doi:10.2489/jswc.64.1.15A

IDEAS & INNOVATIONS

Innovative operational seasonal water supply forecasting technologies

Tom R. Perkins, Thomas C. Pagano, and David C. Garen

Today's Western United States is home to nearly one-third of the American population. The region has experienced rapid population growth in recent years. Continued urban growth, combined with economic changes and federal environmental mandates, are exerting increased pressure on the Western United States' irrigated agriculture. Serving these competing uses, in the face of uncertain and extreme weather, can be difficult. Managers of a fixed water supply must operate their projects within an ever-shrinking margin of error (Western Water Policy Review Advisory Commission 1998). Improved water supply technologies and products will help resource managers to improve water use efficiencies wherever possible. This article describes some of the recent advances by the USDA Natural Resources Conservation Service in its mission to support Western US water managers.

HISTORY OF WATER SUPPLY FORECASTING TECHNOLOGIES

About 100 years ago, Dr. James Church, a professor of the Classics at the University of Nevada, Reno, pioneered snow survey measurements and water supply forecasting in the Western United States. In 1910, using a simple seasonal percentage technique, he was able to relate the water content in the snowpack on Mt. Rose to the spring rise in the Lake Tahoe water level (Church 1917). His early forecasts were used by water managers to make decisions of how the lake water level was to be managed to provide for downstream power development, municipal and irrigation rights, and to prevent flooding of the lakeshore dwellers (Church 1937). Church's early efforts, along with significant contributions from others, led to the eventual funding of the Cooperative Snow Survey Program in 1935, under the USDA Bureau of Agricultural Engineering. The Bureau became the Soil Erosion Service, soon named the Soil Conservation Service (SCS) and eventually the Natural Resources Conservation Service (NRCS) (Helms 1991). The NRCS Snow Survey and Water Supply Forecasting Program is now centered at the National Water and Climate Center (NWCC) in Portland, Oregon, with support from offices in each Western state.

Tom R. Perkins, Thomas C. Pagano, and **David C. Garen** are hydrologists at the National Water and Climate Center, USDA Natural Resources Conservation Service, Portland, Oregon, www.wcc.nrcs.usda.gov.

The NRCS methods using snow tubes to measure the depth and water equivalent of snow have not changed significantly since 1909, when Dr. Church collected the first recognized snow course samples on Mt. Rose. However, there has been substantial change in the automated measurement and reporting of snow and other meteorological and hydrological data. Research that started in the late 1950s has evolved into the current automated snow telemetry (SNOTEL) system. SNOTEL data are transmitted by reflecting signals off of ionized meteor trails in the Earth's ionosphere (Barton and Burke 1977). This "meteor burst technology" is used to transfer snow water equivalent, snow depth, precipitation, temperature, wind, solar radiation, relative humidity, and soil moisture data to a central computer facility in Portland, Oregon. There are currently 758 active SNOTEL sites throughout the Western United States, collectively producing over 16 million observations per year.

Similarly, NRCS water supply (seasonal streamflow volume) forecasting technologies have seen many changes over the more than 70 years that the agency has been providing outlooks for use by Western US water managers. Forecasting methods have evolved from simple snow water/seasonal runoff percentage relations to complex statistical regression models.

Initially, the seasonal percentage of measured snow was applied to the normal runoff from a basin to forecast the expected basin runoff (Church 1917). Gradually, as more data were accumulated, laborious hand calculations of simple linear regressions followed. In the mid-1950s, SCS forecasters partnered with universities to speed up calculations by using some of the earliest mainframe computers. Since that time, NRCS has moved from communicating with remote computers through acoustic modems to using powerful, portable laptop computers to develop statistical runoff models and provide water supply guidance to the water user community.

Crater Meadows SNOTEL site, Idaho. Hypalon snow pillows (white circles about 3 m in diameter) measure snow water equivalent and are the primary basis for most Western US water supply forecasts. Precipitation gauge is in upper left. Other instruments are mounted on the tower between the two snow pillows. All SNOTEL sites are solar powered.

VIPER TECHNOLOGY DEVELOPMENT

Due to increased climate variability (Redmond and Koch 1991; Pagano and Garen 2005) and user demands for more accurate and timely forecast information, NRCS hydrologists were faced with the need to develop new and better forecasting technologies in a timely fashion. Until 2006, the agency was using a forecasting software package developed by NRCS circa 1992 (Garen 1992). This package, despite its limited visualization and interactivity, became somewhat of a standard for water supply forecasting among several agencies in the Western United States and Canada (Smith and Weiss 1996; Wortman 2005).

The new NRCS statistical model development and forecasting environment is called Visually Interactive Prediction and Estimation Routines (VIPER). VIPER's primary features include portability, Internet-based data collection, a richly visual interface, interactive forecast equation development and management, optimized predictor search routines, report generation, and forecast production management.

Restructuring of the agency's information technology infrastructure has posed challenges to innovation, in particular the need to comply with the increasingly complex cyber-security requirements of the federal government. This influenced the decision to use the agency-approved Microsoft Excel 2003, with its VBA macro language and associated integrated development environment. Given the ubiquity of the Microsoft Office suite, VIPER is portable and usable by external partners.

Timely and accurate real-time data are critical to operational forecasting. The NWCC maintains a central Air and Water Data Base (AWDB) of all necessary historical and real-time data for use in forecasting. Anytime a predictor or predictand is added or changed on VIPER's main interface, the program queries the database (using Web services) for the period of record data. Hydrologist-created local comma delimited (CSV) files can also be used for data not contained in AWDB. VIPER can also retrieve real-time and historical streamflow data automatically from the US Geological Survey's National Water Information System. VIPER has extensive capabilities for managing and editing data, and virtually any type of data needed for statistical analysis can be accommodated. In addition, any predictor (e.g., snowpack, precipitation) can become a predictand, making VIPER a useful tool for missing data estimation as well.

VIPER's colorful interface may appear overwhelming at first, but it is visually immersive and designed to give the hydrologist convenient access to a wealth of important information. Scatter plots and time-series charts of forecast versus observed are displayed along with charts that show the availability of data for individual predictors. Tables of numbers describe the quality of predictors, their historical correlation with the predictand, and the forecasts based on each individual predictor or by all predictors as a whole. VIPER's visual output formats replace innumerable pages of printouts and tables and provide an easy way to interpret the effectiveness of the developed statistical models.

VIPER's interactive model development component assists the user in determining the best combination of predictors, places recent hydrologic variability in a historical context, and allows "what if" analysis for the forecasts. Predictors can be quickly added, modified, and deleted to establish the candidate predictor slate. The forecasting equation is recalibrated and results displayed automatically with each change made by the user, providing instantaneous feedback on the quality of the models being developed. The forecast for current conditions is displayed, and changing predictors allows the hydrologist to perform "what if" analysis and test the stability of the results. When the forecaster is satisfied with the developed equation, it can be saved for later use.

A hydrologist may have a large pool of predictors to select from, including a mix of stations, elements (e.g., snowpack, precipitation, temperature), and time periods. VIPER contains optimization tools both for selecting time periods and for selecting optimal combinations of predictors (some of which are described in Garen 1992). The program systematically tests potentially thousands of combinations of candidate variables to identify the 30 most accurate models. These optimizations, however, do not preclude the application of the hydrologist's judgment, although it does remove the laborious guesswork and trial and error of an unaided search.

VIPER is capable of producing several specialized reports. Most frequently, condensed forecast reports are generated for water users. These "quick look" reports contain a succinct summary of the current forecasts. Also, extensive and detailed reports of the input and output of hundreds of real-time forecast equations can be generated.

NRCS forecasters use two computer monitors while forecasting, one containing VIPER and another containing a geographic information system (GIS) interface to map-based data. To allow easy communication between the two applications, VIPER produces GIS-ready, CSV files of data and forecast information. This is particularly useful for visual data quality control and missing data estimation.

Forecasts can be produced using the interactive interface, one station at a time, or for entire river basins in batch mode, evaluating the saved regression equations for each point. Designed for full operational production, VIPER also has tools for generating and organizing forecasts, from developing rough drafts of forecasts, to promoting them through various stages of completion into a finished product to be delivered into machine-readable files or human-readable reports. NRCS forecasts are produced jointly with other agencies, primarily the National Weather Service, and VIPER allows the storage of other agencies' forecasts, comments about the forecasts, and other information.

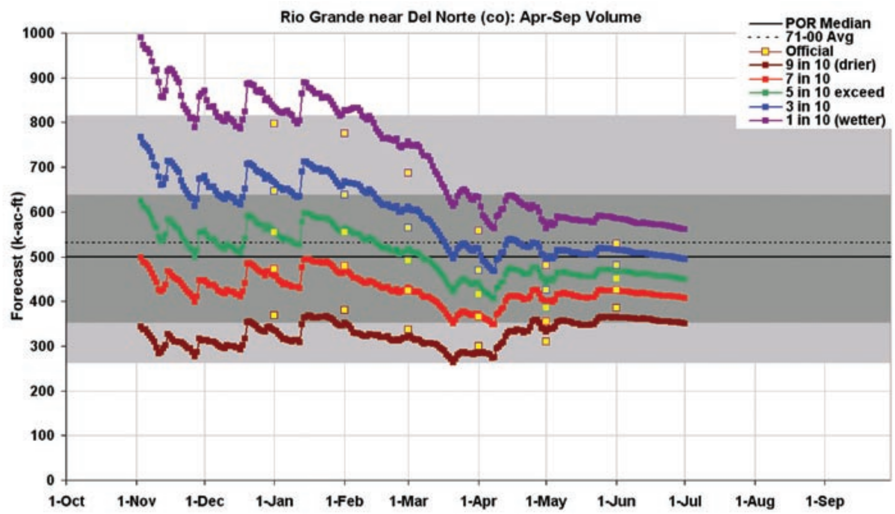
INNOVATIVE REGRESSION ANALYSIS

VIPER uses multiple regression analysis to assess the relative importance of the predictor variables and to formulate and verify the forecasting model. Two types of multiple regression analysis are used: principal components and Z-score regression. Both of these techniques involve preprocessing of the input data to overcome issues of intercorrelation (collinearity) among predictors, varying record lengths, and missing data that preclude a straightforward application of standard multiple regression.

Most predictor variables used in water supply forecasting are highly intercorrelated. If used in stepwise regression, many of these predictors will be eliminated from the model, whereas if used in a standard multiple regression, the coefficients can often be ill-defined or be of the opposite algebraic sign from their correlation with the predictand (McCuen 1985). It is desirable for model robustness not to eliminate many predictors that are highly correlated with the predictand, and it is desirable that the regression coefficients be reliable and have a physically meaningful algebraic sign. Principal components regression addresses these issues and is the technique that has been used in the NRCS for many years and has been incorporated into VIPER (Garen 1992; USDA NRCS 2007b).

Those developing statistical models are often faced with uneven record lengths for candidate predictor variables. Forecasting has the added challenge of missing real-time data. Principal components regression requires a continuous record for all variables used in the analysis. With uneven record lengths, the hydrologist must make a choice of either using a limited set of variables with a long period of record or using more variables with fewer years of record.

This problem can be at least partially overcome by using the Z-score regression technique, which relaxes the requirement for serial completeness. This NRCS-developed technique combines individual predictors into a composite index that then becomes the independent variable used in a regression (Pagano 2005; USDA NRCS 2007b). The compositing is based on first standardizing the data into Z-scores (subtracting the mean and dividing by the standard deviation) and then combining the data into weighted averages, the weights being based on each variable's correlation with the predictand. The calibration sample size of the regression then becomes the union of all predictor record lengths rather than their intersection, allowing the user to mix short period of record stations in with longer ones. Also, if some data are missing during real-time operations, the forecast equation can still be used to produce a forecast without manually estimating the missing data. While one must be mindful of the technique's limitations, it can be an invaluable tool in many operational forecasting situations (USDA NRCS 2007b). Under conditions of serially complete data, principal components and Z-score regression generally produce very similar predictions, although the coefficients on each predictor variable can be quite different.



Daily water supply forecast product. Forecasts are given for five exceedance probability levels (colored lines). Official monthly forecasts (yellow squares) are also shown. Gray shading indicates historical variability for 30% to 70% (dark) and 10% to 90% (light) exceedance probabilities.

DAILY FORECAST UPDATE GUIDANCE

One of the most common requests received by the NWCC from water users is for more frequent updates to the official water supply forecasts. Historically, water supply forecasts have been issued once per month from January through June by a labor-intensive manual process. The schedule of production was tied to the frequency of collection of the manual snow course measurements.

The NRCS SNOTEL network has been providing daily snowpack and precipitation data for nearly 30 years at many sites, and it is now possible to use these data exclusively for the development of statistical models. Therefore, it is no longer necessary to be constrained by the traditional monthly data and forecast time schedule. These daily data are now being used to drive an automated forecast system that provides a daily update to the official monthly water supply forecasts. As storms pass through the watersheds, water managers can now get an instant assessment of how much their water supply situation has improved or deteriorated.

With its ability to handle situations of non-serially complete data, the Z-score regression technique now makes this kind of frequency of forecast updating operationally feasible. This product is based on daily regression models that use SNOTEL snow water equivalent and precipitation as predictor variables. It is a fully automated, self-calibrating system and has built-in filters to optimize predictor variable time periods as the models progress through the water year (Pagano et al. 2008; USDA NRCS 2007a). The current prototype system is running twice daily for about 150 streamflow locations in the Western United States. These graphical products provide a quick way for water users to evaluate the current status and trends of their potential water supply.

CONCLUSION

Water supply forecasting technology has been changing over the past 70 plus years that NRCS hydrologists have been providing products for use by Western US water users and managers. The demographic, physical, and political landscape of the Western United States is changing rapidly, and there will be increased competition over water for irrigation, municipal and industrial uses, and in-stream uses, such as river-based recreation, esthetic enjoyment, fish and wildlife habitat, and hydroelectric power generation. Increasing water demands will require more precise management of this valuable resource. Add to this the fact that climate variability has increased over the past 30 years (Pagano and Garen 2005), which makes it more difficult to forecast seasonal water supplies with the accuracy required for today's complicated water issues. It becomes critical that water managers understand the risks and adaptively modify their operating practices using the best guidance possible. NRCS, through new and innovative technologies, is striving to provide the most accurate and timely information possible to ensure the efficient management and use of the Western United States' lifeblood … water.

REFERENCES

Barton, M., and M. Burke. 1977. SNOTEL: An operational data acquisition system using meteor burst technology. *In* Proceedings of the Western Snow Conference, 82-87.

Church, J.E. 1917. Snow Surveying: Its problems and their present phases with reference to Mount Rose, Nevada and vicinity. Paper presented at the Second Pan American Scientific Congress, Washington, DC, 1-51.

Church, J.E. 1937. The human side of snow. Scientific Monthly 44(2):137-149.

Garen, D.C. 1992. Improved techniques in regression-based streamflow volume forecasting. Journal of Water Resources Planning and Management 118(6):654-670.

Helms, J.D. 1991. Bringing federal coordination to snow surveys. *In* Proceedings of the Western Snow Conference, 119-125.

McCuen, R.H. 1985. Statistical Methods for Engineers. Englewood Cliffs, NJ: Prentice-Hall.

Pagano, T.C. 2005. The role of climate variability in operational water supply forecasting for the western United States. PhD dissertation, Department of Hydrology and Water Resources, University of Arizona.

Pagano, T.C., and D.C. Garen. 2005. A recent increase in western US streamflow variability and persistence. Journal of Hydrometeorology 6:172-179.

Pagano, T.C., D.C. Garen, T. Perkins, and P. Pasteris. 2008. Daily updating of operational seasonal water supply forecasts for the western US. Journal of the American Water Resources Association (forthcoming).

Redmond, K.T., and R.W. Koch. 1991. Surface climate and streamflow variability in the western United States and their relationship to large scale circulation indices. Water Resources Research 27:2381-2399.

Smith, S.E., and E. Weiss. 1996. Application of principal components regression to streamflow forecasting. *In* Proceedings of the Western Snow Conference, 47-58.

USDA NRCS. 2007a. Development and interpretation of daily seasonal water supply forecasts. Electronic Directives System, Technical Notes, Snow Survey and Water Supply Forecasting. http://directives.sc.egov.usda.gov/tn_sswsf_1_a.pdf.

USDA NRCS. 2007b. Statistical techniques used in the VIPER water supply forecasting software. Electronic Directives System, Technical Notes, Snow Survey and Water Supply Forecasting. http://directives.sc.egov.usda.gov/tn_sswsf_2_a.pdf.

Western Water Policy Review Advisory Commission. 1998. Water in the West: Challenge for the Next Century. Springfield, VA: National Technical Information Service.

Wortman, R.T. 2005. Improved water supply forecasting for Kootenay Basin. NDIA Tri-Services Infrastructure Conference, St. Louis, MO.