



HAL
open science

Deep Multilayer Perceptron for Knowledge Extraction: Understanding the Gardon de Mialet Flash Floods Modeling

Bob E. Saint Fleur, Guillaume Artigue, Anne Johannet, Séverin Pistre

► To cite this version:

Bob E. Saint Fleur, Guillaume Artigue, Anne Johannet, Séverin Pistre. Deep Multilayer Perceptron for Knowledge Extraction: Understanding the Gardon de Mialet Flash Floods Modeling. Theory and Applications of Time Series Analysis - Selected Contributions from ITISE 2019, Springer International Publishing, pp.333-348, 2020, Contributions to Statistics, 978-3-030-56219-9. 10.1007/978-3-030-56219-9_22 . hal-03411377

HAL Id: hal-03411377

<https://imt-mines-ales.hal.science/hal-03411377v1>

Submitted on 2 Nov 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.

Deep Multilayer Perceptron for Knowledge Extraction: Understanding the *Gardon de Mialet* Flash Floods Modelling

Bob E. Saint Fleur ^{1,2}, Guillaume Artigue ¹, Anne Johannet ¹, Séverin Pistre ²

¹ LGEI, IMT Mines Alès, Alès, France

² Hydrosociences Montpellier, Univ Montpellier, CNRS, IRD, 34090 Montpellier, France
guillaume.artigue@mines-ales.fr

Abstract. Flash floods frequently hit Southern France and cause heavy damages and fatalities. To enhance persons and goods safety, official flood forecasting services in France need accurate information and efficient models to optimize their decisions and policy in crisis management. Their forecasting is a serious challenge as heavy rainfalls that cause such floods are very heterogeneous in time and space. Such phenomena are typically nonlinear and more complex than classical flood events. This analysis had led to consider complementary alternatives to enhance the management of such situations. For decades, artificial neural networks have been proved very efficient to model nonlinear phenomena, particularly rainfall-discharge relations in various types of basins. They are applied in this study with two main goals: first modelling flash floods on the *Gardon de Mialet* basin (Southern France); second, extract internal information from the model by using the KnoX: knowledge extraction method to provide new ways to improve models. The first analysis shows that the kind of nonlinear predictor strongly influences the representation of information: e.g. the main influent variable (rainfall) is more important in the recurrent and static models than in the feed-forward one. For understanding "long-term" flash floods genesis, recurrent and static models appear thus as better candidates, despite their lower performance. Besides, the distribution of weights linking the exogenous variables to the first layer of neurons is consistent with the physical considerations about spatial distribution of rainfall and response time of the hydrological system.

Keywords: Neural networks, flash floods, knowledge extraction, deep learning.

1 Introduction

In the Mediterranean regions, flash floods due to heavy rainfalls frequently occur and cause numerous fatalities and costly damages. During the last few years, Southern France has been particularly exposed to these catastrophic events. In such cases, in only one event, there can be more than 20 fatalities, and damages that can reach more than one billion euros, in only one event [1]. Facing these issues, authorities need reliable forecasts for early warning purposes. Unfortunately, both the short-term rainfall forecasts and the processes leading to the discharge response remain poorly known at the space and time scales required. It is thus difficult to provide forecasts using the

traditional coupling between a meteorological model and a physically based hydrological model.

Artificial neural networks therefore appear as an alternative paradigm as they are able to provide forecasts of an output (discharge) without making any other hypothesis on the system than the causality between rainfall and discharge. Artificial neural networks have been applied in a wide variety of domains, as they are essentially based on data and training [2]. They appear as particularly suitable for identifying the generating processes in hydrological time series because of their ability to model nonlinear dynamic systems [3,4]. However, due to their statistical origin, it is difficult to associate meaning to their internal parameters, and they are rightly considered as black-box models. For this reason and to enhance the understanding of the behavior of both the model and the physical processes, several works have been done to bring more transparency in the operating mode and introduced concepts of gray-box and transparent-box models [5, 6]. Some other works have been conducted to make neural networks models more hydrologically meaningful [6, 7, 8].

2 Material and methods

2.1 Study area: location and general description

The *Gardon de Mialet* basin covers 220 sq.km in Southern France. It is part of the *Cévennes* range, which is known as a preferential location for the well-known meteorological phenomenon named “cevenols episodes” (**Fig. 1**). These episodes consist in short duration (less than 2 days) very heavy rainfall events.

The elevation of the *Gardon de Mialet* basin ranges from 150 m.a.s.l. to 1170 m.a.s.l. and its mean slope is about 33 %. As for the most of basins of the *Cévennes*, these characteristics lead to limited deep infiltration or deep underground flow, and thus to a high drainage density. Its response time is relatively short: between 2-4 hours [4]. The area is dominated by a metamorphic formation with 95 % of mica-schist and gneiss, which leads to a poorly porous and impermeable rocky sub-soil. The land use is almost homogeneous while covered by natural vegetation (chestnut trees, conifers, mixed forest and bush) for 92 %. The rest is shared between rocks and urban areas.

Typically, in Mediterranean regions, heavy rainfalls sometimes exceed 500 mm in only 24 h, to be compared to the 600 mm that fall on Paris annually. They are mainly produced by convective events, triggered either by relief, by a wind convergence, or by both. For example, in September 2002, the Gard (France) department has registered 687 mm of rainfall in 24h with 137 mm in only one hour at *Anduze* (a few kilometers distant from *Mialet*).

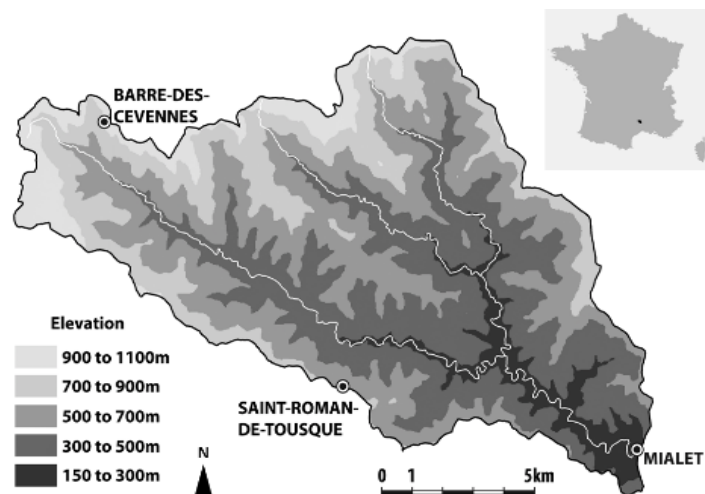


Fig. 1. The study area (Artigue, 2012)

2.2 Database

The database used in this study is essentially compounded with hourly observations from 1992 to 2002, and 5 minutes time-step observations from 2002 to 2008, on three rain gauges and one hydrometric station at the outlet at Mialet (Fig. 1). From upstream to downstream, these stations are: BDC (Barre des Cévennes), SRDT (Saint-Roman de Tousque) and Mialet which coincides with the discharge station. They are all managed by the local Flood Forecasting Service (SPC Grand Delta). 58 events were extracted at 30 minutes time-step (based on rainfall events having at least 100 mm accumulation in 48 h on any of the rain gauges). Data description is synthetized in Tables 1 & 2.

Table 1. Data description

	Rainfall (mm)			Discharge	
	<i>BDC</i>	<i>SRDT</i>	<i>Mialet</i>	(m^3s^{-1})	($m^3s^{-1}km^{-2}$)
<i>Maximum (30 min)</i>	33.3	41.8	62.0	819.3	3.72
<i>Median (30 min)</i>	0.3	0.3	0.2	29.3	0.13
<i>Moy</i>	1.0	1.3	1.2	43.4	0.20
<i>Min</i>	0	0	0	2.13	0.010

Table 2. Test event description

Event	Date	Duration	Maximum of discharge (m^3s^{-1})	Mean discharge (m^3s^{-1})	Cumulative rainfall (mm)	Intensity ($mm.h^{-1}$)
13	Sept. 00	26 h	454	70	230	40

2.3 Artificial neural networks

As widely explained in [9, 4], three kinds of neural networks models have been used in this study: a static model, a recurrent model and a feedforward model. The same references should provide the reader guidance about the implemented methods for the control of the bias-variance dilemma and of overtraining (early stopping, cross-validation, ensemble model) and about the performance criteria used (R^2 criterion and peak analysis). Only the part about knowledge extraction is reminded here, due to its important role in the study.

2.4 Extracting information: KnoX method

First, the KnoX method is applied to a specific architecture, based on multilayer perceptron, which represents the behavior of the physical process, in order to constrain the model to represent this physical behavior [7]. As the rain is essentially added in the first step of the rainfall-runoff transformation, we have introduced one layer of linear neurons implementing the addition of rains fallen at different time-steps (delayed rains). This supplementary layer is called "l" (linear hidden neurons) in Fig 2. The second hidden layer (non-linear hidden layer) calculates a non-linear combination of the "locally added" rains.

The KnoX method [7, 8, 9] allows calculating a "simplified" contribution of each input to the model output. This method is described for the general deep model (two hidden layers) shown in **Fig. 2**. The principle of the method is that a contribution of an individual input variable can be quantified, after training, by the product of the parameter's chain linking this input to the output. The considered parameters are (i) "normalized" by the sum of the parameters linked to the same targeted neuron, and (ii) made independent from the model initialization by calculating the median of absolute values of their values for 20 different random initializations. This regularized value is noted as $^M|C_{ij}|$ for the parameter C_{ij} linking the neuron (or input) j to the neuron i .

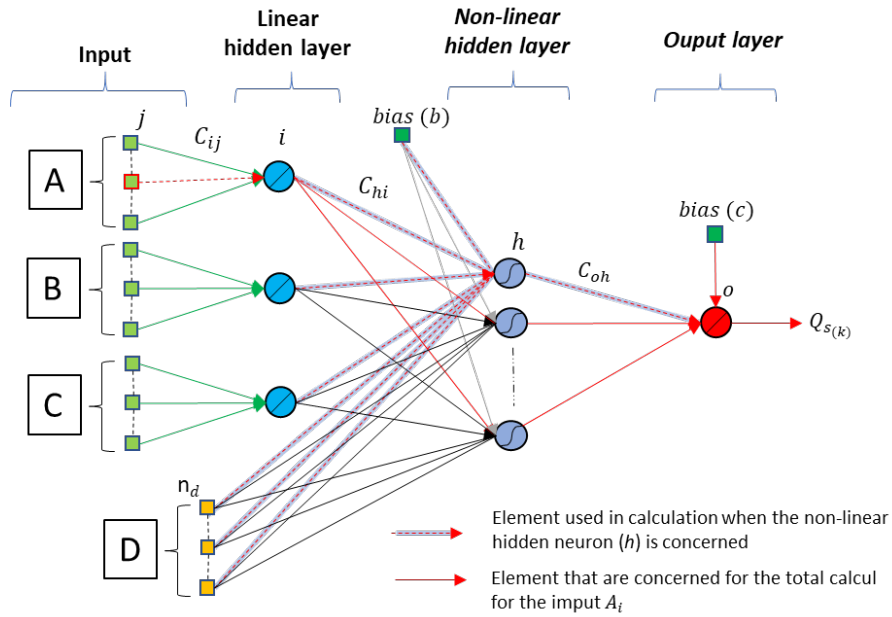
As the value of the sigmoid is not taken into account in eq. 2, this contribution can be seen as the contribution of the "linearized" model. Nevertheless the model is really a non-linear model.

Regarding the model shown in **Fig. 2**, it appears that inputs are applied in several groups, for example A, B, \dots . Each group corresponds to a variable, for example the rain gauge of *Mialet*, or the previous discharge (D). As the output depends dynamically on these inputs, following a complex and unknown multi-scale relation, these inputs are applied at several time-step in order to allow the model to estimate these multi-scale relations. Thus, the contribution (P_A) of the grouped inputs A (including several delayed inputs) is the sum of the contributions of each individual delayed input of the group A . The equation calculating the contribution for just one element (the value for the delay j) of the input A is provided in eq. (1). Unhappily, it is not possible to explain more comprehensively the method in the short present paper, so we suggest to the reader to refer to [7, 8].

$$P_{A(j)} = \frac{M|c_{ij}|}{\sum_{i=1}^{n_A} M|c_{ij}|} \sum_{h=1}^H \left(\frac{M|c_{hi}|}{\sum_{i=1}^{n_i} M|c_{hi}| + \sum_{d=1}^{n_d} M|c_{hd}| + b_h} \right) \left(\frac{M|c_{oh}|}{\sum_{h=1}^H M|c_{oh}| + c_o} \right), \quad (1)$$

$$\text{and: } P_A = \sum_{j=1}^{n_A} (P_{A(j)}), \quad (2)$$

Where the categories of parameters C_{ij} , C_{hi} , C_{oh} , C_{hd} are shown on Fig. 2; n_A is the number of inputs in the group A ; H the number of hidden non-linear neurons; n_j the number of hidden linear neurons (first hidden layer); n_d is the number of delayed inputs of the group D ; b_h is the bias inputted to the non-linear hidden input and c_o is the bias parameter inputted to the output neuron.



3 Results

3.1 Choice of variables

Starting from previous works of [4], we chose the following exogenous variables: (i) *Barre des Cevennes* rain gauge, *Saint-Roman de Tousque* rain gauge and *Mialet* rain gauge, each one with a sliding window length $\{k, \dots, k-n_r+1\}$, (ii) the sum of the mean rain (mean calculated over the three gauges) fallen from the beginning of the event. Of course, a bias input is used; several values were tried in order to evaluate the sensitivity of the KnoX method to its value.

Depending on the kind of considered model, states variables can be added: previous observed discharges for the feed-forward model, and previous estimated discharges for the recurrent one. The static model only takes rains and mean rains into account [9].

3.2 Model selection

Model selection is a key issue of machine learning. The goal is to define accurately the architecture of the model managing the bias-variance tradeoff. This was done in this following work [10] using cross correlation, cross validation, and early stopping using the following rules.

- Hyper-parameters are adjusted for each one of the three kinds of model (static, feed-forward, recurrent): input sliding windows width (n_A, n_B, n_C, n_D), number of non-linear hidden neurons (h).
- Widths of the rainfall windows applied to the model, $\{n_A, n_B, n_C\}$, are selected thanks to cross correlation [11]. Initially proposed by [12], [13] generalizes the application of cross correlation in hydrology. The used equation in this study is presented in eq. (3).

$$C_{xy}(k) = \frac{\text{Cov}(x_i, y_{i+k})}{\sigma_x \sigma_y} = \frac{\frac{1}{n} \sum_{i=1}^{n-k} (x_i - \bar{x})(y_{i+k} - \bar{y})}{\sigma_x \sigma_y} \quad (3)$$

With $k \in \mathbb{N}^+$; the truncation m , which is the maximum value of k , is recommended to be $m=n/3$. [12] indicated that two hydrological variables can be considered as statistically independent if their cross-correlation is superior to 0.2. Starting from this work, we selected three possible lengths for the sliding windows of rain gauges inputs: (i) the number of time-step between $C_{xy} = 0$ and $C_{xy} = 0.2$, that defines the memory effect (called *memory window*); (ii) the sliding window between $C_{xy} = 0.2$ (positive slope) and $C_{xy} = 0.2$ (negative slope) (called *strong correlation window*), and (iii) all the m positive values of C_{xy} (called *full correlation window*). Based on [12], the correlations between gauges and response times are indicated in **Table 3**.

Table 3. Correlation analysis of the data

Rain gauge	Mialet (h)	SRDT (h)	BDC (h)
Average response-time	2	3	4.5
Response-time range	1 – 3.5	2.5 – 4.5	4 – 5.5
Rainfall-discharge average cross-correlation	0.40	0.455	0.44
Rain gauge cross-correlation	Mialet	--	0.58
	SRDT	--	0.61

- The partial cross-validation score was operated on a subset of K events, the 17 most intense events in the database [3].

- The number of hidden neurons was increased from 1 to 10. The best model was chosen according to the highest cross-validation score S_v estimated as following:

$$S_v = \frac{1}{K} \sqrt{\sum_{i=1}^K |E_i|^2} \quad (4)$$

Where E_i is the validation error of the subset i used in partial cross validation.

- An ensemble model is used in order to regularize on the initialization of parameters; moreover, the output values are the result of the median of the outputs of an ensemble of 20 members differing only by their initialization before training [8].

- Three bias values were considered (0.01; 0.1; 1), three depths of sliding windows and three kinds of models (see section 2.3), *i.e.* 27 different models have been designed following the procedure indicated in [9]. The best one in each kind of models has been chosen, regarding the test event, in order to have the most efficient models to analyze.

Architectures presented in **Table 3** were thus selected.

Table 3. Selected models

Input variables	Static	Recurrent	Feed-Forward
Rain-gauge window width (n_r) (BDC/SRDT/Mialet)	32/32/23	27/28/20	32/32/23
Cumulative rainfall window width	3	3	3
Order (r) / Previous observed outputs	x	3	3
Number of hidden nonlinear neurons	2	10	5
Bias value	1	0.01	0.01

3.3 Discharge estimation

As shown in [9, 4], the best results are provided by the feed-forward model. This is usual because the feedforward model uses the previous observed output as a state variable. The recurrent model is not as efficient but exhibits better dynamics, which is also frequently observed [4]. The static model presents an acceptable performance, being able to generate 63% of the peak discharge (**Table 4; Fig 3**).

Table 4. Models performances on the test set

<i>Model</i>	R^2	<i>SPPD %</i>	$P_D(0.5h)$
Static	0.83	63,3	1
Recurrent.	0.89	78.5	0
Feed-Forward	0.99	99.3	1

3.4 Contributions of input variables

After having verified that the models are convenient, it is possible to apply the KnoX method. The extracted contributions are presented in **Table 5** [9].

It is interesting to compare the relative weights of the three rain gauges with a classic method dedicated to distribute rainfall on a watershed and widely used in hydrology: the Thiessen polygons method (or Voronoï polygons). This comparison is presented in **Fig. 4**.

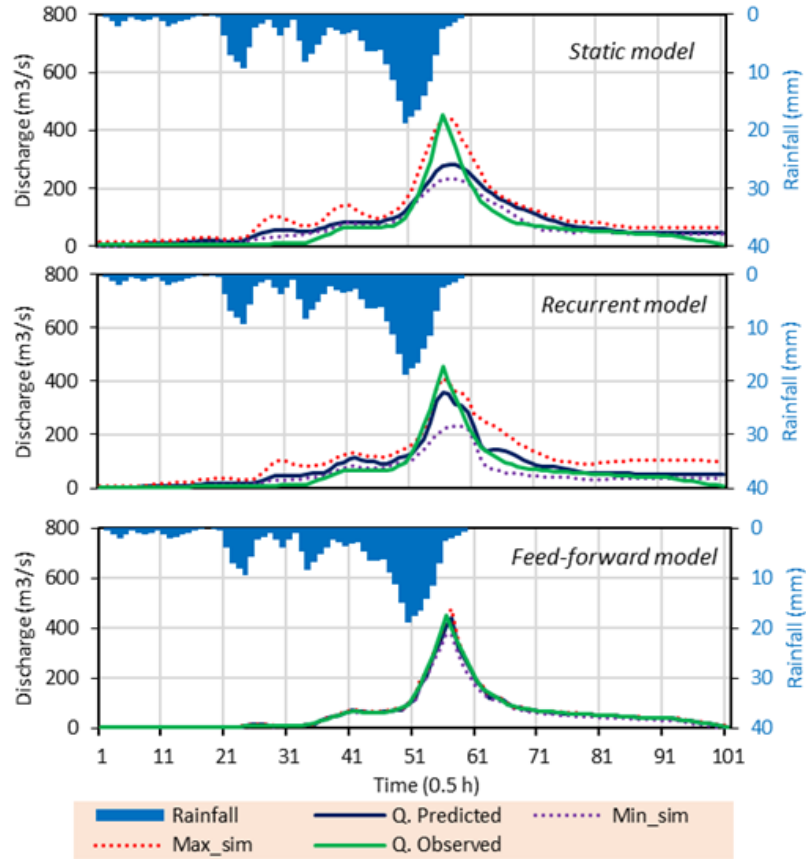


Fig. 3. Hydrographs for the test set. Min_sim and Max_sim correspond to the minimum and maximum values of the ensemble model. Q is the median of the 20 members of the ensemble.

Table 5. Contributions (P_x) for the variables, from each model, expressed in %.

Name of variable X	Static	Recurrent	Feed-forward
<i>BDC</i>	11 %	10 %	9 %
<i>SRDT</i>	31 %	17 %	22 %
<i>Mialet</i>	13 %	12 %	5 %
<i>Cumulated rainfall</i>	31 %	20 %	12 %
<i>Previous Q. obs</i>	--	--	45 %
<i>Previous Q. calc</i>	--	25 %	--
<i>bias</i>	14 %	16 %	7 %
<i>Total</i>	100 %	100 %	100 %

As *Mialet* (MIA) is at the outlet of the basin and *Barre-des-Cévennes* (BDC) at the top of the basin, they are both represented with less contribution than *Saint-Roman-de-Tousque* (SRDT, near the middle of the basin) by the Thiessen polygon method. It is more or less also the case for the neural network models, with a very similar distribution

to Thiessen distribution for the static model, being a little more different for the recurrent model and even more for the feed-forward model (providing the best results).

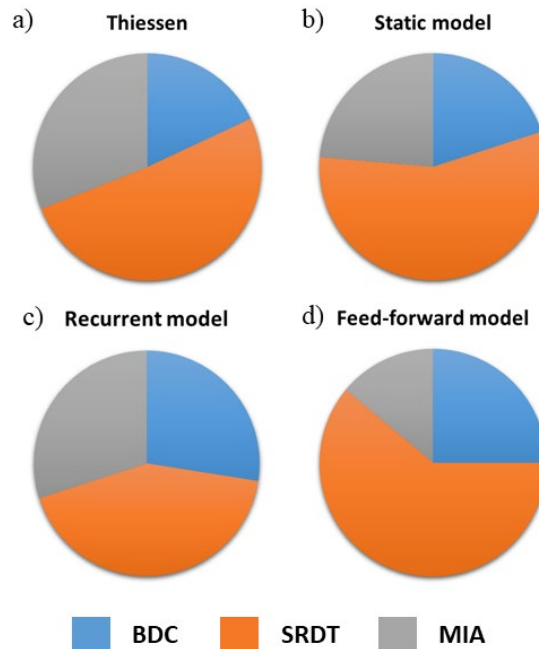


Fig. 4. Thiessen method weights (a) and relative weights from the models of the three rain gauges (b, c, d).

3.5 Results: contributions as a function of time windows

Here, we have considered the distribution of contributions amongst the time delay in the first layer of parameters (arriving to the linear neurons in Fig 2). We compare the sum of these contributions (for the three rain gauges) to the cross-correlogram of the average rainfall (average of the three rain gauges) and the discharge. This comparison involves the three selected models presented in section 3.1 (Fig. 5).

The static model shows the greatest similarity with the cross-correlogram, for the total contributions and for the relative contributions of each rain gauge. Regarding the response time (time corresponding to the peak of the cross-correlogram) the static model seems also to be the best. This result is logical because the variables taken into account by the static models are similar to those considered by the cross-correlogram: only rains.

For the three models, the SRDT rain gauge is the most represented in most of the time lags considered and not only in general, as shown in section 3.4.

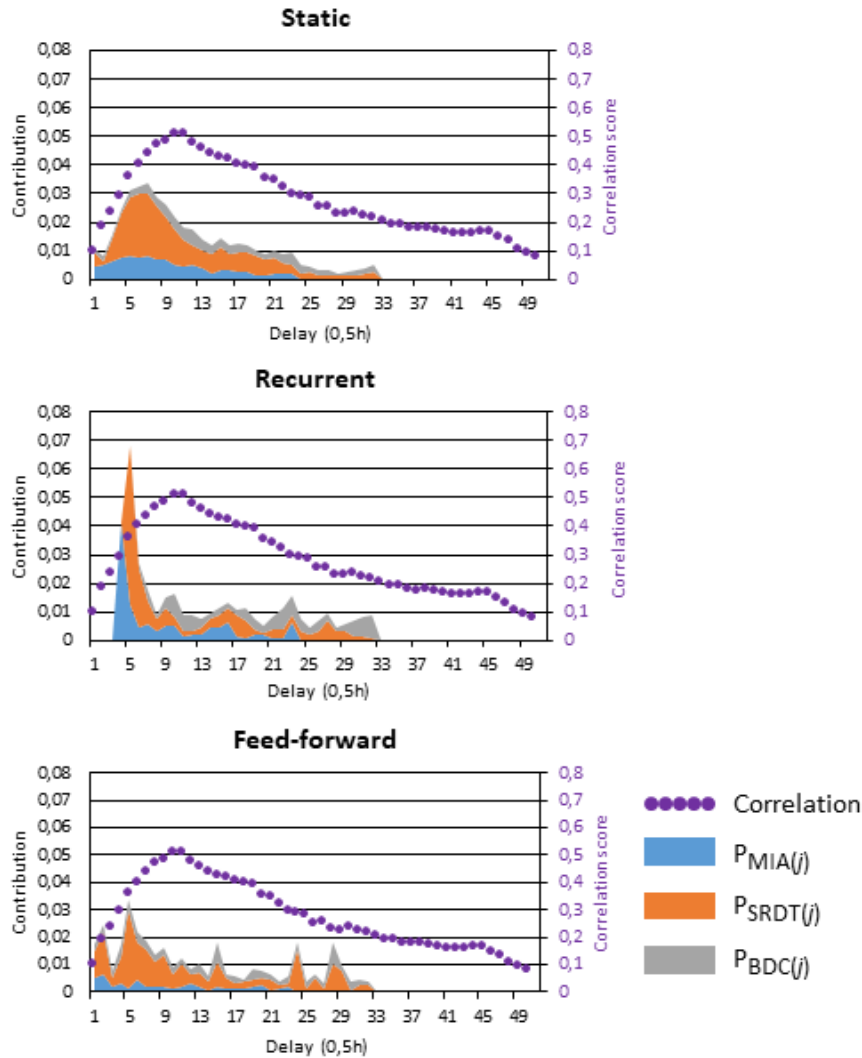


Fig. 5. Cross-correlogram (mean rainfalls-Mialet discharge) and distribution of contributions calculated as indicated in section 2.4.

3.6 Results: effects of the bias

Before obtaining the selected models, many different combinations were tried during optimization. Among these combinations, three values of bias have been experimented, each separated by an order of magnitude: 0.01, 0.1 and 1. The Fig. 6 shows the contributions of *Saint-Roman-de-Tousque* amongst the delays of the input time window, for the three types of models (static, recurrent, feed-forward) and for the three time-windows defined in section 3.2 (memory, strong correlation and full correlation

windows). The other rain gauges have not been presented here due to the large number of figures it would have produced; but the *Saint-Roman-de-Tousque* station is representative of the three rain gauges from this point of view.

It can be noticed that the bias value does not deeply impact the contributions of the input variables. In particular, it does not change the general shape of these contributions even if in some cases, moderate amplitude differences appear.

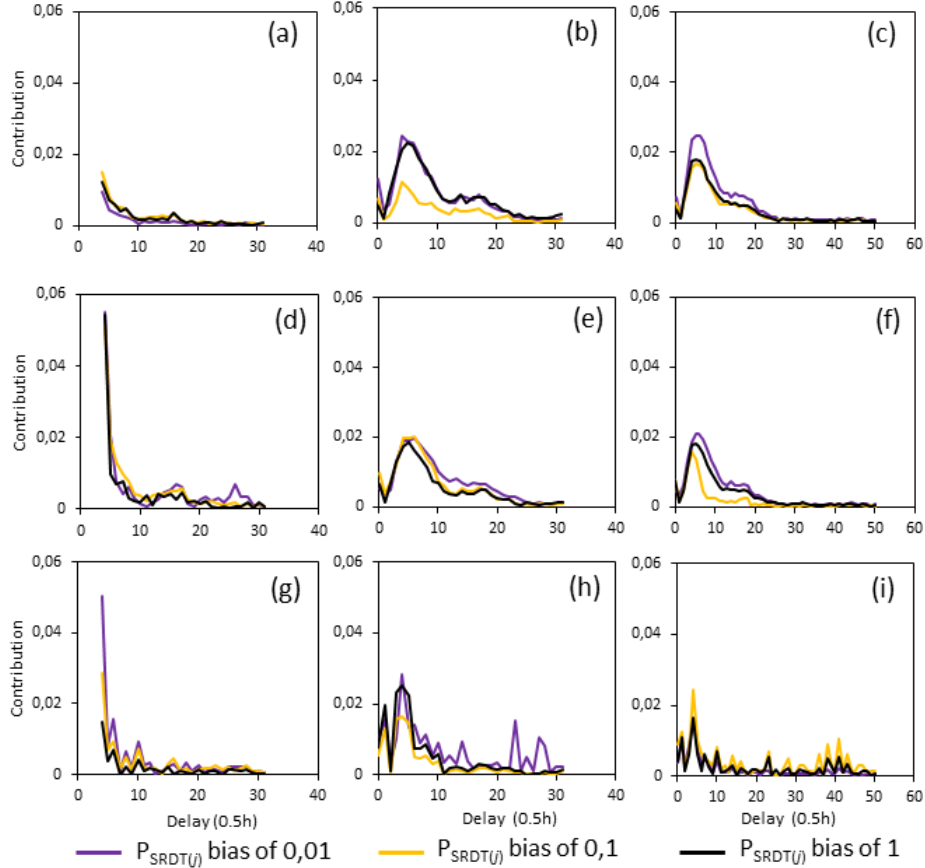


Fig. 6. *Saint-Roman-de-Tousque* contributions calculated as indicated in section 2.4 with different bias and in different modeling configurations: (a), (b) and (c) are for static models; (d), (e) and (f) are for recurrent models; (g), (h), and (i) are for feedforward models whereas (a), (d) and (g) are for memory windows; (b), (e) and (h) are for strong correlation windows and (c), (f) and (i) are for full correlation windows.

4 Discussion

These results show how the kind of model takes into account explanatory variables on an observed phenomenon. Even if they use the same exogenous variables in the same

context, their performances and behaviors are different due to their configuration and architecture.

4.1 Selecting a model type for physical knowledge extraction

Analyzing the contributions assigned to each input variable (Table 5), it appears that:

- The static model strongly uses exogenous variables (total contribution of 55%) and uses an important contribution (31%) to the cumulated rainfall that is useful to represent the soil saturation and could thus be considered as a substitute to a state variable.
- The recurrent model uses mostly previous estimated discharge (25%), whereas the total contribution of cumulated rainfall (20%) and of exogenous variables (40%) is lower than for the static model.
- The feedforward model uses a smaller contribution for rains (12% for cumulated rainfall and 36% for exogenous variables) whereas previous observed values of discharge contribution is predominant (45%).

As foreseen by [14], the optimal type of model is strongly linked to the quality of explanatory information that is given to the model during the training phase. Here, we show that, despite its low performance, the static model is forced to represent, the physical relationship between exogenous variables and the output, whereas the recurrent model and the feedforward model are helped in this task resp. by the previous estimated or observed discharge. Consequently, the total use of the exogenous variables decreases when state variables information increases. If we compare the relative contributions of the three rain gauges with the Thiessen polygons, we observe a decrease of the similarity while state variables are added.

Finally, in this study, the best tradeoff between model performance and knowledge extraction capacities seems to be provided by the recurrent model. Nevertheless, this conclusion is based on one test set, it should be confirmed by further studies.

4.2 Response time and contributions

The cross-correlation provides a simple linear representation of the behavior of the modeled system and allows estimating the response time. Here again, while the contributions of state variables appear, the similarity with the correlogram decreases. This could be interpreted as a confirmation that recurrent and feed-forward models represent well the behavior that takes profit of the rich information provided by the previous discharge input (estimated or simulated): the accumulation of previously fallen rains. The less the model is helped by the previous discharge input, the more it is forced to represent well the role of recent and ancient rainfalls. This appears in Fig. 6.d and 6.g with great values of recent rains contributions; on the equivalent contributions in 6.e and the "noisy" contribution of the Fig 6.i (feed-forward with the maximum window-width).

4.3 Bias input importance

The bias input plays a role that is usually interpreted in hydrology as the base flow (remaining discharge when there is no rainfall). In this case, its contribution is consistent: it is significantly less involved in the calculation of the output when the previous observed discharges are used as input (the previous base flow is thus applied by the inputs). In the other cases, it seems to guide the models to acceptably approximate the discharge information when necessary.

If the bias input seems necessary to guide the model, its value does not deeply change the distribution of the contribution of the rain gauges as a function of the instant of the time window. One could suppose that changing an order of magnitude in the bias input value can easily be counterbalanced during the training step by applying a proportional modification to the weights applied to this input.

5 Conclusions and perspectives

Flash flood forecasting is a very challenging task, especially in the *Cévennes* range. Several examples of robust forecasts using neural networks have been published but the results did not always allow understanding how close the model was to the physical behavior of the basin, in addition of being close to the observed output. The obtained results prove again that when using relevant and properly combined variables on any of the networks used here, an efficient model can be implemented.

Nevertheless, enhancing these models and applying them to an increasing number of basins, in a context of climate change, and with various characteristics, requires a better understanding of the processes involved in their operation as well as in such flood events. For this purpose, the KnoX method, developed to extract information from a neural network model, was applied to the *Gardon de Mialet* basin. This method allows understanding how the variables are handled by the model to approximate the modeled phenomenon. First it appears that the bias input was consistently used to model the base flow. Then, interestingly, there has been evidence that the variables do not express themselves in the same way depending on the different models used. It was known that the choice of a model must be driven by the modeling goal (for example a recurrent model for a long-term prediction). Besides being driven by the modeling goal, it appears that the choice for a model might be guided by the situation: availability (real-time and historical) of data, quality and explanatory nature of the data. In this study, this results in three kinds of model: static, recurrent, feed-forward, showing increasing performances while taking into account more realistic state variables. On the other hand, if assess the performance of a kind of model by the ability to extract physical information from it, the ranking is reversed and the less the model considers state variables, the more the design of its estimator will adopt behaviors that mimic the physical processes.

Finally it appears that the KnoX method shows very interesting capabilities; the next steps will consist in generalizing this method to other sites and other rainfall events in the *Cévennes* range, with an increasing complexity in the physical processes to extract (dams and/or karst systems for example)..

6 Acknowledgements

The authors thank IMT Mines Alès that funded the present study; the SPGD flood-forecasting agency for providing datasets. Our gratitude is extended to B. Janet for the stimulating collaboration shared with the SCHAPI and to R. Moussa and P. Roussel-Ragot for the helpful discussions and support. The constant effort made by D. Bertin and the Geonosis Company to enhance and develop the neural network software RNF Pro are thereby acknowledged as well.

References

1. Rouzeau, M., Martin, X., Pauc, J.C. Retour d'expérience des inondations survenues dans le département du Var les 15 et 16 juin 2010. http://cgedd.documentation.developpement-durable.gouv.fr/documents/cgedd/007394-01_rapport.pdf (2010).
2. Roberts, S. J., Penny, W. Neural networks: friends or foes? *Sensor Review*, 17(1), 64-70 (1997).
3. Toukourou M., Johannet, A., Dreyfus, G., Ayrat, P.A. Rainfall-runoff Modeling of Flash Floods in the Absence of Rainfall Forecasts: the Case of "Cévenol Flash Floods". *App. Intelligence*, 35 2,178-189 (2011).
4. Artigue, G., Johannet, A., Borrell, V., Pistre, S. Flash Flood Forecasting in Poorly Gauged Basins Using Neural Networks: Case Study of the Gardon de Mialet Basin (Southern France). *NHESS*, 12(11): 3307-24 (2012).
5. Oussar, Y., Dreyfus G. How to Be a Gray Box: Dynamic Semi-Physical Modeling. *Neural Networks* 14 (9): 1161-72 (2001).
6. Johannet, A., Vayssade, B., Bertin, D. Neural Networks: From Black Box towards Transparent Box - Application to ETP Modelling. *Int. J. of Comp. Int.* 24 (1): 162 (2007).
7. Kong-A-Siou, L., Cros, K., Johannet, A., Borrel-Estupina, V., Pistre, S. KnoX method, or Knowledge eXtraction from neural network model. Case study on the Lez karst aquifer (southern France), *J. Hydrol.*, 507, 19–32 (2013).
8. Darras, T., Borrel-Estupina, V., Kong-A-Siou, L., et al. Identification of spatial and temporal contributions of rainfalls to flash floods using neural network modelling: case study on the Lez basin (southern France). *Hydrol. Earth Syst. Sci.*, 19, 4397–4410, (2015).
9. Saint-Fleur, B., Artigue, G., Johannet, A., Pistre, S. Knowledge Extraction (KnoX) in Deep Learning: Application to the Gardon de Mialet Flash Floods Modelling. In: *Proceedings ITISE-2019*, pp. 178-189. Granada, 25th-27th September (2019).
10. Kong-A-Siou, L., Johannet, A., Borrell V., Pistre, S., 2012. Optimization of the generalization capability for rainfall–runoff modeling by neural networks: the case of the Lez aquifer (southern France). *Environ Earth Sci* 65, 2365–2375.
11. Kong-A-Siou, L., Johannet, A., Borrell, V., Pistre, S., 2011. Complexity selection of a neural network model for karst flood forecasting: The case of the Lez Basin (southern France). *Journal of Hydrology* 403, 367–380.
12. Jenkins, G.M., Watts, D.G. *Spectral analysis and its applications*. Holden-Day (1969).
13. Mangin, A. Pour Une Meilleure Connaissance Des Systèmes Hydrologiques à Partir Des Analyses Corrélatoire et Spectrale. *Journal of Hydrology* 67 (1-4): 25-43 (1984).
14. Nerrand, O., Roussel-Ragot, P., Personnaz, L., Dreyfus, G., Marcos, S. Neural Networks and Nonlinear Adaptive Filtering: Unifying Concepts and New Algorithms. *Neural Comp* 5 (2): 165-99 (1993).