



**HAL**  
open science

# Flash Floods Forecasting in a Karstic Basin Using Neural Networks: the Case of the Lez Basin (South of France)

L. Kong a Siou, A. Johannet, Séverin Pistre, V. Borrell

► **To cite this version:**

L. Kong a Siou, A. Johannet, Séverin Pistre, V. Borrell. Flash Floods Forecasting in a Karstic Basin Using Neural Networks: the Case of the Lez Basin (South of France). Bartolomé Andreo; Francisco Carrasco; Juan José Durán; James W. LaMoreaux. *Advances in research in karst media*, Springer, pp.215-221, 2010, *Environmental Earth Sciences*, 978-3-642-12485-3. 10.1007/978-3-642-12486-0\_33 . hal-02410504

**HAL Id: hal-02410504**

**<https://hal.science/hal-02410504v1>**

Submitted on 9 Jun 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Flash Floods Forecasting in a Karstic Basin using Neural Networks: the case of the *Lez* basin (south of France)

Line Kong A Siou<sup>1,2</sup>, Anne Johannet<sup>1</sup>, Séverin Pistre<sup>2</sup>, Valérie Borrell<sup>2</sup>

<sup>1</sup>Ecole des Mines d'Alès, CMGD, 6 avenue de Clavières, 30 319 Alès Cedex, France

<sup>2</sup>Université Montpellier II, Hydrosociences Montpellier, Place E. Bataillon, 34 095 Cedex 5, France

**Abstract** The present study focuses on the modeling of the *Lez* karstic system (France) using artificial neural networks. Two methods of variable selection were compared: cross-correlation and cross-validation. In both cases, the artificial neural network forecasts closely matched the measured discharge, giving Nash criteria higher than 0.8, which can thus provide satisfactory 2-day forecasts.

## 1. Introduction

Flash flood forecasting in populated areas is a difficult and crucial task. This paper addresses the case of the *Lez* aquifer, near the city of *Montpellier* (South of France). The first section shows that several effects interact non-linearly, leading to a very difficult modelling problem. The second section therefore presents neural networks, their limitations and how to use them in black-box modelling. The *Lez* outflow prediction is derived, and assessed on a test set, independent of the learning database, thereby obtaining a good forecast quality, which would enable two-day ahead warning of the population.

## 2. Problem statement

### Flash flooding of the *Lez*

The *Lez* is a 25.8 km long coastal river in south of France (Figure 1). It is sustained by a 380 km<sup>2</sup> karstic aquifer (Thiery *et al.* 1983) and rises 10 km north of *Montpellier* (400,000 inhabitants). The aquifer contains two main compartments separated by the *Corconne* fault, whose influence is poorly understood. This fault brings the main karstic formation (West) into contacts with the impermeable formations (East) (Figure 1).

The climate is mediterranean, subject to intense rainfalls during the autumn, delivering huge floods, called *épisodes cévenols*. These events can have catastrophic consequences (during the 2002 event, 25 fatalities and 2 billion Euros of damage occurred in the *Languedoc-Roussillon* Region), worsen by the proximity of the city of *Montpellier* and the complexity of the surface and underground water interactions (Roesch and Jourde 2006).

Three daily rainfall time series are available (rain gauges): *Prades-le-Lez*, near the spring, *Valflaunès* and *Saint-Martin-de-Londres* at the north of the basin.

Moreover, in addition to the complexity of the heterogeneous aquifer, to the surface and underground water interactions, and to the nonlinear flash-flood behavior; the spring is exploited for providing drinkable water to *Montpellier*.

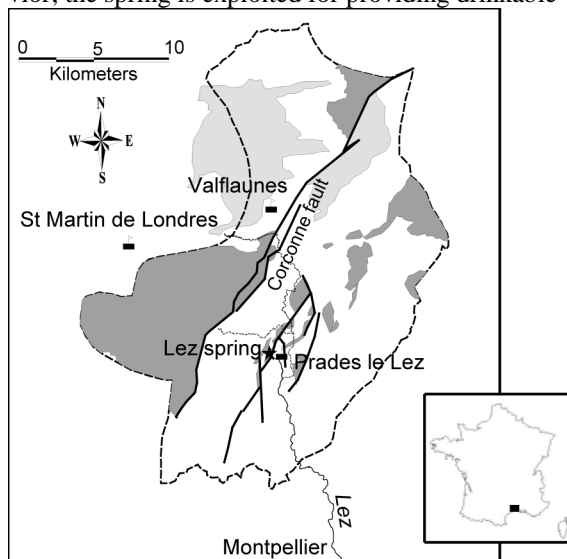


Figure 1: Watershed of the Lez spring and available rain gauges. Dotted line is the limit of the supposed hydrogeologic basin. Dark gray areas correspond to the main karstic aquifer outcrops, light gray areas to other karstic formation outcrop, white areas are impermeable areas.

### Active water management

Since 1982 the *Lez* spring is exploited with an “active water management” in order to satisfy the water demand of the city of *Montpellier*: instead of pumping water inside the spring pool, drillings were carried out and enable pumping inside the main drain of the spring, often drying up the spring and emptying a part of the aquifer. Nevertheless, in order to maintain a minimal outflow at the spring, a Public Utility Declaration imposes to release 160 l/s in the river when the spring is dry. The discharge is therefore, at least for a part, artificial.

In order to build a model able to simulate “natural” outflows at the spring, (Dörfliger *et al.* 2008) have taken into account the relations between: (i) water level in the pool, (ii) overflow discharge, (iii) water table in the boreholes and (iiii)

pumped discharge. Thus, a “natural” discharge chronicle has been generated since 1970 until 2005. Based on this chronicle and in order to characterize the behavior of the *Lez* aquifer, Fleury et al. (2008) have designed a reservoir-based model to simulate spring discharge and water level in the main drain. Its inputs were a combination of rainfall from the three pluviometers cited above. Nash criterion of 0.80 was obtained for the discharge simulation.

### 3. Neural Networks for Flash Flood Forecasting

#### Nonlinear modeling

As pointed out above, there are multiple and complex phenomena which interact to produce outflows at the spring. The rainfall-runoff relation is thus nonlinear, depending on variables that are difficult to measure and/or estimate. Therefore, given the insufficient knowledge about physical processes, machine learning appears suitable for elaborating a nonlinear model of the *Lez* spring flash flooding.

Neural networks are statistical models that calculate their parameters from a calibration dataset; their generalization abilities are evaluated using an independent test set of data. Neural networks have seldom been used to simulate or forecast flash flooding. Several authors note that neural networks cannot generalize learned behavior to the test set (Piotrowski 2006). However, Toukourou *et al.* (2009) showed that *cévenol* flash floods can be forecasted accurately, without rainfall forecasts, provided overfitting is avoided throughout the modeling process by introducing regularization methods: early stopping (Coulibaly *et al.* 2000), model selection and complexity control by cross-validation.

#### Model design

The proposed model is intended, at discrete time  $kT$ , ( $k \in \mathbb{N}^+$ ), to forecast or simulate the outflow at time  $(k + f)T$ , where  $f$  is the forecasting horizon ( $f \in \mathbb{N}^+$ ).

Denoting the predicted value of the quantity of interest by  $q(k)$ , the observed outflow by  $q^p(k)$ , the rainfall vector by  $\mathbf{r}(k)$ , and the nonlinear function implemented by a feedforward neural network by  $g_{NN}$ , the input-output “neural” model was designed based on (Nerrand *et al.* 1993):

$$q(k) = g_{NN}(q^p(k-1), q^p(k-2), \dots, q^p(k-n), r(k), r(k-1), \dots, r(k-w+1))$$

Rainfall information was conveyed to the network as a sliding window of width  $w$ , whose optimal value was chosen as described in Section 4. The information about past outflows was provided by the outflow for the previous day  $q^p(k-1)$

The feedforward neural network used was a multilayer perceptron with one hidden layer of  $N_c$  sigmoid neurons and a linear output neuron. This model was chosen because of its properties of universal approximation and parsimony (Dreyfus 2005).

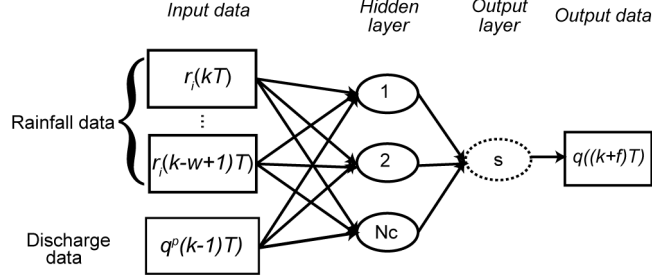


Figure 2: The model is fed by rainfall measurements over a temporal window of width  $w$ . The output is the forecast outflow,  $f$  sampling periods ahead.

The usual squared error cost function was minimized by the Levenberg-Marquardt algorithm (Dreyfus 2005) during calibration. Regularization was performed using early stopping (Coulibaly *et al.* 2000; Toukourou *et al.* 2009)

### Model Selection

#### Model selection by cross validation

Using the database of  $Y=16$  years, one year was selected as the test set, and two years as the stop set (early stopping). On the  $K=13$  remaining years, one year, each one at its turn, was set apart as the validation set; then learning was performed  $K$  times on  $K-1=12$  years (Dreyfus 2005) and the mean squared error was calculated each time on the validation set. The generalization ability of the model was assessed by a Nash criterion cross-validation score:

$$S = \sqrt{\frac{1}{K} \cdot \sum_{i=1}^K Nash_i}, \quad Nash_i = 1 - \frac{MSE_i}{\sigma_i^2}$$

where  $Nash_i$  is the Nash criterion of the forecast outflow, for year  $i$  of the validation set (one complete year,  $(i = 1 \dots 13)$ ),  $\sigma_i^2$  is the variance of the outflows observed at year  $i$  and  $MSE_i$  is the mean squared error calculated on the same year  $i$ . The model complexity was selected in picking up the configuration for which the cross-validation score is the best. The above procedure was used to select the appropriate complexity (*i.e.* the appropriate values of rainfall width  $w$ , and the number of hidden neurons  $N_c$ ). For each model, 50 different parameter initializations were performed. After selecting, a final model was trained for each horizon  $f$ , from 13 sequences: every year except the test year and the early stopping years. Its performance was assessed on the test sequence.

#### Rainfall width selection by cross-correlation

Another way to select the width  $w$  is to use cross-correlation between rainfall and discharge. According to (Mangin 1975), when the correlation is lower than 0.2, we can consider that both variables are independent. Cross-correlations were calculated event by event, the time observed to reach the 0.2 correlation was recorded

for each event: it indicates the period during which rainfall is correlated to discharge, thus determining the required width  $w$ .

## 4. Results and Discussion

Due to the exploitation of the spring since 1982, discharge time series for the *Lez* spring are available from 1988 to 2004, distributed as follows: a 13-year learning set (01/09/1988 - 31/08/1993 and 01/09/1995 - 31/08/2003), a 2-year stopping set (1993-94 and 1994-95) and a 1-year test set (01/08/2003-31/08/2004).

Variables were selected by cross-correlation or cross-validation. Equivalent results were obtained with both methods. Nevertheless, cross-correlation gave slightly more parsimonious models than cross-validation.

### Rainfall width selection using cross-correlation

Rainfall-runoff cross-correlations with the *Lez* spring discharge were calculated for each of the 3 rain gauges. Their maximal cross-correlation values were the same (0.6). One model was then designed for each gauge. Table 1 presents the Nash criteria for simulation ( $f=0$ ) and forecasts at different horizons ( $f=1, 2, 3$ ).

Forecasting horizon $f$	0	1 day	2 days	3 days	Selected architecture
<i>Valflaunès</i>	0.95	0.90	0.84	0.73	8 inputs, 5 hidden neurons
<i>Prades</i>	0.94	0.90	0.84	0.73	9 inputs, 3 hidden neurons
<i>St Martin</i>	0.94	0.90	0.87	0.75	9 inputs, 3 hidden neurons

Table 1 Nash criteria for simulation and forecasting

The Nash criterion was about 0.95 for simulation and 0.84 for two-day forecasting (Fig. 3). Concerning forecasting, peaks remained adequately synchronized whichever the rain gauge used. Both simulation and forecasting proved satisfactory and provided reliable two-day outflow forecasts, respectively.

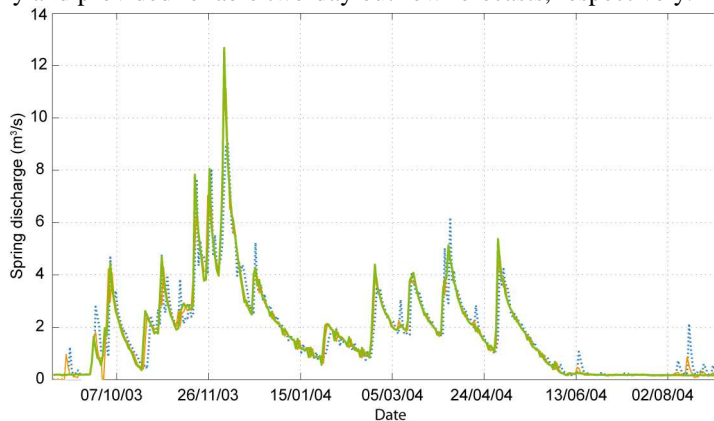


Fig. 3 Simulation and forecasting at a 2-day horizon using *Valflaunès* data: green-bold line - observed discharge, orange-thin line - simulation, blue-dotted line - two-day forecasting.

## Conclusion

Understanding karst behavior during floods is a major issue for improving flood forecasting. The *Lez* karstic system is complex and affected by pumping. Neural networks therefore appeared appropriate for designing a nonlinear model of its behavior. In this study, feedforward neural networks are calibrated for each of the three rain gauges available, using two methods of input selection. Simulation is satisfactory using both methods and all three rain gauges, leading to Nash criteria higher than 0.90. In addition, the selected model is able to correctly predict hydrographs, with a forecasting horizon of two days. Given this encouraging result, a challenging aim is to capitalize on these good modeling performances to obtain information on the physical behavior of the aquifer.

## Acknowledgments

The authors acknowledge METEO-France for providing rainfall datasets, and Gilles LeGac, from the DIREN Montpellier for providing discharge data. We are very grateful to Nathalie Dörfliger, Perrine Fleury, Daniel Diep, Pierre Roussel-Ragot, Bernard Vayssade and Marc Vinches for helpful discussion. We also thank Dominique Bertin for his highly effective collaboration in the design and implementation of the Neural Network simulation tool: RnfPro.

## References

- Coulibaly P, Anctil F, Bobée B (2000) Daily reservoir inflow forecasting using ANN with stopped training approach. *Journal of Hydrology*, 230:244-257
- Dörfliger N, Jourde H, Ladouche B, Fleury P, Lachassagne P, Conroux Y, Pistre S, Vestier A (2008). Active water management resources of karstic water catchment: the example of Le Lez spring (Montpellier, France). In *World Water Congress*. Montpellier
- Dreyfus G, (2005) *Neural networks, methodology and applications*. Springer
- Johannet A, Vayssade B, Bertin D (2008) Neural Networks: From Black Box towards Transparent Box Application to Evapotranspiration Modeling. *Int. Journal of Comp. Int.*, 4(3)
- Jourde H, Roesch A, Guinot V, Bailly-Comte V (2007) Dynamics and contribution of karst groundwater to surface flow during Mediterranean flood. *Environ. Geol.*, 51(5), 725-730
- Mangin A (1975) Contribution à l'étude hydrodynamique des aquifères karstiques. Thèse Université de Dijon. Laboratoire Souterrain du Centre National de la Recherche Scientifique
- Nerrand O, Roussel-Ragot P, Personnaz L, Dreyfus G, Marcos S (1993) Neural networks and nonlinear adaptive filtering: Unifying concepts and new algorithms. *Neural Computation*, 5(2), 165-199
- Piotrowski A, Napiorkowski JJ, Rowinski PM (2006) Flash Flood Forecasting by Means of Neural Networks and Nearest Neighbour Approach – A Comparative Study. *Nonlin. Processes Geophys.* 13: 443-48
- Roesch A, Jourde H (2006) Incidence d'une gestion active de la ressource en eau en milieu karstique sur le risque hydrologique. Exemple du fleuve Lez (Montpellier, France). In *GIRE3D Conference*, Marrakech
- Thiery D, Bérard P, Camus A (1983) Captage de la source du Lez. Etude de la relation entre la source et son réservoir aquifère. Rapport public BRGM/83-SGN-167-LRO. BRGM
- Toukourou MS, Johannet A, Dreyfus G (2009) Flash Flood Forecasting by Statistical Learning in the Absence of Rainfall Forecast: A Case Study. In *European Artificial Neural Networks Conference*. pp. 98-107