Forecast error estimation in the NRL relocatable ocean ensemble forecast system

Clark Rowley[1], Mozheng Wei[1], Emanuel Coelho[2], and Peter Spence[3]

[1]Naval Research Laboratory Oceanography Division [2]University of New Orleans [3]QinetiQ North America

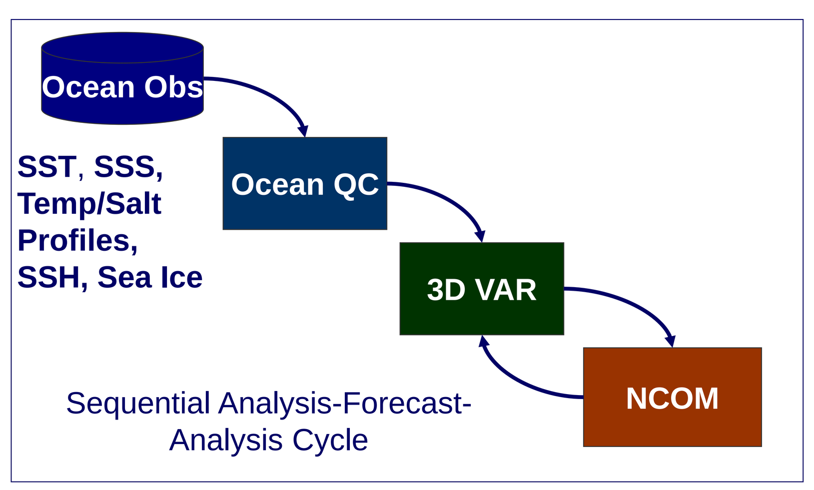
Introduction

A globally relocatable regional ocean ensemble forecast system has been developed. The representation of uncertainty in the ensemble and the growth of uncertainty through the ensemble forecast are controlled by the analysis error estimated by the analysis. Analysis error estimated through model temporal variability or through forecast error via the analysis increment may over- or under-estimate the forecast error due to the limited availability of subsurface observations. We propose a method to estimate analysis error that includes model variability, recent model forecast error, and historical model or climate uncertainty with specified error growth rates, and accounts for the subsurface sampling. By including the growth back to historical error estimates in the long periods between subsurface sampling, we improve the ensemble spread as measured using both verifying analyses and *in situ* and remote sensing observations.

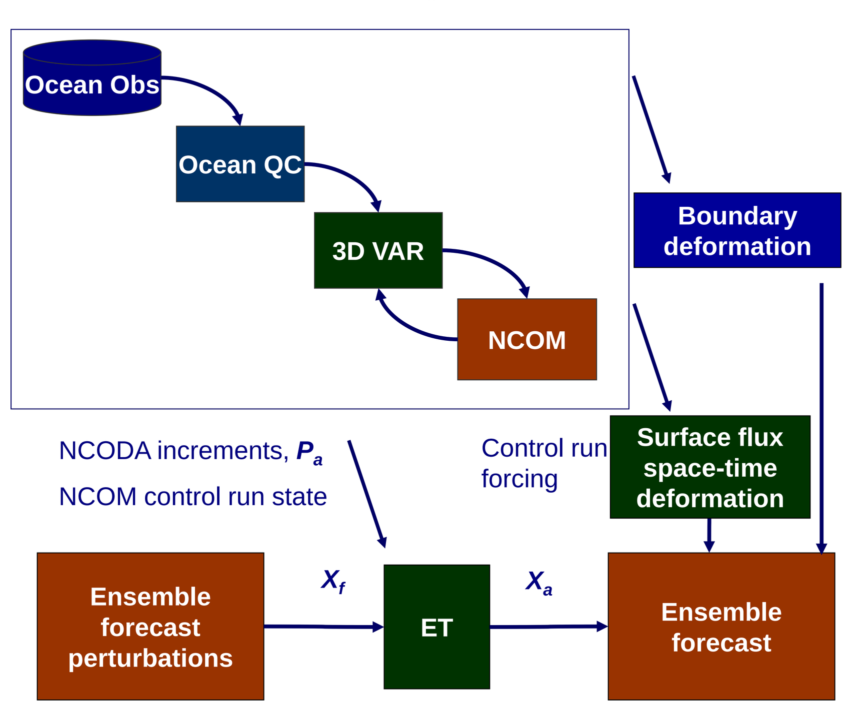
The NRL relocatable ocean ensemble forecast system

NCODA - NRL Coupled Ocean Data Assimilation
Cummings, QJRMS, 2005
NCOM - Navy Coastal Ocean Model
Barron, et al., Ocean Modelling, 2006
ET - Ensemble Transform
Bishop and Toth, JAS, 1999

Deterministic forecast cycle



Ensemble forecast cycle



Forecast error in the Ensemble Transform

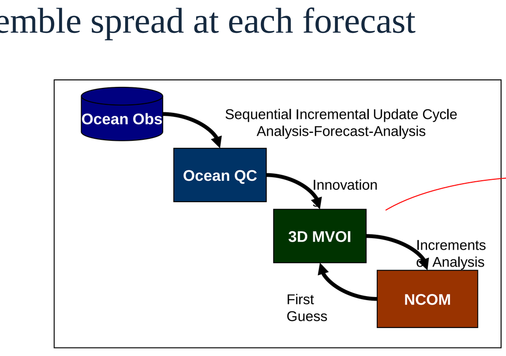
The ensemble transform rescales forecast perturbations X_f to analysis perturbations X_a with a transformation T such that the global perturbation variance normalized by the forecast error variance P_f is 1. In this way the forecast error estimate from the analysis sets the initial ensemble spread at each forecast cycle.

Posterior / Likelihood Function / Prior

$$p(x|y^a) = \frac{p(y^a|x)\cdot p(x)}{p(y^a)}$$

$$X'_a = \left[T \cdot X'_f \right]^T$$

Posterior Ensemble pert. (at analysis time); Prior Ensemble pert. (previous forecast); Perturbations Likelihood/Transform Function



$$(X_a' T)^T B^{-1} X_a' T = N_{ens} \cdot I$$

Forecast error

Approach:
- Estimate forecast error using the recent history of analysis increments

Advantages:
- Analysis corrections to the model reflect the model forecast error where there are observations

Limitations:
- The error estimate is subject to the sparse subsurface observational sampling in the ocean forecast system
- Does not reflect model temporal variability

Time weighting: $w_n = (1-\gamma)^{(n-1)}$

Analysis increments: X'_n

Error variance estimate:

$$e^2 = \frac{\sum_{n=1}^{N} (1-\gamma)^{(n-1)} X'^2_n}{\sum_{n=1}^{N} (1-\gamma)^{(n-1)}}$$

Climate error

Add in climate variance ε^2_c to account for long times between sparse observations. The climatological error can be a long-term model error or (more typically for the ocean) a simple climatological variability. Climate variance may overestimate the forecast error near observations.

Analysis error reduction

The posterior analysis error variance P_a is estimated in NCODA 3DVar as a calculation of the scalar error reduction factor β. This reflects the influence of observations

Approach

Calculate a combined estimate of forecast error variance P_f that accounts for the observing system

- Calculate analysis error reduction factor β and estimate analysis error ε as an initial estimate from the prior background error $\varepsilon_b = P_f^{1/2}$ as $\varepsilon = \beta * \varepsilon_b$
- Calculate forecast temporal variability ε_m. Where ε_m is larger than ε, relax toward ε_m with timescale factor N^τ_m (τ ~ 2d) as
  $w_m = 1/N^\tau_m$
  $\varepsilon = \varepsilon + w_m(\varepsilon_m - \varepsilon)$
- Calculate forecast error from the analysis increment timeseries ε_a. Where ε_a is larger than ε, relax toward ε_a with timescale factor N^τ_a (τ ~ 10d) as
  $w_a = 1/N^\tau_a$
  $\varepsilon = \varepsilon + w_a(\varepsilon_a - \varepsilon)$
- Load climate error ε_c. Where ε_c is larger than ε, relax toward ε_c with timescale factor N^τ_c (τ ~ 30d) as
  $w_c = 1/N^\tau_c$
  $\varepsilon = \varepsilon + w_c(\varepsilon_c - \varepsilon)$

Temporal variability

Approach:
- Calculate model temporal variability using a weighted sum of squares of the recent successive differences in the model forecast state valid at the analysis time

Advantages:
- Model temporal variability may be a useful proxy for forecast error where observations are limited.

Limitations:
- Temporal variability may underestimate true forecast error
- The estimate is sensitive to the observational sampling in the cycling system

Time weighting:

$w_n = (1-\gamma)^{(n-1)}$

Timeseries of differences between successive forecasts:

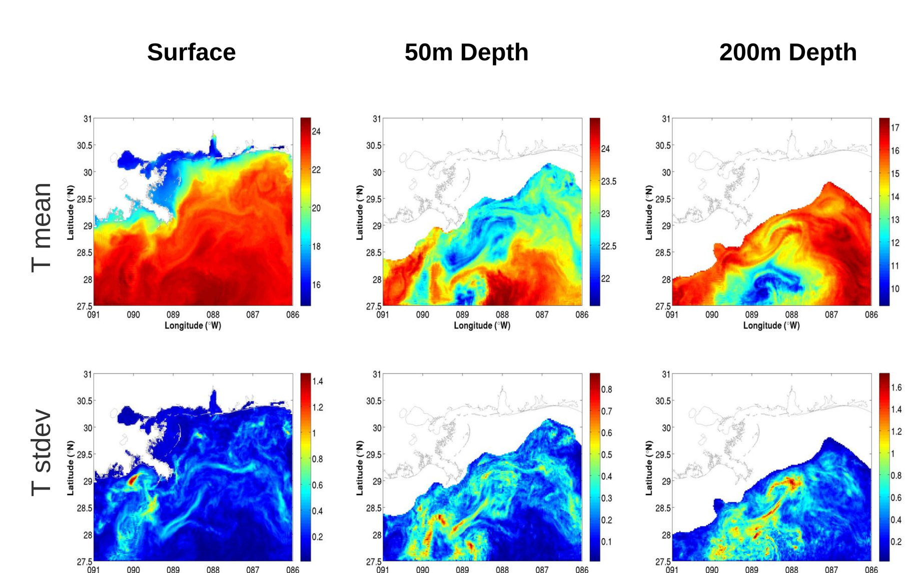
$X'_n = X_n - X_{n+1}$

Error variance estimate:

$$e^2 = \frac{\sum_{n=0}^{N-1} (1-\gamma)^n X'^2_n}{\sum_{n=0}^{N-1} (1-\gamma)^n}$$

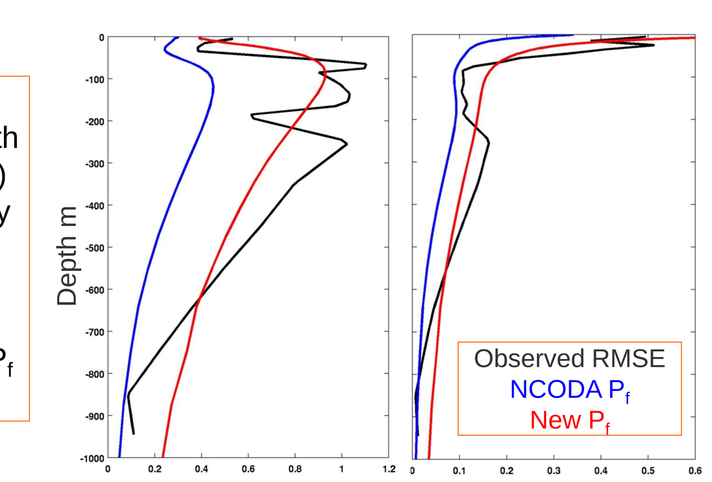
Results

A twin experiment with the new forecast error estimate was configured for a small 12-member ensemble forecast series for the northern Gulf of Mexico. The reference ensemble used the standard NCODA error estimate based on forecast variability. The modified ensemble used the complete P_f estimate above.

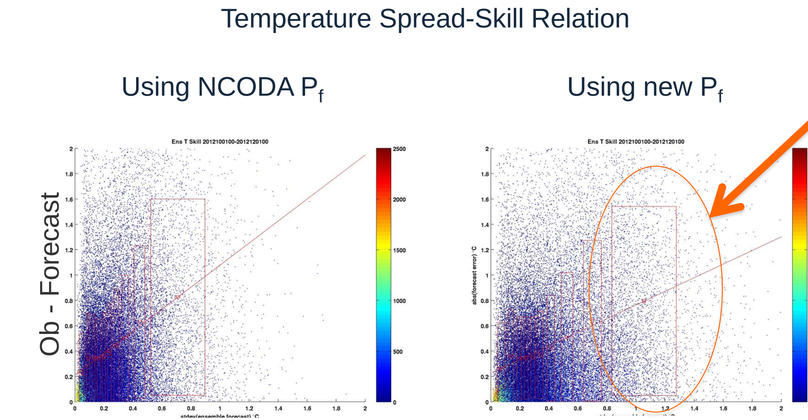


The ensemble temperature mean and spread (standard deviation) from a 12-member GOM ensemble forecast at 00Z 01 Dec 2012. The ensemble was initialized on 02 September 2012 and is run with perturbed surface forcing and perturbed initial conditions using the ensemble transform.

The RMS observed forecast error with depth (01 Oct – 01 Dec 2012) is better represented by the area-average complete P_f estimate (red) compared to the reference experiment P_f (blue) on 01 Dec 2012.

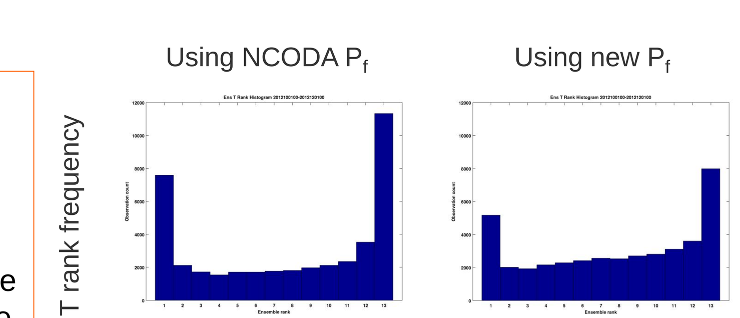


Temperature Spread-Skill Relation



The temperature spread-skill plots for the twin experiment show the improved spread in the ensemble with the complete P_f.

The temperature rank histogram plots for the twin experiment show the ensemble with the complete P_f estimate better captures extreme values compared to the reference ensemble.



Conclusions

The representation of uncertainty in the ensemble analysis and the growth of uncertainty through the ensemble forecast are controlled by the analysis error estimated by the NCODA analysis. We propose a method to estimate analysis error that includes model variability, recent model forecast error, and historical model or climate uncertainty with specified error growth rates, and accounts for the subsurface sampling. By including the growth back to historical error estimates in the long periods between subsurface sampling, we improve the ensemble spread as measured using both verifying analyses and observations. Future work on model uncertainty will address under-prediction of growth in the ensemble spread.

Bibliography

Barron, C. N. , A. B. Kara, P. J. Martin, R. C. Rhodes and L. F. Smedstad (2006): Formulation, implementation and examination of vertical coordinate choices in the Global Navy Coastal Ocean Model (NCOM). *Ocean Modelling*, **11**, 347-375.

Bishop, C. H. and Z. Toth (1999): Ensemble transformation and adaptive observations. *J. Atmos. Sci.*, **56**, 1748-1765.

Coelho, E.F., J.P. Fabre, C. Rowley, G. Jacobs, C. Bishop, J. Cummings, and X. Hong (2009): Targeting Observations to Reduce Acoustic Prediction Uncertainty. *U.S. Navy Journal of Underwater Acoustics*, **60(1)**, 125-152.

Cummings, J. (2005): Operational multivariate ocean data assimilation. *Q. J. R. Meteorol. Soc.*, **131**, 3583-3604.

Rowley, C., P. J. Martin and J.A. Cummings (2009): The Naval Research Laboratory Relocatable Ocean Nowcast/Forecast System. *U.S. Navy Journal of Underwater Acoustics*, **60(1)**, 169-202.

Acknowledgments

This work is supported by the Office of Naval Research.

For further information

Please contact *clark.rowley@nrlssc.navy.mil*. More information on related projects can be obtained at *www7320.nrlssc.navy.mil*.