Study on the effect of rainfall spatial variability on runoff modelling

Jun Zhang, Dawei Han, Yang Song and Qiang Dai

ABSTRACT

Rainfall spatial variability was assessed to explore its influence on runoff modelling. Image size, coefficient of variation (Cv) and Moran's I were chosen to assess for rainfall spatial variability. The smaller the image size after compression, the less complex is the rainfall spatial variability. The results showed that due to the drawing procedure and varied compression methods, a large uncertainty exists for using image size to describe rainfall spatial variability. Cv quantifies the variability between different rainfall values without considering rainfall spatial distribution and Moran's I describes the spatial autocorrelation between gauges rather than the values. As both rainfall values and spatial distribution have an influence on runoff modelling, the combination of Cv and Moran's I was further explored. The results showed that the combination of Cv and Moran's I is reliable to describe rainfall spatial variability. Furthermore, with the increase of rainfall spatial variability, the hydrological model performance decreases. Moreover, it is difficult for a lumped model to cope with rainfall events assigned with complex rainfall spatial variability since spatial information is not taken into consideration (i.e. the VIC model used in this study). Therefore, it is recommended to apply distributed models that can deal with more spatial input information.

Key words | Cv, image size, Moran's I, rainfall spatial variability, runoff modelling

Jun Zhang (corresponding author)
Dawei Han
Yang Song
Qiang Dai
Department of Civil Engineering,
University of Bristol,
Bristol BS8 1TR,
UK
E-mail: *catherine.zjun@gmail.com*

Qiang Dai
Key Laboratory of VGE of Ministry of Education,
Nanjing Normal University,
Nanjing 210000,
China

INTRODUCTION

Undoubtedly, rainfall is one of the most important inputs for runoff modelling. However, due to the climate condition and catchment morphology, rainfall is unlikely to be evenly distributed over the whole catchment, which is known as rainfall spatial variability. Therefore, it is of great significance to retrieve proper input data for hydrological models from recorded data.

In order to determine the relationship between rainfall spatial variability and runoff modelling, assessing the rainfall spatial variability properly is of significance. There are several currently used indicators to assess rainfall spatial variability. Coefficient of variation (Cv) is one of the most commonly used indicators in hydrology, it is easy to apply and is able to describe a certain level of rainfall spatial variability (Pedersen *et al.* 2010). Moreover, some researchers tried to investigate the inter-gauge correlations (Zhang

et al. 2012a) and spatial deviation index (SDI) (Segond *et al.* 2007) to describe the rainfall spatial variability. However, only investigating the relationship between limited rainfall gauges is not capable of describing the rainfall spatial distribution over the whole catchment. The semi-variogram is one index describing the decorrelation distance of rain gauges (Bacchi & Kottegoda 1995), however, it is possible that the distance is larger than the catchment scale, which hinders its application in small catchments. Although there are several indicators being used currently, there is not a generally accepted indicator that can systematically describe the rainfall spatial variability so far.

Studies on the effect of rainfall spatial variability on runoff generation and modelling have been carried out in the past decades. With varied rainfall magnitudes and the routing paths in multiple locations over the catchment, as a

result, runoff is supposed to be unevenly distributed spatially. Some researchers have found that model performance is significantly affected by rainfall spatial variability. It has been tested that the increase of rainfall spatial variability enhanced the runoff variability (Wood *et al.* 1988). It was found that a large uncertainty in estimated model parameters can be expected if detailed variation in the input rainfall is not taken into account (Chaubey *et al.* 1999). Moreover, the estimated peak flow and runoff volume were affected by spatially distributed rainfall (Arnaud *et al.* 2002). The spatial rainfall resolution for runoff estimation has been investigated, indicating that the model performance decreases with the increase of rainfall spatial variability (Paudel *et al.* 2011; Zhang *et al.* 2012b). However, a number of researchers argued that rainfall spatial variability could be mitigated due to the damping effect of the catchment processes. The spatial characteristic scale of runoff was found to decrease compared to catchment rainfall spatial variability resulting from superposition of small-scale variability of catchment (Skøien 2003). Obled *et al.* (1994) determined that in a rural medium-sized catchment, rainfall spatial variability was not significant enough to overcome the dampening of the catchment. It was not always true that higher rainfall spatial resolution could improve the model output (Bell & Moore 2000). A slight improvement on runoff modelling was experienced with the increase of rainfall input data. More researchers indicated that sometimes average rainfall would be enough for the catchment modelling because of the large damping behaviour of the basin.

It was demonstrated that convective storms would have greater runoff variability than stratiform rainfall (Bell & Moore 2000). Not only the rainfall magnitude but also the position of the main rainfall cell could affect the runoff generation (Syed *et al.* 2003). The catchment antecedent soil water condition was demonstrated as a crucial factor for wet conditions and fewer gauges were required, while for dry conditions the role of rainfall spatial distribution was more important (Shah *et al.* 1996). It was found that for catchments with the rainfall spatial variability scale larger than the hillslope scale, flood response was more sensitive to the average rainfall. Moreover, for larger catchments, the spatial distribution of rainfall highly affected the runoff production because of the heterogeneous transport paths (Nicótina *et al.* 2008).

In this study, several indicators were adopted to assess the rainfall spatial variability and further investigation was carried out on their feasibility with a hydrological model. Image size is a novel indicator to assess the rainfall spatial variability by quantifying the size of the rainfall map. Moran's I is an indicator to describe the rainfall distribution, which has been applied to varied research areas to describe spatial autocorrelation (Moran 1950). Cv is also calculated to explore the optimal indicator for variability assessment. This study was performed in the Brue catchment in the UK with a dense rainfall gauge network of 49 gauges. The relationship between rainfall spatial complexity and runoff modelling performance was then explored to further identify the reliability of the indicators.

METHODOLOGY

Spatial variability indicators

To measure the spatial variability of a rainfall field, several indicators were explored in the study, i.e. Cv, image size and Moran's I. Details of these indicators are described in the following.

Cv

The first chosen approach is the spatial coefficient of variance (Cv), which is one of the most frequently used indices in hydrology (Pedersen *et al.* 2010). Cv calculates the variation of relevant rainfall records, defined as:

$$Cv = \frac{\sqrt{\sum_{i=1}^{n} \left(P_i - \bar{P}\right)^2}}{\bar{P}} \qquad (1)$$

in which P_i is the rainfall value at the ith gauge, mm; $\bar{P}$ is the average rainfall of all gauges, mm; n is the number of gauges. The increase of Cv indicates the increase of rainfall spatial variability.

Image size

Image size is a newly introduced indicator to be tested. When drawing a figure, if uncompressed, each pixel is

assigned with one value, which occupies a unit storage space. Therefore, for figures with the same number of pixels, the storage space is exactly the same. However, to save storage space, the figures are usually compressed with certain principles. One of the most widely used compressed image formats is jpeg format. To compress a figure, several procedures are applied, including colour space transformation, down-sampling, block splitting, discrete cosine transform, quantization, and entropy coding (Sayood 2000). Among all the encoding procedures, entropy coding is one of the most important ones to save storage space, in which neighbourhood pixels with similar values are grouped together. Instead of saving similar pixels of the same group individually, a simpler message with the value, locations and number of pixels of the group is saved, which occupies much less storage (Wiegand & Schwarz 2011). In this case, with more neighbourhood pixels with similar values, the image is supposed to be compressed on a greater rate, which occupies less storage. When compressing the images with the same principle, the image size is determined by the pixel value variability of the whole image area (Watson *et al.* 1997). In other words, with the same-size original image, the smaller the compressed image size is, more neighbourhood pixels are grouped due to similar values. Therefore, the larger the compressed image size is, the larger spatial variability of the pixels.

When applying this theory into rainfall spatial variability, a cumulative rainfall contour map is generated for each rainfall event without the catchment boundary and all the lines. To draw the map, the catchment is decomposed into multiple cells, the value of each cell is obtained using the kriging method with known gauge data. After that, a rainfall map is drawn based on the cell values with the same plotting scale and colour map. For those images created by rainfall values, each pixel is assigned with a particular rainfall value. Due to the variable rainfall values, the information carried by all pixels will be vastly different. Therefore, when compressing the rainfall contour map, if neighbourhood pixels are assigned with similar rainfall values, they would be grouped to reduce the storage. In other words, the image size of the contour map after compression will be smaller, indicating less spatial variability. By identifying the image size of each rainfall contour map derived from the same compression rule, the complexity of

the image is determined, which is also identified as the rainfall spatial variability.

Moran's I

Spatial autocorrelation is the co-variation of properties within geographic space: characteristics at proximal locations appear to be correlated, either positively or negatively. Moran (1950) proposed a statistic (Moran's I) to assess the spatial autocorrelation between spatial locations. Moran's I is defined as:

$$I = \frac{N}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(P_i - \bar{P})(P_j - \bar{P})}{\sum_{i=1}^{n}(P_i - \bar{P})^2} \tag{2}$$

in which P_i is the rainfall at the ith gauge, mm; P_j is the rainfall at the jth gauge, mm; $\bar{P}$ is the average rainfall of all gauges, mm; w_{ij} is the spatial weight between ith and jth gauge. There are several ways to calculate w_{ij}. The original approach is $w_{ij} = 1$ if ith and jth are adjacent, and $w_{ij} = 0$ otherwise. Another way to define w_{ij} is the inverse distance method, which is defined as:

$$w_{ij} = r_{ij}^{-b} \tag{3}$$

in which r_{ij} is the distance between the ith gauge and jth gauge, m; b is a distance parameter ($b = 1$ in this paper). Considering all rainfall gauges are correlated in this study area, using the first method is possible to ignore the correlated information. Therefore, w_{ij} is calculated with the inverse distance method.

Moran's I is a test statistic to assess the spatial autocorrelation, such as rainfall gauges in this paper, varying from −1 to 1. A zero value indicates a random spatial pattern. Positive values indicate positive spatial autocorrelation, which means that gauge values are correlated, and vice versa. When positive, Moran's I closing to 1 indicates a strong level of positive spatial autocorrelation, meaning that high values are clustered close to high values and low values are clustered close to low values.

The F-test was adopted to test if there is significant difference between different groups. In the hypothesis, if a p-value is lower than 0.05, then the two groups are supposed to be significantly different (Lomax & Hahs-Vaughn 2013).

Hydrological model

Since this study is about the impact of rainfall spatial variability on the performance of lumped hydrological models, so a lumped model is more suitable than distributed models. The variable infiltration capacity (VIC) model was first introduced by Wood *et al.* (1992), and extended to the widely used VIC-2 L (two-layer) and VIC-3 L (three-layer) by Liang *et al.* (1994). The structure of VIC-3 L is used in the study. By introducing VIC in different areas of the catchment allows the heterogeneity of fast runoff production in the model. There are five parameters to be calibrated in the VIC-3 L model, including β (power law exponent for soil evaporation), b (a shape parameter controlling the form of infiltration capacity distribution), i_m (maximum infiltration storage capacity for the area), k_b (maximum base flow rate at saturation), n_b (base flow exponent). The model was calibrated using the Levenberg–Marquardt algorithm automatically.

Since it is a lumped model, an average rainfall P_{in} calculated by the Thiessen polygon method is used for the whole catchment when adopting multiple rain gauges, shown as:

$$P_{in} = \sum_{i=1}^{n} \alpha_i P_i \tag{4}$$

in which α_i is the Thiessen coefficient associated with the ith gauge.

The model is calibrated with hourly data in 1995 with 49 gauges then applied with the rainfall input derived from different numbers of rain gauges for the same year. For the events analysed later, the modelled runoff is picked from the whole year modelling with the corresponding rainfall input data, instead of simulating runoff for each event separately. The model performance is evaluated in terms of Nash–Sutcliffe efficiency (NSE) as:

$$NSE = 1 - \frac{\sum_{i=1}^{m} \left(Q_{sim,i} - Q_{obs,i}\right)^2}{\sum_{i=1}^{m} \left(Q_{obs,i} - \overline{Q_{obs}}\right)^2} \tag{5}$$

in which $Q_{sim,i}$ is the simulated runoff at time i, m³/s; $Q_{obs,i}$ is the observed runoff at time i, m³/s; $\overline{Q_{obs}}$ is the mean observed runoff over the modelling span, m³/s; m is the total time intervals.

DATA AND STUDY SITE

The Brue catchment is located in the southwest of England, as shown in Figure 1, draining an area of 132 km² to its river gauge at Lovington. There is a specially designed HYREX experimental dense rainfall network with 49 tipping
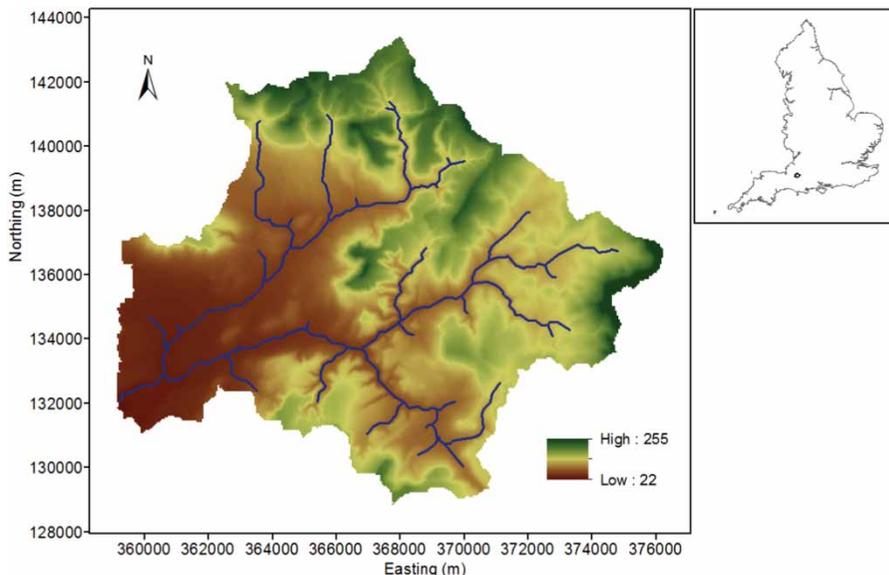

Figure 1 │ Brue catchment and its location in the UK.

bucket rain gauges distributed in the whole catchment, as shown in Figure 2 (Moore *et al.* 2000). The elevation of the catchment varies from 255 m in the upstream to 22 m in the downstream.

The contour map of total rainfall in 1995 is plotted in Figure 2 ranging from 748 to 957 mm, as well as the distribution of rain gauges (black dots). In general, rainfall decreased from the east to the west, which is also identified as from the upstream to the downstream. Hourly data of rain gauge, runoff gauge and climate data are available from 1994 to 1999 for the catchment. The data in 1995 were chosen because of the complete datasets and fewer data errors. Due to problems such as blocking and damage of rainfall measurement instruments, a data quality check was performed before analysis. A cumulative hyetograph was used to determine faulty data, which has been proved to be a valid method (Wood *et al.* 2000). When data from one gauge was found to be faulty, a kriging interpolated rainfall replacement was used to generate for the faulty gauge.

RESULTS AND DISCUSSION

Rainfall spatial variability with a single indicator

In order to determine the relationship between rainfall spatial variability with model performance, three indicators, i.e. Cv, image size, and Moran's I, were calculated for 236 events in 1995. After that, we explored if there was any correlation between the rainfall spatial variability with model performance.
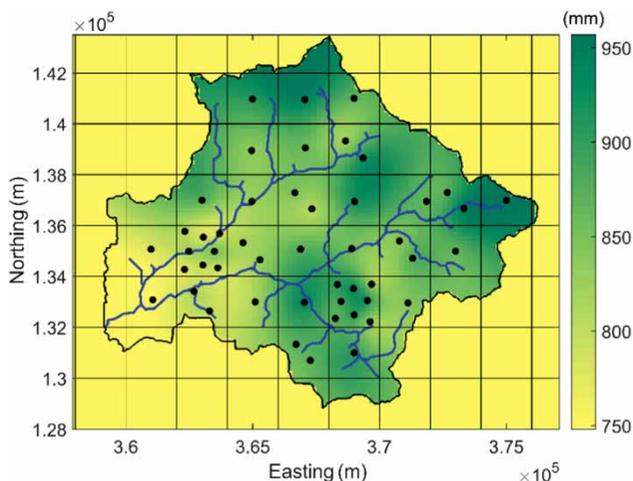
Image size

Multiple pixels were divided for the whole catchment to produce a rainfall contour map. By comparing running time, computation load and carried information, to decompose the catchment into ten thousand pixels was proved to be the optimal choice. All rainfall contour maps were generated with the scale 0–80 mm, and compressed to jpeg format using the same compression principle. The image size varied from 17.95 to 32.66 kb. To simplify the analysis, three groups with different rainfall spatial variability assessed by image size were divided. The events with an image size of 17.95–22.87 kb were treated as simple events, 22.88–27.78 kb as medium events, and 27.79–32.66 kb as complex events. NSE in different groups are plotted in Figure 3.

When applying the F-test, the *p*-value between the simple and medium group is 7.42×10^{-5}, between the simple and complex groups is 5.24×10^{-7}, and between the medium and complex groups is 1.06×10^{-3}. All values are less than 0.05, meaning that the groups are significantly different. According to the result, the average NSE value of three groups varies vastly. Unexpectedly, the NSE of simple events is smaller than the NSE of complex events. Moreover, there are quite large overlapping areas between medium and complex events. Also, when we tried different
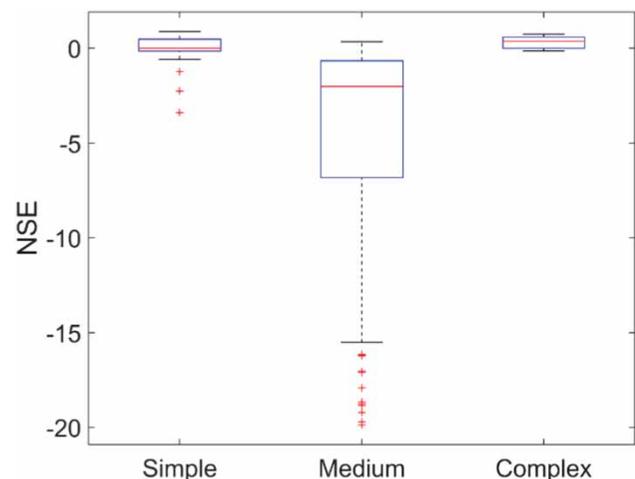

Figure 2 │ Total rainfall of the Brue catchment in 1995.

Figure 3 │ NSE in groups defined by image size.

scales and compressed methods for rainfall contour map drawing, the results varied vastly. Therefore, it is difficult to define an optimal scale and a compression method for rainfall contour map generation, and the corresponding results are unstable.

Moran's I

As mentioned above under 'Spatial variability indicators – Moran's I', Moran's I reflects the level of spatial autocorrelation of a rainfall field. When greater than 0, the complexity of rainfall event increases with the decrease of Moran's I. For 236 events, Moran's I varied from 0.003 to 0.292. Three groups are divided for different rainfall spatial variability where Moran's I less than 0.1 is the complex group, higher than 0.2 is the simple group and the others are in the medium group. When applying the F-test, the *p*-value between the simple and medium groups is 3.25×10^{-3}, between the simple and complex group is 7.24×10^{-10}, and between the medium and complex group is 2.51×10^{-4}. All values are less than 0.05, meaning that the groups are significantly different. As shown in Figure 4, there is a slightly decreasing trend from the simple group to the complex group. However, the difference between the medium and complex groups is hard to find as the median values and boundary lines are close.

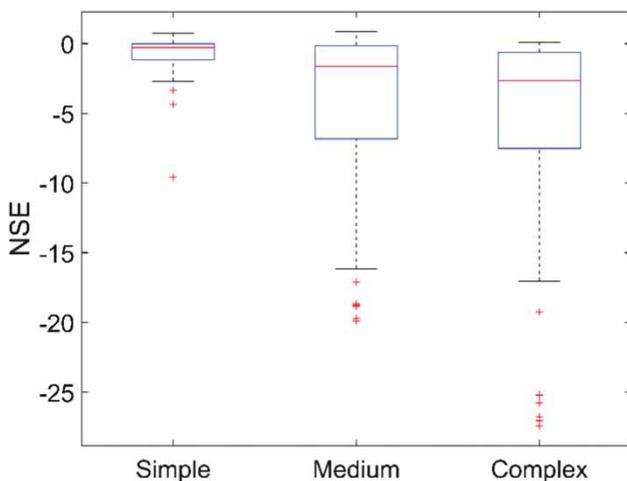There is a main shortcoming of Moran's I that is concerned more about distribution than values. Moran's I remains the same when the distribution is the same, even with different values. For example, for the same chess distribution, no matter what the values are, Moran's I remains the same. However, not only rainfall distribution but also rainfall values have an influence on runoff modelling. Since it is difficult to find two events with exactly the same rainfall spatial distribution, an example is illustrated here. Assume a simple distribution with half higher values in the upstream and half lower values in the downstream. For the first one, the rainfall in the upstream is 30 mm and in the lower stream it is 0 mm, while the other one is the opposite, i.e. rainfall in the upstream is 0 mm and 30 mm in the lower stream. For both events, the average rainfall is the same, which results in the same simulation from a lumped model. However, a late peak and longer recession time would be expected in the first event as the rainfall concentrates farther to the outlet, which is opposite to the second event. Therefore, even with the same spatial distribution, rainfall values have an influence on runoff modelling, which shows that there is information ignored by Moran's I.

Cv

Cv of 236 events were calculated, ranging from 0.064 to 7.00. The larger the Cv is, the more complex the rainfall event spatial variability. Three groups of events, i.e. simple (Cv from 0.064 to 2.38), medium (Cv from 2.39 to 4.69), and complex (Cv from 4.70 to 7.00) events, are divided. When applying the F-test, a *p*-value between the simple and medium groups is 9.21×10^{-7}, between the simple and complex groups is 3.57×10^{-9}, and between the medium and complex groups is 1.21×10^{-3}. The model performance of the three groups is plotted in Figure 5. The average NSE decreases from the simple to the complex group. The events in the simple group perform best. However, there is a large overlap between the medium and complex group. Moreover, part of the blue box of complex events is even better than that of medium events, which is not as expected.

When looking into the principle for calculating Cv, it is easy to find that Cv only considers the variance among all gauge values. Therefore, no matter how the rainfall distributes in the catchment, Cv is not capable of capturing the difference caused by spatial distribution. As shown in Figure 6(a) and 6(b), two rain gauges are assigned with 0.2 mm for both events but with different distributions. Cv
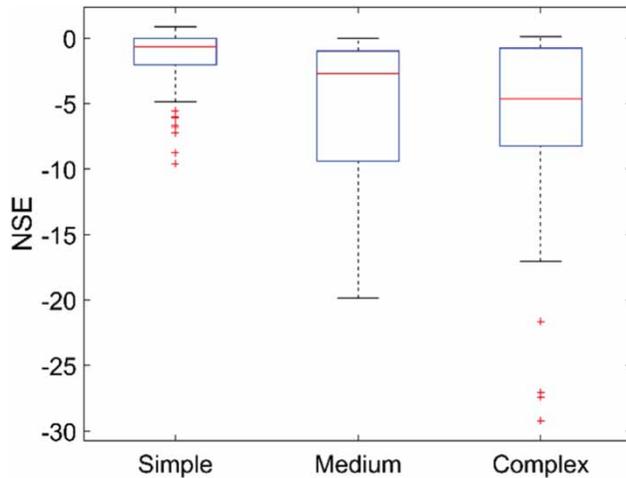

Figure 4 | NSE in groups defined by Moran's I.

Figure 5 │ NSE in groups defined by Cv.

equals 4.89 for both events, which is assumed to be in the same level of spatial variability. However, the NSE of Figure 6(a) is 0.11, and of Figure 6(b) is 0.27.

Assuming the same initial conditions, there are two events with 10 gauges assigned with rainfall of 30 mm, and other gauges are assigned with 0 mm. For the first event, 10 gauges are all located in the upstream, while for the other event 10 gauges are all located in the downstream of the catchment. For a lumped model, the modelled hydrograph is the same with the same average rainfall. However, for the first event, there is likely to be a delay for rainfall travelling from the upstream to the outlet with a longer time in the recession period. For the second event, since rainfall locates close to the outlet, there is likely to be a sharp peak at the beginning of the event while there is less runoff in the recession period. In this case, the real hydrograph is different for the two events.

When calculating NSE with the same modelled hydrograph, the computed model performance for two events are different as well, indicating that rainfall distribution has an effect on runoff modelling performance.

Model performance with combined indicators

Event-based rainfall spatial variability and model performance

As plotted in Figure 7, Cv ranges from 0.064 to 7 and Moran's I ranges from 0.003 to 0.292 for all events. There is an ambiguous decrease trend of Cv with the increase of Moran's I. According to the previous analysis, Cv describes variability between values, while Moran's I describes spatial variability. Therefore, it is possible to describe both value variability and spatial variability with a combination of two indicators (Zhang & Han 2017).

Larger Cv indicates larger values spatial variability and smaller Moran's I illustrates larger spatial variability. Therefore, events in the upper-left part of Figure 7 are supposed to be assigned with the largest rainfall value and spatial variability, categorised in the complex group. Meanwhile, events in the lower-right part of Figure 7 are supposed to be assigned with the smallest rainfall value and spatial variability, categorised in the simple group. Other events are either assigned with simpler value variation or simpler spatial distribution compared to the complex group, which is recognised as the medium group. To specify, the events with Cv larger than 4 and Moran's I lower than 0.1 were in the complex group; the events with Cv smaller than 2 and Moran's I larger than 0.2 were assigned into the simple group; all the other events were put into the medium group.
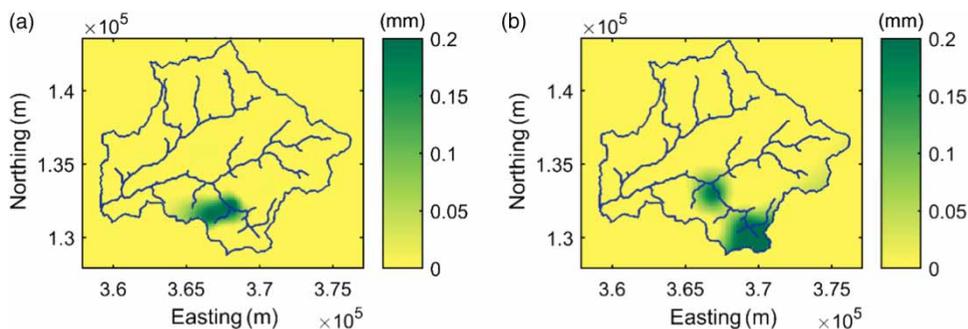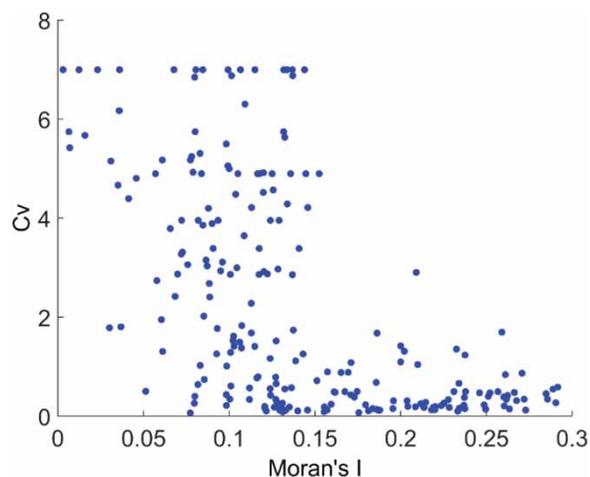

Figure 6 │ Two events with the same Cv.

Figure 7 │ Cv and Moran's I of event in 1995.

Figure 8 │ NSE in groups defined by Cv and Moran's I.

To validate the reliability of the combination of two indicators, the NSE in three groups is plotted in Figure 8. According to the result, it is clear that with the increase of rainfall spatial variability, NSE decreases instead. In general, events in the simple group perform better than the other groups. The average value in the complex group is significantly smaller than the value in the medium and simple groups. Compared to the result in Figure 5, there is less overlapping area between the medium and complex groups. In general, the performance in the medium group is better than the complex group. Moreover, there are more extreme low values in the complex group than the other two groups.

When applying the F-test, the *p*-value between the simple and medium group is 3.61×10^{-14}, between the simple and complex group is 1.07×10^{-15}, and between the medium and complex group is 0.012. All *p*-values are lower than 0.05, indicating that the three groups are significantly different to each other. In other words, groups defined by the combination of Cv and Moran's I have a significantly different influence on runoff modelling performance. Therefore, it is reasonable to use the combination of Cv and Moran's I to describe the rainfall spatial variability.

Poorly performed events

As displayed in Figure 8, there are some events with negative NSE even based on 49 rain gauges. In order to investigate the cause of those poor model performances, the hydrograph
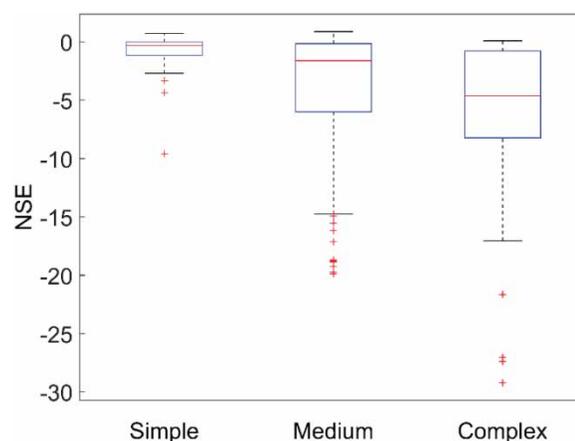
and rainfall distributed were plotted for all the events, and two selected events are shown in Figure 9. Figure 9(a) and 9(c) shows the hydrographs of the two events, Figure 9(b) and Figure 9(d) are the rainfall distribution maps respectively, Figure 9(a) and Figure 9(b) are from the same event, while Figure 9(c) and Figure 9(d) belong to the other one.

The two events displayed here are two typical ones with rainfall distribution. For the first event (as shown in Figure 9(b)), rainfall in the upstream is significantly larger than rainfall in the downstream. Since there is a lag time for the upstream rainfall to reach the flow gauge at the outlet, the peak time is supposed to be later than the modelled peak time with a lumped model. This is because when using a lumped model, rainfall is treated as evenly distributed over the whole catchment. In this case, rainfall input for the model in the downstream is larger than the real rainfall. Therefore, the modelled peak time is earlier than it is supposed to be, as shown in Figure 9(a).

On the other hand, the situation is totally opposite for the second event. In this event, rainfall in the downstream is larger than rainfall in the upstream, as indicated in Figure 9(d). Therefore, the modelled peak time is later than the observed since the downstream rainfall input for the model is lower than the real rainfall, which is as expected in Figure 9(c).

Another reason is worth mentioning for the low NSEs. The runoff simulation was performed for the whole year and the event-based hydrographs were extracted from the whole-year simulation rather than simulated separately. As
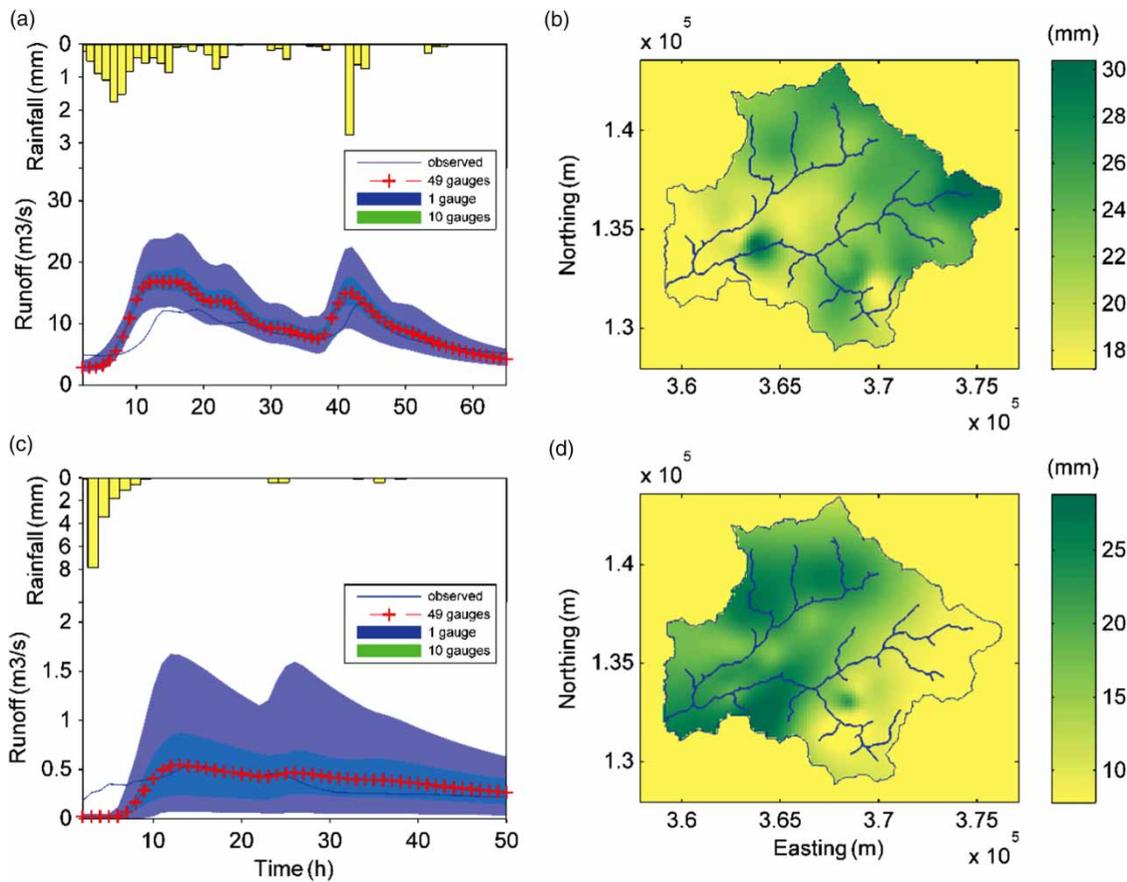
Figure 9 │ Poorly performed events.

a result, the water balance was accounted for over a long period, i.e. one year, rather than individual events, which result in an uneven water balance in some events. Therefore, for those events without a water balance, the assessed performance by NSE was low.

When using NSE for model assessment, the modelled values are compared with the measured values at the same time. Therefore, if there is a mis-estimation of peak time, even with the same peak volume, it is likely to be treated as poorly performed.

DISCUSSION

Three indicators, i.e. image size, Cv and Moran's I, are tested to describe rainfall spatial variability. The main theory of image size is determining the image complexity by grouping similar neighbourhood values. With more nearby pixels

assigned with similar values, image size is smaller. However, when drawing a rainfall contour map, colour type and scale have a significant effect on the image size. Moreover, there are several methods to generate a compressed image from the original image. In the methods, one of the most crucial procedures is to define the grouping threshold. Different choices possibly lead to varied results. As a result, how to define an optimal procedure and parameters in compressing images is not easy to apply and results in large uncertainty due to the varied compressing algorithms.

Cv is one of the most widely used indicators to describe rainfall spatial variability. However, it is found that Cv only takes into account the variation of rainfall values in all gauges, but does not consider the distribution of rainfall events. For runoff modelling, where the rainfall core is located is important, especially for events with complex rainfall spatial variability. When the rainfall core is located in the upstream of the catchment, there is supposed to be a delay of peak runoff

occurrence. When the rainfall core is located in the downstream of the catchment, the peak would appear early while the recession period would be longer. Therefore, for runoff modelling, only considering the values of different gauges is not sufficient to describe the rainfall spatial variability.

Moran's I is an indicator that describes the spatial autocorrelation in the study area. When Moran's I is positive, the larger Moran's I is and the more uniform the rainfall event. However, with the same distribution of a rainfall event, no matter how the rainfall values varied, Moran's I remains the same. In other words, Moran's I only considers the rainfall spatial distribution rather than the rainfall values.

For rainfall events with varied values but uniform distribution, it is likely to be over-estimated by Cv while under-estimated by Moran's I. On the other hand, for rainfall events with similar values but dispersed distribution, it is likely to be under-estimated by Cv but over-estimated by Moran's I.

Due to the limitation of Cv and Moran's I, a combination of Cv and Moran's I is under consideration. In this case, both value variation and spatial distribution variation are taken into consideration. Three groups with different rainfall spatial variability are analysed. The result shows that it is reasonable to define rainfall spatial variability with the combination of Cv and Moran's I. It is considered most complex when both Cv and Moran's I are complex, while simplest when both Cv and Moran's I are simple. For events with complex Cv and simple Moran's I, rainfall cores are uniformly located but with varied values. For events with simple Cv and complex Moran's I, rainfall cores are located in different areas but with similar values. With events with different kinds of spatial variability, it is supposed to lead to different model performance. For rainfall events with simple Cv and complex Moran's I, the peak time is supposed to be affected more significantly than the peak volume because of the damping effect of the catchment. For rainfall events with complex Cv and simple Moran's I, runoff volume is likely to be affected more.

When applying rainfall information into a lumped model, it is likely to ignore some spatial information. For events with different spatial distribution even with the same average values, both peak time and peak volume is possibly different. Therefore, for complex rainfall events, a distributed hydrological model is worthy of being taken into account. Moreover, for events with different spatial variability, required input data is supposed to be different as well. After assessing the rainfall spatial variability, an optimal input data carrying the most crucial information is worthy studying for the model, especially distributed models.

CONCLUSIONS

Rainfall spatial variability has a significant influence on runoff generation and runoff modelling. However, there is no general accepted indicator to describe rainfall spatial variability so far. Image size, Cv and Moran's I are tested in the research. Due to the uncertainty of drawing a rainfall contour map and compression method, it is difficult to apply image size as an assessment statistic as a determined method. Cv is one of the most widely used indicators, considering the variation between different gauges. Moran's I is an indicator describing the spatial distribution. However, there exists information ignored for both Cv and Moran's I.

The results show that with the combination of Cv and Moran's I, both rainfall values and distribution are able to be taken into account and proved to be reliable. Furthermore, model performance decreases with the increase of rainfall spatial variability. The results also indicate the limitations using lumped model for events with complex rainfall spatial variability. Therefore, a semi-distributed and fully-distributed model is under consideration for further work.

ACKNOWLEDGEMENTS

REFERENCES

Arnaud, P., Bouvier, C., Cisneros, L. & Dominguez, R. 2002 Influence of rainfall spatial variability on flood prediction. *J. Hydrol.* **260** (1), 216–230.

Bacchi, B. & Kottegoda, N. T. 1995 Identification and calibration of spatial correlation patterns of rainfall. *J. Hydrol.* **165** (1–4), 311–348.

Bell, V. A. & Moore, R. J. 2000 The sensitivity of catchment runoff models to rainfall data at different spatial scales. *Hydrol. Earth Syst. Sci.* **4** (4), 653–667.

Chaubey, I., Haan, C. T., Grunwald, S. & Salisbury, J. M. 1999 Uncertainty in the model parameters due to spatial variability of rainfall. *J. Hydrol.* **220** (1), 48–61.

Liang, X., Lettenmaier, D. P., Wood, E. F. & Burges, S. J. 1994 A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *J. Geophys. Res.* **99** (D7), 14415–14428.

Lomax, R. G. & Hahs-Vaughn, D. L. 2013 *Statistical Concepts: A Second Course*. Routledge, Abingdon.

Moore, R. J., Jones, D. A., Cox, D. R. & Isham, V. S. 2000 Design of the HYREX raingauge network. *Hydrol. Earth Syst. Sci.* **4** (4), 523–530.

Moran, P. A. P. 1950 Notes on continuous stochastic phenomena. *Biometrika* **37** (1/2), 17–23.

Nicótina, L., Alessi Celegon, E., Rinaldo, A. & Marani, M. 2008 On the impact of rainfall patterns on the hydrologic response. *Water Resour. Res.* **44** (12), WR006654.

Obled, C., Wendling, J. & Beven, K. J. 1994 The sensitivity of hydrological models to spatial rainfall patterns: an evaluation using observed data. *J. Hydrol.* **159** (1–4), 305–333.

Paudel, M., Nelson, E. J., Downer, C. W. & Hotchkiss, R. 2011 Comparing the capability of distributed and lumped hydrologic models for analyzing the effects of land use change. *J. Hydroinform.* **13** (3), 461–473.

Pedersen, L., Jensen, N. E., Christensen, L. E. & Madsenm, H. 2010 Quantification of the spatial variability of rainfall based on a dense network of rain gauges. *Atmos. Res.* **95** (4), 441–454.

Sayood, K. 2000 *Introduction to Data Compression*. Elsevier, Oxford.

Segond, M. L., Wheater, H. S. & Onof, C. 2007 The significance of spatial rainfall representation for flood runoff estimation: a numerical evaluation based on the Lee Catchment, UK. *J. Hydrol.* **347** (1–2), 116–131.

Shah, S. M. S., O'Connell, P. E. & Hosking, J. R. M. 1996 Modelling the effects of spatial variability in rainfall on catchment response. 2. Experiments with distributed and lumped models. *J. Hydrol.* **175** (1–4), 89–111.

Skøien, J. O. 2003 Characteristic space scales and timescales in hydrology. *Water Resour. Res.* **39** (10), WR001736.

Syed, K. H., Goodrich, D. C., Myers, D. E. & Sorooshian, S. 2003 Spatial characteristics of thunderstorm rainfall fields and their relation to runoff. *J. Hydrol.* **271** (1–4), 1–21.

Watson, A. B., Yang, G. Y., Solomon, J. A. & Villasenor, J. 1997 Visibility of wavelet quantization noise. *IEEE Trans. Image Process* **6** (8), 1164–1175.

Wiegand, T. & Schwarz, H. 2011 Source coding: part I of fundamentals of source and video coding. *Foundations Trends Signal Process.* **4** (1–2), 1–222.

Wood, E. F., Sivapalan, M., Beven, K. & Band, L. 1988 Effects of spatial variablity and scale with implications to hydrologic modeling. *J. Hydrol.* **102** (1–4), 29–47.

Wood, E. F., Lettenmaier, D. P. & Zartarian, V. G. 1992 A land-surface hydrology parameterization with subgrid variability for general circulation models. *J. Geophys. Res.* **97** (D3), 2717–2728.

Wood, S. J., Jones, D. A. & Moore, R. J. 2000 Accuracy of rainfall measurement for scales of hydrological interest. *Hydrol. Earth Syst. Sci.* **4** (4), 531–543.

Zhang, J. & Han, D. 2017 Assessment of rainfall spatial variability and its influence on runoff modelling: a case study in the Brue Catchment, UK. *Hydrol. Process.* **31** (16), 2972–2981.

Zhang, N., Xia, Z., Zhang, S. & Jiang, H. 2012a Temporal and spatial characteristics of precipitation and droughts in the upper reaches of the Yangtze River Basin (China) in recent five decades. *J. Hydroinform.* **14** (1), 221–235.

Zhang, X., Hörmann, G., Fohrer, N. & Gao, J. 2012b Estimating the impacts and uncertainty of changing spatial input data resolutions on streamflow simulations in two basins. *J. Hydroinform.* **14** (4), 902–917.