Field Crops Research 121 (2011) 441–449

Contents lists available at ScienceDirect

Field Crops Research

journal homepage: www.elsevier.com/locate/fcr

Simulation for response of crop yield to soil moisture and salinity with artificial neural network

Xiaoqin Dai[a,*], Zailin Huo[b], Huimin Wang[a]

[a] Qianyanzhou Station, Key Laboratory of Ecosystem Network Observation and Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, 100101, PR China
[b] Center for Agricultural Water Research in China, China Agricultural University, Beijing 100083, PR China

ARTICLE INFO

Article history:
Received 30 October 2010
Received in revised form 29 January 2011
Accepted 29 January 2011

Keywords:
Artificial neural network
Soil water
Soil salinity
Sunflower yield

ABSTRACT

In saline fields, irrigation management often requires understanding crop responses to soil moisture and salt content. Developing models for evaluating the effects of soil moisture and salinity on crop yield is important to the application of irrigation practices in saline soil. Artificial neural network (ANN) and multi-linear regression (MLR) models respectively with 10 (ANN-10, MLR-10) and 6 (ANN-6, MLR-6) input variables, including soil moisture and salinity at crop different growth stages, were developed to simulate the response of sunflower yield to soil moisture and salinity. A connection weight method is used to understand crop sensitivity to soil moisture and salt stress of different growth stages. Compared with MLRs, both ANN models have higher precision with RMSEs of 1.1 and 1.6 t ha^{-1}, REs of 12.0% and 17.3%, and R^2 of 0.84 and 0.80, for ANN-10 and ANN-6, respectively. The sunflower sensitivity to soil salinity varied with the different soil salinity ranges. For low and medium saline soils, sunflower yield was more sensitive at crop squaring stage, but for high saline soil at seedling stage. High soil moisture content could compensate the yield decrease resulting from salt stress regardless of salt levels at the crop sowing stage. The response of sunflower yield to soil moisture at different stages in saline soils can be understood through the simulated results of ANN-6. Overall, the ANN models are useful for investigating and understanding the relationship between crop yield and soil moisture and salinity at different crop growth stages.

1. Introduction

Salinity has remained an important threat to global agriculture (Lobell et al., 2007), which is becoming more prevalent with the intensity of land use increase worldwide (Meloni et al., 2003). In arid and semi-arid regions, saline soils are especially abundant, in which the evaporation is intense and the amount of rainfall is insufficient for substantial leaching. Thus, salinity is one of the main limiting factors for agricultural production in these areas.

Reducing root-zone salinity is a beneficial strategy for improving crop emergence and stand establishment in saline fields (Clermont-Dauphina et al., 2010). Water management is the most readily available modifier of salt stress in the root-zone. An understanding of plant responses to water and salinity is of great practical significance. Recently, numerous studies have been carried out to investigate plant response to water and salt stress (Bassil and Kaffka, 2002; Katerji et al., 2003; Ashish et al., 2009; Chen et al., 2009; Clermont-Dauphina et al., 2010; Harris et al., 2010). However, decision-making processes in agriculture often require reliable crop response models to assess the impact of specific management practices and environmental conditions. Mathematical models prove to be a useful tool to define the best water management in saline conditions. There are two distinct modeling approaches, empirical and process models, for identifying crop yield responses to given environmental conditions and management options (James and Cutforth, 1996). Technologically, empirical crop growth models are relatively simple to build or develop, but these models are generally linear and have a lower modeling ability for complex ecological systems. Process-based crop growth models are often preferred to empirical ones, but these deterministic models need to identify many parameters (Sinclair and Seligman, 1996; Matthews, 2002a,b; Ziaei and Sepaskhah, 2003). Notably, deterministic calibration is very difficult in most cases (Ma et al., 2009). So an empirical model may offer an even more reliable method in investigating crop responses than poorly calibrated process models when the necessary data are available (Park et al., 2005). However, the traditional regression-based empirical models lack non-linear modeling ability for complex ecological systems, which is apparent in crop responses to agro-ecological conditions.

* Corresponding author. Tel.: +86 10 64889913.
E-mail address: daixqcas@yahoo.cn (X. Dai).

0378-4290/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.fcr.2011.01.016

Artificial neural network (ANN) methodology is an alternative modeling and simulation tool, which is specially designed for dynamic nonlinear systems. One of the most important traits of ANN models is their ability to adapt to recurrent changes and detect patterns in complex natural systems. During the last decade, there has been a significant increase in agronomic ANN application (Huang et al., 2010), including crop development modeling (Zhang et al., 2009; Fortin et al., 2010), crop yield prediction (Park et al., 2005; Green et al., 2007; Khazaei et al., 2008), evapotranspiration estimations (Dai et al., 2009; Liu et al., 2009), and soil water and salt content assessments (Zou et al., 2010). Kaul et al. (2005) investigated the artificial neural network models performance in predicting corn and soybean yields for typical climatic conditions, and came to the conclusion that ANN models consistently produced more accurate yield predictions than regression models. Alvarez (2009) used an artificial neural network approach to model the effects of soil and climate factors on average regional yield and production of wheat in the Argentine Pampas and found that ANN performed better in regional yield estimation than the regression or the blind guess methods. However, the ANN approach rarely has been used to model crop yield responses to soil environments, especially to soil moisture and soil salinity.

The objectives of this study are: (1) to develop an ANN model to determine the relationship between crop yield and soil moisture and salinity, and (2) to simulate the response of crop yield to different soil moisture and saline environments.

2. Materials and methods

2.1. Field experiment

The study was conducted at the Shahaoqu Experimental Station in the Hetao Irrigation District (40°19′–41°18′N, 106°20′–109°19′E), situated in the west side of the Inner Mongolia Autonomous Region, China. The region has an arid continental climate. Annual average temperature is 8.1 °C, with monthly mean temperature ranging from −10.1 °C in January to 23.8 °C in July. The soil is usually frozen for 5–6 months from late November to middle May. Frost-free days are for 135–150 d and annual sunshine duration is 3100–3300 h. Average annual precipitation is 150 mm with about 60% occurring between July and August. Potential evaporation is about 2200–2400 mm yr^{-1} (Feng et al., 2005). The physical and chemical properties in the experimental fields are listed in Table 1.

An experiment was set up from 2003 to 2005. Sunflower (*Helianthus annuus* L.) was designated as the experimental crop. Experimental plots were arranged in a split plot design with three replications. Three soil salinity levels (0.2–0.5, 0.5–0.8, 0.8–2.1 dS m^{-1}) at crop sowing stage were the main plots. Four irrigation schedules (3 deficits and 1 sufficient) were the subplots. The subplots size was 4 × 10 = 40 m^2. The detailed treatments are listed in Table 2. At the beginning of the experiments, soil saturated moisture content (θ_s) and field capacity (θ_f) was respectively 35% and 26%.

Sunflower (cv. KANGDI115) was planted on May 27, 2003, June 14, 2004 and June 3, 2005 with a spacing of 0.3 × 0.5 m and harvested on September 15, 2003, September 28, 2004 and September 19, 2005, respectively. Fertilization, weeding and other cultural managements were executed following local farmers' practices, which were similar among three years.

Soil samples were collected at crop sowing, seedling, squaring (i.e. floral bud initiation), flowering and maturity growth stages according to a systematic sampling design across the S-shaped transects. Six soil cores were collected from every plot and soil cores were divided into five depths (0–20, 20–40, 40–60, 60–80, and 80–100 cm), composited and mixed by depth. Soil moisture content was determined using gravimetric methods and electrical conductivity (EC) of soil (water:soil ratio of 5:1, dS m^{-1}) was measured by digital conductivity meter. For each plot, soil moisture content and EC of different soil depths at crop every growth stages were respectively averaged as the soil moisture content and EC of its corresponding stages. At sunflower harvest, grain yield was measured from a 9 m^2 area in each plot.

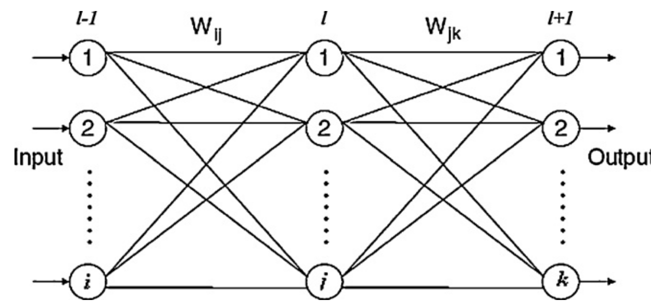

Fig. 1. Three-layer feed-forward ANN architecture.

2.2. Artificial neural networks

Of many ANN architectures reported, the back-propagation network (BP) have a simple structure for simulating complex system, and it is enough robust for the simulation of any non-linear system (Haykin, 1998). So the BP network was used in the study. The network consists of layers of parallel processing elements (neurons), with each layer being fully connected to the preceding layer by interconnection strengths or weights (W). Fig. 1 illustrates a three-layer neural network consisting of layers $l-1$, l and $l+1$ with the interconnection weights W_{ij} and W_{jk} between the neurons from adjacent layers.

Training a network includes a forward propagation of inputs and a backward propagation of errors. In the forward procedure, the effect of an applied activity pattern in input layer was propagated layer by layer through the network. The activation value a_j^l at jth neuron in lth layer is given by the following equation:

$$a_j^l = \sum_{i=1}^{n} W_{ij} O_i^{l-1} + b_j^l, \quad i = 1, 2, 3, \ldots, n, \quad j = 1, 2, 3, \ldots, m \quad (1)$$

where W_{ij} is the interconnection weight between the jth neuron in lth layer and the ith neuron in $(l-1)$ layer, O_i^{l-1} is the output of the ith neuron in the $(l-1)$th layer, b_j^l is the bias of the jth neuron in the lth layer. The activation value of a neuron was used to obtain its output value through a transfer function. The general sigmoidal logistic transfer function that can express any complex relationship was used in this study. It is given by:

$$f(a_j^l) = \frac{1}{1+\exp(-a_j^l)} \quad (2)$$

where exp denotes the natural exponential function. The function value of each neuron in the output layer was obtained by the input effect propagating layer by layer. The goal of ANN is to establish a relation of the form as expressed by:

$$Y^t = f(X^s) \quad (3)$$

where X^s is an s dimensional input predictor vector consisting of $x_1, x_2, \ldots, x_s$ and Y^t is a t dimensional output or target vector consisting of prediction variables of interest $y_1, y_2, \ldots, y_t$. Normally, the network is trained by a back-propagation algorithm and conjugate

Table 1
Soil properties of the experimental field.

	Soil depth (cm)					
	0–20	20–40	40–60	60–80	80–100	100–120
Bulk density (g cm^{-3})	1.51	1.52	1.47	1.46	1.46	1.46
Sand (%)	18.0	13.6	11.2	15.5	18.0	2.8
Silt (%)	66.0	70.4	76.8	73.5	68.0	70.2
Clay (%)	16.0	16.0	12.0	11.0	14.0	27.0

gradient learning algorithms, which adjusts the weights and biases to minimize the error. The error function is

$$E = \sum_{p}\sum_{t}(y_r - o_r)^2 \tag{4}$$

where E is error value, y_r the ANN computed yield of sample r, o_r the observed yield of sample r and p the number of training patterns or data sets.

In this study, we used the resampling method with cross-validation for training ANN. Total 108 samples were randomly averagely divided into 4 groups in which three groups were chosen to train ANN and the remainder was used to test ANN. Every ANN was trained four times using four different data sets. During network training phase, the training samples were processed through the ANN. Afterwards, the connection weights and biases were automatically adjusted until the maximal training times was achieved. Following training, the ANN was tested with the testing data set to assess its ability of generalizing system behavior.

In designing a robust and accurate ANN model, the modeler must address a number of important factors, including of type and structure of neural network (nodes number of hidden layer), input variables used, and data pre-processing, which are generally accomplished through a combination of best professional judgment, heuristic rules, and trial and error. The development of ANNs was performed with Matlab 6.5.

2.3. Connection weight method for quantifying variable importance in ANNs

Prediction accuracy is a major benefit of ANN models, but the ANN models of any physical process are purely black box models, which not to explain the process being simulated, and whose utility is limited without information regarding the relative importance of the parameters in the system. The development of a method to couple input factors to meaningful output in ANN models is of critical importance (Kemp et al., 2007). The data employed for developing ANN models do contain important information regarding the physical process being simulated (Jain et al., 2008).

A connection weight approach was used to evaluate the importance of inputs (soil moisture and salinity) to output (crop yield) in ANNs. The connection weight method is to sum the products of the input-hidden and the hidden-output connection weights between each input neuron and output neuron for all input variables (Olden et al., 2004). The relative contributions of the inputs to the output are dependent on the magnitude and direction of the connection weights. When the signs of the input-hidden and hidden-output connection weights are the same (i.e., either both are positive or both are negative), the input has a positive impact on the output. Contrarily, if the signs of these connection weights are opposite, the specific input has a negative effect on the output. The overall contribution of the input to the output depends on its sum of the positive and negative effect across all different hidden nodes. The larger the sum of the connection weights, the greater the importance of the variable. The relative importance of input variable i is determined through the following formula:

$$\mathrm{RI}_i = \frac{\sum_{j=1}^{m} W_{ij}W_{jk}}{\sum_{i=1}^{n}\sum_{j=1}^{m} W_{ij}W_{jk}} \times 100\%, \quad i = 1, 2, 3, \ldots, n, \quad j = 1, 2, 3, \ldots, m \tag{5}$$

where RI_i is the relative importance (expressed in percentage) of the variable i in the input layer on the output variable, j the index number of the hidden node, W_{ij} is the connection weight between input variable i and hidden node j, and W_{jk} is the connection weight between hidden node j and the output node k. The whole computation was repeated for each output neuron.

2.4. Evaluation criterion of model

To quantify the deviation in simulated results of ANN and multilinear regression (MLR) from the observed data, three statistical parameters, including root mean squared error (RMSE), relative error (RE) and coefficient of determination (R^2), were used in this

Table 2
Treatments in different saline soils.

Treatment[a]	Percentage of field capacity				Soil salinity at sowing stage (EC, dS m^{-1})
	Seedling	Squaring	Flowering	Maturity	
S1D1	70%	100%	100%	100%	0.21–0.5
S1D2	100%	55%	100%	100%	
S1D3	100%	100%	40%	100%	
S1D0	100%	100%	100%	100%	
S2D1	100%	70%	100%	100%	0.5–0.8
S2D2	100%	100%	55%	100%	
S2D3	40%	100%	100%	100%	
S2D0	100%	100%	100%	100%	
S3D1	100%	100%	70%	100%	0.8–2.1
S3D2	55%	100%	100%	100%	
S3D3	100%	40%	100%	100%	
S3D0	100%	100%	100%	100%	

[a] S1, S2, S3 denote three soil salinity level; D1, D2, D3 denote insufficient irrigation, D0 denote sufficient irrigation.

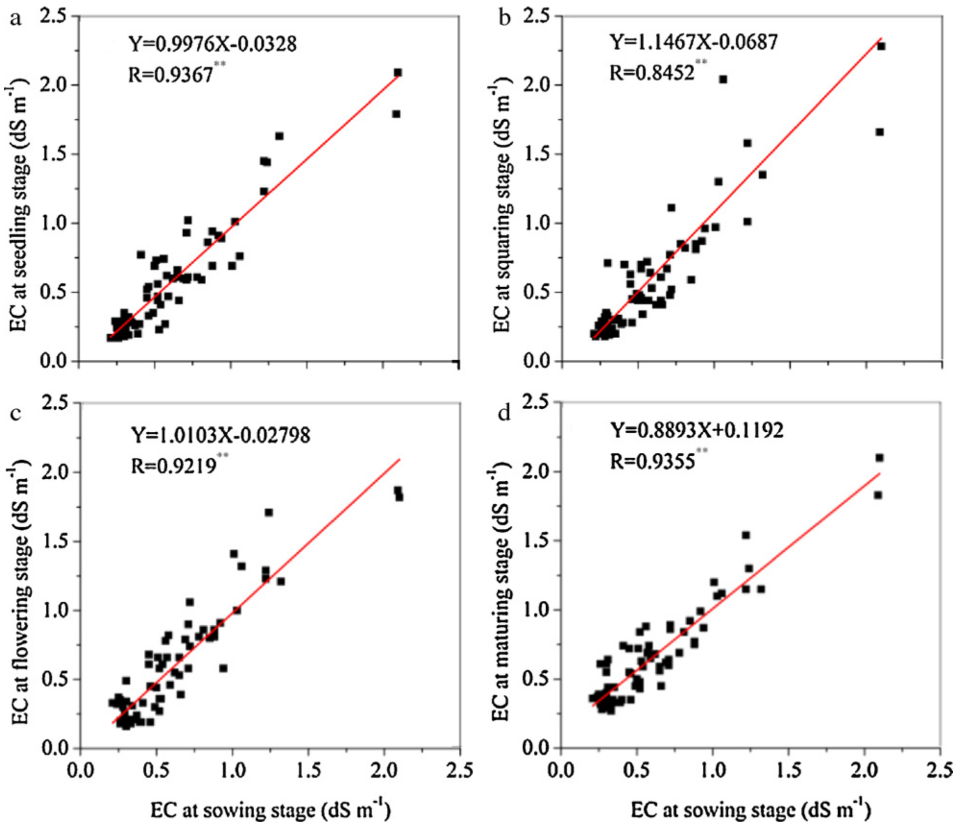

Fig. 2. Correlations of soil salinity at crop sowing stage with other growth stages.

study. The three parameters were calculated as following:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{h=1}^{n}(S_h - M_h)^2} \quad (6)$$

$$\mathrm{RE} = \frac{1}{n}\left(\sum_{h=1}^{n}\left|\frac{S_h - M_h}{M_h}\right|\right) \quad (7)$$

$$R^2 = \frac{\left[\sum(S_h - \bar{S})(M_h - \bar{M})\right]^2}{\sum(S_h - \bar{S})\sum(M_h - \bar{M})} \quad (8)$$

where S_h and M_h are the simulated and measured sunflower yield, respectively, $\bar{S}$ and $\bar{M}$ the average values of the data arrays of S_h and M_h, respectively, and n is the observed sample number.

3. Results

3.1. Artificial neural network model development

3.1.1. Input and output variables

One of the most important steps in the ANN development process is to determine input variables. For saline soil, sunflower growth is significantly influenced by soil moisture and salinity, and its response to water and salinity stress varies in different growth stages (Chen et al., 2009). In this study, soil moisture content and salinity at crop different growth stages were considered as the ANN input variables. As a result, 10 input variables (soil moisture content and EC of sunflower sowing, seedling, squaring, flowering and maturity stages) were included in the ANN. Furthermore, there was a significant positive relationship between soil EC at crop sowing stage and those at other growth stages (Fig. 2). Thus, an ANN with 6 input variables, including soil EC at crop sowing stage and soil moisture content at sowing, seedling, squaring, flowering and maturity growth stages, was also developed in this study. The outputs of the ANNs were sunflower grain yields.

Table 3
Soil moisture content and electric conductivity (EC) ranges in sunflower different growth stages.

		Growth stages				
		Sowing	Seedling	Squaring	Flowering	Maturity
Soil moisture content (%)	Lowest	13.6	10.9	11.1	13.5	13.6
	Highest	30.4	28.9	28.2	29.6	27.1
	Average	24.0	20.6	20.3	21.9	20.2
Soil EC ($dS\,m^{-1}$)	Lowest	0.21	0.17	0.18	0.16	0.27
	Highest	2.10	2.09	3.00	1.87	2.10
	Average	0.60	0.57	0.63	0.57	0.65

Table 4
Error changes with the nodes of the hidden layer for ANN-10.

Nodes		8		10		12		14		16		18		20		22		24		26		28	
		Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Training data	RMSE ($t\,ha^{-1}$)	1.4	0.68	0.8	0.25	0.6	0.17	0.3	0.09	0.3	0.08	0.3	0.06	0.3	0.06	0.3	0.06	0.3	0.06	0.3	0.06	0.3	0.06
	RE (%)	15.1	7.10	8.9	2.93	5.6	1.64	3.8	1.10	3.8	0.84	3.8	0.76	3.8	0.76	3.9	0.76	3.9	0.76	3.9	0.76	3.9	0.76
	R^2	0.62	0.15	0.71	0.12	0.83	0.08	0.85	0.08	0.85	0.08	0.85	0.08	0.85	0.08	0.85	0.08	0.85	0.08	0.85	0.08	0.85	0.08
Testing data	RMSE ($t\,ha^{-1}$)	2.4	0.72	1.8	0.44	1.3	0.25	1.1	0.12	1.2	0.11	1.1	0.11	1.3	0.12	1.6	0.14	2.2	0.18	2.4	0.17	2.5	0.19
	RE (%)	27.9	7.97	22.3	5.15	14.9	2.80	12.9	1.34	12.5	1.26	12.0	1.18	14.5	1.38	19.0	1.67	24.3	1.98	26.7	1.89	27.9	2.11
	R^2	0.54	0.17	0.61	0.15	0.70	0.13	0.81	0.10	0.82	0.1	0.84	0.10	0.71	0.11	0.59	0.11	0.61	0.11	0.59	0.12	0.63	0.12

Table 5
Error changes with the nodes of the hidden layer for ANN-6.

Nodes		4		6		8		10		12		14		16		18		20		22	
		Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Training data	RMSE ($t\,ha^{-1}$)	2.8	0.86	1.7	0.33	0.7	0.20	0.4	0.11	0.4	0.08	0.4	0.08	0.4	0.08	0.4	0.08	0.4	0.08	0.4	0.08
	RE (%)	30.3	10.04	18.0	3.68	8.4	2.52	4.9	1.18	4.1	0.82	4.1	0.82	4.1	0.82	4.1	0.82	4.1	0.82	4.1	0.82
	R^2	0.57	0.16	0.70	0.14	0.79	0.13	0.83	0.10	0.83	0.10	0.83	0.10	0.83	0.10	0.83	0.10	0.83	0.10	0.83	0.10
Testing data	RMSE ($t\,ha^{-1}$)	3.2	0.91	2.7	0.38	1.9	0.34	1.6	0.22	1.7	0.13	2.0	0.11	2.4	0.11	2.7	0.12	3.5	0.39	3.5	0.38
	RE (%)	34.3	10.33	28.8	4.17	20.1	3.70	17.3	2.46	19.9	1.47	22.4	1.25	26.5	1.21	30.9	1.34	33.2	3.72	36.2	3.95
	R^2	0.42	0.18	0.60	0.15	0.72	0.15	0.80	0.12	0.75	0.14	0.64	0.14	0.61	0.16	0.58	0.16	0.50	0.15	0.43	0.16

The ranges of inputs variables are listed in Table 3. The experimental data (108 samples), including sunflower yield, soil moisture content and EC of different growth stages, were randomly averagely divided into 4 groups in which three groups were chosen to train ANN and the remainder was used to test ANN. Every ANN was trained four times using four different data sets.

3.1.2. Network architecture

The neurons of the hidden layer in the BP neural network were determined by trial and error. For training and testing results from ANNs, the RMSE, RE, and R^2 with different node numbers of the hidden layers were compared to select the optimal node numbers of the hidden layers (Tables 4 and 5). For the training data of ANN-10, the RMSE and RE were the lowest and R^2 was the highest when the hidden node was equal to 14 or more. However, for the testing data, the ANN-10 with the node of 18 produced the lowest RMSE and RE and the highest R^2, and when the node number was more or less than 18, the RMSE and RE increased and R^2 decreased. For the training data of ANN-6, the ANN model had the lowest RMSE and RE and the highest R^2 when the node number was 10 or more. For the testing data, the model with the hidden node of 10 had a lowest RMSE and RE, and a highest R^2. Fewer hidden units resulted in underfitting due to the shortage of enough processing units to map the input/output fitting relationship. With more neurons (more than 18 and 10 for ANN-10 and ANN-6, respectively) in the hidden layer, the network became overfitted, which showed that it was capable of fitting the training data very well but not to generalize the unknown inputs (i.e., testing data from all data). In addition, more hidden neurons increased the network training time significantly. Therefore, the optimal network structures were 10-18-1 and 6-10-1 for the two ANNs, respectively.

3.1.3. Training result of ANN model

After training of ANNs, testing data were used to determine the errors of model. The RMSE range of ANN-10 was from 0.1–3.2 t ha^{-1}, and the RE was from 2.9% to 32.4%, whereas the ANN-6 produced a bigger RMSE and RE range, which was 0.2–3.6 t ha^{-1} and 5.5–38.9% for RMSE and RE, respectively (Table 6). The average RMSE, RE, and R^2 were 1.1 t ha^{-1}, 12.0%, and 0.84 for ANN-10, and 1.6 t ha^{-1}, 17.3% and 0.80 for ANN-6, respectively. Comparably, the ANN-10 had a lower RMSE and RE, and a larger R^2 than ANN-6 (Fig. 3 and Table 6).

To evaluate the effect of ANNs, two MLRs (10 and 6 input variables) were developed and tested using the same set of data used for the ANNs developing and calibrating. Similar to the ANNs, the MLR with 10 input variables had a lower RMSE and RE, and a larger R^2 than MLR with 6 input variables (Fig. 3 and Table 6). But the RMSE and RE of MLR-10 were higher than those of ANN-10, and R^2 was lower than that of ANN-10 and MLR-6 had also a higher RMSE and RE and a lower R^2 compared to ANN-6 (Table 6). Even though for the ANN with 6 input variables, it also performed better than MLR with 10 input variables. As a whole, the calibrated ANNs had higher accuracy to simulate the response of sunflower yield to soil moisture and salinity.

3.2. Relative importance of input factors in ANNs

In ANN-10 and ANN-6, the relative importance (RI) of soil moisture and EC at different growth stages to sunflower yield is listed in Table 7. For ANN-10, soil moisture content at squaring and soil EC at seedling had the highest RI among all input variables, which accounted for 18%. Secondly was the soil moisture content at the flowering and maturity stages had a RI of 15% and 10%, respectively, and soil EC at the squaring stage had a RI of 12%. The RI of soil moisture content at sowing and seedling stages and of EC at sowing, flowering and maturity stages was relatively lowest among all input variables. The results indicate that sunflower was more sensitive to soil moisture at the squaring stage and soil salinity at the seedling stage, and with crop growth, its resistance increased and the importance of soil moisture and EC to yield reduced.

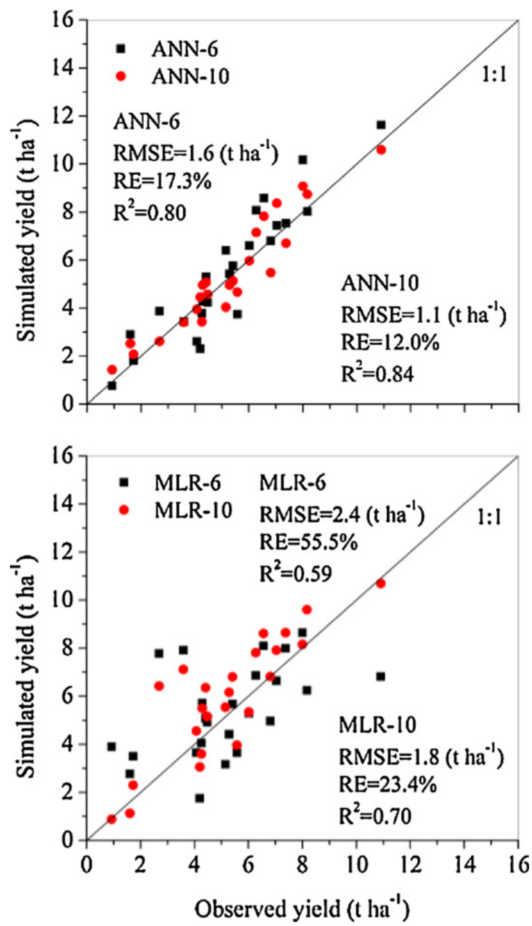

Fig. 3. Observed versus simulated yield from ANN-10 and ANN-6, and MLR-10 and MLR-6.

For ANN-6, soil EC at crop sowing stage had a highest RI (32%), indicating that the input variable is the most important for yield formation (Table 7). Similar to ANN-10, the RI of soil moisture content at crop different growth stages to yield in ANN-6 was arranged in a sequence of squaring, flowering, maturity, seedling and sowing stage. This also proved that soil moisture content at sunflower squaring stage is the most important for grain yield and the RI of soil moisture to yield decreased at crop vegetative growth stages.

3.3. Simulation of sunflower yield using ANN-6

3.3.1. Yield change with soil salinity at crop sowing stage

One ANN-6 was used to simulate sunflower yield change with soil salinity increase of sowing stage and soil moisture content was respectively 100% or 40% of field capacity at every crop growth stage (Fig. 4). The initial soil EC increased from 0.1 to 2 dS m^{-1} with 0.1 dS m^{-1} increments. The simulated results showed that sunflower yield decreased with the initial soil EC increase regardless of soil moisture content. At 100% field capacity, the sunflower yield decreased sharply from 12.0 to 7.0 t ha^{-1} when the initial EC increased from 0.2 to 0.6 dS m^{-1}. Afterward, the sunflower yield declined relatively slowly from 7.0 to 1.8 t ha^{-1} when initial EC increased from 0.6 to 2.0 dS m^{-1}. The simulated yield at 100% field capacity was mostly higher than the observed yield. At 40% field capacity, sunflower yield decreased from 3.5 to 0.3 t ha^{-1} with the soil salinity increase, all of which was lower than the observed yield.

Table 6
Error comparison of the ANN and MLR models.

		10 inputs				6 inputs			
		ANN-10		MLR-10		ANN-6		MLR-6	
		Mean	Std	Mean	Std	Mean	Std	Mean	Std
RMSE ($t\,ha^{-1}$)	Max	3.2	0.18	4.9	0.26	3.6	0.34	6.6	0.37
	Min	0.1	0.02	0.2	0.04	0.2	0.05	0.2	0.04
	Average	1.1	0.11	1.8	0.19	1.6	0.22	2.4	0.21
RE (%)	Max	32.4	5.24	120.5	19.30	38.9	6.31	331.2	45.20
	Min	2.9	0.64	9.6	1.87	5.5	0.84	23.6	3.18
	Average	12.0	1.18	23.4	6.23	17.3	2.46	55.5	9.77
R^2		0.84	0.10	0.70	0.09	0.80	0.12	0.59	0.13

Table 7
Relative importance (RI, %) of the soil moisture and salinity factors to crop yield estimation.

	Crop growth stage	ANN-10	ANN-6
Soil moisture content	Sowing	6	6
	Seedling	7	10
	Squaring	18	24
	Flowering	15	16
	Maturing	10	12
Soil electric conductivity	Sowing	4	32
	Seedling	18	–
	Squaring	12	–
	Flowering	5	–
	Maturing	5	–

–, express no value because the corresponding factor is not input variable of ANN-6.

The results showed that high salt concentrations severely affected the growth of sunflower, but high soil moisture content mitigated the crop yield decrease resulted from soil salinity stress.

3.3.2. Yield change with soil moisture at crop different growth stages

Sunflower yields were simulated with soil moisture content increase respectively at crop seedling, squaring or flowering growth stages using ANN-6, whereas at other growth stages, the soil moisture content was 26%, which was equal to field capacity. Three soil salinity levels of sowing stage (0.3, 0.6 and 1.0 $dS\,m^{-1}$) were chosen (Fig. 5). The results showed that the sunflower yields increased with soil moisture content increase for the different initial soil salinities. When soil moisture content at crop seedling, squaring and flowering stages increased from 12% to 26%, sunflower yield improved from 7.5 to 9.9 $t\,ha^{-1}$ for low saline soil, 2.2–6.8 $t\,ha^{-1}$ for medium saline soil and 1.9–4.8 $t\,ha^{-1}$ for high saline soil. The high initial soil salinity inhibited crop yield formation, but the improvement in soil moisture could decrease the reduction in the yield resulted from high salinity.

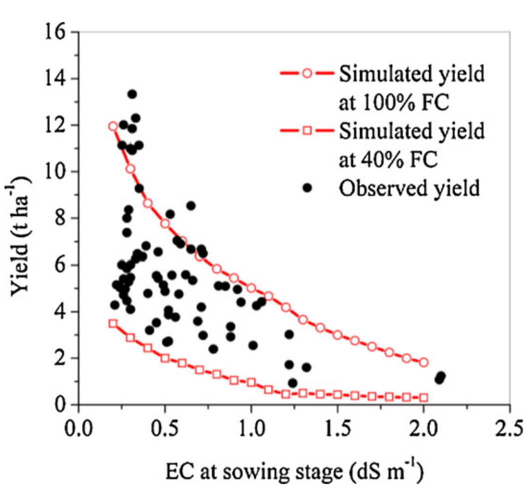

Fig. 4. Observed and ANN-6 simulated yield response to soil salinity of crop sowing stage under 100% and 40% field capacity.

For low (0.3 $dS\,m^{-1}$) and medium (0.6 $dS\,m^{-1}$) saline soils, the sunflower yield was observed to be most sensitive to soil moisture stress at the squaring stage, and the yield decreased by 24% and 68% at soil moisture of 12% compared to those at soil moisture of 26% for low and medium soil salinity, respectively. Secondly, sunflower yield was sensitive to water stress at flowering stage and decreased by 19% and 56% at soil moisture of 12% compared to at soil moisture of 26% for the low and medium soil salinity, respectively. The sunflower yield at the seedling stage was the most insensitive compared to those at the other two growth stages, and the yield at the condition of soil moisture of 12% reduced by 10% and 46% compared to that at soil moisture of 26% for low and medium initial soil salinity.

For high saline soil (1.0 $dS\,m^{-1}$), the response of sunflower yield to soil moisture stress was different from those in the low and medium saline soils (Fig. 5). Sunflower yield was the most sensitive to soil moisture stress at the seedling stage for high initial saline soil. At crop seedling stage, sunflower yield decreased by 60% at soil moisture of 12% compared to at soil moisture of 26%. Relatively, the sunflower yield was insensitive to the soil water content at the squaring and flowering stages, in which the yield reduced by 42% and 30%, respectively. Overall, the effect of soil water and salt stress on sunflower yield is complex and water management needs to be adjusted according to the response of sunflower growth to water stress in the different soil salinities.

4. Discussions

Many previous studies have already shown that the crop response to soil environment is very complex, and should be modeled as cubic or quadratic functions (Kijne, 2003; Jalota et al., 2006; Starr et al., 2008). In our study, the ANN model produced

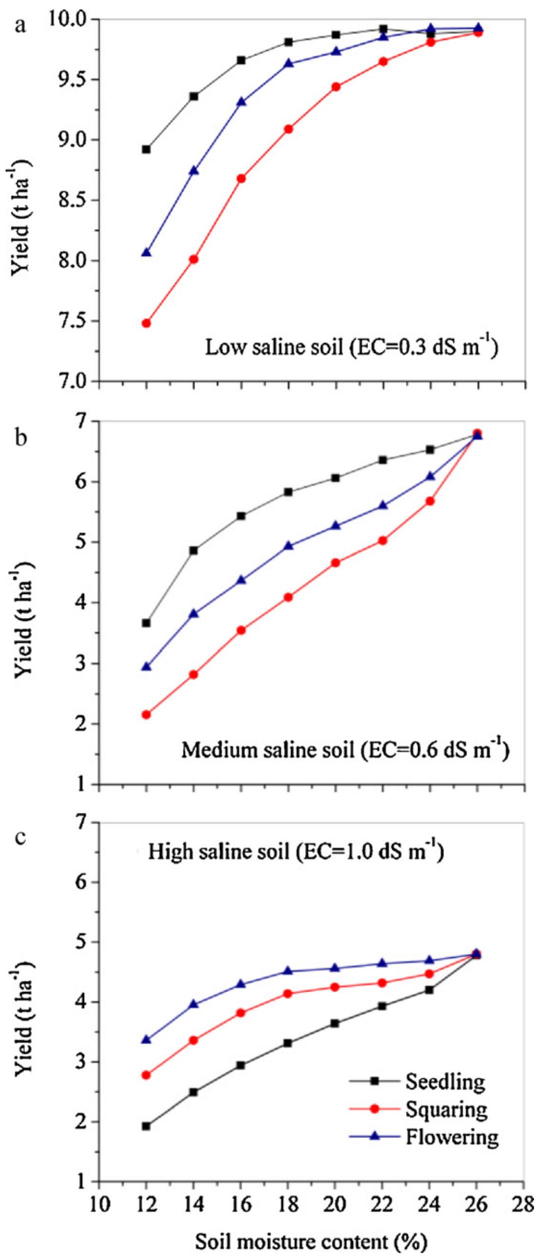

Fig. 5. ANN-6 simulated yield response to soil moisture content at crop different growth stages under low, medium and high saline soil conditions.

a more precise and accurate result than MLR for modeling sunflower responses to soil moisture and salinity. Even for the ANN with 6 input variables, the results also were superior to those of the MLR with 10 input variables. The result was similar to the findings of Kaul et al. (2005) and Miao et al. (2006). Kaul et al. (2005) used both soil productivity rates and climate variables for yield prediction and found that ANN had shown to be better tools than regression methods when analyzing corn and soybean yield data generated in field trials. Miao et al. (2006) employed ANN analysis to evaluate the relative importance of selected soil, landscape and seed hybrid factors on corn yield and grain quality in two Illinois, USA fields, and the results indicated that the response curves generated by the ANN models were more informative than simple correlation coefficients or coefficients in multiple regression equation. The performance of ANN in the study was mainly attributed to the ability of ANNs to capture the nonlinear input–output relationship between crop growth and soil moisture and salinity, whereas MLRs were unable to reflect these complicated relationships due to their linear characteristics. Batchelor et al. (1997) showed that the ANN had the advantage over other empirical modeling techniques that do not assume a priory structure for the data, are well suited for fitting non-linear relationships and complex interactions, and can expose hidden relationships among input variables.

Moreover, ANN-10 in this study performed better than ANN-6, but in practice ANN-6 is more convenient to use in saline soil due to fewer input variables. Because it is impossible to exactly know the soil salinity of crop different growth stages at planting, the soil EC at crop sowing stage could easily be acquired, which was important in tutoring water management in saline soils. According to the different initial soil salinities, it is feasible to adjust soil moisture content at crop different growth stages for acquiring high yield. Dai et al. (2009) also modeled the reference evapotranspiration using three or four climate factors and indicated that the error of ANN model increased with the decrease in the number of input variable. Considering the simplicity and practicality of model, the ANN-6 was more perfect for simulating the sunflower response to soil moisture and salinity.

Using connection weight method, the study results for ANN-10 and ANN-6 indicated that sunflower was the most sensitive to soil moisture at the squaring and flowering stage. For low (0.3 dS m^{-1}) and medium (0.6 dS m^{-1}) soil salinity at crop sowing stage, the response of sunflower yields to water stress also were the most sensitive at squaring stage and secondly was at flowering stage. This is because sunflower was at its reproductive growth phase, when water and salt stress significantly influence its achene differentiation and filling. Flagella et al. (2004) and Di Caterina et al. (2007) had showed that sunflower yield reduction was attributed mainly to a decrease in achene per head and in the 1000 achene weight. But the sensitivity of sunflower yield to soil moisture content varied for the different soil salinities. For high (1.0 dS m^{-1}) soil salinity at crop sowing stage, the yield response was the most obvious at seedling stage and the relative importance of soil EC for ANN-10 also indicated that was the most important at seedling stage, which mainly was owing to severe salt stress resulting in poor emergence and vegetation growth (Chen et al., 2009). High soil moisture could compensate some yield reduction resulting from salt stress. The modeling also proved that increasing the amount of irrigation resulted in sunflower yield improvement reaching 86.9% under high soil salinity level (Gaballah et al., 2006). In the study, the simulated yield at 100% field capacity was higher than that at 40% field capacity, and mostly higher than the observed yield. The soil moisture content at the initial planting and the four growth stages was at the field capacity of $\theta_f = 26\%$, which means that the soil moisture conditions were optimized and there was no water stress to the sunflower. However, several data from the observed yield were higher than the simulated yield mainly due to ANN had higher error for marginal data from training sample. The simulated yield at 40% field capacity was lower than the observed yield owing to the crop yield was simulated at the condition of water stress at sunflower every growth stage. In fact, when the crop was influenced by the dry soil condition, irrigation schemes were generally conducted by local farmers. Accurate knowledge of the relative importance of the input parameters in ANN models to producing output would be useful in guiding irrigation practices in saline soils.

However, in this study, only the relationship between sunflower yield and the average soil water content and EC of soil profile at different growth stages was investigated. However, the crop growth response to soil moisture and salinity are more complex. The soil moisture and salinity dynamics, as well as their distribution at soil profile would change with crop growth and irrigation management. In future, ANN technology ought to be used to model the relationship between crop growth and soil water and salinity of soil profile, which was important for more accurately understanding and describing the yield response, and further better guiding water management in saline conditions.

5. Conclusions

Two ANN models, with 6 inputs (ANN-6) and with 10 inputs (ANN-10), were developed to simulate the response of sunflower yield to soil water content and salinity in saline soil. Compared with MLRs, both ANN models have higher precision with RMSEs of 1.1 and 1.6 t ha^{-1}, REs of 12.0% and 17.3%, and R^2 of 0.84 and 0.80, for ANN-10 and ANN-6, respectively, indicating that both ANN models can accurately describe the complex relationship between sunflower yield and soil moisture and salinity at crop different growth stages. The response of sunflower yield to soil water and salt stress can be understood through the simulation results of the ANN models. For low and medium saline soils, sunflower yield was more sensitive at crop squaring stage, but for high saline soil at seedling stage. High soil moisture content could compensate the yield decrease resulting from salt stress regardless of salt levels at crop sowing stage. However, the squaring stage the most sensitive was, and the flowering stage secondly. The response of sunflower yield to soil moisture at different stages in saline soils can be understood through the simulated results of ANN-6. Therefore, an ANN model is a useful tool in investigating and understanding the complex relationships between crop yields and soil water and salinity.

Acknowledgements

This study was supported by the National Key Technology R&D Program of China (No. 2011BAD25B05).

References

Alvarez, R., 2009. Predicting average regional yield and production of wheat in the Argentine Pampas by an artificial neural network approach. Eur. J. Agron. 30, 70–77.

Ashish, D.P., Hina, B., Amar, N.P., 2009. Effect of salinisation of soil on growth, water status and general nutrient accumulation in seedlings of Delonix regia (*Fabaceae*). Acta Ecol. Sin. 29, 109–115.

Bassil, E.S., Kaffka, S.R., 2002. Response of safflower (*Carthamus tinctorius* L.) to saline soil and irrigation II. Crop response to salinity. Agric. Water Manage. 54, 81–92.

Batchelor, W.D., Yang, X.B., Tshanz, A.T., 1997. Development of a neural network for soybean rust epidemics. Trans. ASAE 40, 247–252.

Chen, M., Kang, Y., Wan, S., Liu, S., 2009. Drip irrigation with saline water for oleic sunflower (*Helianthus annuus* L.). Agric. Water Manage. 96, 1766–1772.

Clermont-Dauphina, C., Suwannangb, N., Grunbergerc, O., Hammeckera, C., Maeght, J.L., 2010. Yield of rice under water and soil salinity risks in farmers' fields in northeast Thailand. Field Crops Res. 118, 289–296.

Dai, X., Shi, H., Li, Y., Ouyang, Z., Huo, Z., 2009. Artificial neural network models for estimating regional reference evapotranspiration based on climate factors. Hydrol. Process. 23, 442–450.

Di Caterina, R., Miuliani, M.M., Rotunno, T., De Caro, A., Flagella, Z., 2007. Influence of salt stress on seed yield and oil quality of two sunflower hybrids. Ann. Appl. Biol. 151, 145–154.

Feng, Z., Wang, X., Feng, Z., 2005. Soil N and salinity leaching after the autumn irrigation and its impact on groundwater in Hetao Irrigation District, China. Agric. Water Manage. 71, 131–143.

Flagella, Z., Giuliani, M.M., Rotunno, T., Di Caterina, R., De Caro, A., 2004. Effect of saline water on oil yield and quality of a high oleic sunflower (*Helianthus annuus* L.) hybrid. Eur. J. Agron. 21, 267–272.

Fortin, J.G., Anctil, F., Parent, L., Bolinder, M.A., 2010. A neural network experiment on the site-specific simulation of potato tuber growth in Eastern Canada. Comput. Electron. Agric. 73, 126–132.

Gaballah, M.S., Ouda, S.A., Mandour, M.S., 2006. Predicting the role of antioxidants and irrigation on sunflower yield grown under saline conditions. In: International Conference on Environmentally Sound Technology in Water Resources Management , September 11–13. Univ. Botswana, Gaborone.

Green, T.R., Salas, J.D., Martinez, A., Erskine, R.H., 2007. Relating crop yield to topographic attributes using spatial analysis neural networks and regression. Geoderma 139, 23–37.

Harris, B.N., Sadras, V.O., Tester, M., 2010. A water-centred framework to assess the effects of salinity on the growth and yield of wheat and barley. Plant Soil, doi:10 1007/s11101-010r-r0489-9.

Haykin, S., 1998. Neural Networks-A Comprehensive Foundation, 2nd ed. Prentice-Hall, Upper Saddle River, NJ.

Huang, Y., Lan, Y., Thomson, S.J., Fang, A., Hoffmann, W.C., Lacey, R.E., 2010. Development of soft computing and applications in agricultural and biological engineering. Comput. Electron. Agric. 71, 107–127.

Jain, S.K., Nayak, P.C., Sudheer, K.P., 2008. Models for estimating evapotranspiration using artificial neural network, and their physical interpretation. Hydrol. Process. 22, 2225–2234.

Jalota, S.K., Sood, A., Harman, W.L., 2006. Assessing the response of chickpea (*Cicer aeritinum* L.) yield to irrigation water on two soils in Punjab (India): a simulation analysis using the CROPMAN model. Agric. Water Manage. 79 (3), 312–320.

James, Y.W., Cutforth, H.W., 1996. Crop growth models for decision support systems. Can. J. Plant. 76, 9–19.

Katerji, N., van Hoorn, J.W., Hamdy, A., Mastrorilli, M., 2003. Salinity effect on crop development and yield, analysis of salt tolerance according to several classification methods. Agric. Water Manage. 62, 37–66.

Kaul, M., Hill, R.L., Walthall, C., 2005. Artificial neural networks for corn and soybean yield prediction. Agric. Syst. 85, 1–18.

Kemp, S.J., Zaradic, P., Hansen, F., 2007. An approach for determining relative input parameter importance and significance in artificial neural networks. Ecol. Model. 204, 326–334.

Khazaei, J., Naghavi, M.R., Jahansouz, M.R., Salimi-Khorshidi, G., 2008. Yield estimation and clustering of chickpea genotypes using soft computing techniques. Agron. J. 100, 1077–1087.

Kijne, J.W., 2003. Water productivity under saline conditions. In: Kijne, J.W., Barker, R., Molden, D. (Eds.), Water Productivity in Agriculture: Limits and Opportunities. CAB International, pp. 89–102.

Liu, X., Kang, S., Li, F., 2009. Simulation of artificial neural network model for trunk sap flow of Pyrus pyrifolia and its comparison with multiple-linear regression. Agric. Water Manage. 96, 939–945.

Lobell, D.B., Ortiz-Monsterio, J.I., Gurrola, F.C., Valenzuuela, L., 2007. Identification of saline soils with multiyear remote sensing of crop yields. Soil Sci. Soc. Am. J. 71, 777–783.

Ma, L., Hoogenboom, G., Saseendran, S.A., Bartling, P.N.S., Ahuja, L.R., Green, T.R., 2009. Effects of estimating soil hydraulic properties and root growth factor on soil water balance and crop production. Agron. J. 101 (3), 572–583.

Matthews, R., 2002a. Crop management. In: Matthews, R., Stephens, W. (Eds.), Crop-soil Simulation Models. CAB International, pp. 69–79.

Matthews, R., 2002b. Regional and national planning. In: Matthews, R., Stephens, W. (Eds.), Crop-soil Simulation Models. CAB International, pp. 29–52.

Meloni, D.A., Oliva, M.A., Martinez, C.A., Cambraia, J., 2003. Photosynthesis and activity of superoxide dismutase, peroxidase and glutathione reductase in cotton under salt stress. Environ. Exp. Bot. 49, 69–76.

Miao, Y., Mulla, D.J., Robert, P.C., 2006. Identifying important factors influencing corn yield and grain quality variability using artificial neural networks. Precision Agric. 7 (2), 117–135.

Olden, J.D., Joy, M.K., Death, R.G., 2004. An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. Ecol. Model. 178, 389–397.

Park, S.J., Hwang, C.S., Vlek, P.L.G., 2005. Comparison of adaptive techniques to predict crop yield response under varying soil and land management conditions. Agric. Syst. 85, 59–81.

Sinclair, T.R., Seligman, N.A.G., 1996. Crop modelling: from infancy to maturity. Agron. J. 88, 698–703.

Starr, G.C., Rowland, D., Griffin, T.S., Olanya, O.M., 2008. Soil water in relation to irrigation, water uptake and potato yield in a humid climate. Agric. Water Manage. 95 (3), 292–300.

Zhang, J.Q., Zhang, L.X., Zhang, M.H., Watson, C., 2009. Prediction of soybean growth and development using artificial neural network and statistical models. Acta Agron. Sin. 35 (2), 341–347.

Ziaei, A.N., Sepaskhah, A.R., 2003. Model for simulation of winter wheat yield under dry land and irrigated conditions. Agric. Water Manage. 58, 1–17.

Zou, P., Yang, J., Fu, J., Liu, G., Li, D., 2010. Artificial neural network and time series models for predicting soil salt and water content. Agric. Water Manage., doi:10.1016/j.agwat.2010.02.011.