Probabilistic flood forecasting

Problem presented by

Mike Vaughan

Environment Agency

Executive Summary

The Environment Agency provides a forecasting and warning service to people at risk from flooding. However, flood forecasts are inherently uncertain. Efforts to quantify the uncertainty based on quantile regression have failed to capture the full extent of the uncertainty associated with significant flooding events.

An investigation into factors that may be correlated with the uncertainty lead to the observation that there are structural biases in the model. It is possible to remove these, and thereby reduce the mean square error of the predictions, but the benefit of this is apparent in the prediction of 'normal' conditions, rather than in flood predictions.

Additionally, a tweak to the linear fit in the quantile regression is suggested which is better suited to the data.

Version 2.1
October 21, 2012
iii+11 pages

Report author

Paul Munday (Industrial Mathematics KTN)

Contributors

Natalie Pearson (University of Oxford)
Georgina Lang(University of Oxford)
Emma Warneford(University of Oxford)
Zubin Siganporia(University of Oxford)

ESGI85 was jointly organised by
The University of East Anglia
The Knowledge Transfer Network for Industrial Mathematics

and was supported by
The Engineering and Physical Sciences Research Council

Contents

1 Introduction

1.1 Background and scope

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

(1.1.2) Historic flood forecast performance data has been analysed to give an estimate of unceratinty of current flood forecasts in real time. This approach assumes that previous error relationships continue to hold. The paper (Weerts et al, 2009) describes a technique for doing this based on quantile regression.This is used to determine non-parametric relationships between the quantiles of the error distribution of the flood forecasts and the forecast magnitude and the lead-time of the forecast.

(1.1.3) An evaluation of the method has found that a greater-than-expected proportion of observed flood peaks fall above the upper uncertainty bounds. For instance, restricting observations to significant events it was found for some locations that typically more than 50% of peaks exceed the 5% level. However, overall the uncertainty bounds were found to contain the correct proportions of observations. The problem is that it is the peaks that are of most interest to forecasters and the public.

(1.1.4) It is thought that this may be due to non-stationarity of the errors. In particular, rivers tend to rise more quickly than they fall, and the errors in prediction are much smaller for falling water levels. Another possible factor is the use of forecasts (which may be inaccurate) for rainfall. However, even when the actual observed rainfall is used as a model input, the problem persists.

(1.1.5) Amongst the quantities that are of most interest are the magnitude of the flood peak, and the time at which a particular threshold is crossed. Currently, the uncertainties in the predictions of these are not quantified.

1.2 Available data

(1.2.1) For this study group problem, the data consisted of observations of river height at a single location over a number of years, together with predictions based on the actual observed rainfall data. The predictions are made every two hours, at lead times of 1 to 40 hours. Observations are taken every 15 minutes.

(1.2.2) A subset of these data were considered. Only lead times of an even number of hours were considered, together with the corresponding observations.

This results in a straightforward, and largely complete, dataset where at time steps of 2 hours we have the observed value at that time, the predicted value 2 hours previously, the predicted value 4 hours previously, . . ., the predicted value 40 hours previously.

(1.2.3) Naturally, the predictions made at longer lead times are more inaccurate. Consideration was given to aspects of modelling that could give rise to increased uncertainty in the predictions.

2 Theoretical sources of variability

2.1 Motivation

(2.1.1) From the data provided, we note that the error predictions are particularly poor when there is a rapid increase in water level, and the model often underestimates large flood events.

(2.1.2) Whilst we have not been given the model used by the Environment Agency to predict river levels following rain, we are aware that the error estimates do not take account of the rate of increase or decrease of water level. Therefore we examine the gradients in water level, with the aim of understanding their effect on the evolution of water levels over time. In this section we look at a well established model for river flow, with the aim of linking greater errors in prediction to physical effects.

2.2 Shallow water equations

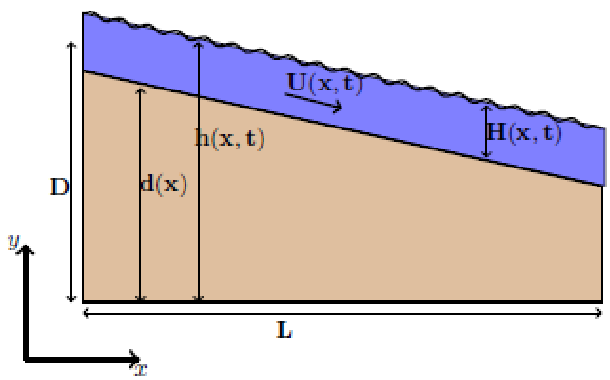

Figure 1: Diagram of modelling set-up including relevant parameters for the shallow water equations

We consider flow along a river with a bed of constant slope. We assume that the section of river under consideration is of typical length L and drops a height D as shown in Figure 1. At any point the top of the water is of height $h(x,t)$, and the bed of height $d(x)$, above the reference height. Thus the water depth at any point is $H(x,t) = h(x,t) - d(x)$. The water velocity along the river is given by $U(x,t)$.

The governing equations for the above system are derived using the principles of conservation of mass and momentum, and then averaged over the depth of the river to give respectively

$$\frac{\partial H}{\partial t} + \frac{\partial (H\bar{u})}{\partial x} = 0, \tag{1}$$

$$\frac{\partial (H\bar{u})}{\partial t} + \frac{\partial (H\bar{u}^2)}{\partial x} = -\rho g H \frac{\partial (d+H)}{\partial x} - \alpha \bar{u}^2. \tag{2}$$

Here, $\bar{u}$ is the depth averaged horizontal velocity, ρ is water density, g is the gravitational constant, and we have assumed that shear stress is proportional to $\bar{u}^2$ with co-efficient α.

We non-dimensionalise using L as the horizontal length scale, D as the height scale, and we choose our velocity scale by balancing gravity with shear stress. We use the shallow water assumption that $\epsilon = D/L \ll 1$, and obtain the equation set

$$\frac{\partial H}{\partial t} + \frac{\partial (H\bar{u})}{\partial x} = 0, \tag{3}$$

$$0 = -H \frac{\partial (d+\epsilon H)}{\partial x} - \bar{u}^2, \tag{4}$$

where we have neglected $O(\epsilon^2)$ terms. We now compare the $\epsilon = 0$ to the $\epsilon \ll 1$ case, in order to determine the importance of this term.

We solved the $\epsilon = 0$ case with humped initial conditions to represent heavy rainfall upstream. Results are shown in Figure 2, which demonstrates the shock behaviour that this system can exhibit.

This shock behaviour is unphysical and hence could be a source of inaccuracy in a simple model. These findings show that it is important to take into account the magnitude of the water gradient when predicting river depth.

2.3 Hysteresis

Next we compare the behaviour of the $\epsilon = 0$ and $\epsilon \ll 1$ models, as the water level increases and decreases. By plotting the flux versus water depth at fixed x over time, we see that the $\epsilon \ll 1$ model exhibits hysteresis, whereas the $\epsilon = 0$ model does not (as shown in Figure 3).

From this we can see that the more complicated model picks up behaviour not seen in the simpler system, indicating it is important to take into account whether water levels are rising or falling to generate more realistic models.

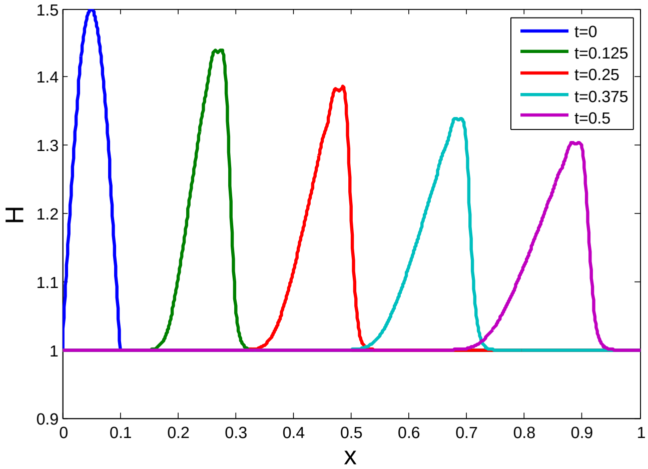

Figure 2: Evolution of the solution of the $\epsilon = 0$ equations with the initial condition shown by the dark blue line

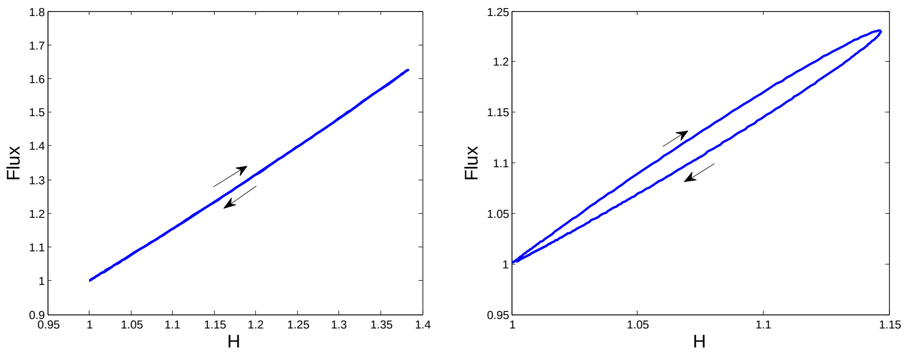

Figure 3: Plots of water flux against depth of the river at a fixed location as a flood passes through (arrows show increasing time). $\epsilon = 0$ is on the left, with $\epsilon = 0.02$ on the right.

3 Analysis of the data

3.1 Factors affecting uncertainty

(3.1.1) An attempt was made to find factors affecting the uncertainty of the 12 hour lead time prediction. Firstly, since it is already known that there is a relationship between predicted river height and the error, a simple plot was produced (figure 1)

(3.1.2) Two things are apparent from this plot. Firstly, it is clear that the error is indeed dependent on the predicted height. Quantile regression seems like an appropriate method to attempt to capture this effect. Secondly, the mean of the error is clearly dependent on the predicted height. That is, there is a bias in the prediction. In particular, low predicted heights are liable to be underestimates and higher predicted heights are liable to be overestimates.

(3.1.3) It is possible to quantify the magnitude of this bias, and to adjust the predictions so that the means of the residual errors are approximately zero irrespective of the predicted river height. This is a very natural step, however whether it is an improvement is very much contingent on how model performance is scored. If it is desired to reduce the mean squared error of the time series (this would be a typical objective) then the model output could be adjusted to remove the bias and the score will improve. However, it is clear that the problem of the predictions being underestimates for events of interest (where the river is observed to be unusually high) this correction will make the problem worse!

(3.1.4) There is reason to suspect, both from theoretical considerations and from plots of the residuals that the rate of change in river height may have an effect on the accuracy of the predictions. There are a number of ways that rate of change can be calculated, we chose to take the difference between the prediction with lead time of 12 hours and the prediction with lead time of 6 hours as an approximation.

(3.1.5) Figure 2 shows a scatter plot of the error against the gradient. It is clear again that there is a systematic error present - when the river level is predicted to rise significantly, the predicted rise is typically an overestimate. Similarly, predicted falls are also typically underestimated. This suggests that the mean square error of the prediction could be improved by applying some kind of smoothing.

(3.1.6) The structural bias is clearly the dominant effect here, it would need to be removed before an assessment can be made as to the impact of gradient on the spread of errors, and how it interacts with the predicted water level.

3.2 Removing bias

(3.2.1) We have at any time a large amount of information which may be helpful for assessing the uncertainty of a particular prediction. We have the recently observed actual values as well as the recent and current predictions (note there is evidence of correlation in the errors of predictions close in time). This gives a a high dimensional problem, which can be reduced by using principle components analysis (PCA).

(3.2.2) Taking the 10 most important dimensions, as found by PCA, we performed a multiple linear regression against the error. This gave a highly significant result ($p < 10^{-10}$) and by removing the projected error from the predictions, we reduced the sum of the square of the errors by 54%.

(3.2.3) It is important to remember that this step by itself is not necessarily an improvement. Although the mean squared error has been reduced, there may be other measurements of error of greater concern. In particular, one of the effects of this step is to make the predictions significantly less volatile. This means that extreme events are predicted less often. Although many of these may be false alarms, some will be correct and the trade-off may be considered undesirable.

(3.2.4) The main advantage of this process is as a cleaning step, and ensuring that predictions genuinely are unbiased. Having done this, the resulting error may prove easier to analyse and model, allowing the risk of flood to be captured with the uncertainty measurement, rather than having a central prediction which is overly cautious.

4 Comment on existing approach

4.1 Quantile regression

(4.1.1) Quantile regression seems to be an appropriate method for establishing a non-parametric relationship between the error and potential correlates such as predicted river height.

(4.1.2) It is important that assumptions made in the regression (such as fitting a straight line) are justified by the available data.

(4.1.3) Figure 3 shows a linear quantile fit. There is some cause for concern here, in particular it is evident that a disproportionate number of points where the error exceeds the 99th percentile occur when the predicted river height is large. This corresponds with the observations made about the uncertainty of the peaks not being captured. Of additional concern is the fact that the 95th percentile has negative slope, which is unexpected given the greater uncertainty at higher predicted values. This shows that the bias in the model is a dominating effect. Indeed a two-tailed 90% prediction interval would not necessarily include the actual predicted value at high river levels.

(4.1.4) There are a number of options for attempting to correct the issue of the poor linear fit. One of the simplest, which was implemented, is to simply divide predicted river heights into a low-regime and a high-regime and to fit different straight lines in the two regimes. An example is shown in figure 4.

(4.1.5) Since the behaviour in the two regimes is clearly different, it is advantageous to split the fits in this way as this allows for more accurate representations of the possible error. In particular, this results in larger estimates for the upper bound in flood-type conditions, which corresponds with the observations from the data.

(4.1.6) It was not possible in the time available to combine this approach with the one taken in subsection 3.2, but this would potentially be a fruitful avenue for exploration.

5 Conclusion and recommendation

5.1 Conclusion

(5.1.1) The analysis undertaken confirms that it is possible to get a better understanding of the uncertainty of a prediction from information that is known at the time the prediction is made. However, the relationships between these known factors and the uncertainty are complex and a certain amount of experimentation will be required to find a good implementation.

(5.1.2) A surprising feature of the data was the bias in the predictions. This bias was in the direction of the model making larger jumps in the predicted values than the actual observed changes in the river height. Given that a false negative with regard to a flood event is more serious than a false negative, it is perhaps understandable that this has arisen.

5.2 Recommendations

(5.2.1) An attempt should be made to remove the systematic bias from the predictions. Having done this it may then be easier to assess the factors affecting the uncertainty and to quantify this.

(5.2.2) A danger in this approach is that the prediction will not have such high peaks - thereby making the prediction of extreme events even worse (though compensated for by a reduction in false alarms). Therefore, it would be particularly important that the uncertainty is quantified accurately.

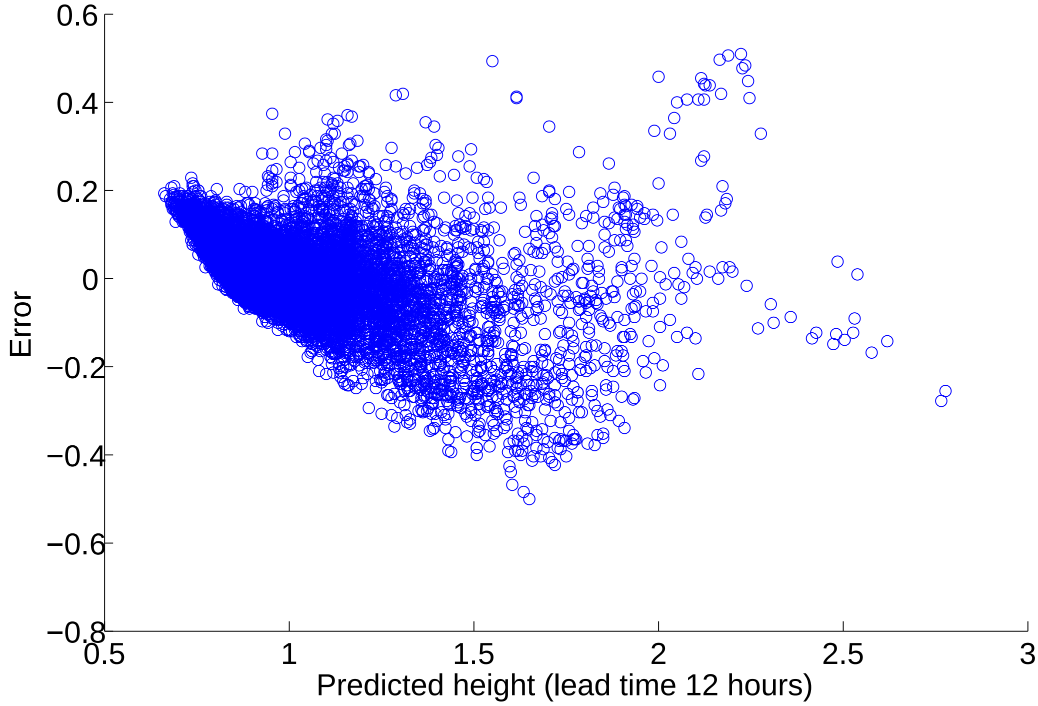

Figure 1: Scatter plot of error against predicted height

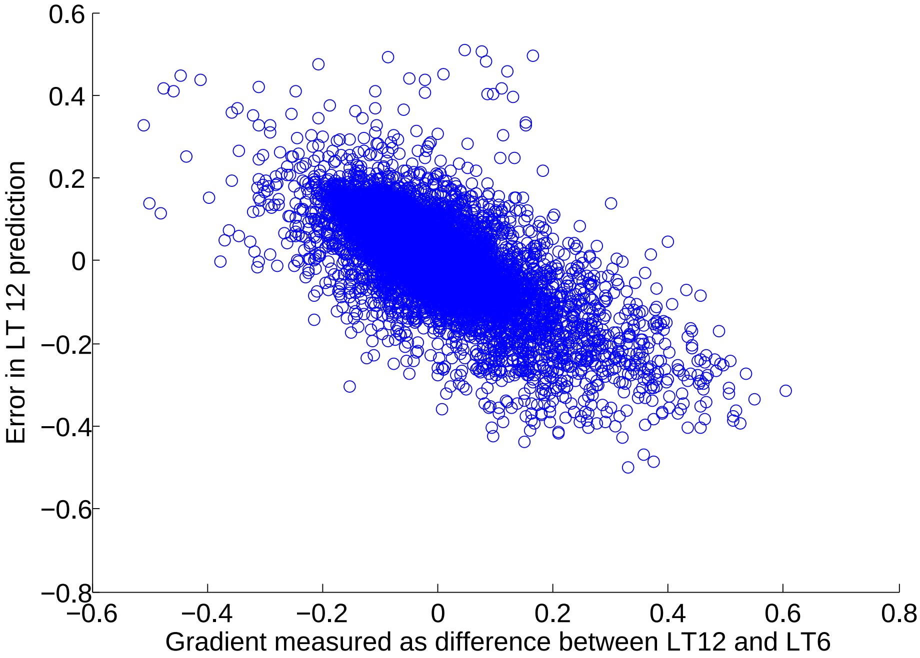

Figure 2: Scatter plot of error against gradient as measured by the difference in the LT12 and LT6 predictions.

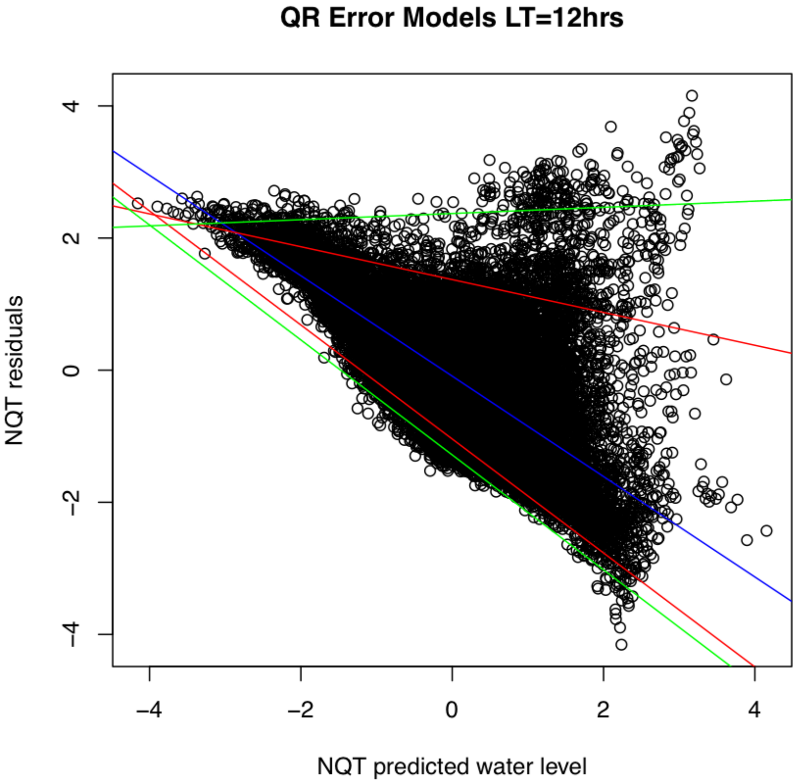

Figure 3: Quantile regression with a linear fit. Showing 1st, 5th, 50th, 95th and 99th percentiles.

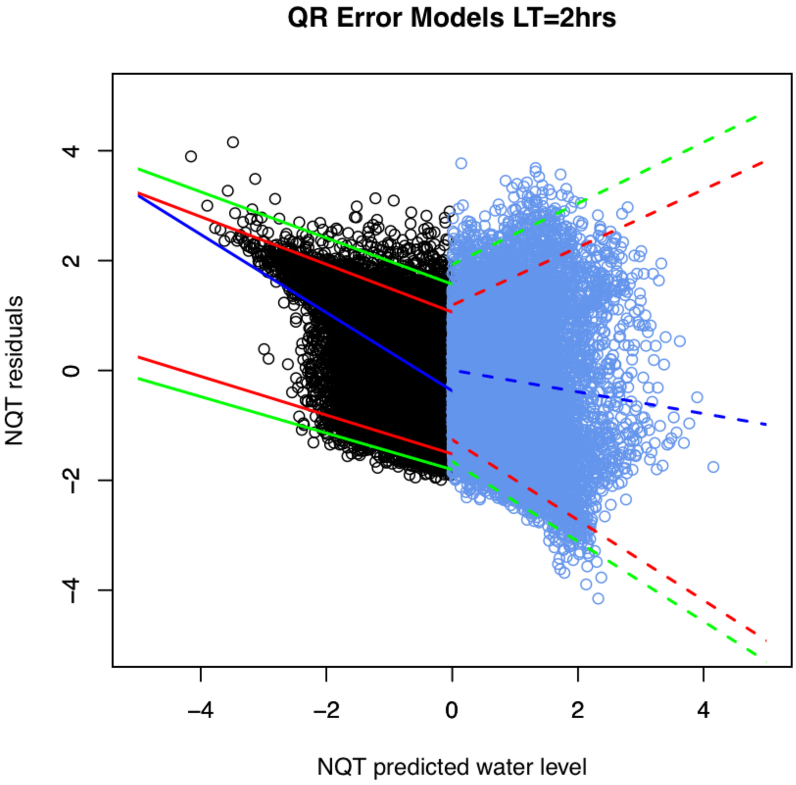

Figure 4: Quantile regression with two separate linear fits. This better captures the behaviour for high river levels.