GÉOSCIENCES DE SURFACE/
SURFACE GÉOSCIENCES
(HYDROLOGIE-HYDROGÉOLOGIE/
HYDROLOGY/HYDROGEOLOGY)

On the applicability of ARMA models to time-series analysis of karstic spring flow

C.R. Acad. Sci. Paris,
t. 321, série II a,
p. 31 à 37,
1995

Alberto Padilla, Antonio Pulido-Bosch and María L. Calvache

A. P. : Investigaciones Geológicas y Mineras SA, Urban. Alcázar Genil. Ed. Azulema, Bajo, Granada, Spain ;

A. P.-B. and M. L. C. : Departamento de Geodinámica, Universidad de Granada, 18071 Granada, Spain.

Abstract ARMA models have been applied to study the flow series of the karstic spring of Aliou. We have analysed the theoretical meaning of the parameters involved in the model by applying it to a simple system of the emptying of a karstic aquifer. We have also indicated the types of transformation necessary to apply these models to the flow series which lack normality and have a strong periodic component. From our results we point out the advantages of this type of model, and discuss the physical significance of the parameters obtained, from a hydrodynamic standpoint.

Keywords: ARMA model, Time-series, Karstic spring, Aliou aquifer.

Résumé **Sur l'applicabilité des modèles ARMA à l'analyse des chroniques de débits de sources karstiques**

Les modèles ARMA ont été appliqués à l'étude des chroniques de débit de la source karstique d'Aliou. Nous avons analysé la signification théorique des paramètres employés dans le modèle, en l'appliquant à un simple schéma de vidange d'un aquifère karstique. Nous avons aussi indiqué les types de transformations nécessaires, lors de l'application de ces modèles à des chroniques de débit qui manquent de normalité et ont une forte composante périodique. A partir de nos résultats, nous établissons les avantages de ce type de modèles et nous discutons la signification physique des paramètres obtenus, du point de vue hydrodynamique.

Mots-clés : Modèle ARMA, Chronique, Source karstique, Aquifère d'Aliou.

Version française abrégée

Les modèles autorégressifs à moyennes mobiles (ARMA), appliqués jusqu'à présent essentiellement à l'analyse de processus hydrologiques, peuvent être également utilisés dans l'étude de processus hydrogéologiques dans le karst. Dans ce travail, on applique un modèle ARMA (p, q) à paramètres constants, dont l'expression générale est (1) (Box et Jenkins, 1976). Suivant un processus déductif, on établit une comparaison entre les paramètres autorégressifs et à moyennes mobiles avec les variables qui interviennent dans le fonctionnement d'un aquifère karstique (2), donnant ainsi un sens physique aux paramètres d'un modèle ARMA. Ainsi, ϕ_i indiquerait la proportion de vidange de la zone saturée qui s'écoule lentement, a_t serait la partie de la précipitation qui circule rapidement à travers de grands conduits à l'instant t et, enfin, θ_j n'a pas de signification physique évidente, bien qu'elle puisse être assimilée à la fraction des précipitations qui circule par de grands conduits à l'instant précédent.

Pour justifier l'applicabilité des modèles ARMA et vérifier le sens physique de leurs paramètres, on a utilisé ce type de modèles à la source d'Aliou (Pyrénées françaises, **fig. 1**) qui présente un fonctionnement extrêmement karstique (Mangin, 1975 et 1984; Mangin et Pulido-Bosch, 1983 ; Padilla *et al.*, 1994 ; **fig. 3**).

On a choisi une chronique de données journalières correspondant à 5 années hydrologiques (1/10/70-30/9/75), à laquelle on applique des transformations dans deux buts différents : d'une part, diminuer le biais, étant donné que les valeurs élevées de celui-ci indiquent un manque de normalité ; et, d'autre

Note
présentée par Georges Pédro.
remise le 20 février 1995, *acceptée après révision le* 10 avril 1995.

part, éliminer la périodicité que présente la chronique de données originales. On constate que, en effet, la périodicité a été éliminée, car le premier coefficient d'autocorrélation, $R_{1,T}$, (Salas *et al.*, 1980) dans les différents intervalles a une valeur d'environ 0,8 (**fig. 4**).

CALAGE ET EXACTITUDE DU MODÈLE

Lorsque la périodicité a été éliminée, on ajuste un modèle ARMA à paramètres constants. Le modèle à appliquer a un seul paramètre autorégressif étant donné que l'autocorrélogramme partiel, à intervalles de confiance de 95 % (Box et Jenkins, 1976 ; **fig. 5**), montre l'existence d'un premier paramètre autorégressif très marqué. Parmi les cinq essais réalisés, on a choisi le modèle ARMA (1, 2) (**tableau**), étant donné que c'est celui qui présente une Sa^2 (variance de résidus) plus petite. On peut voir l'expression du modèle dans (3). Une façon d'éprouver l'exactitude du modèle est de voir l'indépendance des résidus. En effet, au moyen du test de Porte Manteau, on déduit que l'indépendance des résidus est acceptée.

On a généré une chronique synthétique *z*, en l'appliquant au modèle ARMA (1, 2) avec des valeurs aléatoires générées avec la même distribution (lognormale), la même moyenne et le même écart-type des résidus. Enfin, on a réalisé un processus inverse pour obtenir la chronique de débits synthétiques *q* et pouvoir ainsi comparer l'hydrogramme réel et synthétique (**fig. 6**), ainsi que les autocorrélogrammes (**fig. 7**) de deux chroniques de données. Comme on peut l'observer, la ressemblance est évidente : dans l'hydrogramme, le caractère de réaction rapide de l'aquifère est conservé, ainsi que le même type d'évolution des crues. On en déduit que, dans l'autocorrélogramme, les principales statistiques se conservent dans le modèle calé.

DISCUSSION

On peut conclure que l'applicabilité des modèles ARMA à l'étude des processus hydrogéologiques est possible, au moins dans le cas de sources qui drainent des aquifères carbonatés à comportement extrêmement karstique. En outre, dans ce cas, les coefficients autorégressifs peuvent être assimilés à l'écoulement qui constitue le débit de base de l'aquifère ; par contre, les paramètres à moyennes mobiles et résidus sont indicatifs de l'organisation interne de l'écoulement (degré de karstification, existence de grands conduits, etc.), nuancée par le régime de précipitations. En effet, comme à Aliou la karstification est importante, ces paramètres atteignent des valeurs élevées.

INTRODUCTION

Autoregressive moving average (ARMA) models have been used up to now to model univariate seasonal hydrological processes (Salas *et al.*, 1980, 1982; Obeysekera and Salas, 1986; Haltiner and Salas, 1988; Ula, 1990). These univariate models are based on the analysis of the stochastic structure of a temporal series, with the objective of short-term predictions, completing data or generating synthetic series which conserve the mean statistics, above all the temporal correlation (autocorrelation function) of the underlying process in the original series. In this paper, we shall attempt to demonstrate the applicability of the ARMA models to the study of a time series of a spring draining a carbonate aquifer; thus we offer an explanation for the physical meaning of the autoregressive and the moving average parameters of an ARMA model applied to a time series of flows draining a karstic system. For these two objectives, a study has been made of the Aliou spring (French Pyrenees) which drains a system with a surface of 12 km^2 (**fig. 1**). This karstic aquifer is characteristic of a quick-flow system due to the circulation of the water through a network of drains formed during the karstification processes (Mangin, 1975 and 1984; Mangin and Pulido-Bosch, 1983; Padilla *et al.*, 1994).

METHODOLOGY AND PHYSICAL JUSTIFICATION

Here we apply an autoregressive moving average model, ARMA (p, q) of constant parameter, the general expression of which (Box and Jenkins, 1976) is:

$$(1) \quad z_t = \sum_{i=1}^{p} \phi_i z_{t-i} + a_t - \sum_{j=1}^{q} \theta_j a_{t-j}$$

where z_t represents a standard periodic hydrological time series, ϕ_i and θ_j are time-varying autoregressive and moving-average coefficients, respectively, and a_t is an independent and normally distributed variable (white noise).

Let us investigate the physical meaning which an ARMA model can have when applied to the study of the flows of a karstic spring. Let the karstic aquifer be the one sketched in figure 2, where the variables represent discretized values. Let us assume that in this system the functions which link the transference of volumes are linear. The precipitation at a given instant P_t is divided into the infiltration towards the saturated zone, which flows slowly (constituting the baseflow), aP_t; the evapotranspiration bP_t; and the infiltration circulating rapidly through the great conduits (quickflow), which in an isolated system would be equal to $(1 - a - b)P_t = dP_t$. The parameters a, b and d represent the different fractions into which the precipitation is divided. The flow of the spring at an instant t will be given by

$$Q_t = c\,V_{t-1} + dP_t$$

where $c\,V_{t-1}$ indicates the fraction of the stored volume in the previous instant drained by the spring. On the other hand, the stored volume in the instant t is equal to

$$\begin{aligned} V_t &= V_{t-1} + a\,P_t - c\,V_{t-1} \\ &= (1-c)\,V_{t-1} + a\,P_t \end{aligned}$$

Combining these two equations with those obtained by extending the previous ones for Q_{t-1}, V_{t-1} and V_{t-2}, gives

$$\begin{aligned} (2) \quad Q_t &= (1-c)\,Q_{t-1} \\ &+ dP_t - [d\,(1-c) - ca]\,P_{t-1} \end{aligned}$$

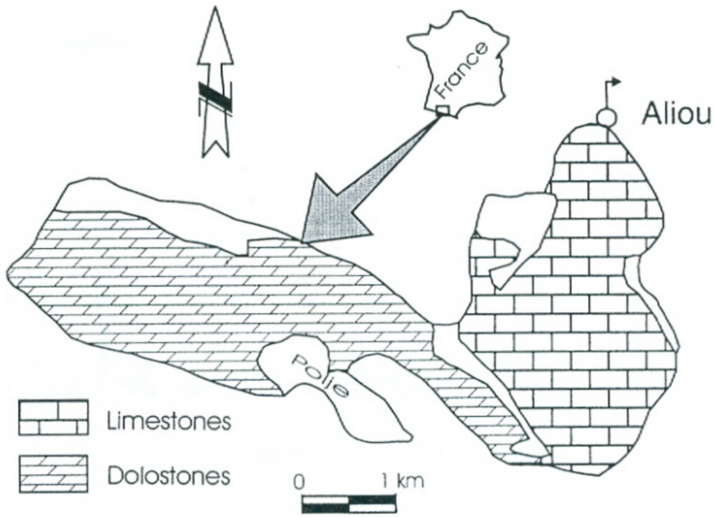

Figure 1

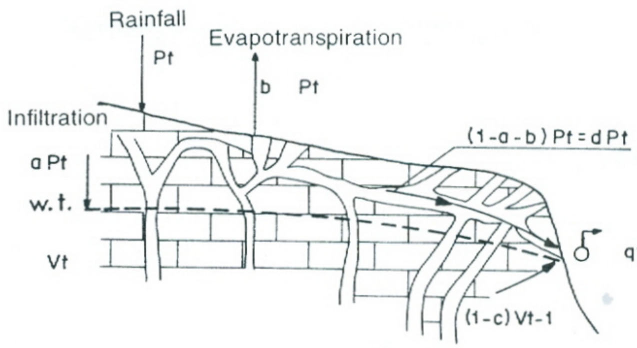

Figure 2

This expression is similar to that of an ARMA (1, 1) model, in which the precipitation is a random and independent variable; $(1-c)$ equals ϕ_1 and indicates the manner in which the saturated zone drains through small discontinuities (diffuse flow?); dP_t equals the random term a_t and represents the fraction of the precipitation which circulates essentially through the great conduits at the instant t; and $[d\,(1-c) - ca]$ equal the term θ_1; this term lacks clear physical meaning, although it could be the precipitation fraction circulating through the great conduits in the previous instant, in which case it can be considered representative of the degree of organization of the existing karstic formation (network of conduits and channels) varying with the precipitation regime.

Fig. 1 Location of the Aliou aquifer.

Localisation de l'aquifère d'Aliou.

Fig. 2 Conceptual representation of the discharge process of a karstic spring (modified from Salas *et al.*, 1980).

Représentation conceptuelle de la décharge d'une source karstique (modifié de Salas *et al.*, 1980).

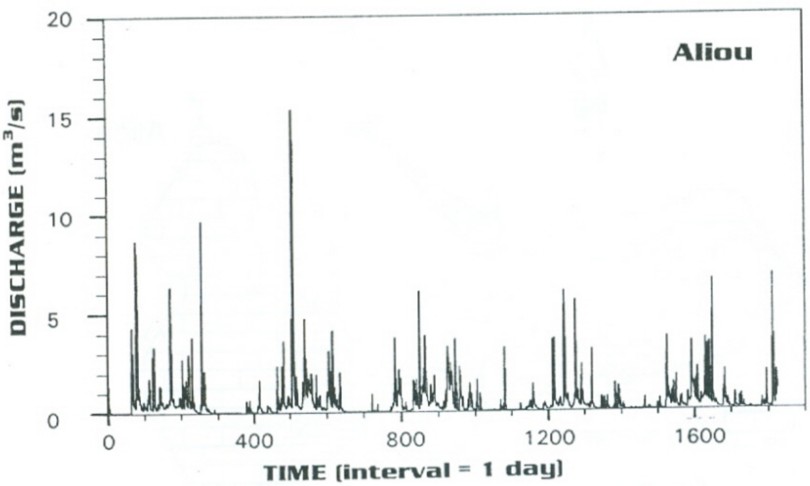

Figure 3

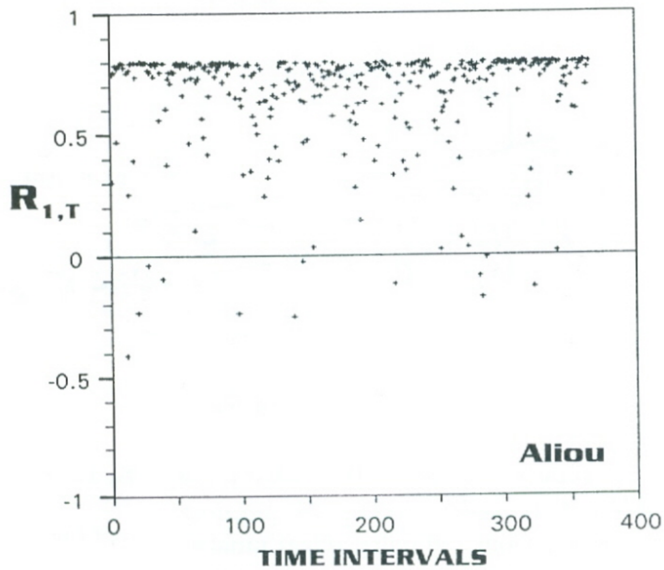

Figure 4

Fig. 3 Flow hydrograph of Aliou.

Hydrogramme de la source d'Aliou.

Fig. 4 First autocorrelation coefficient ($R_{1,T}$) of the series $z_{V,T}$ of Aliou.

Premier coefficient d'autocorrélation ($R_{1,T}$) de la série $z_{V,T}$ d'Aliou.

The procedure for this paper was as follows. Firstly, the spring flow series was transformed to reduce the bias and eliminate the periodic character of the series. Secondly, the autoregressive parameters and those of the moving averages were calculated by applying the least-squares method, resolving the systems of equations with the algorithm of Marquarot (1963). Thirdly, the ARMA model was found to adjust best to the series of data which we have, and, finally, the goodness of fit of the model was demonstrated.

DATA SERIES TRANSFORMATION

The study period (1 October 1970 to 30 September 1975) covered 5 hydraulic years with sampling intervals of of 1 day (1825 data). The hydrograph of the spring flow (**fig. 3**) shows that storage is practically nil and also shows a certain periodicity due to the more intense peaks in the periods of high water. Consequently, there is a direct relationship with the distribution of the precipitation in the sector (Mangin, 1981 *a* and *b*). We have to eliminate this periodicity by transforming the original series for applying a constant-parameter ARMA model.

The mean statistics of the data series are: m = 0.47 (mean), s = 0.94 (standard deviation) and g = 6.36 (high values of bias indicate a lack of normality). To decrease the bias, after assaying a series of transformations, a logarithmic approach proved the most appropriate. The m of the new series, $y_t = \log(q_t)$, is – 0.85, s is 0.71 and g is 0.02; the latter figure is very small with respect to the original series.

To analyse the periodicity of the most significant statistics (mean and standard deviation), we grouped the terms of the series y_t in their corresponding years; this new series was represented by $y_{V,T}$ were V refers to the years and T to the intervals into which the year was divided. The periodicity is eliminated by the transformation

$$z_{V,T} = \frac{y_{V,T} - m_T}{s_T}$$

although, to simplify the process, we have calculated the m_T and s_T by the most significant coefficients, obtained by the Fourier transformation of both statistics (Salas *et al.*, 1980). This method has frequently been used by various authors, notably Chu and Katz (1989) and Wasimi (1990).

Figure 4 confirms that the periodicity in fact has been eliminated, given that many of the $R_{1,T}$ values (Salas *et al.*, 1980) are around 0.8. Therefore, the adjustment of a constant parameter ARMA model is feasible.

The m for the series z_t (obtained from $z_{V,T}$) is –0.02, s is 1.13 and g is 0.03.

Table Parameters of the adjusted models to z_t series at Aliou: *C*, constant; ϕ_p, autoregressive coefficients; θ_q, moving average coefficients; Sa^2, sum of the square residuals; *Q*, value of χ^2 for the first 20 autocorrelation coefficients of residuals; and AIC, information criteria of Akaike (1974) about parameters parsimony.

Paramètres des modèles calés de la chronique z_t d'Aliou : C, constant ; ϕ_p, coefficients autorégressifs ; θ_q, coefficients en moyennes mobiles ; Sa^2, somme du carré des résidus ; Q, valeur de X^2 pour les premiers 20 coefficients d'autocorrelation des résidus ; et AIC, critère d'information d'Akaike (1974) concernant les paramètres de parcimonie.

Model	C	ϕ_1	ϕ_2	θ_1	θ_2	θ_3	Sa^2	Q	AIC
ARMA(1,0)...	0,001	0,798	-	-	-	-	848,4	46,4	-1 396
ARMA(1,1)...	– 0,002	0,744	-	0,149	-	-	839,2	27,3	-1 414
ARMA(1,2)...	– 0,003	0,811	-	– 0,071	0,119	-	832,9	16,3	-1 425
ARMA(1,3)...	– 0,006	0,822	-	– 0,049	– 0,119	– 0,008	832,4	25,5	-1 424
ARMA(2,0)...	– 0,011	0,862	– 0,077	-	-	-	840,4	25,5	-1411

ADJUSTMENT AND GOODNESS OF THE MODEL

The adjustment of the models will be done over the series of 5 years of z_t. To obtain the autoregressive parameters and, to a lesser extent, the moving averages which the ARMA model should have, we carried out the partial autocorrelogram of z_t with reliability limits of 95% (Box and Jenkins, 1976; **fig. 5**). This indicates a strong first autoregressive parameter; therefore, the ARMA model to be adjusted will have one autoregressive parameter. To fix the number of parameters of moving averages, five assays have been carried out (**table**). Of all the models analysed, the ARMA (1, 2) presents the smallest Sa^2. Therefore, the definitive ARMA (1, 2) model is expressed:

$$\hat{z}_t = 0.811\,\hat{z}_{t-1} + a_t + 0.071\,a_{t-1} + 0.119\,a_{t-2} - 0.003$$

One way of testing the goodness of the adjusted models in each case is by proving the independence of the residuals. For this, we chose the test of Porte Manteau, which used the value of *Q*. The statistic $Q = 16.3$ for $L = 20$ is less than the value of $\chi^2 = 27.6$ with $L - p - q = 17$ degrees of freedom and a significant level of 0.05; therefore, the independence hypothesis of the residuals is accepted.

Upon confirming the independence of the residuals, we arranged the synthetic series $\hat{z}_t$. The aleatory values were generated with the same lognormal distribution ($\beta = 1.6$, $\mu_n = 0.49$ and $\sigma_n = 0.36$), mean and standard deviation that have the residuals obtained from adjustment of the ARMA(1,2) model.

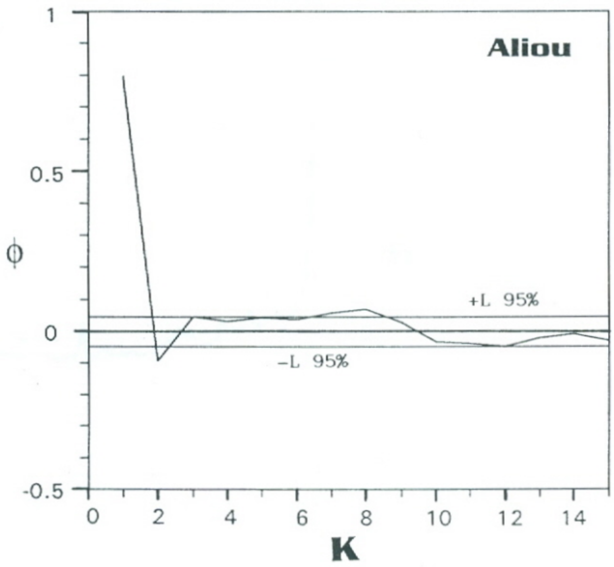

Fig. 5 Partial autocorrelogram of the series z_t of the flow at Aliou.

Autocorrélogramme partiel de la série z_t des débits d'Aliou.

To obtain the series of synthetic spring flows, we have followed the inverse process, carrying out transformations of the series. Once $\hat{z}_t$ values were grouped according to years, the $\hat{z}_{V,T}$ was transformed by the expressions:

$$\hat{y}_{V,T} = \hat{s}_T\,\hat{z}_{V,T} + \hat{m}_T$$

$$\hat{q}_{V,T} = 10^{\hat{y}_{V,T}}$$

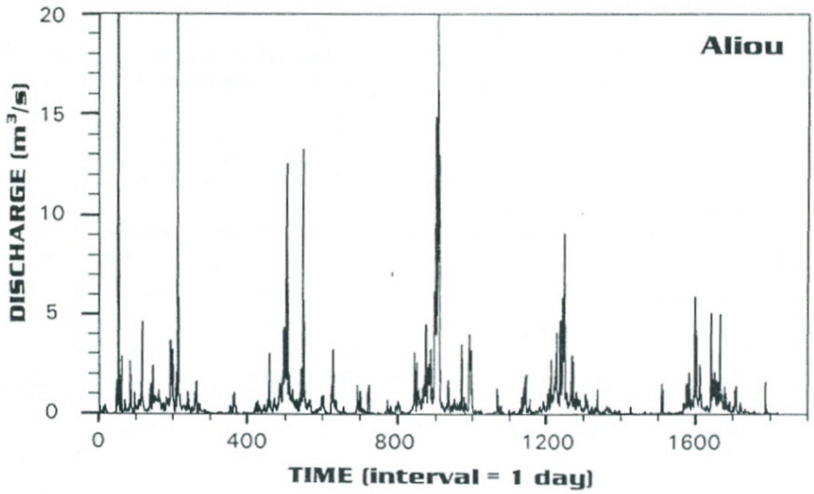

Figure 6

Fig. 6 Synthetic hydrograph obtained with the model adjusted to the flow at Aliou.

Hydrogramme synthétique obtenu avec le modèle calé sur les débits d'Aliou.

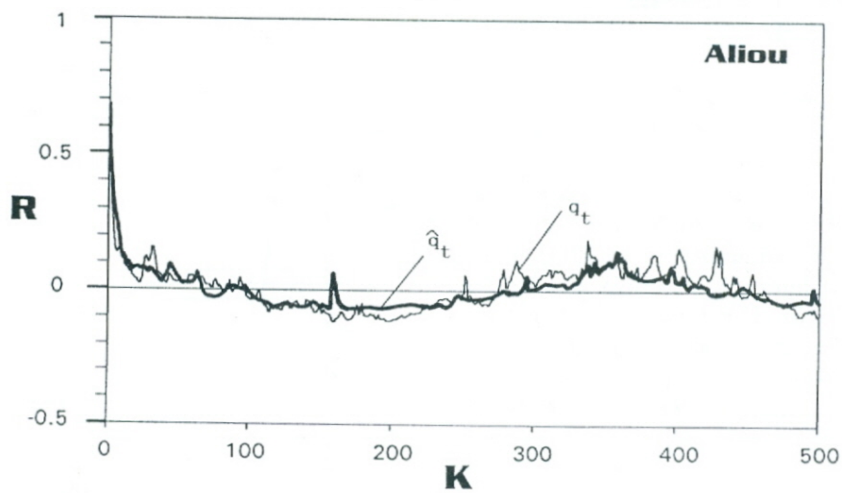

Figure 7

Fig. 7 Autocorrelogram of the real flow (q_t) and the generated ($\hat{q}_t$) at Aliou.

Autocorrélogramme des débits mesurés (q_t) et simulés ($\hat{q}_t$) à Aliou.

The synthetic hydrograph (**fig. 6**), in morphology, quite closely resembles a real hydrograph, with extremely sharp aleatory peaks and a very low storage in the system. The correlograms of the series q_t and $\hat{q}_t$ (**fig. 7**) are very similar, indicating that the principal statistics are preserved in the adjusted model.

FINAL DISCUSSION

With the results obtained from adjusting the ARMA model to the flow of the karstic spring, we have been able to make the following observations.

Autoregressive coefficients can be representative of the drainage coefficient, in the strictest sense of the term – that is, the speed of the emptying of the saturated zone without the influence of precipitations (baseflow). On the other hand, such a high value cannot be explained at Aliou, where the storage as a whole is very low. Nevertheless, the reduced saturated zone can provide a certain storage, although it has very little relative importance in the flow at Aliou. The high autoregressive coefficient obtained would correspond to its depletion.

Consequently, to have a general perspective on the model, we must take into account the moving average coefficients and the residual variance, since both indicate the weight of the randomness in the system. Thus, at Aliou, the first moving average coefficient and the variance are substantially high (in comparison with other karstic springs studied; Padilla, 1990); therefore, the random part of the model has much greater relevance in this system. We interpret these parameters as being representatives of the degree of karstification of the system (presence of great conduits; quickflow) varying with the precipitation regime.

We should point out also that the greater usefulness of the ARMA models is rooted in the generation of the synthetic series which conserve the mean statistics and, above all, those referring to the autocorrelation function, with the aim of using it in other types of models, and for completing series of data.

Acknowledgements: We are grateful to Dr. A. Mangin for help with the data related to the spring at Aliou and for his helpful comments on these manuscripts. The research for this project was financed by CICYT (Project AMB92-0211).

REFERENCES

AKAIKE, H., 1974. A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, AC-19, 6, pp. 716-723.

BOX, G. E. P. and JENKINS, G. M., 1976. *Time Series Analysis: Forecasting and Control*, Holden Day, San Francisco, 575 p.

CHU, P. S. and KATZ, W., 1989. Spectral estimation time series models with relevance to the southern oscillation, *J. Climate, Amer. Meteor. Soc.*, 2, (1), pp. 86-90.

HALTINER, J. P. and SALAS, J. D., 1988. Development and testing of a multivariate seasonal ARMA (1,1) model, *J. Hydrol.*, 104, pp. 247-272.

MANGIN, A., 1975. Contribution à l'étude hydrodynamique des aquifères karstiques, Lab. Sout. du CNRS (Moulis), *Ph. D. thesis*, University of Dijon. In: *Ann. Spéléol.*, 29, 3, pp. 283-332; 29, 4, pp. 495-601; 30, 1: pp. 21-124.

MANGIN, A., 1981 *a*. Utilisation des analyses corrélatoire et spectrale dans la connaissance des systèmes hydrologiques, *C. R. Acad. Sci. Paris*, 293, Series II, pp. 401-404.

MANGIN, A., 1981 *b*. Utilisation des analyses corrélatoire et spectrale croisées dans la connaissance des systèmes hydrologiques, *C. R. Acad. Sci. Paris*, 293, Series II, pp. 1011-1014.

MANGIN, A., 1984. Pour une meilleure connaissance des systèmes hydrologiques à partir des analyses corrélatoire et spectrale, *J. Hydrol.*, 67, pp. 25-43.

MANGIN, A. and PULIDO-BOSCH, A., 1983. Aplicación de los análisis de correlación y espectral en el estudio de los acuíferos kársticos, *Tecniterrae*, 51, pp. 53-65.

MARQUAROT, D. W., 1963. An algorithm for least squares estimation of non linear parameters, *J. Soc. Ind. App. Math.*, 2, pp. 431-441.

OBEYSEKERA, J. T. B. and SALAS, J. D., 1986. Modeling of aggregated hydrologic time series, *J. Hydrol.*, 86, pp. 197-219.

PADILLA, A., 199. Los modelos matemáticos aplicados al estudio de los acuíferos káarsticos, *Ph. D. thesis*, University of Granada, 267 p.

PADILLA, A., PULIDO-BOSCH, A. and MANGIN, A., 1994. Relative importance of baseflow and quickflow from hydrographs of karst spring, *Ground Water*, 32, No. 2, pp. 267-277.

SALAS, J. D., YEVJEVICH, V. and LANE, W. L., 1980. *Applied modeling of hydrologic time series*, Water Resour. Pub. Littleton, 484 p.

SALAS, J. D., BOES, D. C. and SMITH, R. A., 1982. Estimation of ARMA models with seasonal parameters, *Water Resour. Res.*, 18, (4), pp. 1006-1010.

ULA, T. A., 1990. Periodic covariance stationarity of multivariate periodic autoregressive moving average processes, *Water Resour. Res.*, 26, (5), pp. 855-861.

WASIMI, S. A., 1990. A hybrid model for forecasting daily rainfall, *Water Resour. Res.*, 26, (11), pp. 2741-2746.