ISSN 0097-8078, Water Resources, 2017, Vol. 44, No. 4, pp. 568–578. © Pleiades Publishing, Ltd., 2017.

Comparative Study of Different Wavelets for Developing Parsimonious Volterra Model for Rainfall-Runoff Simulation[1]

Mahsa H. Kashani[a], *, Mohammad Ali Ghorbani[a, b], **, Yagob Dinpasho[a], *, Sedaghat Shahmorad[c], ****, and Zbigniew W. Kundzewicz[d], *******

[a] *Department of Water Engineering, University of Tabriz, Tabriz, Iran*

[b] *Engineering Faculty, Near East University, 99138 Nicosia, North Cyprus, Mersin 10, Turkey*

[c] *Faculty of Mathematical Sciences, University of Tabriz, Tabriz, Iran*

[d] *ISRiL PAN, Bukowska 19, 60-809 Poznan, Poland*

**e-mail: m.hkashani@yahoo.com*

***e-mail: m_ali_ghorbani@ymail.com; ghorbani@tabrizu.ac.ir*

****e-mail: dinpashoh@yahoo.com*

*****e-mail: shahmorad@tabrizu.ac.ir*

******e-mail: kundzewicz@yahoo.com*

Received July 17, 2015

Abstract—Although the Volterra models are non-parsimonious ones, they are being used because they can mimic dynamics of complex systems. However, applying and identification of the Volterra models using data may result in overfitting problem and uncertainty. In this investigation we evaluate capability of different wavelet forms for decomposing and compressing the Volterra kernels in order to overcome this problem by reducing the number of the model coefficients to be estimated and generating smooth kernels. A simulation study on a rainfall—runoff process over the Cache River watershed showed that the method performance is successful due to multi-resolution capacity of the wavelet analysis and high capability of the Volterra model. The results also revealed that db2 and sym2 wavelets have the same high potential in improving the linear Volterra model performance. However, QS wavelet was more successful in yielding smooth kernels. Moreover, the probability of overfitting while identifying the nonlinear Volterra model may be less than the linear model.

Keywords: rainfall-runoff process, simulation, Volterra model, wavelet forms

DOI: 10.1134/S009780781704008X

INTRODUCTION

The rainfall-runoff (R-R) relationship is known to be highly non-linear and complex due to large temporal and spatial variability of watershed characteristics and patterns of precipitation as well as the number of input variables of the model. Most hydrologists have adopted the black box models, which often avoid addressing problems of having insufficient data or knowledge on hydrologic system processes [1].

The Volterra model is a general mathematical single input single output (SISO) model (integral operator) for a linear or non-linear black-box system [11]. The Volterra models can fit any complex dynamic systems, without necessity to assume model structure. However, Volterra models are non-parsimonious ones with a large number of unknown parameters. Therefore, large amounts of data are used for the determination of the Volterra model parameters with less error and overfitting problems. The simplest method for the solving and identification of the Volterra model is the ordinary least-squares (OLS) method, which is based on minimizing sum of square errors between measurements and model estimations [16]. If the Volterra model has too many parameters to be estimated, then the OLS method may not perform successfully and overfitting occurs [2]. In order to overcome this problem, one can develop a parsimonious Volterra model and identify the corresponding model from the observed data.

One can construct the parsimonious Volterra model using the compression capabilities of wavelets. By analyzing a sequence using the discrete wavelet transform (DWT), it is localized in frequency and time domains. From the nature of the Volterra coefficients, a large number of wavelet coefficients can be removed, and so there will be fewer coefficients to be estimated.

Some researchers applied the Volterra model in hydrology and R-R simulation. Papazafiriou [17] sim-

[1] The article is published in the original.

ulated R-R process using the time-invariant first-order linear and second-order nonlinear Volterra models. Napiorkowski [12] modeled flow in open channels and surface runoff systems using a model of a cascade of nonlinear reservoirs. The model was equivalent to the first two terms of the second-order Volterra series. Labat et al. [7] applied Volterra first order and second order convolution and a new nonlinear threshold model based on the frequency distribution of runoff at inter-annual mean daily scale over two karstic watersheds, using rainfall and spring outflow data at two semi-hourly and daily scales. Maheswaran and Khosa [9] combined wavelet decomposition with Volterra models to forecast streamflow and compared it with the wavelet-based linear and nonlinear models such as the wavelets-artificial neural networks models.

Wavelets have been used successfully in hydrometeorology and hydrology studies. For example, Labat et al. [7] used the wavelets to explain the non-stationarity in the karstic watersheds. Wavelet analysis of rainfall and runoff values gives important information about the temporal variability of the R-R relationship. Chou and Wang [3] used the DWT to decompose and compress the unit hydrograph. The authors estimated the wavelet coefficients of the UH on-line, using a wavelet-based algorithm. Kamruzzaman et al. [6] compared time series models for streamflow simulation that use lagged rainfall as inputs, with models that use cross product and quadratic terms of the discrete wavelet coefficients of lagged rainfall.

It is important to note that none of the above mentioned studies applies wavelets for constructing the parsimonious Volterra model for R-R simulation. Chou [2] was the first investigator to have proposed a wavelet-based model for simulating nonlinear R-R processes and flood forecasting in a river watershed. He used the DWT to decompose the kernels of the second-order Volterra model and thus modeled the nonlinear R-R processes. The validation results proved the effectiveness of the proposed wavelet-based method. However, the author applied only one mother wavelet, i.e. quadratic spline (QS), and did not evaluate the ability of other, possibly more advantageous, wavelet forms for identifying the Volterra model.

The issue of mother wavelet selection is very important in different applications such as when seeking to apply wavelets for developing the parsimonious Volterra model in hydrologic forecasting studies. The novelty of this study is that it evaluates different wavelet forms for developing a parsimonious Volterra model for R-R simulation. Moreover, it is examined whether using the OLS method for identifying the Volterra kernel may result in data-overfitting. Papazafiriou [17] and Labat et al. [7] have compared performances of linear and non-linear models for the R-R modeling.

This study evaluates the capability of the linear and nonlinear first- and second-order Volterra models for mimicking the essential features of the R-R process. The investigation is implemented on the Cache River watershed and the results are presented herein. This paper is organized as follows. First, the structure of the Volterra model is described and identified using the conventional OLS method. Then the definition of the wavelet transform is presented. It is explained how to apply the DWT for decomposition and compression of the Volterra model. After introducing the case study, discussion of results obtained is offered and conclusions drawn.

MATERIALS AND METHODS

Volterra Model

The Volterra series are an important representation of a nonlinear system. Given the input u and output y variables of the model, the Volterra kernels are identified by solving some linear equations [2]. The discrete form of the Volterra model is written as:

$$\begin{aligned} y_k &= \sum_{j=1}^{m_1} h(j)\,u_{k-j} + \sum_{i=1}^{m_2}\sum_{j=1}^{m_2} g(i,j)\,u_{k-i}u_{k-j} + \cdots + \varepsilon_k \\ &= y_{1k} + y_{2k} + \cdots + \varepsilon_k, \end{aligned} \tag{1}$$

where $h(j)$ and $g(i, j)$ indicate the first-order kernel and second-order kernel of the Volterra model, respectively, which act as the weighting coefficients on the past input to the system. The m_1 and m_2 are the length of data or memory in the linear and nonlinear parts of the Volterra model, respectively. The length of the inputs (u) of the model should be greater than $2m_1$ and $2m_2$ [15]. The y_{1k} and y_{2k} state the Volterra model output obtained at the linear and nonlinear parts of the Volterra model, respectively, and the ε_k represent the random noise (with zero mean) of the model output.

The major drawback of the Volterra model is the exponential growth of the model parameters with the memory and order of the model, which causes problems on identification of the model particularly when the system is strongly nonlinear [5]. Since, an infinite Volterra series is not possible to implement in a practical study, a truncated Volterra model is used. Lattermann [8] corroborated the heuristic finding that including a second term to the linear hydrological model is useful for an accurately approximation, so the present study uses a second-order Volterra model for the R-R process. Moreover, since in some cases the linear models give good results, so in this study the performance of the linear first-order Volterra (FOV) model (analogous to the classical instantaneous unit hydrograph, IUH) is also evaluated for simulation of the R-R process. The first- and second-order Volterra (SOV) models are defined as follows:

$$\text{FOV model: } y_k = \sum_{j=1}^{m_1} h(j) u_{k-j} + \varepsilon_k, \tag{2}$$

SOV model:

$$y_k = \sum_{j=1}^{m_1} h(j) u_{k-j} + \sum_{i=1}^{m_2}\sum_{j=1}^{m_2} g(i,j) u_{k-i} u_{k-j} + \varepsilon_k. \tag{3}$$

Identifying FOV and SOV Models in Time Domain

One of the major tasks in modeling of nonlinear systems using the Volterra model is to determine the Volterra kernels in time domain. In this study, the OLS method is applied to identify the linear and nonlinear Volterra model and its efficiency is evaluated. Assume that discrete convolution structure of the system is as follows [16]:

$$\begin{aligned} y_k &= \sum_{j=1}^{m_1} h(j) u_{k-j} + \sum_{i=1}^{m_2}\sum_{j=1}^{m_2} g(i,j) u_{k-i} u_{k-j} + \varepsilon_k \\ &= \mathbf{u}_{1k}^T \mathbf{h} + \mathbf{u}_{2k}^T \mathbf{G} \mathbf{u}_{2k} + \varepsilon_k, \end{aligned} \tag{4}$$

where $h = [h_1, h_2 \cdots h_{m_1}]^T$, $u_{1k} = [u_{k-1}, u_{k-2} \cdots u_{k-m_1}]^T$, $u_{2k} = [u_{k-1} \cdots u_{k-m_2}]^T$ and the symbol T stands for transposition. The coefficient matrix G of the kernel is assumed to be symmetric with no loss of generality. Therefore, only $M_2 = m_2(m_2+1)/2$ coefficients (of the m_2^2 second-order coefficients) are determined. The linear coefficients number is equal to $M_1 = m_1$. All elements in the lower-triangular part of matrix $\mathbf{G}$ are constructed the vector $\mathrm{vec}_{lt}(G)$. Then y_{2k} can be stated as:

$$y_{2k} = u_{2klt\Omega}^T g_{lt}, \tag{5}$$

where $u_{2klt\Omega} = \Omega_{m_2} \mathrm{vec}_{lt}(U_{2k}), U_{2k} = u_{2k} u_{2k}^T$, $g_{lt} = \mathrm{vec}_{lt}(G)$ and

$$\Omega_{m_2} = \mathrm{diag}\left(\underbrace{1, 2, 2, \cdots 2}_{m_2 \text{elements}} \underbrace{1, 2, 2, \cdots 2}_{m_2-1 \text{ elements}} \cdots \underbrace{1}_{1 \text{elements}} \right).$$

Therefore, the Volterra model as Eq. (4) can be written in the standard linear regression form as:

$$y_k = \phi_k^T \theta + \varepsilon_k, \tag{6}$$

where $\phi_k = \begin{bmatrix} u_{1k} \\ u_{2klt\Omega} \end{bmatrix}$ and $\theta = \begin{bmatrix} h \\ g_{lt} \end{bmatrix}$. Therefore, the OLS method can be used to estimate the unknown parameter θ. The OLS method produces the least-squares error between the observed (y_k) and estimated ($\hat{y}_k$) outputs. The objective function (F) is defined as:

$$F = \sum_{k=1}^{N} (y_k - \hat{y}_k)^2. \tag{7}$$

The minimization of the objective function F leads to an estimator of the kernels:

$$\hat{\theta} = \left[\phi_k \phi_k^T \right]^{-1} [\phi_k y_k]. \tag{8}$$

Note that for the FOV model the OLS method is also used for solving the linear regression form of the model defined as:

$$y_k = u_{1k}^T h + \varepsilon_k. \tag{9}$$

Wavelet Theory

The wavelet function $\psi(t)$ has two properties:

(1) The integral of the function is zero:

$$\int_{-\infty}^{\infty} \psi(t)\, dt = 0; \tag{10}$$

(2) The function has finite energy and is square integrable:

$$\int_{-\infty}^{\infty} |\psi(t)|^2\, dt < \infty. \tag{11}$$

Selection of an appropriate wavelet is very important task and depends on the properties of the wavelet function and the problem at hand [10]. In this study, different wavelet forms namely, Daubechies (dbN), symlets (symN), coiflets (coifN), Haar (haar) and quadratic spline (QS) are used for analysis. Table 1 presents the characteristics of these wavelets. The main reason for choosing these wavelets is the biorthogonality characteristic of these wavelets with their duals.

Identifying the Second-Order Volterra Model Using Wavelet

The 1-D and 2-D DWT can be applied for decomposing and compressing the first- and second-order kernels of the Volterra model, respectively. More information about 1-D and 2-D DWT can be found at [16].

In order to compress a model using wavelets, one must first determine the DWT coefficients of the kernels (η and γ_{lt} of h and g_{lt} kernels, respectively). Then retain only the coefficients with magnitude above a certain threshold value and set the rest to zero. Therefore, one can compress η to η_c and γ_{lt} to γ_{ltc} of smaller dimensions. Then, the second-order Volterra model can be written as [16],

$$y_k = y_{1k} + y_{2k} + \varepsilon_k \approx \varnothing_{kc}^T \theta_c + \varepsilon_k, \tag{12}$$

Table 1. Some properties of different wavelets

Property	haar	dbN	symN	coifN	QS
Compactly supported	•	•	•	•	•
Symmetry	•				•
Near symmetry			•	•	
Asymmetry		•			
Arbitrary number of vanishing moments		•	•	•	
Exact reconstruction	•	•	•	•	•
Discrete transform	•	•	•	•	•
Orthogonal analysis	•	•	•	•	•
Biorthogonal analysis	•	•	•	•	•

where

$$\varnothing_{kc} = \begin{bmatrix} \tilde{v}_{1kc} \\ \tilde{v}^T_{2klt\Omega c} \end{bmatrix} \text{ and } \theta_c = \begin{bmatrix} \eta_c \\ \gamma_{ltc} \end{bmatrix},$$

and $\tilde{v}_{1ck}$ and $\tilde{v}^T_{2klt\Omega c}$ are the compressed DWT of the input vector u_{1k} and $\tilde{v}^T_{2klt\Omega}$ w.r.t. the dual wavelet, respectively.

The Eq. (12) is also a linear equation. However, it has a much shorter vector θ_c (of length equal to $M_{c1} + M_{c2}$) than the vector θ to be identified. The compressed kernels can be identified using the OLS method as [16],

$$\begin{bmatrix} \hat{h}_c \\ \mathrm{vec}\left(\hat{G}_c\right) \end{bmatrix} = \begin{bmatrix} X_1\eta_c \\ X_2\gamma_{ltc} \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix}\hat{\theta}_c, \quad (13)$$

where the matrix X_1 and X_2 are the 1-D and 2-D IDWT, respectively.

True selection of coefficients in η and γ_{lt} to set them to zero is required for accurate identification of the Volterra model [16]. In this study, the way of thresholding was taken from [2] and revised as shown below:

Step 1. Choose resolution level l for the linear and nonlinear parts of the second-order Volterra model based on the length of kernels. Obtain the DWT of the Volterra model. Determine a set of indices for detail coefficients to retain based on certain thresholds.

Step 2. Estimate the DWT retained coefficients using observed R-R data.

Step 3. Perform tests on the prediction residuals and model performance using some criteria such as RMSE.

Step 4. If the criteria values were satisfactory and the kernels shape were in correct and smooth forms then go to step 6, otherwise continue.

Step 5. Adjust the thresholds. Add or drop the detail coefficients according to the adjusted thresholds, and repeat steps 2–4.

Step 6. Stop.

Case Study and Data Used

This study uses R-R data of the Cache River catchment (Fig. 1) located at Forman in the south of Il1inois just north of the confluence of the Ohio and Mississippi rivers. The catchment of the 630 km^2 area has well-developed drainage network and mild slopes. The available data pertain to eight R-R events observed in the period 1935–1951 that had been used in the past by such researchers as Diskin and Boneh [4], Napiórkowski [12], Napiórkowski and Kundzewicz [13] and Napiórkowski and Piotrowski [14]. The data of the first six events are used for calibrating the models and the last two for testing them.

Comparison of Models

In order to quantitatively examine the performance of Volterra models, the one-day ahead predictions were evaluated based on the following five performance criteria:

(1) Coefficient of efficiency (CE):

$$\mathrm{CE} = 1 - \frac{\sum_{i=1}^{n}\left(Q_i - \hat{Q}_i\right)^2}{\sum_{i=1}^{n}\left(Q_i - \bar{Q}\right)^2}, \quad (14)$$

where $\hat{Q}_i$ (mm/day) and Q_i (mm/day) represent the discharge of the simulated and observed hydrographs at the time i, $\bar{Q}$ (mm/day) denotes the average observed discharge at the time i and n is the data points number. The closer the CE is to one, the better the fit.

(2) Root Mean Square Error (RMSE):

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(Q_i - \hat{Q}_i\right)^2}{n}}, \quad (15)$$

(3) The relative error of total volume (EV):

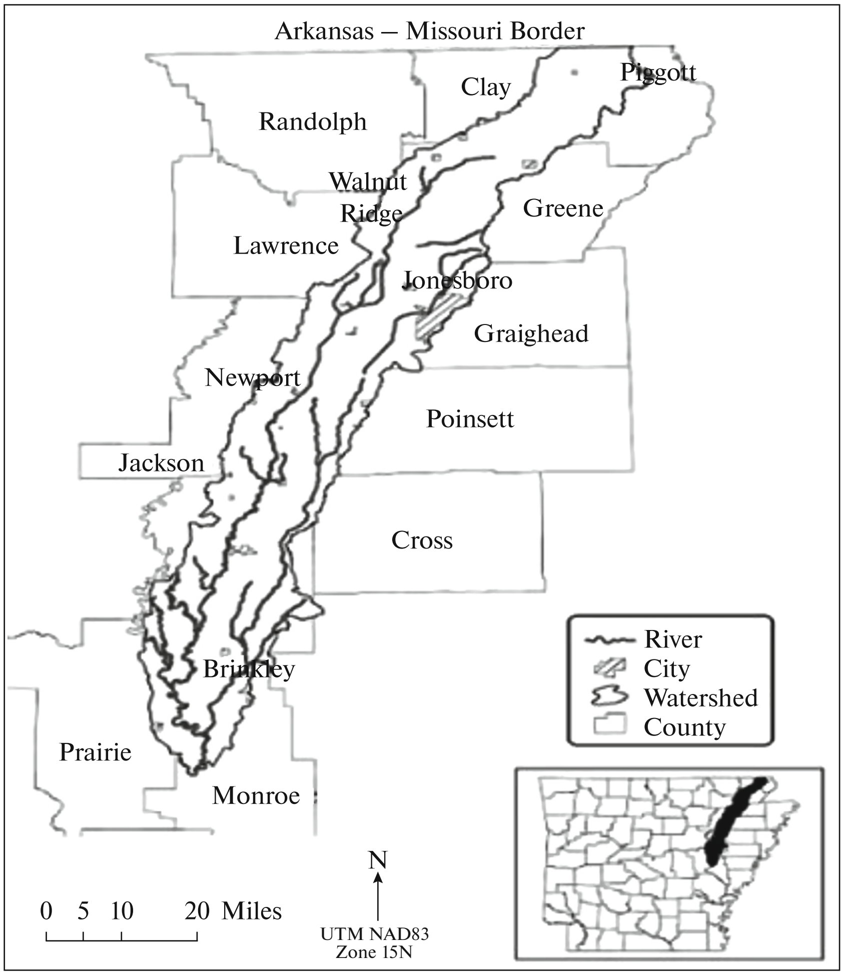

Fig. 1. Cache river catchment.

$$\mathrm{EV}(\%) = \frac{\sum_{i=1}^{n}(\hat{Q}_i - Q_i)}{\sum_{i=1}^{n} Q_i} \times 100\%, \tag{16}$$

(4) The relative error of peak discharge ($\mathrm{EQ_p}$, %):

$$\mathrm{EQ_p}(\%) = \frac{\hat{Q}_p - Q_p}{Q_p} \times 100\%, \tag{17}$$

where $\hat{Q}_p$ (mm/day) and Q_p (mm/day) represent the peak discharge of the simulated and observed hydrographs.

(5) The error of the time-to-peak ($\mathrm{ET_p}$):

$$\mathrm{ET_p} = \hat{T}_p - T_p, \tag{18}$$

where $\hat{T}_p$ (day) and T_p (day) denote the time-to-peak for the simulated and observed hydrograph, respectively.

RESULTS AND DISCUSSION

Identifying FOV and SOV Models in Time Domain

In this study, six storm events over the Cache River catchment were connected in series and considered as a single event for the calibration of the model. The time lag used for the inputs of the model is 1 day, which is the same as the duration of the effective rainfall pulses. The number of parameters of the FOV model was chosen equal to 14 using the OLS and a trial and error method in order to minimize the objective function (F) value. Since the data length (89 days) should be much greater than the memory length ($2m_1$) for noise smoothing, so the number 14 is almost acceptable. Figure 2a displays the first-order kernel of the FOV model calibrated using the six events. It demonstrates that the kernel is smooth; however, there are oscillations at the tail end. The kernel is characterized by a sharp peak, a rapid decrease and an upward trend at the end. The first-order kernels obtained by Labat et al. [7] using different identification methods for the FOV model were also characterized by a sharp peak, a rapid decrease and some oscillations near zero.

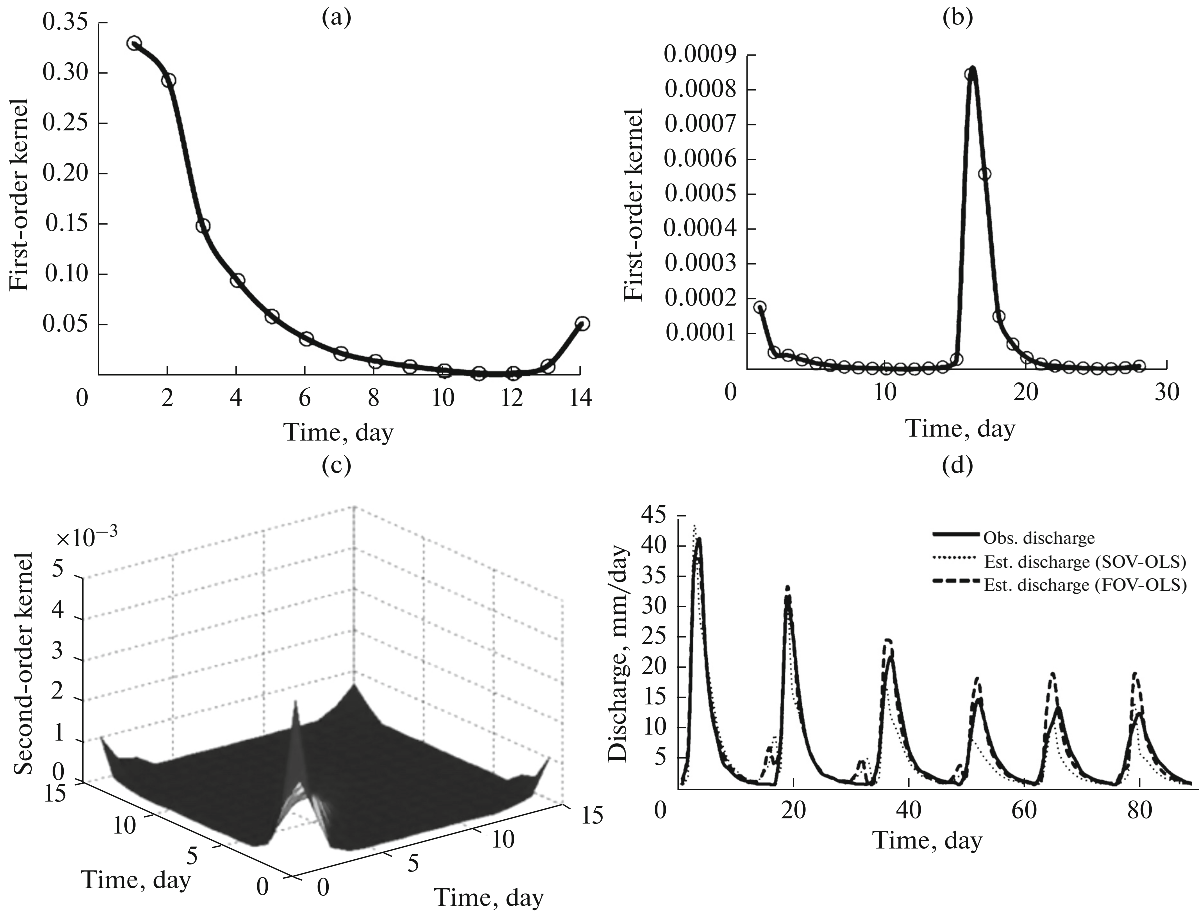

Fig. 2. The first-order kernel of the FOV model (a), two kernels of the SOV model, first-order kernel (b) and Second-order kernel (c), and calibration results obtained from the models using the OLS without wavelet compression (d).

However, Papazafiriou [17] obtained a completely smooth first-order kernel with 96 parameters for the FOV model.

The numbers of parameters at the both linear and nonlinear parts of the SOV model were chosen equal to 28 and 14 using the OLS and a trial and error method. This time, the input data length (89) is not much greater than 56 (2 × 28) and 28 (2 × 14). Figures 2b–2c indicate kernels of the SOV model. It can be seen from Fig. 3b that the kernel has oscillations at the beginning. Such a behavior is seen at the beginning of the first-order kernel with 32 parameters obtained by Chou [2] for a SOV model. The second-order kernel of the SOV model in Fig. 2c is fairly smooth; however it exhibits four peaks. The kernel shape is somewhat similar to that obtained by Labat et al. [7]. In general, the OLS method did not yield completely smooth kernels. This result is in agreement with those obtained by Chou [2] who found that the conventional methods are not accurate for identifying the Volterra model because of many unknown parameters of the model.

Figure 2d shows the calibration results obtained from the FOV and SOV models for the six storm events artificially placed next to each other. The figure indicates that the FOV model overestimates the peak flows for the last five storms. However, it predicts the time to peak quite precisely. Moreover, it preserves the shape of the rising and recession limbs of the hydrographs with less error. The SOV model does not overestimate the peak flows. It also predicts the time to peak and the rising limbs of the hydrographs more or less accurately. However, it predicts the recession limbs with large error. Table 2 summarizes the performance criteria values associated with the Volterra models solved using the OLS method in time domain for the calibration and validation stages. The Table 2 reveals that the FOV model outperforms the SOV based on the all performance criteria values except the EQ_p at the calibration stage. However, difference between the models performances is not large, except for criteria EV and EQ_p.

Figures 4a–4b show the observed and simulated runoff by the FOV and SOV models at the validation stage. The FOV model overestimates the peak flows.

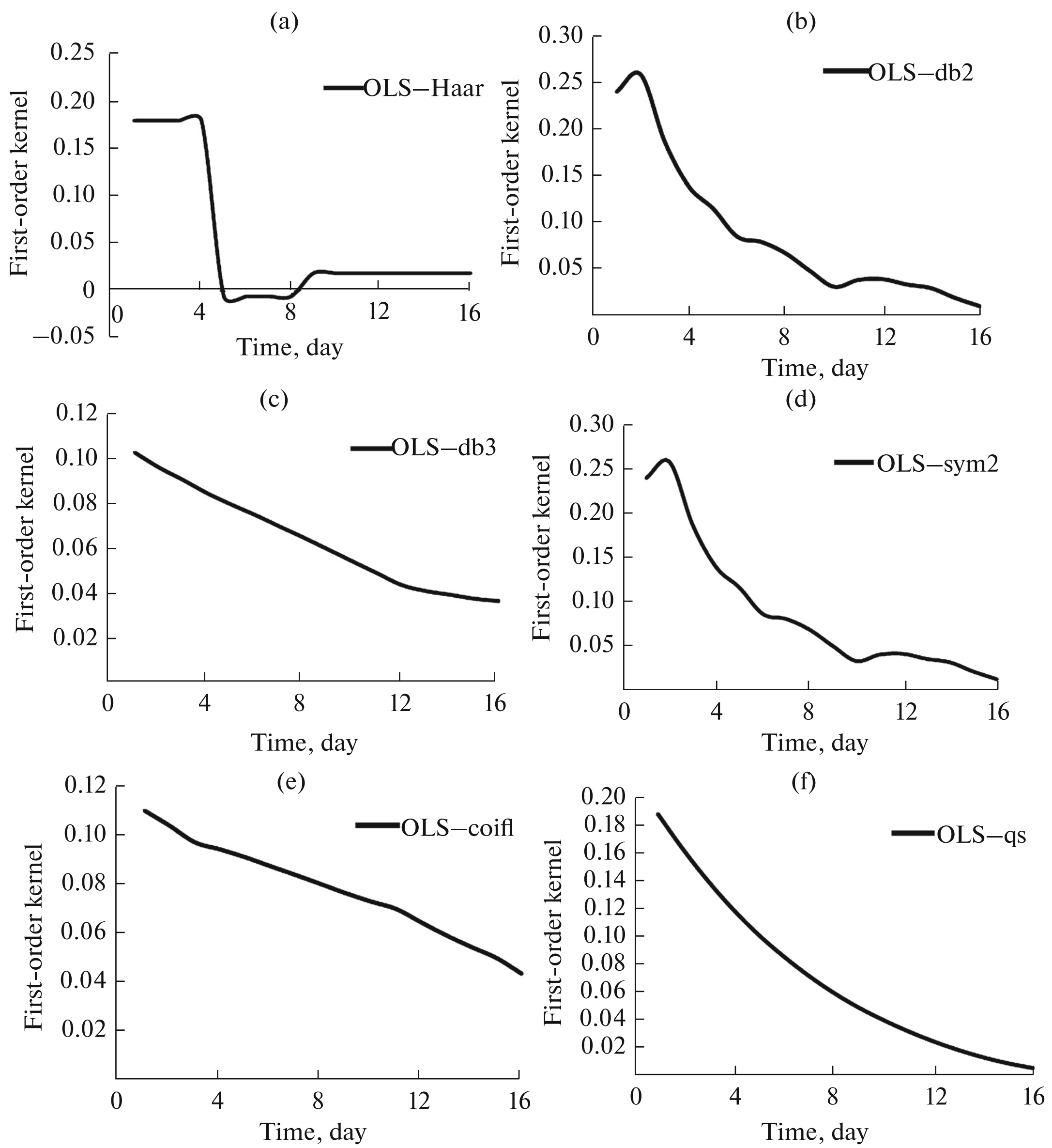

Fig. 3. The first-order kernel of the FOV model calibrated using the OLS without (solid line) and with (broken line) compression of different wavelets, (a) Haar; (b) db2; (c) db3; (d) sym2; (e) coif1; and (f) QS wavelet.

The SOV model predicts the peak discharge more precisely for event 7; but underestimates the peak value in the case of event 8. Both models, estimate the times to peak more or less accurately for both events. Note that the FOV model presents overfitting problem and its error in estimating peak discharges has increased significantly at the validation stage. In other words, the SOV model with more parameters than the FOV (105 and 14 parameters, respectively) does not show over fitting, while both the models use the same number of data samples for calibrating. This is due to the fact that the model structure affects quality of learning and determining the relationship between the inputs and outputs of the model. In other words, the nonlinear SOV model is more capable than the linear FOV in learning the relationship. Generally, it can be concluded that using the OLS method for identifying the Volterra model does not always result in overfitting; it depends on the model structure, number of the calibration data samples and number of the Volterra orders.

According to Table 2, the average values of the performance measures (except the EQ_p) at the validation stage indicate that the FOV model is more capable than the SOV in R-R simulation. However, the difference between the models results is not so large based on the all criteria (except the EV criterion which is high for the SOV). However, one cannot conclude surely about a model performance and potential only

Table 2. The calibration and validation results associated with the FOV and SOV models without wavelet compression (the bold numbers indicate the best model; the average row presents the averages of the absolute values)

Events	CE		RMSE, mm/day		EV, %		EQ_p, %		ET_p, day	
	SOV	FOV	SOV	FOV	SOV	FOV	SOV	FOV	SOV	FOV
Calibration										
In series	0.81	**0.92**	3.45	**2.15**	−21.35	**8.15**	**3.95**	−8.39	−1	**0**
Validation										
7	0.40	0.46	4.63	4.41	−27.71	3.51	7.45	26.59	−1	0
8	0.28	0.49	3.26	2.76	−52.95	3.81	−25.68	28.15	−1	−1
Average	0.34	**0.47**	3.94	**3.58**	40.33	**3.66**	**16.56**	27.37	1	**0.5**

by evaluating performance criteria. Herein, Fig. 4 shows that the FOV model overestimates the peak discharges and this causes some criteria such as the RMSE and EV to show lower values and some others such as the CE to show fairly high values. In fact, the FOV model performance is fairly poor at the validation stage and as mentioned earlier, it faces overfitting problem. Since an overfitted model has high uncertainty, so it cannot be a superior model than the SOV. Therefore, the SOV model is more suitable model than the FOV for R-R process simulation of the catchment. Moreover, since in this study the second-order kernel values of the SOV model is much greater than the first-order kernel, hence the nonlinear part of the SOV model has an important role in simulation; and it can be concluded that the nonlinear SOV model is superior to the linear FOV.

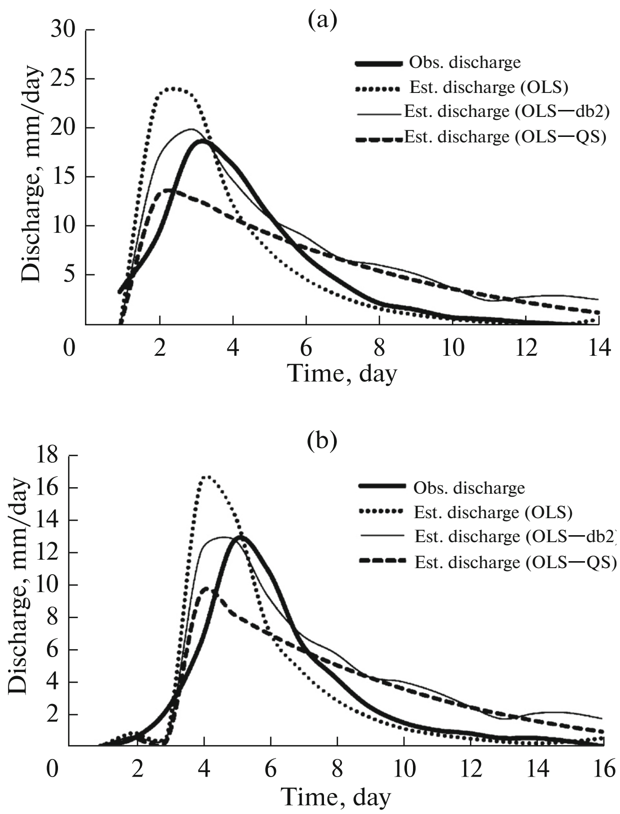

Fig. 4. The validation results obtained from the FOV models using OLS with and without wavelets compression for event 7 (a) and 8 (b).

Identifying FOV and SOV Models Using Wavelet Compression

Since the FOV model showed overfitting problem in its results, so in order to overcome this problem there is a need to use wavelet compression and reduce the number of model parameters. In this investigation, different wavelet functions were applied to decompose and compress the kernel of the FOV model. In this case, we assumed that the memory length of the model kernel is an integral power of 2 according to the characteristics of the wavelet analysis. Figure 3a-f show the first-order kernel of the FOV obtained using the OLS with wavelet compression for the Haar, db2, db3, sym2, coif1 and QS wavelets, respectively. As it can be seen from the figure and as expected, compression using the Haar wavelet did not yield smooth kernel. Moreover, the kernel shape is not acceptable and differs from the original kernel shape. Therefore, it is not recommend using the Haar wavelet for compression of the Volterra kernels. The Haar wavelet is a simple wavelet. It has poor filtering characteristics [2]. The db2 wavelet is more or less successful and yields fairly smooth kernel. However, the suitability of this wavelet will be evaluated certainly by calculating the validation hydrographs. The sym2 wavelet shows a very similar performance to the db2. The db3 and coif1 do not yield acceptable kernel shapes. According to Fig. 3f, it can be seen that the QS wavelet yields a completely

Table 3. The calibration and validation results associated with the FOV using OLS with and without db2-wavelet compression (the bold numbers indicate the best model; the average row presents the averages of the absolute values)

Calibration

Events	CE		RMSE, mm/day		EV, %		EQ_p, %		ET_p, day	
	OLS–db2	OLS	OLS–db2	OLS	OLS–db2	OLS	OLS–db2	OLS	OLS–db2	OLS
In series	0.79	**0.92**	3.61	**2.15**	**0.42**	8.15	−19.35	**−8.39**	**0**	**0**

Validation

Events	CE		RMSE, mm/day		EV, %		EQ_p, %		ET_p, day	
	OLS–db2	OLS	OLS–db2	OLS	OLS–db2	OLS	OLS–db2	OLS	OLS–db2	OLS
7	0.70	0.46	3.28	4.41	0.37	3.51	8.88	26.59	0	0
8	0.72	0.49	2.04	2.76	0.37	3.81	0.49	28.15	0	−1
Average	**0.71**	0.47	**2.66**	3.58	**0.37**	3.66	**4.68**	27.37	**0**	0.5

smooth kernel. The kernel is characterized by a sharp peak, and then a rapid decrease. Therefore, the efficiency of the db2 (or sym2) wavelet will be compared with the QS at the validation stage.

Figures 4a, 4b shows the validation hydrographs obtained using OLS with and without the db2 and QS wavelets. The figure shows that the peak flows and times to peak of the hydrographs obtained by the db2-wavelet compression are very close to the actual values. The accuracy of the rising limbs of the hydrographs has improved; but the recession limbs still show some oscillations. The performance of the model is also improved by the QS-wavelet compression. The rising limbs of the hydrographs are closer to the actual limbs. However, the peak flows, times to peak and the recession limbs indicate errors. Generally, it seems that db2-wavelet is more successful than the QS in improving the FOV performance. However, the QS may be more capable when the model shows high overestimations.

Tables 3, 4 quantitatively compare the three methods using the five selected performance criteria. The model performance with wavelet compression is almost similar to the model without wavelet compression and it is logical to find errors in model results because of eliminating some parameters of the model. However, the validation results show that the CE for the OLS–db2 and OLS–QS (average value of 0.71) is more than that for the OLS (average value of 0.47), and the average value of RMSE for the OLS–db2 and OLS–QS (2.66 and 2.64, respectively) is less than that of the OLS (3.58). According to the average value of EV criteria, the OLS–db2 and OLS–QS (0.37 and 0.09%, respectively) perform significantly better than the OLS (3.66%). The average value of the EQ_p criteria for the OLS–db2 and OLS–QS is 4.68 and 27.35%, respectively which bring improvement to the peak discharge estimated by the OLS (average value of 27.37%). The prediction of peak discharge is very important task for flood protection projects and warning systems. The OLS–db2 method outperforms the OLS according to the ET_p criteria. However, the OLS–QS method shows a bit more error than the OLS. Generally, it can be concluded that wavelet compression improves the FOV model performance by reducing the number of unknown parameters to be estimated, and preventing overfitting.

While using a finite set of observation data, overfitting occurs applying the OLS method for identifying the Volterra model, since several unknown parameters must be estimated. By using wavelet compression, one can generate smoother kernels to get better validation results. This result is in agreement with Chou [2]. The performance of the OLS–db2 method is almost the same with the OLS–QS. However, it shows high ability in precisely estimating the peak flows. Moreover, the OLS–db2 reduces the 16 estimated parameters of the FOV model to 4, while the OLS–QS method reduces the number to 7. In other words, the OLΣ–db2 provides better compression than compression with the OLS–QS method.

Since the SOV model did not show overfitting, hence it seems that that there is no need to use wavelet compression and reduce the model parameters. However, herein, we investigate the effects of using wavelet compression on the kernels obtained. We assumed that the memory length for the nonlinear and linear parts of the SOV model is 16 and 32, respectively. The first-order kernels obtained using the db2 and QS compression are shown in Figs. 5a, 5b. The figure indicates that both of the linear kernels have very large dimensions and negative ordinates which cause large and negative discharges of generated hydrographs. However, the shape of the kernel yielded by the QS-compression is very close to the shape of a hydrograph; while the db2-compression is not more able to derive the kernel in a correct form. It seems that the QS-wavelet is suitable for kernel compression of the

Table 4. The calibration and validation results associated with the FOV using OLS with and without QS-wavelet compression (the bold numbers indicate the best model; the average row presents the averages of the absolute values)

Calibration

Events	CE		RMSE, mm/day		EV, %		EQ_p, %		ET_p, day	
	OLS–QS	OLS	OLS–QS	OLS	OLS–QS	OLS	OLS–QS	OLS	OLS–QS	OLS
In series	0.72	**0.92**	4.20	**2.15**	**0.11**	8.15	−46.33	**−8.39**	**0**	**0**

Validation

Events	CE		RMSE, mm/day		EV, %		EQ_p, %		ET_p, day	
	OLS–QS	OLS	OLS–QS	OLS	OLS–QS	OLS	OLS–QS	OLS	OLS–QS	OLS
7	0.71	0.46	3.21	4.41	0.09	3.51	−28.75	26.59	−1	0
8	0.71	0.49	2.07	2.76	0.09	3.81	−25.96	28.15	−1	−1
Average	**0.71**	0.47	**2.64**	3.58	**0.09**	3.66	**27.35**	27.37	1	**0.5**

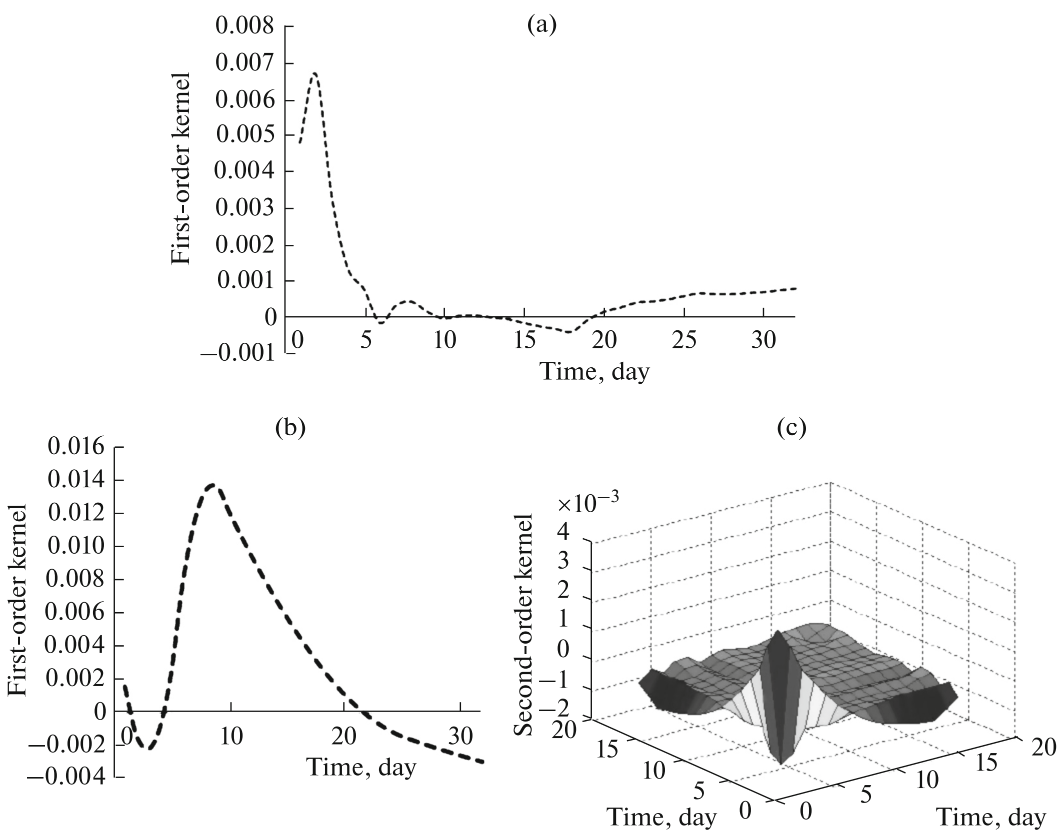

Fig. 5. The first-order kernel of the SOV model calibrated using the OLS with db2 wavelet compression, (a), and two kernels of the SOV model calibrated using the OLS with QS wavelet compression, (b): first-order kernel and (c): second order kernel.

SOV model. The derived second-order kernel using the QS is shown in Fig. 5c. It can be seen from the figure that the kernel is not smoother than the previous uncompressed one. In general, wavelet compression did not yield correct kernels and therefore, cannot improve the SOV model performance. This means that the estimated parameters of the SOV model have not insignificant values. It can be concluded that wavelet compression may not be efficient when the Volterra model does not face overfitting problem.

CONCLUSIONS

This investigation adopted the linear and nonlinear Volterra model of nonlinear R-R system and evaluated different wavelet functions to develop parsimonious first- and second-order Volterra models (FOV and SOV) efficiently. The linear FOV model showed overfitting problem; and hence, its performance was not superior to the nonlinear SOV model. Identifying the Volterra model using OLS method may or may not result in overfiiting; it may depend on number of the model parameters, number of calibration data samples and number of the Volterra model orders. The selection of mother wavelet has an important role in decomposing and compressing the Volterra kernels. Since the Haar, coif1 and db3 wavelets did not yield smooth and correct kernels, they are not suitable wavelets for compressing the Volterra kernels. The db2 and sym2 wavelets were successful for developing parsimonious FOV model and improving the model performance significantly. Moreover, these wavelets provided better compression than the QS-wavelet. However, the QS-wavelet was more capable to yield smoother kernels. This wavelet may be more efficient when the model indicates high over estimations. Wavelet compression eliminates overfitting and increase the efficiency of the model. The method with fewer parameters, compressed using wavelet, showed more precise results at the validation stage. Therefore, it was revealed that wavelet compression can capture the signals characteristics, to model the nonlinear R-R process accurately.

This work constitutes a useful reference which recommends applying the db2 (or sym2) and QS wavelets for compressing kernel of the FOV model and using the QS for developing parsimonious SOV model.

REFERENCES

1. Chang, F.J., Chang, L.C., and Wang, Y.S., Enforeced selforganizing map neural networks for river flood forecasting, *Hydrol. Process.*, 2007, vol. 21, pp. 741–749.
2. Chou, C.M., Efficient nonlinear modeling of rainfall-runoff process using wavelet compression, *J. Hydrol.*, 2007, vol. 332, pp. 442–455.
3. Chou, C.M. and Wang, R.Y., On-line estimation of unit hydrographs using the wavelet-based LMS algorithm, *Hydrol. Sci. J.*, 2002, vol. 47, no. 5, pp. 721–738.
4. Diskin, M.H. and Boneh, A., Determination of optimal kernels for second order stationary surface runoff systems, *Water Resour. Res.*, 1973, vol. 9, no. 2, pp. 311–325.
5. Goulart, J.H.M. and Burt, P.M.S., Efficient kernel computation for Volterra filter structure evaluation, *Signal Process. Lett.*, 2012, vol. 19, no. 3, pp. 135–138.
6. Kamruzzaman, M., Metcalfe, A., and Beecham, S., Wavelet based rainfall-stream flow models for the South-East Murray Darling Basin, *J. Hydrol. Eng.*, 2014, vol. 19, no. 7, pp. 1283–1293.
7. Labat, D., Ababou, R., and Mangin, A., Linear and nonlinear input/output models for karstic springflow and flood prediction at different time scales, *Stochastic Environ. Res. Risk Assess.*, 1999, vol. 13, pp. 337–364.
8. Lattermann, A., *System-Theoretical Modeling in Surface Water Hydrology*, Germany: Springer, 1991, p. 200.
9. Maheswaran, R. and Khosa, R., Wavelet–Volterra coupled model for monthly stream flow forecasting, *J. Hydrol.,* 2012a, vols. 450–451, pp. 320–335.
10. Maheswaran, R. and Khosa, R., Comparative study of different wavelets for hydrologic forecasting, *Comput. Geosci.*, 2012b, vol. 46, pp. 284–295.
11. Muftuoglu, R.F., Monthly runoff generation by nonlinear models, *J. Hydrol.*, 1991, vol. 125, pp. 277–291.
12. Napiorkowski, J.J., Application of Volterra series to modeling of rainfall-runoff systems and flow in open channels, *Hydrol. Sci. J.*, 1986, vol. 31, nos. 2, 6, pp. 187–203.
13. Napiórkowski, J.J. and Kundzewicz, Z.W., A discrete conceptualization of a Volterra series model for surface runoff, *Water Resour. Res.*, 1986, vol. 22, no. 10, pp. 1413–1421.
14. Napiórkowski, J.J. and Piotrowski, A., Artificial neural networks as an alternative to the volterra series in rainfall-runoff modeling, *Acta Geophys. Pol.*, 2005, vol. 53, no. 4, pp. 459–472.
15. Napiórkowski, J.J. and Strupczewski, W.G., Problems involved in identification of the kernels of Volterra series, *Acta Geophys. Pol.*, 1984, vol. XXXII, no. 4, pp. 375–391.
16. Nikolaou, M. and Mantha, D., Efficient nonlinear modeling using wavelet compression, *Progress in Systems and Control Theory*, 2000, vol. 26, pp. 313–334.
17. Papazafiriou, Z.G., Linear and nonlinear approaches for short-term runoff estimations in time invariant open hydrologic systems, *J. Hydrol.*, 1976, vol. 30, pp. 63–80.