AREAL RAINFALL ESTIMATION, FORECAST AND SIMULATION OVER SMALL WATERSHEDS SUBJECT TO MAJOR FLASHFLOODS.

T. LEBEL AND D. CREUTIN

Institut de Mécanique de Grenoble
GRENOBLE, FRANCE

1. INTRODUCTION

A good knowledge of areal rainfall for small time steps (from one to several hours) is needed for flood forecasting for small watersheds. The scarcity of recording raingages as compared to daily raingage records makes it necessary to use a robust method for the calculation of areal rainfalls. Calculation of estimation standard-deviation allows for a better apreciation of model inputs errors. Thus a good estimation of areal rainfalls is necessary but not sufficient for real time flood forescasting.

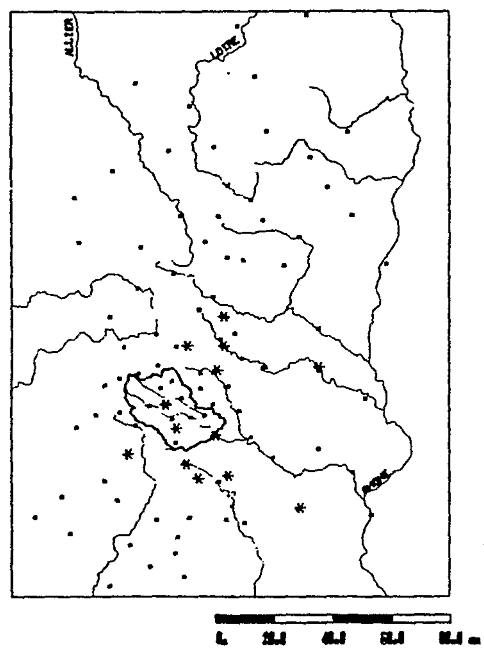
It is necessary to be able to anticipate the rainfall for small watersheds because of the short response time. The determination of analogous situations (DUBAND,1980) is not yet useful for small time steps, therefore we would like to know if there are systematic pattern or preferential trajectory in the dynamics of strong rainfalls.

The lack of results leads us to elaborate a stochastic generator of areal rainfalls taking into account the rainfall already observed.

2. ESTIMATION OF AREAL RAINFALL

2.1. REGION AND WATERSTREDS CONCERNED

Among the different methods which give equivalent results in the evaluation of accumulated daily rainfall (CREUTIN and OBLED 1980), the spline method showed the advantage of simplicity at a low cost. We have therefore compared the spline's results with these given by a classical method (thiessen weighting) in the computation of a set of hourly areal rainfalls on watersheds located on the southeast side of massif central (France). These watersheds are subject to heavy rainfall occurencies which can locally reach 100 mm in one hour (Villefort ; september 1980) ; the region consists roughly of a rectangular area of 150 * 200 kms, which is equipped with 97 recording raingages. The location of the gages is quite irregular and three watersheds have a particularly dense network : the CEZE, the GARD, the VIDOURLE (fig. 1.1).



* Telemeasured station

Figure 1.1

RECORDING RAINGAGES NETWORK.

Two watersheds were chosen in order to perform the comparison :

- The gardon at Anduze (545 km²) flows downstream of two deeply embanked and parallel valleys. Twelve recording raingages have been installed among which 3 are tele measured.

- *Vidourle at quissac (213 km²) lies on the mountain side and is under wide mediteranean influence. It is equipped with 3 recording raingages from which 2 are telemeasured.*

The Vidourle was divided into two subwatersheds and the gardon in four parts: the subareas range between 50 and 500 km² (fig. 2.1).

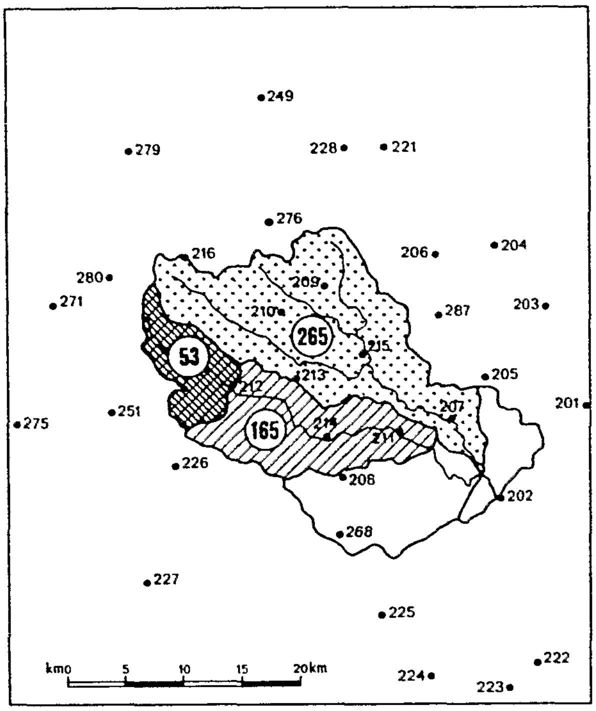


Figure 2.1

RAINGAGES NETWORK AND SUBWATERSHEDS OF GARDON D'ANDUZE (TOTAL AREA : 545 Km²)

For any method, the mean areal rainfall is given by :

$$\hat{P}_s = \sum_{i=1}^{n} \lambda_i Z_i \qquad 2.1$$

where Zi is the rainfall at station i and $\Sigma \lambda = 1$*. However each interpolation method has a different way to compute* λ_i

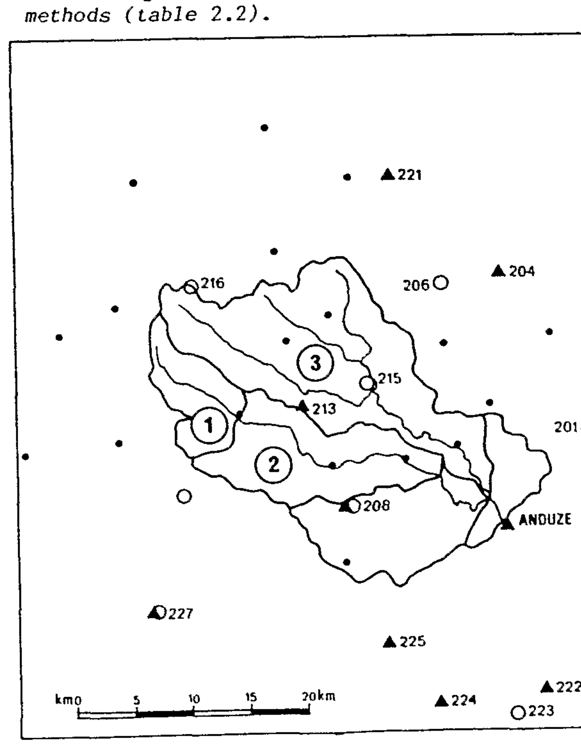
2.2 Complete network :

Spline surface were fitted on a set of 1172 hourly rainfalls (1972-1980) using the complete 97 stations network. The rainfall event occured in autumn which is the season of exceptional waterflood. These surfaces were integrated over the six subwatersheds each discretised with a one square kilometer mesh. The results were correlated with those obtained from the thiessen method (table 2.1).

Table 2.1 : correlation between spline estimate and thiessen estimate on the complete network. Only rainfall of more than 1 MM/H. have been considered.

WATER SHED (Surfaces in km²)	r²	N NUMBER OF COUPLES
GARDON ST ANDRE (53 km²)	.92	686
GARDON ST JEAN (165 km²)	.98	742
GARDON MIALET (265 km²)	.98	738
GARDON ANDUZE (545 km²)	.998	615
VIDOURLE ST HYP.(48 km²)	.96	721
VIDOURLE QUISSAC(213 km²)	.96	695

Both methods give similar results for watersheds with an area of more than 200 km². Where small watersheds are concerned, discrepancies are more noticeable, particulary for heavy, rainfall, but the correlations remain larger than .95. This is well explained by the different weights which were assigned to each station for the two methods (table 2.2).



▲ Telemeasured station ○ Long term station.

① GANDR (53 Km²) ③ GAMIA (265 Km²)

①+② GAJEA (165 Km²) ④ GANDU (545 Km²)

Figure 2.2

SUBNETWORKS ON GARDON D'ANDUZE WATERSHED. A FEW STATIONS ARE LOCATED OUTSIDE OF THE MAP.

Table 2.2 : Comparison of spline and thiessen weighting for two watersheds.

GARDON ST ANDRE (GANDR : 53 km²)			GARDON ANDUZE (GANDU : 545 km²)		
N°	THIES. WEIGHT	SPLINE WEIGHT	N°	THIES. WEIGHT	SPLINE WEIGHT
212	.698	.633	212	.128	.134
251	.038	.219	208	.106	.116
216	.132	.181	207	.097	.084
280	.094	.152	210	.083	.083
213		-.120	211	.083	.081
210		.101	202	.075	.079
279		-.053	215	.083	.073
275		-.046	209	.077	.064
226	.038	.041	216	.059	.057
276		-.034	213	.061	.050
			214	.075	.049
209		-.030	268	.018	.033
214		-.027	226	.018	.027
227		-.024	251	.037	.019
271		-.017	280	.055	.018
211		.012	205		.018
208		-.004	287	.015	.015
268		-.003	276	.009	.014
248		-.003	279		-.010
225		.003	266		.009
207		-.003	225	.002	.008
			275		-.006
			227		-.005
			204		-.005
			222		-.005
			271		-.004
			224		-.004
			223		-.003
			201	.002	.003

2.3 LESS DENSE NETWORKS.

The high cost of telemeasured stations an fairly frequent failures with the others (especially during heavy rainfall) raises the problem of working with incomplete networks. The sensitivity of interpolation methods to the density of the measuring network is therefore important.

It is particurlarly useful to concentrate the study on the following networks:

- telemeasured network (13 on upstream watersheds) installed for flood forecasting.

- long terme stations (more than 15 years of measurement) which are to be used for the temporal distribution analysis of areal rainfall (LEBEL, GUILLOT 1983).

Using the same sample of 1172 hourly rainfalls, new estimates of the areal rainfall were computed by using on the one hand the telemeasured network, on the other hand the network composed of the 13 long term stations. We have correlated the four estimators (two methods, two networks) with the spline estimation usin the complete network, taken as a reference (table 2.3).

It is noteworthy that the spline estimate turns out at least equivalent but more often better than the thiessen estimate. Differences can be low but the Thiessen method may give bad estimate for high values because for small watersheds, this estimate relies on a single station. Any peculiar measure in the phenomenon, or a measurement error fully affects the estimated areal rainfall for these watersheds. Figure 2.3 shows that, by using "non zero" weight coefficients for adjacent stations, the spline method is able to retain more information than the Thiessen estimate.

Table 2.3 : Correlation be tween spline estimation of areal rainfall on a dense network and different estimations on subnetworks of 13 stations.

	SPLINE 97 THIESSEN 97	SPLINE 97 THIES.-LON 13	SPLINE 97 SPLINE-LON 13	SPLINE 97 THIES.TEL 13	SPLINE 97 SPLINE TEL 13
GANDR (53 km²)	.96	.58	.59	.31	.43
GAJEA (165 km²)	.99	.75	.78	.72	.74
GAMIA (265 km²)	.99	.86	.86	.51	.65
GANDU (545 km²)	.999	.89	.89	.87	.89
VIHYP (48 km²)	.98			.71	.72
VIOUI (213 km²)	.98			.96	.96

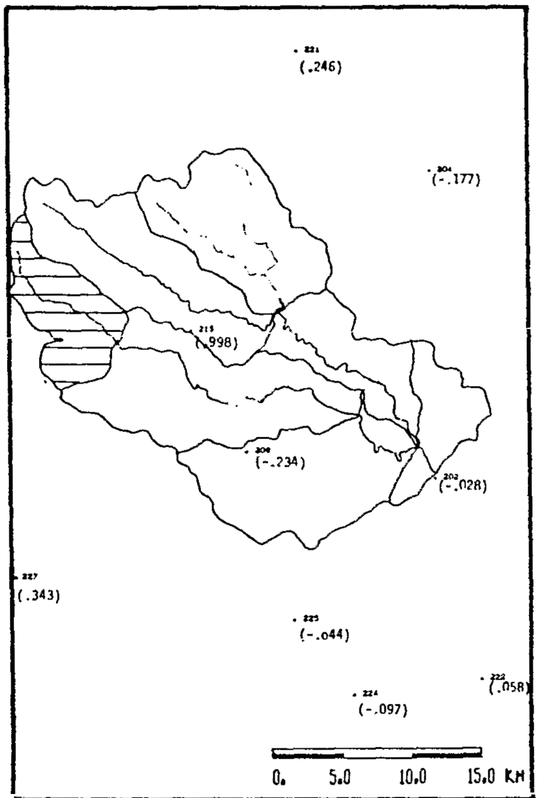


Figure 2.3

SPLINE WEIGHTING USING TELEMEASURED NETWORK FOR GARDON ST ANDRE (53 km²). THIESSEN METHOD GIVES A WEIGHT OF 1.0 FOR STATION N° 213 AND 0.0 ALL THE OTHERS.

3. AREAL RAINFALL STANDARD DEVIATION

3.1 Statement of the problem.

As shown previously, some mean rainfall estimators can yield certain unrealistic results. Calculation of the standard deviation allows for an evaluation of the error; they can also be used to vary the input of flow modeling.

Let $\hat{P}$ be the mean areal rainfall given by 2.1 and P the true mean areal rainfall.

The estimation variance σ_E^2 can be written.

$$\sigma_E^2 = \sigma^2 (\hat{P}_S - P_S) \qquad 3.1$$

$$\sigma^2 = \sigma^2 (\hat{P}_S) + \sigma^2 (P_S) - 2 \operatorname{cov} (\hat{P}_S, P_S) \qquad 3.2$$

$$\sigma^2 (\hat{P}_S) = \Sigma\Sigma \lambda i \lambda j \operatorname{cov}(Zi Zj) \qquad 3.3$$

$$\sigma^2 (Ps) = 1/s^2 \int_S \int_S \operatorname{cov}(Z(M)), Z(M'))dMdM' \qquad 3.4$$

$$\operatorname{cov}(\hat{P}_S P_S) = \frac{1}{S} \Sigma\lambda i \int_S \operatorname{cov}(Zi,\ Z(M))\ dM \qquad 3.5$$

The calculation of 3.2 requires an estimate of the covariance between every couple of points belonging to the network or to the integration surface, respectively (Zi, Zj) and Z (M), Z (M').

A classical way of performing this estimation, especially when the estimation of the covariance is biaised by bad knowledge of the field mean value, may be to use the variogram.

$$\gamma(h) = \tfrac{1}{2} \operatorname{Var} (Z(M) - Z(M + h)) \qquad 3.6$$

where h is a class of interdistance.
It is shown (JOURNEL and HUIJBRECHTS 1978):

$$\gamma(M,M') = \lim_{S \to \infty} \sigma_S - \operatorname{cov}(M,M') \qquad 3.7$$

using 3.7 in 3.3, 3.4, 3.7, 3.2 can be written :

$$\sigma_E^2 = \frac{2}{S} \Sigma\lambda i \int_S \gamma(Mi,M)dM - \Sigma\Sigma\lambda i\ \lambda j\ \gamma(Mi,Mj) - \frac{1}{S^2} \int_S \int_S \gamma(M,M')d_M\, d_M' \qquad 3.8$$

Integrals in 3.8 are calculated by using a discretised surface of M elementary cells.

$$\sigma_E^2 = \frac{2}{M} \sum_{i=1}^{n} \lambda i \sum_{j=1}^{M} \gamma(Mi,Mj) - \sum\sum^{nn} \lambda i \lambda j\ \gamma(Mi,Mj) - \frac{1}{M^2} \sum\sum^{MM} \gamma(Mi,Mj) \qquad 3.9$$

Formula 3.9 is valid whatever the linear estimate $\hat{P}$.

It must be stressed that, although a good estimation of areal mean may be obtained without any assumption on the spatial structure of the phenomenon (see results of table 2.1), for the computation of standard deviation, a model of the covariance function is needed.

3.2 Choice of a variogram model.

Several classes of model may be chosen but it was assumed that any class of model must fit all events. This requires a scaling parameter, which has already been proposed in the litterature (DELHOMME 1978, LORENT & AL 1982). It is based on the strong correlation (.917) between the mean and variance within each seperate field.

This leads to :

$$\gamma(h) = \alpha.\ \gamma_S(h) \qquad 3.10$$

Where γ_S is a scaled variogram while α is related to the particular field variance. Furthermore :

$$\gamma_S(h) = \gamma_S(\beta, h) \qquad 3.11$$

require at least a "shape" parameter.

Two classes of models were selected which cover a wide range of possible variogram shapes :

- spherical models for a finite variance:

$$\gamma(h) = \emptyset \qquad h = 0$$

$$\gamma(h) = \alpha\left|(1,5\,\frac{h}{dp} - 0,5\,(\frac{h}{dp})^3\right| \quad 0 < h < dp \qquad 3.12$$

$$\gamma(h) = \alpha \qquad h > dp$$

where α is the SILL

dp is the RANGE

no nugget effect.

Here dp is the shape parameter and a function of correlation can be calculated.

- Power law models :

$$\gamma(h) = \alpha\, h^{\beta} \qquad 3.13$$

The problem consists in identifying β for the whole set of events as well as the α k associated with the kth event. Two different methods were used to fit the theoretical variogram :

1 - Least square fitting of the experimental variogram.

2 - Minimisation of interpolation error by cross validation of the observed values.

It consists in determining by trial and error or by a simple gradient method which value of β minimises the reconstitution in term of mean square error V(β) over all stations (BASTIN AND GEVERS 1983):

$$V(\beta) = \frac{1}{K}\sum_{k=1}^{K}\ \frac{1}{N}\sum_{j=1}^{N}\ (\hat{Z}_{jK}(M) - Z_{jK}(M))^2 \qquad 3.14$$

Where $\hat{Z}_{jK}(M)$ *is the estimate of* $Z_{jK}(M)$ *given by kriging on the (N-1) other stations of the network.*

- *N is the number of stations*
- *K is the number of events.*

The criteria V (β) is independant of the scale parameter α .

3.4. Application to the gardon d'Andule watershed.

For the identification of the variogram, 34 stations were considered, as shown on figure 2.1, together with the 103 heaviest hourly rainfall field.

The reduction to 34 stations has been made because the weighting coefficient given in table 2.2, show that it is sufficient to take into account the nearest stations to perform a correct estimation of the hourly rainfall over these watersheds.

- Least square method.

The modelling on two sub samples (49 and 54 events) which corresponds to two distinct types of meteorological situation, leads to the results given in table 3.1.

TABLE 3.1 : shape coefficient for two models estimated by L.S method for sub SAMPLES.

	SPHERICAL		POWER TYPE	
	RANGE (KM)	NUGGET	β	NUGGET
SAMPLE N°1	30	0	.69	0
SAMPLE N°2	25	0	.68	0

Parameter α must be recalcuted for each field. The search for priviledged directions was unsuccesful. Thus we make the assumption of an isotropic process. The nugget effect depends on the chosen class division. It seems to be negligible for a number of classes ranging from 10 to 15.

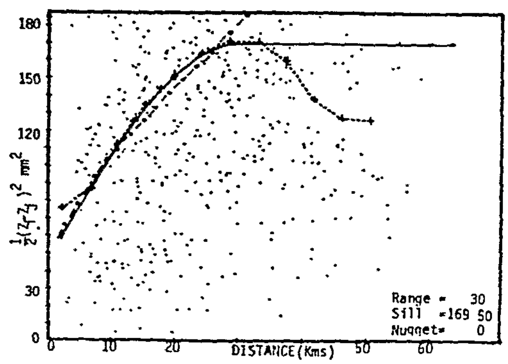


Figure 3.1
Climatological variogram for A subsample of 49 strong events.
Full line : spherical model
power law model
(BETA = .69)

Fitting are very similar for distances less than 25 Kms.

- MIE method :

The computation of 3.14 for every field and various values of β is lengthy. This is the reason why we have considered only the twenty strongest fields since V (β,K) is proportional to the variance of field. Letting β vary for the two classes of model, we seek β such that V (β) be minimum (figure 3.2).

Considering every model belonging to the same class, both criteria (LS and MIE) lead to similar values of β.

On table 3.2 the values of V (β) are very similar at the minimum for both classes of model (90,3 and 90,5), although the

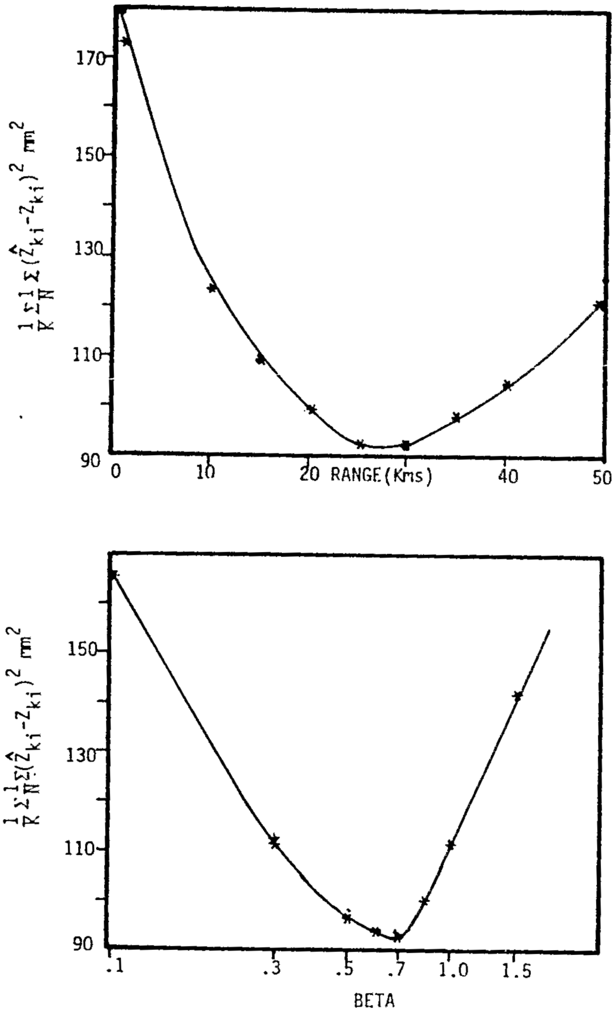


Figure 3.2

MIE METHOD :

UP : SPHERICAL MODEL
DOWN : POWER LAW MODEL.

behaviour of the two classes of model are opposite for infinite distances. This is due to the fact that the interdistances are generally smaller than 30 kms all over the network. In the following, we have selected the class of spherical models for two reasons :

- the existence of an associated model for the correlogram. This can be proven to be useful in various applications (LEBEL - GUILLOT 1983).

- Parameter α is easily estimated. In a spherical model, this parameter represents the sill. When the range is known, one can show (DELFINER and DELHOMME 1975) that the field variance is an unbiased estimator of the sill.

3.5 Calculation of standard deviations :

On the gardon d'Anduze (545 km²) standard deviations were calculated for the 103 fields by using three spherical models (RANGE = 20 ; 28 ; 30). They vary between 6 and 15 % of the estimated areal rainfall when using the dense network (34 stations).

We did the same, using the telemeasured network. In figure 3.3, confidence intervals are shown ($P_{13} - \sigma_E$, $P_{13} + \sigma_E$). We can see that, two times out of three, the rainfall estimated with the 34 stations network, belongs to the confidence interval.

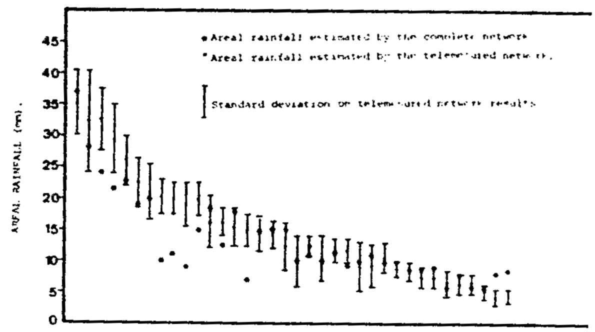


Figure 3.3

COMPARISON BETWEEN 34 STATIONS ESTIMATION AND 13 STATIONS ESTIMATION (TELEMEASURED NETWORK) FOR GARDON D'ANDUZE WATERSHED

TABLE 3.2 : VALUES OF V (β) FOR SPHERICAL AND POWER LAW MODELS.

SPHERICAL	RANGE (β)	1	10	15	20	25	30	35	40	50
SPHERICAL	V (β)	170,3	123,5	109,1	99,0	91,4	92,9	99,1	105,5	126,2
POWER LAW	β	.1	.3	.5	.6	.7	.8	.9	1.0	1.5
POWER LAW	V(β)	159,8	112,6	96,8	94,2	93,5	98,6	105,0	111,4	146,5

4. FORECAST AND SIMULATION

The computation of areal rainfalls over small watersheds (18), allows for the study of concomittances by cross correlation. No classification rain event by rain event could be seen. It does appear neither preferential trajectories nor systematic pattern of the rainfall cells inside watersheds. For the two watersheds under consideration, the results of the cross correlation, rain event after rain event, are shown in table 4.1.

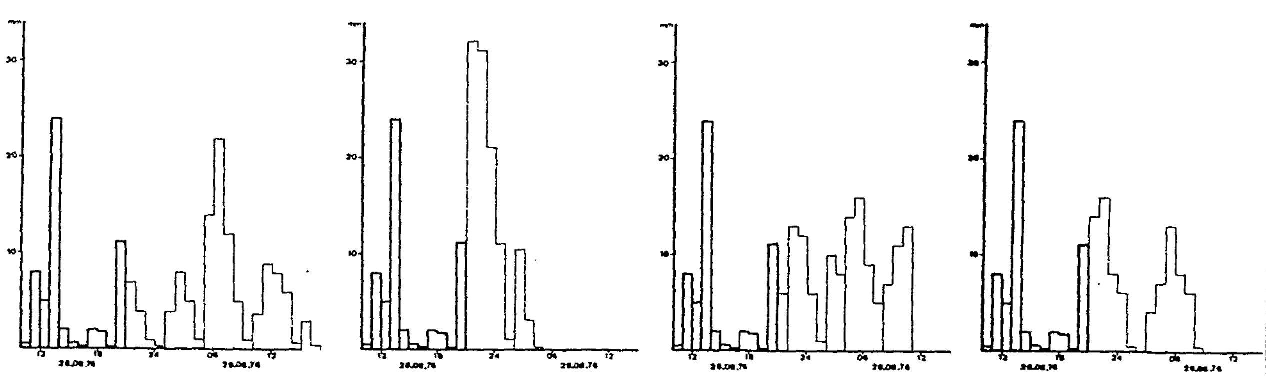
Table 4.1 : Correlations between areal rainfall above 1 MM/H for the neighbouring watersheds, rain event by rain event. Each rain event is made up of 48 to 160 consecutives rain hours.
r is the correlation between simultaneous hourly rainfalls over two watersheds.
r * is the optimum correlation obtained by shifting the rainfall hours for one watershed (cross correlation). The lag is the shiting number of hours.

BULK VALUE CORRELATION

	EVENT N°	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Watersh.1	r	0,75	0,72	0,36	0,16	0,56	0,51	0,80	0,66	0,29	0,54	0,60	0,76	-0,06	+0,07	0,64	0,60
Watersh.2	r*	0,75	0,72	0,36	0,4	0,56	0,51	0,80	0,69	0,29	0,54	0,60	0,76	0,83	0,07	0,69	0,88
Time lag	variation	0	0	0	-1	0	0	0	-1	0	0	0	0	-2	0	-2	-1
Watersh.1	r	0,67	0,22	0,12	-0,28	-0,07	0,22	0,41	0,53	0,08	0,68	0,08	0,37	0,92	-0,55	0,18	+0,39
Watersh.3	r*	0,67	0,42	0,26	0,60	0,31	0,48	0,58	0,54	0,14	0,58	0,53	0,37	0,92	<0	0,18	0,23
Time lag	variation	0	-2	+1	-2	+1	+1	-2	-1	+1	0	-3	-1	0	0	0	-1
Watersh.2	r	0,97	0,77	0,72	0,76	0,46	0,82	0,78	0,72	0,39	0,70	0,46	0,69	-0,35	-0,26	0,05	0,82
Watersh.3	r*	0,97	0,77	0,72	0,81	0,54	0,82	0,78	0,72	0,71	0,70	0,69	0,69	0,88	-0,26	0,22	0,82
Time lag	variation	0	0	0	-1	+1	0	0	0	+2	0	-3	0	-2	0	-1	0

RANK CORRELATION.

	EVENT N°	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Watersh.1	r	0,59	0,51	0,31	0,09	0,35	-0,03	0,60	0,41	0,22	-0,02	0,46	0,54	+0,05	-0,07	0,45	0,43
Watersh.2	r*	0,59	0,51	0,31	0,20	0,35	0,16	0,60	0,46	0,22	-0,02	0,46	0,54	0,49	0,14	0,49	0,68
Time lag	variation	0	0	0	-2	0	-1	0	-1	0	0	0	0	-2	-2	-2	-1
Watersh.1	r	0,44	0,08	0,11	-0,27	0,01	-0,13	0,30	0,33	0,06	0,07	0,27	0,24	-0,38	-0,42	0,26	0,45
Watersh.3	r*	0,44	0,18	0,11	0,39	0,17	0,00	0,34	0,39	0,06	0,07	0,27	0,24	0,59	<0	0,32	0,45
Time lag	variation	0	-2	0	-2	-1	+1	-2	-1	0	0	0	0	-2	0	-1	0
Watersh.2	r	0,67	0,29	0,45	0,47	0,40	0,61	0,55	0,53	0,44	0,22	0,57	0,51	0,56	-0,24	0,28	0,23
Watersh.3	r*	0,67	0,29	0,45	0,47	0,40	0,61	0,55	0,53	0,48	0,22	0,57	0,51	0,56	0,00	0,28	0,32
Time lag	variation	0	0	0	0	+1	0	0	0.	+1	0	0	0	0	-4	0	+2



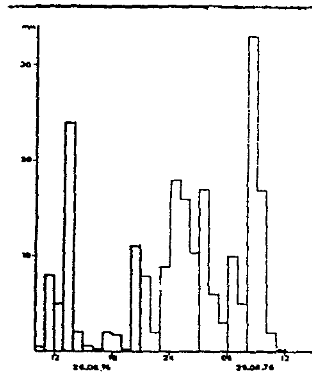


Figure 4.1

Four scenario generated on gardon d'Anduze watershed after eleven hours of rain.

UP : GENERATED RAINFALL
DOWN : OBSERVED RAINFALL.

These results lead us to extend the field of application of the point rainfall generator used by CREUTIN and OBLED(1983). This stochastic model is now able to generate areal rainfall taking into account the rain already fallen over the watershed. The generation of several scenarios (10 to 20) may guide forecasters to estimate the possible floods during the next hours (figure 4.1).

CONCLUSION

In the absence of any assumption on the structure of rainfall fields, it is always possible to estimate the mean areal rainfall over a watershed as arithmetic or weighted mean value of the measured point rainfalls.

Nevertheless, we have shown that over small watersheds, the estimation could be erroneous. The study of climatological variogram, has led to the selection of various classes of models. In each class, the parameters governing the best theoretical variogram can easily be estimated. This estimation is not very dependent on the choice of the minimization method (LS or MIE). Integration of the optimal variogram at the watershed allows for a computation of the standard deviation of areal means.

Moreover, the computation of areal rainfalls allow for a study of their temporal statistical structure and its modeling by a stochastic areal rainfall generator.

REFERENCES

BASTING and GEVERS M. (1983) :

Identification and optimal estimation of random fields from scattered point wise data. working paper. LOUVAIN LA NEUVE, BELGIUM.

CREUTIN J.D. and OBLED CH. (1980) :

"Modeling spatial and temporal characteristics on rainfall as input to a flood forecasting model". IAHS UNESCO-WMO OXFORD 1980. Publ. IAHS N° 129 p 41-49.

CREUTIN J.D. and OBLED CH. (1982) :

"Objective analyses and mapping techniques for rainfall fields : an objective comparison". Water Resour. Res., vol. 18, n° 2, p. 413-431.

DELFINER P. and DELHOMME J.P. (1975) :

"Optimum interpolation by kriging" in display and analysis of spatial data, DAVIS and MACULLOJH Ed., J WILEY.

DELHOMME J.P. (1978) :
"Kriging in the hydrosciences", Advances in water Resources, vol. 1 (5), pp 251-266.

DUBAND D. (1980) :

"Dynamic selection of analogue flow patterns to produce quantitative precipitation forecasts" WMO symposium on probabilistic and statistical methods in weather forecasting. NICE, 8-12 september 1980. Proceedings p. 487-492.

JOURNEL A.G. and AVIJBRECHTS CH.J (1975)

"Mining geostatistics", Academic Press.

LORENT B., BASTIN G., DUQUE C., GEVERS M., (1982) :

"Optimal estimation of the average areal rainfall over a basin and optimal selection of raingage locations, 6 th IPAC symp. on identification and system parameters estimations, Washington, June 7-11, 1982.

MATHERON G. (1965) :

"Les variables régionalisées et leurs estimation", Masson Publ.

TODOROVIC P. and YEVJEVICH V., (1969) :

Stachastic process of precipitation. Hydrology papers n° 35. Colorado state university, Fort Collins, Colorado 1969.

PREPRINTS

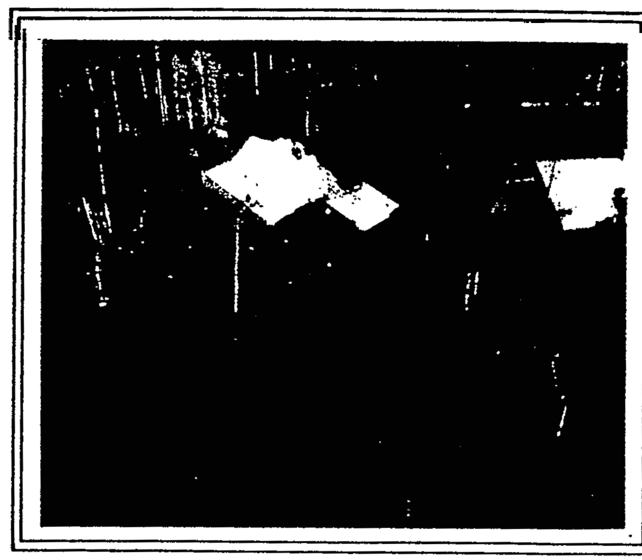
T. LEBEL
J. CREUTIN

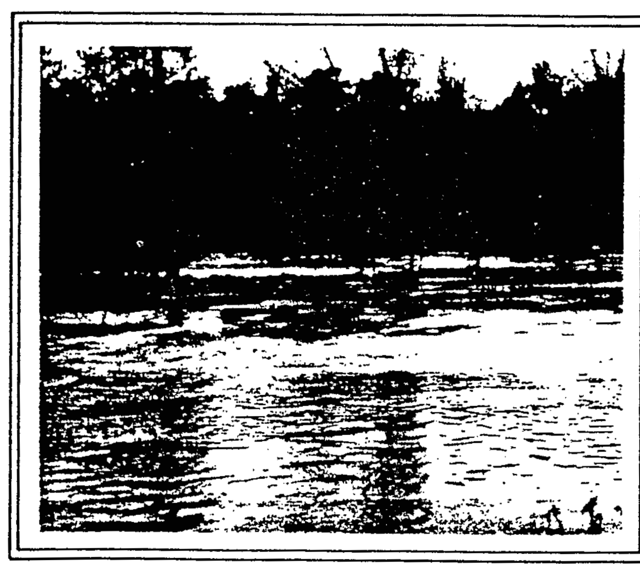
FIFTH CONFERENCE

ON

HYDROMETEOROLOGY

OCTOBER 17-19, 1983 TULSA, OKLA.



AMERICAN METEOROLOGICAL SOCIETY