






RESEARCH ARTICLE

10.1002/2017JC012760

Special Section:

Oceanic Responses and
Feedbacks to Tropical
CyclonesOSSE quantitative assessment of rapid-response prestorm
ocean surveys to improve coupled tropical cyclone predictionG. R. Halliwell Jr.¹ , M. Mehari^{1,2} , L. K. Shay³, V. H. Kourafalou³ , H. Kang³ , H.-S. Kim⁴,
J. Dong⁴, and R. Atlas⁵ ¹NOAA/AOML/PhOD, Miami, Florida, USA, ²CIMAS, University of Miami, Miami, Florida, USA, ³RSMAS, University of Miami, Miami, Florida, USA, ⁴NOAA/NWS/EMC, College Park, Maryland, USA, ⁵NOAA/AOML, Miami, Florida, USA

Key Points:

- Perform ocean observing system simulation experiments to evaluate impacts of ocean observations
- Prestorm ocean profile surveys reduce mesoscale errors and bias in ocean analyses used to initialize coupled TC prediction models

Correspondence to:

G. R. Halliwell Jr.,
george.halliwell@noaa.gov

Citation:

Halliwell, G. R., M. Mehari, L. K. Shay, V. H. Kourafalou, H. Kang, H.-S. Kim, J. Dong, and R. Atlas (2017), OSSE quantitative assessment of rapid-response prestorm ocean surveys to improve coupled tropical cyclone prediction, *J. Geophys. Res. Oceans*, 122, 5729–5748, doi:10.1002/2017JC012760.

Received 31 JAN 2017

Accepted 22 MAY 2017

Accepted article online 25 MAY 2017

Published online 17 JUL 2017

Abstract Ocean fields that initialize coupled TC prediction models must accurately represent the dynamics of mesoscale features and the associated distribution of upper ocean temperature and salinity. They must also provide unbiased realizations of upper ocean heat content and stratification. Ocean Observing System Simulation Experiments (OSSEs) are performed for three storms: Isaac, 2012; Edouard, 2014; and Gonzalo, 2014. These OSSEs assess the impact of rapid-response prestorm ocean profile surveys on improving ocean model initialization. Two types of surveys are evaluated: airborne deployments of expendable profilers and deployments of in situ thermistor chains along lines intersecting predicted storm paths. Assimilation of the existing ocean observing system substantially constrains mesoscale structure in dynamical fields, primarily because of the four available altimeters. However, these observations only modestly constrain mesoscale structure and bias in upper ocean thermal fields. Adding rapid-response airborne surveys to these observing systems produces substantial additional correction in thermal fields, but minimal additional correction in dynamical fields. Without altimetry assimilation, rapid-response profiles produce large additional correction in both dynamical and thermal fields. Airborne CTDs sampling temperature and salinity over 1000 m versus XBTs sampling temperature over 400 m produce additional correction for dynamical fields, but not for upper ocean thermal fields. Airborne surveys are generally more effective than thermistor chain deployments because they can sample a larger area at higher horizontal resolution and because the latter only measures temperature over the upper ~100 m. Both airborne profile surveys and thermistor chain deployments effectively reduce upper ocean thermal biases.

1. Introduction

Tropical Cyclones (TCs) force Sea Surface Temperature (SST) cooling that provides a negative feedback limiting storm intensification [Schade and Emanuel, 1999; Chan et al., 2001; Zhu and Zhang, 2006; Lloyd and Vecchi, 2011; Balaguru et al., 2015]. Increasing wind speed produces faster cooling beneath storms that reduces enthalpy flux from ocean to atmosphere. The strength of this negative feedback depends on the dynamical and thermodynamical (temperature, salinity, and density) structure of the upper ocean. Both temperature and thickness of the surface warm layer, which is defined for TCs as the layer where temperature exceeds 26°C, are important factors governing this feedback [Leipper and Volgenau, 1972; Jacob et al., 2000; Mainelli et al., 2008; Jaimes and Shay, 2009; Shay and Brewster, 2010; Vincent et al., 2012; Cione, 2015]. Thin warm layers permit rapid cooling that increases the negative feedback because the storm-forced ocean mixed layer can efficiently entrain colder water from below. Feedback strength also depends on storm parameters, with larger and slower moving storms experiencing stronger feedback [Lin et al., 2009; Mai et al., 2012; Walker et al., 2014; D'Asaro et al., 2014; Halliwell et al., 2015].

Although prestorm SST tends to be horizontally uniform during the TC season, ocean eddies and boundary currents can be associated with large horizontal differences in upper ocean warm layer temperature and thickness, and thus with the thermal energy available to storms. These features are also associated with differences in upper ocean stratification that affect the rate of mixed-layer deepening and cooling. Dynamical processes are also important because the background vorticity field associated with these ocean features distorts the forced SST cooling pattern [Jaimes et al., 2011, 2015, 2016]. Warm-core anticyclones and western boundary currents have relatively thick warm layers and dynamical properties that tend to limit cooling

while cold-core cyclones have relatively thin warm layers and dynamical properties that tend to enhance cooling [Hong *et al.*, 2000; Shay *et al.*, 2000; Walker *et al.*, 2005; Lin *et al.*, 2005; Wu *et al.*, 2007; Jaimes and Shay, 2010; Vianna *et al.*, 2010; Ma *et al.*, 2013; Jaimes *et al.*, 2015, 2016].

These factors make it necessary to initialize ocean forecast models with three-dimensional fields that accurately represent dynamical ocean features and their associated temperature, salinity, and density structure. Ocean observing systems are important in this context because they constrain errors in ocean analysis products used to initialize ocean forecast models [Pascual *et al.*, 2006; Pun *et al.*, 2007; Goni *et al.*, 2009]. Assimilation of the existing operational ocean observing system substantially constrains initial ocean model errors [Halliwell *et al.*, 2017]. Assimilation of prestorm rapid-response ocean survey observations is expected to further reduce initialization errors and improve prediction of SST cooling rates beneath the inner-core regions of storms, the region where this cooling has the greatest potential impact on intensity [Cione and Uhlhorn, 2003]. The present analysis is designed to quantitatively assess the impact of rapid-response prestorm ocean surveys on reducing initialization errors. It emphasizes the additional positive impact achieved beyond that already realized by assimilating operational ocean observations, particularly satellite altimetry which strongly constrains ocean mesoscale structure. Results are valid only for deep ocean regions away from continental shelves and only for rapid-response observations collected ahead of storms over their projected paths at least 1 day before arrival.

A limited number of prestorm rapid-response ocean profile surveys have been conducted from hurricane research aircraft and evaluated for historical Atlantic Ocean storms. These surveys involve airborne deployments of expendable bathythermographs (AXBT), conductivity-temperature-depth profilers (AXCTD), and/or current profilers (AXCP), with the latter instrument also measuring temperature profiles. (The capability to assimilate current velocity has not yet been developed and tested in our DA system.) Prestorm surveys are typically conducted in a “lawnmower” pattern to map the ocean mesoscale and associated upper ocean temperature-salinity structure and also to reduce bias in upper ocean heat content and stratification over a relatively large area. Surveys have been performed for Isidore and Lili (2001) [Shay and Uhlhorn, 2008; Uhlhorn and Shay, 2012, 2013], Rita (2004) [Jaimes and Shay, 2010], and Gustav and Ike (2008) [Meyers *et al.*, 2016]. Also, an AXBT demonstration project [Sanabia *et al.*, 2013] has deployed profilers in and around a number of storms in the Atlantic and Pacific oceans. The latter project documented ocean model error reduction resulting from profile assimilation using an adjoint-based method that does not require observations to be withheld for evaluation [Cummings and Smedstad, 2014].

Another survey approach has been considered, specifically the prestorm deployment of thermistor chains attached to surface platforms. Two types of surface platforms can be used: surface drifters typically drogued to 15 m that drift with the flow, and powered platforms such as the Liquid Robotics Wave Glider (<https://www.liquid-robotics.com/>) that extract propulsion energy from surface gravity waves and that can be steered to specified locations or instructed to remain stationary. Limited deployments of thermistor chains ahead of storms have been conducted [Hormann *et al.*, 2014; Lumpkin *et al.*, 2016]. Wave gliders have been tested and found to survive within hurricanes [Lenain and Melville, 2014] but have not yet been used in extensive prestorm deployments. One deployment strategy being given serious consideration, and that is evaluated herein, is to deploy several thermistor chains along one or more “picket fence” lines that are approximately normal to the projected storm path.

Rigorous quantitative assessment of one set of in situ airborne lawnmower surveys was conducted using Observing System Experiments (OSEs) [Shay *et al.*, 2011]. These ocean surveys were conducted for an entirely different purpose, specifically to improve prediction of the transport and dispersion of oil from the Deepwater Horizon oil spill. However, impacts can still be assessed with respect to the TC prediction problem. OSEs are twin data-assimilation experiments where one assimilates all observations and the second denies the system being evaluated [e.g., Atlas, 1997]. Impacts are quantified by increased error and bias in ocean analyses, and also in forecasts initialized by these analyses, resulting from denial. The oil spill OSE evaluated combinations of AXBTs, AXCTDs, and AXCPs that were deployed by the NOAA WP-3D hurricane research aircraft in lawnmower patterns on nine flight days between 8 May and 9 July 2010 [Shay *et al.*, 2011]. The OSE demonstrated that denial of the profile observations increased upper ocean RMS temperature errors by ~30% [Shay *et al.*, 2011]. Although these results are encouraging, the ability of OSEs to quantify impact is limited. The truth must be represented by observations withheld from assimilation, or must be

represented by the data-assimilative analysis itself as in the adjoint method of *Cummings and Smedstad* [2014].

In Observing System Simulation Experiments (OSSEs), the truth is represented by an independent, validated, high-quality ocean model simulation. The availability of a high-resolution, three-dimensional representation of the truth permits rigorous analysis of the structure and variability of errors in prognostic and derived model fields that is not possible in OSEs. Furthermore, OSSEs provide a method to evaluate the impact of new observing systems and different deployment strategies for existing systems [*Arnold and Dey*, 1986; *Atlas*, 1997; *Atlas and Riishojgaard*, 2008; *Halliwell et al.*, 2014; *Hoffman and Atlas*, 2015]. These capabilities are exploited herein to evaluate different design strategies for prestorm lawn mower airborne and “picket fence” thermistor chain surveys without the need to actually deploy instruments. To evaluate these strategies in different oceanic conditions, three storms are considered: Isaac (2012) in the Gulf of Mexico along with Edouard and Gonzalo (2014) in the open Atlantic Ocean. Multiple storms are considered because they occurred over regions with different oceanographic conditions. In particular, ocean currents and eddies in the Gulf of Mexico, along with the heat content differences associated with these features, are more energetic than over the open Atlantic regions. Impacts are assessed based on the requirements to accurately represent mesoscale structure and to reduce bias in upper ocean heat content and stratification in the initial fields provided to ocean models. For comparison, a joint OSE-OSSE analysis is presented for an actual airborne profile survey that was conducted prior to Edouard. In the future, impact analysis will be extended to error reduction in coupled intensity forecasts, but that effort is beyond the scope of the present analysis.

The ocean OSSE system is briefly described in section 2 while the experimental approach is discussed in section 3. Section 4 presents the impact of operational ocean observations, focusing in particular on the relative impact of altimetry assimilation versus other observing system components. Impact assessments for rapid-response surveys are presented in section 5. Section 5.1 presents the OSSE impact analysis for synthetic airborne ocean profile surveys. Section 5.2 presents the OSE analysis of a real airborne ocean profile survey during Edouard in comparison to an OSSE analysis assimilating synthetic versions of the same observations. Section 5.3 presents the impact of picket-fence deployment of thermistor chains. Results are summarized in section 6.

2. Ocean OSSE System

Key components of OSSE systems include a Nature Run (NR), a high-resolution unconstrained ocean simulation using a state-of-the-art ocean model that represents the “truth,” a data-assimilative ocean forecast system using an ocean model different from the NR (referred to as the Forecast Model, henceforth FM), and software to simulate ocean observations from the NR and add realistic errors. Detailed design characteristics for OSSE systems are presented in *Hoffman and Atlas* [2015]. The ocean OSSE system used in the present study follows these design recommendations and evaluation procedures to insure that credible assessments are obtained. The present OSSE system uses the fraternal twin approach with the HYbrid Coordinate Ocean Model (HYCOM) used as the NR and FM, respectively. To ensure that realistic errors exist between the two models, they are run with substantially different choices of numerical schemes and subgrid-scale parameterizations. Furthermore, the FM is run at lower resolution than the NR (0.08° horizontal Mercator mesh and 26 vertical layers versus 0.04° horizontal Mercator mesh and 35 vertical layers). Both configurations are adequate to resolve ocean mesoscale features, but the NR is configured to be more realistic. Model configuration and OSSE system design are discussed in detail in *Halliwell et al.* [2017] along with the influence of this system design on observing system impact assessments. Aspects of system design that influence impact assessments for the observing systems examined herein will be discussed as needed in the remainder of this paper.

It is the existence of a validated, three-dimensional, high-resolution representation of the truth that permits detailed rigorous assessments to be performed that are not possible with OSEs. The ocean OSSE system was initially evaluated in the Gulf of Mexico [*Halliwell et al.*, 2014]. The first application of that system evaluated different airborne survey strategies in the Gulf of Mexico during the time of the Deepwater Horizon oil spill [*Halliwell et al.*, 2015]. That evaluation focused on issues such as horizontal profile separation, temporal resolution of surveys, and profiler type (400 m AXBTs versus 1000 m AXCTDs). The results demonstrated the

overall positive impact of these surveys on reducing ocean model initialization errors that can potentially improve TC intensity prediction.

The ocean OSSE system has been extended to cover the Atlantic Ocean hurricane domain (5°S–45°N, extending east to 20°W), which is sufficiently large to limit the influence of boundary conditions over the subdomain directly affected by hurricanes. The suitability of the OSSE system NR for representing the “true” ocean has already been demonstrated. Model mean climatology, seasonal cycles, and variability of dynamical and thermodynamical fields were all deemed realistic [Kourafalou et al., 2016]. Model variability was also deemed realistic with respect to the Gulf Stream response to Hurricane Bill (2009) [Kourafalou et al., 2016]. Androulidakis et al. [2016] further demonstrated that the NR realistically reproduced the barrier layer associated with Amazon-Oricono river discharge, enabling the barrier layer impact on the ocean response to several storms to be studied.

Overall system validation was conducted by comparing OSEs to OSSEs that were identical except that the OSSEs assimilated synthetic versions of the actual ocean observations assimilated by the corresponding OSEs [Halliwell et al., 2017]. The OSSE system does produce credible impact assessments but tends to overestimate impacts based on RMS error reduction by O(10%). Because no correction is applied in the present paper, all assessments contained herein based on RMS error reduction contain this slight overestimate, which was found to not alter fundamental conclusions. Halliwell et al. [2017] also present the first applications of the Atlantic system. First, the impacts of key components of the existing ocean observing system (altimetry, Argo floats, satellite and in situ SST) with respect to reducing ocean model initialization errors were demonstrated. Second, enhancements to the existing ocean observing system achieved by deploying one to several underwater gliders were demonstrated. The present study extends these results to rapid-response, prestorm ocean surveys.

3. Experimental Design

3.1. Prestorm Ocean Survey Experiments

Three reference experiments provide baselines for evaluating observing system impacts (Table 1). A long unconstrained run of the FM beginning late 2008 provides error estimates in comparison to the NR without assimilation. The magnitude and distribution of these errors were deemed realistic compared to errors that presently exist between unconstrained state-of-the-art ocean models and the actual ocean [Halliwell et al., 2017], a key requirement for an OSSE system to produce credible assessments [Atlas, 1997; Hoffman and Atlas, 2015]. With this criterion substantially satisfied, impacts are assessed by the reduction in these error resulting from assimilation of observations into the FM. Rapid-response ocean observations are evaluated in comparison to experiment CONTROL which assimilates four major components of the operational ocean observing system (Table 1), specifically all four altimeters available during the 2014 hurricane season, satellite and in situ SST measurement systems, Argo floats, and XBT profiles. Synthetic versions of these observations are simulated from the NR at the same time, location, and depths as the actual observations. The rapid-response observations are also evaluated in comparison to experiment NOALT (Table 1), which differs from CONTROL by denying all four altimeters. Altimetry is chosen for comparison because these observations provide the largest correction to mesoscale ocean structure compared to other observing system components [e.g., Oke et al., 2015].

Table 1. Table of Reference Experiments Used to Evaluate the OSSE Results

| Experiment | Observations Assimilated | Run Time Interval |
|------------|---|---|
| FM | None (unconstrained) | Late 2008 through 2014 |
| CONTROL | Four altimeters (Jason-2, Cryosat, Envisat, Haiyang-2a) Satellite SST In situ SST Argo profiles XBT profile | Isaac: 1 Mar 2012 to 31 Oct 2012 Edouard, Gonzalo: 1 Mar 2014 to 31 Oct 2014 |
| NOALT | Satellite SST In situ SST Argo profiles XBT profiles | Isaac: 1 Mar 2012 to 31 Oct 2012 Edouard, Gonzalo: 1 Mar 2014 to 31 Oct 2014 |

This large impact has also been documented and verified by comparing OSE and OSSE results using the present Atlantic system in Halliwell et al. [2017]. Comparison of error and bias reduction in CONTROL with respect to NOALT quantitatively assesses the importance of altimetry in comparison to the other components of the operational observing system. By adding rapid-response observations to both CONTROL and

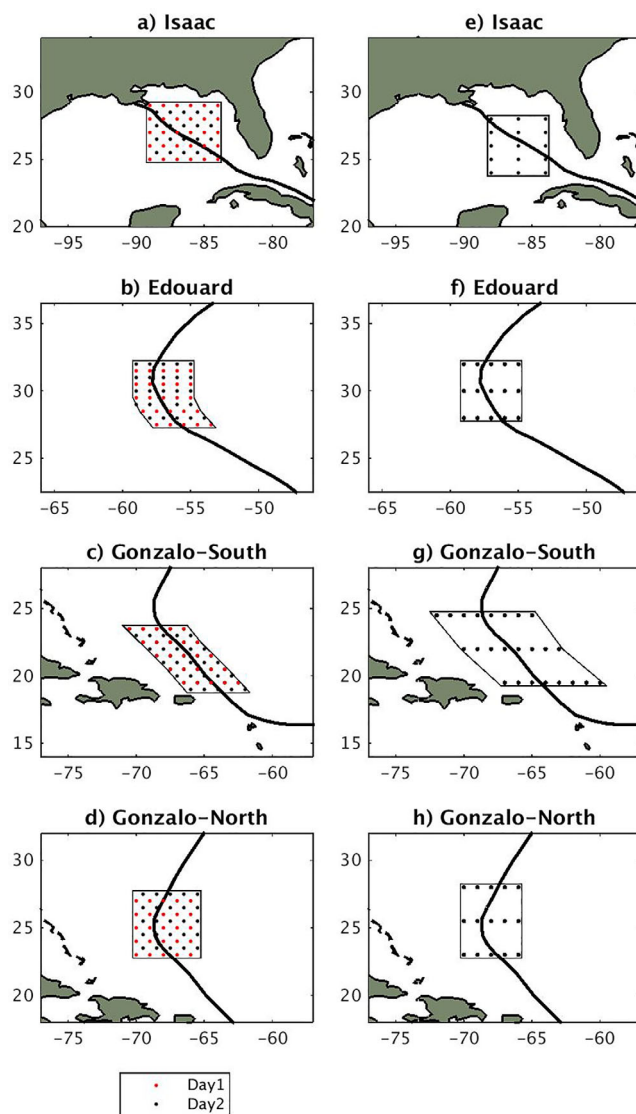


Figure 1. Sampling locations for the synthetic rapid response experiments in four hurricane survey regions. P3 airborne surveys are shown in the left figures while thermistor chains deployed on surface platforms are shown in the right figures. Idealized P3 surveys were run over 2 days, with red points representing the day 1 survey and black points representing the day 2 surveys.

NOALT, the relative importance of satellite altimetry and rapid-response surveys is quantitatively assessed.

OSSEs are performed for three hurricanes: Isaac (2012) in the Gulf of Mexico, Edouard (2014) in the open Atlantic east of 60°W, and Gonzalo (2014) in the open Atlantic between 60°W and 70°W (Figure 1). Evaluations for Gonzalo are performed for surveys conducted in two separate regions along the path (Gonzalo south and Gonzalo north, Figure 1) to illustrate how results can vary as a function of location for a given storm. The resulting four regions representing three storms enable evaluations to be conducted for a variety of ocean conditions. The basic set of OSSE experiments is listed in Table 2, four evaluating synthetic airborne surveys and two evaluating synthetic thermistor chain deployments. This set is run twice in the Isaac, Edouard, Gonzalo north, and Gonzalo south regions to perform the evaluations, once adding the rapid-response profiles to CONTROL and the other adding them to NOALT. For each of the storm regions, the strategy is to sample the ocean between 2 and 4 days before the storm arrives. The errors added to the synthetic observations simulated from the NR are summarized in Table 3.

The subset of experiments to evaluate airborne surveys (Table 2) include two to evaluate AXBTs profiling temperature down to 400 m (AXBT1D and AXBT2D) and two to evaluate AXCTDs

Table 2. Table of the Basic Set of OSSE Experiments Used to Evaluate Both P3 Survey Strategies and Picket Fence Deployments of Thermistor Chains Attached to Surface Platforms^a

| Experiment | Instrument Type | Profiles | Vertical Resolution | Number of Daily Surveys | Nominal Horizontal Resolution |
|------------|---------------------|---------------|---------------------|-------------------------|----------------------------------|
| AXBT1D | Expendable profiler | T to 400 m | 2 m | 1 day | 1° |
| AXBT2D | Expendable profiler | T to 400 m | 2 m | 2 days | 0.5° |
| AXCTD1D | Expendable profiler | T,S to 1000 m | 2 m | 1 day | 1° |
| AXCTD2D | Expendable profiler | T,S to 1000 m | 2 m | 2 days | 0.5° |
| THFIX | Thermistor chain | T to 100 m | 8 m | 3 days | 1° along line; ≥2° between lines |
| THADV | Thermistor chain | T to 100 m | 8 m | 3 days | 1° along line; ≥2° between lines |

^aThis basic set was run for each storm, including both the northern and southern deployments for Gonzalo, by adding the rapid-response observations to both CONTROL and NOALT.

Table 3. Synthetic Observations Simulated to Perform the OSSEs Along With Errors Added

| Observing System Component | Instrument | Instrument Measurement RMS Error | RMS Representation Errors | Other RMS Errors |
|----------------------------|------------------------------|----------------------------------|---|---|
| Satellite altimetry | Cryosat | 0.02 m | 0.02 m; correlation length scale 40 km | Internal tides 0.01 m (length scale 5 km) |
| | Jason-2 | 0.02 m | 0.02 m; correlation length scale 40 km | Internal tides 0.01 m (length scale 5 km) |
| | Envisat | 0.02 m | 0.02 m; correlation length scale 40 km | Internal tides 0.01 m (length scale 5 km) |
| | Haiyang-2a | 0.02 m | 0.02 m; correlation length scale 40 km | Internal tides 0.01 m (length scale 5 km) |
| SST | Satellite MCSST | 0.3°C | 0.2°C | |
| | In situ fixed surface buoy | 0.1°C | 0.2°C | |
| | In situ surface drifter | 0.1°C | 0.2°C | |
| | In situ ship intake | 0.2°C | 0.2°C | |
| Argo | Argo profiling floats | T: 0.005°C S: 0.005 PSU | T: 0.15°C S: 0.08 PSU (taper to zero from surface to 200 m) | Random depth error-2 m (taper to zero above 100 m) |
| In situ XBT | Primarily ship transects | 0.05°C | Random 0.2°C (taper to zero from surface to 200 m) | Random depth error-1.5% of depth fall rate |
| AXBT | Airborne profilers | 0.02°C | Random 0.2°C (taper to zero from surface to 200 m) | Random depth error-RMS amplitude 2 m at 100 m depth |
| AXCTD | Airborne profilers | T 0.02°C S 0.05PSU | Random T 0.2°C S 0.05PSU (taper to zero from surface to 200 m) | Random depth error-RMS amplitude 2 m at 100 m depth |
| Thermistor chains | Attached to surface drifters | T 0.02°C | Random 0.2°C (taper to zero from surface to 200 m) | Random depth error above target depth (2% of depth) |

profiling temperature and salinity down to 1000 m (AXCTD1D and AXCTD2D). For each instrument, either one single-day survey was conducted (AXBT1D and AXCTD1D), or two surveys were conducted on consecutive days (AXBT2D and AXCTD2D). Each individual survey was designed so that instruments were deployed with nominal horizontal separations of about 1°. In experiments where surveys were conducted on two consecutive days, the profilers were deployed between locations sampled on day 1 to achieve nominal horizontal resolution of 0.5° (Figure 1). The FM model 0.08° horizontal resolution is adequate to accept this profile resolution.

The subset of experiments to evaluate thermistor chain deployments (Table 2) include one to evaluate cases where the surface platforms maintain a constant position (THFIX) and another to evaluate cases where the platforms (surface drifters) advect with the 15 m surface current field (THADV). The thermistor chain release points are shown in Figure 1 (right). Thermistors were deployed along “picket fence” lines with horizontal separation of 1° along each line. For all storms and subregions, three lines were deployed with horizontal separation of either 2.0 or 2.5°. The larger separation between lines enables the consequences of increasing horizontal profile separation substantially beyond 1° to be evaluated.

3.2. Quantitative Impact Assessment Procedures

Quantitative impact assessments are based on error reduction in the representation of mesoscale structure, and also on bias reduction in upper ocean heat content and stratification. Several model fields that are important to the hurricane prediction problem are analyzed. To assess impacts on ocean dynamics, dynamic height at the surface relative to 1000 m (D_{1000}), sea surface height (SSH), and the depth of the 20°C isotherm (H_{20}) are considered. In TC regions, H_{20} maps are often used as a proxy for the main thermocline and thus represent the mesoscale structure of upper ocean boundary currents and eddies [e.g., Meyers et al., 2014]. To assess impacts on ocean thermodynamics, two surface fields [SST and sea surface salinity (SSS)] and four subsurface fields [tropical cyclone heat potential (TCHP) relative to the 26°C isotherm, depth of the 26°C isotherm (H_{26}), mean temperature between the surface and 100 m (\bar{T}_{0-100}), and temperature difference at the surface relative to 100 m (ΔT_{0-100})] are considered.

Table 4. RMSE of the Three Reference Experiments With Respect to the NR: the Unconstrained FM Along With CONTROL and NOALT^a

| Storm | Field | Experiments | | |
|---------------|-----------------------------|-------------|--------------|-------------|
| | | FM RMSE | CONTROL RMSE | NOALT RMSE |
| Isaac | D ₁₀₀₀ (m) | 0.278 | 0.035 (87%) | 0.214 (23%) |
| | SSH (m) | 0.267 | 0.032 (88%) | 0.217 (19%) |
| | H ₂₀ (m) | 58.09 | 15.58 (73%) | 46.35 (20%) |
| | SST (°C) | 0.394 | 0.198 (50%) | 0.310 (21%) |
| | SSS (PSU) | 0.625 | 0.444 (29%) | 0.624 (0%) |
| | TCHP (kJ cm ⁻²) | 15.86 | 8.53 (46%) | 12.93 (18%) |
| | H ₂₆ (m) | 27.45 | 10.27 (63%) | 21.18 (23%) |
| | T ₀₋₁₀₀ (°C) | 1.46 | 0.449 (69%) | 1.15 (21%) |
| | ΔT ₀₋₁₀₀ (°C) | 2.15 | 0.488 (77%) | 1.71 (20%) |
| | D ₁₀₀₀ (m) | 0.075 | 0.036 (52%) | 0.076 (-1%) |
| Edouard | SSH (m) | 0.105 | 0.035 (67%) | 0.103 (2%) |
| | H ₂₀ (m) | 73.2 | 44.06 (40%) | 59.69 (18%) |
| | SST (°C) | 0.425 | 0.216 (49%) | 0.244 (43%) |
| | SSS (PSU) | 0.188 | 0.102 (46%) | 0.126 (33%) |
| | TCHP (kJ cm ⁻²) | 6.96 | 3.66 (47%) | 5.32 (24%) |
| | H ₂₆ (m) | 6.99 | 4.65 (34%) | 6.82 (2%) |
| | T ₀₋₁₀₀ (°C) | 0.672 | 0.449 (33%) | 0.550 (18%) |
| | ΔT ₀₋₁₀₀ (°C) | 0.610 | 0.487 (20%) | 0.570 (7%) |
| | D ₁₀₀₀ (m) | 0.092 | 0.055 (40%) | 0.081 (12%) |
| | SSH (m) | 0.132 | 0.048 (64%) | 0.105 (20%) |
| Gonzalo south | H ₂₀ (m) | 32.62 | 23.56 (28%) | 23.80 (27%) |
| | SST (°C) | 0.160 | 0.134 (16%) | 0.152 (5%) |
| | SSS (PSU) | 0.345 | 0.209 (39%) | 0.293 (15%) |
| | TCHP (kJ cm ⁻²) | 9.04 | 6.54 (28%) | 7.67 (15%) |
| | H ₂₆ (m) | 15.44 | 10.57 (32%) | 13.84 (10%) |
| | T ₀₋₁₀₀ (°C) | 0.571 | 0.362 (37%) | 0.480 (16%) |
| | ΔT ₀₋₁₀₀ (°C) | 1.16 | 0.794 (32%) | 1.01 (13%) |
| | D ₁₀₀₀ (m) | 0.088 | 0.034 (61%) | 0.070 (20%) |
| | SSH (m) | 0.099 | 0.031 (69%) | 0.086 (13%) |
| | H ₂₀ (m) | 29.12 | 23.10 (21%) | 26.46 (9%) |
| Gonzalo north | SST (°C) | 0.267 | 0.123 (54%) | 0.152 (43%) |
| | SSS (PSU) | 0.183 | 0.074 (60%) | 0.128 (30%) |
| | TCHP (kJ cm ⁻²) | 10.79 | 4.68 (57%) | 7.71 (29%) |
| | H ₂₆ (m) | 12.49 | 6.87 (45%) | 10.48 (16%) |
| | T ₀₋₁₀₀ (°C) | 0.766 | 0.350 (54%) | 0.514 (33%) |
| | ΔT ₀₋₁₀₀ (°C) | 0.802 | 0.583 (27%) | 0.808 (-1%) |

^aNumbers in parentheses give the reduction in CONTROL and NOALT RMSE with respect to FM RMSE.

TCHP, also referred to as ocean heat content [Leipper and Volgenau, 1972], is calculated by,

$$TCHP = c_p \int_0^{H_{26}} \rho [T(z) - 26] dz, \quad (1)$$

where c_p is specific heat of seawater at constant pressure and H_{26} is 26°C isotherm depth. TCHP derived from satellite altimetry [Meyers et al., 2014] is used in the Statistical Hurricane Intensity Prediction Scheme (SHIPS) to forecast intensity at the U.S. National Hurricane Center [DeMaria et al., 2005; Mainelli et al., 2008]. TCHP serves as an index of the ocean thermal energy potentially available to maintain or intensify storms [e.g., Lin et al., 2013] while the H_{26} field represents the thickness of the upper ocean layer that is sufficiently warm to support storms. \bar{T}_{0-100} was recommended by Price [2009] as an alternate to TCHP in representing thermal energy available to storms because it is defined continuously throughout the ocean as opposed to TCHP and H_{26} which truncate at zero toward higher latitudes. Finally, ΔT_{0-100} represents the difference between SST and temperature at a depth that the ocean mixed layer can reach under strong storm forcing.

Following ocean model initialization requirements, impact assessments are based on reduction of both RMS errors

and mean biases present between ocean analyses fields. Defining x as a model field representing the truth (simulated by the NR) and y as the same field from a model experiment being evaluated, these statistics are calculated by

$$RMSE = \left\{ \frac{1}{n} \sum_{i=1}^n [(y_i - \langle y \rangle) - (x_i - \langle x \rangle)]^2 \right\}^{\frac{1}{2}}, \quad B = \langle y \rangle - \langle x \rangle. \quad (2)$$

RMSE represents the accuracy with which the model field under consideration represents mesoscale structure, which is important for both model dynamical and thermodynamical fields. B represents model bias, which is evaluated only for thermodynamical fields because bias in dynamical fields SSH and D_{1000} does not affect circulation. All statistics are calculated over regions directly influenced by the rapid-response profile observations shown by the boxes in Figure 1 for the reference experiments listed in Table 1 and the OSSE experiments listed in Table 2.

4. Impact of Observational Ocean Observing Systems

Quantitative impact assessment of altimetry assimilation compared to other components of the operational ocean observing system is based in part on reduction in RMSE with respect to the truth (NR) for all nine model dynamical and thermodynamical fields (D_{1000} , SSH, H_{20} , SST, SSS, TCHP, H_{26} , \bar{T}_{0-100} , and ΔT_{0-100}). The RMSE of the FM provides the benchmark error level for each field (Table 4) which, based on the FM evaluation presented in Halliwell et al. [2017], approximately represents the large errors with which present-day

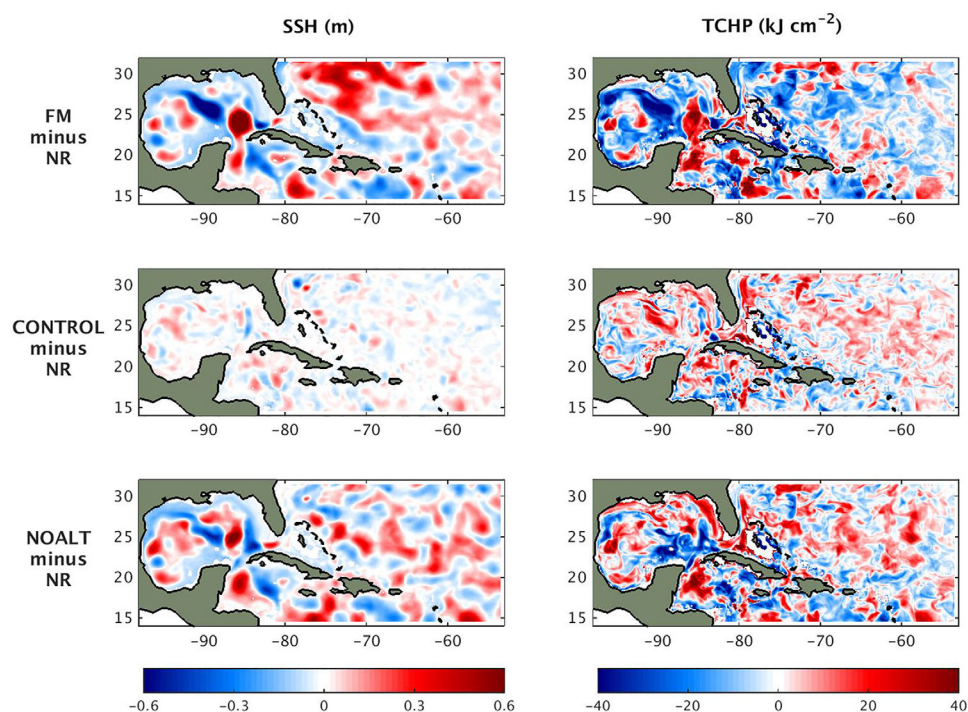


Figure 2. Error maps (difference between each experiment and the truth represented by the NR) for the unconstrained FM (top row) the CONTROL experiment (middle row), and experiment NOALT (bottom row) for SSH (left) and TCHP (right).

unconstrained ocean models represent the chaotic ocean mesoscale. Reduction in RMSE with respect to FM RMSE for experiments CONTROL and NOALT documents the relative impacts of altimetry versus other operational observing system components on these model fields (Table 4). Assimilation of SST, Argo floats, and XBT profiles alone generally produce only modest RMSE reduction, while adding the four altimeters greatly increases RMSE reduction. Altimetry is especially effective for dynamical variables D_{1000} and SSH in all storm regions because it corrects temperature and salinity profiles throughout the water column [Pascual *et al.*, 2006; Pun *et al.*, 2007]. Error reduction in the upper ocean thermodynamical variables due to altimetry assimilation is also large in the Gulf of Mexico prior to Isaac, presumably because altimetry corrects the large temperature and salinity differences associated with the Loop Current and associated eddies. The one exception is for SSS, where RMSE is not reduced at all for NOALT and is reduced by only 29% for CONTROL. This results because of limited availability of Argo floats and the inability of altimetry to provide accurate corrections in SSS. Additional RMSE reduction for thermodynamical variables resulting from altimetry assimilation remains substantial, but is generally modestly smaller in the open Atlantic regions analyzed for Edouard and Gonzalo.

To visualize corrections in mesoscale structure with respect to the truth (NR) resulting from assimilation, errors are mapped for one dynamical variable (SSH) and one thermodynamical variable (TCHP) at the height of the 2014 hurricane season (13 September) over a large section of the Atlantic domain (Figure 2). As expected, altimetry assimilation substantially reduces errors associated with mesoscale structure of dynamical fields. In the Gulf of Mexico, large SSH errors primarily result from displaced locations of the Loop Current path and associated eddies between the NR and experiment FM. In the less-energetic open Atlantic subtropical gyre, errors between the FM and NR associated with displaced mesoscale eddies dominate the error field and these errors are substantially corrected with altimetry assimilation. Errors remain large in NOALT (Figure 2) because Argo and XBT profiles do not have the horizontal resolution to substantially correct the three-dimensional structure of mesoscale features while SST products only correct near-surface temperature. Errors in thermodynamical variable TCHP (Figure 2) display similar behavior as errors in SSH (Figure 2) except that error reduction is not as large in response to altimetry assimilation. As for SSH, Argo and XBT profiles do not sample at sufficient horizontal resolution to produce substantial error reduction in the mesoscale structure of TCHP. Not only does altimetry assimilation produce a smaller correction in TCHP

mesoscale structure compared to SSH, but large-scale regions of either positive or negative bias remain after altimetry assimilation (Figure 2).

The OSSE system assimilates altimetry by correcting subsurface model profiles of layer thickness, temperature, and salinity using the *Cooper and Haines* [1996] approach that preserves water mass structure [Halliwell et al., 2014, 2017]. (HYCOM uses Lagrangian hybrid vertical coordinate layers, not Cartesian z coordinates.) While assimilation of multiple altimeters can substantially correct mesoscale structure in dynamical variables, the resulting correction of model subsurface thermodynamical fields, both mesoscale structure and bias, is less accurate [Halliwell et al., 2017]. Because the *Cooper and Haines* [1996] procedure corrects temperature and salinity profiles throughout the water column emphasizing the first baroclinic mode structure, corrections are usually suboptimal over the upper ~ 200 m including the upper ocean mixed layer. In general, higher-mode vertical structure is associated with a weak SSH signal that is not well constrained by altimetry. Consequently, rapid-response prestorm profile surveys can potentially provide additional correction in subsurface thermodynamical fields over this part of the upper ocean than is provided by altimetry assimilation.

5. Impact of Rapid-Response Ocean Surveys

5.1. OSSEs for Airborne Profile Surveys

The addition of rapid-response profile observations to CONTROL for the P3 experiments listed in Table 2 produces the RMSE reduction statistics graphed in Figure 3, while the addition of the same profile

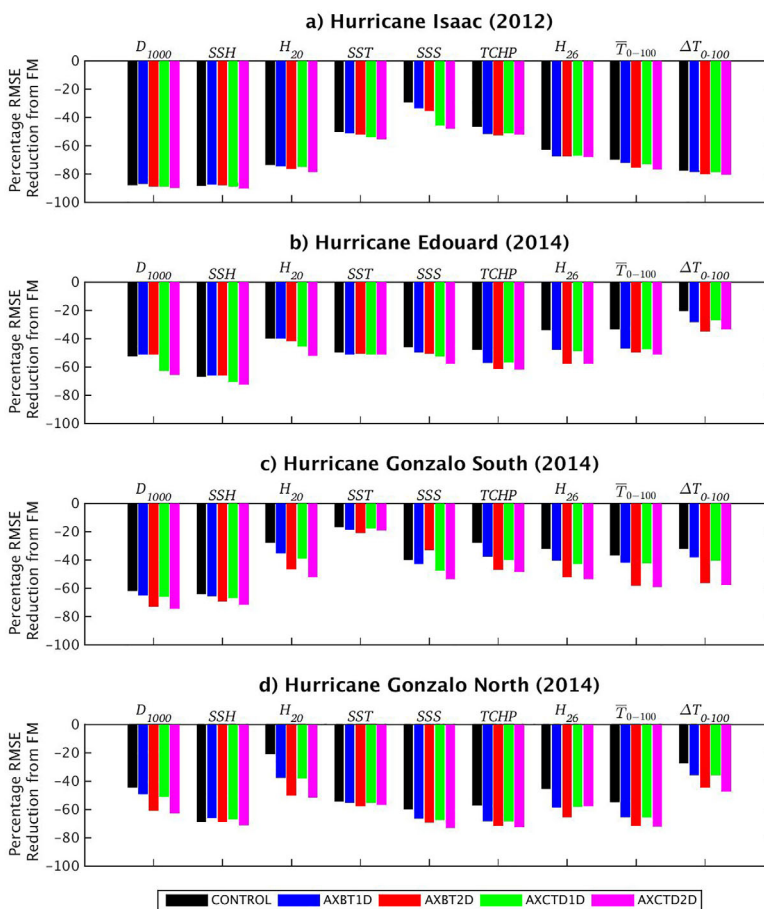


Figure 3. Percentage RMSE reduction in four hurricane survey regions with respect to the unconstrained FM for five experiments including CONTROL along with four experiments where P3 profiles were added to the operational ocean observations that were assimilated in CONTROL.

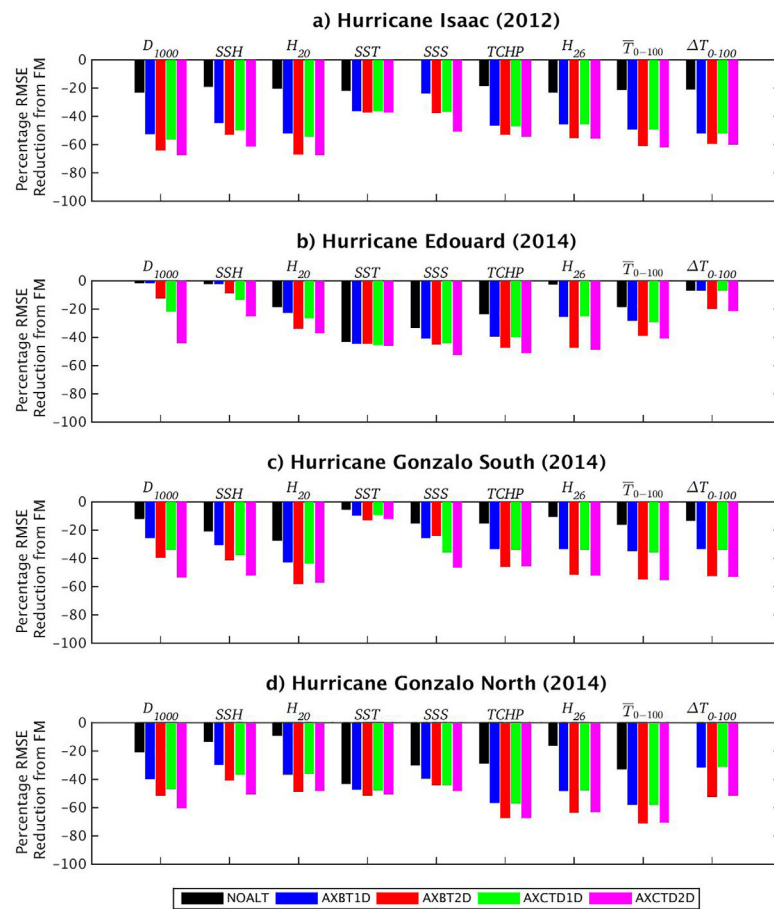


Figure 4. Percentage RMSE reduction in four hurricane survey regions with respect to the unconstrained FM for five experiments including NOALT (which did not assimilate altimetry) along with four experiments where P3 profiles were added to the observations assimilated in NOALT.

observations to NOALT produces the RMSE reduction statistics graphed in Figure 4. Because other components of the ocean observing system besides altimetry produce relatively small RMSE reduction in experiment NOALT for all model fields except SSS, the addition of rapid-response profiles tends to produce large additional corrections (Figure 4). With altimetry assimilation providing substantial additional RMSE reduction, the addition of rapid-response profiles to CONTROL produces substantially smaller additional corrections (Figure 3). Intermediate results are expected if fewer than four altimeters are assimilated. The impact of rapid-response surveys therefore depends on the number of altimeters that are assimilated along with the accuracy of the method used to assimilate altimetry.

With four altimeters assimilated, altimetry correction of mesoscale structure based on RMSE reduction in the Gulf of Mexico for Isaac is large for all fields except SSS (Figure 3). Consequently, additional corrections resulting from profile assimilation are small except for SSS. In the other open-ocean regions, the largest additional corrections are observed for subsurface thermal fields (H_{20} , TCHP, H_{26} , \bar{T}_{0-100} , and ΔT_{0-100}), suggesting that corrections resulting from profiler assimilation are particularly important close to and above the main thermocline where altimetry corrections of upper ocean thermal fields is presumably less accurate. Additional corrections for dynamical field SSH, which tend to be slightly larger for experiments assimilating AXCTDs down to 1000 m, are relatively small. By contrast, the additional corrections for dynamical field D_{1000} tend to be larger than for SSH because the profilers correct density over all (AXCTDs), or a substantial fraction of (AXBTs), the 1000 m depth range over which this field is calculated. Profilers only correct density over a small fraction of the full water column in deep water, so the additional SSH correction is proportionately smaller than for D_{1000} . Without altimetry assimilation, profile assimilation produces large correction in the mesoscale structure of both of these dynamical variables.

Breaking down results by the impact of other factors, comparison of the 1 day and 2 day experiments demonstrates that second-day surveys generally produce additional RMSE reduction whether or not altimetry is assimilated (Figures 3 and 4). Given that day 2 surveys were designed to collect profiles at locations between the locations sampled on day 1, and thus reduce nominal separation distances from 1.0° to 0.5° , the smaller scales resolved when the two consecutive days of profiles are assimilated do not lead to large additional RMSE reduction. Comparing corresponding experiments that assimilate 400 m AXBTs and 1000 m AXCTDs, similar corrections are achieved for all subsurface thermal fields whether or not altimetry is assimilated (Figures 3 and 4) because both profiler types measure temperature within the depth range over which these fields are calculated. However, in the absence of altimetry assimilation, large additional RMSE reduction is achieved for dynamical fields with AXBT assimilation, but a substantially larger reduction is achieved with AXCTD assimilation in all regions except in the Gulf of Mexico where the reduction is only slightly larger (Figure 4). With altimetry assimilation, the advantage of AXCTD assimilation over AXBT assimilation is still present, but greatly reduced (Figure 3). Correction of SSS is always improved with AXCTD assimilation regardless of whether altimetry is assimilated.

The DA method used for the OSSE system [Halliwell et al., 2014] assimilates ocean profiles discretized in the hybrid vertical layered coordinate structure of HYCOM [Bleck, 2002; Chassignet et al., 2003; Halliwell, 2004], which then requires that the observed AXBT profiles in z coordinates first be remapped to model hybrid vertical coordinates [Halliwell et al., 2017]. This in turn requires that synthetic salinity profiles be generated based on climatological temperature-salinity correlations. Present results suggest that this approach displays substantial skill but is still inferior to actually measuring salinity. However, with the assimilation of four altimeters, the advantage of actually measuring salinity is substantially reduced. Overall, an airborne AXBT survey conducted on a single day with nominal 1° profile separation does provide substantial additional RMSE reduction compared to ocean analyses that just assimilate the existing ocean observing system. Enhancements such as sampling over 2 days to cover higher horizontal resolution or deploying deeper AXCTDs does further improve RMSE reduction by an amount that depends on the number of altimeters that are available for assimilation.

Actual corrections achieved in profile assimilation by experiment AXCTD2D are mapped as the difference AXCTD2D minus CONTROL (Figure 5) for one dynamical variable (SSH) and two thermodynamical variables (TCHP and ΔT_{100}) for all storms and regions. One takeaway is that corrections resulting from rapid-response prestorm surveys are strongly confined to the region that is directly sampled. This results because the radius of influence of individual observations, which in the present DA system [Halliwell et al., 2014] depends on both horizontal correlation scale, which is typically $O(100)$ km, and the localization radius of 200 km for ocean profiles that is imposed to terminate influence over distances too large to permit significant correlation. The corrections have insufficient time to spread beyond the sampling area before the storm arrives. [see Halliwell et al., 2017 for a discussion on how these system design parameters affect profile assimilation.] Another takeaway is that the nominal 0.5° horizontal profile spacing is highly adequate to resolve meso-scale structure. The fact that RMSE reduction was only modestly smaller for experiment AXCTD1D with nominal 1° resolution (not shown) indicates that the single-day surveys with 1.0° resolution correct meso-scale structure with nearly as much accuracy. This result is consistent with the Gulf of Mexico OSSEs presented in Halliwell et al. [2015] which demonstrated that mesoscale corrections rapidly degrade only as horizontal profile separations exceed 1° .

Correcting bias in upper ocean thermodynamical fields is also very important to the hurricane prediction problem. Bias with respect to the truth (NR) is summarized for three model thermodynamical fields (TCHP, ΔT_{0-100} , and SST) in all storm regions in Table 5 for experiments FM, CONTROL, and four rapid-response P3 surveys where the profiles were added to CONTROL. Experiment CONTROL reduces bias magnitude in all three fields in the Gulf of Mexico, but does not consistently reduce bias magnitude in the other storm regions, particularly where small bias magnitude exists in experiment FM. When the airborne profiles are added, bias correction for SST is only important for Isaac in the Gulf of Mexico. However, bias correction is more important everywhere for the two subsurface fields, particularly for TCHP. With the exception of Edouard for TCHP and Isaac for ΔT_{0-100} , second-day surveys provide substantial additional bias reduction compared to single-day surveys for both AXBTs and AXCTDs. In the Gonzalo south region, AXCTDs achieved additional bias reduction for both 1 day and 2 day surveys.

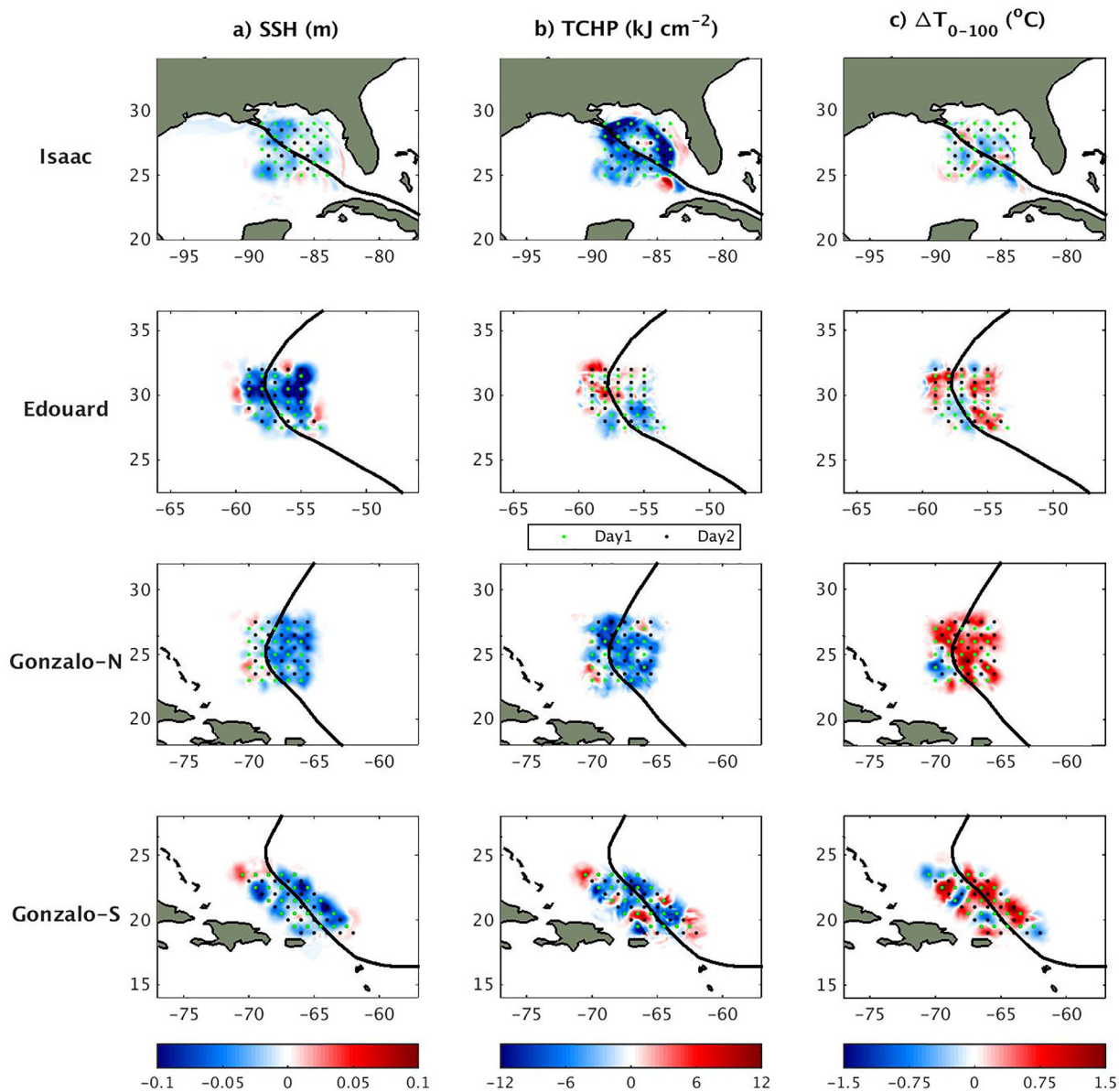


Figure 5. Field corrections resulting from adding the assimilation of P3 profiles from experiment AXCTD2D to the assimilation of operational ocean observations (AXCTD2D minus CONTROL). Each row shows a different hurricane survey region for column (a) SSH, (b) TCHP, and (c) ΔT_{100} .

5.2. OSE for Airborne Profile Survey

An actual prestorm survey was conducted for Edouard on 12 September 2014 by a WP-3D hurricane research aircraft (Figure 6). A mixture of AXBTs, AXCTDs, and AXCPs were deployed. Two lines crossing the projected storm track were run with nominal profile resolution of 0.5° along the lines and with $\sim 3^{\circ}$ separation between lines. An OSE is conducted to study the impact of this survey and is compared to an identical OSSE that assimilated synthetic versions of all observations including the 12 September survey. To perform the OSE, a special CONTROL experiment (OSECONTROL) was run that was identical to the 2014 OSSE CONTROL experiment (Table 1), but assimilated all real operational ocean observations including the four altimeters. Experiment OSE was then initialized from OSECONTROL on 11 September and run over a 4 day interval adding profile assimilation on 12 September to the other observations. Experiment OSSE is identical to OSE but assimilates synthetic observations.

The ability to evaluate the impact of these real profile observations is limited due to the absence of actual observations of the true ocean. In particular, there are no Argo or XBT profiles available from the

Table 5. Bias in Three Model Fields for the Storms Tabulated for Experiments FM, CONTROL, and the Four P3 Surveys^a

| Storm | Experiment | TCHP Bias (kJ cm ⁻²) | ΔT Bias (°C) | SST Bias (°C) |
|---------------|------------|-------------------------------------|-----------------|------------------|
| Isaac | FM | -8.00 | -0.59 | -0.55 |
| | CONTROL | 5.51 | 0.11 | 0.19 |
| | AXBT1D | 3.65 (34%) | -0.16 | 0.12 (37%) |
| | AXBT2D | 2.25 (59%) | 0.25 | 0.09 (53%) |
| | AXCTD1D | 3.34 (39%) | -0.08 (27%) | 0.11 (42%) |
| Edouard | AXCTD2D | 2.02 (63%) | -0.10 (9%) | 0.08 (58%) |
| | FM | -1.68 | -0.43 | -0.21 |
| | CONTROL | 1.12 | -0.33 | -0.22 |
| | AXBT1D | -0.39 (65%) | -0.24 (27%) | -0.23 |
| | AXBT2D | -0.48 (57%) | -0.21 (36%) | -0.22 |
| Gonzalo south | AXCTD1D | -0.36 (68%) | -0.20 (39%) | -0.23 |
| | AXCTD2D | -0.50 (55%) | -0.18 (45%) | -0.23 |
| | FM | -2.31 | 0.05 | -0.15 |
| | CONTROL | 2.42 | -0.32 | -0.02 |
| | AXBT1D | 1.80 (26%) | -0.24 (25%) | -0.02 |
| Gonzalo north | AXBT2D | 0.35 (85%) | -0.14 (56%) | -0.03 |
| | AXCTD1D | 1.27 (48%) | -0.17 (47%) | -0.03 |
| | AXCTD2D | 0.07 (97%) | -0.14 (56%) | -0.04 |
| | FM | -5.55 | -0.09 | -0.38 |
| | CONTROL | 4.61 | -0.51 | -0.06 |
| | AXBT1D | 2.22 (52%) | -0.40 (22%) | -0.05 (17%) |
| | AXBT2D | 0.56 (88%) | -0.19 (63%) | -0.05 (17%) |
| | AXCTD1D | 2.26 (51%) | -0.38 (25%) | -0.06 |
| | AXCTD2D | 0.49 (89%) | -0.17 (67%) | -0.05 (17%) |

^aPercentage bias magnitude reduction with respect to CONTROL is presented in parentheses for all cases where P3 assimilation actually reduced the magnitude.

operational observing system immediately before and after September 12 that are located within the region influenced by the rapid-response profile survey. Consequently, analysis is limited to mapping the actual corrections achieved by the profile assimilation as the difference OSE minus OSECONTROL and comparing these to the difference OSSE minus CONTROL (Figure 6). Maps are presented for the same three fields analyzed for the OSSE experiments in Figure 5. The key result is the impact of having high cross-track resolution but low along-track resolution in the ocean profiles. Eddy structure is well resolved in the cross-track direction, but a gap between the lines exists where little correction is achieved due again to the limited radius of influence of individual observations and the insufficient time for corrections to spread. The OSE and OSSE experiments are in agreement.

5.3. OSSEs for Thermistor Chain Deployments

The addition of rapid-response profile observations to CONTROL for the thermistor chain experiments listed in Table 2 produces the RMSE reduction statistics graphed in Figure 7, while the addition of the same profile observations to NOALT produces the RMSE reduction statistics graphed in Figure 8. With four altimeters assimilated, the thermistor chain assimilation provides little or no additional RMSE reduction (Figure 7). Without altimetry assimilation, modest additional RMSE reduction is achieved for many fields over the four regions (Figure 8), primarily for the subsurface thermal fields TCHP, H_{26} , ΔT_{0-100} , and δT_{0-100} . Whether the thermistor chains advect with the 15 m flow or are steered to remain stationary does not make a significant difference.

Actual corrections achieved by assimilating stationary thermistor chains are mapped as the difference THFIX minus CONTROL (Figure 9) for one dynamical variable (SSH) and two thermodynamical variables (TCHP and ΔT_{100}) for all storms and regions. The same differences are mapped for advecting thermistor chains (THADV minus CONTROL) in Figure 10. One factor limiting the impact of these deployments is the relatively large separation between lines of either 2.0° or 2.5°. As for the OSE and corresponding OSSE that evaluated the actual Edouard prestorm survey (Figure 6), this large separation leaves large gaps in the correction field that then introduces errors in the mesoscale field structure within the region sampled. Two other factors limit the effectiveness of these thermistor chains in comparison to airborne profile surveys. First, temperature is measured only over the upper 100 m extending to a limited distance below. This allows modest improvement of subsurface thermal fields only over this depth range. Second, less-accurate estimated salinity must be assimilated. As a result of all of these factors, RMSE reduction is not realized for dynamical variables, D_{1000} , H_{20} , and SSS.

Despite this difficulty in correcting mesoscale structure, substantial bias reduction is achieved for subsurface thermal fields TCHP and to a lesser extent ΔT_{0-100} (Table 6). Results are similar for fixed and advecting thermistor chains. Bias reduction is generally insignificant for SST, in particular because large bias reduction is already achieved by experiment CONTROL which assimilates satellite and in situ SST. Overall, modest positive impacts are realized by thermistor chain deployments, but due to shallow depth coverage, no salinity measurements, and in particular the large distances between lines in the experiments reported here, they were substantially less effective than the

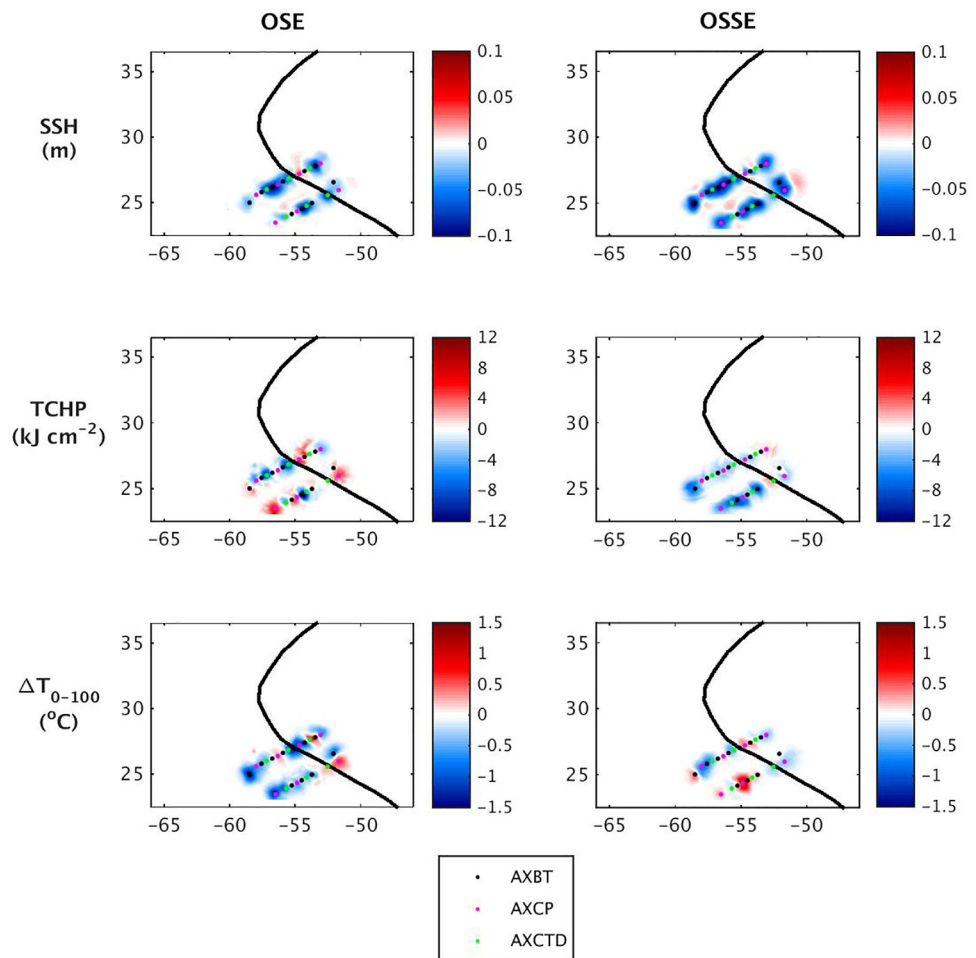


Figure 6. (left) OSE results for three model fields: field corrections resulting from adding the actual WP-3D airborne profiles collected prior to the passage of Edouard on 12 September 2014. (right) Results from an identical OSSE that assimilated synthetic versions of all observations.

airborne survey experiments. Multiple picket fence lines should be separated by no more than $\sim 1^\circ$, which requires a large number of lines to cover large along-track distances and limits cost-effectiveness.

6. Discussion

OSSEs enable quantitative impact assessments to be performed for new ocean observing systems and for alternate deployment strategies for existing systems. Furthermore, because OSSE systems provide a high-resolution, three-dimensional representation of the truth, detailed evaluations of both existing and new ocean observing system impacts with respect to the truth can be performed that are not possible by other means, particularly with regard to improving the analysis and prediction of ocean mesoscale structure. These OSSE system capabilities were exploited to perform quantitative impact assessments of design strategies for airborne profile surveys conducted in a lawnmower pattern, and also for thermistor chains attached to surface platforms and deployed in lines that cross predicted storm paths. The assessments were based on important requirements of the TC prediction problem, specifically reduction in ocean mesoscale errors for all model fields, bias reduction in upper ocean thermal fields, and bias reduction in vertical temperature structure within and immediately beneath the ocean mixed layer. Observations are evaluated only in terms of their ability to improve prestorm initialization of model fields over the

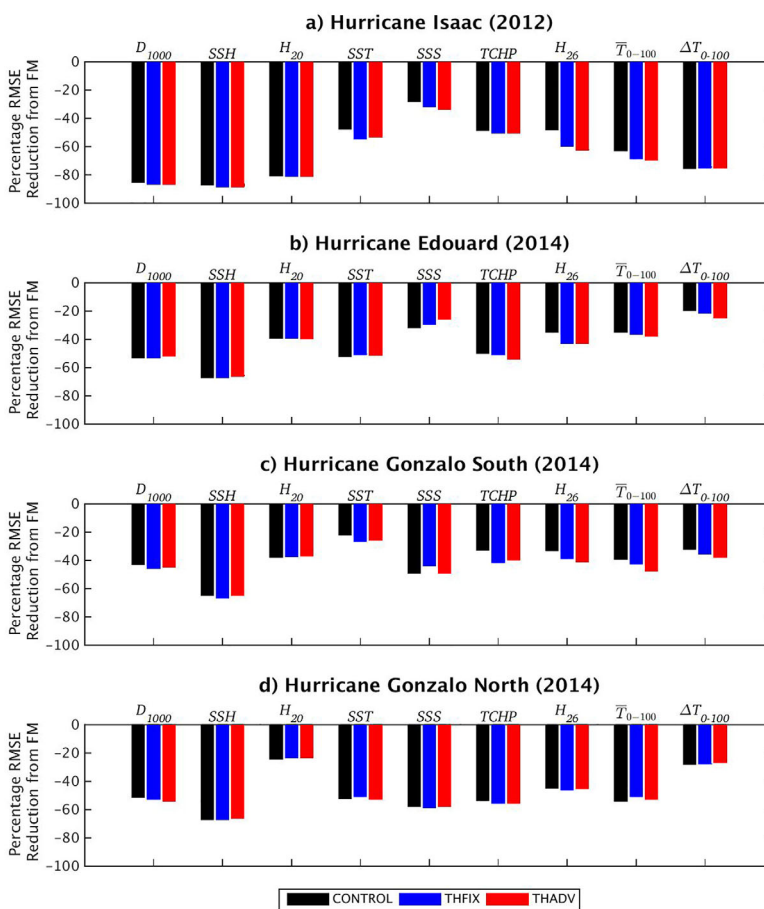


Figure 7. Percentage RMSE reduction with respect to the unconstrained FM for three experiments including CONTROL along with two experiments where thermistor chain profiles were added to the operational ocean observations that were assimilated in CONTROL.

open ocean. Regardless of their impact for this specific application, the rapid-response observations considered herein also have high value in sampling ocean profiles beneath and following storms for model evaluation and scientific analysis.

Because observations from rapid-response surveys are added to existing ocean observations, their impact was evaluated in terms of their added value over the existing operational observing system. Impacts of altimetry versus other FM observing system components were considered separately. Without altimetry assimilation, other components of the observing system produced relatively small corrections to the mesoscale structure of model fields. The horizontal distribution of subsurface Argo and XBT profiles was not sufficiently dense to provide large corrections. Satellite and in situ SST observations were sufficiently dense to resolve horizontal mesoscale structure, but SST is only assimilated within the ocean mixed layer [Halliwell et al., 2017], which prevents the correction of mesoscale features over their full vertical extent. This DA system feature is beneficial for TC-related applications. In tropical regions during TC season, SST and mixed-layer temperature tend to be horizontally uniform due to strong solar heating, which masks the horizontal structure in temperature associated with ocean boundary currents and eddies that exists beneath the mixed layer [e.g., Jaimés and Shay, 2009]. If the DA procedure projects this horizontally uniform structure beneath the mixed layer, it will degrade the representation of mesoscale features. Without altimetry assimilation, adding rapid-response observations where profilers are deployed with sufficient horizontal density to resolve mesoscale structure ($\sim 1^\circ$ separation) produced large additional reduction in mesoscale errors. Adding the assimilation of four altimeters produced a large reduction in mesoscale errors of dynamical fields such as D_{1000} and SSH so that assimilation of rapid-response profiles resulted in only small additional error reduction. However, altimetry assimilation produced smaller mesoscale error reduction in subsurface

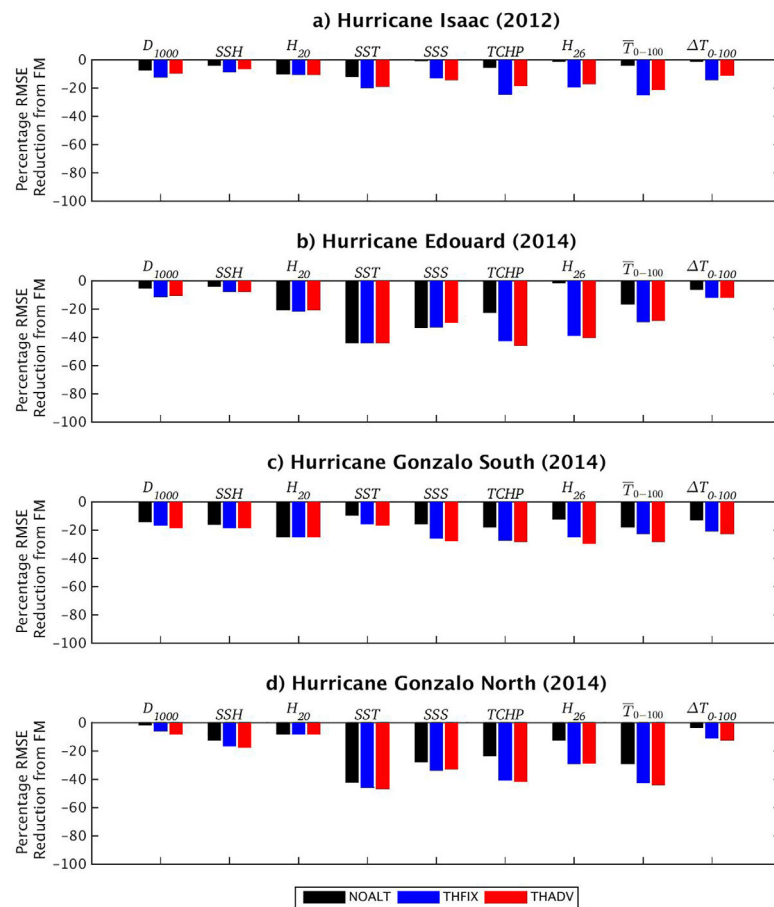


Figure 8. Percentage RMSE reduction with respect to the unconstrained FM for three experiments including NOALT along with two experiments where thermistor chain profiles were added to the operational ocean observations that were assimilated in NOALT.

thermodynamical fields, allowing rapid-response ocean profiles to provide more substantial additional error reduction.

The following recommendations concerning rapid-response ocean survey design are inferred from the present OSSE results:

1. Impact of rapid response surveys depends on the RMSE reduction already achieved by assimilating the existing operational ocean observing systems, in particular satellite altimetry. Increasing the number of altimeters decreases the additional mesoscale error reduction achieved by assimilating the profiles, although error reduction remains important especially for subsurface thermodynamical fields even with up to four available altimeters.
2. Based on the present results and earlier results from the Gulf of Mexico [Halliwell et al., 2015], multiple ocean profiles should be collected at a nominal horizontal resolution of $\sim 1^\circ$ to adequately resolve mesoscale ocean features. Small additional improvement is realized with higher sampling resolution, but $\sim 1^\circ$ represents a cost-effective compromise.
3. In the same vein, multiple picket fence lines should be separated by $\sim 1^\circ$ or less with at least 1° along-line resolution. Line separation $> 1^\circ$ leaves correction gaps that negatively affect the correction of mesoscale structure. This is a disadvantage compared to airborne surveys that can cover larger areas in a single-day survey with adequate resolution between lines.
4. Small radii of influence plus lack of spreading time confines results to the immediate region containing the sampling locations. Coupled with the previous two recommendations, it is necessary to sample relatively large areas at $\sim 1^\circ$ horizontal resolution to maximize the potential impact of observations on intensity forecasts.

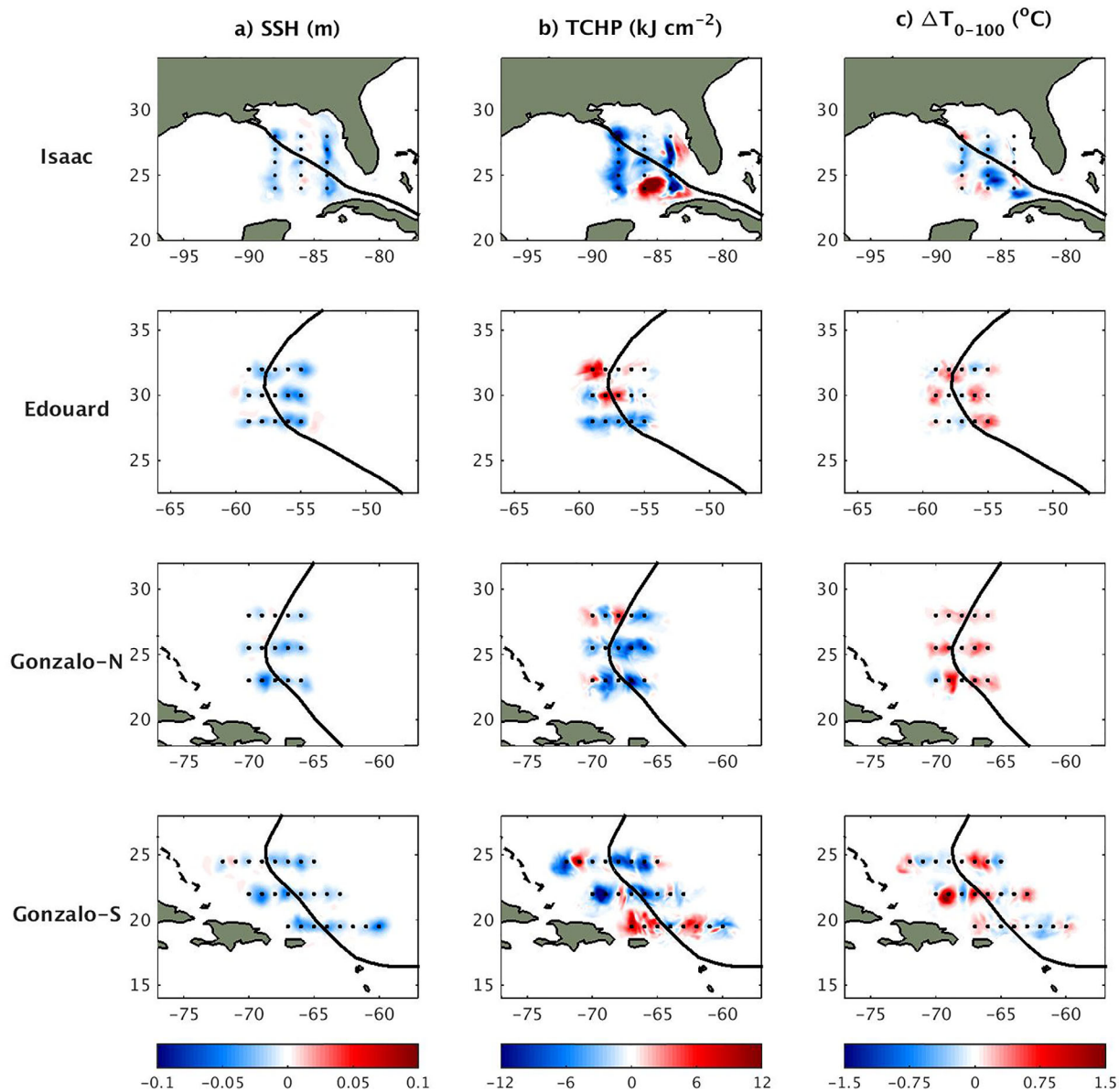


Figure 9. Field corrections resulting from adding the assimilation of thermistor chain temperature profiles from experiment THFIX to the assimilation of operational ocean observations (THFIX minus CONTROL).

5. Profilers that only measure temperature require that salinity be estimated by climatological temperature-salinity correlations. Because estimated salinity has larger errors, RMSE reduction in model dynamical variables is smaller when AXBTs are assimilated compared to AXCTDs. However, RMSE reduction in upper ocean thermal fields is not significantly different whether AXBTs or AXCTDs are assimilated.
6. Although a single-day survey typically achieves substantial correction, additional correction is achieved for most fields in most regions by sampling on two consecutive days.
7. Thermistor chains deployments generally have less impact than airborne surveys: In the present study, this reduced impact resulted from the large ($\geq 2^\circ$) separation between “picket fence” lines, from the limited vertical extent of the profiles (100 m), and by the requirement to estimate salinity based on temperature-salinity correlations. A single line corrects RMSE over only a short distance along the track of the storm, and would likely have minimal impact on intensity. Significant improvement could be realized by deploying multiple lines separated by only $\sim 1^\circ$, but this would be costly.

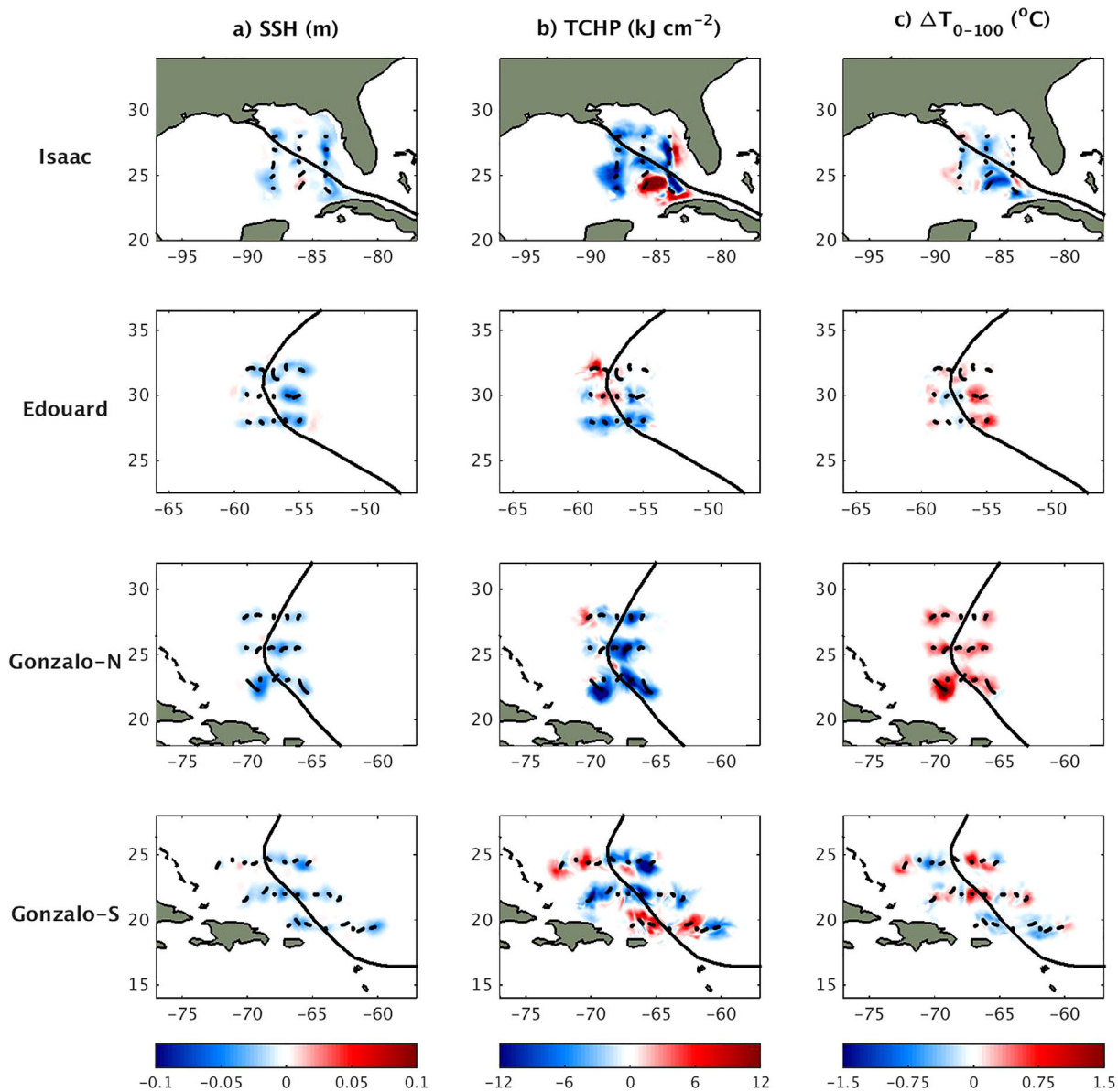


Figure 10. Same as Figure 9, but allowing the surface drifters to advect with the flow field (experiment THADV).

8. Thermistor chains can be deployed on surface platforms that passively drift with the 15 m current (surface drifters) or on powered surface platforms. Differences in RMSE reduction between assimilating drifting and stationary platforms over a 2–3 day interval before storms were not significant.
9. Bias reduction is important for upper ocean thermodynamical fields to avoid making too much or too little thermal energy available to storms, and also to avoid upper ocean stratification that is unrealistically strong or weak. Both airborne surveys and thermistor chain deployments reduced bias in subsurface thermal fields.

It is important to consider that impacts have been assessed based on error and bias reduction in initial ocean fields. Future work is planned to extend impact assessments to error reduction in TC intensity forecasts using coupled prediction systems.

The importance of the ocean response as reported in the literature and highlighted in other papers within this special issue demonstrates the necessity of correctly representing the ocean in coupled prediction systems. Improvements in ocean observation will be an important step in this direction. The present results

Table 6. Bias in Three Model Fields for the Storms Tabulated for Experiments FM, CONTROL, and the Two Thermistor Chain Experiments^a

| Storm | Experiment | TCHP Bias (kJ cm ⁻²) | ΔT Bias (°C) | SST Bias (°C) |
|---------------|------------|----------------------------------|--------------|---------------|
| Isaac | FM | -5.46 | -0.92 | -0.54 |
| | CONTROL | 4.61 | 0.12 | 0.07 |
| | THFIX | 2.88 (38%) | -0.05 (58%) | 0.07 |
| | THADV | 3.08 (33%) | -0.10 (17%) | 0.07 |
| Edouard | FM | -2.63 | -0.59 | -0.36 |
| | CONTROL | 1.03 | -0.09 | -0.08 |
| | THFIX | 0.00 (100%) | -0.12 | -0.09 |
| | THADV | -0.06 (94%) | -0.11 | -0.08 |
| Gonzalo south | FM | -2.48 | -0.05 | -0.20 |
| | CONTROL | 2.45 | -0.31 | -0.08 |
| | THFIX | 1.49 (39%) | -0.28 (10%) | -0.09 |
| | THADV | 1.36 (44%) | -0.32 | -0.08 |
| Gonzalo north | FM | -6.18 | -0.03 | -0.39 |
| | CONTROL | 4.52 | -0.49 | -0.05 |
| | THFIX | 2.62 (42%) | -0.39 (20%) | -0.04 (20%) |
| | THADV | 1/86 (59%) | -0.31 (37%) | -0.05 |

^aPercentage bias magnitude reduction with respect to CONTROL is presented in parentheses for all cases where thermistor chain assimilation actually reduced the magnitude.

document positive value in conducting prestorm rapid-response surveys and adding these observations to the existing operational system. Given the strong dependence of impacts on regional ocean conditions that is documented herein, this work must be expanded to other storms in other global TC regions to provide a comprehensive understanding of observing system impacts. Ocean OSSEs will be a key component of that effort.

Acknowledgments

The authors acknowledge support from the NOAA Hurricane Sandy Disaster Relief Act [OAR-M8R2WHSP01 and NA13OAR4830224] and from the NOAA Quantitative Observing System Assessment Program [QOSAP; OAR-P8R2W02PQF and NA15OAR4320064]. L. K. Shay gratefully acknowledges support from NASA grant NNX15AG43G. G. Halliwell and M. Mehari acknowledge internal support from the NOAA/AOML Physical Oceanography Division. The authors acknowledge the NOAA Research and Development High Performance Computing Program (<http://rdhpcs.noaa.gov>) for providing computing and storage resources that have contributed to the research results presented in this paper. All model output generated in this study is available at NOAA/AOML.

References

Androulidakis, Y., V. Kourafalou, G. Halliwell, M. Le Hénaff, H. Kang, M. Mehari, and R. Atlas (2016), Hurricane interaction with the upper ocean in the Amazon-Orinoco plum region, *Ocean Dyn.*, *66*, 1559–1588.

Arnold, C. P., and C. H. Dey (1986), Observing system simulation experiments: Past, present, and future, *Bull. Am. Meteorol. Soc.*, *67*, 687–695.

Atlas, R. (1997), Atmospheric observations and experiments to assess their usefulness in data assimilation, *J. Meteorol. Soc. Jpn.*, *75*, 111–130.

Atlas, R., and L. P. Riishojgaard (2008), Application of OSSEs to observing system design, in *Proc. SPIE 7087, Remote Sensing System Engineering 708707* (August 25, 2008), doi:10.1117/12.795344.

Balaguru, K., G. R. Foltz, L. R. Leung, E. D’Asaro, K. A. Emanuel, H. Liu, and S. E. Zedler (2015), Dynamic potential intensity: An improved representation of the ocean’s impact on tropical cyclones, *Geophys. Res. Lett.*, *42*, 6739–6746, doi:10.1002/2015GL064822.

Bleck, R. (2002), An oceanic general circulation framed in hybrid isopycnic-Cartesian coordinates, *Ocean Modell.*, *4*, 55–88.

Chan, J. C. L., Y. Duan, and L. K. Shay (2001), Tropical cyclone intensity change from a simple ocean-atmosphere coupled model, *J. Atmos. Sci.*, *58*, 154–172.

Cione, J. J. (2015), The relative roles of the ocean and atmosphere as revealed by buoy air-sea observations in hurricanes, *Mon. Weather Rev.*, *143*, 904–913.

Cione, J. J., and E. W. Uhlhorn (2003), Sea surface temperature variability in hurricanes: Implications with respect to intensity change, *Mon. Weather Rev.*, *131*, 1783–1796.

Chassignet, E. P., L. T. Smith, G. R. Halliwell, and R. Bleck (2003), North Atlantic simulation with the HYbrid Coordinate Ocean Model (HYCOM): Impact of the vertical coordinate choice, reference density, and thermobaricity, *J. Phys. Oceanogr.*, *33*, 2504–2526.

Cooper, M., and K. Haines (1996), Altimetric assimilation with water property conservation, *J. Geophys. Res.*, *101*, 1059–1078.

Cummings, J. A., and O.-M. Smedstad (2014), Ocean data impacts in global HYCOM, *J. Atmos. Oceanic Technol.*, *31*, 1771–1791.

D’Asaro, E. A., et al. (2014), The impact of typhoons on the ocean in the Pacific, *Bull. Am. Meteorol. Soc.*, *95*, 1405–1418.

DeMaria, M., M. Mainelli, L. K. Shay, J. A. Knaff, and J. Kaplan (2005), Further improvements to the statistical hurricane intensity prediction scheme (SHIPS), *Weather Forecasting*, *20*, 531–543.

Goni, G., et al. (2009), Applications of satellite-derived ocean measurements to tropical cyclone intensity forecasting, *Oceanography*, *22*, 176–183.

Halliwell, G. R., Jr. (2004), Evaluation of vertical coordinate and vertical mixing algorithms in the hybrid-coordinate ocean model (HYCOM), *Ocean Modell.*, *7*, 285–322.

Halliwell, G. R., Jr., A. Srinivasan, V. Kourafalou, H. Yang, D. Willey, M. Le Hénaff, and R. Atlas (2014), Rigorous evaluation of a fraternal twin ocean OSSE system in the open Gulf of Mexico, *J. Atmos. Oceanic Technol.*, *31*, 105–130, doi:10.1175/JTECH-D-13-00011.1.

Halliwell, G. R., Jr., S. Gopalakrishnan, F. Marks, and D. Willey (2015), Idealized study of ocean impacts on tropical cyclone intensity forecasts, *Mon. Weather Rev.*, *143*, 1142–1165.

Halliwell, G. R., Jr., M. Mehari, M. Le Hénaff, V. H. Kourafalou, Y. S. Androulidakis, H.-S. Kang, and R. Atlas (2017), North Atlantic Ocean OSSE system: Evaluation of operational ocean observing system components and supplemental seasonal observations for potentially improving coupled tropical cyclone prediction, *J. Oper. Oceanogr.*, *1*–22, doi:10.1080/1755876X.2017.1322770.

Hoffman, R. N., and R. Atlas (2015), Future Observing System Simulation Experiments, *Bull. Am. Meteorol. Soc.*, *96*, 1601–1616.

Hong, X. D., S. W. Chang, S. Raman, L. K. Shay, and R. Hodur (2000), The interaction between Hurricane Opal (1995) and a warm core ring in the Gulf of Mexico, *Mon. Weather Rev.*, *128*, 1347–1365.

- Hormann, V., L. R. Centurioni, L. Rainville, C. M. Lee, and L. J. Braasch (2014), Response of upper ocean currents to Typhoon Fanapi, *Geophys. Res. Lett.*, *41*, 3995–4003, doi:10.1002/2014GL060317.
- Jacob, S. D., L. K. Shay, A. J. Mariano, and P. G. Black (2000), The 3-D oceanic mixed-layer response to Hurricane Gilbert, *J. Phys. Oceanogr.*, *30*, 1407–1429.
- Jaimes, B., and L. K. Shay (2009), Mixed layer cooling in mesoscale oceanic eddies during hurricanes Katrina and Rita, *Mon. Weather Rev.*, *137*, 4188–4207.
- Jaimes, B., and L. K. Shay (2010), Near-inertial wave wake of hurricanes Katrina and Rita over mesoscale ocean eddies, *J. Phys. Oceanogr.*, *40*, 1320–1337.
- Jaimes, B., L. K. Shay, and G. R. Halliwell (2011), The response of quasigeostrophic vortices to tropical cyclone forcing, *J. Phys. Oceanogr.*, *41*, 1965–1985.
- Jaimes, B., L. K. Shay, and E. W. Uhlhorn (2015), Enthalpy and momentum fluxes during hurricane Earl relative to underlying ocean features, *Mon. Weather Rev.*, *143*, 111–131.
- Jaimes, B., L. K. Shay, and J. K. Brewster (2016), Observed air-sea interactions in tropical cyclone Isaac over Loop Current mesoscale eddy features, *Dyn. Atmos. Oceans*, *76*, 306–324.
- Kourafalou, V., Y. Androulikakis, G. R. Halliwell Jr., H. Kang, M. Mehari, M. Le Hénaff, R. Atlas, and R. Lumpkin (2016), North Atlantic ocean OSSE system development: Nature run evaluation and application to hurricane interaction within the Gulf Stream, *Prog. Oceanogr.*, *148*, 1–25, doi:10.1016/j.pocean.2016.09.001.
- Leipper, D. F., and D. Volgenau (1972), Hurricane heat potential of the Gulf of Mexico, *J. Phys. Oceanogr.*, *2*, 218–224.
- Lenain, L., and W. K. Melville (2014), Autonomous surface vehicle measurements of the ocean's response to Tropical Cyclone Freda, *J. Atmos. Oceanic Technol.*, *31*, 2169–2190.
- Lin, I. I., C. C. Wu, K. A. Emanuel, I. H. Lee, C. R. Wu, and I. F. Pun (2005), The interaction of Supertyphoon Maemi (2003) with a warm ocean eddy, *Mon. Weather Rev.*, *133*, 2635–2649.
- Lin, I. I., I. F. Pun, and C. C. Wu (2009), Upper-ocean thermal structure and the western North Pacific category-5 typhoons: Part II: Dependence on translation speed, *Mon. Weather Rev.*, *137*, 3744–3757.
- Lin, I. I., G. A. Goni, J. A. Knaff, C. Forbes, and M. M. Ali (2013), Ocean heat content for tropical cyclone intensity forecasting and its impact on storm surge, *Nat. Hazards*, *66*, 1481–1500.
- Lloyd, I. D., and G. A. Vecchi (2011), Observational evidence for oceanic controls on hurricane intensity, *J. Clim.*, *24*, 1138–1153.
- Lumpkin, R., T. M. Özgökmen, and L. Centurioni (2016), Advances in the application of surface drifters, *Ann. Rev. Mar. Sci.*, *9*(1), 59–81.
- Ma, Z., J. Fei, L. Liu, X. Huang, and X. Cheng (2013), Effects of the cold-core eddy on tropical cyclone intensity and structure under idealized air-sea interaction conditions, *Mon. Weather Rev.*, *141*, 1285–1303.
- Mai, W., C. Pasquero, and F. Primeau (2012), The effect of translation speed upon the intensity of tropical cyclones over the tropical ocean, *Geophys. Res. Lett.*, *39*, L07801, doi:10.1029/2011GL050765.
- Mainelli, M., M. DeMaria, L. K. Shay, and G. Goni (2008), Application of oceanic heat content estimation to operational forecasting of recent category 5 hurricanes, *Weather Forecasting*, *23*, 3–16.
- Meyers, P. C., L. K. Shay, and J. K. Brewster (2014), The development of the systematically merged Atlantic regional temperature and salinity climatology for hurricane intensity forecasting, *J. Atmos. Oceanic Technol.*, *31*, 131–149.
- Meyers, P. C., L. K. Shay, J. K. Brewster, and B. Jaimes (2016), Observed ocean thermal response to Hurricanes Gustav and Ike, *J. Geophys. Res. Oceans*, *121*, 162–179, doi:10.1002/2015JC010912.
- Oke, P. R., et al. (2015), Assessing the impact of observations on ocean forecasts and reanalyses: Part 2: Regional applications, *J. Oper. Oceanogr.*, *8*, s63–s79, doi:10.1080/1755876X.2015.1022080.
- Pascual, A., Y. Faugere, G. Larnicol, and P. Y. Le Traon (2006), Improved description of the ocean mesoscale variability by combining four satellite altimeters, *Geophys. Res. Lett.*, *33*, L02611, doi:10.1029/2005GL024633.
- Price, J. F. (2009), Metrics of hurricane-ocean interaction: Vertically integrated or vertically averaged ocean temperature, *Ocean Sci.*, *5*, 351–368.
- Pun, I. F., I. I. Lin, C. R. Wu, D. H. Ko, and W. T. Liu (2007), Validation and application of altimetry-derived upper ocean thermal structure in the western North Pacific Ocean for typhoon-intensity forecast, *IEEE Geosci. Remote Sens.*, *45*, 1616–1630.
- Sanabia, E. R., B. S. Barrett, P. G. Black, S. Chen, and J. A. Cummings (2013), Real-time upper-ocean temperature observations from aircraft during operational hurricane reconnaissance missions: AXBT demonstration project year one results, *Weather Forecasting*, *28*, 1404–1422.
- Schade, L. R., and K. A. Emanuel (1999), The ocean's effect on the intensity of tropical cyclones: Results from a simple coupled atmosphere-ocean model, *J. Atmos. Sci.*, *56*, 642–651.
- Shay, L. K., and J. K. Brewster (2010), Oceanic heat content variability in the eastern Pacific Ocean for hurricane intensity forecasting, *Mon. Weather Rev.*, *138*, 2110–2131.
- Shay, L. K., and E. Uhlhorn (2008), Loop Current response to hurricanes Isidore and Lili, *Mon. Weather Rev.*, *137*, 3248–3274.
- Shay, L. K., G. J. Goni, and P. G. Black (2000), Effects of a warm oceanic feature on Hurricane Opal, *Mon. Weather Rev.*, *128*, 1366–1383.
- Shay, L. K., et al. (2011), Airborne ocean surveys of the Loop Current complex from NOAA WP-3D in support of the Deep Water Horizon oil spill, in *Monitoring and Modeling the Deepwater Horizon Oil Spill: A Record-Breaking Enterprise*, *Geophys. Monogr. Ser.*, edited by Liu et al., pp. 131–151, AGU, Washington, D. C.
- Uhlhorn, E., and L. K. Shay (2012), Loop Current mixed layer response to hurricane Lili (2002) Part I: Observations, *J. Phys. Oceanogr.*, *42*, 400–419.
- Uhlhorn, E., and L. K. Shay (2013), Loop Current mixed layer response to hurricane Lili (2002) Part II: Modeling results, *J. Phys. Oceanogr.*, *43*, 1173–1192.
- Vianna, M. L., V. V. Menezes, A. B. Pezza, and I. Simmonds (2010), Interactions between Hurricane Catarina (2004) and warm core rings in the South Atlantic Ocean, *J. Geophys. Res.*, *115*, C07002, doi:10.1029/2009JC005974.
- Vincent, E. M., M. Lengaigne, J. Vialard, G. Madec, N. C. Jourdain, and S. Masson (2012), Assessing the oceanic control on the amplitude of sea surface cooling induced by tropical cyclones, *J. Geophys. Res.*, *117*, C05023, doi:10.1029/2011JC007705.
- Walker, N. D., R. R. Leben, and S. Balasubramanian (2005), Hurricane-forced upwelling and chlorophyll a enhancement within cold-core cyclones in the Gulf of Mexico, *Geophys. Res. Lett.*, *32*, L18610, doi:10.1029/2005GL023716.
- Walker, N. D., R. R. Leben, C. T. Pilley, M. Shannon, D. C. Herndon, I. F. Pun, I. I. Lin, and C. L. Gentemann (2014), Slow translation speed causes rapid collapse of northeast Pacific Hurricane Kenneth over cold core eddy, *Geophys. Res. Lett.*, *41*, 7595–7601, doi:10.1002/2014GL061584.
- Wu, C. C., C. Y. Lee, and I. I. Lin (2007), The effect of the ocean eddy on tropical cyclone intensity, *J. Atmos. Sci.*, *64*, 3562–3578.
- Zhu, T., and D.-L. Zhang (2006), The impact of the storm-induced SST cooling on hurricane intensity, *Adv. Atmos. Sci.*, *23*, 14–22.