Recipe for rigorous OSSE assessments – Illustration in the Gulf of Mexico

Matthieu Le Hénaff[1,2], George Halliwell[2], Ashwanth Srinivasan[3], Villy Kourafalou[4], Haoping Yang[1,2], Debra Willey[1,2], Robert Atlas[2]

(1) Univ.Miami-CIMAS, Miami, FL
(2) NOAA-AOML, Miami, FL
(3) Tendral LLC
(4) Univ.Miami-Rosenstiel, Miami, FL

Joint Workshop of the OceanPredict OS-Eval & CP-TT + SynObs Kick-off Workshop, Day 4

Motivation

- **OSSEs** long been implemented in **atmospheric sciences**
- More recent in oceanography
- Expected **increase in ocean OSSE** number (UN Decade of Ocean Science for Sustainable Development projects)
- In recent years, many ocean studies implemented the OSSE methodology, but did not always go through a **rigorous evaluation** of their OSSE system
- This presentation aims at **stressing the steps** for evaluating an OSSE system, which are necessary to further use the OSSE system to obtain **robust impact quantification** of future observing systems
- Criteria explained in **Halliwell et al. (2014) reference study**:

- Halliwell, G. R., A. Srinivasan, V. H. Kourafalou, H. Yang, D. Willey, M. Le Hénaff, and R. Atlas (2014). Rigorous evaluation of a fraternal twin ocean OSSE system for the Open Gulf of Mexico. *Journal of Atmospheric and Oceanic Technology*, 31(1), 105-130

Motivation

OSEs:

- Perform twin data-assimilative experiments:
  - One assimilates all observations
  - One denies only the observing system of interest
- Impact determined by **increased analysis and forecast errors**

OSSEs:

- **Same procedure as OSEs** except for assimilating **synthetic** observations simulated **from a Nature Run**
- Allows estimating the impact of:
  - New operational observing systems
  - Changing the deployment of existing systems
  - Different targeted observing strategies

OSSE challenge: Demonstrating the validity of the observing system impact assessments

Outline

Description of Gulf of Mexico OSSE system components:

- Nature Run
- Second, data-assimilative forecast ocean model
- *(Data-assimilation system)*
- *(Synthetic observation sampling toolbox)*

Evaluation of the OSSE system:

- Separate evaluation of the Nature Run and forecast model
- Evaluation of the errors between the forecast model and the Nature Run
- Perform OSE/OSSE comparisons:
  - 2010 analysis interval (Deepwater Horizon oil spill)
  - Identical, except OSEs assimilate actual obs., and OSSEs assimilate synthetic observations

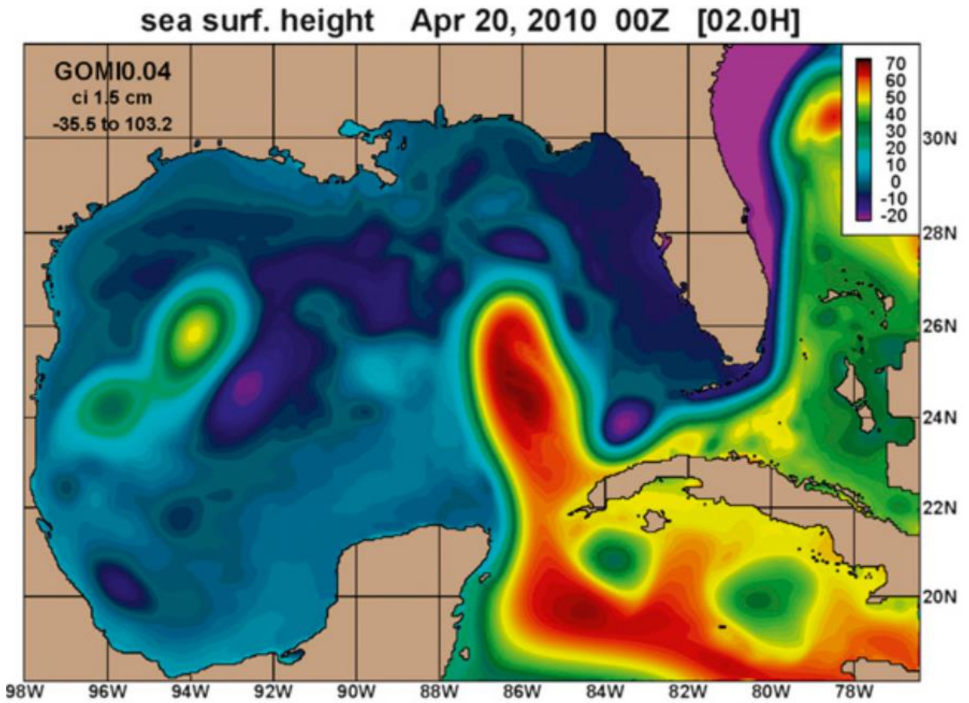
$\Rightarrow$ OSSEs must produce same results as corresponding OSEs

The Nature Run

The Nature Run must be an unconstrained model run that realistically reproduces the climatology and variability associated with ocean phenomena of interest

Gulf of Mexico Nature Run:

- HYCOM model at 0.04º (~4 km) resolution
- 32 vertical layers, sigma-z vertical coordinates
- Unconstrained run performed from 1/1/2004 through 12/31/2010
- Forced by 27 km COAMPS atmospheric model
- Boundary conditions from climatological HYCOM Atlantic simulation



Snapshot of SSH from the Nature Run

Caveat: ocean models are far from perfect

- Nature Run likely to be adequate in some respects, but inadequate in others
- Subsequent evaluation steps can either confirm the adequacy of the Nature Run or expose problems

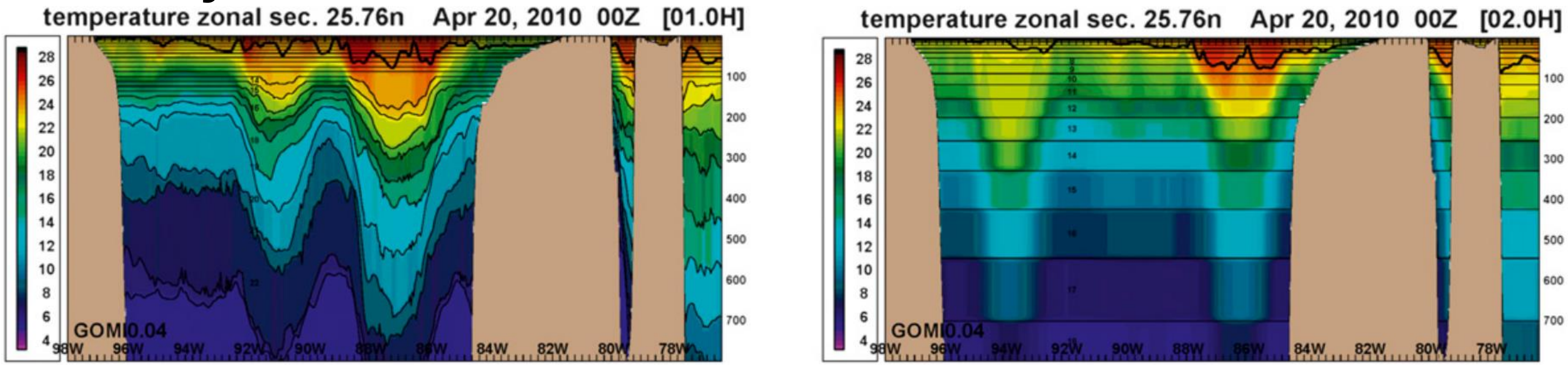
Second, forecast ocean model used in the DA System

Ideally, differences (errors) between the two models should grow to the same magnitude as, and have properties similar to, errors that presently exist between state-of-the-art general circulation models and the true ocean. This can be achieved by using a substantially different model than the Nature Run, with lower resolution to introduce additional truncation errors.

Choice of second, forecast model:

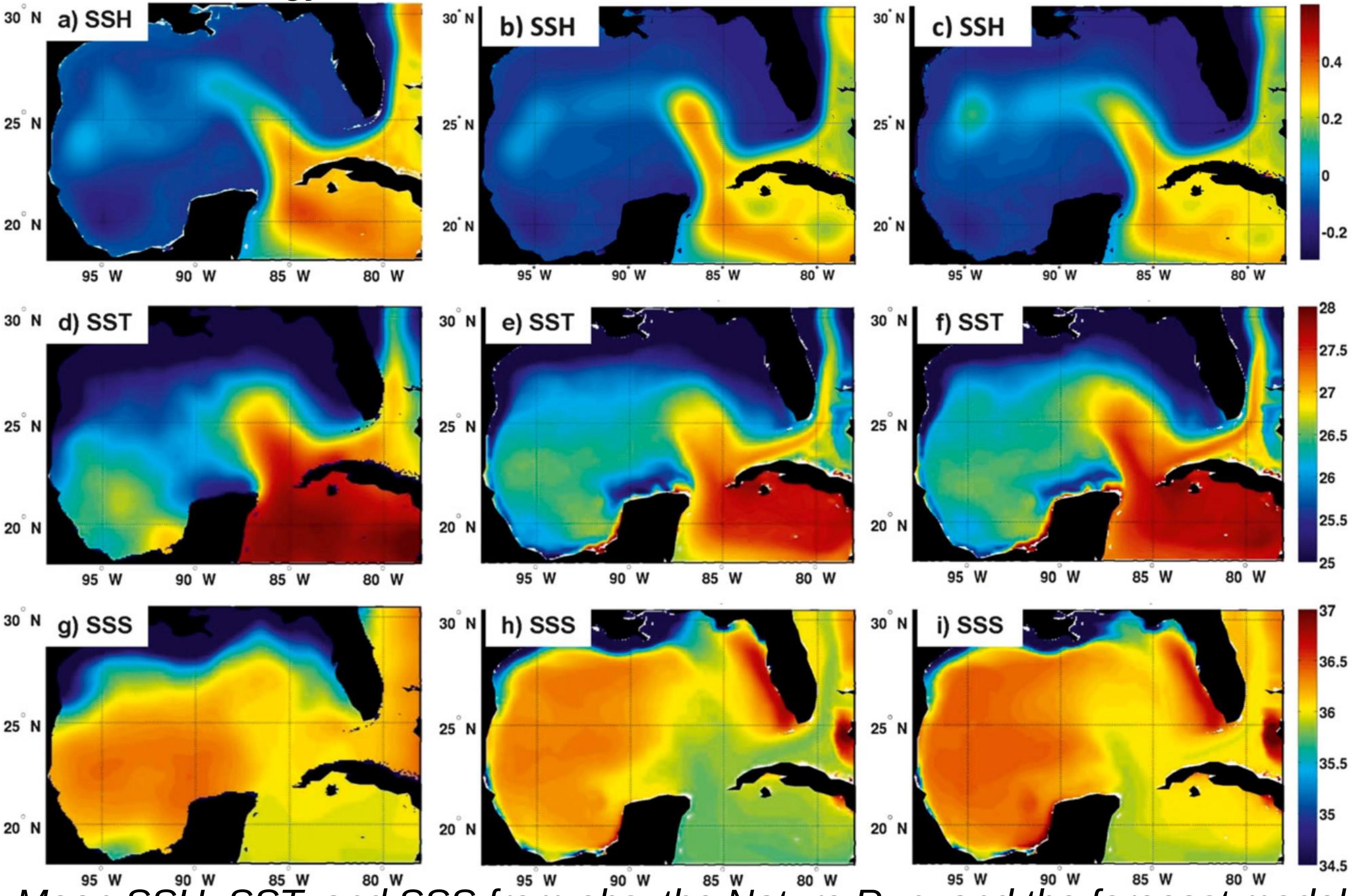
- HYCOM model at 0.08º (~8 km) res., 26 layers (different vertical coordinate system)
- Different vertical mixing scheme (KPP)
- Different diffusion, viscosity and friction parameters

$\Rightarrow$ **Fraternal twin system**



Evaluation of the Nature Run and the forecast model

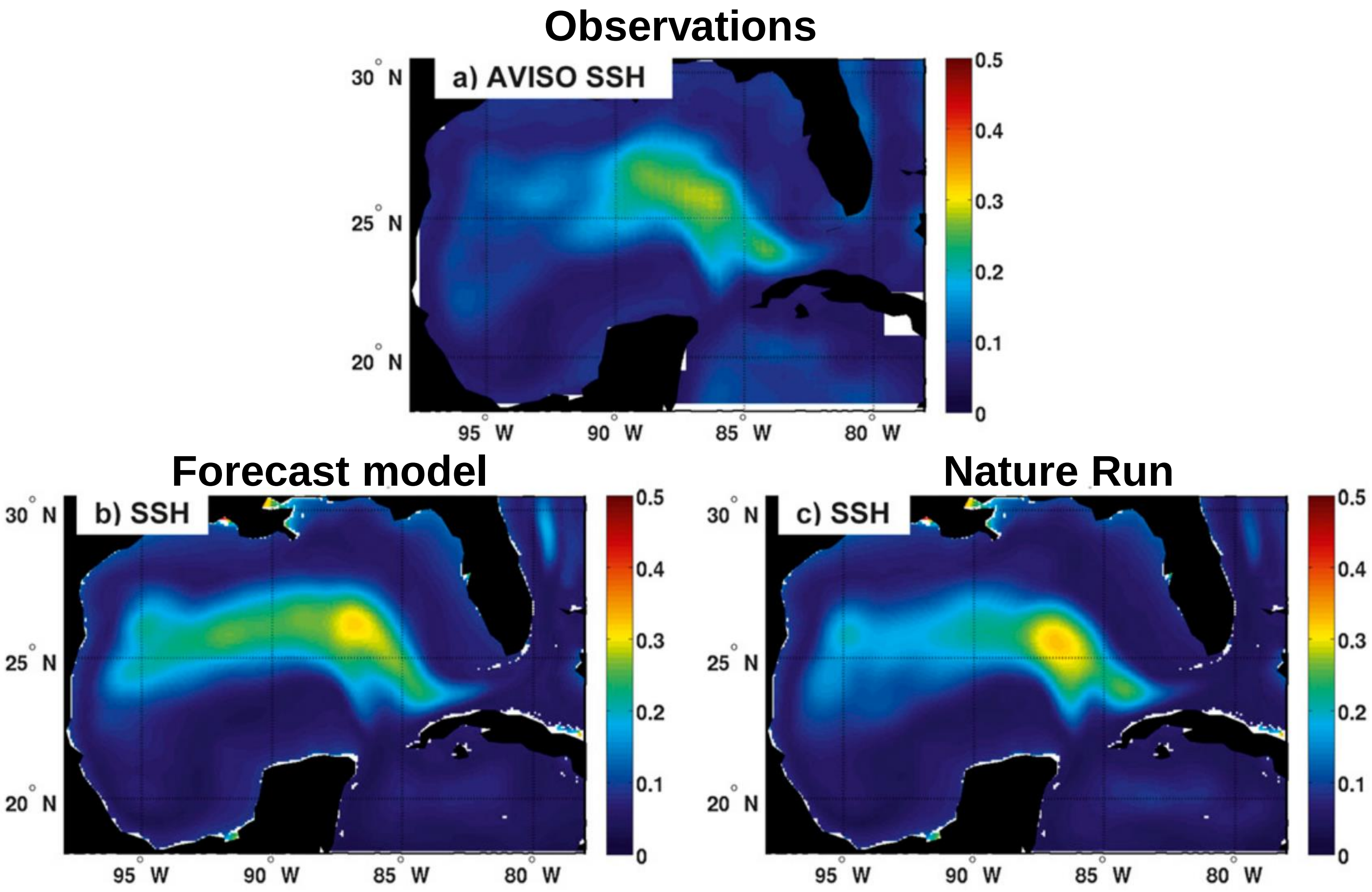
Climatology **Nature Run** **Forecast model**



Mean SSH, SST, and SSS from obs, the Nature Run, and the forecast model

- Mean **SSH realistic** in both the Nature Run and the forecast model
- Mean **SST** overall **realistic**
- Mean **signature** of the **Loop Current realistic** in both models
- Mean **SSS** overall **realistic** (except on West Florida Shelf)

Evaluation of the Nature Run and the forecast model



Standard deviation in SSH (m) in (a) AVISO observations, (b) in the forecast model, and (c) in the Nature Run

- **Largest variability** in the **Loop Current extension region** and eddy shedding region
- Ridge of **larger variability** extends westward along the **pathway of detached Loop Current rings**: narrower and extends somewhat farther to the west in the Nature Run and the forecast model
- Overall, **SSH variability comparable** between AVISO, the forecast model and the Nature Run

Evaluation of errors between both models

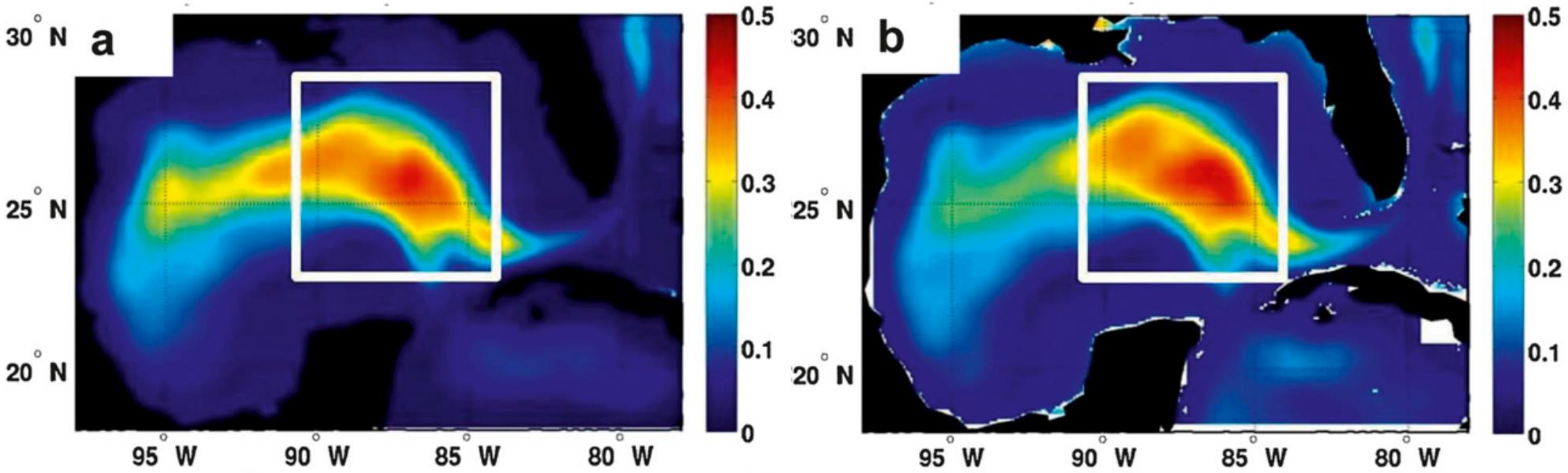
Given identical initialization, the ocean model used for the DA system must develop differences ("errors") with respect to the Nature Run model, as a result of different physical parameterizations, numerical schemes, and resolution. Errors between the models must be similar to the errors that exist between high-quality ocean models and the true ocean.

Key steps in evaluation of the second ocean model :

- Simulation over the same time interval as the Nature Run, with **identical initial conditions**:
    - **Differences** between the two models allow errors to develop in the second ocean model with respect to the Nature Run
    - After errors develop, **compare errors between the two models** to **errors between the Nature Run and the true ocean** (here: compared during 2005-2010)

Evaluation of errors between both models

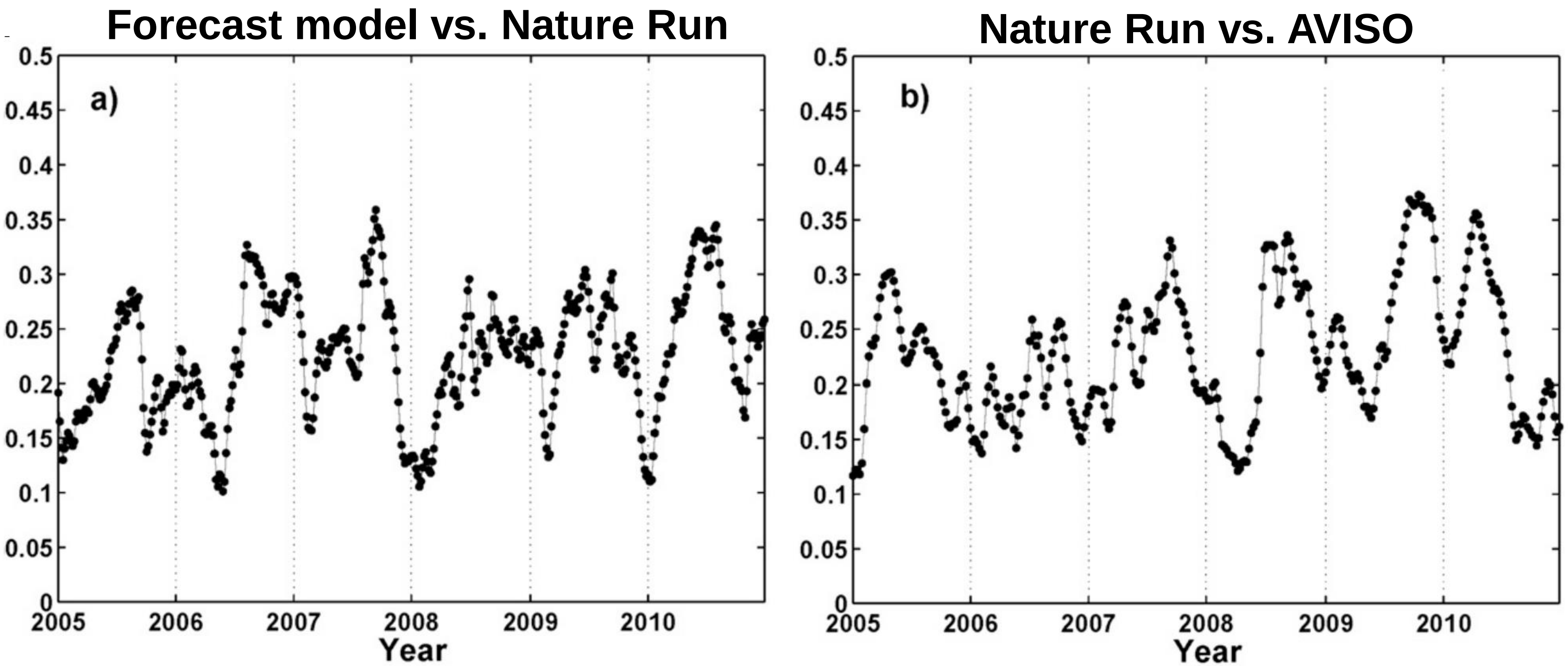
Forecast model vs. Nature Run **Nature Run vs. AVISO**



RMS difference of SSH (m) between (a) the Nature Run and the forecast model, and (b) the Nature Run and AVISO

- The **magnitude** and **distribution** of **RMS-differences between the Nature Run and the forecast model** are very **similar** to the magnitude and distribution of the **RMS-diff between the Nature Run and observed SSH** from AVISO altimetry maps

Evaluation of errors between both models



Time series of SSH RMS difference (m) over the Loop Current region between (a) the Nature Run vs. the forecast model, and (b) the Nature Run vs. AVISO

- The **time variations** of the **RMS-difference between the Nature Run and the forecast model** are very **similar**, in amplitude and frequencies, to those of **the RMS-diff between the Nature Run and AVISO**
- The **errors between the forecast model and the Nature Run** are overall **comparable to those between the Nature Run and the real ocean**

OSSE System Evaluation

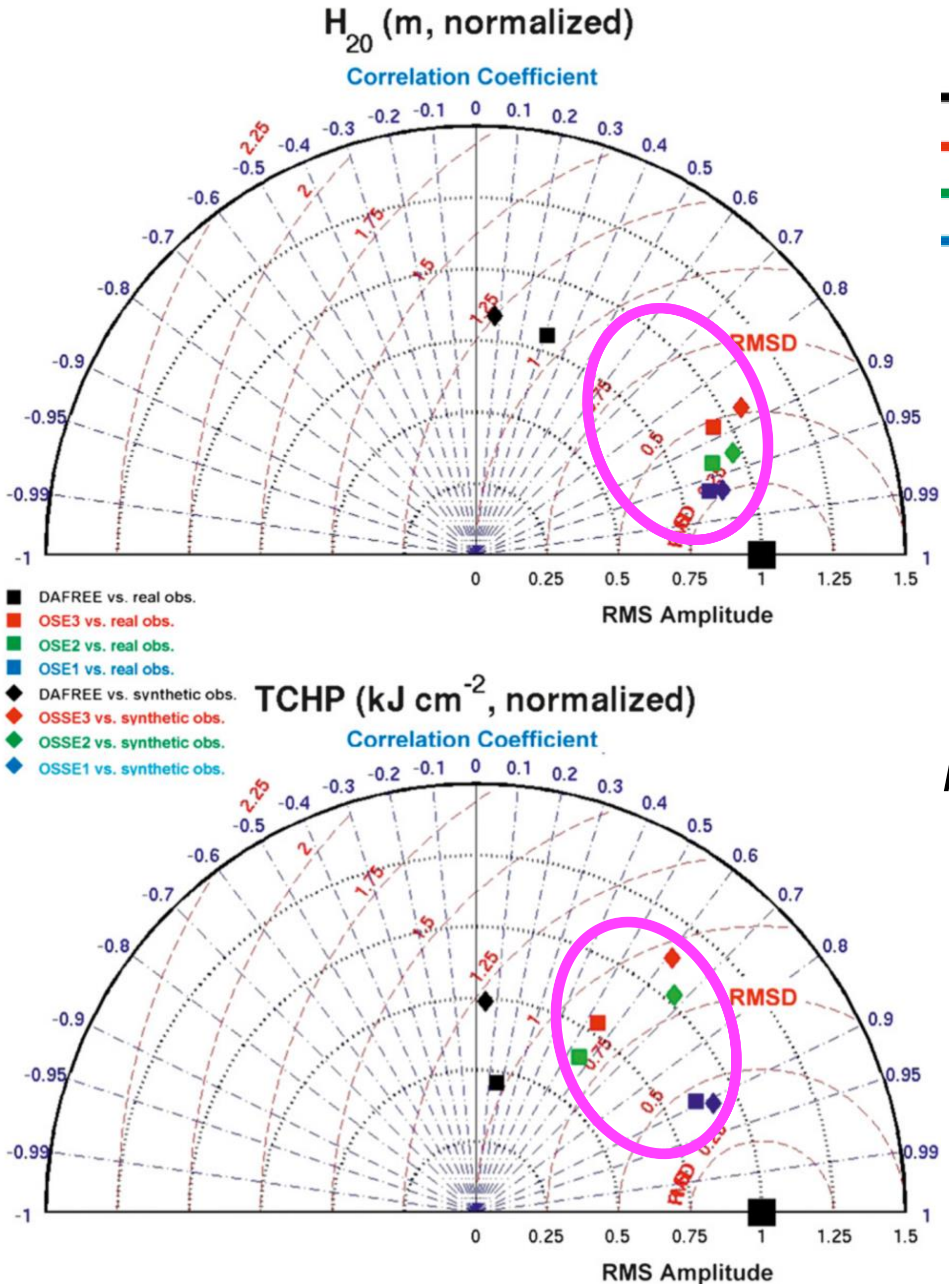
OSSE system errors and biases must be quantified by comparing OSSEs to ref OSEs

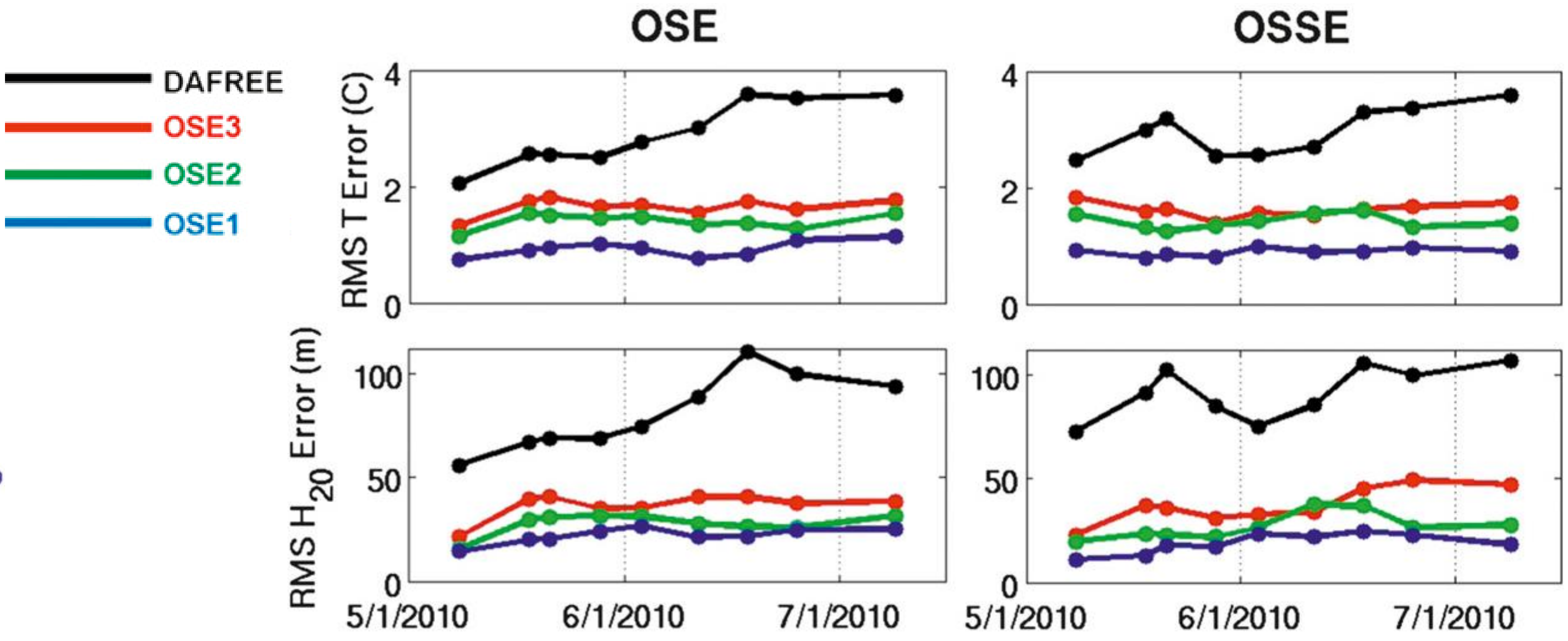
OSEs: 4 experiments using DA model with daily update cycle in 2010 (DWH oil spill)

- **OSE1** – assimilate all real observations:
  - Three altimeters (Jason1, Jason2, Envisat)
  - MCSST SST
  - In-situ SST (ship, surface buoy, surface drifter)
  - Ship XBT profiles
  - Airborne profiles (T from AXBT, T, S from AXCTD, T from AXCPs)
- **OSE2** – deny airborne profiles
- **OSE3** – also deny two of three altimeters
- **DAFREE** – Unconstrained simulation with DA model
- All initialized by unconstrained DA run on 1/1/2010

OSSEs: Experiments **OSSE1**, **OSSE2**, **OSSE3** identical to OSE1, OSE2, and OSE3, but assimilate synthetic instead of real observations

OSSE System Evaluation



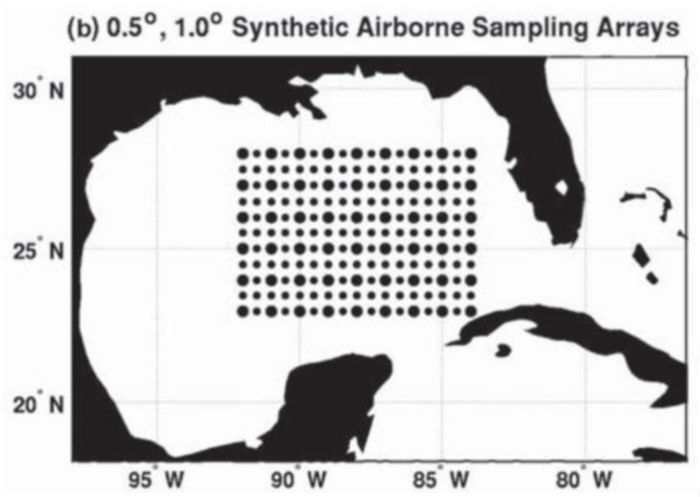


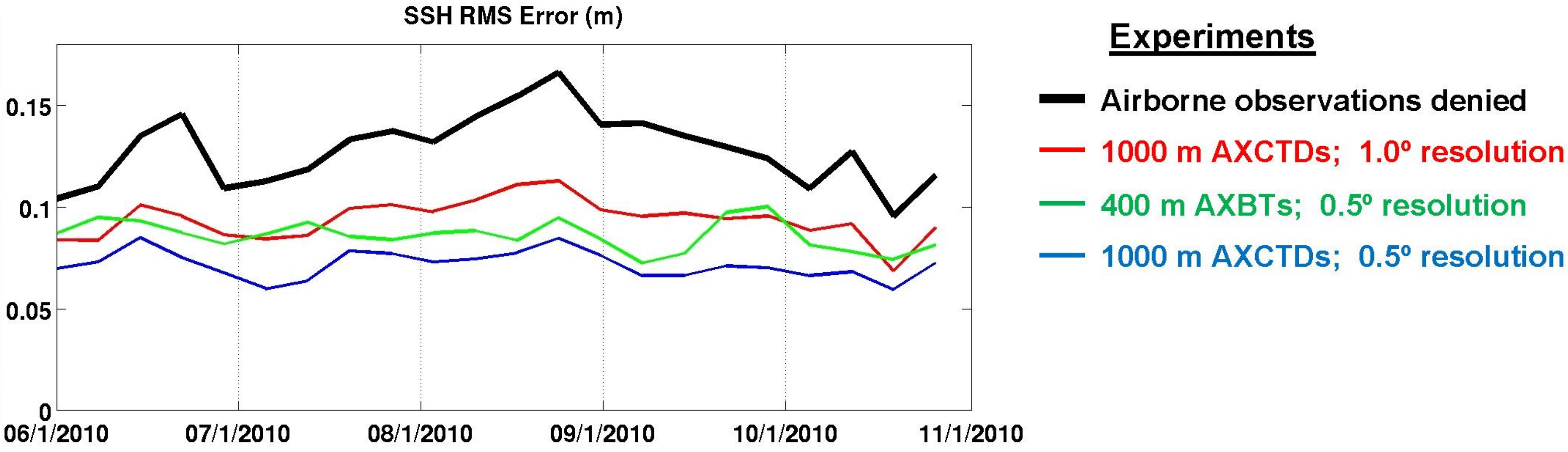
Evolution of RMS error between model and profiles for (top) T in the top 250m, and (bottom) depth of 20°C in 2010

- Denial of airborne obs. and further denial of two of the three altimeters produces essentially the **same impact assessments** between the **OSEs** and **OSSEs**

$\Rightarrow$ <u>Calibration is not required</u>

Example of application: impact of airborne profile resolution and type



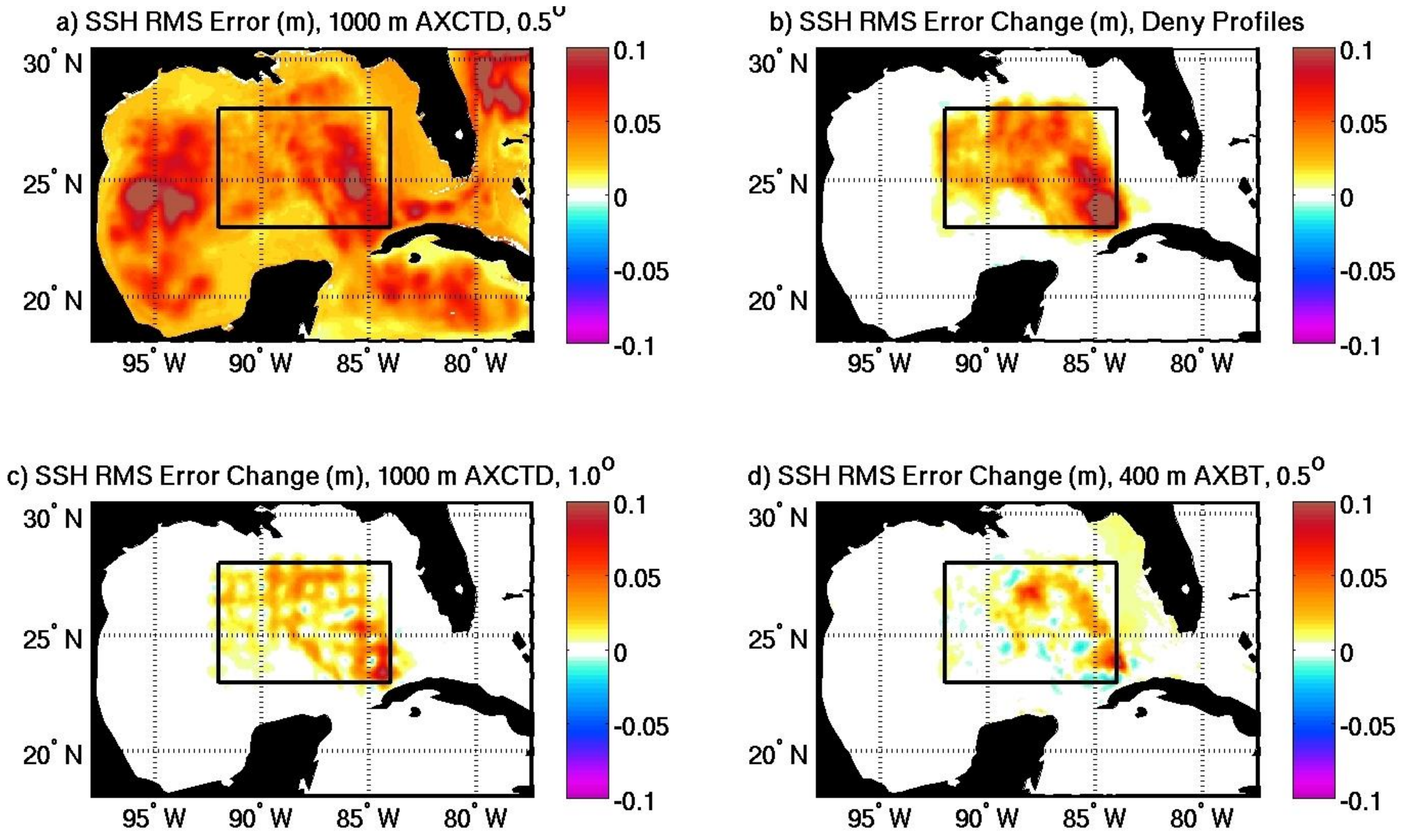


Time series of weekly averaged RMS error in SSH between the Nature Run and 7-day OSSE analysis products (4-day analysis window)

- Impact assessments:

1. **Larger RMS error** when all airborne **profiles denied**
2. **Smallest RMS error** for experiment assimilating **1000 m AXCTDs** at **0.5º resolution:** 46% error reduction
3. RMS **error increases** when 1000 m AXCTDs are assimilated at **1º res**.
4. RMS **error increases** when profiles still have 0.5º resolution, but extend to **400 m** compared to 1000 m, and measure **temperature only** (not salinity) – i.e. AXBTs

Example of application: impact of airborne profile resolution and type



- Largest error increase is the area of airborne profiles deployment
- Error increase in case of **reduced resolution** shows **1º grid pattern**
- Error increase in case of **shallower extension** follows the **edge of the Loop Current**, associated with **frontal dynamics**

Key findings

- The **evaluation of an OSSE system** necessitates:
  - **Separate evaluation** of **each model** (Nature Run and forecast model)
  - Evaluation of the **errors** growing **in the forecast** model because of model differences: they have to be **comparable to errors between state-of-the-art models and the real ocean**
  - Evaluation of the **diagnostics from the OSSE system**: they have to be **similar to the diagnostics from equivalent OSEs** using real observations

  ⇒ **Only after these criteria have been validated** can the **OSSE system be used** to **quantify** the performance of alternative or new observing systems
- Example with airborne surveys:
  - With sufficient spatial sampling, can achieve up to 50% additional RMS error reduction
  - Limitations: horizontal resolution must be substantially smaller than 1°

Key findings

Results can be found:

- **Halliwell, G. R., A. Srinivasan, V. H. Kourafalou, H. Yang, D. Willey, M. Le Hénaff, and R. Atlas (2014). Rigorous evaluation of a fraternal twin ocean OSSE system for the Open Gulf of Mexico. *J. Atmos. Ocean. Technol.*, 31(1), 105-130**

- Halliwell, G. R., V. H. Kourafalou, M. Le Hénaff, L. K. Shay, and R. Atlas (2015). OSSE Impact Analysis of Airborne Ocean Surveys for Improving Upper-Ocean Dynamical and Thermodynamical Forecasts in the Gulf of Mexico. *Prog. Oceanogr.*, 130, 32-46

- Halliwell, G. R., 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 Tropical Cyclone prediction in coupled systems. *J. Oper. Oceanogr.*, 1-22

- Halliwell Jr, 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.*, *122*(7), pp.5729-5748

- Halliwell, G.R., G.J. Goni, M.F. Mehari, V.H. Kourafalou, M. Baringer, and R. Atlas (2020). OSSE assessment of underwater glider arrays to improve ocean model initialization for tropical cyclone prediction. *J. Atmos. Ocean. Technol.*, 37(3), pp.467-487