Effects of network design on climatic maps of precipitation

H. Kutiel[1,*], P. A. Kay[2]

[1]Department of Geography, University of Haifa, Haifa 31905, Israel
[2]Department of Environment and Resource Studies, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

ABSTRACT: We use a simulation study to investigate the effect of network arrangement on the spatial and volumetric estimation of rain fields from a limited number of gauges. Rainfall fields are created as summations of a number of random elliptical events. Each simulation is sampled by 6 different fixed network designs of 9 points. Isohyet maps provide spatial and volumetric estimates, which are then compared to the known characteristics of the events comprising the simulations. Additionally, 100 random network designs for each simulation provide sampling distributions of the number of dry points sampled, from which the efficiency of each network design is inferred. The choice of the 'best' design depends on the criterion which it is most important to satisfy and on the spatial characteristics of the rainfall events across the study area.

KEY WORDS: Climatic maps · Isohyets · Rainfall fields · Network design

1. INTRODUCTION

In a recent article, Kay & Kutiel (1994) discussed the required density of rain gauges for adequate representation of the spatial distribution of rainfall events. Using a simulation of 10 different rain events (partially overlapping) and different densities of networks, they pointed out that a very dense network, on the order of 0.1 of the median event size, is required to depict and represent adequately the complexity of the rain distribution. However, even in the densest national networks (e.g. United Kingdom and Israel) the rain gauge densities are on the order of 0.8 to 1.0 of the median event size (i.e. 8 to 10 times less dense than necessary). Most national networks are not even close to these densities, if gauges are assumed to be uniformly distributed over the national area; in fact, distinct clusters and voids exist due to the nature of population distributions (Table 1). Comparison of the maps for a network density of 1.0 times the median event size, with the 'real' values (from the simulation), revealed that the interpolated spatial pattern was often totally mislead-

ing (Fig. 4 in Kay & Kutiel 1994). These results are in complete agreement with Sumner (1988), who noted the drawbacks of inadequate national networks.

Given that an 8- to 10-fold (at least) increase of the network densities is not realistic for most countries, we may ask if the way the network is arranged has any influence on the results. If the answer is positive, then we may ask what is the best arrangement to sample a certain area using the existing densities. That is, do some network designs seem preferable to a random distribution of gauges?

The purposes of the present study are 2-fold. First, we evaluate the importance of the network design to the interpolation of the rain field (both its pattern and the volume of rainfall it represents). Second, we identify network arrangements that may be better than others in sampling the rainfall field as adequately as possible.

2. METHODOLOGY

2.1. Rainfield simulation. The procedures for the simulations that follow are based on the methodology described by Kay & Kutiel (1994). We tested 6 different arrangements of 9 rain gauges each, to check which

*E-mail: kutiel@geo.haifa.ac.il

Table 1 Selected characteristics of some national or regional networks used in climatological studies. All values refer to an even network distribution

Country	Average area sampled by 1 rain gauge	Square size	Source
UK	50 km^2	~7 × 7 km	Sumner (1988)
Israel	60 km^2	~8 × 8 km	Kutiel (1993)
France	160 km^2	~13 × 13 km	Kutiel (1993)
Tunisia	260 km^2	~16 × 16 km	Kutiel (1993)
Finland	330 km^2	~18 × 18 km	Heino (1995)
Kansas (USA)	500 km^2	~22 × 22 km	Castleberry & Tucker (1995)
USA	576 km^2	~24 × 24 km	Linacre (1992)
Entre-Rios Province (Argentina)	750 km^2	~27 × 27 km	Barrera (1995)
India (excluding mountainous region)	9400 km^2	~97 × 97 km	Sontakke et al. (1993)
Ivory Coast	12300 km^2	~111 × 111 km	Lubès et al. (1995)
Algeria	100000 km^2	~317 × 317 km	Jellouli & Daget (1992)

best samples the rainfall field. The different arrangements were a regular square grid, a circle, a cross, an X-shape, a combination of a cross and an X, and a diamond (a different combination of a cross and an X). All arrangements had a sample point at the centre of the area; the remaining 8 points were placed about the centre in a regular pattern, according to the design. These 6 arrangements are just a very few of an infinite number of distributions of rain gauges. In most real networks, strict geometric patterns such as these are not realized. Therefore, we also used a Monte Carlo simulation of rain-gauge distributions to assess the results we infer for the 6 regular patterns.

The rainfall field was depicted as a set of nested elliptical isohyets representing different proportions of a maximum precipitation amount of 10 mm located at the centre of the ellipse. The eccentricity and location of the events were generated at random. We generated 4 such random fields (Fig. 1a) to gain a sense of whether our interpretations of the performance of the rain-gauge networks were dependent on pattern of precipitation. Our simulated 'rainfall fields' were just 4 from an infinite number of possible rainfall fields, but represented some distinctly different patterns of rainfall concentration. One rainfall field depicted a north-south arrangement of the rain events with a maximum in the centre (this field is identical to the one used in Kay & Kutiel 1994). In a second simulation, the events were clustered close to the centre of the area, without any rain on the perimeter. In a third simulation, the events were located around the perimeter of the area; no event was captured entirely within the area considered. The fourth simulation showed a southwest to northeast arrangement of most of the events. We used a dense, 81-point grid to estimate the values of rain produced by the 10 events in each simulation (Fig. 1b).

The dimensions of the study area were 1000 units by 1000 units. We prefer to use the generic 'units', in order to make the study applicable everywhere regardless of the various storm sizes produced by different mechanisms. This area of 1 000 000 square units was sampled by 9 rain gauges. The average area sampled by each gauge was 111 111 square units, which represents a density of approximately 0.60 to 0.75 of our median storm size (Table 2).

Each event contributing to a rainfall field was approximated by a cone. For a regular cone of radius a and height h the volume is

$$\frac{\pi}{3}a^2 h \tag{1}$$

Our cones have elliptical (horizontal cross-section) bases. Therefore, their bases (the 'storm size') were calculated as

$$ab\pi \tag{2}$$

and the volume of each as

$$\frac{\pi}{3}abh \tag{3}$$

and the total volume given by 10 events as

$$V = \frac{\pi}{3}\sum_{i=1}^{10} a_i b_i h_i \tag{4}$$

where V is the total water volume of the 10 events, a

Table 2. Median storm size and 'real' volumes of water over the entire area in each of the simulations

Simulation	Median storm size (square units)	Water volume (square units × mm)
1	176715	5154188
2	147262	4689482
3	147262	1360051
4	184569	3025609

and b are the major and the minor axes, and h is the maximum precipitation total of each event. As the major axis (a) and the maximum precipitation total (h) were kept constant in all our simulations, at 250 units and 10 mm respectively, we can rewrite Eq. (4) as

$$V = \frac{250 \times 10 \times \pi}{3} \sum_{i=1}^{10} b_i \qquad (5)$$

or

$$V \approx 2618 \sum_{i=1}^{10} b_i \qquad (6)$$

For the cases in which the entire event was not within the outlined study area, the portion of the total volume within the area was estimated. These volumes are listed in Table 2. We use these as the 'real' volumes for comparison of the estimated water volumes obtained from the interpolation scheme.

2.2. Evaluation of the rain field sampling. The 9-gauge networks were positioned on this grid, and the isohyet maps interpreted from the 9 values sampled. Isohyets were plotted using MacGRIDZO software, based on a moving weighted least squares (MWLS) algorithm (MacGridzo 1991, p 120–121). Our interpretation of the results is based on 3 criteria: resemblance of the spatial pattern of the isohyet map to the actual (dense network) rain field; similarity of estimated frequency of wet and dry points to expectation; and, agreement of estimated total water volume represented by the isohyet map with the known volume.

In the real world, rain gauges are not located in symmetric arrangements, but are distributed in irregular patterns across the area being sampled. The placement is not random; it is biased by the needs and distribution of the human population. Therefore, we wanted to compare the results from our regular network designs to the expectation for non-uniform arrangements. We created a Monte Carlo simulation, generating 100 randomly designed networks of 9 gauges for each of the 4 rain fields shown in Fig. 1. For each 9-gauge network, we counted the number of points that received rainfall and the number that were dry. The sampling distribution for the frequency of number of dry points, out of 9, may be modelled as a binomial, with the probability of being dry, p, calculated as the relative frequency of dry points in the dense, 81-point rain field.
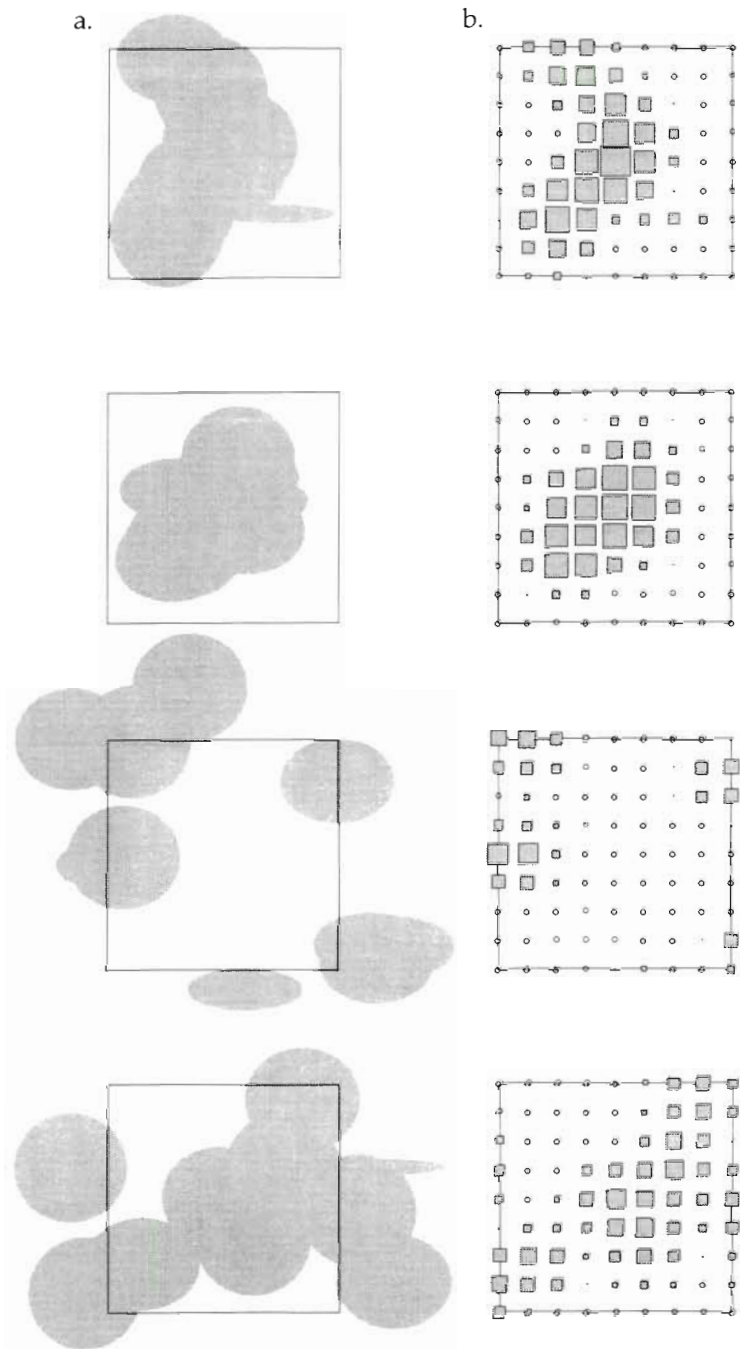

Fig. 1. The 4 different simulated rain fields, and the evaluated rainfall amounts on a rectangular grid. Squares sizes are proportional to the rainfall totals. Open circles are grid points without precipitation

We tested the fit to the binomial distribution as follows. From the dense (81-point) grid, we counted the relative frequency of dry points. We used this value as the probability of any point in the given rainfall field being dry, and its complement as the probability of a point receiving some precipitation. We then generated the theoretical expectation for 0, 1, 2, ..., 9 dry points in
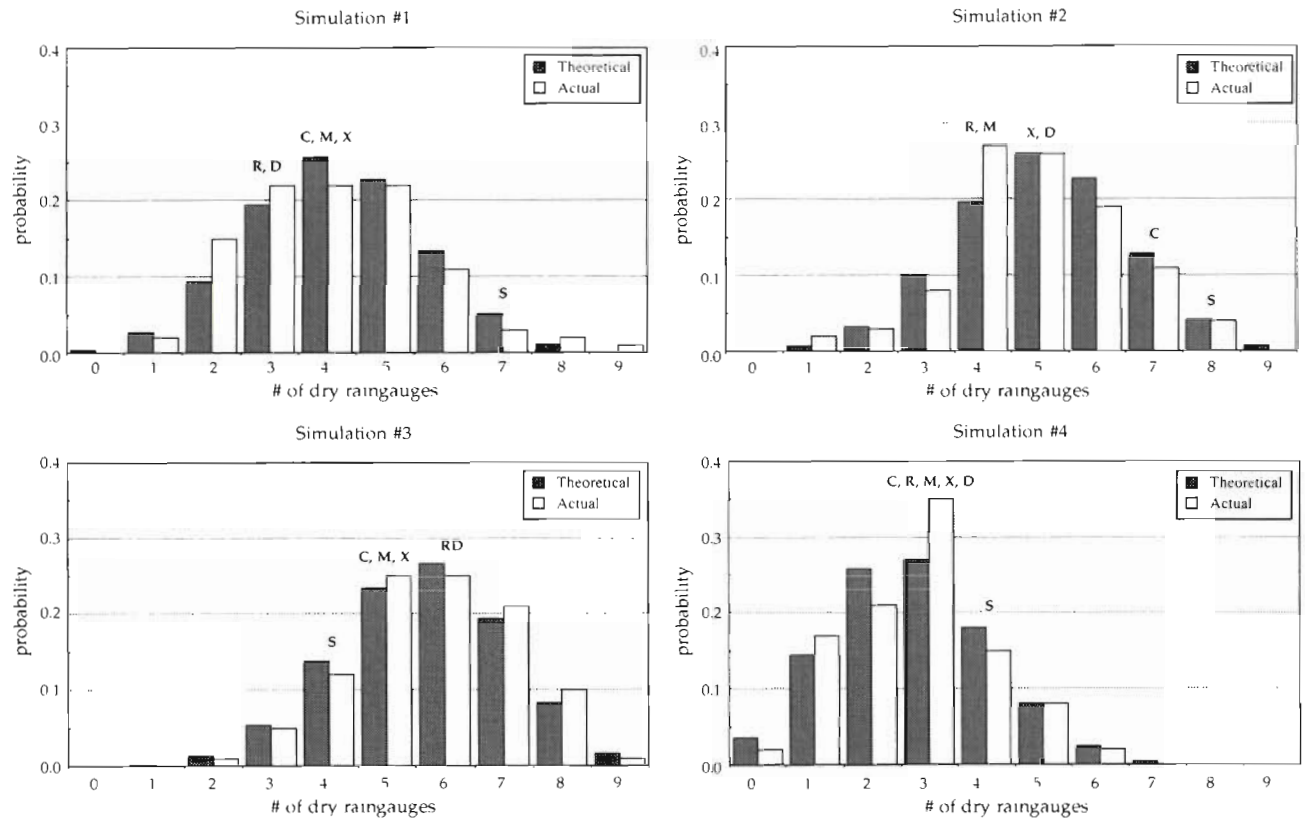
Fig. 2. Theoretical and actual (100 random cases) distributions of dry rain gauges in each simulation. The letters on each graph indicate the number of dry rain gauges in each of the symmetric arrangements: **S**quare, **C**ircle, c**R**oss, co**M**bined, **X**, and **D**iamond

a random sample of 9 points as a binomial probability. The fit of our sampling distribution (from the 100 random networks) to the theoretical binomial was assessed with a 1-sample Kolmogorov-Smirnoff test for the completely specified case. [In this application of the K-S test, the fit of the cumulative distribution function (cdf) for the actual network being evaluated is compared the theoretical cdf for the 100 random networks.] Only in the fourth simulation (the rainfield with southwest-northeast orientation) did the actual frequencies (from the 100 random designs) differ noticeably from the theoretical expectation. However, in all 4 simulations, there was no significant difference (at 0.05 probability) between the expected and actual distributions (Fig. 2). Therefore we can assess each regular network design, in terms of its efficiency in sampling the areas receiving rainfall, in a probabilistic framework generated as these binomial distributions.

3. RESULTS AND DISCUSSION

The first of our criteria for judging the best network design is the ability to produce an interpolated field

that looks like the actual rain field. Fig. 3 shows the isohyets for the 6 regular network designs (horizontally) for the 4 simulated rain fields (vertically). The dots in each diagram indicate the location of the rain gauges at which the rainfall was estimated. Fig. 4 shows the same, but for 6 (of the 100) random network arrangements. It is clear that, in the absence of any other information, the interpolated pattern of isohyets can vary markedly, depending on the distribution of gauges.

On further examination of Figs. 3 & 4, there appears to be a greater resemblance among the isohyet patterns generated by the same network arrangement regardless of actual rainfield pattern (i.e. moving vertically in Figs. 3 & 4) than in isohyet patterns for the same rainfield regardless of network design (moving horizontally). Some examples may be cited for illustration. For Simulation 1, the rainfall field has a NE-SW orientation in the interpolations from the X and Diamond networks, a N-S orientation from the Cross and Combined networks, and appears as a bulls-eye from the Square network. A similar interpretation applies for Simulation 2, with the exception that the Cross and the Combined networks show a noticeable

E-W orientation of the rainfall field (Fig. 3). When the random network designs are considered, the maxima of the rain fields may not even occur in similar locations from one random map to another (Fig. 4). The absence of a fixed gauge at the centre of the study area accounts for this behaviour on the maps from the random networks.

The second criterion for judging 'best' is the ability to provide a representative sampling of wet and dry areas. On the probability distributions for number of dry gauge sites, we have also shown the location for each of the 6 regular network arrangements (Fig. 2). We would expect that a network arrangement, if it were 'representative', would have the number of dry rain gauges with the highest probability. The best arrangement according to that criterion will be an arrangement having 4 dry rain gauges in the first simulation (probability of 0.2574), 5 in the second (0.2594), 6 in the third (0.2660), and 3 in the fourth (0.2698). Although the 'best' network design varies from simulation to simulation, we note that the X pattern often matches, or nearly matches, the maximum expectation for number of dry points. The Square grid usually has many more (or, in Simulation 3, fewer) dry gauges than expected.

To provide an overall index of the performance of each network arrangement, we calculated the average ratio of the probability for the number of dry gauges to the maximum probability. For example, the Square design in the first simulation had 7 dry gauges; there is a probability of 0.0507 for this number of dry gauges. The most probable number of dry gauges was 4, with a probability of 0.2574. The ratio, therefore, was 0.20. The average of such ratios for all 4 simulations was 0.39. In fact, this average ratio was the lowest of the 6 regular network designs (Table 3). Most networks had ratios over 0.80; the X network was best, and the Diamond was a close second.

Our third criterion for judging the 'best' pattern is the ability to estimate volume of precipitation. In

Table 3. Average ratio of probability for number of dry gauges to maximum probability, in each of the network arrangements

Arrangement	Ratio	Arrangement	Ratio
Square	0.39	Cross	0.88
Circle	0.85	X	0.97
Combined	0.91	Diamond	0.94

nearly 80 % of the comparisons, the interpolated field overestimated the actual volume of rain in the simulation (Table 4). Estimations ranged between just one-half of actual volume (the X network for Simulation 3) to nearly twice the volume (the Square for Simulation 3). An index of the overall performance of each network arrangement may be the average deviation (as percentage of actual). By this index, the Diamond network appears best, with the Circle and the Cross close behind. Closer examination indicates that the Cross network provides overestimation for all 4 simulations, whereas all other patterns underestimate either Simulation 2 or 3. Further, the mean absolute deviation is smallest for the Cross.

For each simulation, the 100 random networks provided a sampling distribution of estimated volumes. When these volumes are expressed as proportions of the known value, the sampling distribution is approximately normal, according to the central limit theorem. This property allowed us to assign probabilities to each volume, based on the standard normal distribution and using the mean and standard deviation from each set of 100 random networks (Table 4). Although all designs for all simulations fail to estimate true values (1.00) very closely, only one (pattern X in Simulation 4) exceeds a 0.05 probability of being so extremely different that it might be judged not to have occurred by random sampling error. Thus, we cannot say that any one design, for any simulation, is worse than any random arrangement.

Table 4. Estimated volumes of water, as proportions of known volume for each simulation, and the normal probability (in parentheses) of a value at least that extreme occurring solely due to random sampling error

| Simulation	Square	Circle	Cross	Combined	X	Diamond	Mean for 100 random networks
1	1.22 (0.26)	1.28 (0.22)	1.31 (0.20)	1.43 (0.13)	1.61 (0.06)	1.43 (0.13)	0.95
2	0.92 (0.44)	0.72 (0.24)	1.42 (0.18)	1.61 (0.09)	1.31 (0.24)	1.13 (0.38)	0.99
3	1.84 (0.05)	1.55 (0.12)	1.04 (0.37)	0.80 (0.48)	0.51 (0.30)	0.79 (0.48)	0.83
4	1.28 (0.21)	1.39 (0.12)	1.16 (0.34)	1.19 (0.30)	1.62 (0.03)	1.51 (0.06)	1.03
Mean absolute deviation (%)	36	38	23	36	51	32	
Mean deviation (%)	+ 32	+ 24	+ 23	+ 26	+ 26	+ 22	−5

Square   Circle   Cross
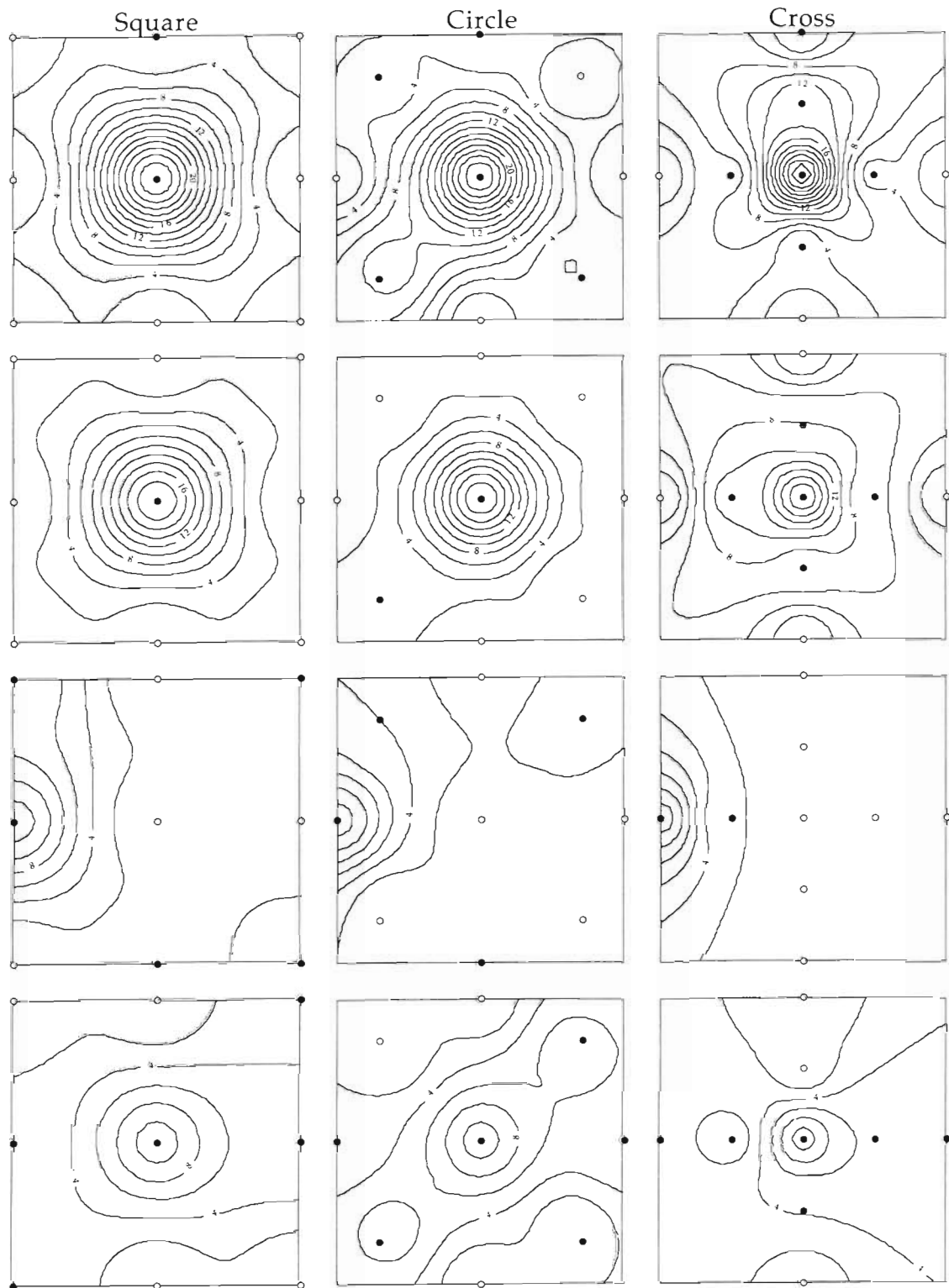
Fig. 3. Rainfall fields obtained using MacGRIDZO, for each of our 4 simulations in the various arrangements. The circles in each diagram indicate the location of the rain gauges for which the rainfall was estimated (open: no rain; filled: a measurable amount of rainfall)
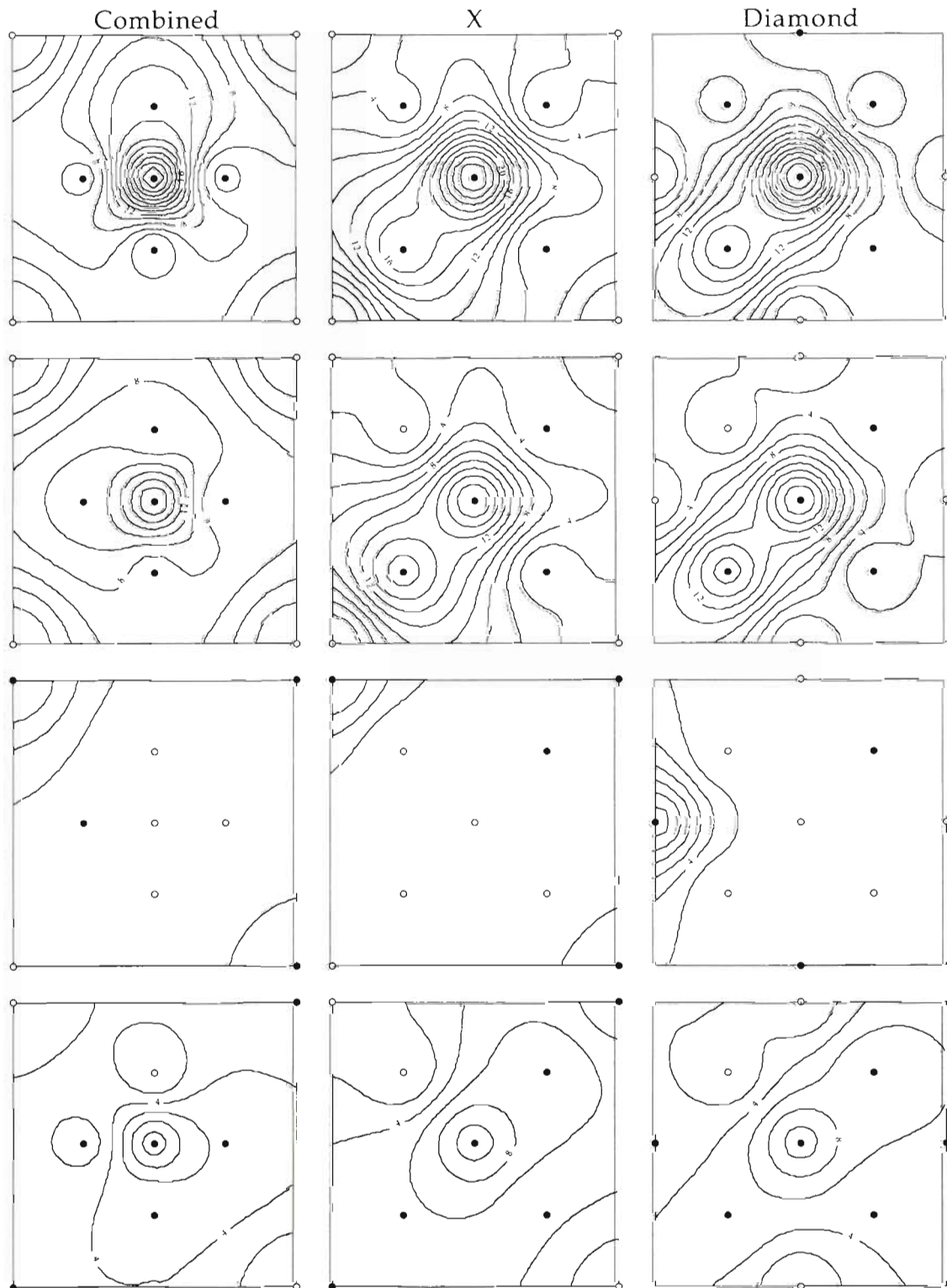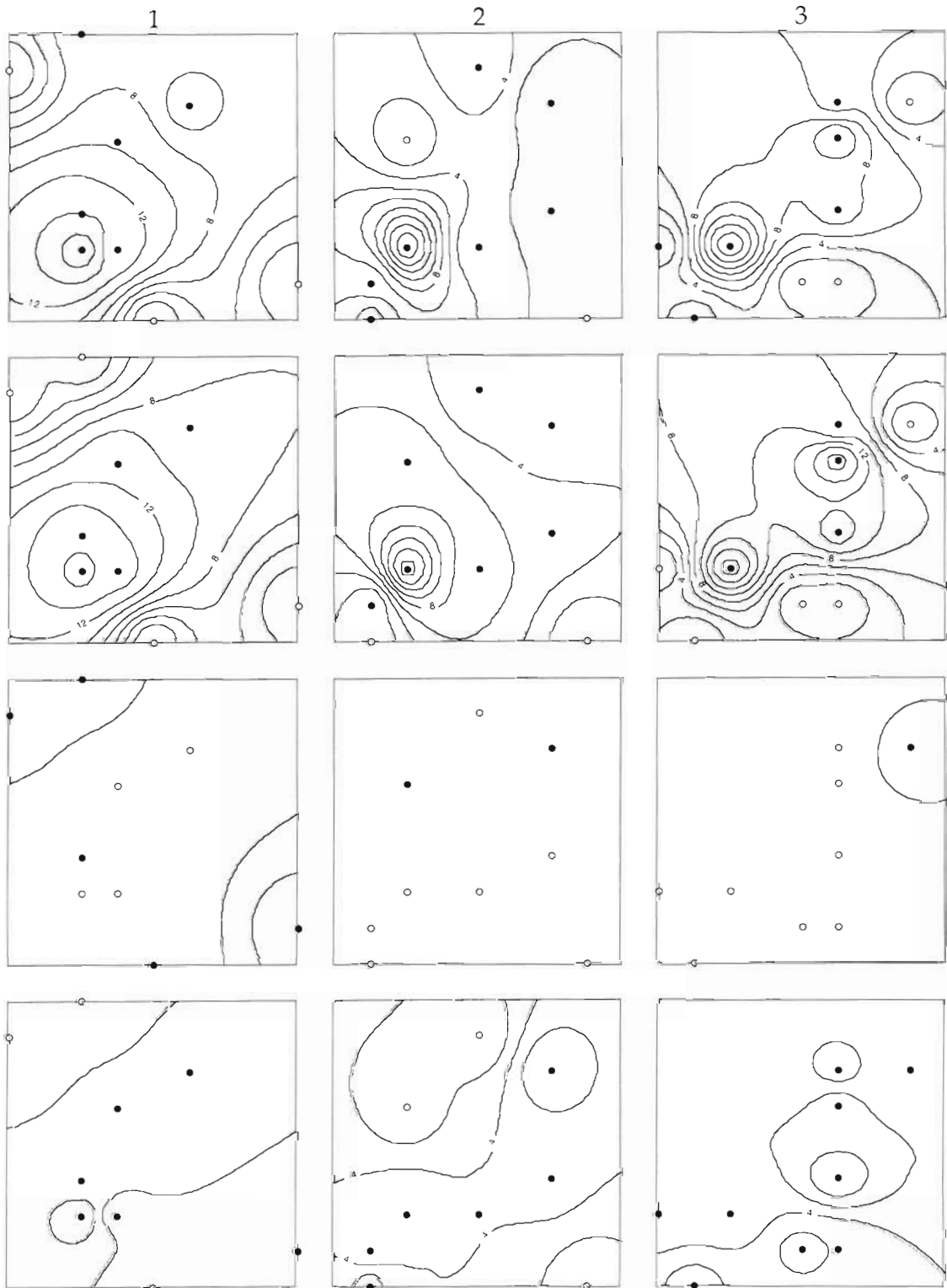
Fig. 3 (continued)

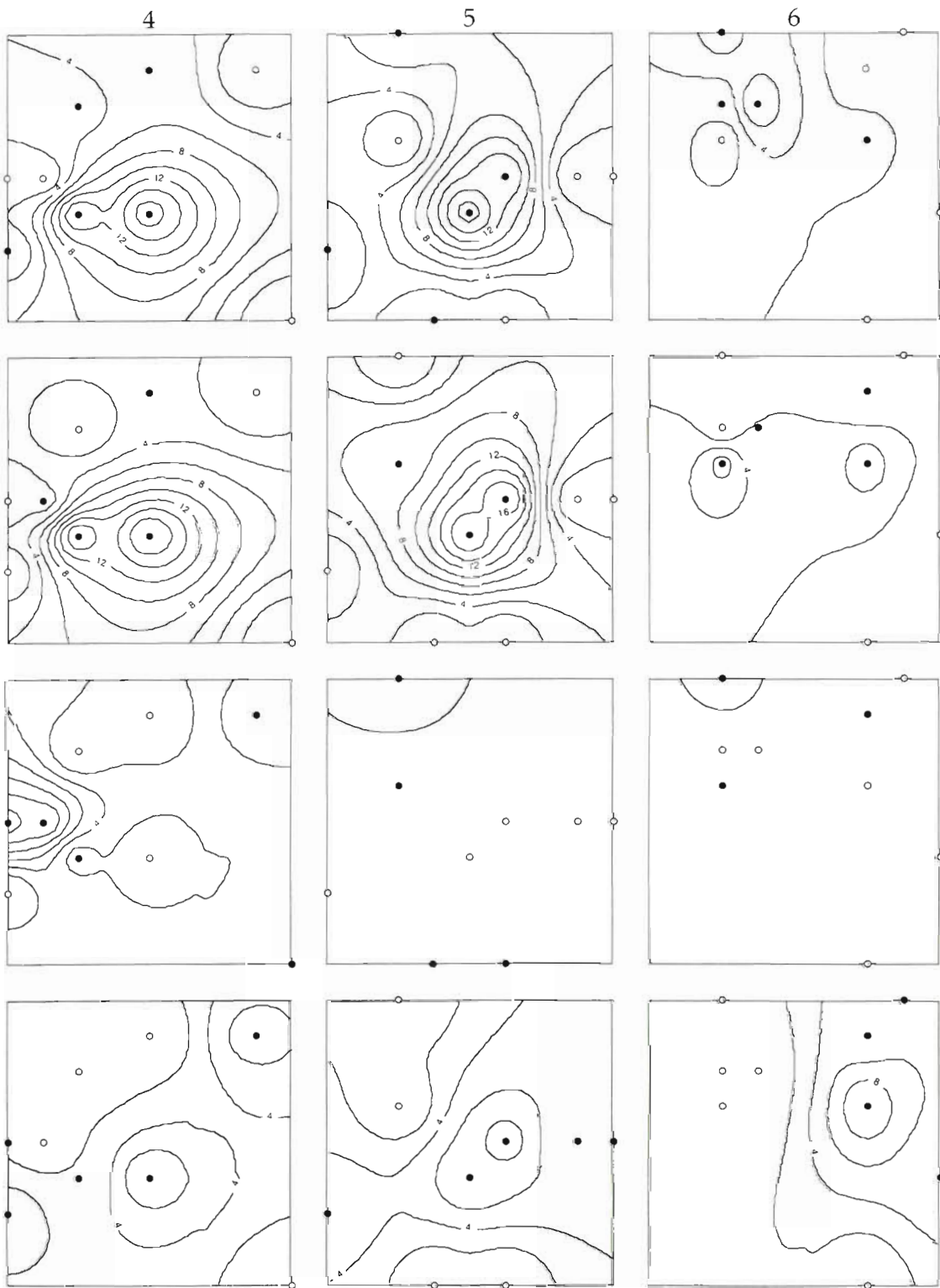Fig. 4. Rainfall fields as in Fig. 3, but for 6 random arrangements

Fig. 4 (continued)

4. CONCLUSION

The main implication of this study is that the arrangement of the network, and not only its density, has a major influence on the maps that one might generate for a series of events. This is true for all arrangements of gauges, symmetric or random. The network's spatial geometry may influence the pattern of the interpolated field, the efficiency of sampling wet and dry areas, and the estimation of total rainfall volume. Moreover, the results of our simulations suggest that the 'errors' may be quite large. For example, volumes were on average overestimated by more than 20%. Some reassurance may be found in the fact that the regular arrangements were, by probabilistic considerations, no worse at estimating volume than any random arrangement. However, without prior knowledge of the nature of the rainfield, there appears to be little guidance from geometry alone to the choice of a network arrangement. Therefore, great care is needed in the interpretation of data from existing networks, and in the design of new networks. Our simulation studies so far yield no consistent recommendation that one or another network arrangement is best for all purposes. One therefore should attempt to utilize available evidence that suggests the theoretical distribution of rainfall. Such evidence may be found in vegetation patterns or in topography (orographic and rain-shadow effects). One should also be clear as to the main purposes to which the network data will be put. It may not be possible to satisfy simultaneously both pattern and volume estimation criteria, without a much denser network than is usually feasible.

Editor: V. Meentemeyer, Athens, Georgia, USA

LITERATURE CITED

Barrera D (1995) On the connection of precipitation systems with spatial correlation fields in the province of Entre-Rios. Proc 6th International Meeting on Statistical Climatology, 19–23 June 1995, Galway, Ireland. University College, Galway, p 371–374

Castleberry PS, Tucker D (1995) The identification and verification of precipitation fields in the Eta model using multivariate randomized block permutation procedures. Proc 5th International Conference on Precipitation Space-Time Variability and Dynamics of Rainfall. June 14–16, 1995, Elounda, Crete, Greece, 3.2. University of Minnesota

Heino R (1995) Changes of precipitation characteristics in Finland. Proc 6th International Meeting on Statistical Climatology, 19–23 June 1995, Galway, Ireland. University College, Galway, p 359–362

Jellouli Y, Daget P (1992) Etude Climatique du Sahara Algerien. In: Perard J, Dalby M (eds) Publications de l'Association Internationale de Climatologie, Vol 5, Contraintes geoclimatiques et environnement tropical. Université de Bourgogne, Dijon, p 235–248

Kay PA, Kutiel H (1994) Some remarks on climatic maps of precipitation. Clim Res 4:233–241

Kutiel H (1993) L'incertitude du régime pluviométrique en Méditerranée orientale. In: Maheras P (ed) Publications de l'Association Internationale de Climatologie, Vol 6, La sécheresse en Méditerranée et dans les pays environnants. Université Aristote de Thessaloniki, p 151–158

Linacre E (1992) Climate data and resources. Routledge, London

Lubès H, Paturel JE, Servat E, Masson JM, Kouamé B (1995) Analysis of annual rainfall series in Ivory Coast in order to detect climatic variations. Proc 6th International Meeting on Statistical Climatology, 19–23 June 1995, Galway, Ireland. University College, Galway, p 455–458

MacGridzo (1991) The Contour Mapping Program for the Macintosh (version 3.3). RockWare-Earth Science Software, Wheat Ridge, CO

Sontakke NA, Singh N, Pant GB (1993) Optimization of the rain gauges for a representative all-India and subdivisional rainfall series. Theor Appl Climatol 47:159–173

Sumner G (1988) Precipitation: process and analysis. John Wiley and Sons, Chichester.