Group #1

ITR/AP: An Ensemble Approach to Data Assimilation in the Earth Sciences

CoPrincipal Investigator(s): Alan S. Willsky; Paola Malanotte-Rizzoli; Dara Entekhabi; Kerry A Emauel
Senior Personnel(s): James Hansen; Carl Wunch; Jacob White
Graduate Student(s): Yuhua Zhou; Greg Lawson; Susan Dunne; Virat Chatdarong; Vikram Khade
Post-doc(s): Wes Dripps; Xiaoyung Zang; Steven Margulis; Patrick Heimbach; Sai Ravela
Graduate Students(s): Ahanin Adel
Post-doc(s): Sang Jin Lyu
Graduate Student(s): Gene Ng; Yousef Jafapour; Piyatida Hoisungwan

Other Collaborators: Dr. Eni Njoku, Jet Propulsion Laboratory, Pasadena, CA
Prof. Arnold Heeminck, Technical University of Delft, Delft, Netherlands
Prof. Fabio Castelli, University of Florence, Florence, Italy
Prof. Eric Wood, Princeton University, Princeton NJ
Dr. Wade Crow, Agricultural Research Service, USDA, Beltsville, MD
Dr. Rolf Reichle, NASA Goddard Space Flight Center, Greenbelt, MD
Prof. Steven Margulis, UCLA, Los Angeles, CA.

Project Summary

New data sources are beginning to have a dramatic impact on our ability to understand the earth as an integrated system. Our prospects for dealing with the environmental issues of the 21st century -- climate change, population pressures on natural resources, and major modifications in global element cycles -- depend largely on this new information. However, our ability to process and interpret environmental data is not keeping pace with the dramatic increase in available information, especially information from airborne and orbital remote sensing platforms. If we are to realize the potential benefits of new sensing technologies we will need to develop intelligent environmental data assimilation procedures that are able to efficiently extract useful information about the earth from a diverse set of data sources. This is the overall objective of our ITR project.

Environmental data assimilation can be posed as a problem of estimating a large number of unobservable or highly uncertain variables (e.g. sea surface heights, atmospheric pressures hydrologic fluxes, etc.) from a large number of related but noisy measurements (e.g. microwave radiances or backscatter detected by a satellite sensor). The estimation procedure relies on mathematical models that relate unknowns to measurements. Environmental estimation problems are challenging because the systems of interest: 1) are often nonlinear, even chaotic, 2) are spatially distributed and highly variable over a wide range of space and time scales, leading to high dimensionality simulation and estimation problems, 3) are difficult to describe with precision, making model uncertainty particularly important.

In this project we are particularly concerned with very large nonlinear problems which are not amenable to traditional data assimilation techniques but are of crucial interest to researchers in the earth sciences. The most successful and operationally promising solutions reduce problem size by focusing on key observable spatial and/or dynamical features. So-called ensemble methods provide a particularly informative way to identify these key features. These methods are based on an ensemble of model simulations that sample the range of possible system states. Our experience suggests that informed

ensemble sampling may be an effective way to obtain efficient low dimensionality approximations to large spatially distributed systems. They are also efficient, flexible, robust, and relatively easy to implement.

The project will include two related initiatives. In the first initiative, we will work as an interdisciplinary team to develop a better understanding of the issues of nonlinearity, high dimensionality and model uncertainty that are crucial to large-scale data assimilation. This work will focus on the ensemble approach. In the second initiative, we will develop a new generation of "intelligent" data assimilation methods that build on the understanding gained in the first initiative. We will investigate the applicability of these methods to problems of broad interest in the earth sciences, including problems that 1) deal with coupled systems, 2) cut across traditional disciplines, and 3) work with remote sensing data sets.