



## **Towards realistic data assimilation for the coastal ocean: Observing network sensitivity experiments in New York Harbor**

**R. Hoffman** (1), R. Ponte (1), E. Kostelich (2), A. Blumberg (3), I. Szunyogh (4), and S. Vinogradov (1)

(1) Atmospheric and Environmental Research, Inc., Lexington, MA; (2) (Dept. of Mathematics and Statistics, Arizona State University, Tempe AZ; (3) Stevens Institute of Technology, Hoboken, NJ; and (4) (University of Maryland, College Park, MD).

A coastal ocean data assimilation system is being developed. The goal is to combine large and disparate datasets with ocean numerical models, producing accurate analyses, forecasts, and respective uncertainty estimates for any littoral region. A modular interface combines the Estuarine and Coastal Ocean Model (ECOM) and the Local Ensemble Transform Kalman Filter (LETKF) into a highly scalable, portable and efficient ocean data assimilation system. LETKF, a recent adaptation of ensemble Kalman filtering techniques, works particularly well for very large non-linear dynamical systems in both sparse and dense data regimes, and provides efficient algorithms for error estimation and quality control. In simulation experiments reported elsewhere for highly idealized data distributions in the the New York Harbor Observing and Prediction System (NYHOPS) the filter quickly converges, eliminating bias and greatly reducing rms errors. This behavior is robust to changes in ensemble size, data coverage, and data error.

As our near-term goals include applying the system to real data in operational environments and evaluating the skill of analyses and forecasts under different flow regimes and boundary conditions, with diverse data streams, and in various model configurations, here we study the behavior of the system with more realistic observing networks—in particular observing networks where most of the observation are at or

near the surface. Preliminary results show the data assimilation is more sensitive to the specification of the localization parameters in this situation. In particular, it appears that we must extend the influence of surface data through the water column. Also a fixed observing network does not perform as well as a randomized observing network. One reason is that some regions with no observations are not affected by the flow of information from other regions with data. An example of this occurs in the upper reaches of some of the rivers in the domain that have no upstream data.