Tweak

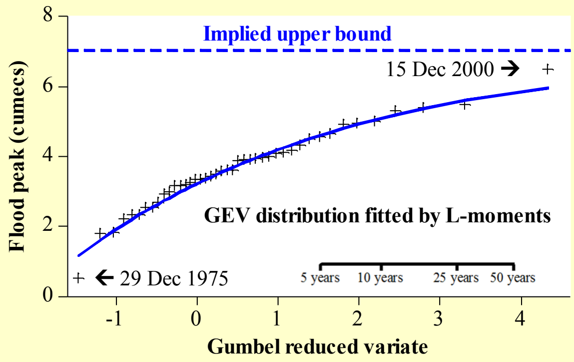
tweak (v): make fine adjustments to. bound (n): upper/lower limit to magnitude something can have. binding constraint (in mathematics): bound or other condition placed on parameter that ends up determining [often unsatisfactorily] solution reached in optimisation. nobbut (adv): merely.

Eee, vee, three: the rhyme that needs a reason

In single-site analyses, it is crucial to inspect the extreme value plot. Distributions fitted to annual maxima sometimes imply an upper bound. This may result from fitting a 3-parameter distribution to an annual maximum (AM) series that lacks any notably large flood. If the GEV distribution has been used, the outcome will be an EV3.

One scenario for such a *bounded above* distribution is the presence of *unusually small* AMs. Even if these are genuine values, it is wrong to allow them to drive the modelling of *floods*! The appropriate approach is to *censor* the analysis to exclude *non-flood years*. Although statisticians will favour a more formal approach, a tweak can suffice (see box➔).

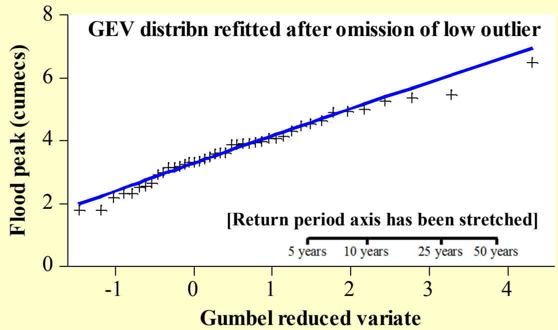
Another scenario is when the largest floods in the AM series all take about the same value. Such data series should always be scrutinised. Has an in-bank rating been applied at a stage when river flow at the gauging station is out of bank? Above a certain threshold, might some of the flood flow be bypassing the gauged section? If major floods are routinely attenuated by lake or flood-plain storage, can it be demonstrated (by physics not statistics!) that this will apply just as much to the very largest incoming floods?



Tweaking to remove a low outlier

Application of the GEV distribution to the Coln at Bibury has led to an unsatisfactory result. Fitted by L-moments, the model implies an upper bound of 7.0 cumecs: which is nobbut 8% bigger than the flood peak of 15 December 2000. The analysis has been distorted by the unusually small AM recorded in the drought-ridden water-year of 1975/76. The tail is wagging the dog!

The frequency analysis is repeated omitting non-flood years. Within the resulting extreme value plot, the frequency axis is *stretched* in compensation. Were three of 30 years omitted, a return period of 45 years (based on *flood years* only) would be taken to be equivalent to 50 real years.



In the Coln example, only one of 43 years had to be excluded, so the tweak to the frequency axis (e.g. a return period of 48.8 flood years has been stretched to 50 true years) is indiscernible. But the flood frequency curve itself is much changed!

Much binding in the marshland of parameter optimisation

Calibration of a rainfall-runoff model is always a challenge. If the model is based on broad concepts only, it is essential to test the calibrated model using additional observations. But is this quite so vital with a physically-based model? On the upside, one already has a reasonable idea ... e.g. from physical testing of soil properties or accumulated experience ... of the *range* within which particular parameters should lie. On the downside, physics-based models tend to have a lot of parameters, so there is extra scope for parameter interaction to confound the calibration. On balance, testing seems no less important for such a model.

Some modellers *constrain* parameters in the optimistic belief that this will help reach an optimum set of parameters. Although in principle a good policy, it is not uncommon for a parameter to bump up against the bound during optimisation. In essence, the parameter wants to take a value outside the stipulated range and is prevented from doing so only by the constraint, which is said to be *binding*. The behaviour is best taken as an appeal for greater parametric parsimony. The model should be simplified.

When evaluating a model, at least two broad standards are relevant. One is whether the model is consistent with the data. The other is whether the model is consistent with the 'real world'.

Kenneth A Bollen, *Structural equations with latent variables,* Wiley, 1989