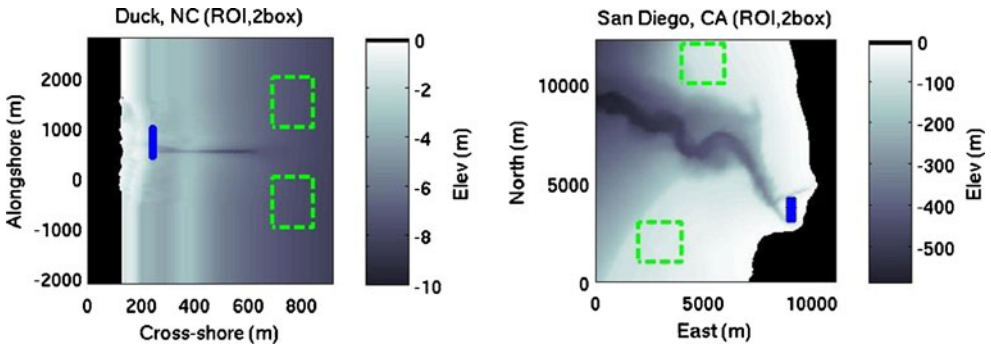

Fig. 6 Configuration for test set #5. Regions of interest (*ROIs*; marked as *blue strips*) on original bathymetry (i.e., same as Fig. 3), with two separate offshore "accessible regions" (*dashed green boxes*). One high- and one low-sensitivity alternate assimilation location are selected from each accessible region

Table 1 Correlations (and biases) of modeled vs. observed spectral statistics at LOI for Duck, NC, test set #1

	Est. Type	Assim. Locn	H_s (m)	$\overline{T}$ (s)	$\overline{\theta}$ (°)	σ_θ (°)
1	Low monthly	Best	1.00 (−0.09)	0.99 (−0.44)	1.00 (−4.1)	0.36 (−0.08)
2		Worst	1.00 (−0.10)	0.99 (−0.52)	1.00 (−4.1)	0.46 (0.13)
3		LOI	0.99 (0.00)	0.49 (−0.02)	0.74 (0.77)	0.08 (0.07)
4		No Assim	1.00 (−0.09)	1.00 (−0.59)	1.00 (−4.0)	0.58 (0.17)
5	High monthly	Best	1.00 (0.06)	0.99 (0.55)	1.00 (3.9)	0.81 (−0.22)
6		Worst	1.00 (0.07)	1.00 (0.63)	1.00 (4.0)	0.79 (−0.32)
7		LOI	0.73 (2.3)	0.49 (−0.94)	0.91 (−2.9)	0.28 (−2.9)
8		No Assim	1.00 (0.07)	1.00 (0.58)	1.00 (4.0)	0.79 (−0.31)
9	3-Yr Avg	Best	0.97 (0.03)	0.02 (0.04)	−0.48 (−0.03)	0.13 (−0.03)
10		Worst	0.70 (0.03)	−0.35 (0.03)	0.80 (−0.05)	0.18 (0.02)
11		LOI	0.91 (−0.01)	−0.26 (−0.79)	0.64 (−1.3)	0.00 (−2.6)
12		No Assim	– (0.03)	– (0.03)	– (−0.02)	– (0.04)

Bias units are shown at top of each data column

site. Assimilations from the best alternate location generally have higher skill than the other two estimates, but they are only slightly higher. Assimilations from the worst alternate location rarely improve on the original model estimates at all. Skill scores are relatively constant and fairly low for the monthly model estimates, while those based on the 3-year-averaged conditions vary much more widely and are zero (or less) for the months of June and July.

Examining time series of observed and modeled wave statistics at the Duck LOI, it will be noted that monthly model estimates track observed values fairly well with a consistent positive or negative bias (Fig. 9; *cf* Table 1, rows 1–8), while the roughly constant 3-year-averaged estimates do not track or exhibit a fixed bias (Fig. 10; *cf* Table 1, rows 9–12). While errors in estimated wave statistics based on monthly data remain relatively constant, they vary more widely when 3-year-averaged estimates are used. The greatest errors occur in 3-year-averaged estimates of wave height, and these appear to have the greatest effect on model skill scores. Indeed, the

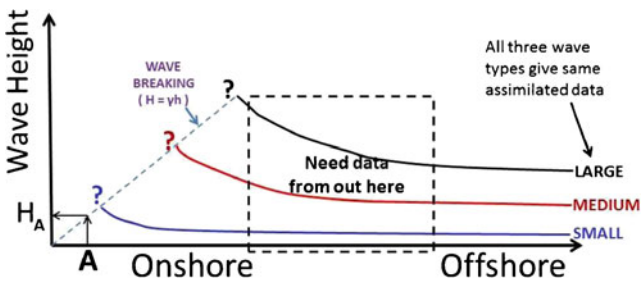

Fig. 7 A problem with spectral assimilation in the surf zone. Because the height of breaking waves is a linear function of the depth at all surf-zone locations (i.e., $H=\gamma h$), it is not possible to determine a given wave's maximum breaking height (or, consequently, its offshore height and other statistics) from a single spectral assimilation (e.g., H_A) at a location in the surf zone (e.g., A)

variations in the 3-year-averaged skill scores of Fig. 8 (bottom panel) qualitatively resemble the variations in observed wave height in Fig. 10 (top panel). (Note, however, that these large wave height errors are not apparent from the correlation and bias values for H_s in Table 1, rows 9–12.) In contrast, monthly observed values of $\overline{T}$, $\overline{\theta}$, and σ_θ are relatively constant over the 12 months, so that 3-year-averaged estimates of these values are often as good or better than monthly estimates. Although their variations are not highly correlated with observed variations, biases remain small.

The low 3-year-averaged skill scores in June and July coincide with a 65 % overestimate of observed wave height by the model. For these 2 months, boundary waves at Duck tend to be significantly smaller ($H_s \sim 0.5\ m$) than average ($H_s \sim 0.86\ m$). This severe mismatch between estimated and observed spectra causes the skill score to drop to zero. As mentioned in Section 3, the RMS skill score is sensitive to differences in total energy and is most effective when they are small. It will quickly decline if these differences become substantial, particularly when observed wave heights are smaller than estimates. Other than in these 2 months, skill scores for the 3-year-averaged cases are generally higher than those for the monthly estimate cases. This contrasts with the lower correlation values but is consistent with the lower bias values obtained for the 3-year-averaged cases in Table 1.

Although high sensitivity assimilation locations do consistently outperform low sensitivity ones in these results, the influence of assimilations from the offshore accessible region on spectral estimates at the Duck LOI is generally small. In Fig. 9, the post-assimilation spectral statistics (thin solid and dashed lines) remain consistently above and below the observed values, never correcting enough to match them. In Fig. 10, post-assimilation statistics deviate only slightly from their (constant) pre-assimilation values, indicating that the

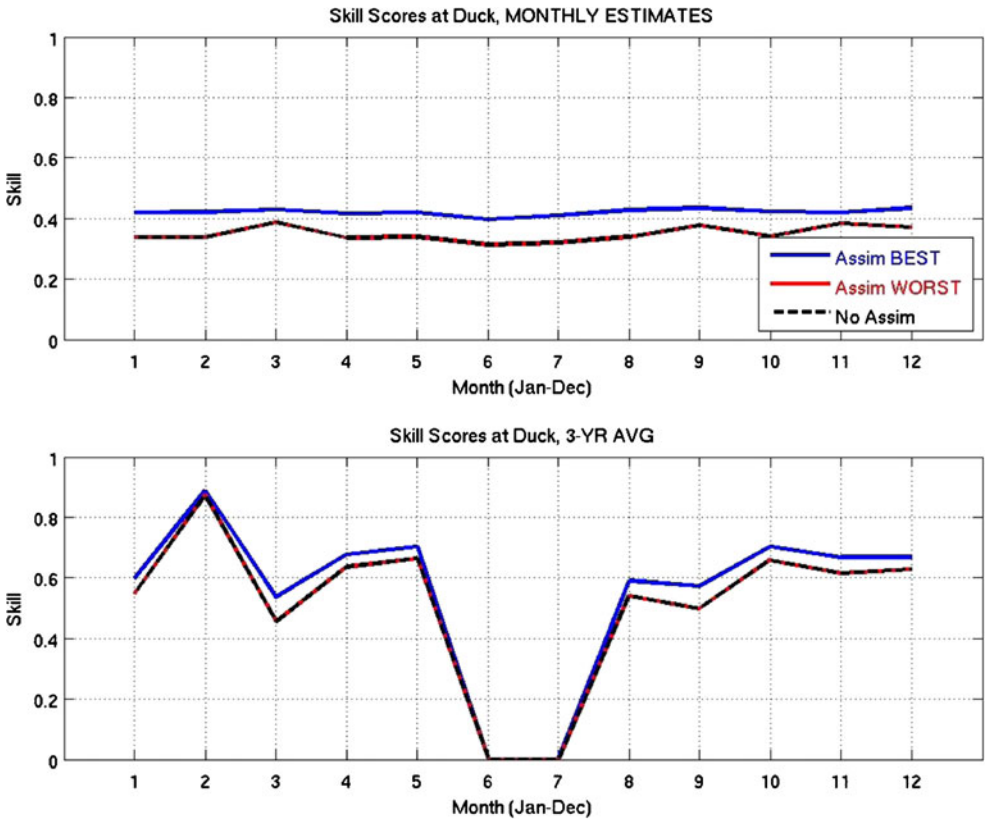

Fig. 8 Skill scores at inaccessible location of interest for Duck site for each of 12 months, based on assimilation from "best" and "worst" alternate locations as well as "no assimilation" (model-only) case. For *top panel*, model "estimate" boundary conditions were created using mean monthly "observed" boundary H_s ±10 %, $\overline{T}$ ±10 %, and $\overline{\theta}$ ±10° ("high" and "low" error cases are combined). For *bottom panel*, a single model estimate was used for all cases, with boundary conditions based on 3-year-averaged H_s, $\overline{T}$, and $\overline{\theta}$. Score is computed by comparing post-assimilation estimated spectrum to observed spectrum at the LOI

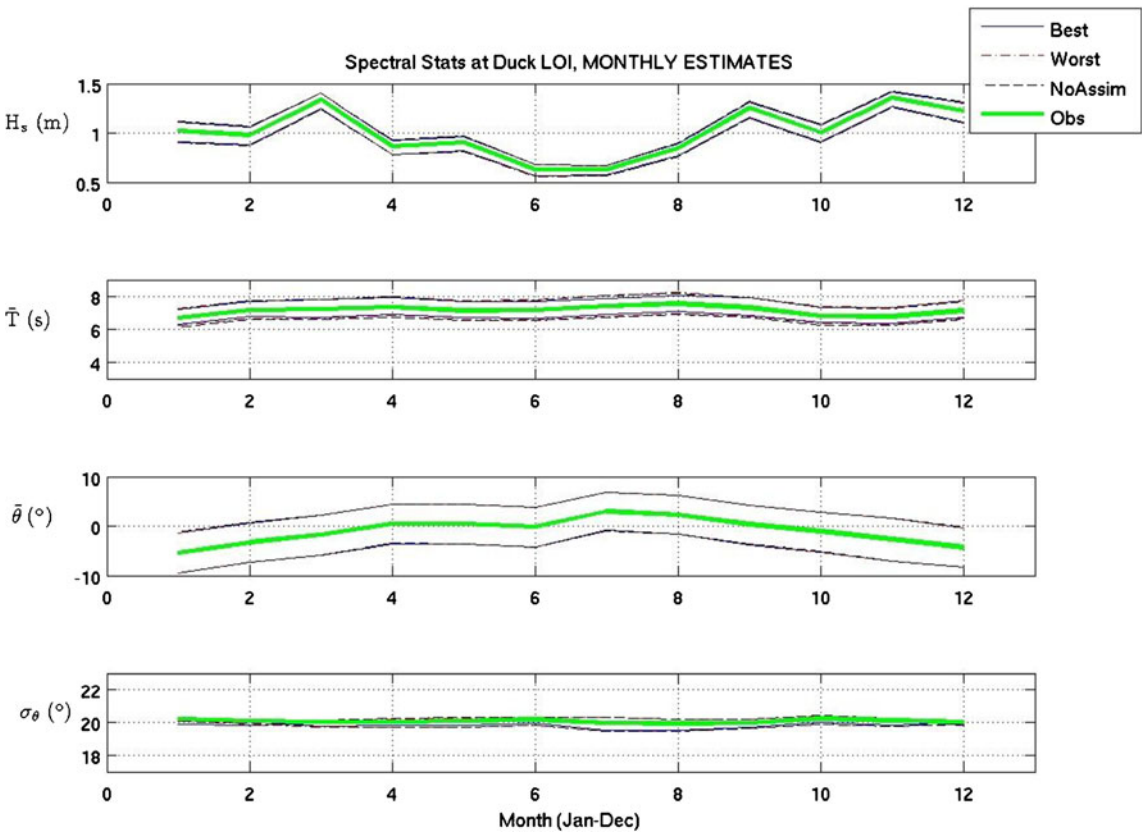

Fig. 9 Spectral statistics at Duck LOI based on monthly estimates, plotted versus month of year. *Solid green line* is observed values while others represent estimates based on best, worst, and non-assimilation cases for both low and high waves

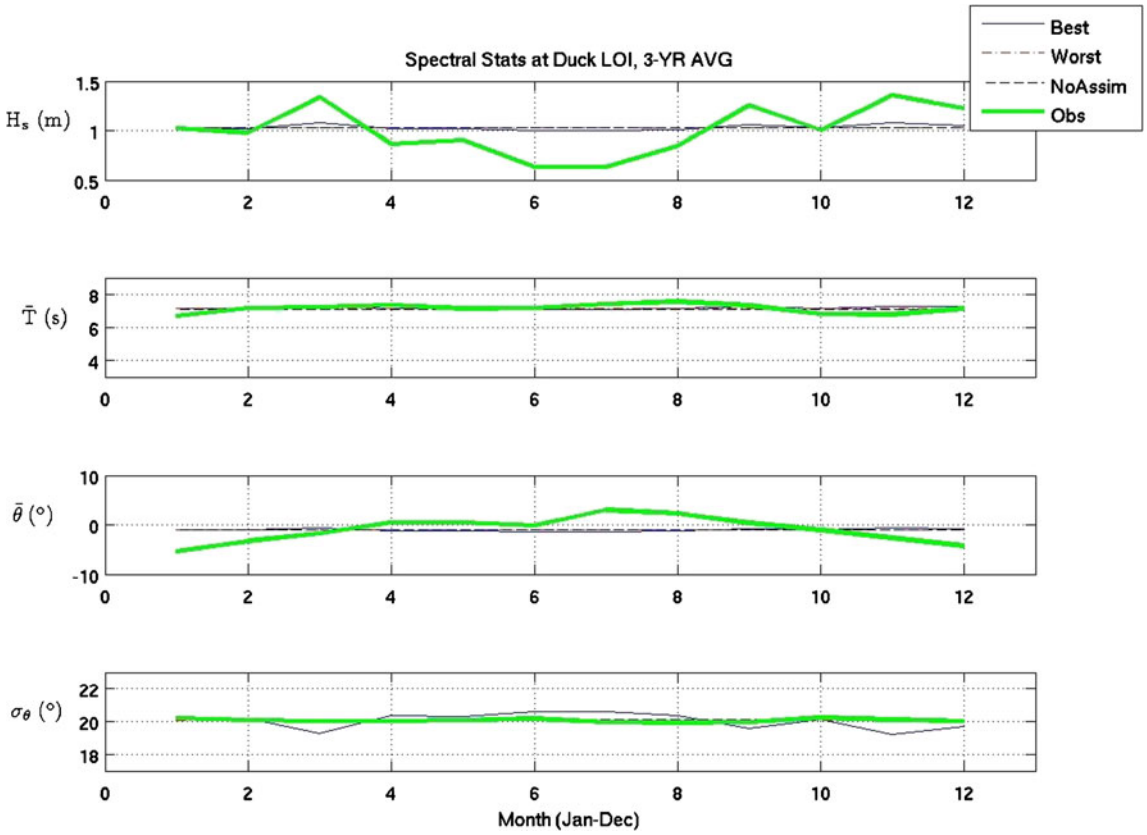

Fig. 10 Spectral statistics at Duck LOI based on 3-year-averaged values for model estimates, plotted versus month of year. *Solid green line* is observed values while others represent estimates based on best, worst, and non-assimilation cases

final estimated LOI spectra were only slightly changed by the assimilation.

Why is the corrective effect at the LOI not larger? For a given assimilation, the spectral correction estimated by SWANFAR is focused on the assimilation point; the corrective effect increases in magnitude as one approaches that point and diffuses/declines as one moves away (e.g., Fig. 11, right panel). This tapering is further accelerated in the surf zone, where all spectra gradually tend toward the same shape (i.e., energy proportional to $k^{-5/2} \rightarrow k^{-4/3} \rightarrow f^{-2}$ as water depth decreases; *cf.* Kaihatu et al. 2007). At the offshore assimilation locations, in contrast to the LOI, the final spectral estimate is much more accurate, achieving skill levels that range from 0.94 to 0.99 for all cases in this test set. This tapering correction effect is primarily a consequence of the unit-valued error covariance used in the present version of SWANFAR (Section 2.3). As mentioned earlier, a five-dimensional covariance function is currently being developed for adjoint output within the SWANFAR assimilation system; when it is complete, spectral corrections will extend considerably farther from the assimilation locations.

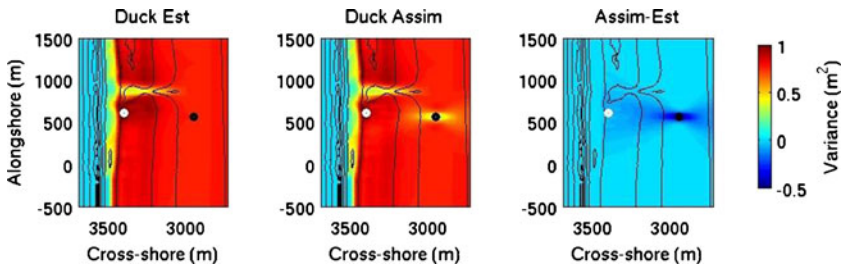

Fig. 11 Sample assimilation result for Duck ("avg" best case, month of July), displaying total spectral variance values at each domain location with bathymetry contours superimposed. *Left panel* shows initial model estimate, *middle panel* shows post assimilation estimate, and *right panel* shows difference (i.e., correction). *White circle* indicates LOI and *black circle* indicates assimilation location. Correction magnitude is greatest at assimilation location, then rapidly declines with distance

Table 2 Correlations (and biases) of modeled vs. observed spectral statistics at LOI for San Diego, CA, simulations, test set #2

	Est. Type	Assim. Locn	H_s (m)	$\overline{T}$ (s)	$\overline{\theta}$ (°)	σ_θ (°)
1	Low monthly	Best	0.88 (−0.02)	0.99 (0.07)	0.95 (−0.39)	0.90 (−0.58)
2		Worst	0.82 (−0.03)	0.99 (−0.03)	0.99 (−0.23)	0.93 (−0.46)
3		LOI	0.90 (−0.01)	0.96 (0.17)	0.99 (0.29)	0.97 (−0.16)
4		No Assim	0.78 (−0.02)	0.99 (−0.16)	0.96 (−0.11)	0.90 (−0.03)
5	High monthly	Best	0.85 (0.00)	0.99 (0.19)	0.94 (0.20)	0.91 (−0.21)
6		Worst	0.78 (0.00)	0.99 (0.28)	0.99 (−0.10)	0.93 (−0.28)
7		LOI	0.98 (−0.01)	1.00 (0.19)	1.00 (0.12)	0.98 (−0.44)
8		No Assim	0.73 (0.01)	0.99 (0.13)	0.94 (−0.13)	0.92 (0.17)
9	3-Yr Avg	Best	0.92 (−0.07)	1.00 (0.20)	0.94 (0.78)	0.94 (−1.8)
10		Worst	0.91 (−0.08)	0.99 (0.27)	0.87 (0.13)	0.71 (−1.6)
11		LOI	1.00 (−0.03)	1.00 (0.17)	1.00 (0.12)	1.00 (−0.92)
12		No Assim	– (−0.08)	– (0.12)	– (0.95)	– (−1.5)

Bias units are shown at top of each data column

4.2 Test set #2: San Diego LOI

Correlations and biases of spectral statistics (H_s, $\overline{T}$, $\overline{\theta}$, and σ_θ) at the San Diego LOI are presented in Table 2, in the same format as seen earlier in Table 1. Skill scores for each model estimate type and assimilation location are presented in Fig. 12, in the same format as Fig. 8. For H_s, data assimilation from the "best" alternate location always results in better correlations, lower bias, and higher skill scores than assimilation from the "worst" alternate location or for the non-assimilation case. For $\overline{T}$, $\overline{\theta}$, and σ_θ, results are somewhat more mixed, but correlations are generally high and biases low to moderate for nearly all scenarios. In this test set, unlike the preceding one, assimilation at the deeper water LOI now

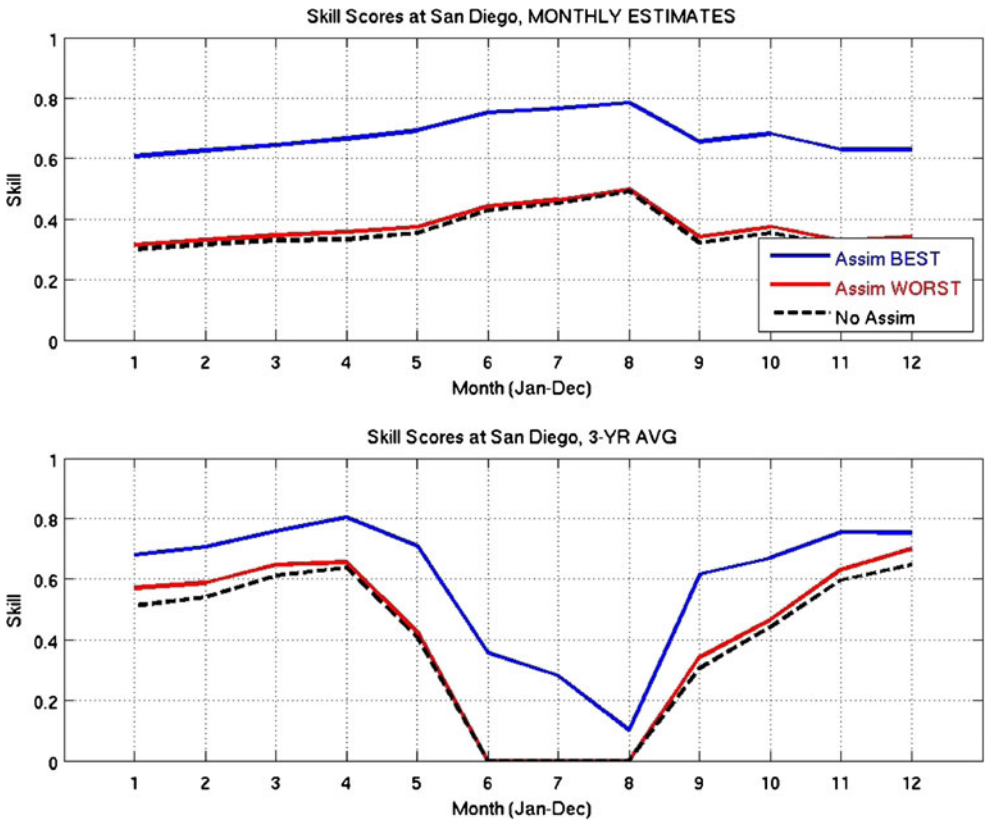

Fig. 12 Skill scores at San Diego site, plotted versus month of the year. Same format as Fig. 8

gives the highest correlations for nearly every test. While correlations of wave heights are somewhat lower at San Diego than at Duck, corresponding biases at San Diego are consistently smaller. Correlation values for $\overline{T}$ and $\overline{\theta}$ are comparable (with smaller biases) in the monthly cases and higher (with larger biases) for the average case at San Diego. Correlations for directional spreading at the deeper San Diego LOI are also generally better than those seen for the shallower Duck domain, but biases are generally larger at the deeper site. Complicating nonlinear effects such as wave breaking may have less effect on wave heights in the deep water, but refraction over the larger, more irregular bathymetry may add greater variability to spectral directions.

Skill scores at the San Diego LOI based on assimilation at the best and worst alternate locations are plotted as time series for months Jan–Dec in Fig. 12, with non-assimilation skill scores included for comparison. As with the Duck simulations (test set #1), the highest skill values are consistently seen from assimilation at the best alternate location, while values based on the worst alternate location are generally little better than the non-assimilation case. For this test set, model skill scores based on the best alternate location are significantly higher than the other two sets of skill scores. In contrast to the Duck domain, the larger scale and significant non-uniformity in the alongshore depth of the San Diego bathymetry allows for greater variations in spectral energy among the assimilation locations and less uniform conditions at the LOI. Similar to Duck, a significant dip in skill scores again occurs for the 3-year-averaged cases in the summer (Fig. 12, bottom panel).

Model statistics for the monthly and 3-year-averaged model estimates at the San Diego LOI are presented in Figs. 13 and 14, respectively. Again, statistics for monthly estimates follow observed statistics fairly closely, while the 3-year-averaged statistics exhibit greater differences. In this case, these greater differences are matched by larger biases for 3-year-averaged data in Table 2 (rows 9–12). In contrast to Duck, there is somewhat more variability in the post-assimilation statistics for the monthly and 3-year-averaged cases, indicating a greater effect of the assimilation on the spectral estimate at the LOI. In this deeper water environment, the post-assimilation spectral correction again is greatest at the offshore assimilation location and tapers toward the shoreline (e.g., Fig. 15, right panel). However, because this LOI is located outside the surf zone, the correction's influence has not completely dissipated and has a larger effect on estimated spectra there. This effect is particularly beneficial for assimilations at the "best" alternate locations, as indicated by the skill scores of Fig. 12.

The low skill values for the 3-year-averaged estimates in June–August coincide with a 50–80 % overestimate of observed $\overline{T}$ by the model (Fig. 14, second panel) and a 25 % underestimate of directional spread (Fig. 14, bottom panel). For these 3 months, the mean approach angles for observed boundary spectra were 320–330° (relative to true N), while the 3-year-averaged value used for the estimates was 298°. Observed boundary wave heights and mean periods were roughly the same ($H_s \approx 1\ m$, $\overline{T} \approx 12\ s$). At the LOI during these months, observed spectra have less energy at their peak frequency and more energy at higher harmonic frequencies than the

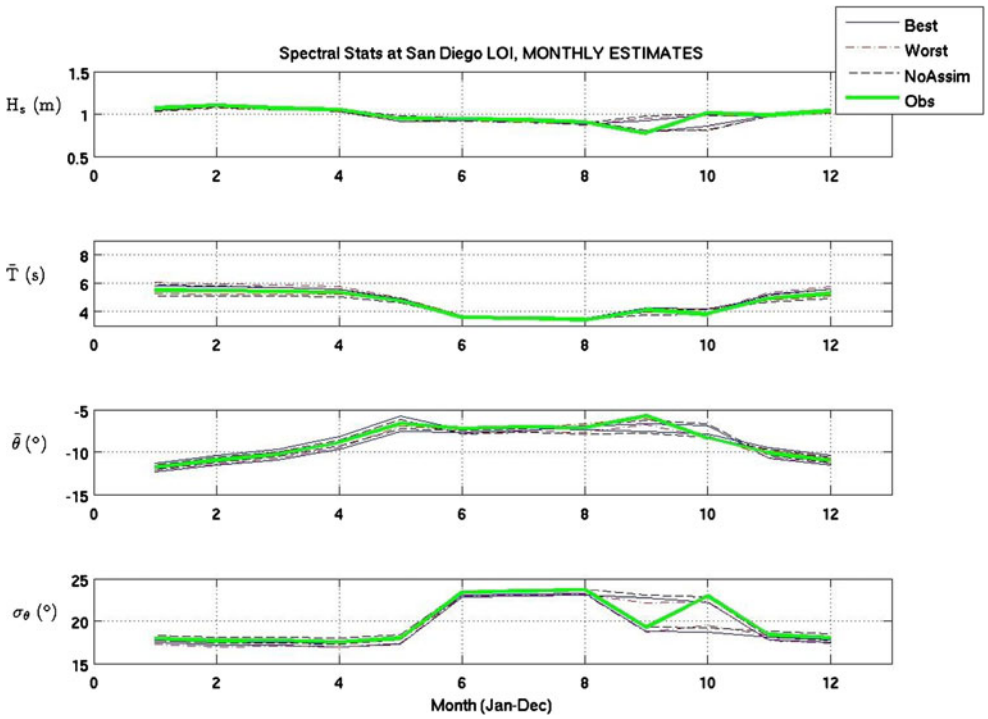

Fig. 13 Spectral statistics at San Diego based on monthly estimates, plotted versus month of year. *Solid green line* is observed values while others represent estimates based on best, worst, and non-assimilation cases for both low and high waves

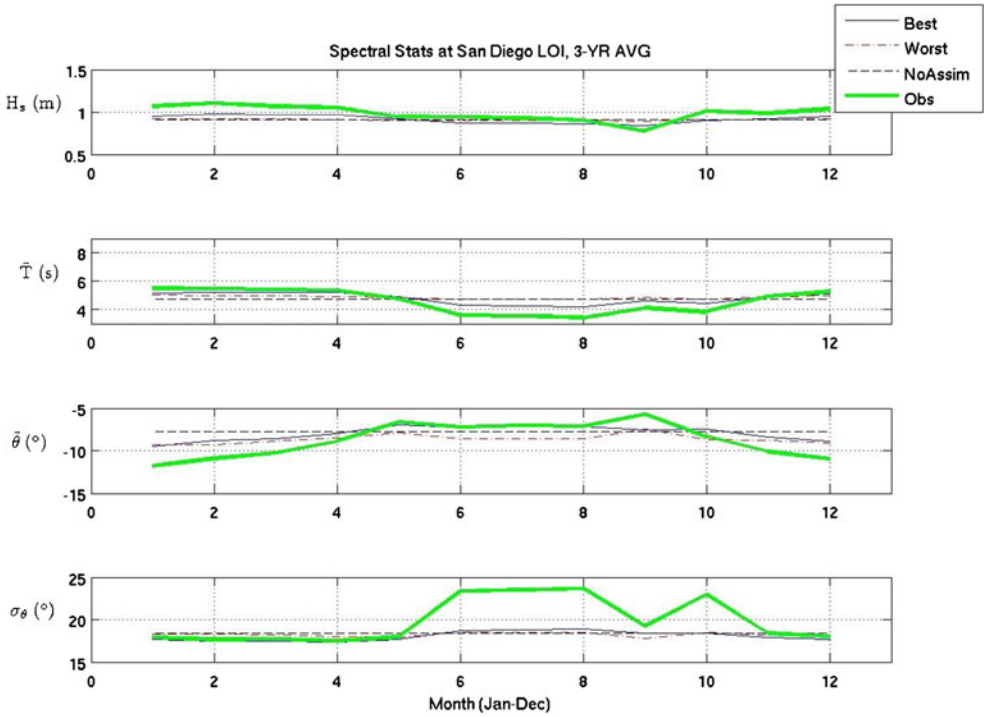

Fig. 14 Spectral statistics at San Diego based on 3-year-averaged values for model estimates, plotted versus month of year. *Solid green line* is observed values while others represent estimates based on best, worst, and non-assimilation cases

corresponding post-assimilation estimated spectra (e.g., Fig. 16). This suggests that the highly oblique offshore approach of the observed waves resulted in a longer, more curved path to the LOI and more significant nonlinear energy transfers. While estimated and observed wave heights are similar in these months, the model–data mismatch in energy levels at both peak and harmonic frequencies is substantial, as clearly reflected by the skill score.

4.3 Test set #3: low-resolution bathymetry

The limitations imposed on bathymetry resolution in test set #3 appear to have a relatively small effect on sensitivity map shape at both the Duck and San Diego sites. As an example, sensitivity map contours based on mean January conditions at the two sites are displayed in Fig. 17. Contours are for "high" and "low" resolution bathymetry are very similar, and optimal alternate instrument locations are also nearly the same.

Assimilations conducted using the low-resolution bathymetry at Duck, initialized with 3-year-averaged conditions, result in skill values at the LOI that are nearly the same as those obtained with the high-resolution data (Fig. 18, top panel). For low-resolution bathymetry at San Diego, skill scores vary considerably from those obtained with the high-resolution bathymetry (Fig. 18, bottom panel). The low-resolution scores actually exceed the high-resolution scores for several months at both sites. For this site, assimilation from highly correlated locations consistently and significantly improves model skill at the LOI relative to assimilation from poorly correlated ones.

4.4 Test set #4: region of interest

This test set examines model performance in a region of interest using 3-year-averaged boundary data with assimilation at a single best offshore alternate location. Results are summarized for the 3-year period at Duck and San Diego with

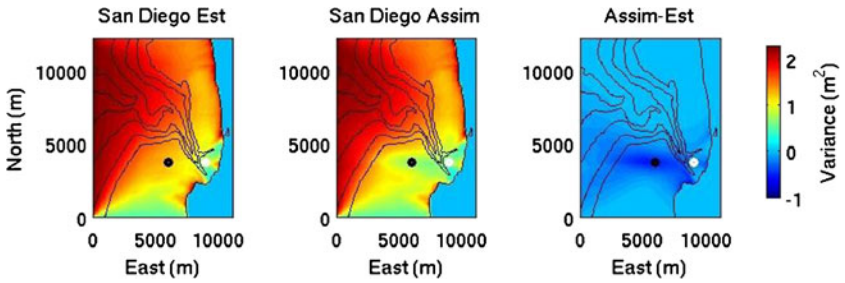

Fig. 15 Sample assimilation result for San Diego ("avg" best case, month of August), displaying total spectral variance values at each domain location with bathymetry contours superimposed. Same format as Fig. 11. Correction magnitude again is greatest at assimilation location, then declines with distance, more gradually than at Duck. Shape of correction area is more clearly affected by bathymetry

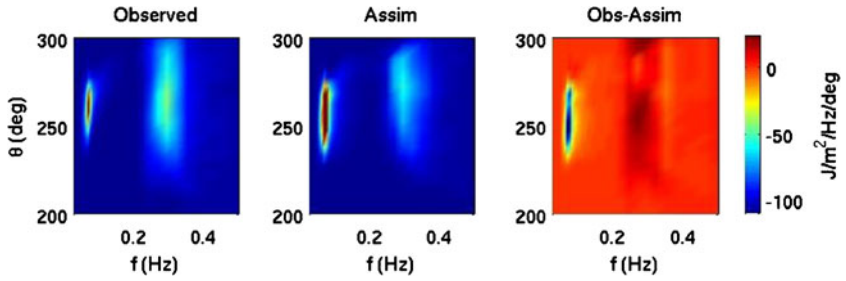

Fig. 16 Sample observed (left) and post-assimilation estimated (center) spectrum at LOI for San Diego ("avg" best case, month of August), with difference spectrum in *right panel*. Observed and assimilation spectra have same color scale. *Colorbar* applies only to difference spectrum. Comparison of energy levels at peak and higher harmonic suggests that observed waves experienced greater nonlinear energy transfers while traveling to the LOI

a color-coded box of mean skill scores for each site. Scores in the ROI box at Duck range from 0.3 to 0.5 (Fig. 19, left panels, box "TS#4"), which is somewhat lower than most skill values obtained at the single Duck LOI using 3-year-averaged boundary spectra (Fig. 8, bottom panel). Scores for the ROI are consistently lower in the vicinity of the Duck pier (~500 m alongshore, southern end of ROI), which likely introduced additional error into adjoint and forward estimates.

The skill score box can be utilized to guide field experiment planners to the best estimated locations in the selected region. Those preparing for such activities at Duck would be advised to focus on more northern portions of the ROI in order to maximize the likelihood of accurate wave predictions incorporating data assimilation at offshore locations.

At San Diego, mean ROI skill scores range from roughly 0.4 to 0.7 when assimilating from the best alternate offshore location (Fig. 19, right panels, box "TS#4"). These scores are in roughly the same range as those seen for the single LOI (Fig. 12, bottom panel). The model achieves consistently higher skill along the southern end of the ROI. As mentioned earlier, the waves at this site generally tend to approach from the northwest, with mean monthly wave directions of 280°–330° (relative to true North) and a 3-year-averaged direction of roughly 300°. Waves reaching the northern end of the ROI thus will generally pass over larger sections of the variable canyon bathymetry and will need to refract more severely than waves that reach the southern end. Existing errors in boundary spectra will be more greatly magnified over the circuitous northern route, leading to lower overall skill scores. Those planning activities or deployments in this region would be advised to select locations nearer to the southern canyon (i.e., La Jolla Canyon) in order to be more confident in SWAN-based wave estimates.

4.5 Test set #5: region of interest with two-point assimilation

The ROI distributions of model skill scores based on assimilation from two separated offshore locations are illustrated in the boxes labeled "TS#5" in Fig. 19. The skill scores in each ROI are distributed in roughly the same manner as those in the preceding set of tests, in which just one alternate assimilation location was selected from a single, larger region. It might be

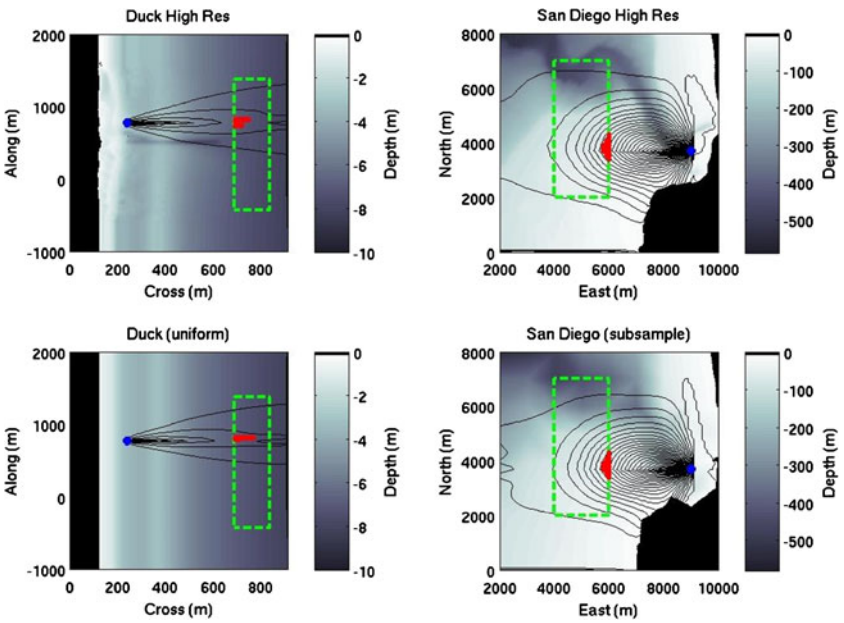

Fig. 17 Sensitivity map contours on high (*top*) and low (*bottom*) resolution bathymetry at Duck (*left panels*) and San Diego (*right panels*) sites. Accessible regions (*green boxes*) and optimal alternate assimilation locations (*red asterisks*) are also shown for each case

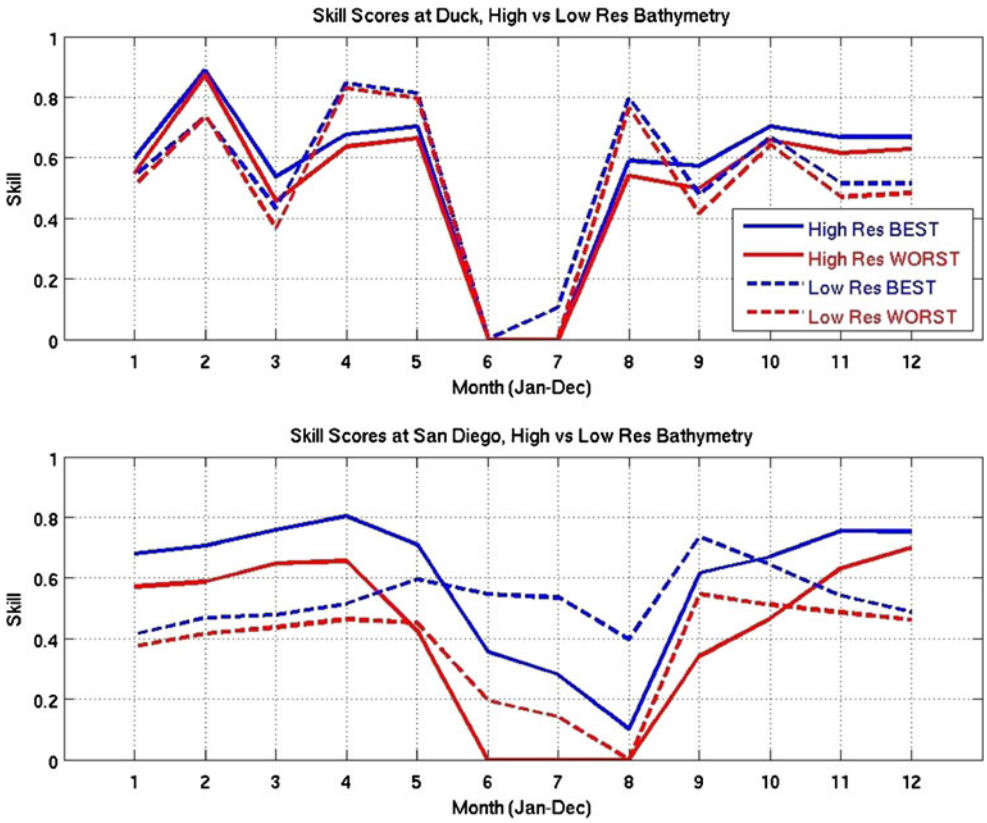

Fig. 18 Spectral skill values at Duck (*top*) and San Diego (*bottom*), including assimilations from "best" (*blue*) and "worst" (*red*) alternate locations for high-resolution (*solid*) and low-resolution (*dashed*) bathymetry. Model estimates are initialized with 3-year-averaged conditions for 2010–2012

expected that two-point assimilations would improve performance in these cases, as assimilating more data generally allows a model to better approximate actual conditions. However, skill scores at Duck are just comparable to those from the single-point assimilation of test set #4, while scores for the San Diego ROI are consistently lower.

These results illustrate the importance of sensitivity when selecting offshore assimilation sites. At Duck, the highest sensitivity contours extend directly seaward from the LOI/ROI, roughly normal to the shoreline (e.g., Fig. 17, top left). At San Diego, these contours generally reach west-northwestward, slightly below the southern edge of the main

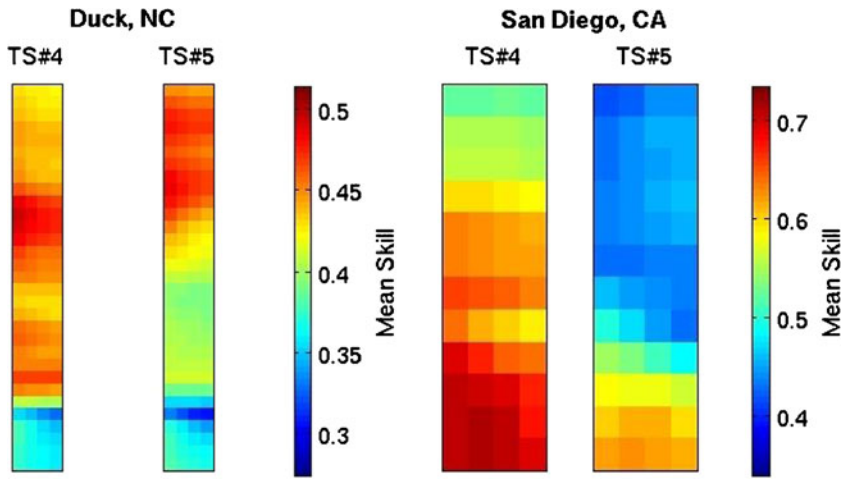

Fig. 19 Mean skill scores plotted as colormaps for each region of interest, using high-resolution bathymetry at Duck and San Diego sites (i.e., *box areas* correspond to blue rectangles in Fig. 5 and 6, respectively). Results are based on data assimilation at the "best" alternate instrument locations for test set #4 ("TS#4") and test set #5 ("TS#5"). Alternate instrument locations are determined from adjoint-generated sensitivity maps initialized with spectral data from all points within the specified ROI. Model estimates are initialized with 3-year-averaged conditions, while quasi-observations are again generated from average monthly conditions for Jan–Dec, 2010–2012

canyon (e.g., Fig. 17, top right). At both sites, the larger single offshore accessible regions used in test sets #1–4 consistently intercept the highest sensitivity contours, allowing the selection of more highly correlated assimilation locations that significantly influence and improve model estimates for the LOI/ROI. In the present test set, the two smaller accessible regions (Fig. 6) do not overlap these high sensitivity contours at either site. Spectra from locations within these areas are less well correlated with those in the ROI and have a smaller corrective effect on model estimates there, even when two locations are included in the assimilation.

Thus, availability of more than one accessible alternate region, allowing instrument deployment and assimilation at multiple separated points, is not guaranteed to improve model skill for a nearshore ROI. If the multiple alternate regions are not in locations of high sensitivity, model skill in the inaccessible ROI may in fact decline. The two-point assimilation of this test set appears to just compensate for the lower sensitivity at Duck, while it is not enough to do so at San Diego. Field teams facing these scenarios would be encouraged to identify additional offshore assimilation locations with higher sensitivity levels in order to improve on the spectral estimates in the ROI. If such locations were not available, the teams would again be advised to focus activities along the southern end of the ROI in San Diego, but farther north in the ROI at Duck, expecting to have somewhat less confidence in the accuracy of their post-assimilation spectral estimates.

5 Discussion and conclusions

Adjoint-based sensitivity maps can be an important tool for improving spectral wave estimates at inaccessible nearshore locations using assimilated data from elsewhere in the domain. As illustrated in the preceding test sets, the maps make it possible to select assimilation locations where data will most effectively increase model accuracy at an inaccessible LOI.

When dealing with inaccessible locations, the sensitivity maps methodology is limited to using only *model-estimated* spectra for initializing the adjoint. This means that the accuracy of adjoint-generated sensitivity map contours and resulting optimal alternate assimilation locations is dependent on the pre-assimilation accuracy of forward model "first guesses" of conditions at the LOI. For maps based on H_s, the preceding test results suggest that map shape is not very sensitive to these guesses, particularly in larger-scale, deeper-water domains such as the San Diego bathymetry used here (e.g., Figs. 13 and 17). Variations in first-guess mean direction $\overline{\theta}$ can shift the orientation of map contours somewhat, but because these variations tend to be smaller for most nearshore LOIs, this generally causes only a small shift in optimal assimilation locations offshore. For deeper water LOIs, variations in first-guess H_s or mean period $\overline{T}$ have little effect on map shape or optimal assimilation locations.

The RMS skill score is an alternative, concise tool for evaluating model accuracy that directly compares each modeled and observed spectral bin, unlike traditional correlations and biases of statistics such as significant wave height or mean period. By design, the skill score is affected by differences in total energy as well as the distribution of energy with frequency and/or direction. Skill scores for a single location may be plotted versus time, or skill scores for a region may be presented as a colormap to show the variation of model performance with location. Skill scores alone only provide an "overview" of model performance. A low skill score can indicate that model estimates are poorly correlated , biased, or both; the score by itself does not identify the sources of model error. To fully understand why a given score is "high" or "low", it is still necessary to examine and compare other modeled and observed statistics and sometimes to look directly at the spectra themselves. However, as illustrated in Fig. 19, the skill score map can often help even novice users to quickly and easily identify areas of optimal model performance in a given region.

The five sets of tests described in the preceding sections investigated different factors influencing sensitivity-map-guided spectral wave data assimilation for inaccessible nearshore locations. Test sets #1 and 2 compared model estimates to quasi-observations at Duck and San Diego LOIs using correlations and biases of statistics (H_s, $\overline{T}$, $\overline{\theta}$, and directional spread σ_θ) as well as skill scores, with assimilation from a single alternate offshore location in each case. Skill scores were reduced when estimates were poorly correlated or biased, with the lowest skill values occurring for model overestimates of low energy observed spectra. The contrast between results at the two sites emphasized the important spectral effects of refraction and wave breaking inside the surf zone and several limitations imposed on the SWANFAR spectral wave assimilation system in shallow water. At present, it is not possible to create a fully consistent adjoint to the wave breaking process in the surf zone. The adjoint model has no information on where the waves are breaking (e.g., Fig. 7); that information is only available to the forward model. For this reason, data assimilation based solely on spectra from inside the surf zone will not be accurate. Also, spectral corrections presently estimated by SWANFAR are largest (and most accurate) at the assimilation location; their effectiveness and accuracy decline with distance from it. A 5-D covariance function is being developed to extend this range in both space and time dimensions.

Test set #3 investigated the effects of bathymetry resolution on sensitivity map shape and alternate assimilation locations. At Duck, the sensitivity map was almost unchanged by the use of an alongshore uniform bathymetry. The nearly shore-normal approach angle of shoaling waves at the shallow water

LOI resulted in similarly shore-normal sensitivity contours extending seaward from it. Neglecting the scoured region under the Duck pier slightly modified the outer contours of the sensitivity map for the alongshore uniform beach but had only a small effect on optimal offshore assimilation locations (Fig. 17, left panels). Oscillations in monthly skill scores from assimilation on the approximate bathymetry roughly paralleled those from the original measured bathymetry (Fig. 18, top). The relative scarcity of rip channels and other surf-zone irregularities near the LOI made alongshore uniform bathymetry a reasonably effective approximation for creating sensitivity maps on this beach.

Skill scores for the low-resolution bathymetry at San Diego varied more greatly from the original values (Fig. 18, bottom), but this did not appear to be due to the sensitivity maps. Scores were significantly higher for assimilation from "best" alternate locations, suggesting that it may be even more important to use data from highly correlated regions when a larger scale, deeper bathymetry is poorly resolved. Sensitivity contours and optimal assimilation locations were again similar to those obtained with higher resolution (Fig. 17, right panels). Here, however, the bathymetry changed more substantially (e.g., LOI depth shifted from 25 to 54 m) and waves traveled longer distances through the modified domain than they did at Duck. These combined factors had a stronger effect on the spectral estimates of the assimilation system than occurred for the smaller-scale, more uniform Duck domain. The sensitivity maps themselves, which are based on a single representative statistic (H_s) at each location, appear to be relatively insensitive to changes in bathymetry resolution. This suggests that optimal alternate assimilation sites may be identified with outdated or approximate depth data. However, subsequent use of assimilated data to estimate spectra throughout the domain will likely require highly resolved, accurate bathymetry.

Test sets #4 and 5 compared model performance in a nearshore ROI based on assimilation at one and two offshore locations, respectively. In both tests, the color-coded skill score box displayed clear variations in model performance within the ROI. Such skill maps can be used to guide nearshore field researchers to areas of greatest model accuracy. Somewhat counterintuitively, the two-location assimilation resulted in lower skill scores than the one-location assimilation at San Diego and gave just comparable skill scores at Duck (Fig. 19). The important conclusion in these cases is that the effectiveness of a given assimilation location is strongly influenced by its sensitivity with respect to the LOI or ROI. While the single larger alternate assimilation regions of test set #4 each overlapped with high sensitivity contours from their respective ROIs, the multiple regions of test set #5 did not. Adding more offshore assimilation locations with poor sensitivity does not guarantee better model estimates in inaccessible nearshore ROIs.

In summary, the best alternate assimilation locations will be along contours of highest sensitivity, but not all such contours will be accessible along many coastlines. The effectiveness of a given assimilation can drop rapidly as sensitivity decreases. Whenever possible, the accessible alternate offshore region must be extended to intersect the highest sensitivity contours for a given nearshore LOI or ROI. Generally speaking, optimal observation points will be as close as possible to the LOI/ROI and in the "upwave" direction (i.e., the direction of propagation of maximum adjoint wave energy). As wave climate changes throughout the year, this direction may vary. If this variability causes significant shifts in optimal assimilation locations, investigators should consider deploying multiple instruments at seasonal high-sensitivity points to maximize year-round model performance.

Acknowledgments This work was funded by a Karles Fellowship grant from the Naval Research Laboratory and through the 6.2 NRL Core Project "Improving Wave Predictions Using Data Assimilation in a Spectral Wave Model", Program Element #0602435N. The authors thank the two anonymous reviewers for their many constructive suggestions and Dr. Hans Ngodock for his consistently helpful advice, input, and tutelage.

References

Booij N, Ris RC, Holthuijsen LH (1999) A third-generation wave model for coastal regions: 1. Model description and validation. J Geophys Res 104(C4):7649–7666

Bugnion V, Hill C, Stone PH (2006) An adjoint analysis of the meridional overturning circulation in an ocean model. J Clim 19: 3732–3750

Erwig M, Fu Z, Pflaum B (2007) Parametric fortran: program generation in scientific computing. J Softw Maint Evol Res Pract 00:1–7

Field Research Facility (2013) Field Data Collections and Analysis Branch, US Army Corps of Engineers, Duck, North Carolina (http://www.frf.usace.mil)

Flampouris S, Veeramony J, Orzech MD, Ngodock HE (2013) Validation of a wave data assimilation system based on SWAN. Geophys Res Abst, (15), EGU2013-5951-1, EGU General Assembly 2013

Hill C, Bugnion V, Follows M, Marshall J (2004) Evaluating carbon sequestration efficiency in an ocean circulation model by adjoint sensitivity analysis. J Geophys Res 190, C11005. doi:10.1029/2002JC001598

Kaihatu JM, Veeramony J, Edwards KL, Kirby JT (2007) Asymptotic behavior of frequency and wave number spectra of nearshore shoaling and breaking waves. J Geophys Res 112, C06016. doi: 10.1029/2006JC003817

Moore AM, Arango HG, Di Lorenzo E, Miller AJ, Cornuelle BD (2009) An adjoint sensitivity analysis of the southern California current circulation and ecosystem. J Phys Oceanog 39:702–720

National Data Buoy Center (2013) Bldg. 3205, Stennis Space Center, MS (http://www.ndbc.noaa.gov)

Ngodock HE, Carrier MJ (2013) A weak constraint 4D-Var assimilation system for the Navy coastal ocean model using the representer method. In: Park SK, Xu L (eds) Data assimilation for atmospheric,

oceanic and hydrologic applications, vol II. Springer, Berlin. doi:10.1007/978-3-642-35088-7_15

Orzech MD, Veeramony J, Ngodock HE, Flampouris S (2012) A 4DVAR data assimilation system for SWAN. In: Wiedling J, Einsporn MH (eds) Recent Impulses to Marine Science and Engineering—Proceedings of 3rd International YOUMARES Conference (Lübeck, Germany), p 52

Orzech MD, Veeramony J, Ngodock H (2013) A variational assimilation system for nearshore wave modeling. J Atmos Ocean Technol 30: 953–970. doi:10.1175/JTECH-D-12-00097.1

Tolman HL (2009) User manual and system documentation of WAVEWATCH III™ version 3.14. NOAA / NWS / NCEP / MMAB Technical Note 276, 194 pp+Appendices

Veeramony J, Orzech M, Ngodock HE, Flampouris S (2012) Towards an operational nearshore wave data assimilation system (Invited), Abstract OS21F-02 presented at 2012 Fall Meeting, AGU, San Francisco, Calif., 3–7 Dec

Zhang WG, Wilkin JL, Levin JC, Arango HG (2008) An adjoint sensitivity study of buoyancy- and wind-driven circulation on the New Jersey inner shelf. J Phys Oceanog 39:1652–1668