available at www.sciencedirect.com

SCIENCE DIRECT®

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

Subsurface drainage performance study using SALTMOD and ANN models

A. Sarangi*, Man Singh, A.K. Bhattacharya, A.K. Singh

Water Technology Centre, Indian Agricultural Research Institute, New Delhi, India

ARTICLE INFO

Article history:
Accepted 23 February 2006
Published on line 18 April 2006

Keywords:
Artificial neural network
Subsurface drainage
MARS
SALTMOD
Drainage effluent salinity
Root zone soil salinity

ABSTRACT

Relative performance of artificial neural networks (ANNs) and the conceptual model SALTMOD was studied in simulating subsurface drainage effluent and root zone soil salinity in the coastal rice fields of Andhra Pradesh, India. Three ANN models viz. Back Propagation Neural Network (BPNN), General Regression Neural Network (GRNN) and Radial Basis Function Neural Network (RBFNN) were developed for this purpose. Both the ANNs and the SALTMOD were calibrated and validated using the field data of 1998–2001 for 35 and 55 m drain spacing areas. Data on irrigation depth, evapotranspiration, drain discharges, water table depths, mean monthly rainfall and temperature and drainage effluent salinity were used for ANN model training, testing and validation. It was observed that the BPNN model with feed forward learning rule with 6 processing elements in input layer and 1 hidden layer with 12 processing elements performed better than the other ANN models in predicting the root zone soil salinity and drainage effluent salinity. Considering coefficient of determination, model efficiency and variation between the observed and predicted salinity values as the evaluation parameters, the SALTMOD performed better in predicting root zone soil salinity and the BPNN performed better in predicting the drainage effluent salinity. Therefore, it was concluded that the BPNN with feed forward learning algorithm was a better model than SALTMOD in predicting salinity of drainage effluent from salt affected subsurface drained rice fields.

© 2006 Elsevier B.V. All rights reserved.

1. Introduction

The estimate of waterlogged and saline lands in India reported by different sources varies from 4.75 to 16 million ha and 3.3 to 10.9 million ha, respectively (Bhattacharya, 1999). Also, high water table areas cover about 2.6 and 3.4 million ha suffers from surface water stagnation (Tyagi et al., 1993). There is a need to reclaim the saline lands for bringing them under production process and in achieving agricultural sustainability. In this context, application of advanced tools in modeling the salinity of these problematic areas would help in future projections and in chalking out suitable remediation measures and best management practices. Coastal agricultural lands often face the twin problems of waterlogging and salinity for which, subsurface drainage is an appropriate and proven solution (Singh et al., 2002a). In the coastal agricultural lands of southern India, two crops of paddy rice, namely, *rabi* (January to May) and *kharif* (July to November) are usually taken. Both suffer due to either of the two problems of excessive soil salinity and waterlogging. In a small but representative coastal rice land in Andhra Pradesh, India, subsurface drainage systems were installed at different spacings during the mid-eighties. The present study is based on the analysis of various data pertaining to the installed systems.

* Corresponding author. Tel.: 9811400885 (mobile).
E-mail addresses: arjamadutta.sarangi@elf.mcgill.ca, asarangi@iari.res.in (A. Sarangi).
0378-3774/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.agwat.2006.02.009

Singh et al. (2002a) simulated the performances of subsurface drained fields, using SALTMOD (Oosterbaan, 1998) for computing soil salinity in the root zone soil and of drainage effluent, drain flow rates, water table and several water balance components for different water management options. SALTMOD, a seasonal water and salt balance model for agricultural fields has been successfully calibrated and validated in Nile delta of Egypt (Oosterbaan and Abu Senna, 1989) and in Tungabhadra Irrigation Project, Karnataka, India. SALTMOD requires input parameters viz. soil properties, climatic and hydrologic parameters, drainage system parameters and defined initial and boundary conditions for model operation (Singh et al., 2002b). Soft computing tools were used in this study for prediction of salinity of the drainage effluent and root zone depth soil of the coastal rice fields in Andhra Pradesh and compare the results with those obtained using SALTMOD. The multivariate adaptive regression splines (MARS) tool (Friedman, 1991) was used to select parameters that are relatively more important in influencing drainage effluent salinity and these were used in the artificial neural networks (ANNs) models to simulate the same and compare the predictive performance of ANN models with the SALTMOD.

Artificial neural networks are data processing systems comprising a large number of simple, highly interconnected processing elements (artificial neurons) in an architecture inspired by the structure of the cerebral cortex (Tsoukalas and Uherg, 1996). In a study to investigate the applicability of ANNs in subsurface drained fields located in St-Dominique, Quebec, Canada, Yang et al. (1996) used back propagation neural network (BPNN) to imitate DRAINMOD in simulating water table depths. They observed that, use of time lag procedure in feeding the input values to the ANN resulted in better ANN performance. They advocated considering the saturated hydraulic conductivity and actual distance from the soil surface to impermeable layer while considering the optimal time dependency period as an input to the ANN. Salehi et al. (2000) used BPNN with one hidden layer having eight processing elements and different learning rules to predict the annual nitrate-nitrogen losses via the drain outflow at Agriculture Canada's Woodslee Research Station, Ontario. The BPNN could effectively predict the loss of nitrate-nitrogen via drain outflows, but the model itself is not transportable to any other site. However, ANN can be used as an identifying tool to discard unnecessary parameters used for modeling to save time and resources in data collection. Sharma et al. (2003) developed two ANN models viz. a fast back propagation neural network model (FBPNN) and a self-organizing radial basis function neural network model (RBFNN) to simulate the subsurface drain outflow and nitrate-nitrogen concentration in the tile drainage effluent from Greenbelt Research Farm, Ontario, Canada. In their study the performance of RBFNN was superior.

It has been demonstrated that neural networks are a competitive alternative to traditional classifiers for many practical classification problems (Zhang, 2000). ANNs have the advantage over deterministic models, as the ANNs require lesser data and are capable of long term forecasting. The disadvantages of ANN are that it is based on a 'black box' approach and the solution is obtained through a trial and error process (Sharma et al., 2003). However, inclusion of system parameters as processing elements (PEs) or as mathematical association with the PEs in the input layer will reorient the ANNs from a complete 'black box' to a 'gray box' approximation (Sarangi and Bhattacharya, 2005). There has been a growing trend of application of ANNs in hydrology and water quality modeling (ASCE, 2000; Sudheer et al., 2002; Yitian and Gu, 2003; Zhang and Govindaraju, 2003) and land drainage engineering (Shukla et al., 1996; Yang et al., 1996). The MARS can also be used as a tool to identify redundant parameters in predictive analysis and select the sensitive parameters for consideration in model development (Abraham and Steinberg, 2001). The use of MARS tool in selection of the sensitive input parameters and development of ANN models to predict the salinity of drainage effluent in this study is a new approach for simulating the complex salt dynamics in artificially drained soils.

2. Materials and methods

2.1. Data acquisition

The data for the study were acquired from the subsurface drained rice fields of *Endakuduru* village in *Krishna* district of Andhra Pradesh in India. It is located within 15°43′ and 17°10′N latitude and 80°0′ and 81°33′E longitude, situated 18 km to the west of Bay of Bengal at an elevation of 1.5 m amsl. The land is flat and is dyked in small units for rice cultivation. The groundwater is shallow and highly saline due to seawater intrusion. The site experiences a moderate coastal climate with mean annual rainfall of 975 mm and mean annual maximum and minimum temperatures of 36.6 and 19.3 °C, respectively. The subsurface drainage system layout consisted of five laterals with three laterals at 35 m spacing and two laterals at 55 m spacing covering the experimental site of 4 ha. The average depth of the lateral drains was 1 m. The climatological data used for both the ANN and SALTMOD models are given in Table 1. Further details about the cropping pattern, data collection and ranges of input data values for SALTMOD were adopted from Singh et al. (2002a). The drainage water quality and quantity, depth to water table and root zone salinity at different depths were measured fortnightly during the *rabi* rice season of each year for a period of 4 years (from year 1998 to 2001) resulting in 40 measured data sets. The *rabi* rice crop season was chosen for field data collection, as most farmers grow rice in this cropping season due to assured canal water availability for irrigation.

2.2. Multivariate adaptive regression spline (MARS) for sensitivity analysis

The MARS 2.0 software (Friedman, 1991) was used to estimate the relative significance of the available data in predicting the salinity of drainage effluent. The concepts used in MARS tool permits the user to analyze the data set and generate the input parameters, which are of more significance in generation of the desired output and display the empirical model of best fit. Finally, the validation input data set is entered into the model and the output is compared with the observed values to decide on the future model applicability.

Table 1 – Climatological parameter of the experimental site used in model comparisons

Year	Parameters	January	Febuary	March	April	May	June	July	August	September	October	November	December
1997	Mean monthly temperature	23.8	26.05	28.2	29.85	33.05	33.9	30.75	30.4	29.35	28.25	27.25	26.35
	Rainfall (mm)	9.4	0.0	9.6	7.7	1.2	56.9	247.9	135.0	491.6	125.1	194.1	87.1
1998	Mean monthly temperature	26.2	27.4	29.05	30.75	33.95	33.75	30	29.5	29.5	28.3	27.35	24.4
	Rainfall (mm)	19.6	0.0	0.4	19.8	0.2	78.2	162.5	181.5	216.5	311.1	57.5	0.0
Normal	Rainfall (mm)	7.9	8.3	7.6	4.2	28.1	86.8	169.7	182.0	166.6	153.9	140.7	19.2
	PET[a] (mm)	109	122	166	176	193	167	134	136	123	118	108	102
Daily	PET[a] (mm)	3.5	4.2	5.3	5.8	6.2	5.5	4.3	4.4	4.1	3.8	3.6	3.3

[a] PET: potential evapotranspiration.

In this study, the data of rainfall (mean monthly and normal), temperature (maximum, minimum and average), mean daily potential-evapotranspiration (PET), water table depths, quantity and salinity of irrigation water, saturated hydraulic conductivity and drainable porosity and leaching efficiency under different zones were used as input. Root zone salinity for 0–30 cm and 30–60 cm soil layers and salinity of drained effluent were used as output parameters. After running the MARS tool with the data sets of 35 and 55 m spaced subsurface drained fields, importance-wise, the input parameters: salinity of irrigation water (IWS), quantity of irrigation water (Q_{Irrig}), mean monthly rainfall (MMR), mean daily potential-evapotranspiration (PET), water table depths (WTD) and the average mean monthly temperature (AMT) topped the list. The rest of the parameters were towards the bottom in the list of relative importance. Therefore, the input data of IWS, Q_{Irrig}, MMR, WTD, PET and AMT were used as the input parameters for ANN models against the output parameters of root zone and drainage effluent salinity.

In general, the ANN models are operated by using the available input and output responses without considering the inherent system parameters. Therefore, in an attempt to elevate the complete ‘‘black box’’ approximation of the ANN models, in this study, the MARS tool was used to consider most of the system response parameters as detailed in SALTMOD and the relatively important parameters obtained were used as input nodes (PEs) to selected ANN models. So the completely ‘‘black box’’ nature of prediction through ANN models were minimized by inclusion of the MARS derived system response parameters as discussed in this section to simulate the performance of subsurface drained rice field of the experimental site. Therefore, the inclusion of MARS generated parameters as PEs of the input layer was an effort towards elevating the ANNs from ‘‘black box’’ modeling approach to a ‘‘gray box’’ approximation.

2.3. *Artificial neural network (ANN) model architecture selection and simulation*

The BPNN, GRNN and RBFNN with different architectures were used in the present study. These ANNs were chosen due to their variant structure and interpretation in multidimensional spaces with different mathematical foundations for solving ill-conditioned problems. Also, it was revealed from the literature that these three ANNs were adopted by the researchers in modeling the drainage effluent quality and quantity under sub surface drained agricultural fields (Yu et al., 2004; Sharma et al., 2003; Salehi et al., 2000; Shukla et al., 1996; Yang et al., 1996). Due to smaller data size, these three ANNs were tested to ascertain which was better suited over the others to accurately predict the subsurface drainage performances in drained rice fields. A major advantage of ANN is that it can be used for smaller data sets without any fixed rules for developing its architecture (Sudheer et al., 2002). So, in this study, different ANN architectures were attempted with the experimental data through trial and error approach to improve the prediction accuracy. The Neural Works Professional II/PLUS 5.23v and the ANN module of MATLAB 6.5v software were used to develop the ANN models with optimal architecture and to perform subsequent validation.

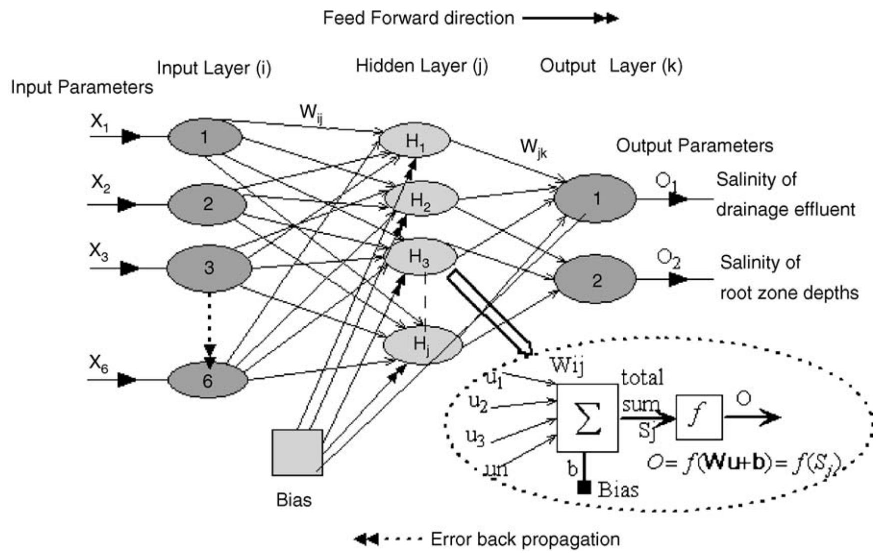

Fig. 1 – The architecture of ANN models used for modeling the salinity of subsurface drained rice fields.

2.3.1. *The back propagation neural network (BPNN)*

Among different ANN algorithms, feed-forward network with the BP training is widely used and is capable of recognizing the nonlinear pattern and memory association (Zhang and Govindaraju, 2003). BPNN is a multilayer perceptron network in which each neuron is connected with a number of input arcs (u_1 to u_n). The network is associated with each neuron (i) having weight W_{ij}, which represents a multiplication factor to a value passing to the neuron. Finally, a neuron sums the values of all inputs and represented as:

$$S_j = \sum_{i=1}^{n} W_{ij} Ui + b \tag{1}$$

In Fig. 1, Wu corresponds to the summation term used in Eq. (1). The term b is called bias. Finally, an activation function is applied to S_j for obtaining the final output from the neuron. When a BPNN training algorithm is used, the sigmoid activation function is most often preferred (Sivakumar et al., 2002). The sigmoid function (φ) is given by

$$\varphi(S_j) = \frac{1}{1 + e^{S_j}} \tag{2}$$

In the present study, the BPNN had one input layer, hidden layer and output layer each. Each neuron in the input layer is connected to each neuron in the hidden layer by weight W_{ij}. After an input neuron receives a signal Ui, it transmits it to the hidden neuron. Each hidden neuron then computes the sum UiW_{ij} entering from each input neuron, and transforms this value to an output signal using sigmoid function.

The ANN input layer in this study consisted of six processing elements (IWS, Q_{Irrig}, MMR, WTD, PET and AMT) and the processing elements (PEs) of the hidden layer were finalized by comparing the root mean square error (RMSE) of the network learning using different numbers of PEs. The output layer consisted of two processing elements representing the salinity of the soil at the root zone depth and of the drainage effluent. The entire data was divided to learning, testing and validation sets and the learning and testing data were fed to the BPNN to select the optimal architecture based on the RMSE values. During the process of learning, the actual output value was compared with the desired output and the error was calculated. The error values were then propagated back into the network to update connection weights between the different layers. These processes were repeated until the network has been trained to the lowest RMSE.

2.3.2. *Radial basis function neural network (RBFNN)*

Like the back-propagation network, the RBFNN has a feed-forward architecture, which consists of three layers viz. one input layer, one hidden layer and one output layer with a number of PEs or nodes in each layer. This ANN derive its structure and interpretation from the theory of interpolation in multidimensional spaces and have a mathematical function embedded in regularization theory for solving ill-conditioned problems (Zhang and Kushwaha, 1999). However, the structure of an RBFNN is one of self-organized characteristics, which allows for adaptive determination of the hidden neurons during training of the network (Sharma et al., 2003). Each input PE or neuron is completely connected to all hidden neurons, and hidden neurons and output neurons are also interconnected to each other by a set of weights. Information fed into the network through input neurons is transmitted to hidden neurons. Each hidden neuron then transforms the input signal using a transfer function *f*. The output of hidden neurons has the form of an RBFNN. For the present model, the Gaussian function was selected as the RBFNN. It is a positive radial symmetric function (kernel) with a center μ and a spread σ. The spread is the radial distance from the center of the kernel, within which the value of the function is significantly different from zero. This is called the receptive field ($\mu \pm \sigma$) of a hidden neuron. An input pattern falling within the receptive field will cause a significant response. For each input pattern, the hidden neurons compute the distance between the input signal and the center of the receiving field. For Gaussian function, the response is unity if this distance is zero, and decays to zero when the distance is greater than the spread. The basic difference between BPNN and RBFNN is that the latter model represents the inputs presented to the network during training phase in local spaces with each local space being represented by a hidden neuron. Therefore, any input to the model in the testing phase that lies near a local space is closely predicted. On the other hand, the BPNN model maps the relationships between the inputs and outputs in global space for the training scenarios. Therefore, the model fails to predict the localized variation in data in the testing phase and this is also reflected in model validation phases. In relation to the simulation time, the RBFNN is faster than the BPNN for larger data sets (Kim et al., 2003).

In training an RBF network, the hidden neurons are self-organized with the training process (Zhang and Kushwaha, 1999). For this purpose, the orthogonal least squares (OLS) algorithm proposed by Chen et al. (1991) was employed. According to this algorithm, the number of hidden neurons at the beginning of training is zero. The hidden neurons are added one by one with training until the output of the network is within a target precision. For every iteration, the RMSE from the network is computed. If the error is lower than a pre-defined tolerance (selected from the lowest RMSE in BPNN training), the training is stopped and the number of neurons added to the hidden layer represent the number of hidden neurons required.

2.3.3. *General regression neural network (GRNN)*

The GRNN algorithm has little resemblance to the more widely used BPNN but it is one of the variants of RBFNN (Huang and Williamson, 1994). It has, at its root, one of the most commonly used statistical techniques, i.e. regression analysis and no iteration are involved in computing with the GRNN algorithm. Another attractive feature is that, unlike BPNN, a GRNN does not converge to local minima, can handle incomplete patterns and approaches the problem on the basis of the probability density function (pdf) of the training data (Huang and Williamson, 1994). GRNN uses one-pass learning algorithm which can be used for the estimation of continuous variables, and converge to the underlying regression surface. Mathematically, GRNN uses a standard statistical algorithm for calculating the conditional mean Y of a scalar random variable y given a measurement X of a vector random variable x. The vector x corresponds to the input to the network and the random variable y corresponds to the output of the network. If there is more than one output node, the same algorithm is used on each output node. GRNN is dominated by the estimation of pdf of x and y and is also used as a static regression technique. GRNN can be used in situations where the statistics of the data are changing over time. This is achieved by specifying a time constant and a threshold, which are used to reset a pattern node if it has not been used recently. The principal advantages of GRNN are its quick learning and fast convergence to optimal regression surface with large numbers of data sets when compared with BPNN and RBFNN (Kim et al., 2003).

2.3.4. *Neural network simulations and optimum network configuration*

The available data were divided into training (50% of data), testing (30% of data) and validation (20% of data), with the training and testing files comprising six inputs and two outputs, and the validation file comprising only the input parameters that were not used for the training and testing processes. The data were partitioned as per the indicated percentages to prepare separate data sets for training, testing and validation processes of the ANN models. However, the percentage of data used for partitioning is based on the concept that major share of it should be used for training processes followed by testing and validation processes. The data were further shuffled within the spreadsheet and 20 data sets were prepared for analysis to nullify the presence of any existing trend and inherent properties within the data (Sarangi and Bhattacharya, 2005; Zhang and Govindaraju, 2003; Patel et al., 2002).

Sensitivity analysis was done to determine the optimum network configuration for BPNN and GRNN by varying two network parameters, learning rate and number of hidden neurons that minimized the error of estimation. Learning rate indicates the rate of change of connection weights during training. A high learning rate causes oscillation of the connection weights resulting in large generalization error, while a low learning rate results in a significant increase in training time. It was observed that, with use of more neurons ($>$20) in hidden layer of BPNN, the network becomes over fitted, in which case it is capable of fitting the training data very well but incapable of generalizing for unknown inputs, i.e. out-of-sample data. Also, a large number of hidden neurons significantly increase the network training time. A small number of hidden units results in under fitting due to the lack of enough processing units to map the input/output relationship. Sensitivity analysis was also done to determine the optimum value of tolerance and receptive field for the RBFNN. Also, for each iteration, the sum of squared error from the network was computed. When the error became lower than a predefined tolerance, the training was stopped. At this stage, the numbers of neurons added to the hidden layer represented the number of hidden neurons required. If the sum of squared error was above the tolerance then the input pattern with largest error was identified and added to the hidden layer. This process was continued till the network error was minimum and within the tolerance limit. It was also observed that with the increase in the number of hidden neurons the computational time increased. Keeping this in view, the trial and error approach was employed in assigning the number of neurons in RBFNN. It was revealed that, with the change of receptive field from 5 to 35 with an increment of 5, keeping the tolerance constant at 5, the number of neurons in hidden layer increased and the statistical parameters (coefficient of determination (R^2) and model efficiency (E)) improved, indicating better network performance. But, with further increase in receptive field ($>$15), there was no improvement in statistical parameters (R^2, E). In this study, receptive field of 15 with tolerance constant of 10 leading to 25 numbers of neurons in the hidden layer was found optimal (RMSE = 0.012) for the operation of RBFNN. The number of neurons of the hidden layer in BPNN, GRNN and RBFNN models and the optimal nodes were selected based on the

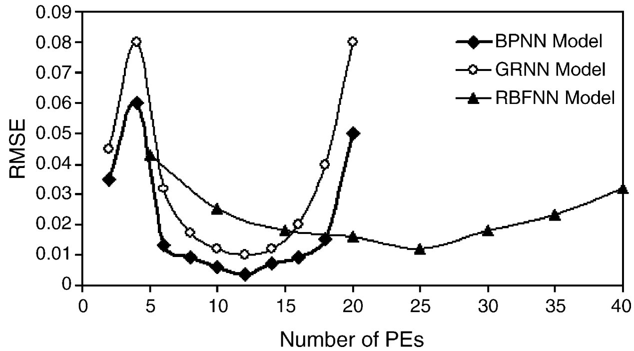

Fig. 2 – The RMSE of BPNN, GRNN and RBFNN models adopted for simulating the salinity of subsurface drained rice fields.

RMSE of the learning results (Fig. 2). It is seen from Fig. 2 that the RMSE was lowest for the BPNN model (0.0045) for 12 hidden layered PEs than the GRNN (0.01) and the RBFNN (0.012) with 25 layered PEs for modeling the drainage effluent salinity for both the spacing. Finally, the BPNN model with the optimal architecture (1 input layer with 6 PEs, 1 hidden layer with 12 hidden neurons and 1 output layer with 2 PEs) resulted in the statistical parameters (R^2, E) significantly higher than the RBFNN and GRNN algorithms. Also, the operation time of the RBFNN was more than BPNN due to more number of neurons in the hidden layer. Similarly, it was observed by comparing the RMSE of BPNN, GRNN and tolerance value of RBFNN that the prediction error of BPNN in simulating the root zone depth soil salinity was the lowest (0.065). In this study, the learning rate was varied from 0.01 to 0.09, hidden neurons varied from 10 to 25 and receptive field varied from 5 to 35. The low ranges of learning rate were chosen because high fluctuations in error were observed at higher learning rates.

The ANNs used in this study were diverse in their structures, network algorithms and applicability to specific field problems. Keeping this in view, the sensitivity of different ANN model training parameters were meticulously adopted to select the optimal architecture for these three ANNs. It was revealed that, the BPNN model performed better than the rest, which can be attributed to its nonlinear pattern recognition capability with specified learning rules and activation transfer functions applied to the data sets of subsurface drained rice fields. Also, the experimental data of *rabi* season exhibited unexplained variations between the climatic parameters, soil moisture regimes, drainage spacings, drained effluent salinity and quantity and salinity variations in root zone depths. Therefore, the smaller data sizes coupled with inherent variability could be the reason for better training of the BPNN model over the RBFNN and GRNN model algorithms.

2.3.5. *SALTMOD simulations*

SALTMOD is a simulation model, which predicts root zone soil salinity, drainage water quality and water table depth in agricultural land under different geo-hydrological conditions and varying water management scenarios. The model runs with hydrologic data, soil strata information, water balance components, drainage criteria and system parameters and initial and boundary conditions as listed in Table 2 and detailed in Singh et al. (2002b). The technical terms used in Table 2 for SALTMOD model calibration are also discussed in Singh et al. (2002b). The model assumed uniform distribution of the cropping, irrigation and drainage characteristics over the 4.0 ha experimental site and uses Hooghoudt's steady state formula to obtain the flow components from above the drain and from below the drain when the water table is below the soil surface. The minimum and maximum time step of computations was 1 and 12 months, respectively. The SALTMOD was calibrated and validated for the experimental site for

Table 2 – Summary of input parameters of SALTMOD (Singh et al., 2002b)

Input parameters	Parameter values
1. Soil properties	
Fraction of irrigation or rain water stored in root zone	0.65
Total porosity of root zone	0.60
Total porosity of transition zone	0.45
Total porosity of aquifer (assumed)	0.35
Drainable porosity of root zone	0.05
Drainable porosity of transition zone	0.08
Drainable porosity of aquifer	0.25
Leaching efficiency of root zone (calibrated)	0.60
Leaching efficiency of transition zone (assumed)	0.80
Leaching efficiency of aquifer (assumed)	1.00
2. Water balance components	
Irrigation in the season	1.25
Rainfall in the season	0.04
Evapotranspiration in the season	0.76
Incoming groundwater flow through aquifer during the season	0.0
Outgoing groundwater flow through aquifer during the season	0.0
Surface runoff in the season (calibrated)	0.35
3. Drainage criteria and system parameters	
Root zone thickness (m)	0.30
Depth of subsurface drains (m)	1.00
Thickness of transition zone between root zone and aquifer (m)	1.60
Thickness of aquifer, assumed (m)	5.00
Ratio of drain discharge and height of the water table above drain (m/(d m))	0.0011–0.015
Rate of drain discharge and squared height of the water table above drain (m/(d m^2))	0.00015–0.002
Drainage reduction factor in the season	0.2
4. Initial and boundary conditions	
Depth of the water table in the beginning of the season	0.30
Initial salt concentration of soil moisture in root zone at field saturation (dS/m)	35.0
Initial salt concentration of the soil moisture in transition zone (dS/m)	40.0
Average salt concentration of incoming groundwater (dS/m)	50.0
Average salt concentration of incoming Irrigation water (dS/m)	1.5

m: meter; d: days; dS: deci-siemens.

which the ANN models were developed in this study. The SALTMOD assumed the solute movement to take place as mass flow and the location of the subsurface drain to be anywhere in the transition zone. The overall functioning of the model was based on the principle of mass conservation. However, the ANN models are based on the observed data of the experimental site only without any detailed physical process involved.

2.3.6. Evaluation of the ANN models

All the three ANN models were run with the selected network architectures (Fig. 2) using the independent training and testing data sets. Then the trained model was validated with the unexposed data followed by estimation of R^2 and E, using the relation:

$$E = 1 - \frac{\sum_{i=1}^{n}(p_i - o_i)^2}{\sum_{i=1}^{n}(o_i - o)^2} \tag{3}$$

where n, total number of observations; o_i, ith observed value; o, mean of observed values and p_i, ith predicted value. The best model was selected based on the R^2 and E value approaching 1.0 (James and Burgess, 1982).

3. Results and discussions

The results of the ANN models were compared with the SALTMOD results as presented in Singh et al. (2002b). The SALTMOD and ANN models were simulated using the available data of the same study area.

3.1. Performance of ANN models

Based on the statistical parameters (R^2, E) as discussed, the BPNN neural network with 12 neurons in the hidden layer and a learning rate of 0.02 was found optimum for simulation of salinity of root zone depth soil and drainage effluent. It was also observed that the change in the epoch numbers for normalized cumulative delta learning rule in BPNN did not affect the prediction accuracy significantly. The epoch is the number of sets of training data sets presented to the learning cycles during weight updates. The variation of epoch from 4 to 16 tried for training the BPNN model did not yield any significant variation of the RMSE value. Therefore, the developed BPNN model was validated using the model efficiency factor E and R^2 of observed and model predicted values (Fig. 3). It was observed from the figure that the BPNN model performed well for both the 35 and 55 m drain spacing with R^2 and E values approaching 1. However, the BPNN performed poorly for predicting the root zone soil salinity (R^2 = 0.45; E = 0.4). This may be attributed to the improper representation of the salt balance in soil profiles by the neural network approach.

3.2. Comparison of the SALTMOD and BPNN models

3.2.1. Subsurface drainage water quality

The observed and predicted salinity of subsurface drainage water from the lateral drains of 35 and 55 m spacing for 4 years using SALTMOD and BPNN models are presented in Fig. 4. It

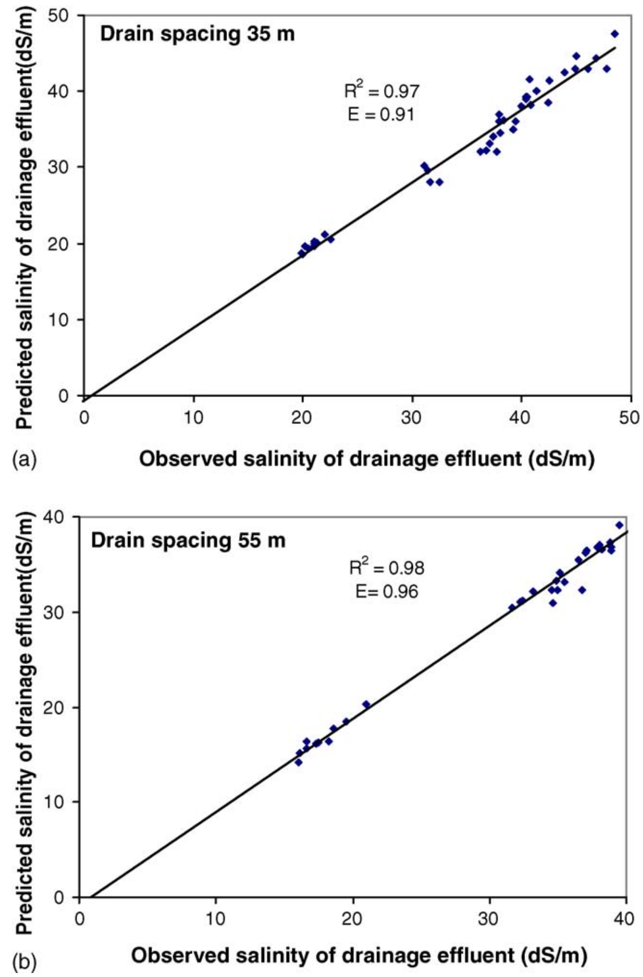

Fig. 3 – The observed and predicted drainage effluent salinity using back propagation neural network (BPNN) model (*rabi* season (January to May) of each year).

was observed that for SALTMOD predictions, the deviation ranged from 21 to 27% in 35 m spacing and 1.5 to 25% in 55 m spacing. While, the BPNN model predictions were more close to the observed values with variation of 5–15% in 35 m spacing and 2–12% in 55 m spacing. Thus, the BPNN model performed better than the SALTMOD in predicting the salinity of the drainage effluent. Besides, it was observed that the drainage rate had increased significantly by over 20 and 10% in the 35 and 55 m drain spacing areas, respectively, after 2 years of operation of subsurface drainage system. In the same study, it was also observed that the improvement in soil physical condition was faster in 35 m spacing as compared to the 55 m spacing. The data presented in Fig. 4 indicated that salt concentrations in effluents and drainage rate were always much higher in 35 m spacing as compared to 55 m spacing. This signified the higher pace of reclamation with 35 m drain spacing. Application of SALTMOD and BPNN models revealed that the land with 35 and 55 m drain spacing for the existing agro-climatic condition might be reclaimed for rice–rice crop rotation within 4–6 years, respectively. This inference drawn from the simulation is in good agreement with the finding based on field monitoring.

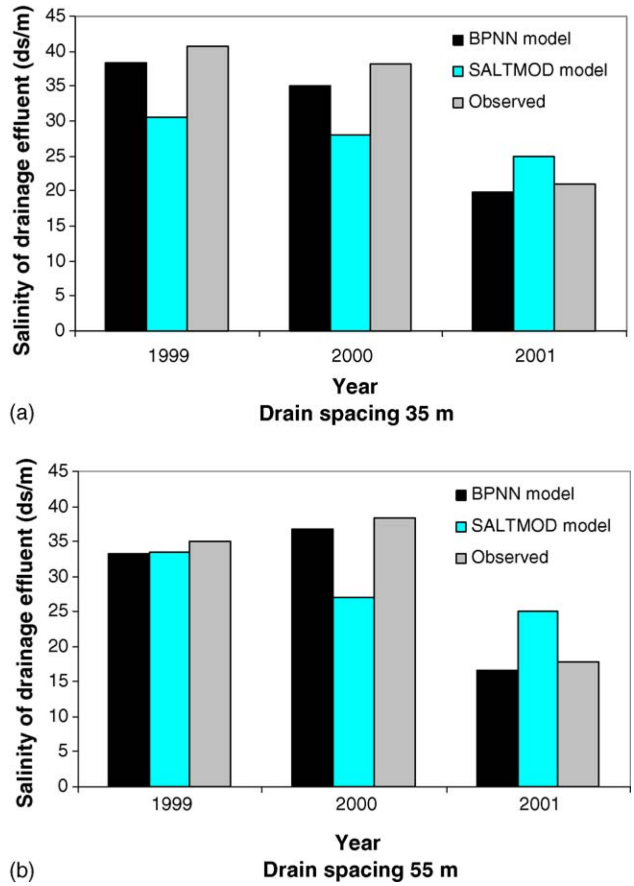

Fig. 4 – Comparison of seasonal (January to May) prediction of drainage effluent using BPNN and SALTMOD models.

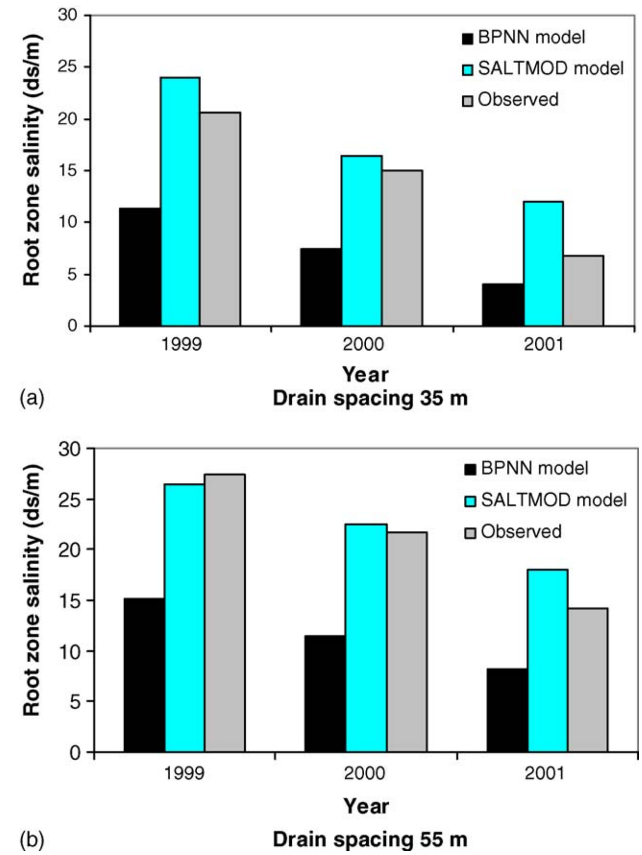

Fig. 5 – Comparison of seasonal (January to May) prediction of root zone salinity (0–30 cm) using BPNN and SALTMOD models.

3.2.2. *Root zone soil salinity*

The results of simulation of the root zone salinity by SALTMOD, using the parameters listed in Table 2 together with the observed data for 4 years (i.e. from the year 1998 to 2001) for the season 1 are presented and compared with the BPNN model (Fig. 5). The results indicate that the SALTMOD performed well and the deviation between the model predicted and the observed root zone salinity varied from 5.3 to 8.9% in 35 m drain spacing and from 2.6 to 15.3% in 55 m drain spacing. The model overestimated the root zone salinity in both 35 and 55 m drain spacing. However, just after 1 year of operation of the subsurface drainage system, the model under-estimated the root zone salinity in 55 m drain spacing. The simulated values of root zone salinity stabilized in a period of 6 years. Therefore, the predictions made by the model suggest that the land with 35 and 55 m drain spacing, for existing soil, water and climatic conditions, may be reclaimed for rice–rice cultivation within 4–6 years. The data presented by Singh et al. (2002b) suggest that there was faster removal of salts in 35 m spacing as compared to 55 m spacing. The same study further concluded that coastal clay soil with given initial conditions would be reclaimed in 3–4 and 6–7 years with 35 and 55 m drain spacing, respectively, provided no additional salts are added to the top 1 m depth of the soil profile from external sources. The results of simulation by SALTMOD are in good agreement with the field estimated values of soil salinity in the root zone as reported by Singh et al. (2002b). However, the BPNN failed to correctly predict the root zone depth salinity and the deviation varied from 45 to 60% with very low R^2 (0.45) and model efficiency ($E = 0.4$). The failure of BPNN can be attributed to lack of specific model algorithm within the ANN to account for the salt balance in the soil profile.

4. Conclusions

This study was done to investigate the applicability of ANN approaches in modeling the root zone soil salinity and salinity of drainage effluent from subsurface drained rice fields in the coastal clay soils of Andhra Pradesh, India. The input parameters and the ANN model architecture was decided based on use of MARS tool and trial and error approach leading to optimal error statistics. The results of the predictability of best ANN model (BPNN model) were then compared with those of SALTMOD model to evaluate the model performances. Performances of both the models were evaluated by comparing the model generated predictions with the recorded data of the study area. It was concluded that, BPNN model performed better for prediction of the drainage effluent salinity for both the drain spacings but failed to predict the root zone soil salinity properly. Therefore, BPNN modeling approach detailed in this study can be applied to similar subsurface drained fields to predict the salinity of the drainage effluent and the reclamation period with acceptable accuracy using minimal input data. The SALTMOD requires some specific field observations for model calibration, which can be avoided with the ANN approaches. The finding of Singh et al. (2002b) regarding the SALTMOD prediction of drainage effluent salinity to be independent of the root zone soil salinity was confirmed in this study as the BPNN model predictions for both salinities were divergent. The ANN approaches are simpler, relatively faster in model development and simulation and can operate on minimal data structure in comparison to conceptual model SALTMOD. Also, consideration of the system based parameters in preparation of the PEs of ANN input layers resulted in elevating the complete ''black-box'' approximation of ANN towards ''gray box'' model representations. However, SALTMOD may be preferred over ANN models for detailed salt balance component estimation in subsurface drained fields.

REFERENCES

Abraham, A., Steinberg, D., 2001. Is neural network a reliable forecaster on earth? A MARS query. In: Jose, M., Alberto, P. (Eds.), Bio-Inspired Applications of Connectionism, Lecture Notes in Computer Science, vol. 2085. Springer-Verlag, Germany, Spain, pp. 679–686.

ASCE, 2000. Artificial neural networks in hydrology. II. Hydrologic applications. ASCE J. Hydrol. Eng. 5 (2), 124–137.

Bhattacharya, A.K., 1999. Drainage of agricultural lands. In: Singh, G.B., Sharma, B.R. (Eds.), 50 Years of Natural Resource Management Research. Division of Natural Resource Management, ICAR, Krishi Bhavan, New Delhi, India, pp. 347–362.

Chen, S., Cowan, C.F., Grant, P.M., 1991. Orthogonal least squares learning algorithm for radial basis function networks. IEEE Trans. Neural Networks 2, 302–309.

Friedman, J.H., 1991. Multivariate adaptive regression splines. Ann. Stat. 19, 1–141.

Huang, Z., Williamson, M.A., 1994. Geological pattern recognition and modeling with a general regression neural network. Can. J. Explor. Geophys. 30 (1), 60–68.

James, l.D., Burgess, S.J., 1982. Selection, calibration and testing of hydrologic models. In: Haan, C.T., Johnson, H.P., Brakensiek, D.L. (Eds.), Hydrological Modeling of Small Watersheds. American Society of Agricultural Engineers, St. Joseph, MI, pp. 215–257.

Kim, B., Kim, S., Kim, K., 2003. Modelling of plasma etching using a generalized regression neural network. Vaccum 71, 497–503.

Oosterbaan, R.J., 1998. SALTMOD ver 1.1: Description of Principles and Applications. ILRI, Wageningen, The Netherlands, 106 pp.

Oosterbaan, R.J., Abu Senna, M., 1989. Using saltmod to predict drainage and salinity in the Nile Deltas. In: Annual Report, ILRI, Wageningen, The Netherlands, pp. 63–74.

Patel, R.M., Prasher, S.O., Goel, P.K., Bassi, R., 2002. Soil salinity prediction using artificial neural networks. J. Am. Water Resour. Assoc. 38 (1), 91–100.

Salehi, F., Prasher, S.O., Amin, S., Madani, A., Jebelli, S.J., Ramaswamy, H.S., Tan, C., Drury, C.F., 2000. Prediction of annual nitrate-N losses in drain outflows with artificial neural networks. Trans. ASAE 43 (5), 1137–1143.

Sarangi, A., Bhattacharya, A.K., 2005. Comparison of artificial neural network and regression models for sediment loss prediction from Banha watershed in India. Agric. Water Manage. 78, 195–208.

Sharma, V., Negi, S.C., Rudra, R.P., Yang, S., 2003. Neural networks for predicting nitrate-nitrogen in drainage water. Agric. Water Manage. 63, 169–183.

Shukla, M.B., Kok, R., Prasher, S.O., Clark, G., Lacroix, R., 1996. Use of artificial neural networks in transient drainage design. Trans. ASAE 39, 119–124.

Singh, M., Bhattacharya, A.K., Nair, T.V.R., Singh, A.K., 2002a. Nitrogen losses through subsurface drainage effluent in coastal rice fields from India. Agric. Water Manage. 52, 249–260.

Singh, M., Bhattacharya, Singh, A.K., Singh, A., 2002b. Application of SALTMOD in coastal clay soils in India. Irrigation Drainage Syst. 16, 213–231.

Sivakumar, B., Jayawardena, A.W., Fernando, T.M.K.G., 2002. River flow forecasting: use of phase-space reconstruction and artificial neural networks approaches. J. Hydrol. 265, 225–245.

Sudheer, K.P., Gosain, A.K., Ramasastri, K.S., 2002. A data-driven algorithm for constructing artificial neural network rainfall-runoff models. Hydrol. Process. 16, 1325–1330.

Tsoukalas, L.H., Uherg, R.E., 1996. Fuzzy and Neural Approaches in Engineering. John Wiley & Sons Inc., New York, pp. 196–267.

Tyagi, N.K., Tyagi, K.C., Pillai, N.N., Willardson, L.S., 1993. Decision support for irrigation system improvement in saline environment. Agric. Water Manage. 23 (4), 285–301.

Yang, C.C., Prasher, S.O., Lacroix, R., 1996. Applications of artificial neural networks to land drainage engineering. Trans. ASAE 39, 525–533.

Yitian, L., Gu, R.R., 2003. Modeling flow and sediment transport in a river system using an artificial neural network. Environ. Manage. 31 (1), 122–134.

Yu, C., Northcott, W.J., McIsaac, G.F., 2004. Development of an artificial neural network model for hydrologic and water quality modeling of agricultural watersheds. Trans. ASAE 47 (1), 285–290.

Zhang, B., Govindaraju, R., 2003. Geomorphology-based artificial neural networks (GANNs) for estimation of direct runoff over watersheds. J. Hydrol. 273, 18–34.

Zhang, G.P., 2000. Neural networks for classification: a survey. IEEE Trans. Syst. Man Cybernetics C: Appl. Rev. 30, 451–462.

Zhang, Z.X., Kushwaha, R.L., 1999. Applications of neural networks to simulate soil–tool interaction and soil behavior. Can. J. Agric. Eng. 41, 119–125.