Spatial rainfall distribution in flood modelling

Problem presented by

Adam Baylis

Environment Agency

Report author

Joanne A. Waller (University of Reading)

Executive Summary

When modelling flood risk the Environment Agency usually assume spatially uniform rainfall inputs and spatial uniformity in parameters that control the transformation of rainfall to runoff or river flow. These assumptions may result in over estimating the flood risk, leading to increased costs; conversely critical rainfall patterns that lead to the most significant flooding may be missed. The techniques of Monte Carlo estimation or continuous simulation may be able to alleviate these problems. However these methods are costly both in time and computational resource. It is undetermined whether the additional confidence in the flood risk is worth the extra cost.

The challenged posed to the study group was to identify an approach that gets some benefit of using the more detailed techniques, with only a minimal increase in modelling effort.

The approach taken was to design a catchment screening process allowing the resource heavy methods to be applied only in areas of high vulnerability and where the assumption of uniform rainfall distribution is a poor approximation. Areas of non-uniform rainfall can be determined using available high resolution (1km) radar rainfall data and by considering a combination of two indicators, the coefficient of variance and the skewness, of data over the spatial domain. As no data was available for testing, the approach was to look into the problem on a theoretical/conceptual basis only.

Once the areas of greatest interest have been determined these could, simply, be modelled using the present approach but with higher resolution. Alternatively Monte Carlo techniques could be used to run multiple scenarios run from different spatial rainfall distributions.

Version 1.0
March 14, 2016
iv+7 pages

Contributors

Elena Zanini (Lancaster University)
Paul Sharkey (Lancaster University)

Contents

1 Introduction and problem statement

(1.1) The Environment Agency aim to reduce the risk of flooding from all causes, including rivers, the sea, groundwater, reservoirs and surface water. To determine flood risk profiles the Environment Agency match knowledge of land topography with the modelled behaviour of the sea and river basins in different weather and tidal conditions. Unfortunately many assumptions are required in the flood modelling process and these can lead to uncertainty in the flood risk profiles.

(1.2) Two of the assumptions used when flood modelling are that rainfall inputs are spatially uniform as are the parameters that control the transformation of rainfall to runoff or river flow. The total volume of rain that falls (one possible way of determining the magnitude of the rainfall event) is assumed uniform in space and the distribution in time is represented using a hyetograph. Different hyetographs (see Figure 1 for an example) are chosen for different flood modelling purposes. For any rainfall event, the total volume of rainfall is kept constant, while the duration of the rainfall event is varied to find a 'critical storm duration' that maximises either risk or cost. A combination of this critical storm duration, the rainfall volume and the rainfall probability become the 'design event'. This design event is applied uniformly throughout a catchment area within the modelling process. To give a better understanding of the profile of risk, several (typically three) design events are used to represent different probabilities of rainfall. However, in reality rain falls unevenly across catchments and throughout time. The impact of assuming the uniformity of the design event is not captured in the risk profile.

(1.3) Possible solutions to understanding the impact of assuming uniform rainfall are to use either a Monte Carlo analysis or Continuous simulation. In a Monte Carlo analysis hundreds, or possibly thousands, of either historical or artificially generated rainfall events are selected from a distribution of events, from each event the flood risk is estimated. This gives a flood risk probability that covers a range of events that may occur from different rainfall distributions. With Continuous Simulation flood models are run using historical rainfall data series, any resulting flooding is recorded providing an overall estimate of flood risk. This method provides better information on the impact of the temporal as well as the spatil rainfall distribution. Both these methods could provide additional information on the impact of the spatial distribution of rainfall in flood modelling. In turn this could improve the estimate of flood risk.

(1.4) There is one main problem with both the suggested approaches; they are costly both in time and computational resource. Furthermore, detailed knowledge is required to design and run flood models using these techniques. Hence, the Environment Agency are looking for ways to gain the benefits of

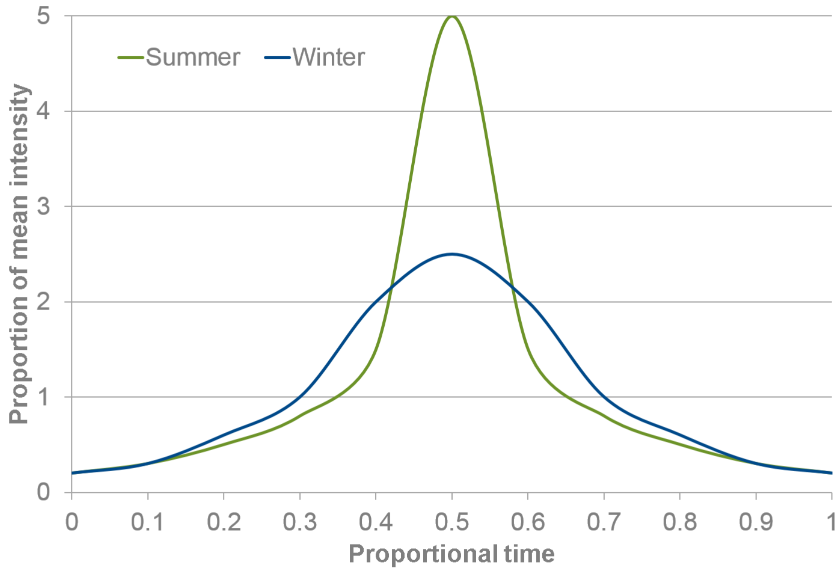

Figure 1: Two standard hyetographs showing the distribution of rainfall over time.

using these more sophisticated approaches but with only a minimal increase in the modelling resource and time.

2 The solution

(2.1) The impact of the spatial distribution of rainfall on flood modelling will likely differ dependent on the type of flooding to be modelled. The proposed solution considers how to understand the impact of the spatial distribution of rainfall when modelling surface water flooding. When modelling surface water flooding for the production of its national-scale flood maps the Environment Agency divide England and Wales into 5km x 5km 'tiles'. For each of these tiles the surface water flooding is modelled using the uniformly distributed rainfall distributed in time by a given hyetograph. The solution presented here attempts to provide a screening of these tiles to determine the 'tiles of interest'. The computational resources can then be focused on these tiles.

2.1 Tile screening

(2.2) The risk of flooding is a product of both the flood event itself and the vulnerability of the person, property or environment exposed to the event. The areas of greatest risk based on this information may be determined by considering:

- Likelihood of flooding;

- Land use;
- Topography;
- Hydrographic network.

However, in this case we are also concerned about the impact of the spatial distribution of rainfall. Therefore it is necessary to determine whether there could be plausible non-uniform rainfall scenarios that would be better moddeled using a more spohisticated approach. If, for any given tile, the rainfall that leads to flooding is uniform, then it may not be beneficial to alter the current modelling procedure. If, however, flooding in a particular tile is a result of non-homogeneous rainfall then the flood risk may be being over or underestimated. Therefore, these tiles affected by non-homogeneous rainfall must be modelled more accurately to better understand the flood risk. Determining the tiles of non-homogeneous rainfall is the main focus of this report and is detailed in 2.2.

(2.3) Once the areas of greatest vulnerability and non-homogeneous rainfall have been determined the computational resources can be allocated. Areas where the rainfall that results in flooding is homogeneous can be treated as they are currently, it may also be sensible not to waste resources on areas where, although the rainfall is heterogeneous, the area is of low vulnerability. The most resources should be focused on the areas with highest vulnerability to heterogeneous rainfall. A schematic diagram of how resources should be allocated is given in Figure 2, where tiles falling in areas with green shading would demand less computational resources than tiles that are described by the areas shaded red. The tiles that fall into the red area are the 'tiles of greatest interest' these could, simply, be modelled with higher resolution rainfall distributions, alternatively Monte Carlo techniques could be used further details are given in section 2.3.

2.2 Determining spatial distribution of rainfall in a tile

(2.4) The screening method relies on understanding, for each tile, if the rainfall distributions that result in flooding are uniform or not. One way to determine the nature of the rainfall distribution for a given tile is to use the existing archive of radar rainfall data, available at 1km resolution every 5 minutes. Two metrics will be calculated to assess the homogeneity of the rainfall over a tile, the skewness and the coefficient of variance.

(2.5) **Skewness** The skewness is a measure of the degree of asymmetry of a probability distribution about its mean. The sample skewness, g can be calculated using

$$g = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{[\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2]^{\frac{3}{2}}}, \tag{1}$$

where n is the number of samples, x_i is the i^{th}sample and $\bar{x}$ is the sample mean.

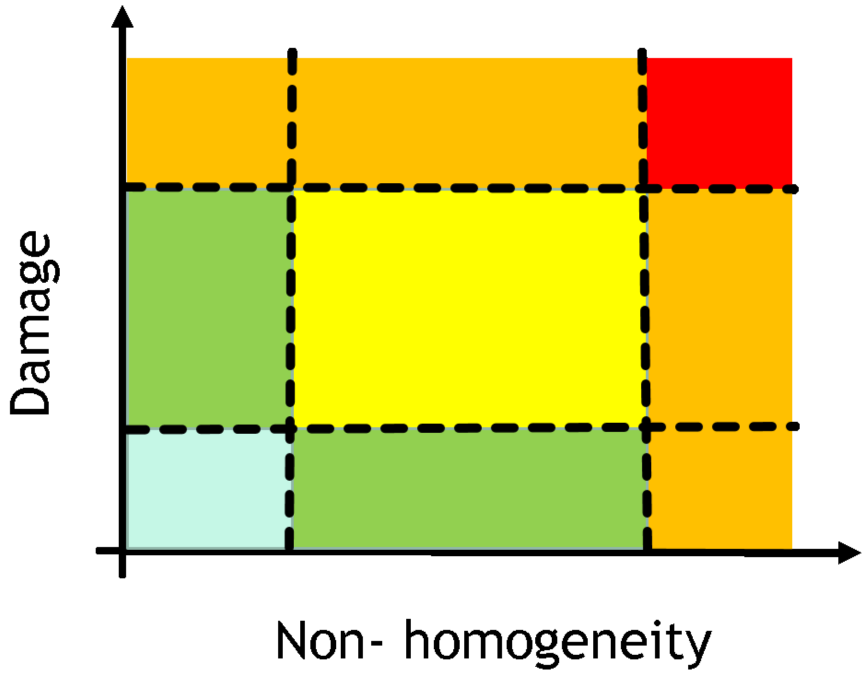

Figure 2: Areas to focus computational resources.

(2.6) **Coefficient of variation** The coefficient of variation is a standardised measure of the spread that describes the amount of variability in a data set relative to the mean of the data set. The population coefficient of variation, cv, can be calculated using,

$$cv = \frac{s}{\bar{x}} \tag{2}$$

the ratio of the sample standard deviation $s = \sqrt{(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2)}$ to the sample mean.

(2.7) We now describe the process of determining the homogeneity of the rainfall distribution using these metrics. The process for each tile is as follows:

1. Select only the radar rainfall that has led to flood events, (if this is not possible then instead consider either a set percentage of the largest rainfall events or all rainfall events that exceed a total rainfall threshold). It is the distribution of this rainfall that is of interest. It may be beneficial to include data that nearly caused a flood event, though how to determine this is unclear.
2. For each rainfall event leading to flooding:
    - From the beginning of the rainfall event group the radar data into hourly blocks ('windows') with 30 minute overlapping windows (all data will appear in two of the hour blocks, with the exception of the data at the beginning and end of the rainfall event). The overlapping windows are required so that no slow passing homo-

geneous band of rainfall is mis-interpreted as a shorter period of heterogeneous rainfall.

- For each window compute the skewness g_w and coefficient of variation cv_w.
- Any window with low skewness ($|g_w| < 1$) or low variability ($cv_w < v$, where v is a predetermined threshold determined by considering the value of cv_w for some given cases) can be flagged as homogeneous. Any window with $|g_w| > 1$ or $cv_w > v$ must be further investigated.
- For each window that needs further investigation aggregate the data from the previous hour window, the current hour window and the following hour window.
- For each aggregated data set compute the skewness, g_a, and coefficient of variation, cv_a.
- Any window with the agregated skewness less than the window skewness $|g_a| < |g_w|$ and the agregated coefficient of variance less than the window coefficient of variance $cv_a < cv_w$ (i.e. the initial window skewness and variation was caused by an incouming or outgoing large band of rainfall) can be flagged as homogeneous. Any window with $|g_a| > |g_w|$ and $cv_a > cv_w$can be flagged as heterogeneous.

3. For data from all events calculate the percentage of heterogeneous windows (100 x number of flagged windows/number of tested windows), if this percentage is large then the flooding in this tile is caused by non-homogeneous rainfall and it may be benificial to apply a more sophisticated moddeling technique for this tile.

2.3 Flood modelling for tiles of interest

(2.8) Once the tiles of greatest interest are determined, more resource can be put into modelling these tiles. We suggest five ways that these tiles could be better modelled. All suggestions include increasing the resolution of the rainfall distribution used. Two simple suggestions involve only increasing the rainfall resolution:

1. Initially it may be beneficial to split the tile equally into smaller regions using a more appropriate hyetograph for each region (i.e. carry out the same procedure as at present, but using a reduced tile size).
2. A more sophisticated, yet potentially cheaper, approach would be to split the tile into areas based on the rainfall distribution, e.g. As in Figure 3 areas of high (red), medium (yellow) and low (blue) rainfall, using an appropriate hyetograph for each region.

The remaining solutions follow a Monte Carlo approach and require running multiple scenarios. These include:

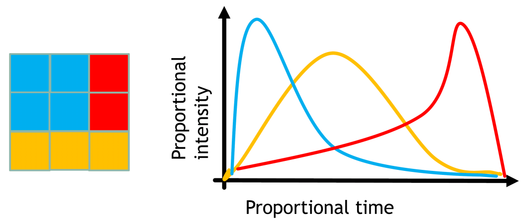

Figure 3: A tile split into three diferent areas based on volume of rainfall. Each tile would take its appropriate (colour matched) hyetograph.

3. Use a similar modelling procedure as suggestion 2, but run multiple scenarios from a number of different rainfall distribution patterns while keeping the same hyetograph. E.g. as in Figure 4

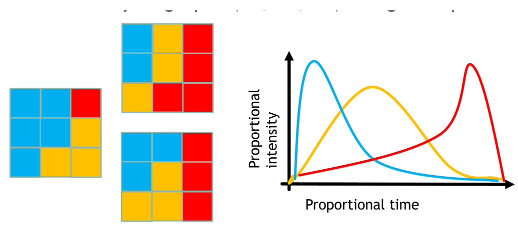

Figure 4: Differnt possible ways to split tiles into three diferent areas based on volume of rainfall. Each tile would take its appropriate (colour matched) hyetograph.

4. Use a similar modelling procedure as suggestion 2, use a single rainfall distribution pattern, but run multiple scenarios using different hyetographs. E.g. as in Figure 5

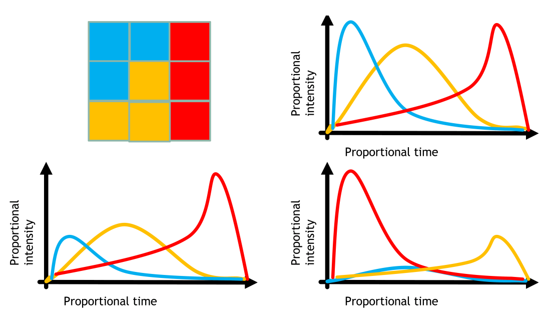

Figure 5: A tile split into three diferent areas based on volume of rainfall. Each tile would be moddled multiple times using diffent appropriate (colour matched) hyetographs.

5. Use a similar modelling procedure as suggestion 2, but run multiple scenarios from a number of different rainfall distribution patterns using a number of different hyetographs (a combination of approaches 3 and 4).

The final suggestion would be the most resource intensive, though should provide the largest range of possible flooding scenarios

3 Conclusions

(3.1) The impact of the spatial distribution of rainfall on flood modelling is not well understood. A number of techniques are available to investigate the impact that the rainfall distribution has, however these methods are costly both in time and computational resource. The challenge posed was to identify an approach that gets some benefit of these techniques, with only a minimal increase in modelling effort. It is likely that spatial rainfall patterns will have different impacts on different types of flooding. The approach developed here focused on the impact of rainfall distribution on surface water flooding. The suggested solution involves screening the catchment to determine the areas of greatest vulnerability and the areas with the most heterogeneous rainfall. Areas of non-homogeneous rainfall can be determined by calculating two metrics, the skewness and the coefficient of variation, for existing radar rainfall data. Once the screening has been determined, more computational resources can be focused on the tiles of interest. These area could be modelled with higher resolution rainfall distributions, alternatively Monte Carlo techniques with multiple scenarios of rainfall distributions and rainfall durations could be used.