



Assessment of prediction performances of stochastic and conceptual hydrological models: monthly stream flow prediction in northwestern Algeria

Ouafik Boulariah¹ · Mohamed Meddi¹ · Antonia Longobardi²

Received: 14 January 2019 / Accepted: 12 September 2019 / Published online: 13 December 2019
© Saudi Society for Geosciences 2019

Abstract

Determining the link between rainfall and flow for a watershed is one of the most imperative problems and challenging tasks faced by hydrologists and engineers. Conceptual and Box-Jenkins hydrological models represent suitable tools for this purpose in circumstance of data Scarce and climate complexity. This research consists in a comparative study between conceptual models and Box-Jenkins model, namely, GR2M, ABCD, and the autoregressive moving average (ARIMA) which has a numerical design. The three models were applied to three catchments located in the north-west of Algeria. Basins have been selected according to the availability of long-time series of hydrological and climatic data (more than 30 years) to calibrate parsimonious models, taking into account the climatic variables and the stochastic behavior of the natural stream flow. Overall, the conceptual models perform similarly; whereas the results show that the GR2M model performed better than the ABCD in the validation stage, the stochastic model shows better results as opposed to conceptual models in the case of the Mellah Wadi which presents high permeability in its behavior. This is due to the simplicity of the model needed for data (only runoff data) and the ability of the stochastic model to produce stream flow in complex catchments. Such circumstance could be caused by different motivations. On the one hand, the diverse number of model parameters that make the ABCD the less parsimonious approach, with four parameters to be calibrated. On the other hand, the inability of the ABCD and the ARIMA model to capture and describe the groundwater processes, important for the cases study. Moreover, the validation period includes a large drought period, started in the late 1980s, which makes difficult model adaptation to different hydrological regimes.

Keywords Statistical comparison · GR2M · ABCD · ARIMA

Introduction

Understanding and modeling hydrological behavior is an important challenge for managing the water resources and for the investigation of extreme hydrological events, such as floods and droughts. However, a substantial topic in many areas is that meteorological and hydrological data accessibility which is frequently rare. Several problems associated with obtaining

reliable long-term hydrological data in semi-arid regions include limited economic resources for monitoring and climates complexity (Beaumont et al. 2016; Wheater 2002). The application of accurate modeling procedures to predict and simulate water availability is an indispensable way for flood protection, irrigation, drought management and water needs. According to Refsgaard and Knudsen (1996) and Wagener et al. (2004), two major categories of rainfall-runoff models can be notable—deterministic models and stochastic models—the first family can be classified into black box models, such as artificial neural networks ANN and statistical models such as Box-Jenkins ARIMA model and conceptual or physically based models such as the GR family. Rainfall-runoff models differ generally in underlying design philosophy. The prior contains forecasting applications (flood warning and hydropower operation) and catchment management purpose (climate impact studies and water distribution); the

Responsible Editor: Fernando Al Pacheco

✉ Mohamed Meddi
m.meddi@ensh.dz

¹ LGEE–Ecole Nationale Supérieure d’Hydraulique de Blida, B.P.31, 09000 Blida, Algeria

² Department of Civil Engineering, University of Salerno, Fisciano, SA, Italy

conclusions are geared on the way to understanding the catchment function and how the individual processes combine to produce catchment response (Parajka et al. 2007). The choice of suitable hydrological modeling for a particular catchment with a special hydrological environment and a distinct physical background should be established on a diversity of aspects containing data availability, model complexity, and prediction uncertainty among others. Consequently, for making a decision on model selection a comparison between the pros and cons of a variety of hydrological models is necessary. In the literature, several studies elaborate the subject of comparative analysis of different hydrological models (Bai et al. 2015; bin Shaari et al. 2017; Machekposhti et al. 2018; McCuen 2016; Mohammadi et al. 2005; Singh et al. 2005; Valipour et al. 2013). In the current research, various models will be used: rural monthly engineering model with two parameters (GR2M), Thomas ABCD model and the non-seasonal autoregressive integrated moving average (ARIMA) models, are involved to predict flow in three basins in the northwestern Algeria.

Conceptual models (occasionally called gray-box models) are transitional between theoretical and empirical models. Generally, conceptual models consider physical laws but in extremely simplified form. There is a big variety of models which belong to this class; examples which are familiar to several researchers are HBV, GR, and the Thomas ABCD model (Xu 2002). Conceptual models, especially GR2M, have been widely employed for predicting and simulating different hydrological processes: drought forecasting (Belarbi 2017), assessing hydrologic changes (Lyon et al. 2017), and quantifying the impact of human activities and climate change on stream flow (Ahn and Merwade 2014). Researchers sub-cited in references argued that this kind of models presents good capacities to reproduce the flows from the rainfall data under climate variability.

Traore et al. (2014) apply the GR2M and the GR4J models for evaluating water resources of the Koulountou River Basin, a tributary of the Gambia River where they found that the GR2M shows the best performance and easier for calibration and application, and they conclude that it can be used to restore gaps in series of flows. Bai et al. (2015) compared the performance of 12 monthly water balance models in different climatic catchments of China. They found the GR model to be more suitable for monthly river predicting. They conclude that the GR2M two-parameter model is sufficient to give a good performance for monthly runoff simulation. Sharifi (2018) used the GR2M two-parameter monthly water balance model for simulating monthly seasonal and annual runoff in Kalu and Mahaweli rivers of Serilanca. He found that the model with the simple structure and two parameters proved to be a very efficient model when simulating runoff in different time scales. It was concluded too that this two-parameter monthly water balance model can be easily and

efficiently used for the water resources planning and management; this is due to its plainness and great productivity in performance.

The ABCD model is a conceptual water balance model well known in modeling different hydrological processes. Alley (1984) studied the comparison of three water balance models; the results showed that the ABCD model had the best performance, especially in its capacity to reproduce the observed behavior of ancillary catchment states (e.g., groundwater). Wang et al. (2011) used the ABCD model for simulating monthly stream flow in the Australian catchments; they found that the ABCD model is a good performer at simulating low flow compared with daily water balance model. They concluded that monthly water balance is more advantageous compared with daily water balance model for the reason that the monthly one requires less data with a low computational cost. In order to estimate the impact of climatic variation on the stream flow in the Yiluo River in China (Liu et al. 2013) carried the ABCD water balance model to see the attribution of change of stream flow, the results showed that the ABCD model proved a high accuracy for long-term simulation of stream flow. Currently, mathematical models have taken over the most important tasks of problem answering in hydrology. A mathematical model expresses the system behavior by a set of equations, perhaps together with logical declaration stating relationships between variables and parameters without taking into account the groundwater behavior of the catchment (Xu 2002). Box-Jenkins models became one of the universal models in estimating hydrological time series, particularly after publishing the work of (Series 1976). An ARIMA non-seasonal model is an overview of an ARMA model. This model is applied in some cases where data clarify suggestion of non-stationary, where an initial differencing step (corresponding to the “integrated” part of the model) can be applied one or more times to remove the non-stationary. Therefore, ARIMA models are non-static and cannot be used to estimate the missing data. However, these models are more suitable for estimating changes in a process (Karamouz and Araghinejad 2012). The non-seasonal autoregressive moving average (ARIMA) models can be employed in a varied range of requests in many hydrological modeling problems such as stream flow prediction (Al-Juboori and Guven 2016; Mohammadi et al. 2005), rainfall–runoff modeling (Castellano-Méndez et al. 2004; Farajzadeh et al. 2014; Machekposhti et al. 2018), and drought forecasting (bin Shaari et al. 2017; Mishra and Desai 2005). Ahsan and O’Connor (1994) stated that Box-Jenkins model appears quiet reliable for predicting flow time series with a minimum mean square error when the flow forecasting model is expected to be autoregressive moving average model and the corresponding flow data are free of measurement errors. Abudu et al. (2010) applied autoregressive integrated moving average (ARIMA), seasonal ARIMA (SARIMA) and Jordan-Elman artificial

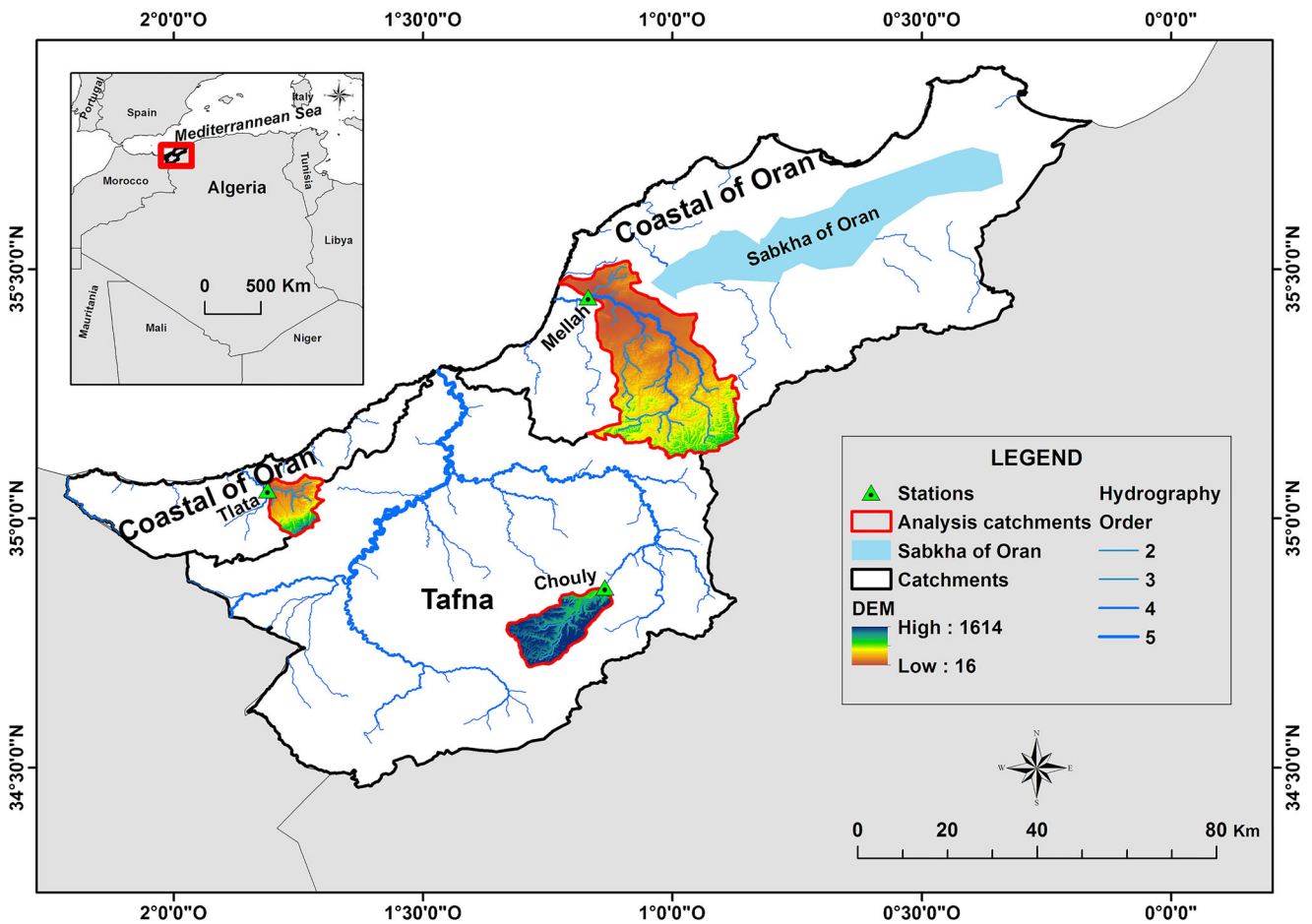


Fig. 1 Location of study basins rainfall, flow, and climate stations

neural networks (ANN) models in forecasting monthly stream flow in the Kizil River, Xinjiang, China. The results indicated that ARIMA model performed similarly to ANN models. Each of these different hydrological models has its own potential, advantages, and disadvantages. Accordingly, it is important to compare them from the standpoint of these procedures in terms of prediction accuracy and efficiency on the one hand, and their implantation in different climate conditions and catchments characteristics, on the other hand. The purpose of this study is, therefore, to evaluate whether the parsimonious hydrological models are able to simulate historical stream flow time series in northern Algeria. The main objective of the research is to compare conceptual rainfall-runoff hydrological models (ABCD and GR2M) with a stochastic modeling approach (ARIMA). Relevant to this, the following research

questions are put forward:

- (i) Which model will be suitable to predict flow in this special type of complex climate?
- (ii) Have geological structures of the catchments an important effect in the simulation of these models?

- (iii) Does the degree of complexity of the model affect the accuracy of the prediction?

In Algeria, GR models have been widely used (Belarbi et al. 2017; Merabet-Baghli et al. 2018; Otmane et al. 2017; Paturel et al. 2017). They gave varying results in terms of robustness according to the geographical regions of the country. On the other hand, ABCD and ARIMA models have not been tested before; hence, the importance of this work that will allow comparison of these different approaches to find out the best alternative. Flow forecasting is a very important step in addressing the information needed for hydraulic studies, engineering structures, and protection studies.

Materials and methods

Study area

Algeria is part of the semi-arid areas which cover over 40% of the world's land surface and which are more exposed to the phenomenon of climate change (Boudjadja et al. 2003; Meddi

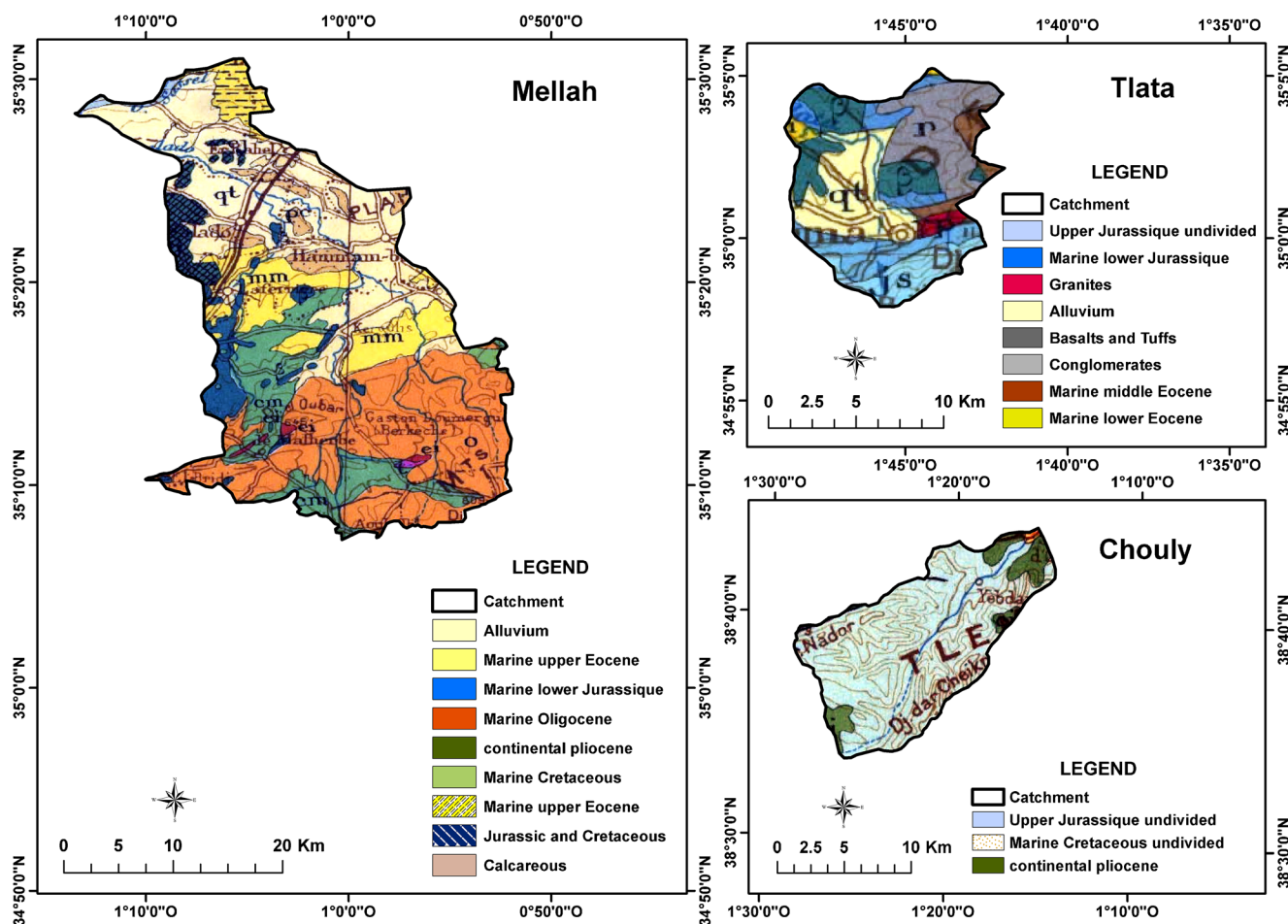


Fig. 2 General map of the massif Oranian coastal

and Hubert 2003). The management of the water resource in North Africa is more complex than it is in tropical zones due to the deficiency of perennial wadis and other readily available water sources (Souza et al. 2016). In this study, we are interested in the northwestern part of the Algerian territory. From 1970 to 2004, the country underwent the most severe drought period, with increases in the annual average temperature (Bakreti et al. 2013) and reduction of rainfall about 25% (Meddi and Hubert 2003). The location of the studied basins extends from $34^{\circ}, 24'$ to $35^{\circ}, 48'$ north latitude and $2^{\circ}, 12'$ to $0^{\circ}, 12'$ east longitude. The highest elevation of the study area is located in the Mont of Telemcen with 1614 m and the lowest is 16 m at the Mellah hydrometric station in the Tlata basin. The present research has focused on three sub-basins, two located in the Oranian coastal area (Wadi Tlata and Wadi Mellah) and the third in the Tafna basin (Wadi Chouly) (Fig. 1). Their drainage areas are 95 km^2 , 720 km^2 , and 170 km^2 respectively.

Geologically, the study area is occupied by a geological series ranging from Primary to the Quaternary period, and it is hollowed out in a material of high variable resistance. Based on the primary schisto-quartzitic substratum and secondary

carbonate formations, tertiary sediments were deposited, predominantly Miocene clays, sandstones, and quaternary alluviums occupying the valley bottom and plains (Benest 1985). Heading to the Mediterranean Sea and in the east of the study area, according to Fig. 2, we notice the presence of the Large Sabkha of Oran. This Sabkha consists of a vast salty lake surrounded by successive strata to the north (Misserghin, Amria, Bou-Tlelis, etc.) and the great plain of Mleta to the south.

The northern part of Algeria is characterized by a Mediterranean climate with a relatively cold and rainy winter and a hot and dry summer. The annual rainfall reaches 400 mm in the west, 700 mm in the center, and 1000 mm in the east for the coast. This type of climate is also found in the Tellian Atlas chains where we record totals ranging from 800 to 1600 mm in the eastern summits, while the values are lower in the center (700 to 1000 mm) and in the west (600 mm). In the plains of the Tellian Atlas, rainfall varies between 500 mm in the west, 450 mm in the center and 700 mm in the east. The pluviometric regime is characterized by a decrease of about 25% since 1975 (Meddi and Hubert 2003) and confirmed by Taïbi et al. (2016). The hydrological regime is variable. The

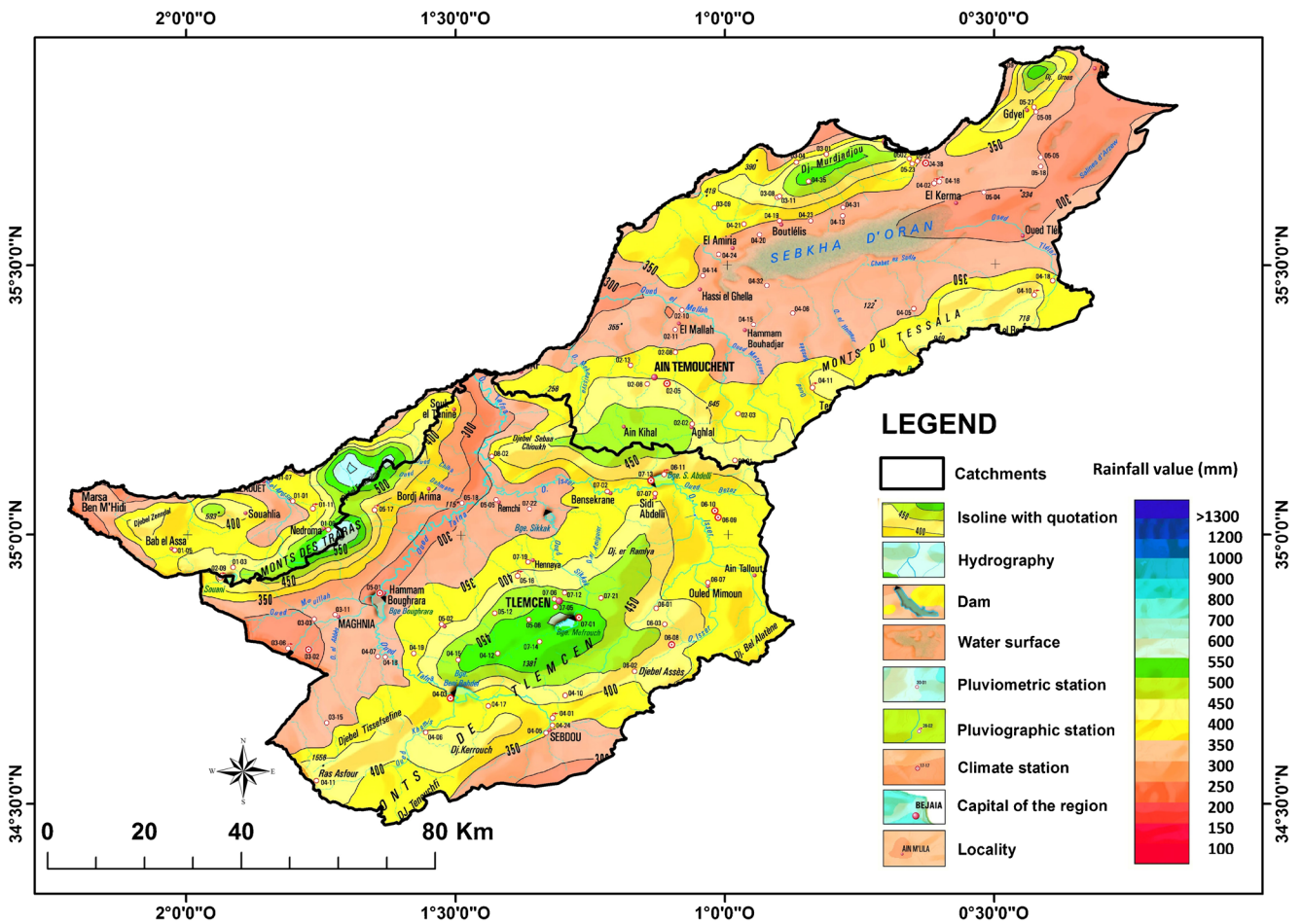


Fig. 3 Rainfall map of the study area (the National Agency of the Water resources of Algiers (Agence Nationale des ressources hydrométrique, ANRH)

spatial evolution of the flow follows the spatial variability of rainfall. The latter decreases from the east to the west, and from the south to the north.

Figure 3 illustrates the spatio-temporal variability of rainfall in the northwestern Algeria.

Data

The collected data concern rainfall, runoff, and temperature. Conceptual models (ABCD and GR2M) require evapotranspiration as an essential input; for this purpose, the Thornthwaite model for estimating the evapotranspiration will be used. The data used were obtained from the two agencies responsible of rainfall network, i.e. the National Agency of Water Resources (N.A.W.R) and the National Meteorological Office (NMO). The characteristic of set data are regrouped in Table 1.

GR2M

The GR2M model (agricultural engineering with 2 monthly parameters) is a conceptual model with two parameters. It is developed by the CEMAGREF (France) for applications in the field of water resources and the base flow. This model has

known several versions, proposed successively by (Kabouya 1990; Makhlouf 1994; Mouelhi 2003; Mouelhi et al. 2006) which made it possible to improve gradually the performances of the model. The version presented here is that of Mouelhi et al. (2006) which appears the best performing. The schematic of the GR2M is presented in Fig. 4.

At the beginning, the production store has a capacity to collect the moisture defined by parameter X1. The initial level S specifies the initial state of water/moisture in the production store of X1 capacity. Rainfall (P) is distributed such that total of the rainfall increases the level in the production store (infiltration) and the rain that does not infiltrate (excess) becomes surface flow (P1). The water level in the production store also varies due to evapotranspiration (ET). The level in the storage tank is also affected by percolation (P2). The remaining water becomes the storage water level S that is used as the starting level for the next time step (month). The routed water consisted of the infiltrated water and excess rainfall. The routed water (R1) is affected by the exchange coefficient X2 that represented groundwater exchange. X1 and X2 are free parameters that are determined through calibration. According to Mouelhi et al. (2006), the routing store has a fixed capacity of 60 mm.

Table 1 Calibration and validation periods of the models

Catchments	Hydrometric stations	Period of calibration	Period of validation
TAFNA	Chouly Wadi (160601)	1971–1988	1995–2006
Oranian coastal	TLATA Wadi (040401)	1979–1989	1990–2010
	Mellah Wadi (040220)	1979–1990	1990–2011

ABCD

The ABCD model is a conceptual model with four parameters, it presents a diverse formulation in the evapotranspiration process, and takes into account an excess of water whenever the soil moisture reservoir (S) is not full yet (Mouelhi et al. 2006). There are four parameters governing the model behavior—A controls the amount of runoff and recharge that occurs when the soils are under-saturated; B controls the saturation level of the soils; C defines the ratio of groundwater recharge to surface runoff; and D controls the rate of groundwater discharge (https://abcd.walkerenvres.com). As displayed in Fig. 5, the ABCD contains two compartments: the first one is the soil storage (S) and the second is the ground water storage (R). The upper (S) has two productions: runoff and infiltration. Therefore, the ABCD has two inputs: rainfall (P) and evapotranspiration (ET). The outputs will be the runoff (Q_s), infiltration, monthly available water, real evapotranspiration (ET), satisfied soil moisture at the last of month (S), total runoff (Q), groundwater runoff (Q_R) and monthly groundwater storage

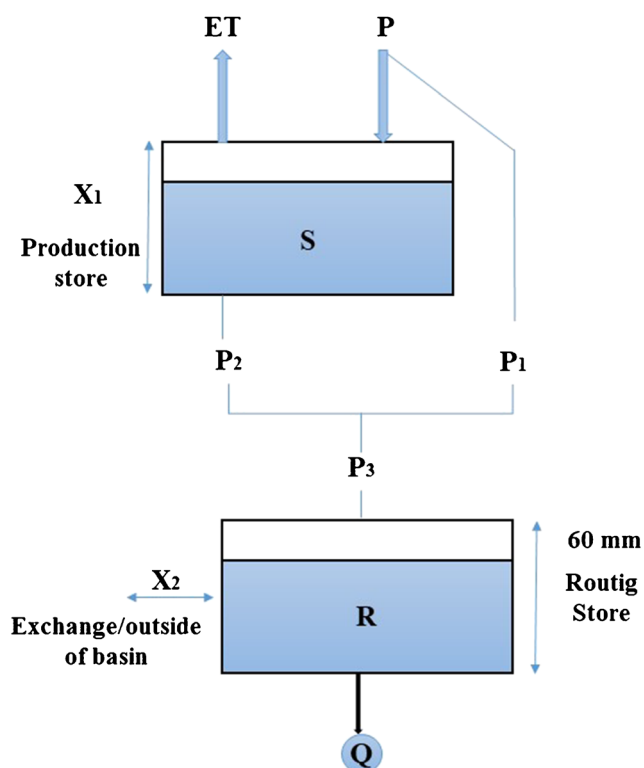


Fig. 4 Schematic of the GR2M model (Mouelhi et al. 2006)

(R) (Fig. 3). The following equations describe the catchment runoff production (storage, soil moisture and ground water).

$$\begin{aligned} \text{Storage}_t + \sum \text{Outflow}_t &= \text{Storage}_{t-1} + \sum \text{Outflow}_t \\ \text{SoilMoisture}_t + ET_t + \text{Runoff}_t + \text{Recharge}_t &= \text{SoilMoisture}_{t-1} \\ \text{Groundwater}_t + \text{Discharge}_t &= \text{Groundwater}_t + \text{Recharge}_t \\ &+ \text{Precip}_t \end{aligned}$$

Box-Jenkins model

According to the previous works of Yule and Wold (Wold 1938; Yule 1926), Box-Jenkins (Series 1976) contracted a convenient approach to constructing ARIMA models, which have the basic effect on the forecasting applications and the time series analysis. As demonstrated by Valipour et al. (2013), such a model can be created using a combination of moving average (MA) and AR processes. To apply the ARIMA model, some processes have to be realized first. It proposes a three-phase practical technique to achieve a suitable model; this technique is based on identification, estimation, and diagnosis (Fig. 6).

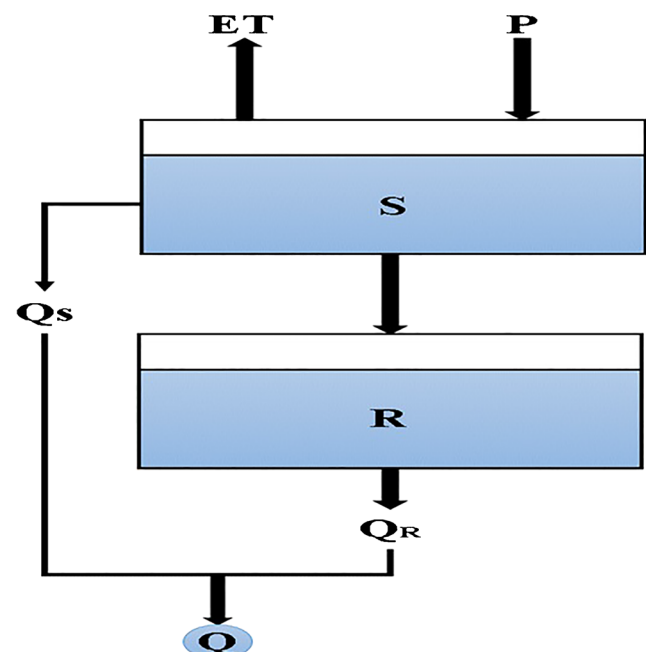


Fig. 5 Schematic of the ABCD model (Thomas 1981)

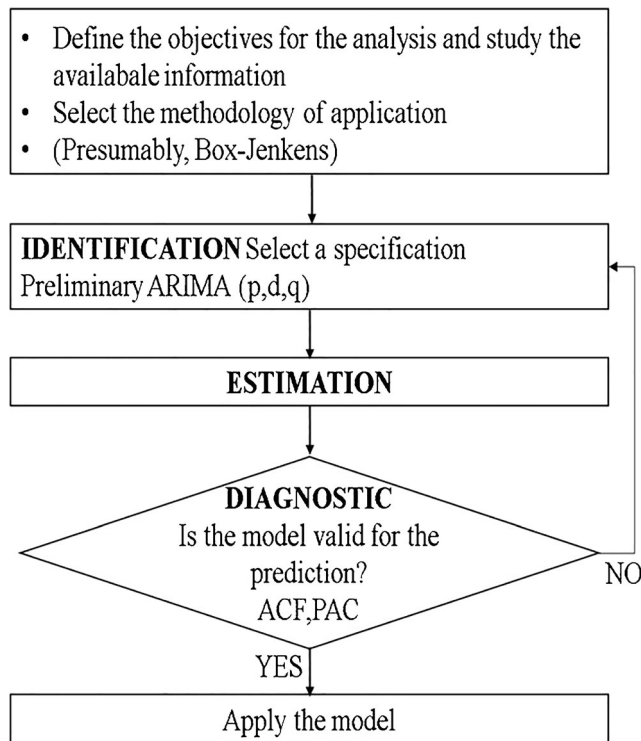


Fig. 6 Schematic of the ARIMA forecasting model (Series 1976)

The first step is to check the stationarity of the data set. “Stationarity” designates that the series remain at a constant level over time. Data should also show a constant variation in their variability over time. This is easily appreciated with a series that is spectacular seasonal and grows at a faster rate. Many accounts associated with the process cannot be determined without gathering the stationarity conditions. The ARIMA model function is represented by (p, d, q); these parameters are successively the order of the autoregressive components, the number of differencing operators, and the highest order of the moving average term. According to Series (1976), the Box-Jenkins methodology proposed an extensive variety of ARIMA models to select from; for the present study, the projected ACF and PACF were used to select one or more applicable ARIMA models. The basic idea is that each ARIMA model should be linked with theoretical ACF and PACF. During the identification step, we obtained the estimated ACF and PACF, deliberate from the data derived from different theoretical ACF and PACF. Consequently, the most resembled model to the theoretical one will be chosen, this model is then adapted to the available data by calculating the

essential coefficients. If these estimated coefficients do not answer some mathematical conditions, the model is rejected (Valipour et al. 2013).

In time series predicting, Y_t , can be reflected as the output of the linear system, where the input ε_t is a Gaussian white noise, Y_t is defined by Eq. (1) (Hamilton 1994);

$$Y_t = \mu + \psi(B)\varepsilon_t \tag{1}$$

μ is the mean of Y_t , the back shift operator B is defined by Eq. 2; $B^k\varepsilon_t = \varepsilon_{t-k}$, for $K = 1, 2 \dots$, and $\psi(B)$ is the transfer function, which links the input, ε_t to the linear system output, Y_t , such as $\psi(B) = 1 + \psi_1B + \psi_2B^2 \dots$

The autoregressive (AR) one is a particularly interesting class of these models. In such a model, every value Y_t is the weighted and finite sum of previous values plus a random term ε_t . The model is then designated by AR (p), p is the corresponding order. Therefore, for $\mu = 0$, it is expressed as follows (Valipour et al. 2013):

$$\phi(B)Y_t = \varepsilon_t \tag{2}$$

Where $\phi(B) = 1 - \phi_1B - \phi_2B^2 - \dots - \phi_pB^p$ is the AR (p) operator, which is a p order polynomial in B , which converges for $|\phi_i| < 1$, the moving average (MA) approach is another interesting model to insure the stationary conditions, in which each value is the sum of $q + 1$ previous values of a white noise, hence Y_t is defined in the following equation (Eq. 3) (Valipour et al. 2013);

$$Y_t = \theta(B)\varepsilon_t \tag{3}$$

Where $\theta(B) = 1 - \theta_1B - \theta_2B^2 - \dots - \theta_qB^q$ is the MA (q) operator, which is a q order polynomial in B , and it converges for $|\theta_i| < 1$ in order to insure for invertibility condition insurance (Valipour et al. 2013).

The linear combination of AR (p) and MA (q) models generate a new mixed process called autoregressive and moving average of order p and q , ARMA (p, q) and it is expressed as follows (Valipour et al. 2013).

$$\phi(B)Y_t = \theta(B)\varepsilon_t \tag{4}$$

In this case, Y_t , can be considered as the output of a linear filter whose transfer function is the ration of the two polynomials $\theta(B)$ and $\phi(B)$:

Table 2 Range of coefficient of efficiency NSE, BIAS (%), and coefficient of determination (Moriassi et al. 2007)

	NSE	BIAS (%)	R ²
Very Good	0.75 < NSE ≤ 1.00	PBIAS < ± 10	0.75 < R ² ≤ 1.00
Good	0.65 < NSE ≤ 0.75	± 10 ≤ PBIAS < ± 15	0.65 < R ² ≤ 0.75
Satisfactory	0.50 < NSE ≤ 0.65	± 15 ≤ PBIAS < ± 25	0.50 < R ² ≤ 0.65
Unsatisfactory	NSE ≤ 0.50	PBIAS ≥ ± 25	R ² ≤ 0.50

Table 3 Calibration and validation performance for GR2M model

Catchment	X1	X2	Calibration				Validation			
			RMSE	NSE	BIAS	R^2	RMSE	NSE	BIAS	R^2
Chouly Wadi	5.6	0.8	1.44	0.85	-16.82	0.81	0.34	0.7	16.18	0.76
Tlata Wadi	4.53	0.55	0.016	0.84	-0.68	0.78	0.46	0.73	24.08	0.74
Mellah Wadi	7.13	0.66	0.37	0.29	-24.4	0.11	1.15	-0.1	-77.63	0.08

$$Y_t = \frac{\theta(B)}{\phi(B)} \varepsilon_t \quad (5)$$

This type of modeling assumes data to be stationary (Mohammadi et al. 2005), but this is not the case in most of hydrological time series (Hamilton 1994). In this case, it is always possible to make them stationary using some mathematical transformations such as the differences and the normal logarithm. Hence, ARMA model becomes ARIMA; more details on the method can be found in several references such as (bin Shaari et al. 2017; Mohammadi et al. 2005; Valipour 2015; Valipour et al. 2013). The general basis of the used methods in this investigation revolves around the identification of the hydrological models that best fit each catchment considered (Hamilton 1994). The assessment of the goodness of fit for the three models was mainly performed using the ratings suggested (McCuen 2016); Pearson's correlation coefficient between observed and predicted runoff in the studied period was also analyzed. The conceptual hydrological models are defined with two parameters for the GR2M and four parameters for the ABCD, they present different moisture balances according to the different processes in a hydrological system through all phases of the hydrological cycle. The ARIMA model with three parameters takes into account only the observed runoff as the exclusive variable as a time series; the goodness of fit of this type of model based on the Akaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC), depends also on the coefficient of determination R^2 .

Goodness of fit

The model parameters were calibrated using different efficiency indices such as the Nash-Sutcliffe Efficiency (NSE) (Nash and Sutcliffe 1970), the coefficient of determination R^2 , the root mean square error (RMSE) (Singh et al. 2005) and the percent bias (%). Moriasi et al. (2007) proposed a rating system to classify the models into four types: very good, good, satisfactory, and unsatisfactory.

For the Box-Jenkins model, to analyze the performance of these models, another two popular measures were applied such as the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). They are used for comparing maximum likelihood models.

$$NSE = 1 - \frac{\sum_{i=1}^n (Q_{oss,i} - Q_{mod,i})^2}{\sum_{i=1}^n (Q_{oss,i} - \bar{Q}_{oss})^2} \quad (6)$$

$$RMSE = \left[\frac{1}{n} \sum_{i=1}^n (Q_{mod,i} - Q_{oss,i})^2 \right]^{1/2} \quad (7)$$

$$BIAS(\%) = \frac{\sum_{i=1}^n (Q_{mod,i} - Q_{oss,i})}{\sum_{i=1}^n Q_{oss,i}} * 100 \quad (8)$$

$$AIC = -2 * \ln(\text{likelihood}) + 2 * k \quad (9)$$

$$BIC = -2 * \ln(\text{likelihood}) + \ln(N) * k \quad (10)$$

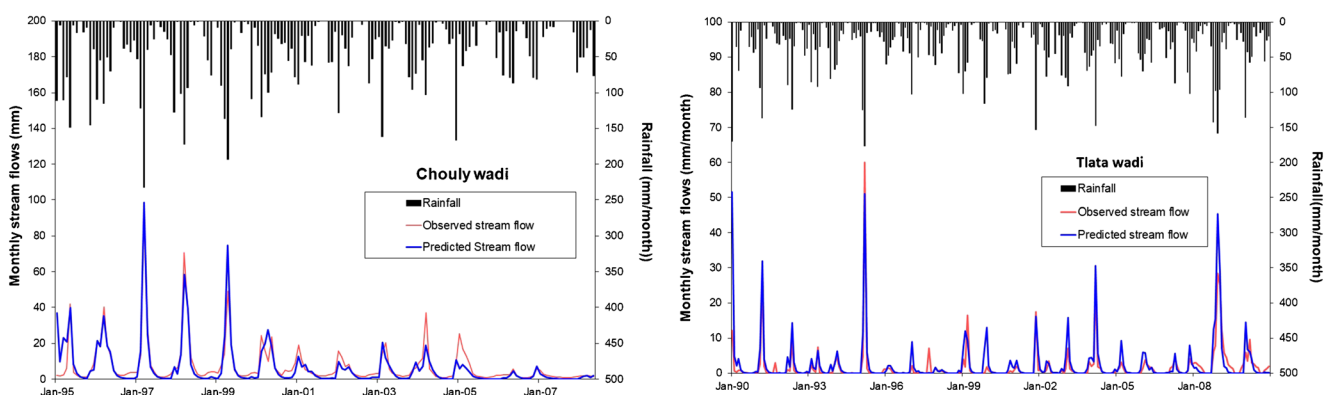


Fig. 7 The comparison of the observed and predicted monthly stream flow using the GR2M procedures for Chouly and Tlata wadi watershed for validation

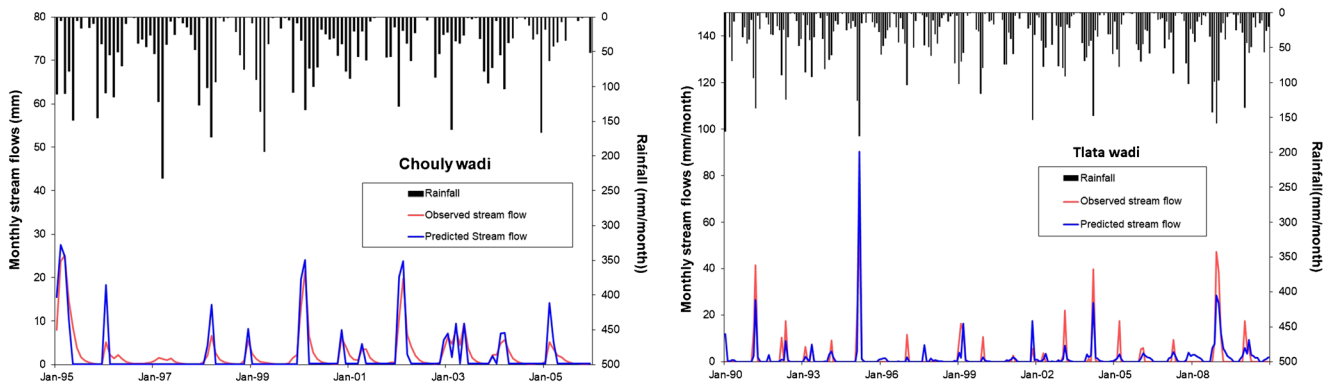


Fig. 8 The comparison of the observed and predicted monthly flow using the ABCD model for Chouly and Tlata basins in validation

Where n is the number of data observations, Q (oss) is the observed stream flow, Q (mod) is the estimated stream flow, and \bar{Q} (oss) is the average of observed stream flow. The modeling procedure having lower RMSE values can be supposed to be the most accurate model for stream flow prediction. A value of NSE equals to 1 means an optimal performance of the model. When the simulated values correspond exactly to the observed values, I , t established the model's ability to simulate the flow values from the mean values. Correlation coefficient values near unity indicate a linear relationship between the observed and predicted stream flows (Table 2).

Results and discussion

Calibration and validation of the models

The GR2M model The calibration of the GR2M model is based on two parameters: X1 and X2, coefficient of production and exchange, respectively. The calibration of the model was done by comparing the observed and predicted stream flow data in the three catchments for different periods (Table 1). All the calibrated parameters of the selected model were identified using iterative procedure for several times based on high value of R^2 and NSE criteria and low value for RMSE. Table 3 presents the calibrated parameter values for the model considering a high value for R^2 and for the NSE criteria. Furthermore, a graphical verification was done to ensure the

reliability of the rainfall–runoff behavior in catchments. The performance of the GR2M model in the validation period was evaluated using a graphical illustration to compare the observed and simulated flows (Fig. 7).

Considering Fig. 7, it can be assumed that there is a relatively good agreement between the observed and simulated flows. A comparison of the observed and simulated monthly flows for the GR2M model is shown in Fig. 8, which confirms the robustness of the calibrated model to predict flow (Table 3).

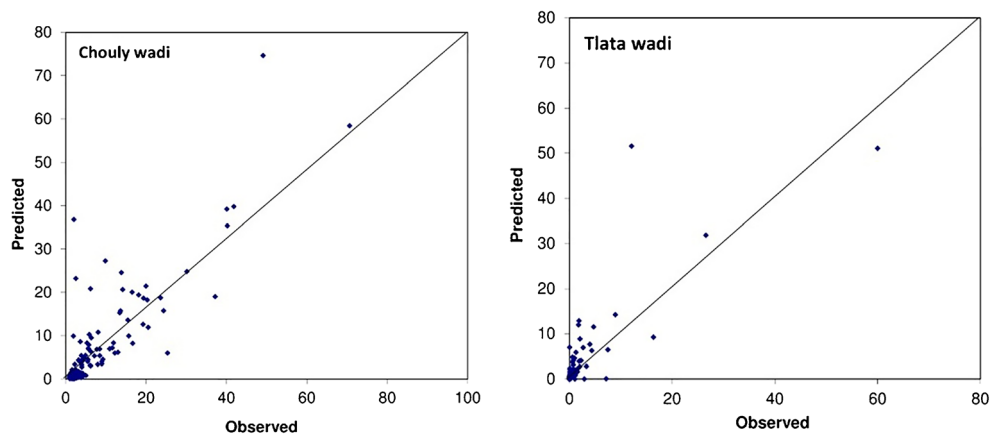
The ABCD model For predicting runoff using the ABCD model, a process was developed for calibrating the model automatically and for predicting runoff. For each catchment, the spatial average of monthly rainfall and PE is computed over the entire studied period, the model was used with these data as input, while their respective four parameters were improved within a defined range (Table 4). The objective of these systematic test runs was to optimize the statistical criterion between the predicted and observed values of flows for the calibration and the validation period. In order to obtain a successful calibration situation by using automatic optimization procedures, it will be necessary to formulate numerical performance measures that reflect the prediction objectives. Calibrated and validated values for the model were presented in Table 4, where results showed that in calibration period the agreement between the observed and predicted monthly flows is widely satisfactory.

The calibration performance was evaluated using plots of the observed and predicted data series and analysis of the

Table 4 Calibration and validation performance for ABCD model

Catchment	Calibration period				Validation period							
	a	b	c	d	RMSE	NSE	BIAS	R^2	RMSE	NSE	BIAS	R^2
Chouly Wadi	0.39	18	0.44	0	1.38	0.9	-19.31	0.91	0.21	0.66	-0.48	0.48
Tlata Wadi	0.35	6	0.38	0	0.64	0.84	-14.37	0.85	0.14	0.65	7.51	0.70
Mellah Wadi	0.52	12.2	0.94	0	0.03	0.31	1.65	0.32	0.3	0.18	-21.5	0.06

Fig. 9 The scatter plots of the observed and predicted monthly flows using the GR2M in the Chouly and Tlata wadi watersheds for the validation data set



statistical measures. Table 4 presents the calibration and the validation performance of the ABCD model based on the most accurate model run taking into account the best fit of the model's parameters. As can be seen in Table 4, the agreement between the observed and simulated monthly flows is generally more satisfactory in the calibration period compared with the validation period. The performance of the ABCD model in the validation period was evaluated using a plot of the observed and predicted flows (Fig. 9). It can be noted that the ABCD model predicted the flow better in the Tlata basin than Chouly basin. As a result, the underestimation of the flow by the ABCD model in Chouly wadi could be seen in Fig. 10. In addition, Fig. 10 shows a comparison of the observed and simulated monthly stream flows with respect to the 1:1 line.

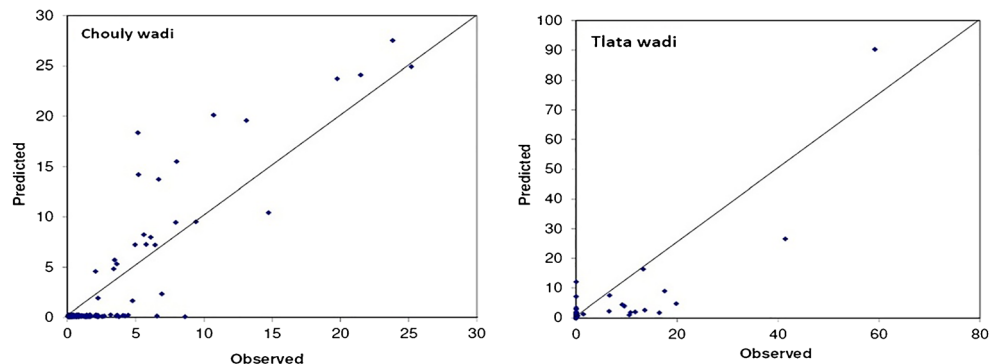
ARIMA models There are five stochastic models to model the monthly flows with the Box-Jenkins model. The most suitable one will be selected based on the Akaike's information, the Bayesian criterion and the root mean square error. The lesser is the RMSE, AIC, and BIC, the better is the model. The monthly time series need some transformations to be stationary as indicated in "Box-Jenkins model." Figure 11 shows the auto-correlation function and the partial correlation function that are applied on data. The first two displays show the ACF and the PACF for initial non-stationary data before application of the first-order transformation ($d=1$). Figure 11, of ACF and PACF after transformation of data, confirms that the data

are compliant to be used in ARIMA model where we have a very significant value in lag one for the ACF and two significant values in the first lags in PACF.

In this way, data are ready for an ARIMA modeling with ($d=1$), several researchers such as Hamilton (1994), Mohammadi et al. (2005), and Valipour et al. (2013) say that an order of differences of more than one is not recommended for time series because data with ($d>1$) results have to be further from reality and the model to be overestimated even with this order of difference we got the perfect correlation between the observed and the predicted observations.

As shown in "Box-Jenkins model," three steps of modeling were followed for selecting the most appropriate model among the different ARIMA models to the time series of stream flows in the three catchments located in the study area. To define the insistence system in the series the ACF and the PACF were tested. Different models were identified based on the ACF and PACF analysis. Table 5 presents the 9 selected ARIMA models which might be suitable for predicting the monthly stream flow in the three catchments under study. For selecting the best-fit model out of the various tested models, the AIC and BIC criteria were used; however, the BIC was intended to select one of the competitive models. The ARIMA (2, 1, 3), ARIMA (1, 1, 1) models performed better than other models. The ARIMA (1, 1, 1) model had the lowest value of BIC in the three cases of the study (Table 5).

Fig. 10 The scatter plots of the observed and predicted monthly flows using the ABCD model in the Chouly and Tlata basins for the validation period



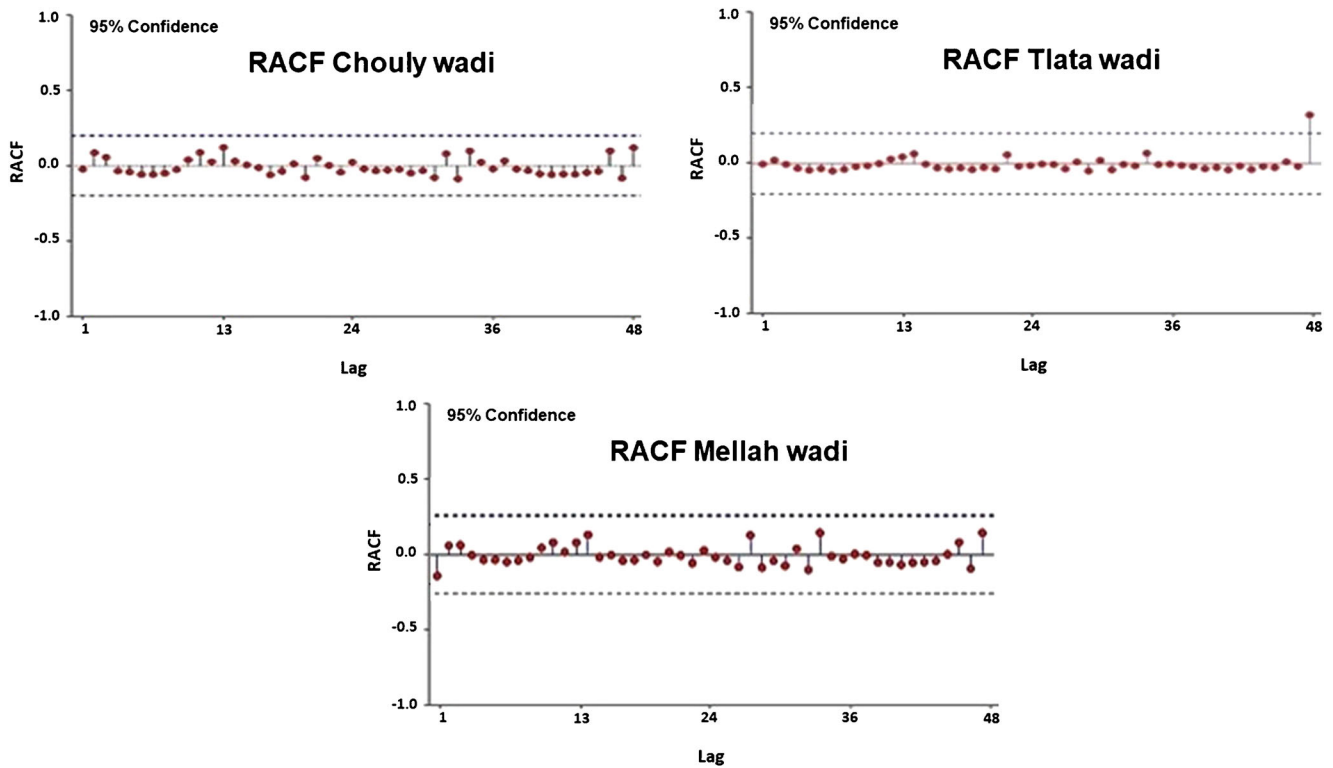


Fig. 11 The RACF for the selected ARIMA (1, 1, 1) model; (a) display of the quality of the calibration (b) display of the quality of the validation

The residual autocorrelation function tests were employed for checking the adequacy of the models fitted to the time series. Depending on the RAFC tests and residual correlation, the ARIMA (1, 1, 1) model was the most suitable one. Figure 11 shows a plot of the RACF for the selected ARIMA (1, 1, 1) model in the three catchments under study. It can be noted that the estimated values fall within the 95% confidence interval. Therefore, the residuals are normally distributed and white noise. The final calibrated model must be tested for the next step to predict stream flow data series in the period of validation. The selected ARIMA (1, 1, 1) model predicted moderately the stream flow in the three catchments in the validation period. Figure 12 shows the comparison between the observed and predicted stream flows in the validation step. However, the

underestimation of the stream flow by the selected model during high flow periods is seen in Fig. 12 especially for the Tlata wadi. Figure 13 shows the accuracy of the ARIMA model in the validation stage as a scatter plot of the observed and predicted stream flows. Figure 12 shows the calibration and the validation with the Box-Jenkins ARIMA model with one parameter of autoregressive and moving average. It was the simplest and the quickest to predict stream flows in a semi-arid catchment. It is so clear also that the Box-Jenkins model is more powerful to predict low flow, but it was so weak in high flow or peak point such as extreme events that are often observed in the area of study. On the other hand, the trend period has no influence during the simulation and the model is more stable after some iterations.

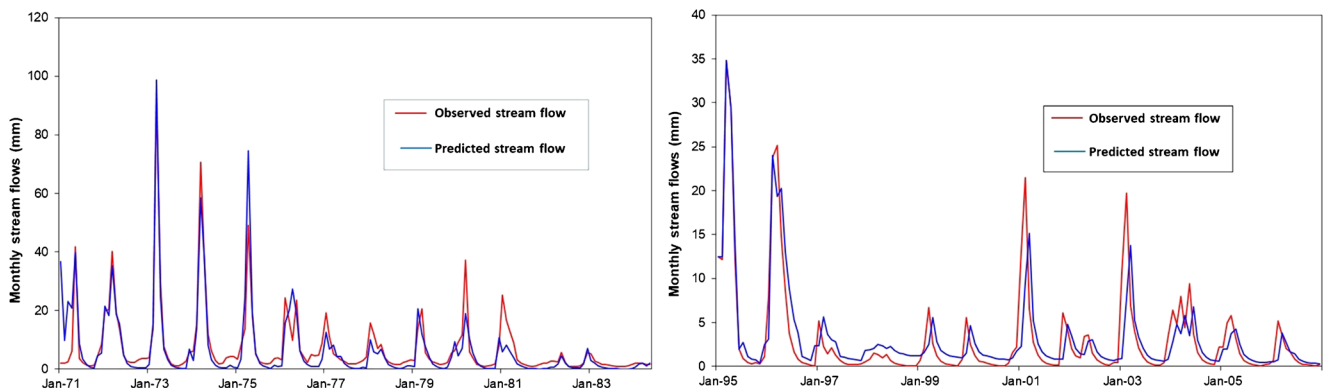


Fig. 12 Results of calibration obtained by the ARIMA (1, 1, 1) model

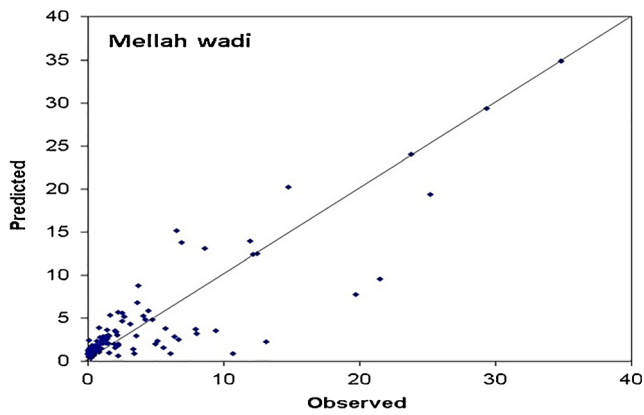


Fig. 13 The scatter plots of the observed versus the predicted monthly stream flows using the ARIMA model in the Mellah wadi watersheds in the validation data set

Hydrological models comparison

Table 6 presents the comparison between the stochastic model (ARIMA) and the conceptual models (GR2M and ABCD) in the validation stage based on different statistical indexes namely, RMSE, NSE, BIAS, and R^2 .

According to these criteria, the GR2M model performed better than the other models. The ABCD model gave competitive results of prediction compared with the ARIMA model in the case of Chouly and Tlata wadi but it is not the case in the third catchment where the ARIMA model prove its ability for modeling stream flow in the Mellah wadi. GR2M and ABCD models show small values of prediction errors with reference to the RMSE criterion. It can be noted from Table 5 that the statistical evaluation of the predicted stream flow provided the less values of RMSE for the ABCD model in the three catchments under study. On the other hand, the GR2M model gave little more value than the ABCD model, 0.34, 0.46 respectively for the Chouly and Tlata Wadi, noting that the Mellah wadi is a special case regarding the complexity of the physical behavior of the catchment. This indicates that the stream flow prediction of the ABCD gave less error than those of the

Table 5 Calibrated model parameters for the Box–Jenkins ARIMA models

Catchment	Model	AIC	BIC	RMSE
Mellah wadi	ARIMA (1,1,1)	571.4	582.9	0.407
	ARIMA (2,1,3)	570.9	588.2	0.336
	ARIMA (0,1,1)	874.3	886.9	0.653
Tlata wadi	ARIMA (1,1,1)	874.1	885.6	0.766
	ARIMA (1,1,3)	875.9	890.3	0.766
	ARIMA (2,1,3)	879.4	899.5	0.732
Chouly wadi	ARIMA (1,1,1)	1617.8	1631.3	0.629
	ARIMA (1,1,3)	1620.2	1640.3	1.341
	ARIMA (2,1,3)	1617.8	1638.9	1.526

Table 6 The comparison of different hydrological models based on accuracy measures

Catchment	Model	RMSE (%)	NSE (%)	BIAS (%)	R^2 (%)
Chouly Wadi	GR2M	0.340	0.70	16.18	0.76
	ABCD	0.210	0.66	-8.48	0.48
	ARIMA	0.629	0.59	21.17	0.68
TLATA Wadi	GR2M	0.460	0.73	24.08	0.74
	ABCD	0.140	0.65	7.51	0.70
	ARIMA	0.766	0.54	15.23	0.62
Mellah Wadi	GR2M	1.150	-0.1	-77.63	0.08
	ABCD	0.300	0.18	-21.45	0.06
	ARIMA	0.407	0.68	-13.29	0.73

GR2M model. Based on the RMSE criteria the GR2M and the ABCD performed better than ARIMA model in the Chouly and Tlata wadi.

The BIAS index was used to evaluate the quality of prediction for the different models used in the paper. The ABCD model shows least values of BIAS of 7.51 and -8.48 for the Tlata and Chouly wadi respectively.

This result shows the small range of overestimation and underestimation in stream flow prediction using the ABCD model. Then, again, the results given by the GR2M model show the large underestimation in stream flow prediction in the Tlata wadi and the high overestimation in both of Chouly and Tlata wadi (Table 5), this value is still in an acceptable standard (Moriassi et al. 2007). However, the accuracy measures for the ARIMA model in Tlata wadi showed a relatively acceptable performance (RMSE = 0.407, NSE = 0.68, BIAS = -13.29, R^2 = 0.73). Nevertheless, it was not the case for the two other models in this catchment. According to the scatter plots of the predicted and observed stream flow for each model in previous sections, it can be noted that the GR2M model has the superiority of predicting in respect to the 1:1 line. Model parameters for the GR2M appear similar to all the analyzed catchments. The ABCD parameters are instead rather variable for the three catchments. This circumstance is probably caused by the adaptation of the different hydrological processes embedded in the model structure to the different cases of the study. What appears quite clear is the inability of the ABCD model to capture and describe the groundwater processes ($d=0$), which is of particular importance for the cases studied.

For evaluating the prediction accuracy for the models, the NSE criterion was used; the NSEs of the ABCD and the ARIMA model were ranging between 0.54 and 0.66 in the Chouly and Tlata wadi catchment. The results reveal that the simulation of these models is not satisfactory but still in an acceptable standard. On the other hand, the GR2M model showed a satisfactory prediction accuracy founded on the NSE criteria with an average value superior than 0.70 which

sets the GR2M model better than the other two models. This implies that the GR2M model is the most accurate modeling procedure based on NSE criteria.

For the complex catchment, which is the Mellah wadi, the ARIMA proves good ability for modeling this type of catchments with $NSE = 0.68$, where the input data were confined only on stream flow as predictor, as indicated in “Study area.” This basin is characterized by a specific geology represented, in certain areas of the basin, by Karsts. Many authors have found that the ARIMA model gives good results in basins with a similar geology (Ghanbarpour et al. 2010; Graupe et al. 1976; Pallard et al. 2009).

The correlation coefficient R^2 was used to compare the different models based on the quality of the agreement between the observed and predicted stream flow. The average value of the correlation coefficients for the GR2M and ABCD models were 0.70 and 0.60, respectively (Table 5), which shows a good agreement between the observed and simulated stream flow values. On the other hand, the ARIMA model shows a satisfactory agreement and correlation in the Chouly and Tlata wadi catchments and a good agreement in the Mellah wadi catchment. This result confirms the robustness of the prediction from the GR2M and the ABCD modeling procedures for stream flow simulation in the study area. However, the GR2M model accurately predicted the stream flow in high flow and low flow. The ABCD model performed very well but not as accurately as the GR2M model. Yet, the ARIMA model predicted the flow rates very well in the Mellah wadi catchment.

Conclusion

In order to identify a valid modeling tool for stream flow predictions in the northernwest of Algeria, conceptual (GR2M and ABCD) and stochastic (ARIMA) modeling approaches have been investigated and compared in this research. The different approaches have been used to simulate stream flow at three catchments featured by different properties, particularly in terms of the hydrogeological aspect. In fact, among the three, the Mellah catchment represents the most complex catchment because of important and diffused Karst phenomena.

As an overall finding, all the applied methods are characterized by similar hydrological performance in terms of goodness of fit indices, although a rank in model performance can be identified. Conceptual models appear indeed to moderately outperform stochastic models. In particular, the GR2M has been identified as the best modeling solution for the case study, probably due to the fact that this specific approach is able to model the water exchange term with neighboring catchments. This is done through the parameter X_2 which, according to the values found during the calibration period,

leads to effectively important contributions of groundwater exchange. The second best model for the case study is represented by the ABCD model which, differently to the GR2M model, was not able to properly characterize the groundwater store. This can be noted in the results of the parameters calibration which has led to a value of zero for the parameter d that accounts for this particular hydrological process. The poor performance of conceptual models in the case of the Mellah catchment can probably be ascribed to its relevant hydrogeological processes which cannot be correctly identified and described by the conceptual parameters.

As a conclusion, this study can be considered as the first step toward a regional scale analysis where the ability of each of the considered models may be validated over a larger number of catchments and where furthermore an investigation of regional model parameters can be achieved.

Acknowledgments The authors wish to thank the National Agency of Water Resources (ANRH) for providing material help and data on which reported analyses are based.

References

- Abudu S, C-I C, King JP, Abudukadeer K (2010) Comparison of performance of statistical models in forecasting monthly streamflow of Kizil River, China. *Water Sci Eng* 3:269–281
- Ahn K-H, Merwade V (2014) Quantifying the relative impact of climate and human activities on streamflow. *J Hydrol* 515:257–266
- Ahsan M, O'Connor KM (1994) A reappraisal of the Kalman filtering technique as applied in river flow forecasting. *J Hydrol Eng* 161: 197–226
- Al-Juboori AM, Guven A (2016) A stepwise model to predict monthly streamflow. *J Hydrol* 543:283–292
- Alley WM (1984) On the treatment of evapotranspiration, soil moisture accounting, and aquifer recharge in monthly water balance models. *Water Resour Res* 20:1137–1149
- Bai P, Liu X, Liang K, Liu C (2015) Comparison of performance of twelve monthly water balance models in different climatic catchments of China. *J Hydrol* 529:1030–1040
- Bakreti A, Braud I, Leblois E, Benali A (2013) Analyse conjointe des régimes pluviométriques et hydrologiques dans le bassin de la Tafna (Algérie Occidentale). *Hydrol Sci J* 58:133–151
- Beaumont P, Blake G, Wagstaff JM (2016) *The Middle East: a geographical study* vol 13. Routledge
- Belarbi H, Touaibia B, Boumechra N, Amiar S, Baghli N (2017) Sécheresse et modification de la relation pluie-débit: cas du bassin versant de l'Oued Sebdu (Algérie Occidentale). *Hydrol Sci J* 62: 124–136
- Benest M (1985) Evolution de la plate forme de l'Ouest algérien et du Nord-Est Marocain au cours du Jurassique supérieur et au Crétacé: Stratigraphie, milieux de dépôts et dynamique sédimentaire Doc Lab Géol, Univ, Lyon1 Fasc I et II
- bin Shaari MA, Samsudin R, bin Shabri Ilman A (2017) Comparison of drought forecasting using ARIMA and empirical wavelet Transform-ARIMA. In: International Conference of Reliable Information and Communication Technology. Springer, pp 449–458
- Boudjadja A, Messahel M, Pauc H (2003) Assessment of water resources in Northern Algeria. *Revue des Sciences de l'Eau/J Water Sci* 16: 285–304

- Castellano-Méndez MA, González-Manteiga W, Febrero-Bande M, Prada-Sánchez JM, Lozano-Calderón R (2004) Modelling of the monthly and daily behaviour of the runoff of the Xallas river using Box–Jenkins and neural networks methods. *J Hydrol* 296:38–58
- Farajzadeh J, Fard AF, Lotfi S (2014) Modeling of monthly rainfall and runoff of Urmia lake basin using “feed-forward neural network” and “time series analysis” model. *Water Resour Ind* 7:38–48
- Ghanbarpour MR, Abbaspour KC, Jalalvand G, Moghaddam GA (2010) Stochastic modeling of surface stream flow at different time scales: Sangsoorakh karst basin, Iran. *J Cave Karst Stud* 72:1–10
- Graupe D, Isailovic D, Yevjevich V (1976) Prediction model for runoff from karstified catchments. In: United States-Yugoslavian Symposium on Karst Hydrology and Water Resources. Dubrovnik (Yugoslavia). 2-7 Jun 1975
- Hamilton JD (1994) Time series analysis vol 2. Princeton university press, Princeton
- Kabouya M (1990) Modélisation pluie-débit aux pas de temps mensuel et annuel en Algérie septentrionale
- Karamouz M, Araghinejad S (2012) Advance hydrology. Amirkabir University of Technology Press
- Liu X, Dai X, Zhong Y, Li J, Wang P (2013) Analysis of changes in the relationship between precipitation and streamflow in the Yiluo River, China. *Theor Appl Climatol* 114:183–191
- Lyon SW, King K, Polpanich O-U, Lacombe G (2017) Assessing hydrologic changes across the Lower Mekong Basin. *J Hydrol Reg Stud* 12:303–314
- Machekposhti KH, Sedghi H, Telvari A, Babazadeh H (2018) Modeling climate variables of rivers basin using time series analysis (case study: Karkheh River Basin in Iran). *Civ Eng J* 4:78–92
- Makhlouf Z (1994) Compléments sur le modèle pluie-débit GR 4 J et essai d’estimation de ses paramètres
- McCuen RH (2016) Modeling hydrologic change: statistical methods. CRC press
- Meddi M, Hubert P (2003) Impact de la modification du régime pluviométrique sur les ressources en eau du Nord-Ouest de l’Algérie IAHS publication: 229-235
- Merabet-Baghi N, Bouanani A, Billaud J (2018) Comment traiter la question de l’eau en Algérie. *Techn Sci Méthodes*:33–46
- Mishra A, Desai V (2005) Drought forecasting using stochastic models. *Stoch Env Res Risk A* 19:326–339
- Mohammadi K, Eslami H, Dayani DS (2005) Comparison of regression, ARIMA and ANN models for reservoir inflow forecasting using snowmelt equivalent (a case study of Karaj)
- Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans ASABE* 50:885–900
- Mouelhi S (2003) Vers une chaîne cohérente de modèles pluie-débit conceptuels globaux aux pas de temps pluriannuel, annuel, mensuel et journalier. ENGREF, Paris
- Mouelhi S, Michel C, Perrin C, Andréassian V (2006) Stepwise development of a two-parameter monthly water balance model. *J Hydrol* 318:200–214
- Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models part I—A discussion of principles. *J Hydrol* 10:282–290
- Otmame A, Baba-Hamed K, Bouanani A, Safa A (2017) Prédétermination des valeurs de crues extrêmes et contribution de la plaine de Sidi Bel-Abbès dans le bilan hydrologique du bassin versant de l’oued. *Mekerra Techniques Sciences Méthodes*: 27-48
- Pallard B, Castellarin A, Montanari A (2009) A look at the links between drainage density and flood statistics. *Hydrol Earth Syst Sci* 13: 1019–1029
- Parajka J, Blöschl G, Merz R (2007) Regional calibration of catchment models: potential for ungauged catchments. *Water Resour Res* 43. <https://doi.org/10.1029/2006WR005271>
- Paturel JE et al (2017) Using land cover changes and demographic data to improve hydrological modeling in the Sahel. *Hydrol Process* 31: 811–824
- Refsgaard JC, Knudsen J (1996) Operational validation and intercomparison of different types of hydrological models. *Water Resour Res* 32: 2189–2202
- Series GBGJT (1976) Analysis: forecasting and control, revised edn. Holden Day, San Francisco
- Sharifi M (2018) Calibration and verification of a-two parameter monthly water balance model and its application potential for evaluation of water resources-a case study of Kalu and Mahaweli rivers of Sri Lanka
- Singh J, Knapp HV, Arnold J, Demissie M (2005) Hydrological modeling of the Iroquois River watershed using HSPF and SWAT JAWRA. *J Am Water Resour Assoc* 41:343–360
- Souza JO, Correa AC, Brierley GJ (2016) An approach to assess the impact of landscape connectivity and effective catchment area upon bedload sediment flux in Saco Creek Watershed, Semiarid Brazil. *Catena* 138:13–29
- Taïbi S, et al. (2017) Etat de la variabilité pluviométrique du nord de l’Algérie entre 1936 et 2016: un retour vers la normale. 7ème conférence Eauclimat-santé
- Thomas HA (1981) Improved Methods for National Water Assessment. Report, Contract: WR15249270. Washington, D.C.: US Water Resource Council
- Traore VB, Sambou S, Tamba S, Diaw A, Cisse M, Fall S (2014) Calibrating the rainfall-runoff model GR4J and GR2M on the Koulountou river basin, a tributary of the Gambia river. *Am J Environ Prot* 3:36–44
- Valipour M (2015) Long-term runoff study using SARIMA and ARIMA models in the United States. *Meteorol Appl* 22:592–598
- Valipour M, Banihabib ME, Behbahani SMR (2013) Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. *J Hydrol* 476:433–441
- Wagener T, Wheeler H, Gupta HV (2004) Rainfall-runoff modelling in gauged and ungauged catchments. World Scientific
- Wang Q, Pagano T, Zhou S, Hapuarachchi H, Zhang L, Robertson D (2011) Monthly versus daily water balance models in simulating monthly runoff. *J Hydrol* 404:166–175
- Wheeler H (2002) Progress in and prospects for fluvial flood modelling. *Philos Trans R Soc Lond A Math Phys Eng Sci* 360:1409–1431
- Wold H (1938) A study in the analysis of stationary time series. Almqvist & Wiksell
- Xu C-Y (2002) Hydrologic models textbooks of Uppsala University. Department of Earth Sciences Hydrology
- Yule GU (1926) Why do we sometimes get nonsense-correlations between Time-Series? A study in sampling and the nature of time-series. *J R Stat Soc* 89:1–63