Geosci. Model Dev. Discuss., 8, C2281–C2283, 2015
www.geosci-model-dev-discuss.net/8/C2281/2015/

Geoscientific Model Development Discussions

Interactive comment on "DasPy 1.0 – the Open Source Multivariate Land Data Assimilation Framework in combination with the Community Land Model 4.5" *by* X. Han et al.

Anonymous Referee #1

Received and published: 30 September 2015

General Comments

This paper has developed DasPy, an open source data assimilation framework based on the Ensemble Kalman filter method by using Python scripting language. The authors have demonstrated their framework for the Community Land model. This is an important work and quite a tremendous undertaking to provide this capability to the data assimilation community. However, the paper is limited in two important fronts: scientific contribution, and computational capability. As it stands, the DasPy program is not a framework but single data assimilation method applied to one land surface model and a standardized observation model, CMEM.

Scientific contribution:

a) In the data assimilation (DA) community, the assimilation methods are many and varied. DA methods including Ensemble Kalman fitler (EnKF), particle filter, evolutionary data assimilation, etc. Each method is designed to account for model and observation uncertainties in different ways. Even within each approach there are shades of differences. So what the DA community needs from a program like this is an identification of the common DA modules, for example, ensemble generation methods, error estimation methods, etc. This DasPy program does not have this feature, and does not qualify as a framework. DasPy has hard coded modules for only a single DA method, the Local Ensemble Transform Kalman Filter.

b) The DA community uses several prediction models: different land surface models, different hydrological models, numerous environmental models and weather prediction models. Again, there are commonalities from all these models from a data assimilation viewpoint. These models have state variables, model parameters, forcing variables, etc. A DA framework about how to access and treat these components from any prediction model is needed. The DasPy program does not the capability and the flexibility to treat these model components.

c) Error estimation. There are several approaches to error estimation in the DA community. Numerous error estimation methods for prediction model, observation, state variables, forcing variables, model parameters, etc. At least an identification of the commonly used ones from the literature is needed from a paper like this.

d) Update procedure. Again, commonly used updating strategies from the literature need to be identified and be included into a program like DasPy.

e) One sound contribution the authors' made in this paper is the inclusion of the CMEM model. The CMEM is a standardized observation model in the land surface DA community. Yet, its inclusion into DasPy is not standardized from a DA standpoint. Different DA strategies for handling observation model, and its assimilation.

Computational capability:

a) The authors' choice of scientific language of Python is very sound and their use of modules such as numpy and matplotlib is very good. What make modules like numpy and matplotlib so popular is their modularity, they provide tools independent of the problem at hand, allowing flexibility for users to adapt these modules to different problems. Yet, this DasPy lacks modularity and the flexibility to adapt to different assimilation problems.

b) The DasPy is very limited to only one DA method, the Local Ensemble Transform Kalman Filter. Even the modules developed for this one DA method are not transferrable to other shades of EnKF. No inclusion of different ensemble generation approaches, different error estimation methodologies, etc. The modules need be independent so that they can be plug-and-played into another approach.

c) Computationally, what the DA community needs from a program like DasPy is a standardized collection of modules so that users can pick and choose specific tools they need to address their problem. At this stage, the DasPy program is simply a specialized application of one DA method for a specific problem. DasPy is hardly extensible to other problems, e.g. streamflow, weather prediction, etc. Even with soil moisture, there are still lots to be desired. Assimilation with calibrated ensemble of model parameter values, for example.

Interactive comment on Geosci. Model Dev. Discuss., 8, 7395, 2015.