Impact of sea surface salinity assimilation on coupled forecasts in the tropical Pacific

Eric Hackert,1 Joaquim Ballabrera-Poy,2 Antonio J. Busalacchi,1 Rong-Hua Zhang,1 and Raghu Murtugudde1

Received 6 October 2010; revised 21 February 2011; accepted 3 March 2011; published 14 May 2011.

In this paper, we assess the impact of sea surface salinity (SSS) observations on seasonal variability of tropical dynamics as well as on dynamical El Niño–Southern Oscillation (ENSO) forecasts using a hybrid coupled model (HCM). The HCM is composed of a primitive equation ocean model coupled with a singular value decomposition–based statistical atmospheric model. An Ensemble Reduced Order Kalman Filter (EROKF) is used to assimilate observations to constrain tropical Pacific dynamics and thermodynamics for initialization of the HCM. Rather than trying to produce the best possible operational forecasts, point-wise subsurface temperature (sTz) has been assimilated separately and together with gridded observed sea surface salinity (SSS) from optimal interpolation to more efficiently isolate the impact of SSS. Coupled experiments are then initiated from these EROKF initial conditions and run for 12 months for each month, 1993–2007. The results show that adding SSS to sTz assimilation improves coupled forecasts for 6–12 month lead times. The main benefit of SSS assimilation comes from improvement to the spring predictability barrier (SPB) period. SSS assimilation increases correlation for 6–12 month forecasts by 0.2–0.5 and reduces RMS error by 0.3°C–0.6°C for forecasts initiated between December and March, a period key to long-lead ENSO forecasts. The positive impact of SSS assimilation originates from warm pool and Southern Hemisphere salinity anomalies. Improvements are brought about by fresh anomalies at the equator which increases stability, reduces mixing, and shoals the thermocline which concentrates the wind impact of ENSO coupling. This effect is most pronounced in June–August, helping to explain the improvement in the SPB. In addition, we show that SSS impact on coupled forecasts is more pronounced for the period 1993–2001 than for the period 2002–2007 due to the improved inherent predictability associated with the strong 1997–1998 ENSO. Rather than being the final say for the issue of SSS assimilation, this study should be considered as a necessary first step. Future work is still required to assess issues such as SSS satellite data coverage and the complementary nature of satellite/in situ assimilation. However, these results foreshadow the important positive potential impact that gridded satellite SSS provided by missions such as SMOS and Aquarius/SAC-D will have on coupled model predictions.


1. Introduction

The El Niño–Southern Oscillation (ENSO) phenomena has a significant impact on climate variability throughout the world and so has been the key focus for improving coupled forecasts for the tropical Pacific [e.g., Bjerknes, 1966; Rasmusson and Carpenter, 1982; Wallace et al., 1998]. Assimilation of satellite products such as sea level (SL) from altimetry [e.g., Zheng et al., 2007; Ferry et al., 2007; Chen et al., 1998] and sea surface temperature (SST) measurements [e.g., Zhou et al., 2009; Deng et al., 2009; Zheng et al., 2006] have shown a unique ability to improve short-term forecasts of the coupled system. In addition, assimilation of in situ data such as subsurface temperature and salinity from XBT, CDT, TAO, and more recently Argo has also contributed to improving the initialization of ENSO forecasts for the tropical Pacific [Vidard et al., 2007; Cazes-Boezio et al., 2008].

[3] Ballabrera-Poy et al. [2002] have shown previously that, in the context of statistical prediction of ENSO, sea
surface salinity (SSS) observations, together with SST and SL, have a significant impact on the 6 to 12 month lead predictions of ENSO. A multiple regression analysis for ENSO was conducted that explored the impact of SSS on prediction of equatorial SST. The observed SSS product from the work of Delcroix [1998] was utilized along with COAPS wind stress [Stricherz et al., 1997], SST [Reynolds and Smith, 1994], sea surface height from ocean model hindcasts [Behringer et al., 1998], and precipitation [Xie and Arkin, 1996] minus evaporation [Kalnay et al., 1996]. The results of their multiple regression analysis indicate that SSS would have little impact on ENSO nowcasts. However, statistical ENSO forecasts are improved with lead times of 6–12 months in advance. In particular, key regions of variability included the central basin near 20°S at 6 month leads and the western equatorial Pacific including the South Pacific Convergence Zone (SPCZ) from 9 to 12 month lead times (see Ballabera-Poy et al. [2002, Figures 2–6] for details).

[4] More recently, Yang et al. [2010] used the coupled bred vector technique to assimilate gridded subsurface salinity for the 2006 El Niño. They showed that improvements in the early forecast can be traced to the equatorial central Pacific while the longer-term forecast was related to improved salinity assimilation at the depth of the thermocline in the western Pacific. With the recent launch of the Soil Moisture and Ocean Salinity (SMOS) and the targeted launch date of Aquarius/SAC-D in 2011, important work still needs to be done investigating the impact that surface salinity (SSS) may have on ENSO forecasts.

[5] Whereas the work of Ballabera-Poy et al. [2002] established the impact of observed SSS to statistical ENSO forecasts, here we perform coupled ocean-atmosphere experiments to test the impact of SSS assimilation over a 15 year period. Our methodology uses assimilation of SSS and subsurface temperature to initialize the ocean component of a coupled model as previous studies have established the key importance of initial conditions on coupled forecasts [e.g., Hackert et al., 2007; Moore et al., 2006; Yu et al., 2009]. Several examples of using gridded fields of salinity for assimilation can be found in the literature. At NCEP, empirical orthogonal functions (EOFs) were used to convert observations of subsurface salinity and temperature into dynamic height [e.g., Vossepoel et al., 1999; Maes et al., 2000; Maes and Behringer, 2000; Vossepoel and Behringer, 2000; Vossepoel et al., 2002; Huang et al., 2008]. At ECMWF, the methodology for producing synthetic salinity to fill data voids was based on the conservation of the T-S relationship of the background states and required the vertical shifting of the background salinity [e.g., Cooper and Haines, 1996; Alves et al., 2001; Troccoli et al., 2002; Balmaseda et al., 2007]. In a third general methodology, multivariate statistics from a long simulation of a free model experiment are used to derive corrections to the salinity field. For example, assimilation of subsurface temperature profiles can influence all state variables (including salinity) through EOF basis functions derived from the model [e.g., Borovikov et al., 2005; Sun et al., 2007; Verron et al., 1999; Ballabera-Poy et al., 2007; Yang et al., 2010]. This research utilizes this last approach.

[6] Since NINO3 SST is our validation metric of choice, an experiment that assimilates subsurface temperature profiles (ASSIM_sTz) has been purposely chosen as a baseline since it does not overly constrain SST. Several statistical model studies have shown the importance of including subsurface temperature and thermocline depth for ENSO predictions [Ruíz et al., 2005; Drosdowsky, 2006; Lima et al., 2009]. Since fields such as SL and subsurface salinity may be correlated with SSS, we have chosen not to assimilate all available observed data in order to best isolate and analyze the impact of fields of SSS observations. Therefore, in a separate assimilation experiment, SSS data are assimilated along with sTz data (ASSIM_sTz_SSS). From these experiments, anomalies are formulated with respect to the seasonal cycle (from 1993 to 2007) and then are added to model climatology and used as initial conditions for freely coupled experiments that are run for 12 months starting from the first of each month for 1993–2007. Additional experiments are performed to highlight the contributions of SSS from different geographical regions and time periods.

[7] The organization of this paper is as follows. The data processing and optimal interpolation technique are described in section 2. Section 3 details the ocean model, assimilation technique, and the hybrid coupled model. Section 4 contains the results of this paper highlighting the improvement in the forecasts from SSS assimilation and the spatial distribution of this contribution. In section 5, the impact of SSS assimilation for different periods is discussed and section 6 contains the summary.

2. Data and Data Processing

[8] Temperature and salinity profile data for assimilation are obtained from three sources; Global Temperature-Salinity Profile Program (GTSP), World Ocean Database 2005 (WOD05), and TAO. For the GTSP “best copy” data set [National Oceanographic Data Center, 2006] both real-time data from the Global Telecommunications System (GTS) and delayed mode data received by the NODC are included in a continually managed database which maintains all available subsurface information removing duplicate entries. This data set includes profiles from instruments such as CTD and XBT measurements from ships, TAO buoys, and Argo profiling floats. Only data classified as “good,” “probably good” or “modified” are included in our data set after location, date, gradient, density validation, climatological, and profile consistency tests were performed (http://www.nodc.noaa.gov/GTSP/access_data/gtspbc.html). Daily mean TAO data [McPhaden et al., 1998] are also included in this data set since mooring data were handled inconsistently in both of the other data sets. All TAO data (http://www.pmel.noaa.gov/tao/data_deliv/) except “lower quality” data are included. An additional source of data, the WOD05 [Boyer et al., 2006] data set (http://www.nodc.noaa.gov/OC5/SELECT/dbsearch/dbsearch.html), includes research quality salinity and temperature profile data on standard levels. Extensive quality controls were performed, and only data of “highest quality” (i.e., depth and salinity/temperature error flag = 0) are retained in our database. In addition to the profile data, grids of the seasonal cycle for temperature and salinity provided by World Ocean Atlas (WOA05 [Boyer et al., 2006]) are used to formulate anomalies. Prior to assimilation, each profile is linearly
in order to extend the influence of the limited number of temperature and salinity profiles, and to remove the residual seasonal cycle after anomalies are formulated, the optimal interpolation (OI) technique of Carton and Hackert [1989] is employed to convert point-wise information to grids. This is necessary since salinity information is particularly sparse during the early part (1993–2001) of the simulation period (see Figure 1). For SSS, all available salinity observations within 10 m of the surface were averaged onto 1° × 1° × 1 month bins for the period 1993–2007. Combining SSS data within 10 m as surface observations is a reasonable assumption based on studies that have shown that over 84% of the time, salinity differences between 1 m and 10 m are less than 0.05 psu for the TAO moorings [Henocq et al., 2010].

[10] Gridding the data effectively deals with redundant data from different data sources like GTSPP, WOD05, and TAO. OI was performed on the binned data using decorrelation scales of 15° longitude, 3° latitude, and 1 month matching the values for SST estimated by Meyers et al. [1991]. The resulting SSS product is an excellent surrogate for potential future satellite SSS since estimated SSS errors through Monte Carlo studies (e.g., performing several analyses by randomly withholding 10% data) show estimated error of ~0.22 psu which is comparable to the expected accuracy of Aquarius/SAC-D retrievals at 0.2 psu over 1 month [Le Vine et al., 2007]. The process is repeated to obtain temperature anomaly profiles and the residual mean seasonal cycles (from the differences between the gridded observations and WOA05 gridded data) were removed for both the SSS and subsurface temperature gridded OI data so that the EROKF assimilation procedure would remain stationary.

[11] Plots of OI SSS anomaly data are presented in Figure 2 and show sparse data examples corresponding to the peaks of the 1997 El Niño (Figure 2a) and the 1998 La Niña (Figure 2b) and for better data coverage examples, 2006 El Niño (Figure 2c) and 2007 La Niña (Figure 2d). Note the increasing details of the fields as the coverage improves. The general SSS features for the El Niño, namely the negative SSS anomaly in the western Pacific extending into the intertropical convergence zone (ITCZ), are evident in 2006. As described by Delcroix and Picaut [1998], the big event in 1997 displaces the SSS pattern to the east with negative values extending from 180° and 140°W. In addition, the La Niña patterns of 1998 and 2007 (Figures 2b and 2d) have common aspects such as the positive SSS anomaly in the western Pacific warm pool which extends into the ITCZ. For 2007, the positive SSS anomaly in the South Pacific Convergence Zone (SPCZ) is more prominent.

[12] Observations of SST come from the blended analysis of Reynolds and Smith [1994], an OI of all in situ temperature reports from ships, buoys, and satellite SST. The OI analysis produces a weekly mean gridded product on a 1° × 1° grid and anomalies of these data are used to validate the coupled model results for the Niño3 region (5°N–5°S, 150° W–90°W) for 1993–2008. Anomalies of observations are calculated from the 1993–2007 seasonal cycle climatology.

3. Model and Data Assimilation Description

3.1. Model

[13] The reduced gravity, primitive equation, sigma coordinate model with variable depth oceanic mixed layer is described by Gent and Cane [1989] and Murtugudde et al. [1996]. This ocean model has been validated in a series of simulation studies of circulation in all three tropical ocean basins [Hackert et al., 2001; Murtugudde et al., 1996; Murtugudde et al., 1998] and explicitly accounts for a complete upper ocean hydrology [Murtugudde and Busalacchi, 1998]. The ocean model is coupled to the atmosphere using the atmospheric mixed layer (AML) model of Seager et al. [1995] which calculates surface fluxes of latent and sensible heat, longwave radiation, and evaporation using the bulk transfer formula. This version of the AML formulation requires solar radiation (Earth Radiation Budget Experiment (ERBE)), cloudiness from the NCEP reanalysis [Kalnay et al., 1996], and precipitation from a combination of Xie and Arkin [1998] and Global Precipitation Climate Project (GPCP) [Adler et al., 2003] which are all specified externally. Monthly anomalies of the cloud and precipitation data calculated from the full period of the data set are added to the Interannual Satellite Cloud Climatology Project (ISCCP) seasonal cycle [Rossow and Schiffer, 1991] in order to combine different data sets in a consistent fashion. The AML also calculates the fresh water flux which is treated as a natural boundary condition [Huang, 1993] allowing SSS and SST to vary freely with no explicit relaxation to prescribed Levitus values (unlike other coupled models). Thus, this model has been shown to be particularly appropriate for SSS studies [e.g., Murtugudde and Busalacchi, 1998; Zhang and Busalacchi, 2009].

[14] The OGCM uses the hybrid vertical mixing scheme of Chen et al. [1994] which combines the advantages of the traditional bulk mixed layer of Kraus and Turner [1967] with the dynamic instability model of Price et al. [1986]. This allows simulation of all three major processes of oceanic vertical turbulent mixing: atmospheric forcing is related to mixed layer entrainment/detrainment, gradient Richardson number accounts for shear flow, and instantaneous
adjustment simulates high-frequency convection in the thermocline. It is important to note that implementation of this mixing scheme has led to accurate simulation of the mixed layer, barrier layer, and subduction pathways [Luo et al., 2005]. In addition, this improved mixed layer formulation is particularly well suited to diagnose changes in forcing since surface heat and freshwater fluxes are calculated interactively by coupling the OGCM to a thermodynamic AML model [Murtugudde et al., 1996] thus allowing feedbacks between SST, SSS, and surface fluxes.

The model configuration used for the present simulations covers the tropical Pacific basin (124°E–76°W, 30°N–30°S) with a homogeneous longitudinal grid spacing of 1° and a stretched latitudinal grid (down to 1/3° within 10° of the equator). The boundaries are treated as a sponge layer within 5° of the meridional boundaries smoothly relaxing to WOA05 climatological values. The vertical structure consists of a variable depth mixed layer and 19 active sigma layers with a deep motionless boundary being specified as \( T_{\text{bottom}} = 6^\circ \text{C} \) and \( S_{\text{bottom}} = 35 \text{ psu} \).

The model is spun up from rest using climatological winds (i.e., seasonal cycle calculated from 1978 to 2007) with the initial conditions derived from WOA05 data and is allowed to come to equilibrium after 30 years of forcing by the wind product of Bourassa et al. [2001]. This pseudostress wind field, which objectively combines various in situ observations, has similar wind curl and divergence patterns as scatterometer data [Bourassa et al., 2005] and has the advantage of having a consistent formulation for an extended period from 1978 to present. The bulk formula is used to convert these pseudostress values to wind stress \( (\mathbf{C}_d = 1.2 \times 10^{-3} \text{ and } \rho = 1.2 \text{ kg/m}^3) \). Interannual runs are initialized from this climatological spin-up and the wind speeds required for the AML sensible and latent heat fluxes are computed from interannual wind stresses.

3.2. Ensemble Reduced Order Kalman Filter Data Assimilation

Since pure ocean model hindcasts are tainted by errors in the forcing and deficiencies in the model, we exploit data assimilation to get the best possible representation of the ocean state for initial conditions for various sets of coupled experiments. The Ensemble Reduced Order Kalman Filter (EROKF) is used here. The equations of the EROKF are obtained by projecting the equations of the Kalman Filter upon a basis of multivariate EOFs of the model from 1985 until 2004. See additional details and bibliographical references of Ballabrera-Poy et al. [2001].

Preliminary experiments have shown that 30 MEOFs provide a reasonable compromise between accuracy, overfitting, and computational cost. To reduce the pervasive effects of neglecting the complementary of the EOF subspace, which underestimates the analysis error covariance [see, e.g., Cane et al., 1996], an ensemble technique is used to estimate the analysis error covariance each month. The filter stability is maintained, in part, by adding a constant diagonal matrix, \( \mathbf{Q} \), to the reduced order background error covariance which is empirically estimated from the residuals.
of the data-EOF fitting during the same period [Verron et al., 1999].

[19] Prior to assimilation, vertical profiles of temperature are subsampled from the OI field at observation locations and are then interpolated to the nearest model grid point. This requires the knowledge of layer thickness, which is given by the model forecast. Projecting observations onto the corresponding model numerical grid point simplifies the forward observational operator, H, which now becomes a simple horizontal interpolation. Multiple subsurface observations falling within the same model grid point are averaged.

[20] For the baseline experiment, all available subsurface temperature data subsampled from OI (sTz) are assimilated using a 10 day assimilation cycle starting in October 1992. Assimilating point-wise data rather than gridded fields of temperature serves to minimize the impact of this variable within the innovation vector while still retaining the subsurface information, so as not to overconstraint the subsurface thermal field. The observational error values for sTz and SSS were optimized by running a series of assimilation experiments assimilating each variable individually. The best results were obtained when using error values of 0.75°C and 0.31 psu for temperature and salinity, respectively.

[21] For the period 1993–2007, validation of experiments that assimilate subsurface temperature and sea surface salinity (ASSIM_sTz_SSS) and subsurface temperature (ASSIM_sTz) are presented in Figure 3. Figure 3 highlights the overall good agreement between assimilation results and observations. For example, SL correlation is greater than 0.6 for ASSIM_sTz results in the waveguide climbing to greater than 0.7 for ASSIM_sTz_SSS showing the good temporal agreement between model and observed sea level variability. In addition, the ASSIM_sTz_SSS experiment shows slight improvement in SL at 10°S, in the central basin as well as off the equator in the Northern Hemisphere. The very good agreement of the modeled mean depth of the thermocline (dashed line) versus observations (solid) in Figures 3g and 3h reinforces the quality of the assimilation products. For subsurface temperature validation, the ASSIM_sTz_SSS results show slight improvement below and degradation above the thermocline with respect to the ASSIM_sTz results. Surface temperature RMS between model and observation anomalies are generally less than 1°C except in the far eastern Pacific where values climb as high as 1.5°C for ASSIM_sTz_SSS and 1.35°C for ASSIM_sTz. The SSS anomalies are validated using the gridded SSS product described in section 2 and confirm the quality of the assimilation products. Note that the SSS anomalies are better represented by ASSIM_sTz_SSS (Figures 3e and 3f) than for the ASSIM_sTz experiment especially in the high-rainfall regions of the Intertropical Convergence Zone (ITCZ), the western Pacific and South Pacific Convergence Zone (SPCZ).

3.3. Hybrid Coupled Model

[22] The hybrid coupled model (HCM) has been developed based on the Gent–Cane ocean model combined with the statistical atmospheric model [Zhang et al., 2006]. This model is used to derive a wind field from the SST anomaly field and is constructed from a singular value decomposition (SVD) analysis between these two fields [Syu et al., 1995; Chang et al., 2001; Zhang and Zebiak, 2004; Zhang et al., 2005]. The SST anomalies are from the work of Reynolds et al. [2002], and the wind stress anomalies come from the Max Planck Institute for Meteorology (MPI) Atmospheric GCM (ECHAM4.5 [Roeckner et al., 1996]), more precisely from the ensemble mean of a 24 member ECHAM 4.5 simulation for the period 1950–1999, forced by observed SST anomalies. As demonstrated by Barnett et al. [1993] and Syu et al. [1995], the seasonality of the atmosphere can have an important effect on the onset and evolution of El Niño. Thus, 12 different submodels, one for each calendar month, are utilized to derive surface wind responses to SST anomalies.

[23] Operationally, the coupled model is initialized from the data assimilation results starting on the first of each month, 1993–2007. For each time step, the HCM calculates the model SST, and then looks up the climatological seasonal cycle SST for the current model day using the cubic spline of the last year of the climatological experiment (i.e., climate year 30). This climatology is subtracted from the current model SST and this residual is used to convert SST anomaly to wind stress anomaly using the appropriate monthly SVD wind submodels. The resulting wind stress anomaly is then added to the climatological winds and the process repeats for the next time step. All other forcing, such as precipitation, cloudiness and solar, are prescribed to seasonal cycle climatological values. Persistence is defined as repeating the observed value from the first time step and comparing against real observed NINO3 SST anomaly time series for each month from 1993 to 2007 using either correlation or root-mean-square error (RMS) difference. A forecast is considered to be skillful if it is significantly greater (smaller) than the correlation (RMS) value for persistence.

4. Results

4.1. Impact of SSS Assimilation on Coupled Forecasts

[24] The most straightforward way to assess the impact of SSS assimilation is to compare the results of a baseline experiment with and without SSS assimilation. Therefore, we present the results of HCM runs initialized from experiments that assimilate subsurface temperature (ASSIM_sTz) with one that also includes SSS assimilation (ASSIM_sTz_SSS). Here it is important to remember that assimilation of subsurface temperature has an impact on all model variables including subsurface salinity through the MEOF nature of the assimilation scheme. Figure 4 shows correlation and root-mean-square error (RMS) with observed NINO3 SST anomaly (5°N–5°S, 150°W–90°W) for these two coupled forecast experiments. To put the correlation plot into context, a correlation of 0.31 is significant at the 95% confidence level assuming 30 independent observations, corresponding to the number of coupled forecasts for each lead time (i.e., 180) divided by an assumed NINO3 SST anomaly decorrelation scale of 6 months. Individually, both the ASSIM_sTz and ASSIM_sTz_SSS results pass the 95% confidence limits for all forecasts up to 9 and 11 months, respectively. Prior to 3 months, persistence of the SST observations has higher correlation than the coupled forecasts. However, after 3 months, both the ASSIM_sTz and ASSIM_sTz_SSS experiments exceed persistence for correlation (Figure 4a). This corresponds to the time it
Figure 3
significant at the 95% confidence limit. For 30 independent observations, a correlation of 0.31 is significant, and ASSIM_sTz_SSS (solid line) assimilation. For experiments initiated from ASSIM_sTz assimilation (dotted line) are formulated with respect to 1993–2007 seasonal cycle. 3h correspond to the average depth of the 20°C isotherm for observations and model results, respectively. All anomalies from the OI gridded field using all available subsurface temperature information. The solid and dashed lines in Figures 3g and (h) RMS of equatorial longitude versus depth. Sea level data are from Archiving, Validation, and Interpretation of Satellite Oceanographic data (AVISO) gridded multisatellite product, SST data are from Reynolds and Smith gridded SST, SSS data are from gridded product discussed in section 2, and temperature longitude versus depth comes from the OI gridded field using all available subsurface temperature information. The solid and dashed lines in Figures 3g and 3h correspond to the average depth of the 20°C isotherm for observations and model results, respectively. All anomalies are formulated with respect to 1993–2007 seasonal cycle.

Figure 3. Validation of (left) ASSIM_sTz_SSS and (right) ASSIM_sTz model results versus observed values. (a, b) Correlation of sea level anomaly, (c, d) root-mean-square (RMS) of sea surface temperature (SST) anomaly, (e, f) RMS of SSS anomaly, and (g, h) RMS of equatorial longitude versus depth. Sea level data are from Archiving, Validation, and Interpretation of Satellite Oceanographic data (AVISO) gridded multisatellite product, SST data are from Reynolds and Smith gridded SST, SSS data are from gridded product discussed in section 2, and temperature longitude versus depth comes from the OI gridded field using all available subsurface temperature information. The solid and dashed lines in Figures 3g and 3h correspond to the average depth of the 20°C isotherm for observations and model results, respectively. All anomalies are formulated with respect to 1993–2007 seasonal cycle.

Figure 4. NIÑO3 SST anomaly statistics for (a) correlation and (b) root-mean-square (RMS) error versus forecast lead time for observation persistence (dashed line), coupled experiments initiated from ASSIM_sTz assimilation (dotted line), and ASSIM_sTz_SSS (solid line) assimilation. For 30 independent observations, a correlation of 0.31 is significant at the 95% confidence limit.

takes for an equatorial Kelvin wave, present in the IC, to traverse the Pacific. By month 6, both ASSIM_sTz and ASSIM_sTz_SSS correlations exceed persistence at the 99.9% confidence level using the Fisher Z test confirming these as skillful forecasts. In addition, both these HCM results show lower RMS than persistence after 6 month forecasts (Figure 4b). To summarize, both the HCM forecasts, ASSIM_sTz and ASSIM_sTz_SSS, are generally significant on their own and are validated with respect to persistence after 3 (6) month forecasts for correlation (RMS).

[25] To test the impact of SSS assimilation, one needs to compare the results for the two coupled experiments shown in Figure 4. ASSIM_sTz_SSS correlation consistently exceeds ASSIM_sTz after 3 months by as much as 0.13 at month 12. For correlation, the Fisher Z one-sided test, which tests whether one correlation is significantly higher than another, indicates that at 6 months, ASSIM_sTz_SSS is significantly larger than ASSIM_sTz at the 79% level. However, by 9 months, this value climbs to 90% and by month 12, the significance stands at the 92% level. In addition, the ASSIM_sTz_SSS RMS is lower than ASSIM_sTz by as much as 0.22°C RMS by month 9. Thus, assimilation of SSS into the initial conditions for coupled experiments leads to a statistically significant improvement with respect to using ASSIM_sTz alone. These results reinforce those of Ballabrera-Poy et al. [2002] who determined that SSS observations would not improve the short-term statistical ENSO forecast (i.e., less than or equal to 3 months), but that SSS would help improve the results after 6 month forecasts. Figure 4 indicates that SSS assimilation significantly improves the coupled forecasts after 6 month lead times.

[26] Another way to examine the improvement in coupled forecasts brought about by SSS assimilation is to break down the statistics by time of year of the initial state. HCM forecasts are compared to observations using plots showing forecast month versus lead time for correlation and RMS (Figures 5 and 6, respectively). The observation persistence (Figures 5c and 6c) is included to provide a background reference for assimilation results and shows that the correlation (RMS) persistence remains high (low) for at least 11 month lead times for May–November forecasts. However, for December–March, the relatively quick degradation of skill, with a correlation drop from 1 to below 0 and RMS rise from 0°C to 1.2°C over 0–6 month lead times, is the well-known spring prediction barrier (SPB) [Barnston et al., 1994; Latif et al., 1994; Webster and Yang, 1992].

[27] For both the ASSIM_sTz_SSS (Figures 5a and 6a) and ASSIM_sTz (Figures 5b and 6b) experiments, the SPB problem is somewhat ameliorated as compared to persistence. This is not surprising in light of previous work that showed that specifying the upper ocean heat content through assimilation of subsurface temperature [McPhaden, 2003; Ruiz et al., 2005; Xue et al., 2000] and the use of ensemble initial conditions [e.g., Duan et al., 2009] improved spring predictions. Comparing Figures 5a and 5b to Figure 5c shows that coupled forecasts have increased correlation of approximately 0.4–0.8 for forecasts longer than 4 months for December–March initial states. In addition, Figures 6a and 6b versus Figure 6c shows the RMS is reduced by greater than 0.5°C. This reduction as compared to the observed persistence is apparent for coupled forecasts initiated from unassimilated, ensemble model results (not shown) especially between March and May for forecasts.
longer than 2 months where the correlation is increased by as much as 0.4–0.6 and the RMS is reduced by approximately 0.2°C.

[28] But what additional benefit comes from assimilating SSS into the initial conditions? Figure 5d shows the correlation differences, ASSIM_sTz_SSS versus observations minus ASSIM_sTz versus observations (i.e., Figure 5a minus 5b). Over most of the plot, positive differences indicate where SSS assimilation improves forecasts. Correlation improvements from 0.1 to 0.5 can be seen for the key SPB region from December–March for forecasts from 6 to 12 months. This reinforces the results from above since forecasts are generally improved for long-term forecasts from 6 to 12 months. For RMS (Figure 6d), negative values show where the ASSIM_sTz_SSS experiment is improved with respect to ASSIM_sTz. Overall RMS differences are generally negative. In particular, RMS values are improved by as much as 0.5°C for the December–March initiated forecasts and 6–12 month lead times. Such improvement indicates that much of the impact of SSS assimilation is going toward diminishing the impact of the SPB. This improvement goes above and beyond what assimilation of subsurface temperature also provides. This is particularly noteworthy for longer-lead El Niño forecasting as the December–March time of year is normally well before full-blown Bjerknes feedback has locked in with warm SST in the east and relaxed trade winds to the west.

[29] The physical explanation for the improvement in the coupled forecasts brought about by including SSS assimilation can be illustrated by analyzing differences between the ASSIM_sTz_SSS model and ASSIM_sTz (Figure 7). Figure 7a shows the annual mean difference for salinity (colors) and ASSIM_sTz_SSS experiment Bernoulli function (contours). The Bernoulli function can be thought of as the geostrophic streamlines along a specified isopycnal [Luo et al., 2005]

\[
B(\sigma) = \frac{1}{2} \rho \left( \mu^2 + v^2 \right) + \rho g \eta + g \int_{\sigma}^{0} |\rho - \rho(\sigma)| \, dz
\]

where \(\mu, v\) are zonal and meridional currents along an isopycnal, \(\rho\) is density, and \(\eta\) is the sea level. In our case, we choose the isopycnal, \(\sigma = 23.5\), because it corresponds to outcrop regions within 25° north and south of the equator and represents mixed layer transport well above the thermocline where \(\sigma \approx 25.5\). A fresh water anomaly can be seen over most of the Pacific extending from 15°S to about 5°N east of 160°E. In the same region, the mean flow, represented by the Bernoulli function, shows that there is a general confluence of fresh water toward the equatorial upwelling region east.
of 180° with the southern feature being more prominent. North of 5°N, a large region of positive salinity anomaly is cut off from the equator by the vorticity island centered at 10°N at the western boundary [Rothstein et al., 1998]. West of 160°E weak positive salinity leaks to the equator through western boundary currents.

The relatively fresh anomaly along the equator for the ASSIM_sTz_SSS versus ASSIM_sTz experiments leads to a more stratified mixed layer depth (MLD), a more stable water column, and upwelling. This feature is illustrated in Figures 7b–7f. The barrier layer thickness (BLT), which is formed because of subduction of central Pacific salty water beneath western Pacific fresh water [Shinoda and Lukas, 1995], is defined as the difference of the isothermal MLD (i.e., depth where $T_{\text{mld}} = T_{\text{sfc}} - 0.5^\circ\text{C}$) minus the density-defined MLD (i.e., depth where $\rho_{\text{mld}} = \rho(T_{\text{sfc}} - 0.5^\circ\text{C}, S_{\text{sfc}})$) as in the work of Sprintall and Tomczak [1992]. Positive values in Figure 7b indicate that salinity assimilation increases the salinity barrier thickness at the base of the mixed layer throughout most of the equatorial waveguide except for west of 170°E along the equator and north of 10°N west of the dateline. BLT plays a key role in MLD [Lukas and Lindstrom, 1991] and positive BLT corresponds with a decreased MLD (Figure 7c) for all longitudes within 5° of the equator. The depth of the 20°C isotherm, representing the depth of the thermocline, shows that the ASSIM_sTz_SSS is shallower than the ASSIM_sTz experiment throughout most of the equatorial waveguide. The stronger upwelling for ASSIM_sTz_SSS versus ASSIM_sTz leads to cooling just above the thermocline (Figure 7e) which surfaces west of 120°W near the equator (Figure 7f). Near the equator, east of 100°W, the positive SST anomaly along the coast is most likely due to downwelling imposed by specification of the minimum mixed layer depth in the model.

In our experiments, the shoaling of the mixed layer leads to more accurate coupled predictions since any wind perturbations associated with equatorial Kelvin and Rossby waves are concentrated within a shallower layer for the ASSIM_sTz_SSS IC coupled experiments as compared to ASSIM_sTz. For example, Vialard et al. [2002] presented forced-only experiments for 1993–1999 that contrasted a control experiment with one that eliminated salinity effects, including BLT. During the second half of 1997, thick BLT, corresponding to thinning MLD in central Pacific, led to increased eastward currents and increased SST in central and eastern Pacific. Therefore, vertical stratification of salinity impacts the surface layer momentum budget, decreasing the mixed layer, thus increasing the response to wind forcing [Vialard and Delecluse, 1998].

The features represented in Figure 7 are most intense in June–July–August (JJA) leading to especially improved forecasts for 6 months following December–January–February (DJF) initialized forecasts. In other words, the
increased upwelling along the equator is most intense during JJA (not shown) which concentrates wind forcing in a shallower mixed layer and improves the SPB for 6 month forecasts initialized in DJF. These results are confirmed by Maes et al. [2005] who found that the coupled response of eliminating the barrier layer by adjusting the coupled model mixing is especially sensitive from October to March because of the timing of Kelvin waves arrival in the east during the upwelling season.

[33] To summarize, we compared the results of assimilation of SSS and stTz, ASSIM_sTz_SSS, versus ASSIM_sTz in order to assess the impact of SSS assimilation on coupled forecasts. ASSIM_sTz_SSS and ASSIM_sTz are individually significant, and these correlation values are significantly better than persistence. In addition, the major result is that

Figure 7. Mean annual values for ASSIM_sTz_SSS minus ASSIM_sTz for (a) SSS (color) and ASSIM_sTz_SSS Bernoulli function using 23.5 °C (contours), (b) barrier layer thickness (BLT) (as defined in text), (c) density criteria mixed layer depth (MLD), (d) depth of the 20°C isotherm (i.e., thermocline), (e) equatorial longitude versus depth of temperature (solid line indicates mean depth of observed 20°C isotherm, and (f) SST. Bernoulli streamlines are contoured every 0.2 cm and flow counterclockwise (clockwise) in the Northern (Southern) Hemisphere around low values. Missing contours correspond to regions where the 23.5 °C surfaces. MLD and BLT are defined using same criteria as used by Sprintall and Tomczak [1992].
lags provide independent information. However, SSS for between 3°N–equator at 180°W and extend into the ITCZ region centered Namely, minimum negative values are found near the pattern looks very similar to that found in Figure 2c. For linear prediction of NIÑO3 SST anomaly. For 0 anomaly that provided significant independent information multiregression technique to highlight the regions of SSS regions of significant usefulness for predictability. By 6 month leads, the third most significant predictor is SSS anomaly located near 5°S, 140°W, centered on a region of positive correlation extending between 5°S–10°S, and west to the western boundary encompassing the South Pacific Convergence Zone (SPCZ). To the east, this particular data set [Delcroix, 1998] lacks enough information to grid the data. For 9 and 12 month leads, the pattern of SSS anomaly looks generally similar to 6 months but the highest predictor region (now ranked as second most important) is located near 5°S, 165°E.

[35] In order to assess the spatial contribution of the SSS anomaly, a new series of experiments is performed by limiting SSS data assimilation to particular regions. Experiments that highlight the longitudinal impact of SSS are presented and include assimilating SSS anomaly within the equatorial waveguide for all longitudes (5°N–5°S, 124°E–76°W; e.g., ASSIM_sTz_SSS_5N–5S), for the western Pacific (5°N–5°S, 124°E–180°E; e.g., ASSIM_sTz_SSS_5N–5S_125–180E), and for the eastern Pacific (5°N–5°S, 180°E–76°W; e.g., ASSIM_sTz_SSS_5N–5S_180E–76W). For these and all subsequent experiments, subsurface temperature profiles for the entire basin are assimilated (just like the previous experiments) and the assimilation results are used to initialize the HCM model starting on the first day for each month, 1993–2007.

[36] Figure 8 shows the results of the experiments that assimilate SSS for different longitudinal regions. The experiment in which all SSS observations are assimilated shows the best overall statistics. As the number of observations decreases, the performance of the assimilation is reduced. These results, by themselves, are of interest because they illustrate the overall positive impact of SSS monitoring as envisioned by satellite SSS. Moreover, a priori expectations from the results of Ballabrera-Poy et al. [2002] would suggest that the experiment assimilating SSS data in the western Pacific should outperform the experiment assimilating SSS data in the eastern part of the basin. Figure 8 shows that these results do follow as expected: ASSIM_sTz_SSS has the best correlation and lowest RMS followed by ASSIM_sTz_SSS_5N–5S and ASSIM_sTz_SSS_5N–5S_125–180E and the least is ASSIM_sTz_SSS_5N–5S_180E–76W. Each of these experiments is significant at the 95% level for forecasts up to 10 months (assuming 30 independent observations corresponding to a correlation value of 0.31). However, maximum values of the Fisher Z statistic never exceed the 77% significance level for the correlation differences between the best (ASSIM_sTz_SSS) and least (ASSIM_sTz_SSS_5N–5S_180E–76W) experiments. Breaking these results into monthly start times (as in Figures 4 and 5) to examine the impact on the SPB shows no particular differences between the full spatial SSS assimilation and the various longitudinal assimilation experiments (not shown) with the lone exception being the ASSIM_sTz_SSS_5N–5S_180E–76W experiment that limits improvements brought about by SSS assimilation to January for 6–12 month forecast lead times.

[37] In order to assess how the salinity is impacted for each of the regional assimilation scenarios, Figures 9a–9d shows the SSS differences with the ASSIM_sTz experiment. The full assimilation shows the largest amplitude followed by the ASSIM_sTz_SSS_5N–5S and ASSIM_sTz_SSS_5N–5S_125–180E. The ASSIM_sTz_SSS_5N–5S_180E–76W (Figure 9d) shows a fresh anomaly in the eastern Pacific but also has a strong signal west of the assimilation region. The western anomaly is due to the high correlation between the southwest Pacific and the salinity peak at 5°N, 100°W in the assimilation basis MEOF structure. For all these experiments, the ranking of the

**Figure 8.** NIÑO3 SST anomaly (a) correlation and (b) RMS versus forecast lead time statistics for coupled experiments assimilating different regions of SSS OI data; full grid 20°N–20°S, 125°E–76°W (solid line); equatorial band 5°N–5°S, 125°E–76°W (dotted); warm pool 5°N–5°S, 125°E–180°E (short dash); and eastern Pacific 5°N–5°S, 180°E–76°W (long dash line).

ASSIM_sTz_SSS significantly improves upon ASSIM_sTz coupled forecasts as validated by NIÑO3 SST anomalies and assimilation of SSS contributes to minimizing the SPB.

### 4.2. Spatial Contribution of SSS

[34] The work of Ballabrera-Poy et al. [2002] utilized a multiregression technique to highlight the regions of SSS anomaly that provided significant independent information for linear prediction of NIÑO3 SST anomaly. For 0–3 month SSS leads NIÑO3 SST anomaly correlations, the spatial pattern looks very similar to that found in Figure 2c. Namely, minimum negative values are found near the equator at 180°W and extend into the ITCZ region centered between 3°N–10°N. For SSS, neither 0 nor 3 month lags provide independent information. However, SSS for 6–12 months leading the NIÑO3 SST anomaly shows regions of significant usefulness for predictability. By 6 month leads, the third most significant predictor is SSS anomaly located near 15°S, 140°W, centered on a region of positive correlation extending between 5°S–10°S, and west to the western boundary encompassing the South Pacific Convergence Zone (SPCZ). To the east, this particular data set [Delcroix, 1998] lacks enough information to grid the
improvement brought about by various regional SSS assimilation depends on the salinity signal within 5° of the equator. For example, the SSS pattern of ASSIM_sTz_SSS_5N–5S looks similar to the full assimilation but with reduced amplitude (compare Figure 9b with Figure 9a). With the reduction in amplitude of the SSS signal, a concomitant reduced upwelling with cooling above the thermocline predominates within 5° of the equator. For each of these regional assimilation examples, cooling above the thermocline surfaces between 140°W to 110°W but is lower than for the full assimilation.

[38] Next, the latitudinal impact of SSS assimilation on coupled model forecasts is assessed by performing two additional experiments which utilize SSS in different latitude bands, one assimilates for all longitudes between 5°S and 20°S (ASSIM_sTz_SSS_5S–20S) and another assimilates SSS data between 5°N and 20°N (ASSIM_sTz_SSS_5N–20N). The former is designed to test the impact of the South Pacific Convergence Zone (SPCZ) and the latter is meant to test how SSS variability in the ITCZ impacts coupled forecasts. From the work of Ballabrera-Poy et al. [2002] we should expect that the SPCZ SSS assimilation contributes to the signal, since the observations in this region show a significant, independent SSS contribution to multiregression of the NINO3 SST time series, especially at long lead times. In addition, we know this region to be the source of the western Pacific barrier layer fresh water [Shinoda and Lukas, 1995]. On the other hand, the northern
or ITCZ SSS assimilation should result in a relatively poor forecast since there is no significant contribution from the Northern Hemisphere. The results of the ASSIM_sTz_SSS_5°N–20°S and ASSIM_sTz SSS 5°N–20°N are compared against the full region SSS (ASSIM_sTz_SSS) and the equatorial band (ASSIM_sTz_SSS_5°N–5°S) assimilation results in Figure 10.

[39] The order of the correlation and RMS after 6 month lead times follows expectations. Namely, full spatial assimilation has the highest (lowest) correlation (RMS) followed by the equatorial band (ASSIM_sTz_SSS_5°N–5°S); and the SPCZ SSS assimilation (i.e., ASSIM_sTz_SSS_5°S–20°S) outperforms the Northern Hemisphere results. ASSIM_sTz_SSS_5°N–20°N, with higher correlation and lower RMS. Plots of start month versus forecast lead time (not shown) indicate a similar pattern as the full assimilation for ASSIM_sTz_SSS_5°N–5°S and ASSIM_sTz_SSS_5°S–20°S experiments except that the contribution of the SPCZ is centered on January–February, later in the forecast period, after 8 month lead times. The ASSIM_sTz_SSS_5°N–20°N experiment is the clear outlier from this analysis and has a maximum SSS assimilation improvement in February–March for 10 month forecasts.

[40] The reason for the poor performance of ASSIM_sTz_SSS_5°N–20°N can be determined through analysis of the SSS signal. Figure 9f shows the salinity anomaly with respect to the ASSIM_sTz results. Unlike the other regional assimilation products, ASSIM_sTz_SSS_5°N–20°N has a strong positive SSS anomaly extending from the western boundary to 160°W. The lack of a Southern Hemisphere source for fresh water allows the positive anomaly from the Northern Hemisphere to dominate, the barrier layer to be eliminated, and the mixed layer to be increased (not shown). The positive SSS is associated with downwelling that extends from 160°E to 160°W between 5°S to 10°N along the equator and so the upwelling signal that improves the coupled forecasts for the other scenarios is thus trapped west of 180° limiting the potential positive impact.

[41] It is interesting to note that when the SPCZ contribution (ASSIM_sTz_SSS_5°S–20°S) is compared to the western Pacific results (ASSIM_sTz_SSS_5°N–5°S) shown in Figure 8, these two are very similar with the results statistically indistinguishable (i.e., significance of the difference never exceeding 63%). This indicates that the contribution of the western Pacific and SPCZ SSS variability are equally important for coupled model forecasts at long lead times. In addition, the SSS signals for both experiments (Figures 9d and 9e) look remarkably similar. The high correlation between the equatorial southwest Pacific and the fresh anomaly centered at 5°N and 100°W in the assimilation MEOF basis means these two features are strongly linked in the model results.

[42] In summary, the results of the coupled model experiments initialized from various assimilation experiments shown in Figures 8 and 10 indicate that (1) SSS assimilation from the entire region gives the best overall results, (2) both the western Pacific and SPCZ are important for contributing to the positive impact of SSS assimilation has on coupled forecasts, and (3) assimilation of SSS in the eastern Pacific and Northern Hemisphere has considerably less impact.

5. Discussion

[43] The results of this work show that assimilation of SSS leads to improved coupled forecasts for the period, 1993–2007. Soon, additional near-global SSS information will become available from SMOS and Aquarius/SAC-D satellite missions. With the potential opportunity for weekly global gridded SSS fields, important questions need to be asked about how data coverage might impact these results. For the current example, time series of the number of observations presented in Figure 1 indicate that, prior to 2002, only about 100 bins are filled with SSS observations for OI for each month. After that year, the expansion of the Argo network provide for nearly complete monthly data coverage for the OI procedure. Figure 2c shows an example of well defined grids of SSS anomaly versus the poor data coverage example of Figure 2a. Clearly, Figure 2c shows more detailed SSS anomaly fields for December 2006 as compared to Figure 2a (December 1997), which contains gridding artifacts in the north at 150°W, 20°N and near the southern boundary. Unfortunately, our current data set does not have the resolution to allow weekly assimilation of SSS data as envisioned once Aquarius/SAC-D and SMOS data become available. However, this data set does allow for
experiments to be performed that split the results into poor (1993–2001) and good (2002–2007) data coverage periods.

To address the data coverage issue four additional experiments were completed following the same experiment design. Namely, gridded subsurface temperature is subsampled at observation locations and used as the baseline experiment (ASSIM_sTz). In addition, interpolated fields of SSS anomalies are assimilated along with sTz. These assimilation results are used to initialize coupled experiments which are identical to those presented in Figure 4 except for one set of experiments, assimilation anomalies for 1993–2001 and for the other, 2002–2007, is used for initializing coupled experiments.

The resulting forecasts for all these assimilation experiments beat persistence after three months (all pass 91% significance difference using Fisher Z test). However, the main result of Figure 11 is that for both periods, the SSS assimilation improves forecasts with respect to baseline forecasts at lead times greater than 6 months. SSS assimilation correlation results are significantly greater than sTz for 1993–2001, greater than 80% significance at 6 months climbing to 94% for 12 month lead times. At month 11, the SSS assimilation improves forecast NINO3 SST by greater than r = 0.2 and the RMS is 0.3°C smaller. In summary, the 1993–2001 period shows that SSS assimilation significantly improves correlation and RMS with respect to the baseline experiment.

For 2002–2007 the results are less conclusive. Although the Fisher Z test indicates that the differences are less significant for ASSIM_sTz_SSS versus ASSIM_sTz for 2002–2007, the SSS assimilation results still provide higher skill than the coupled experiments that assimilate only subsurface temperature. The maximum correlation differences of the results with and without SSS assimilation are for 7 month forecasts but never exceed r = 0.06. For 2002–2007 RMS, ASSIM_sTz_SSS has consistently smaller values than ASSIM_sTz of about 0.05°C. Although not as statistically significant as the earlier period, 2002–2007 shows that SSS assimilation improves the coupled forecasts using validation against observed NINO3 SST anomaly.

The fact that the data poor period outperforms the data-rich period is counterintuitive for the prospect of improving coverage brought about by satellite SSS observations. The 1993–2001 results, which have less information, show a bigger impact of SSS assimilation than for 2002–2007. In order to rule out the possibility that data coverage, 1993–2001 versus 2002–2007, might cause these apparent inconsistencies, an additional experiment was performed in which the data coverage of 2002–2007 is degraded to match 1993–2001. For each month 2002–2007, the data were masked prior to the OI step using data coverage from a random month, 1993–2001. These degraded SSS OI data were then assimilated into the forced ocean model and then used as initial conditions for a coupled experiment following the exact procedure as before. The results (not shown) match closely the original coupled results for 2002–2007, from Figure 11, with slightly worse statistics (lower correlation, higher RMS), but statistically indistinguishable from one another. Since the degraded OI SSS results were so similar to the original, the relatively poor data coverage from 1993 to 2001 cannot be responsible for the differences in the impact of SSS assimilation for the different periods.

Another reason for the degraded impact of SSS for the data-rich period may be due to the different signal variance for different analysis periods. The differences in the predictability between one period and another are highlighted by Figure 12 which shows the persistence of the NINO3 SST anomaly observations for the SSS data-poor period, 1993–2001, and the Argo period, 2002–2007. For the period 1993–2001, both correlation and RMS show a higher degree of persistence than for 2002–2007. The 1993–2001 correlation results are statistically larger (greater than 94% confidence level) than 2002–2007 following 6 month forecast times. These results are expected since Balmaseda et al. [1995] and Yu and Kao [2007] showed that the strength of the SPB depended on decadal variability and periods with stronger interannual variability have fundamentally better predictability [e.g., Kirtman and Schopf, 1998; Tang et al., 2008; Deng and Tang, 2009; Jin et al., 2008]. For example, Tang et al. [2008] showed that the range between persistence calculated for various 20 year
subsections over 1881–2000 was 0.2 for correlation and 0.8–1.0°C RMS at 10 month leads [see Tang et al., 2008, Figure 7]. The values in Figure 12 fall within this range. In our case, the large magnitude ENSO of 1997–1998 leads to strong predictability for this period as opposed to the weaker El Niño in 2006 and La Niña in 2007 for the Argo period (2002–2007).

[49] In summary, it is difficult to prove that improved data coverage afforded by satellite SSS would lead to significantly improved coupled forecasts for weak ENSO events. However, even though 2002–2007 is not significant, we have shown that SSS assimilation always improves coupled forecasts. This is especially true for assimilation of sparse-gridded SSS observations during the large event in 1997–1998 which leads to improved forecasts for 1993–2001 as well as for the entire period of our study, 1993–2007. Note however that we have not considered the spatiotemporal coverage that Aquarius/SAC-D would provide. Along with the recently demonstrated improvements in ENSO simulations and forecasts with high-resolution coupled climate models and associated process and predictive understanding [Neale et al., 2008], we can expect that satellite SSS observations and assimilation will lead to direct and indirect improvements in ENSO understanding and forecasts and related teleconnections, especially those associated with ENSO-driven freshwater flux anomalies over the global ocean [Zhang et al., 2010].

6. Summary and Conclusions

[50] All available subsurface temperature and near-surface salinity observations were collected from various sources. These data were converted to anomalies and then were gridded onto standard Levitus depths using the OI technique of Carton and Hackert [1989]. The temperature OI was subsampled at all available locations and depths and assimilated into the reduced gravity, primitive equation, sigma coordinate model for the tropical Pacific using the EROKF technique. These results were then used to initialize HCM experiments for each month from 1993 to 2007 and served as the baseline experiment for subsequent coupled experiments. In a similar manner, near-surface salinity was gridded, assimilated into the forced experiment, and used as initial conditions for coupled experiments.

[51] Both HCM results for baseline and SSS assimilation experiments were validated with respect to the observed NIÑO3 SST anomaly. Individual correlation values exceed the 95% confidence limits for nearly all forecast lengths. In addition, correlation and RMS for both experiments were improved at a significance of 98% level with respect to persistence. Therefore, both experiments have significant correlations on their own right and both are validated as skillful since they exceed persistence values.

[52] The experiment with SSS assimilation was compared to the baseline that assimilated only subsurface temperature. Correlation of the SSS assimilation versus observations consistently exceeded those of the baseline experiment by as much as 0.13. Using the Fisher Z test, differences in correlations are significant at levels varying from 79% to 92% depending on the month. Following 3 month forecast, the RMS of the SSS assimilation experiment beats persistence. SSS is consistently lower than the baseline for all lead times. The improvement of the SSS experiment is as big as 0.22°C by month 9.

[53] The experiments were further broken down into forecast month versus lead time. We showed that the SPB was significantly improved by SSS assimilation. Correlation increased and RMS decreased for 6–12 month forecasts initiated from December to March initial conditions, a period important to long-lead El Niño forecasts since it precedes full-blown Bjerknes feedback. Therefore, assimilating SSS into the initial conditions for coupled experiments invariably led to a significant improvement when validating against observed NIÑO3 SST anomaly. The improvements brought about by SSS assimilation was explained by the increase in fresh water available in the mixed layer. BLT was increased leading to thinning of the MLD, upwelling and increased impact of ENSO wind anomalies in a thinner Ekman layer.

[54] Additional experiments were presented that explored the importance of SSS assimilation in various regions. SSS data were assimilated for the equatorial band, for the western Pacific and for the eastern Pacific to test the longitudinal contribution of SSS. These results show that the full region SSS assimilation has the best overall correlation and RMS, followed by the equatorial region and the western Pacific (which are statistically similar), and the eastern Pacific has
the worst validation with observations. Taken together these results confirm the strong impact of the western Pacific SSS for forecasting NIÑO3 SST anomalies, as was first hypothesized by Ballabrera-Poy et al. [2002]. [55] Experiments that assimilated SSS between 5°S–20°S and 5°N–20°N differentiated the influence of the SPCZ versus ITCZ SSS variability. These results confirmed that the SPCZ assimilation significantly outperformed the ITCZ example for all forecast periods. In addition, the SPCZ experiment was statistically indistinguishable from the western Pacific indicating that both these regions are particularly important for coupled model forecasts at long lead times. To summarize these experiments, SSS assimilation for the full region gives the best overall results, the western Pacific and SPCZ regions contribute key information for coupled forecasts and the eastern Pacific and ITCZ add little information for extending the accuracy of coupled HCM forecasts.

[56] Last, the impact of SSS data coverage was considered by splitting the coupled experiments into poor (1993–2001) and good (2002–2007) data coverage periods delineated by the start of the implementation of the Argo observing system. For these two periods, the SSS experiments outperformed the subsurface temperature experiments for both 1993–2001 and 2002–2007 periods. Individually only the 1993–2001 SSS assimilation experiment generally exceeded significance at the 95% confidence limits. The impact of the assimilation of SSS was more statistically significant for 1993–2001 versus 2002–2007. For 1993–2001, SSS assimilation correlation results are significantly greater than stTz for 1993–2001 varying from 80 to 94% significance whereas, for 2002–2007, the correlation differences never exceed r = 0.05. For RMS, SSS assimilation always reduced the error by as much as 0.31°C at month 11 for 1993–2001 and 0.05°C for 2002–2007. In some sense, these results are counterintuitive. Better data coverage should give better demonstrated impact of SSS assimilation. However, these results are interpreted in terms of the higher predictability of 1993–2001 because of the large ENSO signal.

[57] Satellite-derived SSS will soon be available from SMOS and Aquarius/SAC-D and this paper has demonstrated the important role of gridded SSS data assimilation for improving coupled forecasts. In particular, regions of key SSS variability in the western Pacific and SPCZ contribute to improving coupled NIÑO3 SST forecasts. The authors acknowledge that most operational coupled forecast systems include all available observational data such as SL, SST, stTz, and sSz assimilation. It is to be expected that the inclusion of all such data would lessen the impact of SSS assimilation. The purpose here was not to arrive at the best possible operational forecast, but rather to isolate and analyze the impact of the incorporation of fields of SSS observations. Therefore, rather than being the final say for the issue of SSS assimilation, this study should be viewed as a necessary first step and serves to highlight the importance of buoyancy processes in ENSO initialization. Renumbing these sorts of experiments with more complicated models utilizing fully coupled atmospheric models and more extensive data assimilation is required to assess issues such as asymmetry of ENSO forcing, the impact of freshwater flux, and the complementary nature of satellite/in situ assimilation. In addition, this study necessarily does not take into account high-frequency temporal or spatial SSS features that will be present in gridded satellite-derived SSS products. Smoothing is currently required because of the lack of data and prevents such high-frequency experiments. However, this paper has shown that assimilation of the fields of SSS is an important factor in improving coupled model forecasts.

[58] Acknowledgments. We would like to thank the many universities and Government agencies, without whose efforts the compilation and development of the Argo, GTSPP, WOD05, TAO, and COAPS Objective pseudostream wind data would not have been possible. In addition, we would like to thank the anonymous reviewers who helped improve this manuscript. This research is supported by NASA Physical Oceanography grant NNX08A176G with salinity processing support from the Ocean Salinity Science Team (NNX09A174G). In addition, Zhang’s work has been supported, in part, by NSF grant (AGS-1061998).

References


