HEPEX Testbed Progress Report #1

Statistical Re-scaling and Downscaling Project

Martyn Clark (*National Institute for Water and Atmospheric Research (NIWA) Christchurch, New Zealand*)
mp.clark@niwa.co.nz

John Schaake (*Consultant, NOAA NWS Office of Hydrologic Development Silver Spring, MD*)
john.schaake@noaa.gov

Robert Hartman (*NOAA NWS California Nevada River Forecast Center Sacramento, CA*)
robert.hartman@noaa.gov

1 Project Purpose

1.1 Scientific Question

What are the advantages and limitations of different methods for extracting information from Numerical Weather Prediction models, for the purposes of forecasting streamflow.

1.2 Objective(s)

- Identify the space-time scales for which forecast skill is present, for different variables, and develop methods to extract and combine information at different space-time scales
- Identify the atmospheric forecast variables that can be used to provide sub-grid information—for example, wind and humidity output can be used in a statistical model to replicate orographic precipitation processes, and provide local-scale information that is not present in the raw precipitation output [Clark and Hay (2004) have a table that summarizes the frequency that different variables are used in their regression equations].
- Identify of the sample size required to reliably forecast precipitation, temperature, and streamflow, for different thresholds.

2 Accomplishments

2.1 Background

Hydrologic forecast models are highly non-linear. As a result they are very sensitive to biases in the climatology of input forcing, especially precipitation forcing, as well as biases in the spread of ensemble forecasts of input forcing. Because hydrologic models integrate the effects of input forcing in time and space, joint distribution properties of input forcing between different times and places, hydrologic models are also sensitive to covariance properties of the input forcing. These covariance properties are scale dependent in space and time.

There are two fundamental problems in using information from atmospheric forecasts to create input forcing data sets for hydrologic prediction. One is that there is a scale mis-match between the spatial scales at which atmospheric models operate and spatial scales at which hydrologic models operate. Typically atmospheric models operate at larger spatial scales than the computational scale of hydrologic models. As a result variables such as precipitation that are highly variable in both space and time do not have the same variability characteristics as are required by hydrologic models. In addition, the local climatology of atmospheric models typically is biased relative to the observed local climatology. As a result, atmospheric forecasts must be both re-scaled and downscaled to be most useful for hydrologic prediction.

Research on how to assure that forecast information from atmospheric forecast models is used well to create input forcing for hydrologic models is in its infancy. For example, there is a long history of work done to understand the nature of precipitation variability in space and time. But there is a wide gap between that work and the work that is needed to re-scale and downscale precipitation forecasts for hydrologic application.

The work that has been done for this test bed so far has focused on building an initial, "reference" ensemble preprocessor for use by the National Weather Service to create ensemble precipitation and temperature forcing for input to its Ensemble Streamflow Prediction (ESP) system. Much of this work has built on earlier work on this problem (Clark and Hay, 2004; Clark et al., 2004; Werner et al, 2005).

2.2 NWS Ensemble PreProcessor Development

During the past year our main accomplishment has been to develop procedures to use ensemble mean precipitation and temperature forecasts from a fixed version of the National Weather Service (NWS) Global Forecast System (GFS) to produce ensemble forcing for input to hydrologic forecast models at the space and time scales at which the models operate. These procedures were implemented in a prototype operational pre-processing system that is running in an experimental mode at the California Nevada River Forecast Center (CNRFC). These preprocessor procedures are part of what is called the "GFS Subsystem". The GFS Subsystem allow use of both short and medium range single-value forecasts to produce ensemble forecasts for specific hydrologic basin areas. These procedures are operating at several pilot locations at the CNRFC. Earlier versions of the preprocessor that operate only for very short range forecasts are operating at the Arkansas Basin River Forecast Center (ABRFC) and the Middle Atlantic River Forecast Center (MARFC).
An article describing the procedures used in the GFS Subsystem has been submitted for publication in a special issue of the European Geophysical Union's (EGU) Hydrology and Environmental Sciences (HESS) journal. (Schaake et al, 2006).
These procedures require an historical archive of single-value forecasts and corresponding observations. A fixed version of NCEP's GFS ensemble forecast system was used by the NOAA Climate Diagnostic Center (now part of the Earth System Research Laboratory (ESRL) as part of the Physical Sciences Division of NOAA's Office of Ocean and Atmopheric Research (OAR)) to make ensemble reforecasts from 1979 to

date (Hamill et al, 2006; Hamill et al, 2004). The reforecast data as well as current forecasts can be found at http://www.cdc.noaa.gov/reforecast/. These data were used to create an archive of ensemble mean forecasts of precipitation and temperature. The archive of corresponding observations was available for the RFC forecast basin segments at each RFC.

Because atmospheric ensemble forecasts do not reliably estimate the conditional uncertainty (i.e. the spread is typically underestimated), only the ensemble mean forecast is used at the present stage of preprocessor development. Empirical statistical relationships between the single-value forecasts and the corresponding observations are used to account for uncertainty and to control the generation of ensemble members.

2.3 Supporting Data Set Development

In order to encourage potential research community participation in the activities of this test bed, the data sets being used by NWS to develop its preprocessor procedures are available via anonymous ftp. This includes: (i) an archive of GFS ensemble mean forecasts and (ii) an archive of corresponding observations.

Documentation of the supporting data sets is located at ftp://hydrology.nws.noaa.gov/pub/gcip/hepex/documentation/

2.3.1 GFS Ensemble mean Forecast Archive

Historical ensemble mean precipitation and temperature forecasts from the fixed version of NCEP's GFS ensemble forecast system are located under the directory ftp://hydrology.nws.noaa.gov/pub/gcip/hepex/gfs_archive/. There are separate subdirectories for precipitation and temperature. In each of these subdirectories there is a file named ensmean.zip that contains the data.

The original GFS ensemble forecast data included 15 members. The forecasts were made once a day at 0000z. Twice daily data values for both precipitation and temperature were given for a 15-day lead time (i.e. 30 values for each member). The forecasts were on a global 2.5 x 2.5 degree grid (144 rows and 73 columns).

The mean forecast data provided in this directory are for a domain that only covers the U. S. There are 300 grid points in this domain. The location of row and column grid point centroids and the corresponding grid element indices in the global grid are given in the file gfs_fcst_data.doc in the HEPEX/documentation/ directory.

The precipitation and temperature directories contain a file for each grid point. There is a record in each file for each forecast day. The date in the forecast record is the date the forecast was created. Precipitation data were aggregated to daily values for the period 12z-12z. (NWS RFC's operate on a 12z-12z forecast day). So the forecasts are for 14 days. There are 14 precipitation forecast values and 28 temperature forecast values. Precipitation units are mm. Temperature units are degrees Celsius.

2.3.2 Observations Data Archive

Historical 6hr Mean Areal Precipitation (MAP) values for more than 430 U.S. river basins are located in directory ftp://hydrology.nws.noaa.gov/pub/gcip/hepex/gfs_archive/mopex/US_Data/6hr_MAP/z12_z12/ascii/. These river basins are being used in the Model Parameter Estimation Experiment (MOPEX) (Schaake et al, 2006). The locations of these basins and information about the USGS stream gages are located in the subdirectory gage locations are in /mopex/basin_characteristics/. Documentation about the preparation of the data and availability of all of the MOPEX data can be found in /mopex/documentation/. The date stamp on these data are for the day on which the valid 12z-12z period ends. Historical mean areal daily precipitation data and max/min temperature data for more than 430 U.S. MOPEX basins are located in the directory ftp://hydrology.nws.noaa.gov/pub/gcip/mopex/us_data/us_438_daily. These data are on a local 24hr clock. As a result the daily precipitation data are out of phase with the daily GFS 12z-12z precipitation data. Therefore the 6hr MAP data described above should be used with the GFS precipitation forecasts. But the daily max min temperature data can be correlated with the 12hr GFS temperature forecasts.

2.4 Example Results

As explained in section 2.2 above an article describing the procedures used in the GFS Subsystem has been submitted for publication in a special issue of the European Geophysical Union's (EGU) Hydrology and Environmental Sciences (HESS) journal. (Schaake et al, 2006). This article presents selected results for an application of the procedures to the North Fork of the American River basin in California. An example result presented in that publication is presented here. Two different sources of precipitation forecasts for lead times out to 5 days were analyzed. These sources were short-term, single-value forecasts produced by the River Forecast Center (RFC) using a wide range of operational forecast guidance products. The other was the ensemble mean forecast of precipitation from a fixed version of the GFS ensemble forecast system. One of the summary verification statistics used to analyze the results is the continuous rank probability skill score (CRPSS). More about this verification statistic can be found in the paper. But the statistic was applied to each of the two the raw single value forecasts and to ensemble forecasts genated by the EPP for each of them. This was applied to forecasts for 6hr periods out to 5-days (20 periods) plus an additional set of 8 periods were created by accumulating precipitation for different aggregate forecast periods during the 5-day forecast window (see Schaake et al, 2006). This was applied for re-forecasts made during a 5-year period and for all days of the year. The values of the CRPSS are presented in Figure 1 below. This figure shows that the CRPSS varies seasonally, with forecast lead time and with length of temporal accumulation (forecast periods 21-28). This figure also shows that the EPP consistently produced better CRPSS values than the input single-value forecast.

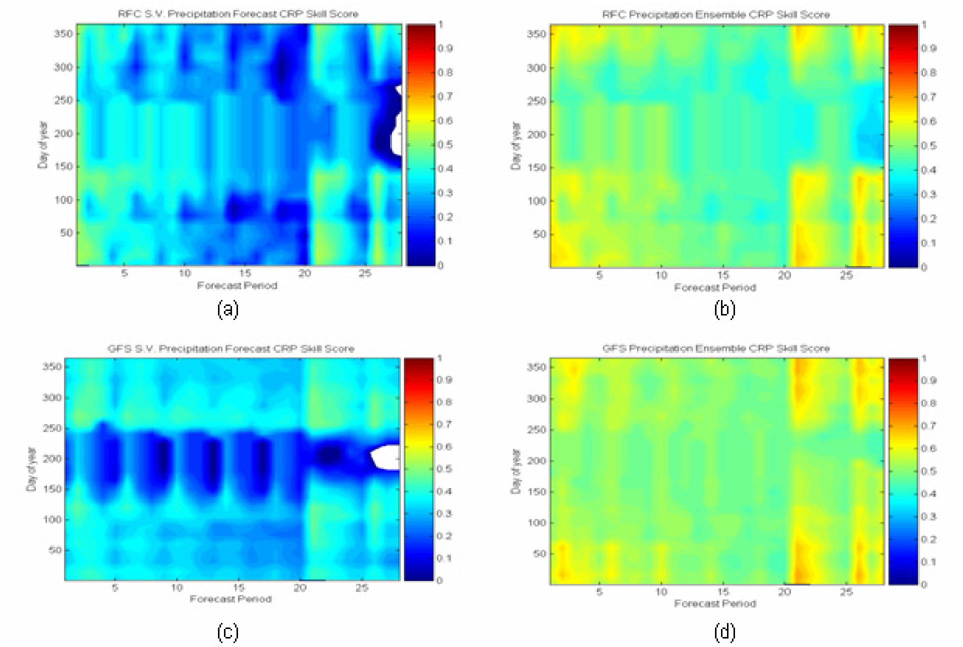

Figure 1 – Continuous Rank Probability Skill Score: (a) Single-value RFC forecasts, (b) Ensemble forecasts based on RFC single-value forecasts, (c) Single-value GFS forecasts, (d) Ensemble forecasts based on GFS single-value forecasts.

3 Future Plans

- Our future plans include the following:
- Continued development of the NWS Ensemble PreProcessor
- Algorithm development
- Testing of alternative algorithms
- Collaboration with other testbeds
- Supporting data set development

3.1 Continued Development of the NWS Ensemble PreProcessor

Development of the NWS ensemble pre-processor is just beginning. Our strategy has been to keep the pre-processor as simple as possible using existing operational products as input. Our immediate goal is to create a "reference" system that could be implemented quickly, be easy to use and that could serve as a basis of comparison for future improvements. Some of the future improvements include:

3.1.1 Include Long Range Forecasts (8 months) from NCEP's Climate Forecast System (CFS)

This will be done during the coming year. It will extend the forecast lead time from 14 days to about 8 months. It will use an ensemble mean forecast from the CFS forecast system. It will be possible to run the pre-processor as frequently as every day.

3.1.2 Gridded EPP

The initial EPP produces ensemble forcing for hydrologic basins directly from the input atmospheric forecast input. This makes it difficult to consider the spatial scale dependent uncertainty in the forecasts. And it means that EPP parameters must be estimated separately for each hydrologic forecast segment. An alternative approach would be to produce gridded ensemble forcing and then process the grids to create the hydrologic segment forcing. We plan to begin working on this during the next year.

3.1.3 Additional Algorithms

The initial algorithms were chosen because they were simple, parsimonious with parameters that could be estimated with limited data, even for arid areas. Potentially better algorithms might include: explicit estimation of the probability of precipitation, analog techniques, alternative scale dependency procedures, procedures to use individual members from the atmospheric ensemble forecasts to account for the conditional uncertainty (existing procedures produce the same uncertainty for a given day of the year that is independent of uncertainty estimates available the atmospheric forecast information for that day), procedures to integrate information from multiple forecast models and procedures to integrate long-range empirical statistical forecast information. More careful attention needs to be given to the definition of the probability of precipitation as a function of space and time scale and for precipitation observations/analyses vs precipitation forecasts and to the uncertainty inherent in estimation of the occurrence of precipitation from observations.

3.1.4 Verification Statistics

We plan to work toward having a unified suite of verification statistics that include summary statistics as well as diagnostic statistics such as found in reliability diagrams and in statistics decompositions that measure effects of bias, resolution and climatological uncertainty. We expect to make progress on this during the coming year.

3.1.5 Forecaster Options

We plan to consider the possible ways in which the forecaster might control the operation of the EPP. This might include real time control of EPP options, real time diagnostic information about EPP output and possible forecaster over-ride of EPP output.

3.2 Algorithm Development

We will continue to monitor work being done by the international meteorological and hydrological communities on new algorithms.

3.3 Testing of Alternative Algorithms

It is essential that we work to evaluate alternative algorithms and compare alternative algorithms with each other. Ideally this would be done using a common data base and using common metrics. In the near term, the testbed project will do this for algorithms that are being considered for use in the NWS EPP. In the longer run it may be desirable to have a HEPEX intercomparison of alternative EPP algorithms. We plan to pursue this idea first though possible collaboration with other testbeds. But an intercomparison project involving the broader meteorological and hydrological community might be desirable. This is a topic that could be discussed at the 3rd HEPEX workshop in June, 2007.

3.4 Collaboration with other Testbeds

Most of the HEPEX testbeds are using procedures to pre-process atmospheric forecasts before using them as input to their hydrological forecast models. A range of procedures are being used. Some test-beds are developing innovative new and sophisticated procedures. Others are using very basic procedures. During the next year we will develop plans to test some of these procedures as part of the work of this testbed. We also will see if any of the work being done by this test-bed would be useful in other testbeds.

3.5 Supporting Data Set Development

We will continue to make data sets that are being used for the development and application of the NWS EPP available to the science community. We are looking forward to working with the WWRP/THORPEX/TIGGEE project to test ensemble forecasts that will be produced by that project. And we look forward to being able to include long range forecasts

4 References, Publications and Presentations

Clark, M. and L. Hay, 2004,: Use of medium-range weather forecasts to produce predictions of streamflow. *Journal of Hydrometeorology,* 5, 15-32.

Clark, M., S. Gangopadhyay, L. Hay, B. Rajagopalan, and R. Wilby, 2004: The Schaake Shuffle: a method for reconstructing space-time variability in forecasted precipitation and temperature fields. *J. of Hydrometeor.*, **5** , 243-262.

Hamill, T. M.**,** J. S. Whitaker, and S. L. Mullen, 2006: Reforecasts: An important dataset for improving weather predictions. *Bulletin of the American Meteorological Society*, **87**, 33-46.

Hamill, T. M.**,** J. S. Whitaker**,** and X. Wei, 2004: Ensemble reforecasting: Improving medium-range forecast skill using retrospective forecasts. *Mon. Wea. Rev.*, **132**, 1434-1447

Schaake, J., S. Cong and Q. Duan, 2006: The U. S. MOPEX Data Set, IAHS Red Book #307.

Schaake, J., C., J. Demargne, R. Hartman, M. Mullusky, E. Welles, L. Wu, H. Herr, X. Fan, and D.J. Seo, 2006: Precipitation and temperature ensemble forecasts from single value forecasts. *Hydrology and Earth System Science, hessd-2006-0094.*

Werner, K, D. Brandon, M. Clark and S. Gangopadhay, 2005: Incorporating medium-range numerical weather model output into the ensemble streamflow prediction system of the National Weather Service, *J. of Hydrometeor.*, **5** , 101-114