COMMISSION OF EUROPEAN COMMUNITIES

AFORISM

A COMPREHENSIVE FORECASTING SYSTEM FOR FLOOD RISK MITIGATION AND CONTROL

Contribution of the National Technical University of Athens
to the 3rd annual progress report

Th. Xanthopoulos, D. Koutsoyiannis and I. Nalbantis

September 1994
Research Contract no.: EPOC-CT90-0023

Contract no.: EPOC-CT90-0023

Contractor Institution: NTUA-DWR, National Technical University of Athens, Division of Water Resources

Project leader: Prof. Themistocle S. XANTHOPOULOS

Other institutions co-operating: University of Bologna
Ente Regionale di Sviluppo Agricolo
University College Cork
University of Newcastle Upon Tyne
Ecole Polytechnique Federale de Lausanne
Instituto Superior de Agronomia, Lisbon
Institut National Polytechnique de Grenoble

Starting date: June 1st, 1991

Title of the project: A comprehensive FOrecasting system for flood RISk Mitigation and control (AFORISM)

1. OBJECTIVES

The research of NTUA-DWR has two basic lines within the framework of AFORISM: the study of intense, flood producing, rainfall and the study of rainfall-runoff models.

The study of intense rainfall includes:

(a) analysis and modelling of the temporal structure of storm events at a point and areal (lumped) basis,

(b) construction of a stochastic rainfall generator by using disaggregation techniques, and

(c) application of the rainfall generator for generation of storm scenarios and testing of the model results.

The second line of research includes:

(a) evaluation of the performance of several rainfall-runoff models with emphasis on their runoff production part by treating the transfer to the watershed's outlet in a uniform way,

(b) addressing the problem of data inadequacy that arises from the gauging networks with non-recording devices; a methodology for integrating, in a continuous-time model, information from both daily and shorter time-step data is sought,

(c) investigation of the usefulness of the FDTF-ERUHDIT method in tackling the above problem described in b,

(d) examining the possibility of initialising continuous-time rainfall-runoff models through models of the same structure operating on a daily time basis, and

(e) participation in testing of different rainfall-runoff models with the framework of the common experiment of the project.

2. METHODS, DATA AND MATERIALS

2.1 Methods

Modern techniques, based on the theory of self-similar (scaling) processes are used for modelling of the time distribution of rainfall. In addition, the meteorological conditions of the area are considered in order to classify storm events by weather type. New disaggregation techniques with simple structure and reduced parameter sets, which can be combined with the stochastic rainfall model are developed.

For rainfall-runoff modelling two well-known conceptual rainfall-runoff models were applied. An analysis framework for model testing was set up to perform different model calibrations of the same model on different kinds of data and to link model operation on more than one time steps.

2.2 Hydrologic data

The Evinos River basin, Middle Greece, at Poros Righaniou with a total area of 884 km^2 was selected as the main study area. A database was constructed for the rainfall and runoff data for that basin. For the rainfall model development and parameter estimation the data of three recording rain gauges for a 20 year period were digitised on an hourly time basis. In addition, weather maps at the surface and 500 mb level for the same period were used. Some 500 intense rainfall events were isolated from the time series by using certain criteria for event identification and selection.

For rainfall-runoff modelling, the following data of the Evinos River basin were obtained:

(a) Hourly rainfall data of three rain recorders for a 20-year period.

(b) Daily rainfall of six rain gauges for an about 30-year period.

(c) Hourly runoff data at the outlet for 28 flood events.

(d) Hourly runoff data at the outlet and maximum and minimum daily temperatures for a 2-year period.

(e) Daily runoff at the outlet for a 8-year period.

(f) Daily mean temperature of 3 measuring stations, for a 8-year period (as in e).

In addition, data from two other basins were used in the rainfall study. A rainfall data set of the Aliakmon River basin, province of Macedonia, Greece was used for the first period of rainfall model testing. Finally, a data set including two-year rainfall records of four raingauges (Firenzuola, Poreta Terme, Montecatone, Bologna o.s.i.) at the Reno river basin was used in the last year in order to calibrate and apply the rainfall model in this principal study area of AFORISM.

2.3 Materials

The FDTF-ERUHDIT identification package was obtained from French EDF within the frame of collaboration with the French partner. Initial FORTRAN code of the TANK model was provided by UNIBO-ICI and served as a basis to recode the model All the developed models have been programmed and run on a PC/DOS environment in Pascal programming language. No other software packages have been used.

3. RESULTS TO DATE

3.1 Rainfall modelling

A stochastic model describing the temporal distribution of rainfall within a storm event, based on the theory of self-similar (scaling) processes, was the theoretical background of the rainfall modelling. This model, developed in the framework of AFORISM, has been named the Scaling Model of Storm Hyetograph (Koutsoyiannis and Foufoula-Georgiou, 1993). The basic hypothesis of the model is that the instantaneous rainfall intensity process at any time position in the interior of a storm event of a certain duration *D* depends on that duration in a manner expressed by a simple scaling law with a constant scaling exponent *H*. Thus, the instantaneous intensities of two events with different durations, after appropriate scaling of time (determined by the ratio of the durations) and intensity (determined by the ratio of the durations raised to the scaling exponent), can have identical distributions, as expressed by

$$\{\xi(t, D)\} \overset{d}{=} \{\lambda^{-H}\, \xi(\lambda t, \lambda D)\}$$

where t denotes time ($0 \le t \le D$) and $\xi(t, D)$ is the instantaneous intensity. As a consequence of this hypothesis, the mean and standard deviation of total storm depth increase with duration each according to a power law with the same exponent; also, the mean and standard deviation of the incremental (e.g. hourly) depth increase with duration according to the same power law; the lag-one correlation coefficient of incremental depths increases with duration; and the decay rate of the autocorrelation function of the incremental rainfall depth decreases with the increase of duration. The model was implemented in its simplest possible mathematical structure with four parameters only, and in a modified form using five parameters. It was found to explain reasonably well the statistical properties of historical data, thus providing an efficient parameterisation of storms with varying durations and total depths. Also, it is consistent with, and provides a theoretical basis for, the concept of normalised mass curves. Furthermore, it was found that the scaling model is superior to other simple temporal rainfall models, which were unable to capture important statistical properties of storm rainfall. At the stage of initial model testing the rainfall data of the Aliakmon River basin (Macedonia, NW

Greece) were used, and it was found that the model is in good agreement with those data. The model was subsequently applied to the data of Evinos River basin and calibrated separately for the storms of each of the two seasons (rainy and dry). The conclusion is that the scaling model fits the intense rainfall data in both seasons, thus providing a basis for modelling of intense rainfall events. Finally, the model was applied to the data of the Reno River basin in an areal basis where it was found to fit all storms of the year regardless of the season (wet or dry) using only one parameter set. It is remarkable that the fitting is also good for rainfall characteristics that were not used explicitly in the estimation of model parameters, such as the autocorrelation functions of hourly depth for lag > 1.

A common property of rainfall data, which was also validated from all data examined in this study, is the high coefficient of variation (usually greater than unity) of all variables associated with a rainfall event and mainly of the hourly depth. This property is apparently a serious obstacle in building a stochastic rainfall forecasting model. A part of the variance of these variables can be explained by the storm duration as inferred by the scaling hypothesis. This part can be as high as about 50% for the total depth but it drops to 2-4% for the hourly depth. As an attempt to further lower the unexplained part of the variance of these variables, we examined the general meteorological patterns that produce the specific events, applying a certain classification of storm events by weather type.

In this analysis, the data set of intense rainfall events at the Evinos River basin was used. The rainfall events of both rainy and dry seasons were initially classified into different weather types. The specific weather types examined were introduced by Maheras (1982). Their definition depends upon weather characteristics such as the location of centres of anticyclones, the main trajectories of cyclones and some special characteristics at the surface and at the 500 mb level. It was found that two of the weather types, namely the cyclonic types SW1 and NW1 give rise to the majority of intense rainfall events (about 30% each one), while other four cyclonic types (W1, W2, SW2, NW2) and one special type (DOR) produce intense rainfall less frequently. The main characteristics of the rainfall events (mean and variance of event duration and of total and hourly depth; lag-one autocorrelation coefficient of hourly depth) were calculated for each class of events that belongs to a specific weather type. By comparison of different classes, it was concluded that apart from the probability of occurrence of a storm, only few significant differences appear between characteristics of different weather types (Mamassis et Koutsoyiannis, 1993). Thus, the introduction of weather types does not explain high portion of the variability of rainfall (for example, it explains only a 10-20% of the variance of the total depth, depending on the season, wet or dry).

As another attempt to lower the unexplained part of the variance of the rainfall process (and thus to built better stochastic forecasting models), we examined the possibility to utilise the lagged cross correlations between rainfall depths at several points at a basin or neighbouring basins. If these lagged correlations are strong then it could be possible to convey information from rainfall occurring at neighbouring areas to improve the forecast.

To investigate this possibility we performed some tests using rainfall data of the Reno River basin to which several simple stochastic models were fitted. Examples of such tests are given in Table 1. We observe that the use of the lag-one autocorrelation of the hourly rainfall depth by means of an AR(1) model (Model 1) can lead to a variance of residuals significantly lower than the total variance of the hourly depth. However, incorporating the lag-one cross correlation with a neighbouring station (Model 2) does not reduce the variance further. This remains true even in the case where we also include the lag-zero cross correlation (Model 3), although there is a significant lag-zero correlation between the reference station and the neighbouring station (as indicated by the reduced variance of residuals of Model 4 for large rainfall durations).

Table 1 Gain from the use of various types of information for modelling hourly point rainfall at a site in terms of reduction in variance.

Reference rain gauge (with hourly rainfall depth Y_t)	Neighbouring rain gauge (with hourly rainfall depth X_t)	Duration of rainfall events examined	Total variance of hourly depth Var[Y_t]	Variance of residuals (Var [W_t]) for Model			
				1	2	3	4
Firenzuola	Poreta Terme	≤ 22 h	1.47	0.95	0.95	0.95	1.46
Firenzuola	Poreta Terme	> 22 h	2.53	1.31	1.24	1.23	2.09
Montecatone	Bologna o.s.i.	≤ 22 h	1.39	1.13	1.12	1.12	1.35
Montecatone	Bologna o.s.i.	> 22 h	1.88	1.27	1.24	1.17	1.38

Model 1: $Y_t = a_1 Y_{t-1} + W_t$, , Model 2: $Y_t = a_2 Y_{t-1} + c_2 X_{t-1} + W_t$, ,
Model 3: $Y_t = a_3 Y_{t-1} + b_3 X_t + c_3 X_{t-1} + W_t$, , Model 4: $Y_t = b_4 X_t + W_t$,.

Thus, we conclude that it is impractical to convey information of neighbouring stations to improve the real-time stochastic forecast of rainfall. This justifies well the use of univariate rainfall models, in case that a lumped or semi-lumped rainfall-runoff framework has been chosen. The univariate rainfall models are convenient and simple tools that provide adequate input to lumped or semi-distributed rainfall-runoff models.

The next topic studied is the implementation of the scaling model for generation of synthetic rainfall events either in a marginal or a conditional manner. In the case of marginal generation, i.e. simulation of rainfall without any former information, the generation involves two steps. First, we generate the duration of the event. Second, we can either generate directly the consecutive incremental depths within the event or generate the total depth and disaggregate it into incremental depths. To this aim, an event-based rainfall generation scheme was developed which embodies two different generation forms: a typical sequential form and a disaggregation form (Koutsoyiannis, 1994). Both forms of the generation model are compatible and can be combined with either the scaling model or any other appropriate rainfall model and can perform with arbitrary time step less than the duration. The sequential form of the model is based on the generalised matrix relation

$$\mathbf{X} = \mathbf{\Omega} \mathbf{V}$$

where **X** is the vector of incremental depths inside the event, **V** is a vector of independent variables and **Ω** is a matrix of coefficients. Given the marginal and joint moments of the incremental depths (e.g., as a consequence of the scaling model) the moments of **V** and the coefficients **Ω** can be easily determined.

On the other hand, the disaggregation model can divide the total depth of an event (with known duration) into incremental depths. The disaggregation technique is characterised by simplicity and parsimony of parameters. It assumes a random shape of the hyetograph and it is compatible with various rainfall models. It is well known that most disaggregation models in the literature (e.g. Valencia and Schaake, 1972, 1973; Mejia and Rousselle, 1976; Todini, 1980; Stedinger and Vogel, 1984; Pereira et al., 1984; Lane and Frevert, 1990; Grygier and Stedinger, 1990; Koutsoyiannis, 1992) are not applicable to short scale rainfall disaggregation. Other models such as the one by Koutsoyiannis and Xanthopoulos (1990) are especially designed for short scale rainfall disaggregation, but they are not so generalised as to be combined with any rainfall model, as they include certain hypotheses about the stochastic structure of the rainfall process. The developed model is generalised in a high degree as the only hypothesis it uses is that the incremental rainfall depths are approximately gamma distributed and not very highly serially correlated. With this assumption a simple two-step disaggregation method was established. At the first step the method uses the above mentioned sequential model without reference to the total depth, while at the second step an appropriate adjusting procedure is applied. It was found that the method gives good approximations of the important statistics of interest. Furthermore, under some ideal conditions the disaggregation method was shown to be exact in a strict sense, i.e., it preserves the complete distribution of the variables (Koutsoyiannis, 1994). Both generation techniques were combined with three alternative rainfall models (the scaling model, a Markovian in continuous time and a Markovian in discrete time). The results of the model application and testing at these three cases indicated very satisfactory resemblance of the important statistics of incremental rainfall depths (first, second and third order marginal moments, marginal distributions, and joint second order statistics).

Finally, the generation model was modified so as to form a conditional generation scheme of the future evolution of a storm, given the situation at the current time step k and the previous ones. This scheme can also incorporate (as a condition) any available information about the event. A conditional simulation scheme with two main steps was used for this purpose. At the first step the total duration D is generated from its conditional distribution, given any condition that is known for duration. The second step involves the generation of the sequence of incremental depths X_j (j = k+1, ..., D), given any condition that is known for depths. The latter step is performed either in a typical sequential manner or by disaggregation.

The conditions examined fall into two categories: The first category encompasses the information that is known from the past and it includes (a) the obvious condition $D \geq k$, and (b) the observed series of incremental depths X_1, ..., X_k. The second category includes information that possibly could be provided from meteorological predictions such as approximate estimates of (a) the total duration and (b) the total depth of the event or incremental depths (e.g., every 6 hours). Such estimates can be deduced from the quantitative precipitation forecasts of the European Centre Medium Range Weather Forecast (ECMWF). Since these forecasts have a great degree of uncertainty, they can be treated by the generation scheme in a probabilistic manner, i.e., the generation scheme can directly add random components to the ECMWF forecasts.

The above described conditional simulation scheme was applied in several cases and the obtained stochastic forecasts were compared with historical data. It was concluded that in case that we use the conditions of the first category only (known past) and perform the conditional simulation for an unlimited lead time, the stochastic forecast of the evolution of the rainfall process is poor. This is due to the high coefficient of variation (> 1) and the low autocorrelation function of hourly depths. The stochastic forecast is improved if the lead time is limited to 1 hour (adapting the conditions of the recorded hyetograph every 1 h), even when the information of the first category is known only. The situation is also improved at the case that the information of the second category (estimates for the total duration and depth) is available.

3.2 Rainfall-runoff modelling

In steep headwater basins with a size ranging from several tens to several hundreds km^2, as it is very frequently the case in the Mediterranean zone, very short response times (up to a few hours) are encountered. This obliges the modeller of the basin's response to adopt a time step for discretisation of the hydrologic data within the range from a fraction of 1 hour to a few hours. This kind of data necessitates the existence of recording devices such as recording rain gauges, stage recorders and recording temperature gauges. However, during the design phase or even the very early stages of operation of a flood forecasting system a common situation for many catchments consists in the following: (a) a rather dense network of non-recording devices is present providing data for a long period (e.g. 20 or 30 years or more) on a daily basis in the most usual case, (b) there are a few recording devices with short periods of operation providing discontinuous records in most cases related to some water resources study, and (c) a proper network of recording devices is planned to be installed. In such a situation the modeller of the basin faces two alternatives: (a) either use only the limited set of pieces of continuous charts to extract some flood events and then follow an event-based modelling approach, or (b) use the above charts in combination with the daily data in the hope to embody all available information in a continuous-time model. It is the latter approach that we have chosen to explore within the AFORISM project.

We limited our analyses in the calibration of continuous-time lumped rainfall-runoff models which are widely used in flood forecasting for small and medium-sized headwater basins. Among the models of this category we have chosen those that have the common structure used in the Unit-Hydrograph modelling context, namely they comprise of one first part called production function and a second part, the transfer function. The first part summarises all hydrologic processes involved in an idealised soil column representative of the whole basin and yields the runoff volume or effective rainfall for each time step. The transfer function encompasses all transfer processes within the catchment and is reasonably assumed linear thus satisfying the assumptions of the Unit Hydrograph. In a recent paper (Duband et al., 1993) a new approach for calibrating lumped rainfall-runoff models, called FDTF-ERUHDIT was presented. Based on the Unit Hydrograph concept, the method performs a simultaneous identification of the effective rainfall series and the First Differenced transfer function or Unit Hydrograph through an alternate iterative procedure without presupposing any runoff production function or applying any arbitrary baseflow removal. The FDTF-ERUHDIT method was a key element in our analyses. We set up a framework for model validation which is then applied to a Greek basin. Two well-known rainfall-runoff models were selected for the analyses, the version of the SACRAMENTO model adopted by the U.S. National Weather Service known as the Soil Moisture Accounting (SMA) of the U.S. National Weather Service River Forecast Service or SMA-NWSRFS (Burnash et al., 1973), and the TANK model (Sugawara et al., 1983).

Our approach consisted in the following:

(a) For each model structure (e.g. that of the TANK model) calibrate a daily model based the long continuous-time data set that is available.
(b) Calibrate the transfer function of the model on the time step suitable for the dynamics of the basin which in our case was equal to 1 hour; the calibration is made on the event-based data set available through the FDTF-ERUHDIT method without presupposing any production function.
(c) Based on the above identified parameters construct a new continuous-time model, called the derived model; its production function parameters are derived from those of the daily model while the parameters of its transfer function are the only ones which are directly identified (as described in b).

The problem is how to perform this transformation of the parameters of the production function from one time step to another. We studied a certain number of cases related to the model structures selected; these cover, to our opinion, a broad spectrum of cases appearing in conceptual rainfall-runoff models commonly used in practice.

The framework for validation of our approach comprised of the following steps:

(1) Identify, for comparison purposes, a reference model on the continuous-time data set available on a short time basis; of course, this data set is available only in our case study and not in the real-world situation that faces our approach.
(2) Based on the continuous daily and the event-based data sets obtain the derived model.
(3) Evaluate and compare the performances of the two models.
(4) Examine the possibility of initialising the derived model through the daily model; a composite scheme is tested on several flood events: first the daily model is run up to the day before the flood and then the derived model is launched.
(5) Detect any deficiencies in the composite scheme and propose improvements.

The above framework was applied on the Evinos watershed through an extensive series of tests. The transfer function identified through the FDTF-ERUHDIT method over a length of 20 time steps through 5 iterations is found to have a time-to-peak equal to 6 hours. The overall determination coefficient starts from 0.823 in the first iteration and reaches 0.918 in the 5th iteration; for the 5th iteration it varies from 0.723 to 0.980 for different events.

The performance criteria for model calibration and verification periods were compared for both the daily and the hourly reference models. The results are practically the same for the two model structures as reported by others (see, for example, Franchini and Pacciani, 1990). The daily models were less good than the hourly ones and their performance criteria were lower by 10 to 20% from those of the hourly models. Then, a comparison was made between the derived models and the corresponding reference models. We observed that the performance of the derived model is very close to that of the daily models, that is only 10 to 20% lower than the best (i.e. the reference) models.

To investigate the possibility of initialising the derived model through the daily model we examined the transferability of the values of state variables from the daily to the derived hourly model. This was done by performing a linear regression of the end-of-day values of these states for the two models. For the SMA-NWSRFS model very high values of R^2 (higher than 0.85) were found for all state variables of the lower zone of the model while for the upper zone R^2 took rather poor values. The same result was shown by the regression coefficients α and β. For the TANK model the results were completely analogous: high values of R^2 for the variables of the lower zone and rather poor values for the upper zone; in the latter zone only the contents of the two subtanks and were used in the analysis because of the fact that the other variables (the contents of the two free water storages) took values always very close to zero; in fact their depletion coefficients summed practically to unity thus excluding any storage functioning of these elements and in any case a zero value was always a good prediction. Of course, this result is particular to the dynamics of the basin chosen as compared to the time steps used, but it does not pose any problem to the validation of our methodology nor reduces the generality of the approach.

After having been assured that a direct transfer of the states from the daily model to the so called derived model is possible, we tested, for some flood events, a simulation scheme with the daily model providing initial conditions for the derived model; the latter model is activated only in flood periods. Comparisons with the continuous run of the reference model showed that this composite scheme is satisfactory especially if the contents of the upper storage elements are tuned at the moment of shifting from the daily to the derived model.

The main conclusions drawn from our study are:

(1) It is not an uncommon situation for many countries to dispose inadequate hydrologic data during early stages of design and development of a flood forecasting system; in particular, large amounts of daily data are usually available together with event-based data sets; while the most straightforward approach is to calibrate event-type model, we propose an alternative approach to calibrate a continuous-time rainfall-runoff model integrating both daily and shorter time-scale information

(2) The calibration procedure recommended utilises the FDTF-ERUHDIT method for Unit Hydrograph identification to estimate the transfer function while the runoff-production function is derived by calibrating the whole model on daily data.

(3) The resulting derived model performs well in continuous simulation and can be easily initialised through the daily model; the performance of this scheme is very much improved through tuning of the states directly related to the quickflow component of the runoff.

(4) The methodology for integrating information on different time steps in a continuous-time rainfall-runoff model was validated in a modelling context with noisy data similar to that motivated the whole study; further insight can be expected by making extra tests on generated error-free data; an attempt to do this is planned for the future

REFERENCES

Grygier, J. C. & Stedinger, J. R., SPIGOT, A synthetic streamflow generation software package, Technical description, School of Civil and Environmental Engineering, Cornell University, Ithaca, NY., Version 2.5, 1990.

Koutsoyiannis, D., A stochastic disaggregation method for design storm and flood synthesis, *Journal of Hydrology*, 156, 193-225, 1994.

Koutsoyiannis, D., A nonlinear disaggregation model with a reduced parameter set for simulation of hydrologic series, *Water Resour. Res.*, 28(12), 3175-3191, 1992.

Koutsoyiannis, D. and Foufoula-Georgiou, E., A scaling model of storm hyetograph, *Water Resources Research*, 29(7), 2345-2361, 1993.

Koutsoyiannis, D. and Xanthopoulos, Th., A dynamic model for short-scale rainfall disaggregation, *Hydrol. Sci. J.*, 35(3) 303-322, 1990.

Lane, W. L. and Frevert, D. K., Applied Stochastic Techniques, User's Manual, Bureau of Reclamation, Engineering and Research Center, Denver, Co., Personal Computer Version 1990.

Maheras, P., Synoptic situations and multivariate analysis of weather in Thessaloniki (in Greek), Publication of the Laboratory of Climatology, University of Athens, Greece, 1982.

Mamassis, N. et Koutsoyiannis, D., Structure temporelle de pluies intenses par type de temps, *6ème Colloque International de Climatologie*, Thessaloniki, 1993.

Mejia, J. M. and Rousselle, J., Disaggregation models in Hydrology revisited, *Water Resour. Res.*, 12(2), 185-186, 1976.

Todini, E., The preservation of skewness in linear disaggregation schemes, *J. Hydrol.*, 47, 199-214, 1980

Valencia, D. and Schaake, J. C., Disaggregation processes in Stochastic Hydrology, *Water Resour. Res.*, 9(3) 211-219, 1973.