Characterization of dynamic evolution of the spatio-temporal variation of rain-field in Hong Kong

Peng Liu and Yeou-Koung Tung

ABSTRACT

A significant part of Hong Kong has hilly terrain with relatively short flow concentration time and, hence, is susceptible to the threat of flash floods and landslides during intense convective thunderstorms and tropical cyclones. For places like Hong Kong, a rainfall model that could adequately capture small-scale temporal and spatial variations would be highly desirable. The main challenge in rain-field modeling is to capture and describe the dynamic time-space evolution of the rainfall during rainstorm events. In this study, radar data with a high spatial (1 km^2) and temporal (6 min) resolution of four rainstorm events in Hong Kong are analyzed. A geostatistical approach based on indicator variograms of rain-fields is used. The spatial structure of a rain-field is found to be highly anisotropic and should be adequately considered in the model. Variability of the spatial structure of a rain-field was described well by the main features of the variograms. Moreover, it is possible to identify whether multiple rainstorm centers exist by comparing the mean length and range. In order to establish reliable statistics on the spatial and temporal structure of rain-fields in Hong Kong, this approach could be applied to a large set of rainstorm events in this same region in the future.

Key words | geostatistics, mean length, precipitation, spatial structure, variogram

Peng Liu (corresponding author)
State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering,
Nanjing Hydraulic Research Institute,
223 Guangzhou Road,
Nanjing 210029,
China
E-mail: *pliu@nhri.cn*

Yeou-Koung Tung
Department of Civil and Environmental Engineering,
The Hong Kong University of Science and Technology,
Clear Water Bay,
Kowloon,
Hong Kong

INTRODUCTION

Hong Kong is small in size but complex in topography. There are extensive mountain ranges with substantial altitude variations scattered throughout the Territory. These areas and catchments have a relatively short flow concentration time and are particularly susceptible to the threat of flash floods and landslides during intense convective thunderstorms and tropical cyclones. For places like Hong Kong, a rainfall model that could adequately capture small-scale temporal and spatial variations of rain-fields would be highly desirable. In rain-field modeling, the main challenge is to capture and describe the dynamic time-space evolution of the rainfall during rainstorm events.

However, most current models make simplified assumptions about the way rainfall fields evolve, like space homogeneity and weak time dependency. A false assumption of a stationary model may lead to unreliable estimation. This kind of model cannot be expected to preserve important statistical features of the actual rainstorm events, in particular, at finer temporal and spatial scales. Furthermore, the orographic effect on the rainfall amount has been clearly observed across the Hong Kong Territory. A rainfall model that can consider non-stationarity in space would be desirable and useful.

It has long been acknowledged that the intensity of precipitation is highly variable over space and time in mountain environments (Castro *et al.* 2014). An understanding of the space-time structure of rain-fields is of great importance for storm modeling, runoff prediction, and subsequent hydraulic structural design (Abedini *et al.* 2013; Kang & Merwade 2014; Shafiei *et al.* 2014; Xu *et al.* 2015). Arnaud *et al.* (2002) estimated the peak flows and runoff volumes from the distributed hydrological models using different patterns of rainfall as input (namely, spatially averaged uniform rainfall and non-uniform rainfall) and studied the influence

doi: 10.2166/nh.2015.129

of the spatial variability of rainfall on the estimation of runoff characteristics. They concluded that runoff volumes and peak flows could be affected considerably by the spatial variation of rainfall patterns, as well as the runoff production model, the size of the catchment, and the frequency of the event. Le Lay & Saulnier (2007) evaluated the impact of the space-time structure of the rainfall on the distributed hydrological response and suggested that flash flood events were primarily controlled by the space-time structure of the rainfall. The entire circulation of water in a basin is governed by the spatial and temporal distribution of rainfall. Hence, it is crucial to quantitatively characterize the variability and the structure of intense precipitation in rainfall forecasting, flood warning, and real-time operation of major hydrosystem infrastructures to mitigate potential threats and losses caused by floods.

Although ground measurements of rainfall at raingauge are more accurate, radar measurements have gained increasing attention and become an essential source of the rain-field information because they can provide rainfall data with high spatio-temporal resolution (Smith *et al.* 2004; Viglione *et al.* 2010; Li *et al.* 2013; Finsen *et al.* 2014; Adjei *et al.* 2015). Quite a number of works have been carried out to analyze the structure of precipitation based on radar rainfall observations (Austin & Houze Jr 1972; Crane 1990; Steiner *et al.* 1995; Gebremichael & Krajewski 2004; Wu *et al.* 2015).

In this study, a geostatistical view of the precipitation field is adopted. Originally for mining, geostatistics is preferred because it offers a useful theoretical and mathematical description of spatially structural properties of natural phenomena and a practical technique to solve estimation problems (Berne *et al.* 2004). The geostatistics framework has also been applied in water sciences (Hevesi *et al.* 1992; Bacchi & Kottegoda 1995; Prudhomme & Reed 1999; Goovaerts 2000; Lloyd 2005; Buytaert *et al.* 2006; Skøien & Blöschl 2006). To quantify the variability of precipitation, Barancourt *et al.* (1992) addressed the issue of spatial intermittence of rainfall by developing a geostatistical model using a binary random function to describe the intermittency and, a second random function to represent the inner variability of rainfall inside the rainy areas. The procedure showed its clear superiority for delineating and estimating rainfall fields as compared with a classic global kriging. However, a procedure such as this has certain drawbacks: (1) independence between within-storm variability and variability due to fractional coverage must be assumed; (2) estimation variance cannot be produced; and (3) it is difficult to identify exactly what conditional estimates are sought and the impacts of independence assumption on the estimation due to the lack of mathematical precision of the estimation procedure.

Spatial variation of alpine precipitation was quantified by Germann & Joss (2001) by means of variogram analysis using high-resolution radar reflectivity data. Even under the difficult conditions in a mountainous region, the results gave quantitative answers to practical questions related to the spatial continuity. Berne *et al.* (2004) estimated the minimum resolutions of rainfall required for hydrological applications, based on quantitative investigations of the space-time scales of urban catchments and rainfall. The approach was proven to be relevant for various regions. Using the variogram within the geostatistics framework, the space-time structure of rainfall was studied. Berne *et al.* (2009) analyzed the variability of the spatial structure during intense Mediterranean precipitation using the concept of mean length, which was shown to be an efficient tool to represent the spatio-temporal variation of rain-fields. However, the indicator variogram, which could provide additional information about the structure and the correlation of the studied random field, was not used. Owing to the complexity of the spatial structure of precipitation fields, the identification of the range of a variogram remains a difficult task.

The main objective of this paper is to contribute to a better description of the dynamic time-space evolution of the rainfall during rainstorm events by combining the indicator variogram with the mean length. The paper is organized as follows. First, a brief description of the study area and the data preparation is presented. Then, the methodology to quantify the structure of rainfall is introduced. After discussions of the results, the conclusions that can be drawn from this study are presented.

STUDY AREA AND DATA PREPARATION

Study area

Hong Kong, with a population of about seven million people, is situated on China's south coast and is enclosed

by the Pearl River Delta and South China Sea. Heavy rain in summer in Hong Kong is typically associated with the southwest monsoon and tropical cyclones. The threat of flooding is greatest during storm surges generated by the passage of typhoons due to high wind. Many parts of the Territory are densely populated urban areas surrounded by or situated within catchments that are typically small with relatively steep slopes. These urban areas and catchments have a relatively short time-of-concentration and are particularly susceptible to the threat of flash floods and extensive flooding rendering significant economic losses and possible fatalities. Apart from urban flash flooding, another disaster caused by persistent heavy rain in Hong Kong is landslides because of the hilly terrain.

Ground rainfall data and radar rainfall data

The rainfall data used in this study were obtained from the Hong Kong Observatory (HKO). To make radar measurement on reflectivity useful for rainfall estimation, a transfer function is needed to allow conversion of radar reflectivity readings to ground-based rainfalls. The establishment of this transfer function requires an empirical correlation study between the two types of measurements during a storm event. A total of 132 raingauge stations, 46 from the HKO and 86 from Geotechnical Engineering Office (GEO), over the Territory of Hong Kong were used. Figure 1 shows the location of the raingauge stations. The HKO developed a rainstorm nowcasting system termed SWIRLS (Short-range Warning of Intense Rainstorms in Localized Systems) in 1997 (Li & Lai 2004). To convert the radar echo intensity (i.e., reflectivity Z in linear units mm^6/mm^3) to rainfall intensity R in mm/h, the standard Marshall–Palmer relationship ($Z = aR^b$ with $a = 200$ and $b = 1.6$) is used in many operational nowcasting systems. However, it is well-known that many factors can affect the accuracy of the radar-rainfall estimation and this relationship is too generalized, which may not be suitable in many cases. Another fact is that the $Z - R$ relationship changes with time as the precipitation system evolves. SWIRLS makes use of both radar and raingauge data over Hong Kong to calibrate radar reflectivity in real-time for a reasonable $Z - R$ relationship.

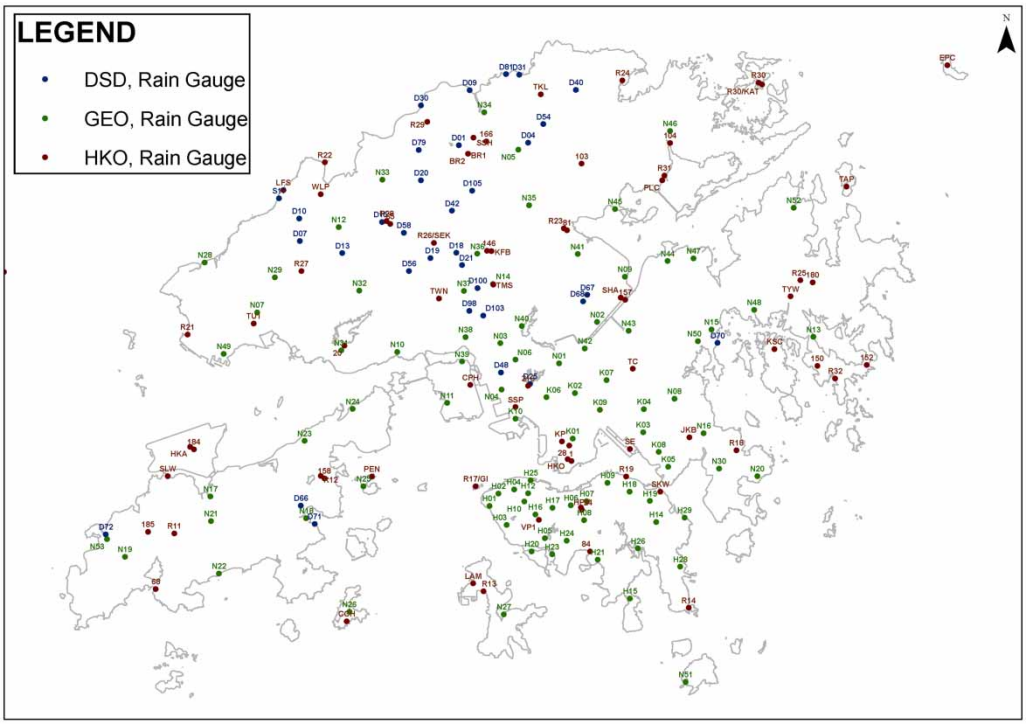

Figure 1 | Location of raingauge stations in Hong Kong.

To determine a and b in the $Z-R$ relationship between radar–raingauge pairs, linear least square analysis is used.

$$\mathrm{dBZ} = 10\log a + b \times \mathrm{dBG} \quad (1)$$

where dBZ is the radar reflectivity expressed in decibel; a and b are parameters of $Z-R$ relationship; and dBG is the ground rainfall expressed in decibels. All the reporting radar–raingauge pairs once accepted will be retained for the linear least square analysis throughout the entire rainstorm episode to ensure statistical significance.

The radar reflectivity with a high spatial (1 km^2) and temporal (6 min) resolution which is detected at 1 km height is correlated every 5 min with the rainfall recorded by the raingauges underneath to gain the optimal parameter a and b. Owing to the existence of winds, raindrops may deviate horizontally from their starting positions, so a certain adjusted approach of searching area from each raingauge position is devised and the radar image with low altitude is chosen to make sure the raindrops preserve their volume when they reach the ground. If there is no ground true value available as the rain has not fallen over any of the raingauges, such as a rainstorm is still outside Hong Kong, either the standard Marshall–Palmer relationship $a = 200$, $b = 1.6$ or the latest $Z-R$ relationship will be used as first guess. As more radar–raingauge pairs are available, the $Z-R$ relationship will be updated accordingly. Analysis strategies are also proposed to deal with the non-synchronous problem arising from the 6 min radar volume scans and the 5 min ground raingauges distributed over the Territory of Hong Kong (Li 2000). For a 6 min radar sweep, the radar data approximately corresponds to two 5 min raingauge observations and the apportioning of weights between the two latest available raingauge observations is specified when applying temporal interpolation. Figure 2 shows the approaches adopted to circumvent the non-synchronous problem and adjust the $Z-R$ relationship.

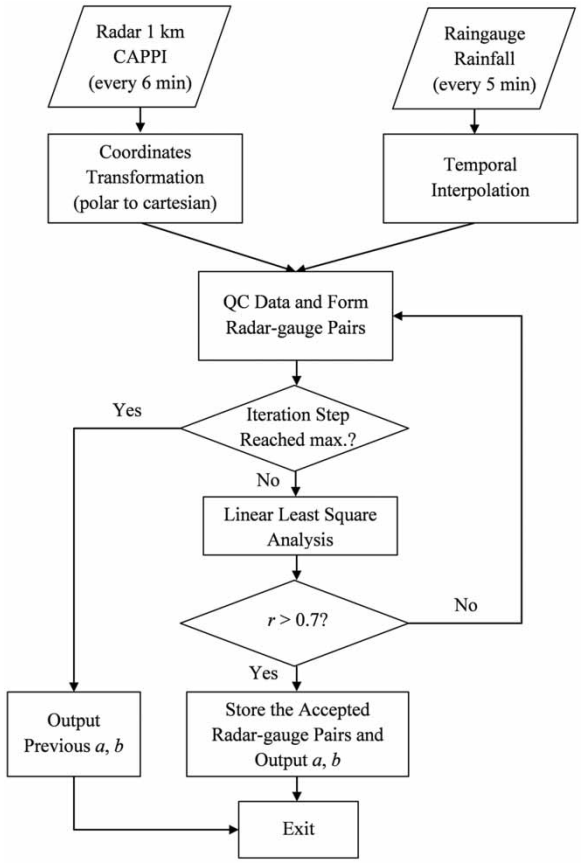

Figure 2 | Flow diagram for the automatic radar-raingauge adjustment (Li 2000).

General features of four rainstorm events

In this study, an area of 64×64 km^2 region covering the entire Territory of Hong Kong is considered. Four rainstorm events were chosen.

2007-05-18 and 2007-05-19 rainstorm events

On 18 May 2007, a trough of low pressure developed over inland Guangdong and moved across the south China coast. A squall line swept across Hong Kong from northwest to southeast that evening, bringing heavy showers and severe gusts to the Territory. A peak gust over 100 km/h was recorded at Waglan Island. With the trough of low pressure lingering along the south China coast, there were heavy showers and thunderstorms between 19 and 22 May.

2008-04-19 rainstorm event

Under the combined effect of Typhoon Neoguri and the northeast monsoon, local winds started to pick up on 18 April 2008. Local winds became generally strong on the

afternoon of 19 April 2008 when Neoguri was about 200 kilometers to the west-southwest of Hong Kong. When Neoguri approached the coast of western Guangdong, the warm southerly winds associated with Neoguri met the relatively cooler northeast monsoon and formed a warm front with severe convective activities over the coastal waters of Guangdong. The warm front moved from south to north across the coast of Guangdong and brought heavy rain to Hong Kong on that day. The total daily rainfall recorded at the HKO on that day was 233.4 mm, the highest daily rainfall amount recorded in April since records began. After making landfall and moving further inland, Neoguri weakened rapidly. Locally, rain also eased off rapidly with just a few showers on 20 April 2008.

2009-09-15 rainstorm event

A tropical depression over the western North Pacific entered the northern part of the South China Sea on 13 September 2009. It intensified into a tropical storm that evening and was named Koppu. The outer rainbands of Koppu brought a few thundery showers to Hong Kong during that evening. Koppu intensified into a typhoon on 14 September and the weather became cloudy with squally showers. Local winds also strengthened gradually towards the evening. Gale force southeasterly winds with maximum gusts up to 151 km/h were recorded at the Cheung Chau Beach in the late evening of 14 September. Koppu made landfall over the western coast of Guangdong on the morning of 15 September and weakened into a tropical storm in the afternoon. The rainbands associated with Koppu brought heavy squally showers to the Territory on that day. With Koppu moving further inland and dissipating, the showers gradually eased off on 16 September.

For the purposes of illustration, the rainstorm that occurred on 19 April 2008 is used herein.

The 2008-04-19 rainstorm event lasted 22 hours and occurred between 02:00 am and 24:00 pm on the 19 April 2008. The contour map of total amounts of rainfall over the study area is shown in Figure 3. Territory-wide, the maximum total rainfall amount for the 2008-04-19 rainstorm event was 157.9 mm.

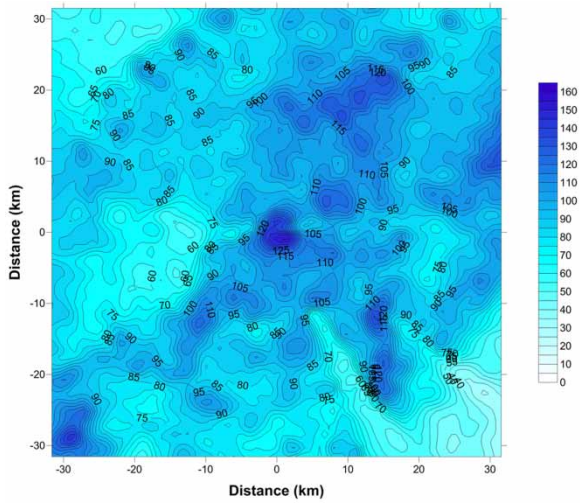

Figure 3 | Contour of total rain amount (in mm) for the 2008-04-19 rainstorm event.

METHODOLOGY

Theoretical variogram model

Geostatistics is used to investigate the spatial structure of rainfall at each time step during the entire rainstorm events. A main concept in geostatistics is the variogram, which quantifies the spatial continuity of regionalized variables. An important advantage of the variogram over the covariance is that no information about the variance is required for its calculation. Therefore, the variogram is used more often than the common covariance function. For an intrinsic random function, its variogram is defined as half of the variation between two points in a spatial field as a function of their distance separation lag vector.

$$\gamma(\vec{h}) = \frac{1}{2}\mathrm{Var}[R(x+\vec{h}) - R(x)] \tag{2}$$

where $\gamma(\vec{h})$ is the variogram, which is also called semi-variogram; $\vec{h}$ is the distance separation lag vector; and R is the random variable under consideration (e.g., rainfall intensity). The semi-variogram does not depend on the mean of the random function, therefore is more robust than the covariance.

Isotropic variogram describes a random field in which spatial correlation does not vary with directions but only

on the separation distance. It is then a function of the modulus of the vector $\vec{h}$. When the variogram shows different behavior along different directions, the random field is anisotropic.

The best possible estimate of point or mean areal rainfall can only be provided by interpolation method (e.g., kriging) when the set of variables is multivariate Gaussian (Barancourt *et al.* 1992). However, the distribution of rain rates is significantly asymmetric and skewed toward high values and hence far from normality at short accumulation periods. A simpler answer to circumvent the non-Gaussian distribution issue is by thresholding which offers an application of the indicator method to map the probability distributions of rainfall (Steffens 1987). In the classical approach, the rainfall process $R(x, t, w)$ is analyzed as a random field with R being a non-negative value of rainfall rate (or amount) that fell at location x at time t during event w. In this case, the actual rain process is transformed into a binary process to distinguish whether the rainfall intensity R at location x at time t is above a given threshold k under consideration. For this, a binary function denoted $i(x, t, w, k)$ is defined by:

$$i(x, t, w, k) = \begin{cases} 1, R(x, t, w) > k \\ 0, R(x, t, w) \leq k \end{cases} \quad (3)$$

The indicator function $i(x, t, w, k)$ will be 1 if the rainfall intensity $R(x, t, w)$ is above a given threshold k. The indicator variograms of the rain-field for each threshold will be calculated based on these indicator values $i(x, t, w, k)$. These indicator variograms can provide useful information about the spatial intermittency of the rain-field for the considered thresholds as well as the size and shape of the rainy areas above/below the thresholds.

To check the data for anisotropy, several indicator maps are constructed, from which indicator variograms are calculated for different directions to examine the anisotropy of the rain-field (see Figure 4). It is clear to see that the spatial structure of the rain-field reveals strong anisotropy. Range varies with the direction and sill is also not a constant. For a specific direction, the variogram may display a lower sill

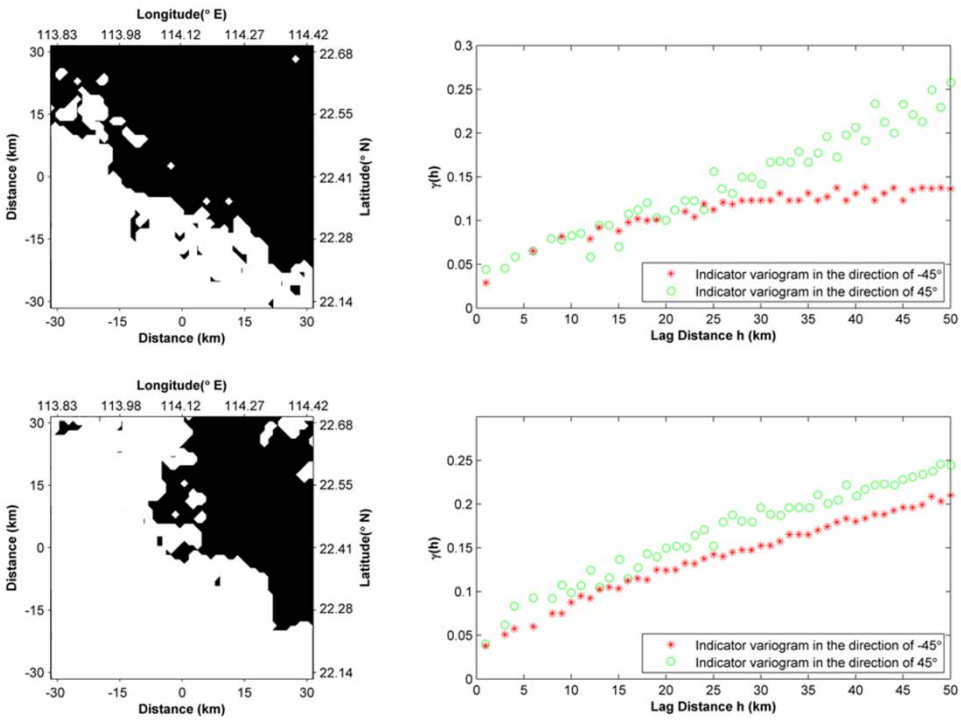

Figure 4 | Rain-field at $t = 20{:}00$ pm during the 2008-04-19 rainstorm event at 6-min time resolution, with black regions indicating the values above the threshold of 2 mm/h (top left) and 10 mm/h (bottom left) and corresponding indicator variogram in the direction of -45° and 45° (right).

than the others and show a larger range. Therefore, the indicator rainfall data need to be treated according to different directions separately to preserve anisotropy. As the variogram is symmetrical with respect to the origin, the angles between -90° and 90° are sufficient. For each time step, a total of 24 directions along which the indicator variograms are fitted and mean lengths are estimated in the study.

Since strong anisotropy exists, pairs are grouped into space lag 'bins' for different directions θ and the sample indicator variogram value for that bin is calculated. The mean lag of all the pairs in a particular space lag bin is used as the representative lag distance for that bin. The sample indicator variogram $\gamma^*(h, t, w, k, \theta)$ is calculated as:

$$\gamma^*(h, t, w, k, \theta) = \frac{1}{2N(h, t, w, k, \theta)} \sum_{i=1}^{N(h,t,w,k,\theta)} [i(x + h, t, w, k, \theta) - i(x, t, w, k, \theta)]^2 \quad (4)$$

where $N(h, t, w, k, \theta)$ denotes the number of pairs of points $(x + h, x)$ in the direction θ within the domain. The general rule is to choose the maximum bin as half of the maximum distance, and also make sure each bin has more than 30 pairs of data (Journel & Huijbregts 1978).

Fitting the theoretical variogram model to experimental data

To estimate nugget effect, sill, and range for each sample indicator variogram, it is necessary to fit a model to the experimental variogram. Some of the models that are often used in practice include linear, spherical, exponential and Gaussian variogram models.

The fitting of the theoretical variogram model to the experimental data $\gamma^*(h, t, w, k, \theta)$ has to be done for each time step and each direction, separately. Least squares criterion is used to determine the optimal parameters of a considered variogram model for a particular time instant t and direction θ, $d(t, \theta) = (C_0(t, \theta), C(t, \theta), a(t, \theta))$, by minimizing

$$Q[d(t, \theta)] = \sum_{i=1}^{N_h} [\gamma^*(h_i, t, w, k, \theta) - \gamma(h_i, d(t, \theta))]^2 \quad (5)$$

where N_h is the total number of considered distance lags. From the visual inspection of the indicator variograms, nugget effects are negligible for most directions and time instants, thus C_0 is set to 0 during the fitting procedure. In most cases, variogram value increases with distance but there is an erratic behavior in some directions when distance lag between pixels is large. Hence, high weights to the small distances are given by the weighted least squares technique (Cressie 1985) in this study to determine optimal variogram model parameters as:

$$Q[d(t, \theta)] = \sum_{i=1}^{N_h} w_i[\gamma^*(h_i, t, w, k, \theta) - \gamma(h_i, d(t, \theta))]^2 \quad (6)$$

where the weights w_i is given by h_i^{-1}. The distance bin that has the maximum number of pairs is about 25 km in the majority of the rainfall intensity thresholds and time instants considered. Thus, distance separation less than or equal to 50 km is considered in the process of fitting variogram models in attempting to identify the best-fit variogram model for further analysis.

Mean length

To quantify the spatial characteristic scale of rain events, another spatial characteristic of rain-field, called mean length $\bar{l}(t, w, k, \theta)$, is calculated. This length corresponds to the average length of the segments defined by rain rate values above the threshold along a given direction in a particular time during a rainfall event (Carle & Fogg 1996). In this study, the mean length $\bar{l}(t, w, k, \theta)$ is calculated as:

$$\bar{l}(t, w, k, \theta) = \frac{p(t, w, k, \theta)}{\frac{\partial \gamma(h \to 0, t, w, k, \theta)}{\partial h}} \quad (7)$$

where $\bar{l}(t, w, k, \theta)$ denotes the mean length of the segments defined by rainfall intensity values at a particular time t during a rainstorm event w above the threshold k in a given direction θ; $p(t, w, k, \theta)$ is the proportion of rainfall intensity values above the threshold k in the direction θ; and $\gamma(h, t, w, k, \theta)$ is the indicator variogram associated with the rainfall threshold k in the direction θ.

Analysis of the sample indicator variograms in many directions suggests that they are linearly increasing for short distance lag from 1 to 9 km. Therefore, the derivative of $\gamma(h, t, w, k, \theta)$ in Equation (7) is estimated as the slope of the regression line between the sample indicator variogram values and distance lag from 1 to 9 km. Although only the points within 9 km are considered in the calculation of the derivative of the variogram which may not be as reliable as using all the data, by limiting calculations to sufficiently short distance lags makes the simple linear model become valid. In addition, only local stationarity is required to estimate the indicator variogram considering short distance lags (Berne *et al.* 2009). Furthermore, this linearity assumption is also confirmed by the high correlation coefficients (which measures the strength of linear dependence) between the sample data from 1 to 9 km.

The domain size that is rainy must be suitable to the mean length calculation. The sample variogram and the mean proportion are representative of the rain-field only if the domain is large enough and appropriately located to observe the entire precipitating system (Berne *et al.* 2009). Based on that, mean length can be derived. In addition, a proportion above a given threshold close to 1 means that the entire study domain is rainy above the threshold whereas a proportion above a given threshold close to 0 means that there is nearly no rainfall above the threshold within the studied area. Hence, in this study, the mean length of the rain-field is not calculated if the proportion of area where rainfall intensity is above the given threshold is higher than 99% or lower than 1%.

Range of the indicator variogram and mean length are two different characteristic scales of a random function, and both of them offer important information about the spatial structure of rain-field. Range is the limiting distance beyond which two points are uncorrelated, which means that within the range rainfall intensities at two points have some correlation. Mean length is the average length of the segments defined by the rainfall intensity value above the considered threshold, which means, on average, two points separated by a distance that is less than the mean length tend to belong to the same category (above or below the considered threshold). It is interesting to investigate the relationship between these two since mean length has been proved to be an efficient measure to represent the variability of the spatial structure of rain-field by Berne *et al.* (2009). In this study the mean length is used to check the automatic fitting procedure of indicator variogram, to study its correlation with range, and to characterize the evolution of spatial structure of rain-field for the four studied rainstorm events along with the range.

RESULTS AND DISCUSSION

Spatial structure of rain-field

To choose the thresholds, the quantiles of the rainfall intensity values are obtained for each rainstorm event (see Figure 5). The quantiles 10–40% are almost equal for the 2007-05-18 rainstorm event and 2009-09-15 rainstorm event, while larger quantiles (60–90%) vary considerably. The same situation can be found between the 2007–05-19 rainstorm event and 2008-04-18 rainstorm event. Table 1 lists the values of the 10, 30, 50, 70, and 90% quantiles. From these quantiles, two thresholds were chosen for each rainstorm event (1 mm/h and 5 mm/h for the 2007-05-18 rainstorm event, 2 mm/h and 10 mm/h for the 2007-05-19 and 2008-04-19 rainstorm events, 1 mm/h and 10 mm/h for the 2009-09-15 rainstorm event) to characterize the spatial structure of rain-field and examine the relationship between range of the variogram and the mean length. The two selected thresholds for each rainstorm event are different but correspond to similar quantiles (the smaller one is between 10 and 30%, and the other one is between 70 and 90% representing the intense rain cells).

The indicator maps for different threshold k at different time t during event w are plotted and indicator variograms for each indicator map are calculated the same way as standard variograms, except that the indicator values $i(x, t, w, k)$ are used in place of the actual rainfall intensity values. The analysis result of the rainstorm that occurred on 19 April 2008 is shown herein to illustrate the method.

Distance separation greater than 50 km is not included in the fitting procedure to avoid the erratic behavior of the variogram in some direction when distance lag gets large. Among the four theoretical variogram models, the

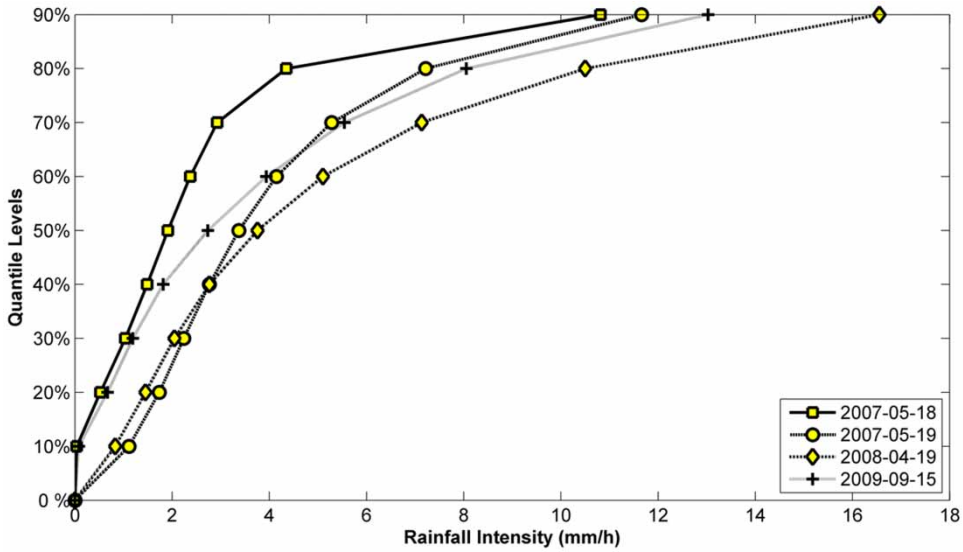

Figure 5 | Quantiles of the distribution of the rain rate values for the four studied rainstorm events.

Table 1 | Rainfall intensity quantiles (in mm/h) for the four studied rainstorm events

	Rain rate quantile levels				
Event	**10%**	**30%**	**50%**	**70%**	**90%**
2007-05-18	0.036	1.035	1.914	2.928	10.813
2007-05-19	1.112	2.239	3.373	5.285	11.66
2008-04-19	0.827	2.047	3.756	7.135	16.552
2009-09-15	0.078	1.19	2.735	5.545	13.03

exponential variogram was found to best fit the sample variograms (see example results in Table 2) in the majority of the rainfall intensity thresholds and time instants considered with the highest mean $R^2 = 0.928$ (0.905 for spherical variogram, 0.886 for Gaussian variogram and 0.834 for linear variogram, respectively).

Spatial structure of the rain-field is examined using the indicator maps for different thresholds at each time step during the entire rainfall events. Two example indicator maps at 20:00 pm during the 2008-04-19 rainstorm event and the corresponding variograms in the direction of maximum range and its perpendicular direction are shown in Figure 6. For the threshold of 2 mm/h, the maximum range exists in the direction of 45 ° while for the threshold of 10 mm/h, the maximum range exists in the direction of −50.19 °. Range varies with the direction for both thresholds and the value as well as the corresponding direction could give most information about the spatial structure of the instantaneous rain-field. Figure 7 provides the corresponding ranges for different directions.

Estimated ranges for the threshold of 2 mm/h and 10 mm/h in different directions indicate that the shapes of the rain-fields as well as the magnitude of the ranges are different for the two considered thresholds. For the threshold of 2 mm/h, large ranges exist in the northeast–southwest direction and ranges are comparatively small along the northwest–southeast direction. While for the threshold of 10 mm/h, relatively larger ranges can be found in both northeast–southwest and northwest–southeast directions. The value of the maximum range for the threshold of 10 mm/h is about 80 km, which is much smaller than the maximum value (400 km) for the threshold of 2 mm/h. For each time step, the spatial structure of rain cells above the considered threshold more or less exhibits an elliptic shape just like these two examples shown. Therefore, it is possible and sufficient to describe the shape of the rain-field by the values and the directions (quantified as an angle) of the maximum range and the minimum range instead of the large amount of information (ranges and sills for 24 different directions) of the spatial structure at each time step.

Table 2 | Sample range, sill, and R^2 of fitted exponential model for different directions under thresholds of 2 and 10 mm/h at 20:00 pm during the 2008-04-19 rainstorm event at 6 min time resolution

	Threshold = 2 mm/h			Threshold = 10 mm/h		
Angle (°)	Range (km)	Sill	Coefficient of determination R^2	Range (km)	Sill	Coefficient of determination R^2
−80.54	138.33	0.36	0.986	48.48	0.14	0.954
−71.57	116.63	0.3	0.988	45.99	0.15	0.895
−63.43	85.56	0.24	0.984	57.36	0.17	0.91
−56.31	59.57	0.19	0.964	69.25	0.2	0.929
−50.19	48.44	0.16	0.92	81.06	0.22	0.94
−45	33.58	0.14	0.951	77.95	0.23	0.941
−39.81	20.91	0.12	0.885	56.34	0.21	0.889
−33.69	30.96	0.13	0.885	65.08	0.22	0.919
−26.57	25.96	0.12	0.876	50.62	0.2	0.895
−18.43	29.15	0.13	0.841	39.98	0.19	0.875
−9.46	40.98	0.14	0.86	35.17	0.17	0.878
0	45.29	0.16	0.883	28.99	0.17	0.863
9.46	116.81	0.26	0.928	33.52	0.17	0.901
18.43	170.39	0.37	0.945	34.82	0.18	0.892
26.57	227.3	0.49	0.955	37.64	0.2	0.892
33.69	256.85	0.58	0.955	45.23	0.22	0.907
39.81	327.85	0.71	0.961	55.5	0.23	0.928
45	377.84	0.79	0.953	47.8	0.23	0.908
50.19	237.67	0.6	0.954	66.47	0.24	0.928
56.31	326.44	0.72	0.963	61.54	0.23	0.932
63.43	244.26	0.58	0.972	55.03	0.22	0.918
71.57	221.23	0.54	0.985	53.27	0.2	0.918
80.54	217.19	0.53	0.985	49.62	0.18	0.903
90	188.13	0.46	0.985	33.76	0.15	0.836

Relationship between variogram range and mean length

To analyze the evolution of the spatial structure during a rainstorm event, the maximum range of the semi-variogram, minimum-to-maximum range ratio, and the direction of the maximum range are calculated for each time step. The time resolution considered is 6 min. The relationship between range and mean length is also studied by comparing the maximum and minimum values at the same time instant and the corresponding directions. They are plotted as a function of time in Figures 8 (threshold of 2 mm/h) and 9 (threshold of 10 mm/h) for the 2008-04-19 rainstorm event. The ranges and mean lengths quantitatively characterize the spatial structure of rain-fields.

It is interesting to investigate the behavior of range and mean length in relation to time. There are some features that can be observed. First, maximum range and maximum mean length share a similar trend, while the angle of the maximum range may not be identical to that of the maximum mean length for some time instants. According to rain-field measurements, it is not surprising to observe the above-mentioned features because: (1) although the angle of the maximum range and the angle of the maximum mean length are different, the values are very close, so are the values of maximum range and maximum mean length; (2)

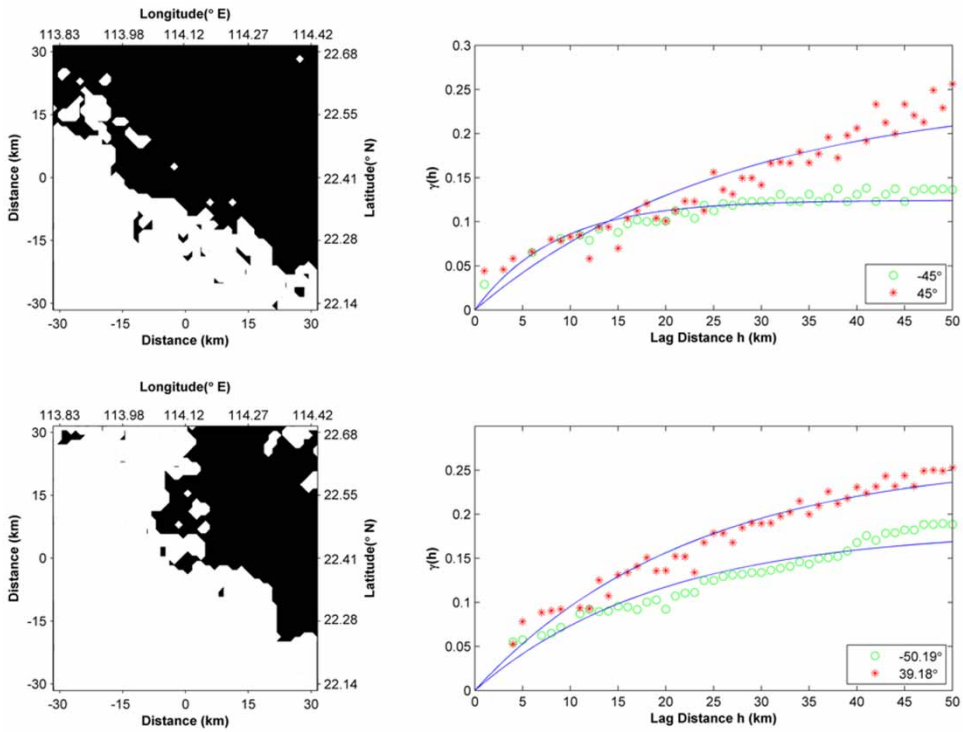

Figure 6 | Sample rain-fields at 20:00 pm during the 2008-04-19 rainstorm event at 6-min time resolution, with black regions indicating the values above the threshold of 2 (top left) and 10 mm/h (bottom left) and corresponding indicator variogram in the direction of max range and its perpendicular direction (right).

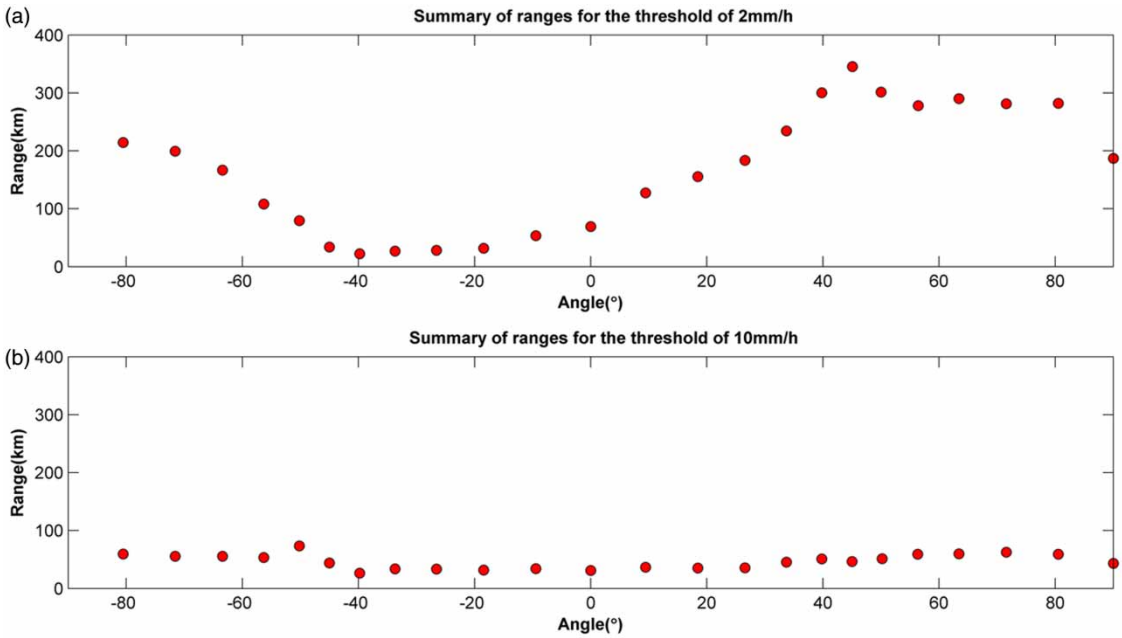

Figure 7 | Variation of estimated ranges with angle for rainfall threshold of (a) 2 mm/h and (b) 10 mm/h at 20:00 pm during the 2008-04-19 rainfall event.

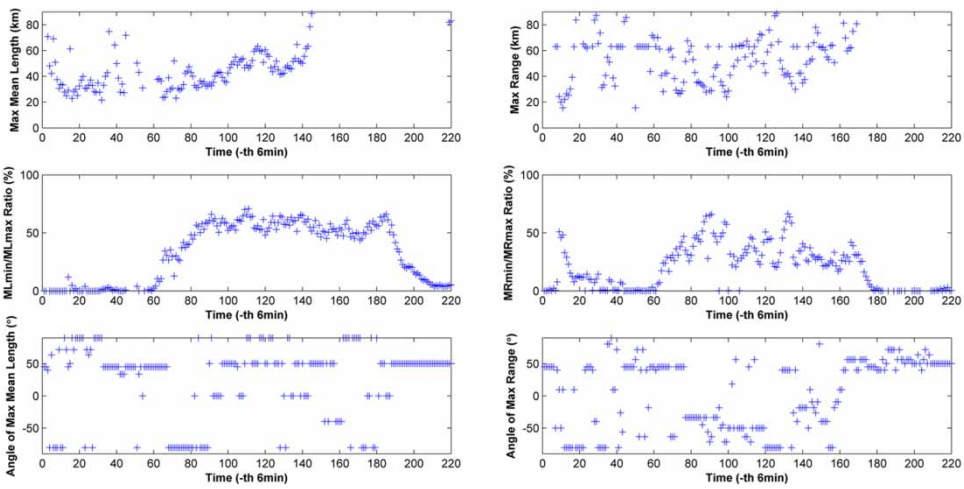

Figure 8 | Time series of max mean length (top left), min/max mean length ratio (middle left) and direction of max mean length (bottom left) as a function of time, time series of max range (top right), min/max range ratio (middle right) and direction of max range (bottom right) as a function of time during the 2008-04-19 rainstorm event, with a threshold of 2 mm/h.

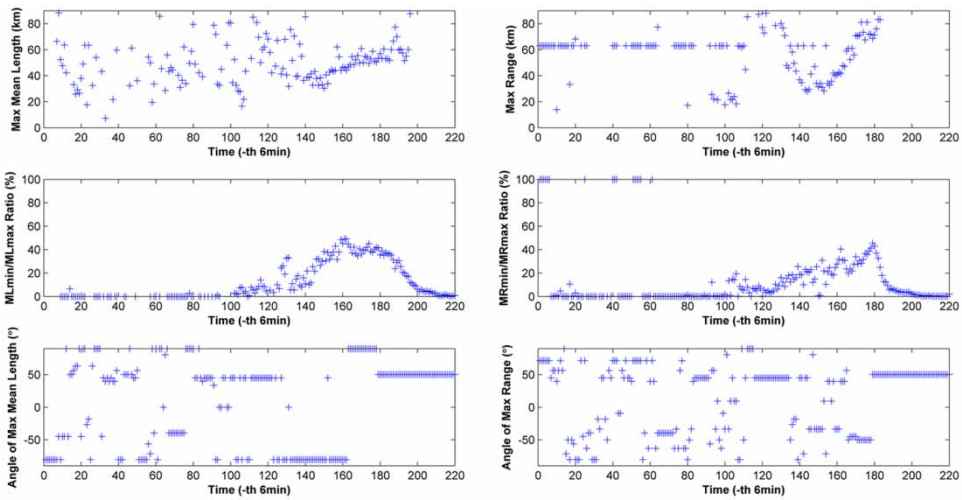

Figure 9 | Time series of max mean length (top left), min/max mean length ratio (middle left) and direction of max mean length (bottom left) as a function of time, time series of max range (top right), min/max range ratio (middle right) and direction of max range (bottom right) as a function of time during the 2008-04-19 rainstorm event, with a threshold of 10 mm/h.

when the spatial structure of the rain-field displays more or less a circular shape, the two angles may differ a lot but the values of maximum range and maximum mean length are similar; (3) mean lengths of a category α representing the rain cells above the threshold k, where the corresponding $i(x, t, w, k) = 1$, can be readily calculated from continuous data as the total length of α in the direction θ divided by the total number of embedded occurrences of α in the direction θ. When multiple embedded occurrences of α exist in some direction θ, mean length of this direction will be much smaller than the range which is not affected that much by the number of the embedded occurrences of α.

The correlations between the maximum (minimum) range and maximum (minimum) mean length for the selected rainstorm events are shown in Table 3. The correlation coefficients of ranges and mean lengths for the 2007-05-18 rainstorm event (0.7561 for rainfall intensity threshold of 1 mm/h and 0.8391 for 5 mm/h, respectively) are the highest among all considered rainstorm events. These values of correlation coefficients give useful

Table 3 | Correlation coefficient between range and mean length

Rainstorm event	2007-05-18	2007-05-19	2008-04-19	2009-09-15
Threshold between quantile level 10–30%	0.7561	0.6707	0.7437	0.7541
Threshold between quantile level 70–90%	0.8391	0.7504	0.5471	0.783

information about the spatial structure of rain-field from a different aspect. Mean length is affected by the number of embedded occurrences of rain cells where rainfall intensity is greater than the threshold. It can be seen from the radar rainfall intensity images that if there are multiple rainstorm centers above the considered threshold at some times during the rainstorm event, the correlation coefficient of ranges and mean lengths is lower. 'Rainstorm centers' here refer to the areas where the rainfall intensity is greater than 10 mm/h. The instantaneous rain-field of the 2007-05-18 rainstorm event at the 33th 6 min (see Figure 10) clearly indicates that the only one rainstorm center is located in the northwest. For the 160th 6 min of the 2008-04-19 rainstorm event (see Figure 11), the rainstorm centers occurred at several different locations. Hence, it is possible to identify whether multiple rainstorm centers exist by comparing the mean length and range.

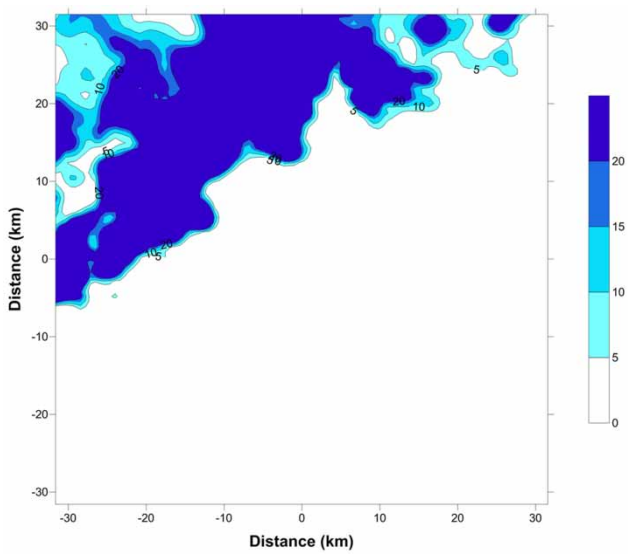

Figure 10 | Instantaneous rainfall intensity contour (in mm/h) over the study area during 2007-05-18 rainstorm event at 33th 6-min.

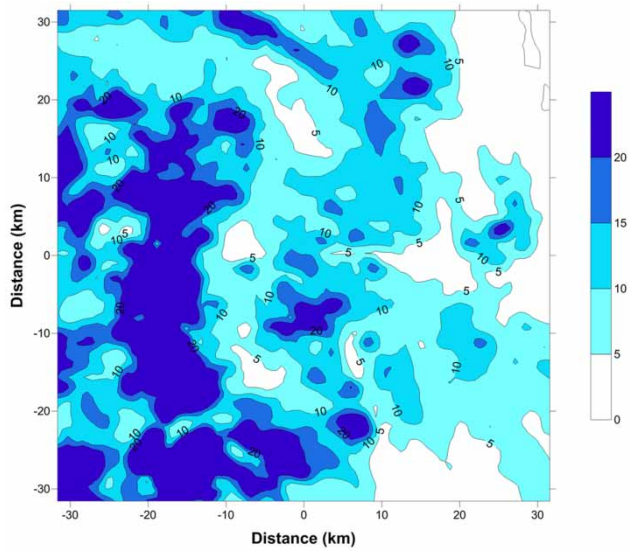

Figure 11 | Instantaneous rainfall intensity contour (in mm/h) over the study area during 2008-04-19 rainstorm event at 160th 6-min.

Second, it is difficult to identify the periodicity of spatial structure parameters for the 2008-04-19 rainstorm event. However, the peaks of the maximum range, maximum mean length, minimum-to-maximum range ratio, and minimum-to-maximum mean length ratio exist for the 2007-05-18 rainstorm event with a periodicity of about 3 hours which is not shown in this paper. The general feature for all considered rainstorm events is the negative correlation between the maximum range (mean length) and the minimum-to-maximum range (mean length) ratio for both thresholds, which indicates that the spatial structure of the rain-field is stretched for the large maximum range and mean length, but circular for small maximum range and mean length.

CONCLUSIONS

During a rainstorm, the precipitation intensity is highly variable in time and space. Rainfall data with high space-time resolution collected by radar have recently been used more frequently as inputs for distributed hydrological modeling of floods, erosion, and other processes.

In this paper, a geostatistical approach was adopted to analyze the variability of the spatial structure of the rain-field. Four rainstorm events that occurred in 2007, 2008,

and 2009 in Hong Kong were used to illustrate the method. Indicator variograms of rain-fields had been examined which revealed substantial anisotropy for all rainfall thresholds considered. Based on indicator variograms, the ranges and mean lengths above different thresholds in 24 directions were estimated for each 6 min time step.

The conclusions that could be drawn from the study of spatial structures of observed rainfall in Hong Kong are listed below:

1. The behavior of variograms is not regular with large distance in some directions. Higher weights to the small distances should be given and distances less than or equal to 50 km are considered when fitting the variogram model.
2. The spatial structure of rain-field for the four rainstorm events is affected by strong anisotropy. The calculation of range and mean length in different directions is needed to capture the spatial structure of anisotropic rain-fields. The influence of this anisotropy condition on generating rain-fields according to the variogram should be investigated.
3. The spatial structure of the four chosen rainstorm events exhibits a more or less elliptic shape. Therefore, it is possible and sufficient to describe the shape of the rain-field by the values and the directions (quantified as an angle) of the maximum range and the minimum range instead of the large amount of information (ranges and sills for 24 different directions) of the spatial structure at each time step. The negative correlation between the maximum range (mean length) and the minimum-to-maximum range (mean length) ratio for both thresholds can be identified for all considered rainstorm events, which indicates that the spatial structure of the rain-field is stretched for the large maximum range and mean length, but circular for small maximum range and mean length.
4. The values of correlation coefficients between the maximum (minimum) range and maximum (minimum) mean length for the selected rainstorm events give useful information about the spatial structure of the rain-field. When multiple rainstorm centers above the considered threshold exist, the correlation between range and mean length of the corresponding rainstorm event becomes lower.

ACKNOWLEDGEMENTS

The authors would like to thank the two anonymous reviewers and the editor who helped us improving the quality of the original manuscript. The authors wish to acknowledge the support of the Hong Kong Research Grant Council (Project no. 620110: Development of time-space rainstorm model for Hong Kong) and Nanjing Hydraulic Research Institute (Project no. Y515014) for the study. Gratitude is extended to the Hong Kong Observatory for providing the rainfall radar data used in the study.

REFERENCES

Abedini, M., Md Said, M. A. & Ahmad, F. 2013 Integration of statistical and spatial methods for distributing precipitation in tropical areas. *Hydrol. Res.* **44** (6), 982–994.

Adjei, K. A., Ren, L. L., Appiah-Adjei, E. K. & Odai, S. N. 2014 Application of satellite-derived rainfall for hydrological modelling in the data-scarce Black Volta trans-boundary basin. *Hydrol. Res.* **46** (5), 777–791.

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

Austin, P. M. & Houze Jr, R. A. 1972 Analysis of the structure of precipitation patterns in New England. *J. Appl. Meteorol.* **11**, 926–935.

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

Barancourt, C., Creutin, J. & Rivoirard, J. 1992 A method for delineating and estimating rainfall fields. *Water Resour. Res.* **28**, 1133–1144.

Berne, A., Delrieu, G., Creutin, J. D. & Obled, C. 2004 Temporal and spatial resolution of rainfall measurements required for urban hydrology. *J. Hydrol.* **299**, 166–179.

Berne, A., Delrieu, G. & Boudevillain, B. 2009 Variability of the spatial structure of intense Mediterranean precipitation. *Adv. Water Resour.* **32** (7), 1031–1042.

Buytaert, W., Celleri, R., Willems, P., Bièvre, B. D. & Wyseure, G. 2006 Spatial and temporal rainfall variability in mountainous areas: a case study from the south Ecuadorian Andes. *J. Hydrol.* **329**, 413–421.

Carle, S. F. & Fogg, G. E. 1996 Transition probability-based indicator geostatistics. *Math. Geol.* **28**, 453–476.

Castro, L. M., Salas, M. & Fernández, B. 2014 Evaluation of TRMM multi-satellite precipitation analysis (TMPA) in a mountainous region of the central Andes range with a Mediterranean climate. *Hydrol. Res.* **46** (1), 89–105.

Crane, R. K. 1990 Space-time structure of rain rate fields. *J. Geophys. Res Atmos.* **95**, 2011–2020.

Cressie, N. 1985 Fitting variogram models by weighted least squares. *J. Int. Assoc. Math. Geol.* **17**, 563–586.

Finsen, F., Milzow, C., Smith, R., Berry, P. & Bauer-Gottwein, P. 2014 Using radar altimetry to update a large-scale hydrological model of the Brahmaputra river basin. *Hydrol. Res.* **45** (1), 143–164.

Gebremichael, M. & Krajewski, W. F. 2004 Characterization of the temporal sampling error in space-time-averaged rainfall estimates from satellites. *J. Geophys. Res.* **109**, D11110.

Germann, U. & Joss, J. 2001 Variograms of radar reflectivity to describe the spatial continuity of Alpine precipitation. *J. Appl. Meteorol.* **40**, 1042–1059.

Goovaerts, P. 2000 Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *J. Hydrol.* **228**, 113–129.

Hevesi, J. A., Istok, J. D. & Flint, A. L. 1992 Precipitation estimation in mountainous terrain using multivariate geostatistics. Part I: structural analysis. *J. Appl. Meteorol.* **31**, 661–676.

Journel, A. G. & Huijbregts, C. J. 1978 *Mining Geostatistics*. Academic Press, London, UK.

Kang, K. & Merwade, V. 2014 The effect of spatially uniform and non-uniform precipitation bias correction methods on improving NEXRAD rainfall accuracy for distributed hydrologic modeling. *Hydrol. Res.* **45** (1), 23–42.

Le Lay, M. & Saulnier, G. M. 2007 Exploring the signature of climate and landscape spatial variabilities in flash flood events: case of the 8–9 September 2002 Cévennes-Vivarais catastrophic event. *Geophys. Res. Lett.* **34**, L13401.

Li, P. 2000 *SWIRLS – An Evolving Nowcasting System*. Hong Kong Observatory, Hong Kong.

Li, P. & Lai, E. S. 2004 Short-range quantitative precipitation forecasting in Hong Kong. *J. Hydrol.* **288**, 189–209.

Li, L., Ngongondo, C. S., Xu, C. Y. & Gong, L. 2013 Comparison of the global TRMM and WFD precipitation datasets in driving a large-scale hydrological model in Southern Africa. *Hydrol. Res.* **44** (5), 770–788.

Lloyd, C. 2005 Assessing the effect of integrating elevation data into the estimation of monthly precipitation in Great Britain. *J. Hydrol.* **308**, 128–150.

Prudhomme, C. & Reed, D. W. 1999 Mapping extreme rainfall in a mountainous region using geostatistical techniques: a case study in Scotland. *Int. J Climatol.* **19**, 1337–1356.

Shafiei, M., Ghahraman, B., Saghafian, B., Pande, S., Gharari, S. & Davary, K. 2014 Assessment of rain-gauge networks using a probabilistic GIS based approach. *Hydrol. Res.* **45** (4–5), 551–562.

Skøien, J. & Blöschl, G. 2006 Catchments as space-time filters – a joint spatio-temporal geostatistical analysis of runoff and precipitation. *Hydrol. Earth Syst. Sci.* **10**, 645–662.

Smith, M. B., Seo, D. J. & Koren, V. I. 2004 The distributed model intercomparison project (DMIP): an overview. *J. Hydrol.* **298** (1–4), 4–26.

Steffens, F. 1987 Spatial distribution of probabilities. APCOM 87. In: *Proceedings of the Twentieth International Symposium on the Application of Computers and Mathematics in the Mineral Industries, Volume 3, Geostatistics*. Johannesburg, South Africa, 83–92.

Steiner, M., Houze Jr, R. A. & Yuter, S. E. 1995 Climatological characterization of three-dimensional storm structure from operational radar and rain gauge data. *J. Appl. Meteorol.* **34**, 1978–2007.

Viglione, A., Chirico, G. B., Woods, R. & Blöschl, G. 2010 Generalised synthesis of space-time variability in flood response: an analytical framework. *J. Hydrol.* **394** (1), 198–212.

Wu, S.-J., Lien, H.-C., Hsu, C.-T., Chang, C.-H. & Shen, J.-C. 2015 Modeling probabilistic radar rainfall estimation at ungauged locations based on spatiotemporal errors which correspond to gauged data. *Hydrol. Res.* **46** (1), 39–59.

Xu, H. L., Xu, C.-Y., Sælthun, N.-R., Zhou, B. & Xu, Y. P. 2015 Evaluation of reanalysis and satellite-based precipitation datasets in driving hydrological models in a humid region of Southern China. *Stoch. Env. Res. Risk A.* **29** (8), 2003–2020.

First received 24 June 2015; accepted in revised form 1 September 2015. Available online 23 October 2015