PROBABILISTIC FLOOD FORECASTING

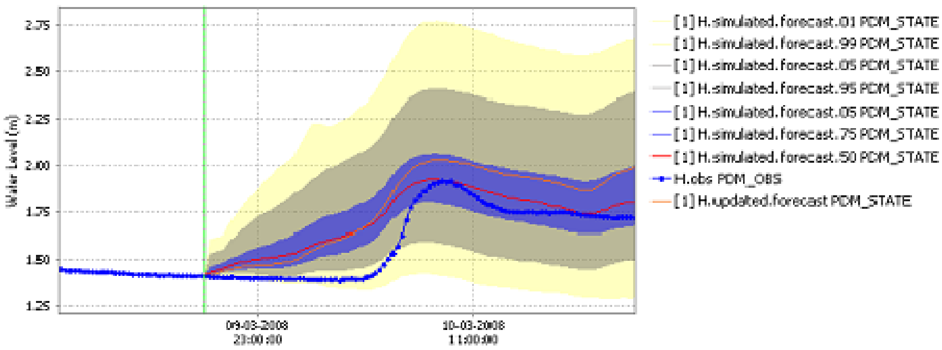
Mike Vaughan
Flood Forecasting Technical Specialist, Environment Agency

The Environment Agency provides a forecasting and warning service to people at risk from flooding. However, flood forecasts are inherently uncertain. There are differences (errors) between forecast time series of river level and subsequent observations. These differences can be relatively large, so understanding uncertainties is useful when interpreting forecasts for decision support.

We recently investigated whether information on historic flood forecast performance could be analysed to give an estimate of uncertainty around current flood forecasts in real time. This 'rear mirror' view assumes that previous error relationships continue to hold. Since it is relatively easy to compute, we saw it as a quick way to quantify errors using existing historic performance information. In the longer term, ensembles of forecast rainfall may be introduced to complement this approach.

A published, peer reviewed paper (Weerts et al., 2011) describes the technique. It uses quantile regression to determine non-parametric relationships between the quantiles of the error distribution of the flood forecasts (estimated from many historic forecasts) and forecast magnitude and lead-time. These relationships are used to calculate estimates of the uncertainty in real-time forecast time series in the form of error quantile bands. Quantiles are presented as a plume around the forecast, with lines for different percentiles (Figure 1).

Figure 1 : An example forecast plume



Recent evaluation of the method, on a larger dataset, has found that a higher-than-expected proportion of observed flood peaks fall above the upper uncertainty bound of the plume. Specifically, forecasts were made, with uncertainty bands, for every significant observed event in a 20 year period for 10 locations in the Midlands. We calculated the number of observed peaks that exceeded the 1% and 5% quantile and found typically more than 50% of peaks going over the 5% level. Although the bands were found to encapsulate the right proportion of observations overall, it is the peak magnitude that is of most interest to forecasters and the public.

One likely reason for this problem is that the underlying statistical approach does not capture the full non-stationarity of the errors. In particular, errors associated with the observed peaks are typically larger than for the flow range considered as a whole.

This may be caused by rivers' tendency to fall more slowly than they rise, giving flood recessions undue weight in the quantile estimation. As model errors are much smaller for falling water levels, this tends to yield narrower uncertainty bands.

Another contributing factor is that rainfall inputs to forecasting models may themselves be forecast leading up to a peak, making them less certain. However, the effect is still clearly visible when models are fed observed rainfall instead of forecast data.

Another relevant factor is that we commonly derive quantities from the forecast time series to support decisions we make. The magnitude of the flood peak and the timings of any threshold crossings (e.g. the level at which property flooding commences) as the river rises are the most important of these. Timings of the flood peak and of threshold crossings as the river falls are of lesser interest. In some cases, such as managing artificial flood storage areas, the total volume of river flow is also of interest. Consequently, we would like to be able to interpret the uncertainty estimates associated with the forecast time series to allow us to infer the uncertainties of the derived quantities of interest. We would, for example, like to be able to use the technique to make statements like ‘we are 95% confident that River A will not reach Level 1 in the next 12 hours’. This is not possible using the method as it stands.

Our question to the study group

How can the Quantile Regression approach of Weerts et al. (2011) be improved to provide representative and meaningful estimates of the uncertainty of flood forecasts?

Available data sets

We propose to provide a time series forecast dataset for one location in a simple ASCII format to allow analysis by delegates. We will also supply the dataset that was used to develop the quantiles.

Contact:

Mike Vaughan
(01903-832116)
michael.vaughan@environment-agency.gov.uk

References

Weerts, A.H., Winsemius, H.C. and Verkade, J.S. (2011). *Estimation of predictive hydrological uncertainty using quantile regression: examples from the National Flood Forecasting System (England and Wales)*. Hydrology and Earth System Sciences 15:255:265. http://www.hydrol-earth-syst-sci.net/15/255/2011/hess-15-255-2011.html