Verificaiton-Based Model Tuning – Part II

Caren Marzban
Applied Physics Laboratory, University of Washington
1013 N.E. 40th St.
Seattle, WA 98105-6698
phone: (206) 221-4361 fax: (206) 545-4404 email: marzban@stat.washington.edu

Scott A. Sandgathe
35574 Camp Creek Rd.
Springfield, OR 97478
phone: (541) 988-0289 fax: (541) 988-0289 email: sandgathe@apl.washington.edu

David W. Jones
Applied Physics Laboratory, University of Washington
1013 N.E. 40th St.
Seattle, WA 98105-6698
phone: (206) 543-3236 fax: (206) 545-4404 email: dwjones@apl.washington.edu

Award Number: N00014-12-G-0078/0029

LONG-TERM GOALS

In our ongoing work we have been developing frameworks in which the relationship between model parameters and forecast quantities can be quantified. An example of the former is the fraction of available precipitation in Kain-Fritch fed back to the grid scale in an NWP model, and an example of the latter is the mean (across spatial domain) of 24-hr accumulated precipitation. The result can be used for two purposes: knowledge of the “forward relationship” can be used to determine the effect, on all of the forecast quantities, of changes in any given model parameter; by contrast, knowledge of the “inverse relationship” can aid in setting the values of the model parameters for the purpose of implementing a certain desirable effect in any given forecast quantity. As such, both the forward and the inverse relations can aid in "tuning" the NWP model.

OBJECTIVES

The main objective is to develop a framework which informs a forecaster or model developer how the various model parameters effect the forecast quantities (forward relation), and vice versa (inverse relation), by inferring mathematical/statistical relationships between them. We will examine features of the forecasts which spatial in nature (e.g., the number and size of clusters in a precipitation field), as well as non-spatial features (e.g., mean precipitation across a spatial domain). Additionally, the developed framework will account for interactions between the model parameters, as well as between forecast quantities.
APPROACH

The framework we have been developing is based on two statistical methods: a) variance-based Sensitivity Analysis (SA), and b) Canonical Correlation Analysis (CCA). Although each of these two methods independently addresses the main task of this proposal, they also work in conjunction; for example, the “inputs” of the CCA have been selected through the SA approach.

It is important to point out that all of the above approaches require only model forecasts, and no analysis or observations. This is an important feature of the method because model forecasts are generally and significantly easier to generate than observations. Said differently, this project does not deal with verification (see the Related Projects section).

WORK COMPLETED

The SA has now been applied to the Lorenz ‘63 model, and the results have been compared to an adjoint-based approach to sensitivity analysis (Marzban 2013). As of the last report, two papers in which SA and CCA are applied to COAMPS where in the review process; they are now both published (Marzban, Sandgathe, Doyle, Lederer 2013; Marzban, Sandgathe Doyle 2014).

RESULTS

The project has not yet fully started, because the PI had committed to teaching responsibilities prior to the funding of the project. Work will resume in December, 2014. Meanwhile, two graduate students have been hired: 1) Ning Li is a graduate student in the department of Statistics, in second year of studies for a Masters degree. She has expertise in sensitivity analysis and cluster analysis, which are the main tools involved in this work. She is supported at 50% by this project. 2) Natalia Hryniw, a graduate student in the department of Atmospheric Sciences, received a bachelor's in physics from the University of Chicago, and spent a year working at the Los Alamos in the Computational Physics Division and the Earth and Environmental Sciences Division. She has experience running the Weather Research & Forecast (WRF) modeling system, along with experience in statistical and computational physics. She, too, is supported at 50% by this project.

We will begin the project by addressing the following specific tasks:

Our past work has revealed how the model parameters effect some scalar features of the forecasts (e.g., mean precipitation across a spatial domain). But now we would like to know how the parameters affect the distribution of precipitation? For example, can one set the model parameters so that the mean precipitation is increased, but heavy precipitation is not (or, vice versa)?

Thus far, for each forecast quantity, two summary measures are examined: the intensity, and the coordinate of the center-of-mass of the forecast field. How do the model parameters effect other measures which are more closely related to spatial features of forecasts, e.g., number of clusters, or the texture of the forecast field (e.g., measured by variogram)?

The CCA method is inherently a multivariate technique, but the SA method is not, at least not the version used by us. We aim to develop a multi-output SA method, where the effect of the parameters is examined on multiple forecast quantities, simultaneously. In addition to the three types of
precipitation we have been using, we will also use air temperature and water vapor, because we currently have data on them.

The CCA method does not currently take into account interactions between model parameters and interactions between forecast quantities. This is relatively straightforward to generalize because interactions are easy to incorporate in regression; (the underlying method in CCA is the same as that of regression.)

We have learned that many of our results regarding COAMPS may be generalizable to WRF. To test that hypothesis, we are planning to investigate the use of the Stochastic Kinetic Energy Backscatter (SKEBS) scheme as a complement to a basic set of WRF model parameters used in the SA. This approach aims to represent model uncertainties arising from interactions with unresolved scales. The technique consists of introducing random streamfunction and temperature perturbations from a prescribed kinetic energy spectrum during the course of model integrations. Control over the spatial and temporal characteristics of the perturbations allows the generation of a variety of plausible model solutions.

IMPACT/APPLICATIONS

All of these results can be used to better set model parameters for the purpose of affecting forecasts. This may be for the purpose of improving forecasts (as compared with observations), or for the purpose of having a desirable feature in the forecasts (not necessarily based on observations).

RELATED PROJECTS

Two of the PIs (Marzban and Sandgathe) have received NSF funding on a related project. Sensitivity Analysis (SA) is common to both projects. However, the NSF project aims to develop a verification method, based on cluster analysis (CA), for automatic identification of “objects,” their locations, and their orientation, in both forecast and analysis fields. Progress in the NSF project will aid this project in better assessing how the model parameters affect the number of clusters in the forecast field.

PUBLICATIONS

