Coupling soil moisture and precipitation observations for predicting hourly runoff at small catchment scale

Gokmen Tayfur [a,*], Graziano Zucco [b], Luca Brocca [b], Tommaso Moramarco [b]

[a] Dept. Civil Engineering, Izmir Institute of Technology, Urla, Izmir, Turkey
[b] Research Institute for Geo-Hydrological Protection, CNR, Perugia, Italy

SUMMARY

The importance of soil moisture is recognized in rainfall–runoff processes. This study quantitatively investigates the use of soil moisture measured at 10, 20, and 40 cm soil depths along with rainfall in predicting runoff. For this purpose, two small sub-catchments of Tiber River Basin, in Italy, were instrumented during periods of October 2002–March 2003 and January–April 2004. Colorso Basin is about 13 km^2 and Niccone basin 137 km^2. Rainfall plus soil moisture at 10, 20, and 40 cm formed the input vector while the discharge was the target output in the model of generalized regression neural network (GRNN). The model for each basin was calibrated and tested using October 2002–March 2003 data. The calibrated and tested GRNN was then employed to predict runoff for each basin for the period of January–April 2004. The model performance was found to be satisfactory with determination coefficient, R^2, equal to 0.87 and Nash–Sutcliffe efficiency, NS, equal to 0.86 in the validation phase for both catchments. The investigation of effects of soil moisture on runoff prediction revealed that the addition of soil moisture data, along with rainfall, tremendously improves the performance of the model. The sensitivity analysis indicated that the use of soil moisture data at different depths allows to preserve the memory of the system thus having a similar effect of employing the past values of rainfall, but with improved GRNN performance.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

The importance of soil moisture on runoff, infiltration, and evapotranspiration is well recognized in the literature (Goodrich et al., 1994; Merz and Plate, 1997; Scipal et al., 2008; Brocca et al., 2009a, among others). Specific monitoring programs (Merz and Bardossy, 1998; Aubert et al., 2003; Castillo et al., 2003; Brocca et al., 2009a; Matgen et al., 2012; Morbidelli et al., 2012) and modelling studies (Gautam et al., 2000; Anctil et al., 2004; Komma et al., 2008; Berthet et al., 2009; Sheikh et al., 2009; Brocca et al., 2012; Tramblay et al., 2012; Van Steenbergen and Willems, 2013) were carried out to investigate the influence of soil moisture on producing runoff hydrographs.

Several studies investigated the benefit of using soil moisture observations within rainfall–runoff models. First applications considered soil moisture data for the improvement of the calibration and verification of rainfall–runoff models (Wooldridge et al., 2003; Koren et al., 2008; Parajka et al., 2009). Other studies directly used the observations for the assessment of the antecedent

wetness conditions through in situ (Meyles et al., 2003; Huang et al., 2007; Tramblay et al., 2010; Zehe et al., 2010) and remotely sensed (Jacobs et al., 2003; Brocca et al., 2009b; Beck et al., 2009) estimates. For instance, Goodrich et al. (1994) monitored two sub-catchments of 0.044 km^2 and 6.31 km^2 size in Walnut Gulch experimental watershed of the U.S. Department of Agriculture. They pointed out that a basin-wide remotely sensed average initial soil moisture can be sufficient for rainfall–runoff modelling in semiarid regions provided that the spatial–temporal variability of rainfall is accurately known. Brocca et al. (2009a), at five nested catchments (13–137 km^2) in central Italy, found that the integration of in situ soil moisture observations into a simple event-based rainfall–runoff model improved the prediction of runoff hydrographs with respect to the use of antecedent precipitation and baseflow indices. Grayson and Western (1998) suggested that a network of a limited number of soil moisture sensors can provide reliable estimates of areal mean soil moisture time series that can potentially be used as antecedent conditions data. All these studies suggest that in situ soil moisture data from a small number of locations can provide useful information for improving runoff prediction at the basin scale.

Several studies employed artificial intelligence techniques (or data driven techniques) for modelling the rainfall–runoff transformation process (see Elshorbagy et al., 2010 for a recent review of

* Corresponding author. Tel.: +90 533 5565339; fax: +90 232 750 6801.
E-mail addresses: gokmentayfur@iyte.edu.tr (G. Tayfur), graziano.zucco@irpi.cnr.it (G. Zucco), luca.brocca@irpi.cnr.it (L. Brocca), t.moramarco@irpi.cnr.it (T. Moramarco).

hydrological application). However, only few of them attempted to use (modelled or observed) soil moisture data for runoff prediction (Gautam et al., 2000; Anctil et al., 2004, 2008; Casper et al., 2007). Gautam et al. (2000) used a back propagation Neural Network (NN) to predict runoff employing only soil moisture data measured at 20 and 40 cm, without rainfall, in Tono catchment in Japan (0.52 km^2). The catchment had steep slopes and is located in a humid region where most of the time period it is saturated with very low evaporation rates (Gautam et al., 2000). The authors found the 40 cm soil moisture at a downslope location provides the most important information for runoff estimation. However, Gautam et al. (2000) concluded that every basin has its own characteristics and the use of different location, in space and in depth, might be more appropriate for soil moisture monitoring. Anctil et al. (2004) employed data from a river catchment (Seine River near Paris) that drains 1949 km^2 area with a travel time of about 3 days. They employed daily values of discharge, soil moisture index, rainfall and potential evaporation for predicting runoff through a multilayer perceptron NN. The use of the soil moisture index as additional input into the NN resulted in increased simulation skills. In another study, Anctil et al. (2008) used a similar dataset from 104 km^2 catchment in France for forecasting daily flow rates, but using in situ soil moisture observations at 5, 35 and 95 cm depth still obtaining improved performance.

Considering that the previous analyses are based on daily data at most, a Generalized Regression Neural Network (GRNN) was adopted in this study for predicting runoff by using half-hourly rainfall and in situ soil moisture observations at 10, 20 and 40 cm depth. In fact, the GRNN model is also found to slightly outperform other NN techniques (Kim et al., 2013) and, for that, it is selected for this study. Two different catchments, subjected to the Mediterranean climate, with drainage area of 13 and 137 km^2 were used in order to investigate the effect of the catchment area

on the GRNN performance along with to test the GRNN model at hourly time scale differently from Anctil et al. (2004, 2008).

2. Experimental catchments and dataset

The Niccone catchment (Fig. 1) is located in central Italy and is a subcatchment of the Tiber River Basin. The Niccone catchment has a size of 137 km^2 at Migianella river section with elevation between 249 and 887 m above sea level and a mean catchment slope of 25%. The Colorso stream is a tributary of the Niccone stream and the catchment has an area of 13 km^2 at Pian di Marte river section (see Fig. 1). The Niccone catchment is characterized by the Mediterranean climate with precipitation mostly in the autumn–spring period, when floods generally occur. Based on the period 1989–2012, the average annual precipitation was about 870 mm equally distributed on the whole catchment. The maximum mean monthly precipitation occurs in November (on average 125 mm) and the minimum in July (32 mm). The mean annual runoff is 270 mm ranging between 85 mm (in the dry year of 2007) and ~700 mm (in 2010). The peak runoff occurs during December and April. In the summer period, the minimum and maximum monthly temperature values, on average, are 13.5 °C and 22.5 °C, respectively; whereas in the winter period they are 4.0 °C and 10.5 °C, respectively.

The catchment lithology is mainly characterized by layered sandstones interleaved with thin marls and pelitic levels (72.8%), second by clays with chaotic structure (15.0%) and third by alluvial deposits constituted by gravel and sand (11.7%). The predominant land use, that covers nearly 60% of the total catchment area, is wooded and of the remainder 32% is cropland and 10% is rangeland.

The study area has been operating a dense hydrometeorological network (Fig. 1) with 8 raingauges, 2 hydrometric gauges and 2
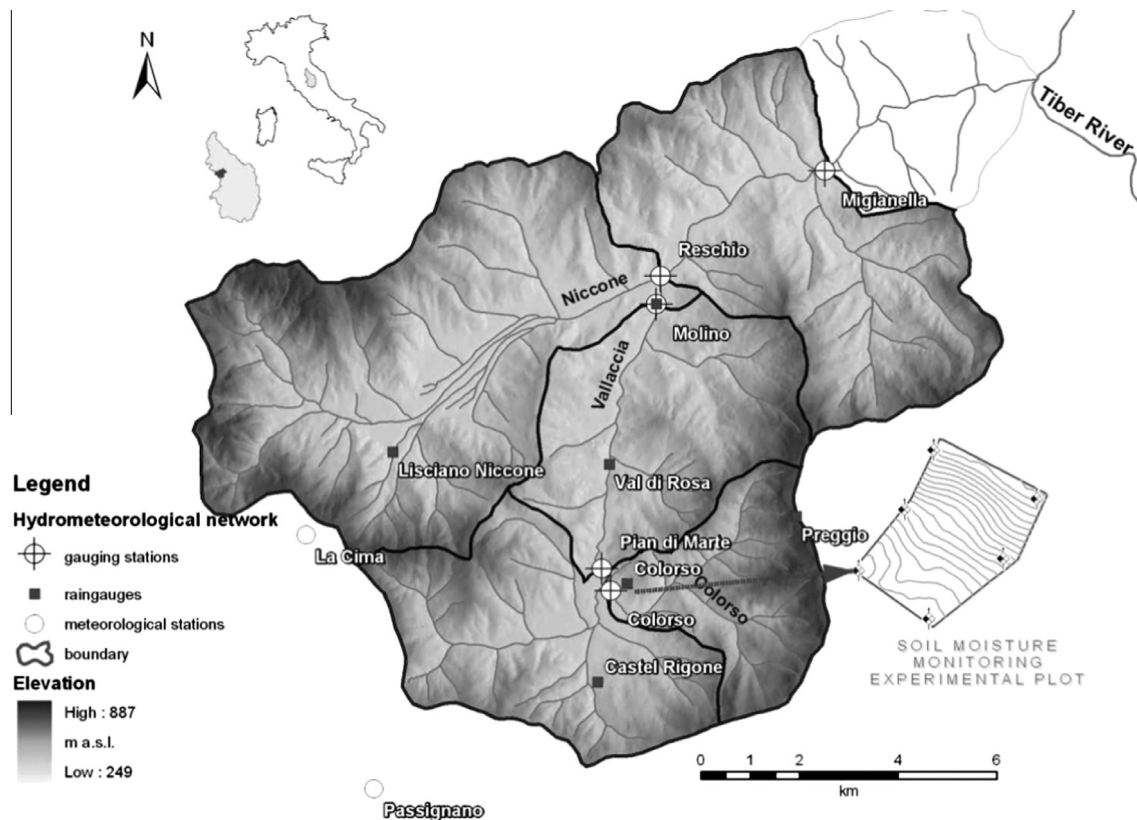

Fig. 1. Location of the catchments and hydrometeorological monitoring network.

Table 1
Main characteristics of the two nested study catchments.

Catchment	Area (km^2)	Mean slope (%)	River length (km)	Land use (%)				
				Wood land	Crop land	Range land	Pasture	Urban
Niccone	137	25	18.1	55	32	10	2.7	0.3
Colorso	13	28	5.4	71	17.9	0	10.9	0.1

meteorological stations and data have been collected since 1989 with a time resolution of 30 min. These hydrometric stations allow analyzing the hydrological behaviour of two different catchments (Niccone and Colorso) whose main geomorphological characteristics are given in Table 1.

The experimental soil moisture data set selected for this study covers the period from October 2002 to April 2004 except from April to December 2003, due to drawbacks with acquisition. Soil moisture was monitored at depths of 10 cm, 20 cm, and 40 cm with six continuous Frequency Domain Reflectometry (FDR) sensors located in an experimental plot of 1 ha (Sentek Sensor Technologies, 1997). The average of the six sensors at different depths is used as 'observed' soil moisture data here in after. Fig. 2 shows temporal variation of measured soil moisture data, along with precipitation and discharge for the Niccone and Colorso catchments subtended by the equipped section of Migianella and Pian di Marte, respectively. As it can be seen, soil moisture at 10 cm depth shows a rapid rise during heavy rainfall, followed by a recession which differs according to the period of the year and to the meteorological condition. The sensors at 20 and 40 cm depth show a delayed response to rainfall and also saturate conditions during very wet periods.

Previous studies in the Niccone and Colorso catchments (Brocca et al., 2009a; Camici et al., 2011) highlighted the strong influence of the initial soil moisture conditions in the runoff generation mechanism of the study area, as it usually occurs for Mediterranean catchments (Tramblay et al., 2012). Specifically, the basins are characterized by a quick runoff response (Brocca et al., 2011) likely dominated by shallow subsurface flow taking place near the surface layer of soil. Therefore, it would be of interest to investigate which soil layer needs to be monitored for the estimation of runoff in these basins. This issue is addressed by the application of the ANN technique in this study.

3. Generalized regression neural network (grnn)

GRNN is a method for estimating a joint probability density function (*pdf*) of x and y given only a training set. If $f(x,y)$ represents a joint continuous *pdf* of a vector random variable x and a scalar random variable y, then the estimation can be expressed as follows (Cigizoglu, 2005):

$$E[y|x] = \frac{\int_{-\infty}^{\infty} y f(x,y) dy}{\int_{-\infty}^{\infty} f(x,y) dy} \tag{1}$$

The density function $f(x,y)$ must usually be estimated from a sample of observations of x and y (Cigizoglu, 2005):

$$\hat{f}(x,y) = \frac{1}{(2\pi)^{(p+1)/2} \sigma^{(p+1)}}$$
$$\times \frac{1}{n} \sum_{i=1}^{n} \exp\left[-\frac{(x-x_i)^T(x-x_i)}{2\sigma^2}\right] \exp\left[-\frac{(y-y_i)^2}{2\sigma^2}\right] \tag{2}$$

where $\hat{f}(x,y)$ is the probability estimator, which is based upon sample values x_i and y_i of the random variables **x** and **y**, n is the number of sample observations, T is the transpose of vector **(x−xᵢ)**. σ is the smoothing parameter and p is the dimension of

the vector variable **x**. Note that x_i and y_i being the input and output data respectively in Eq. (2) forms a training pattern.

A physical interpretation of the probability estimate $\hat{f}(x,y)$ is that it assigns sample probability of width σ for each sample x_i and y_i, and the probability estimate is the sum of those sample probabilities. Defining the scalar function d_i as:

$$d_i^2 = (x-x_i)^T(x-x_i) \tag{3}$$

and substituting into Eq. (2) and performing the indicated integrations in Eq. (1) yields the following for the estimation of output variable, y:

$$\hat{y}(x) = \frac{\sum_{i=1}^{n} y_i \exp\left(-\frac{d_i^2}{2\sigma^2}\right)}{\sum_{i=1}^{n} \exp\left(-\frac{d_i^2}{2\sigma^2}\right)} \tag{4}$$

In Eq. (4), y_i represents output data while d_i is defined by Eq. (3) involving input data. When σ is made large, the estimated density is forced to be smooth and in the limit becomes a multivariate Gaussian with covariance $\sigma^2 I$ where I is the identity (Cigizoglu, 2005). On the other hand, a smaller value of σ allows the estimated density to assume non-Gaussian shapes, but with a disadvantage that extreme points may have a negative effect on the estimate. Its optimal value is often determined experimentally by using input–output data in Eq. (4) (Kim et al., 2013). Eq. (4) is directly applicable to problems involving numerical data.

The structure of the GRNN consists of four layers (Fig. 3): Input layer, Pattern layer, Summation layer, and Output layer (Tsoukalas and Uhrig, 1997). Neurons in each layer are fully connected to each other. The number of input units in the first layer is equal to the total number of input variables. The second layer has the pattern units and the outputs of this layer are passed onto the summation units in the third layer. Each unit in the pattern layer represents a training pattern and its output is a measure of the distance of the input from the stored patterns. The summation layer has two units: S-summation neuron and D-summation neuron (Seckin et al., 2013). The S-summation neuron computes the sum of the weighted outputs of the pattern layer while the D-summation neuron calculates the non-weighted outputs of the pattern neurons. The connection weights are set unity between pattern (hidden) layer neurons and the D-summation neuron (Kim et al., 2013). The output layer divides the output of each S-summation neuron by that of each D-summation neuron, yielding the predicted value to an unknown vector.

More details on GRNN are given elsewhere (Tsoukalas and Uhrig, 1997; Cigizoglu and Alp, 2006; Tayfur, 2012; Seckin et al., 2013; Kim et al., 2013). The GRNN, in the scientific literature, was employed for estimations of monthly flow rates (Cigizoglu, 2005), daily suspended sediments (Cigizoglu and Alp, 2006), annual peak flows (Seckin et al., 2013) and daily evaporation rates (Kim et al., 2013).

In this study, the package program of NeuroTools of Palisade Corporation was employed. The details of this package program can be obtained from the manual book of the NeuroTools (Palisade Corporation, 2012). The training, testing, and prediction procedures can be carried out in GRNN. Available data is generally first divided into two sets as training and prediction. The training set
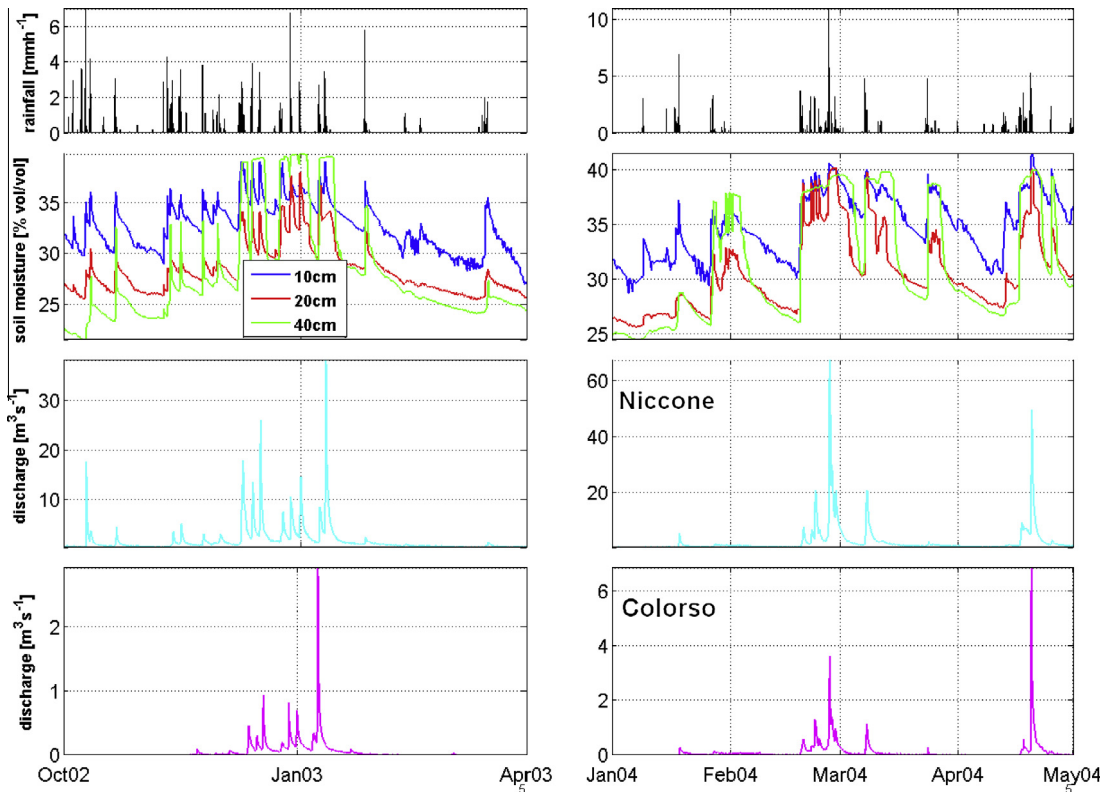
Fig. 2. Temporal variation of rainfall, soil moisture, and discharge for Niccone and Colorso catchments.
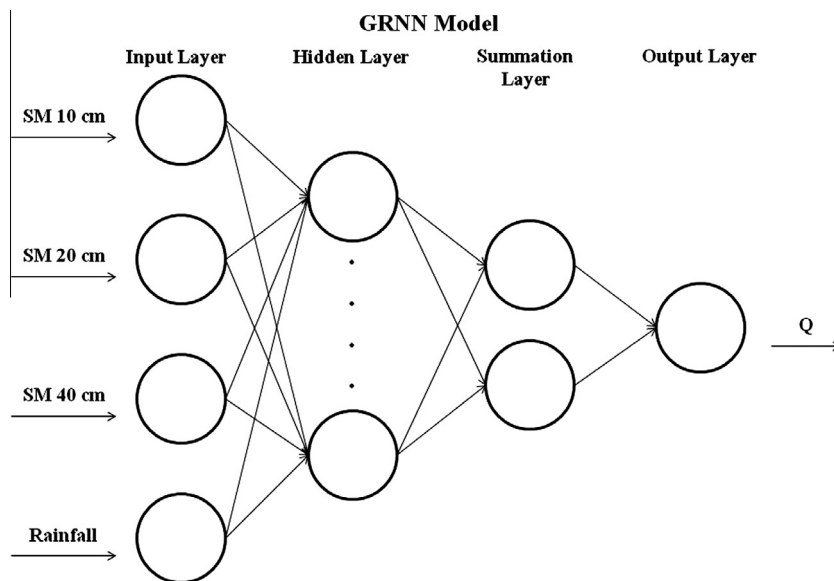

Fig. 3. Schematic representation of the Generalized Regression Neural Network (GRNN).

is then randomly divided into training and testing data sets. This portion is generally 80% for training and 20% for testing which is suggested in the literature (Palisade Corporation, 2012). When there is limited number of data sets, one can then reduce the portion of testing data sets or not to use it at all. This is because; the objective is, with sufficient data sets, to carry out the training successfully whereby the model learns generalization. Often, when there is enough data, which is the case in this study, in order to test the generalization of the model, data from the same training set (randomly chosen 20%) is used for testing the performance of the

training. The calibrated (and often tested) model can then be used to predict the data spared for the prediction stage.

4. Results

4.1. Investigating influence of soil moisture on runoff prediction

In order to investigate the importance of soil moisture data for runoff prediction, the whole dataset (October 2002–March 2003

plus January 2004–March 2004) was initially used in the GRNN model. First, rainfall was the only input. Then, in each model, soil moisture at 10 cm (denoted by SM10 hereafter), then at 20 cm (denoted by SM20 hereafter), and then at 40 cm (denoted by SM40 hereafter) were included in the input vector. For each model configuration, the performance of the model was evaluated in terms of coefficient of determination, R^2, and Root Mean Square Error, RMSE. Fig. 4 summarizes the results for Niccone and Colorso catchments. As it can be seen, by employing only the rainfall as input to predict runoff provided very low performance ($R^2 = 0.04$ and 0.09 for Niccone and Colorso, respectively). By including SM10 along with rainfall tremendously improved the model performance ($R^2 = 0.69$ and 0.78). Then, adding SM20 further improved the results ($R^2 = 0.92$ and 0.97) and finally, SM40 slightly improved the runoff predictions only for Niccone catchment ($R^2 = 0.99$). Moreover, for Niccone catchment the higher improvements in the GRNN model results were obtained by considering SM20, differently from Colorso for which SM10, in general, provided better performance. These differences might be explained by the soil layer depth that is more influencing the runoff generation in the two catchments. Specifically, a shallower soil layer is found to be more representative for the smallest Colorso catchment that is characterized by a faster runoff response. Interestingly, the model performance was found to be quite similar for the two catchments characterized by different size (13 and 137 km^2). This highlights that local scale (1 ha) soil moisture measurements contains useful information that can be scaled up to ~150 km^2 without losing in GRNN model performance (at least for the case study analyzed here).

4.2. Model training, testing, and prediction

A more robust assessment of the GRNN model was carried out by training the model with data (SM10, SM20, SM40, rainfall and discharge) of the period of October 2002–March 2003 for both catchments. Note that when the GRNN model was trained, 20% of the training data were randomly chosen for testing the performance of the model (see below). This is done to test the generalization capability of the trained network. The more data are used in training of the network; the better would be the generalization

of the model (Tayfur, 2012). Hence, as pointed out earlier, generally 20% of the data is spared for testing purpose considering that the model is further tested in the prediction stage.

Figs. 5 and 6 present the model simulations at the end of the training and testing stages for Niccone and Colorso catchments. Table 2 summarizes the performance measures for the respective stages for both the catchments.

The training were successfully accomplished for both the catchments with $R^2 = 0.97$ for Niccone and $R^2 = 0.81$ for Colorso catchment. The capturing of the temporal variation of runoff, including low and high values was satisfactory for Niccone catchment (Fig. 5). For Colorso catchment, the model could not capture the peak value during the training stage, although the temporal variation of discharge was well simulated (Fig. 6). The related scatter diagram shows the data more closely and evenly distributed along the regression line in the case of Niccone catchment (Fig. 5) than that of Colorso catchment (Fig. 6).

The testing was carried out against 20% of the randomly chosen training data (October 2002–March 2003) for both the catchments. The results are shown in Figs. 5 and 6 for the two catchments. As it can be seen, the models performed satisfactorily with $R^2 = 0.77$ for Niccone and $R^2 = 0.75$ for Colorso catchment (see Table 2). The model almost captured the temporal variation of runoff including the peak values in the case of Niccone catchment while a slight underestimation was observed for Colorso catchment.

The trained and tested models were then employed for predictive purpose for the period January 2004–March 2004. Fig. 7 show the prediction results for each catchment. The model performance is satisfactory for both catchments with $R^2 = 0.87$ and NS = 0.86 (see Table 2). The models well captured the temporal variation of runoff including the peak values. As before (see Fig. 4), the model performance is the same for both catchments.

4.3. Sensitivity analysis

The hourly cross-correlation values between the input variables (SM10, SM20, SM40, rainfall) and the output variable (discharge) were obtained for both the catchments. Fig. 8 shows the sample cross-correlation versus the lag (in hours) for both catchments.
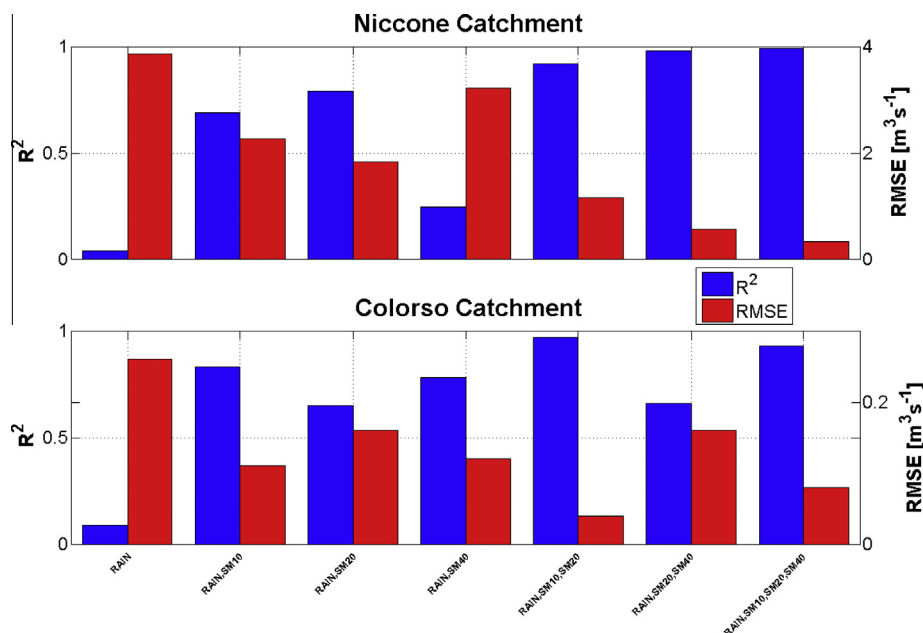

Fig. 4. Determination coefficient, R^2, and Root Mean Square Error, RMSE, for predicting runoff through the GRNN model and by using different configurations for the input vector (RAIN: rainfall, SM10, SM20, SM40: soil moisture at 10, 20 and 40 cm).
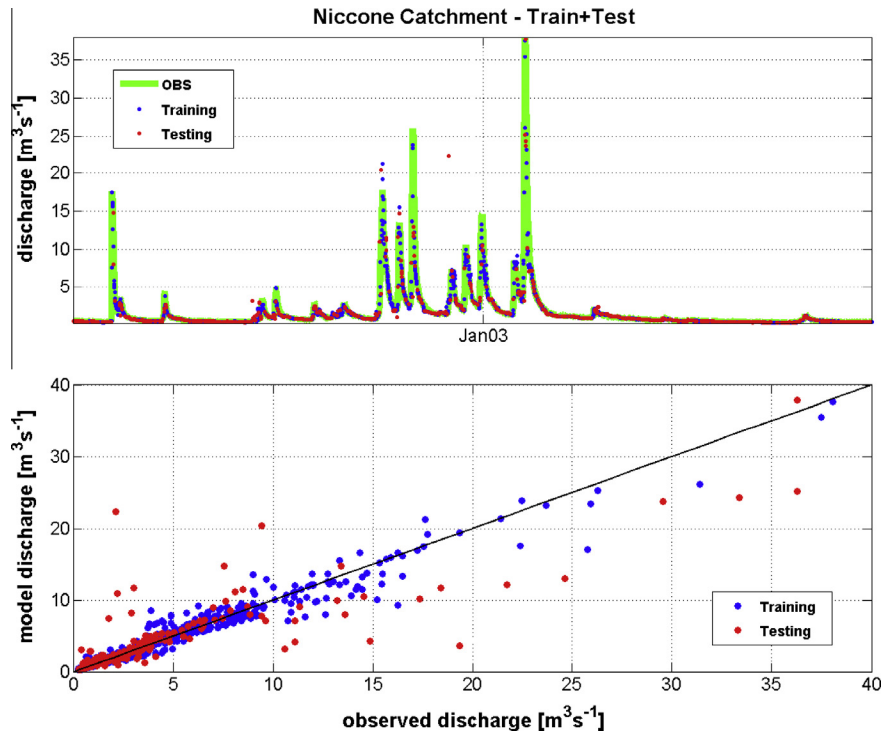
Fig. 5. Training and testing results for Niccone catchment at Pian di Marte: timeseries (upper panel) and scatter diagram (lower panel).
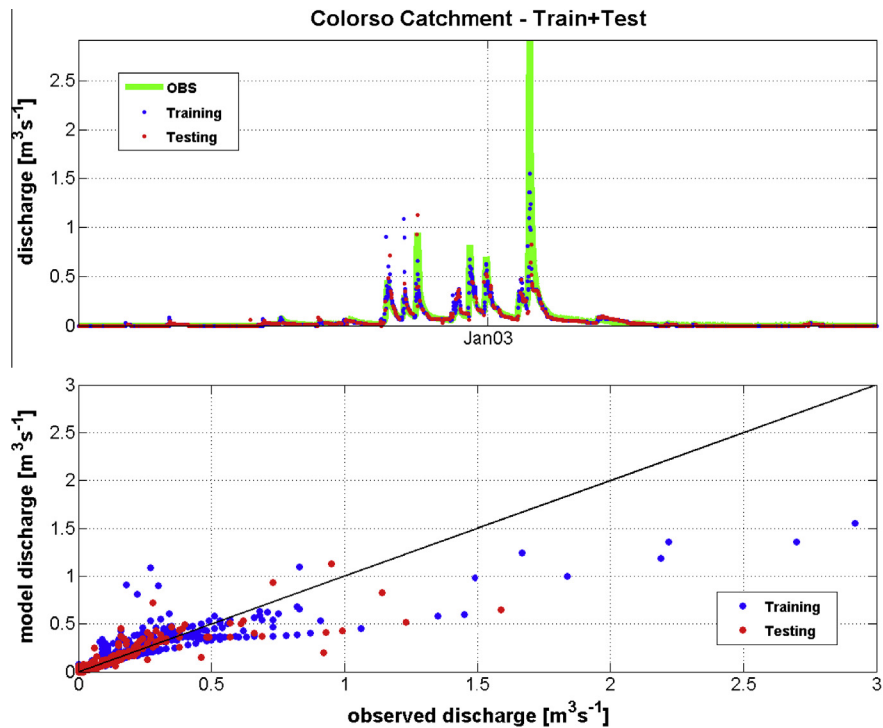

Fig. 6. As in Fig. 5 for Colorso catchment.

The lag values did not change much for the soil moisture content data at different depths for both the catchments with, more often than not, maximum values at lag zero. For rainfall data, a recognizable change of the cross-correlation with lag was obtained. Hence, several test scenarios were carried out to investigate whether the inclusion of past rainfall data in the input vector would improve the results.

Table 3 presents obtained R^2-values for different test cases. As discussed earlier in Fig. 4, SM20 plays a major role in Niccone catchment while SM10 for Colorso basin. In other words, SM20 significantly improves GRNN model performance for Niccone catchment while SM10 does it for Colorso basin. Hence, as a first choice of input vector; SM20 plus Rainfall, RAIN, for Niccone catchment; and SM10 plus Rainfall for Colorso basin were employed. As
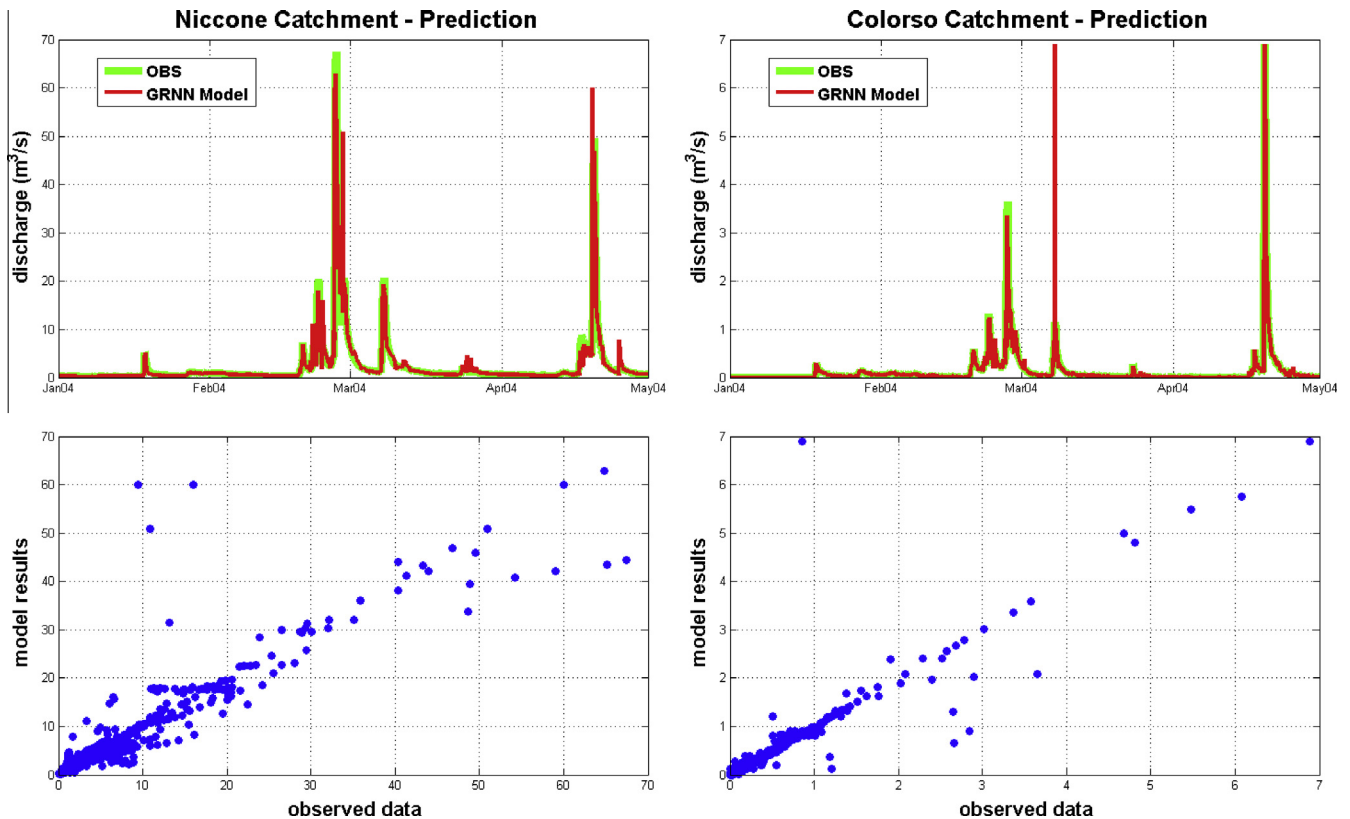
Fig. 7. Validation results for Niccone (left) and Colorso (right) catchments: timeseries (upper panels) and scatter diagram (lower panels).

Table 2
GRNN model performance, in the calibration and validation periods, in terms of determination coefficient, R^2, Root Mean Square Error, RMSE, and Nash–Sutcliffe efficiency, NS. The input data vector contains rainfall and soil moisture data at the depths of 10, 20 and 40 cm.

	R^2	RMSE ($m^3\ s^{-1}$)	NS
Niccone			
Training	0.97	0.49	0.97
Testing	0.77	1.66	0.70
Prediction	0.87	1.89	0.87
Colorso			
Training	0.81	0.07	0.80
Testing	0.75	0.07	0.62
Prediction	0.87	0.14	0.86

seen in Table 3, these input vectors provided the same performance for both catchments with $R^2 \sim 0.78$. When past values of rainfall, RAIN($t - 1$), ..., RAIN($t - 5$), were added in the each input vector, the prediction results improved from ~ 0.78 to ~ 0.87 for both catchments. On the other hand, the results did not further improve after adding more moisture information (see Table 3).

When one compares results in Tables 2 and 3; input vector of SM10, SM20, SM40 plus present RAIN(t) produces the same performance of $R^2 = 0.87$ (Table 2) as the input vector of SM20 plus present and 5 past values rainfall for Niccone catchment (Table 3). Similarly, input vector of SM10, SM20, SM40 plus present RAIN(t) produces the same performance of $R^2 = 0.87$ (Table 2) as the input vector of SM10 plus present and 5 past values of rainfall for Colorso basin (Table 3). That means using soil moistures measured at 10, 20 and 40 cm along with a present rainfall acts like employing only SM20, present rainfall and past values of rainfall for Niccone catchment. Similarly, employing soil moistures measured at 10, 20 and 40 cm along with a present rainfall acts like employing only SM10,

present rainfall and past values of rainfall for Colorso basin. These results imply that; the use of soil moisture data at different depths allows to preserving the memory of the system thus having a similar effect of employing the past values of rainfall.

5. Discussion

In setting up the ANN model; soil moisture information at 10 cm (SM10), 20 cm (SM20), and 40 cm (SM40) were used as variables, along with rainfall (RAIN). For each variable, one neuron was added to the input layer. The number of neurons in the hidden layer was then decided as twice plus one of the number of neurons in the input layer (ASCE 2000, Sen 2004, Tayfur, 2012). For instance, when SM10 and RAIN were included in the input vector; there were 2 and 5 neurons in the input and hidden layers, respectively. Addition of more variables in the input vector, as discussed above, resulted in change in the structure of the network and better performance of the model.

ANN is a black box model which cannot reveal insight into the physics of the process. On the other hand, it is quite practical and can provide solutions in an easy and non-costly way. If one were to apply the ANN model set up in this study to predict runoff in any other region or within the same catchments subject to climate and/or land use change; the network has to be restructured, recalibrated and retested. Although this seems a shortcoming of ANN, it is not really a major disadvantage. The restructuring, retraining and retesting can be done in a very short period of time, provided that data is available. In fact, we note that even for physics-based models the recalibration and revalidation procedures have to be performed, like the case in ANN, when they are applied for different regions and climates. The calibration and validation of the physics-based models, compared to ANN, would be timely and costly. This is because; the physics-based models, in addition to
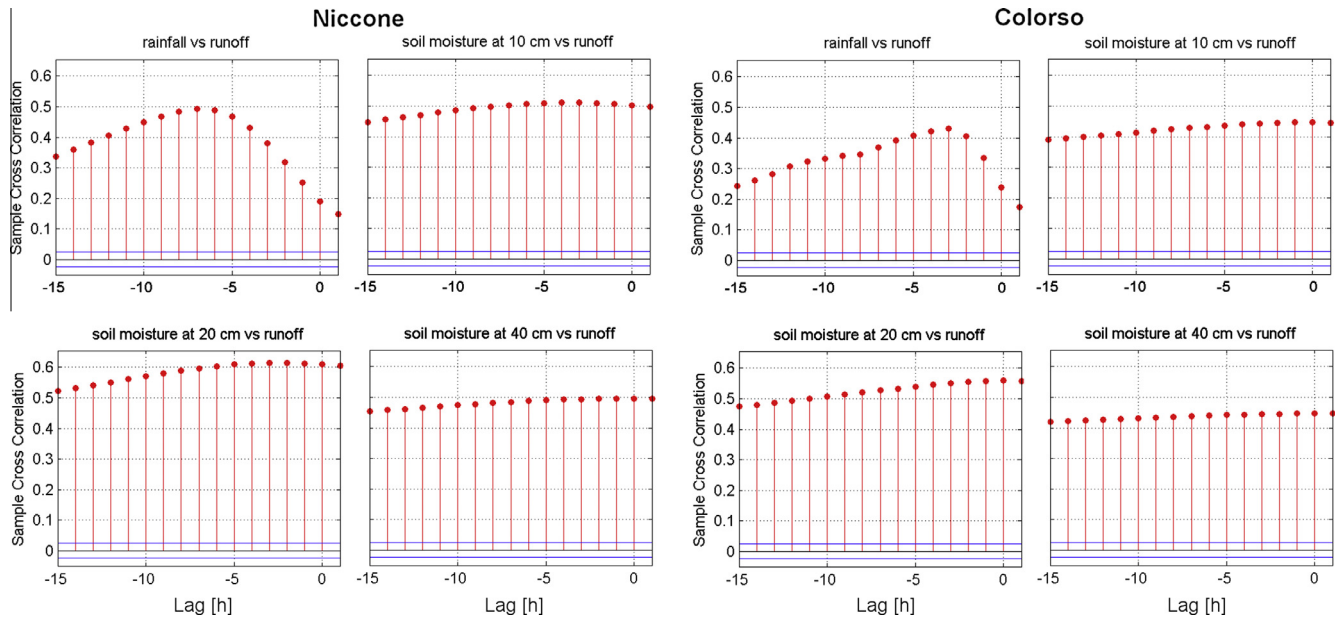
Fig. 8. Cross-correlation plots between rainfall and soil moisture at 10, 20 and 40 cm against runoff for Niccone (left) and Colorso.

Table 3
GRNN model performance in the prediction stage for different input vector (RAIN: rainfall, SM10, SM20, SM40: soil moisture at 10, 20 and 40 cm).

Input data	Prediction
Niccone catchment	R^2
SM20, RAIN(t)	0.79
SM20, RAIN(t), ..., RAIN($t - 5$)	0.87
SM10, SM20, SM40, RAIN(t), ..., RAIN($t - 5$)	0.87
Colorso catchment	R^2
SM10, RAIN(t)	0.78
SM10, RAIN(t), ..., RAIN($t - 5$)	0.88
SM10, SM20, SM40, RAIN(t), ..., RAIN($t - 5$)	0.87

inherent numerical stability and convergence problems, require substantial parameter estimation and distributed data. It can be then stated that ANN model could be an alternative to the physics based models in providing quick timely practical answers.

The importance of this study is that the use of soil moisture measured at 10, 20 and 40 cm soil depths in a 1 ha area (0.01 km^2) along with rainfall in ANN model provides fairly good estimates of runoff at 13 and 137 km^2 area watersheds. Therefore, for accurate estimations of runoff by ANN, it would be sufficient to monitor soil moisture in 1 ha area. This is an important outcome in the sense that there would be no need to have a dense soil moisture monitoring system to be installed all over the watershed, which would be too time consuming and costly.

Another important outcome of this study is that as the watershed size grows, the soil moisture monitoring depth increases. For 13 km^2 size Colorso basin, monitoring soil moisture at 10 cm soil depth may be sufficient for ANN model to predict runoff. On the other hand, for 137 km^2 Niccone catchment, this monitoring needs to be done in 20 cm soil depth. These would imply that if one were to predict runoff at much larger scale, then the soil depth might need to be increased.

Lastly, as it is presented in this study, the use of soil moisture in GRNN model improves predictions. Researchers in other parts of the world can use this information when building a neural network model. By considering the widespread availability of satellite soil moisture data (on a global scale and with daily temporal resolution, Massari et al., 2013), besides from in situ networks, the possibility to incorporate soil moisture data into rainfall–runoff modelling with neural network models could be easily done anywhere. This underlines the potential broader impact of this study.

6. Conclusions

This study employed GRNN model to predict hourly runoff from two small sizes (13 km^2 and 137 km^2) catchments in central Italy, employing rainfall and soil moisture data measured at 10, 20, and 40 cm soil depths in the input data vector. The following conclusions can be drawn from this study:

1. By employing only rainfall as an input into the GRNN model to predict runoff gives inadequate results (Fig. 4).
2. The addition of soil moisture data into the input data vector (along with rainfall) tremendously improves the runoff predictive capability of the GRNN model (Fig. 4).
3. For both catchments, the GRNN model was found to satisfactorily reproduce the observed discharge variations with determination coefficient, R^2, equal to 0.87 and Nash–Sutcliffe efficiency, NS, equal to 0.86 in the validation phase (Fig. 7 and Table 2).
4. The similar model performance for the two catchments, that are characterized by different size (13 and 137 km^2), highlights that local scale (1 ha) soil moisture measurements contains useful information that can be scaled up to ~150 km^2.
5. The use of present and past values of rainfall data, along with surface soil moisture, only slightly improves the GRNN performance.
6. Using soil moisture data measured at different depths together with present rainfall to predict discharge performs the same as employing only the shallow soil moisture data together with present and past (lag) rainfall values in the input vector. Thus, soil moisture data at different depths has the effect to preserve the memory of the system.

This study employed soil moisture and rainfall data from two catchments in central Italy, which is subject to the Mediterranean climate, having generally dry period from June to September and a rainy season from October to May. Hence, the conclusions drawn in

this study should be considered within this context. For different regions under different climate conditions, the GRNN model needs to be recalibrated and validated and for those areas studies will be addressed as soon as data will be made available. The assessment of the capability of soil moisture data to give useful information to predict runoff at different scales will be also analyzed in order to determine up to which spatial scale point soil moisture information can be employed.

Acknowledgements

References

Anctil, F., Michel, C., Perrin, C., Andreassian, V., 2004. A soil moisture index as an auxiliary ANN input for stream flow forecasting. J. Hydrol. 286, 155–167.

Anctil, F., Lauzon, L., Filion, M., 2008. Added gains of soil moisture content observations for streamflow predictions using neural networks. J. Hydrol. 359 (3–4), 225–234.

Aubert, D., Loumagne, C., Oudin, L., 2003. Sequential assimilation of soil moisture and streamflow data in a conceptual rainfall runoff model. J. Hydrol. 280, 145–161.

Beck, H.E., de Jeu, R.A.M., Schellekens, J., van Dijk, A.I.J.M., Bruijnzeel, L.A., 2009. Improving Curve Number based storm runoff estimates using soil moisture proxies, Selected Topics in Applied Earth Observations and Remote Sensing. IEEE Journal 2 (4), 250–259.

Berthet, L., Andréassian, V., Perrin, C., Javelle, P., 2009. How crucial is it to account for the Antecedent Moisture Conditions in flood forecasting? Comparison of event-based and continuous approaches on 178 catchments. Hydrol. Earth Syst. Sci. 13, 819–831.

Brocca, L., Melone, F., Moramarco, T., Singh, V.P., 2009a. Assimilation of observed soil moisture data in storm rainfall–runoff modelling. J. Hydrol. Eng. ASCE 14 (2), 153–165.

Brocca, L., Melone, F., Moramarco, T., Morbidelli, R., 2009b. Antecedent wetness conditions estimation based on ERS scatterometer data. J. Hydrol. 364 (1–2), 73–87.

Brocca, L., Melone, F., Moramarco, T., Penna, D., Borga, M., Matgen, P., Heitz, S., 2011. Investigation of the hydrologic response of three experimental basins across Europe. Die Bodenkultur 62 (1–4), 31–37.

Brocca, L., Moramarco, T., Melone, F., Wagner, W., Hasenauer, S., Hahn, S., 2012. Assimilation of surface and root-zone ASCAT soil moisture products into rainfall–runoff modelling. IEEE Trans. Geosci. Remote Sens. 50 (7), 2542–2555.

Camici, S., Tarpanelli, A., Brocca, L., Melone, F., Moramarco, T., 2011. "Design soil moisture" estimation by comparing continuous and storm-based rainfall–runoff modelling. Water Resour. Res. 47, W05527. http://dx.doi.org/10.1029/2010WR009298.

Casper, M., Gemmar, P., Gronz, O., Johst, M., Stüber, M., 2007. Fuzzy logic-based rainfall–runoff modelling using soil moisture measurements to represent system state. Hydrol. Sci. J. 52 (3), 478–490.

Castillo, V.M., Gomez-Plaza, A., Martinez-Mena, M., 2003. The role of antecedent soil water content in the runoff response of semiarid catchments: a simulation approach. J. Hydrol. 284, 114–130.

Cigizoglu, H.K., 2005. Generalized regression neural network in monthly flow forecasting. Civil Eng. Environ. Syst. 22 (2), 71–84.

Cigizoglu, H.K., Alp, M., 2006. Generalized regression neural network in modelling river sediment yield. Adv. Eng. Softw. 37 (2), 63–68.

Elshorbagy, A., Corzo, G., Srinivasulu, S., Solomatine, D.P., 2010. Experimental investigation of the predictive capabilities of data driven modeling techniques in hydrology – Part 2: Application. Hydrol. Earth Syst. Sci. 14, 1943–1961.

Gautam, M.R., Watanabe, K., Saegusa, H., 2000. Runoff analysis in humid forest catchment with artificial neural network. J. Hydrol. 235, 117–136.

Goodrich, D.C., Schmugge, T.J., Jackson, T.J., Unkrich, C.L., Keefer, T.O., Parry, R., Bach, L.B., Amer, S.A., 1994. Runoff simulation sensitivity to remotely sensed initial soil water content. Water Resour. Res. 30 (5), 1393–1406.

Grayson, R.B., Western, A.W., 1998. Towards areal estimation of soil water content from point measurements: time and space stability of mean response. J. Hydrol. 207, 68–82.

Huang, M., Gallichand, J., Dong, C., Wang, Z., Shao, M., 2007. Use of soil moisture data and curve number method for estimating runoff in the Loess Plateau of China. Hydrol. Process. 21 (11), 1471–1481.

Jacobs, J.M., Myers, D.A., Whitfield, B.M., 2003. Improved rainfall/runoff estimates using remotely sensed soil moisture. J. Am. Water Resour. As. 4, 313–324.

Kim, S., Shiri, J., Kisi, O., Singh, V.P., 2013. Estimating daily pan evaporation using different data-driven methods and lag-time patterns. Water Resour. Manage 27, 2267–2286.

Komma, J., Blöschl, G., Reszler, C., 2008. Soil moisture updating by Ensemble Kalman Filtering in real-time flood forecasting. J. Hydrol. 357 (3–4), 228–242.

Koren, V., Moreda, F., Smith, M., 2008. Use of soil moisture observations to improve parameter consistency in watershed calibration. Phys. Chem. Earth 33 (17–18), 1068–1080.

Massari, C., Brocca, L., Barbetta, S., Papathanasiou, C., Mimikou, M., Moramarco, T., 2013. Using globally available soil moisture indicators for flood modelling in Mediterranean catchments. Hydrol. Earth Syst. Sci. Discus. 10, 10997–11033.

Matgen, P., Heitz, S., Hasenauer, S., Hissler, C., Brocca, L., Hoffmann, L., Wagner, W., Savenije, H.H.G., 2012. On the potential of METOP ASCAT-derived soil wetness indices as a new aperture for hydrological monitoring and prediction: a field evaluation over Luxembourg. Hydrol. Process. 26, 2346–2359.

Merz, R., Bardossy, A., 1998. Effects of spatial variability on the rainfall runoff process in a small loess catchment. J. Hydrol. 212–213, 304–317.

Merz, R., Plate, E.J., 1997. An analysis of the effects of spatial variability of soil and soil moisture on runoff. Water Resour. Res. 33 (12), 2909–2922.

Meyles, E., Williams, A., Ternan, L., Dowd, J., 2003. Runoff generation in relation to soil moisture patterns in a small Dartmoor catchment. Southwest England. Hydrol. Processes 17, 251–264.

Morbidelli, R., Corradini, C., Saltalippi, C., Brocca, L., 2012. Initial soil water content as input to field-scale infiltration and surface runoff models. Water Resour. Manage 26 (7), 1793–1807.

Palisade Corporation. 2012. NeuroTool6, Sophisticated Neural Networks for Excel. User Manual. 798 Cascadilla Street, Ithaca, NY 14850, USA.

Parajka, J., Naemi, V., Bloschl, G., Komma, J., 2009. Matching ERS scatterometer based soil moisture patterns with simulations of a conceptual dual layer hydrologic model over Austria. Hydrol. Earth Syst. Sci. 13, 259–271.

Scipal, K., Drusch, M., Wagner, W., 2008. Assimilation of a ERS scatterometer derived soil moisture index in the ECMWF numerical weather prediction system. Adv. Water Resour. 31, 1101–1112.

Seckin, N., Cobaner, M., Yurtal, R., Haktanir, T., 2013. Comparison of Artificial Neural Network Methods with L-moments for Estimating Flood Flow at Ungauged Sites: the Case of East Mediterranean River Basin, Turkey. Water Resour. Manage. 27, 2103–2124.

Sentek Sensor Technologies. Enviroscan: Hardware Manual, version 3.0. Sentek Pty Ltd., Australia, 1997.

Sheikh, V., Visser, S., Stroosnijder, L., 2009. A simple model to predict soil moisture: Bridging Event and Continuous Hydrological (BEACH) modelling. Environ. Model Softw. 24, 542–556.

Tayfur, G., 2012. Soft Computing Methods in Water Resources Engineering. WIT Press, Southampton, England, 267 p..

Tramblay, Y., Bouvier, C., Martin, C., Didon-Lescot, J.F., Todorovik, D., Domergue, J.M., 2010. Assessment of initial soil moisture conditions for event-based rainfall–runoff modelling. J. Hydrol. 380 (3–4), 305–317.

Tramblay, Y., Bouaicha, R., Brocca, L., Dorigo, W., Bouvier, C., Camici, S., Servat, E., 2012. Estimation of antecedent wetness conditions for flood modelling in Northern Morocco. Hydrol. Earth Syst. Sci. 16, 4375–4386.

Tsoukalas, L.H., Uhrig, R.E., 1997. Fuzzy and Neural Approaches in Engineering. Wiley, NY.

Van Steenbergen, N., Willems, P., 2013. Increasing river flood preparedness by real-time warning based on wetness state conditions. J. Hydrol. 489, 227–237.

Wooldridge, S.A., Kalma, J.D., Walker, J.P., 2003. Importance of soil moisture measurements for inferring parameters in hydrologic models of low-yielding ephemeral catchments. Environ. Model. Softw. 18 (1), 35–48.

Zehe, E., Graeff, T., Morgner, M., Bauer, A., Bronstert, A., 2010. Plot and field scale soil moisture dynamics and subsurface wetness control on runoff generation in a headwater in the Ore mountains. Hydrol. Earth Syst. Sci. 14, 873–889.