American Society of
Agricultural and Biological Engineers

An ASABE Meeting Presentation

Paper Number: 062124

Neural Approach of Sub-pixel Rice Landuse Classification for Optimized Irrigation Scheduling

Manoj Karkee

Agricultural and Biosystems Engineering, Iowa State University, karkee@iastate.edu.

Brian L. Steward

Agricultural and Biosystems Engineering, Iowa State University, bsteward@iastate.edu.

Lie Tang

Agricultural and Biosystems Engineering, Iowa State University, lietang@iastate.edu.

Written for presentation at the
2006 ASABE Annual International Meeting
Sponsored by ASABE
Portland Convention Center
Portland, Oregon
9 - 12 July 2006

Abstract. *Irrigation scheduling optimization is carried out in the context of a complex system of agricultural practices and crop calendars. Remote sensing is being used for the monitoring of crop development, crop health, and cropping practices. However, this is possible only if the resolution is sufficiently high to classify patches of different types of crops and cropping practices. MODIS imagery is essential for national or regional scale studies, but has a spatial resolution of 1 km and thus results in sub-pixel mixing of different land covers. In the case of rice farms, one pixel may consist of some proportions of rice grown under different cropping system such as one, two and three crops per year as well as other land covers. Classification of the land area covered by individual pixels is of the great importance for irrigation scheduling. A method was developed for classifying sub-pixel rice land area using a neural network. Temporal patterns of NDVI, which can easily be remotely sensed, depend on and result from the complex relationship between NDVI and cropping practices associated with a pixel. These parameters consist of the proportions of different types of rice and their emergence dates. An artificial neural network (ANN) was used as a model inverter to estimate these parameters. The data for this research were produced using the SWAP crop growth model. The ANN produced up to 95.7% accuracy in crop proportion quantification with an average emergence date error of 9 days. This method had a low computational cost taking 1.22 microseconds per pixel classification in a candidate experiment conducted in a laboratory personal computer.*

Keywords. pixel mixture, artificial neural network, remote sensing, temporal vegetation signature, rice cropping practice.

Introduction

Optimized use of the irrigation water is essential for increased crop production under limited water resources. According to Seckler et al. (1999), 70% of the total developed water of the world is consumed only by irrigation. Due to the increasing pressure to improve water utilization, there is a strong need for sustainable water use among multiple water users. Irrigated agriculture, being the major user of water, could offer much potential in this regard (Ines, 2002). Optimal water use is an important issue to be addressed while dealing with efficient water utilization in crop production (Toung and Bhuiyan, 1999). Irrigation schedule optimization is carried out in the context of a complex system of agricultural practices.

Over the years, remote sensing (RS) has been developed into a promising technology to monitor agricultural and water management activities because both the spatial and temporal characteristics of a region can be easily obtained by RS imagery (Droogers and Bastiaanssen, 2002). As a result, seasonal changes of vegetation activities at the regional level can be monitored to help policy makers and farm/water managers make better operational decisions (Ines and Honda, 2005). Crop monitoring from RS has conventionally been practiced with high resolution remote sensing data. However, high resolution techniques are beneficial only to small scale applications and have higher cost and low temporal coverage. Low spatial resolution data have attributes that clearly outweigh the limitations associated with the use of high spatial resolution data for larger scale agricultural monitoring - they have higher temporal resolution, wider spatial coverage, and minimal cost (Takeuchi et. al., 2003). One big challenge with the low resolution technique, however, is the problem of land cover mixing inside a pixel. Since the spatial resolution is coarser, several agricultural landuses could be embedded in one pixel (Ines and Honda, 2005).

The temporal pattern of vegetation signatures has been used for landuse classification over the past decades. Several studies have been carried out to classify landuse based on one year-long patterns of composite Normalized Differential Vegetation Index (NDVI). A composite NDVI is the average of a number of NDVIs over a period of time. So, a 10 day composite NDVI is the average of daily NDVIs over 10 days period. Composite NDVIs are used to shorten a temporal NDVI pattern to a reasonable length. Lacoul et al. (2000) prepared the monthly composites of NDVI patterns which represented the specific land covers and used ISODATA clustering to classify the landcover based on these patterns. Similar studies were done by Defries and Townshend (1994) and Sugita and Yasuoka (1996). Keiko et al. (1999) compared vegetation classification based on the time series NDVI patterns from NOAA/AVHRR with the actual vegetation data to investigate the feasibility of vegetation classification from time series NDVI.

Ines and Honda (2005) addressed the mixed pixel problem using simulated data and a genetic algorithm. The SWAP (Soil Water Atmosphere Plant) model was used to generate the simulated dataset for the study. SWAP is a commonly used computer model that simulates transport of water, solutes and heat in variably saturated top soils (SWAP, 2006). The evolution of the temporal LAI patterns to closely match the simulated LAI patterns was successful. However, the technique proposed by them had two limitations. First, the genetic algorithm took approximately ten minutes to evolve the temporal pattern of Evapotranspiration (ET) or Leaf Area Index (LAI) of a single pixel. It is difficult to apply this technique, in its given form, to a practical image, which typically consists of millions of pixels. Second, they used biophysical parameters like ET and (LAI) as the inputs of the system. It is difficult to measure these parameters from the satellite images as compared to the vegetation indices like NDVI.

Extensive work has been reported in the literature addressing the mixed-pixel problem, but the primary foci have been on other areas than agricultural monitoring and water management

(Shimabukuro and Smith, 1991). In agricultural monitoring programs, the mixed pixel problem is not just the quantification of landuse proportions, but also the estimation of sowing or emergence dates associated with different cropping practices. With the wealth of information embedded in the temporal patterns of low spatial resolution data, it may be possible to derive such information at the sub-pixel level (Ines and Honda, 2005).

The objective of this study was to develop a technique that was applied to low spatial resolution remote sensing data to quantify agricultural practices within a pixel using an Artificial Neural Network (ANN). An ANN is an information-processing paradigm that was inspired by the way biological nervous systems, such as the brain, process information. Human beings learn by experience; neural nets learn by setting connection weights between nodes that will produce a specified output. When the ANN is presented with new data, it applies the weights, and the resulting output is consistent with previous experience (Diane et al. 1995). ANNs have been applied to remote sensing image interpretation. Typically, there are two categories of applications: classification (Diane et al., 1995; Chen et al., 1997; Ito and Omatu, 1997; Brivio et al., 2003) and modeling and/or inverse modeling (Chen et al., 1993; Delfrate and Wang, 2001).

The technique presented in this paper also uses the SWAP model to generate the dataset for the study. The inputs to the system, however, are an NDVI pattern, which can be calculated directly from remote sensing images.

Materials and Methods

At the pixel level, RS data could be mixed; it can be composed of signatures coming from several land features including a variety of crops and bare soil as well as other landuses. This research used the temporal vegetation signature of a pixel to classify different cropping practices within a pixel. To address the problem, a typical example of rice field has been chosen. Three types of cropping practices were used: one crop per year rainfed rice (rice1), two crops per year irrigated rice (rice2) and three crops per year irrigated rice (rice3). These practices are commonly used in the rainy plains of south and south-east Asia. The NDVI pattern of a pixel to be observed in this environment typically depends upon the sowing or emergence dates, fractions of the landuse of three different types of rice, soil parameters and weather parameters (Fig. 1).

The method of classifying sub-pixel agricultural practices started with the generation of the training and testing datasets (Fig. 2). A temporal NDVI pattern and the corresponding cropping parameters defined an input-output pair to be used by an ANN. Next, an ANN was created to estimate the input-output relationship described by the available dataset. Finally, accuracy achieved by the ANN were assessed. In what follows, these steps are briefly described.

Generating Crop Parameters:

A pattern of crop parameters consists of emergence date for rice1 (e_{11}), two emergence dates for rice2 (e_{21} and e_{22}), three emergence dates for rice3 (e_{31}, e_{32} and e_{33}) and three area fractions covered by rice1, rice2 and rice3 (a_1, a_2 and a_3) in that pixel. However, to simplify the problem, emergence dates for rice3 were assumed to be fixed, which is close to actual practice. This simplification was necessary to reduce the computational time taken by the ANN in both training and testing phases.

One hundred patterns of crop parameter were generated from the random combination of emergence dates and the crop fractions subject to the constraints:

$$a_1+a_2+a_3= 1.0 \qquad (1)$$

$e_{22} \geq e_{21}+100$ (days) (2)

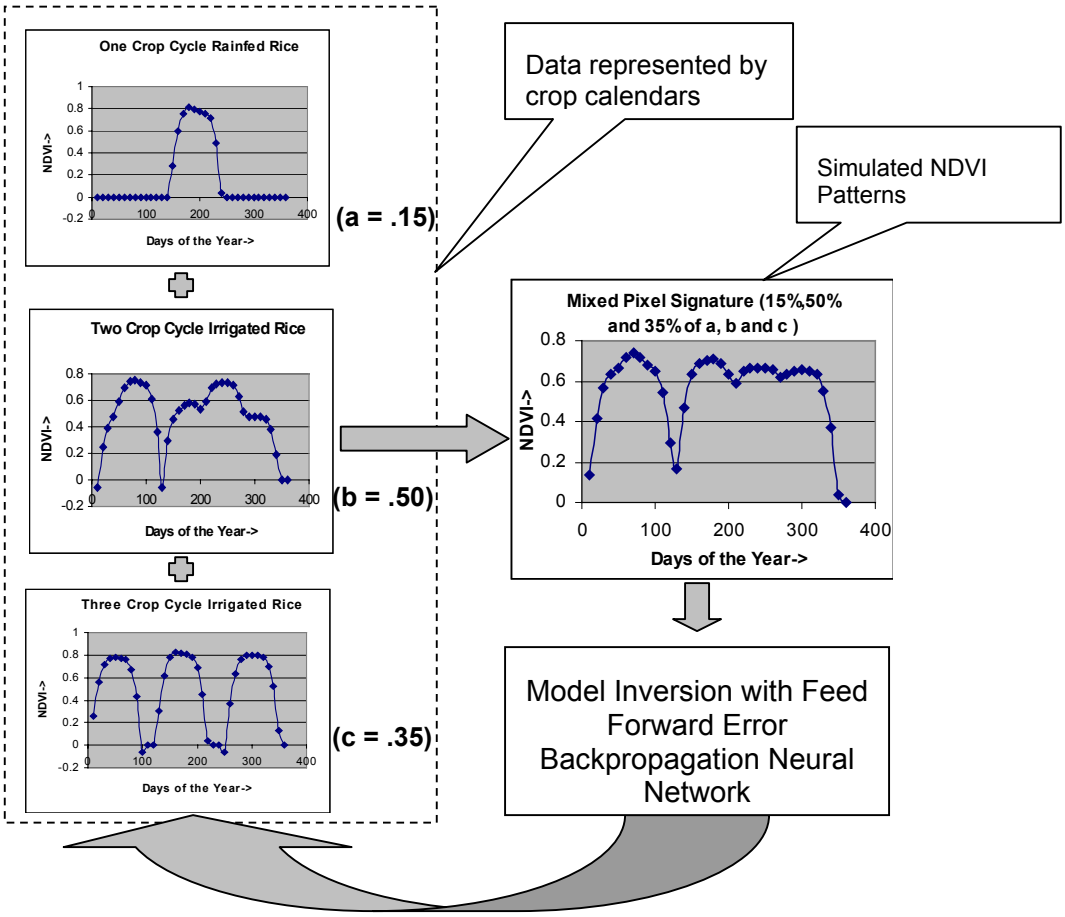

Fig. 1: Data flow schematic of the research. The box in the left represents the simulated NDVI patterns for three classes of rice. The pattern in the right is the composite signature.

Generating Temporal NDVI Patterns:

The temporal NDVI patterns used in this research were generated numerically using the SWAP model and a relationship between NDVI and LAI. Provided the crop parameters (those discussed in previous paragraph), soil parameters (residual moisture content, saturated moisture content etc.) and weather parameters (temperature, humidity, wind, rainfall etc.), the SWAP model generated temporal patterns of ET and LAI. Climatic information and soil parameters fed into the SWAP model were collected in Bangkok, Thailand. Since the input to the system is NDVI patterns, the LAI profiles generated by SWAP model were transformed into NDVI profiles using simple logarithmic relationship given by (NASA, 2004),

$$NDVI = 0.1395\ln(LAI) + 0.5816 \qquad (3)$$

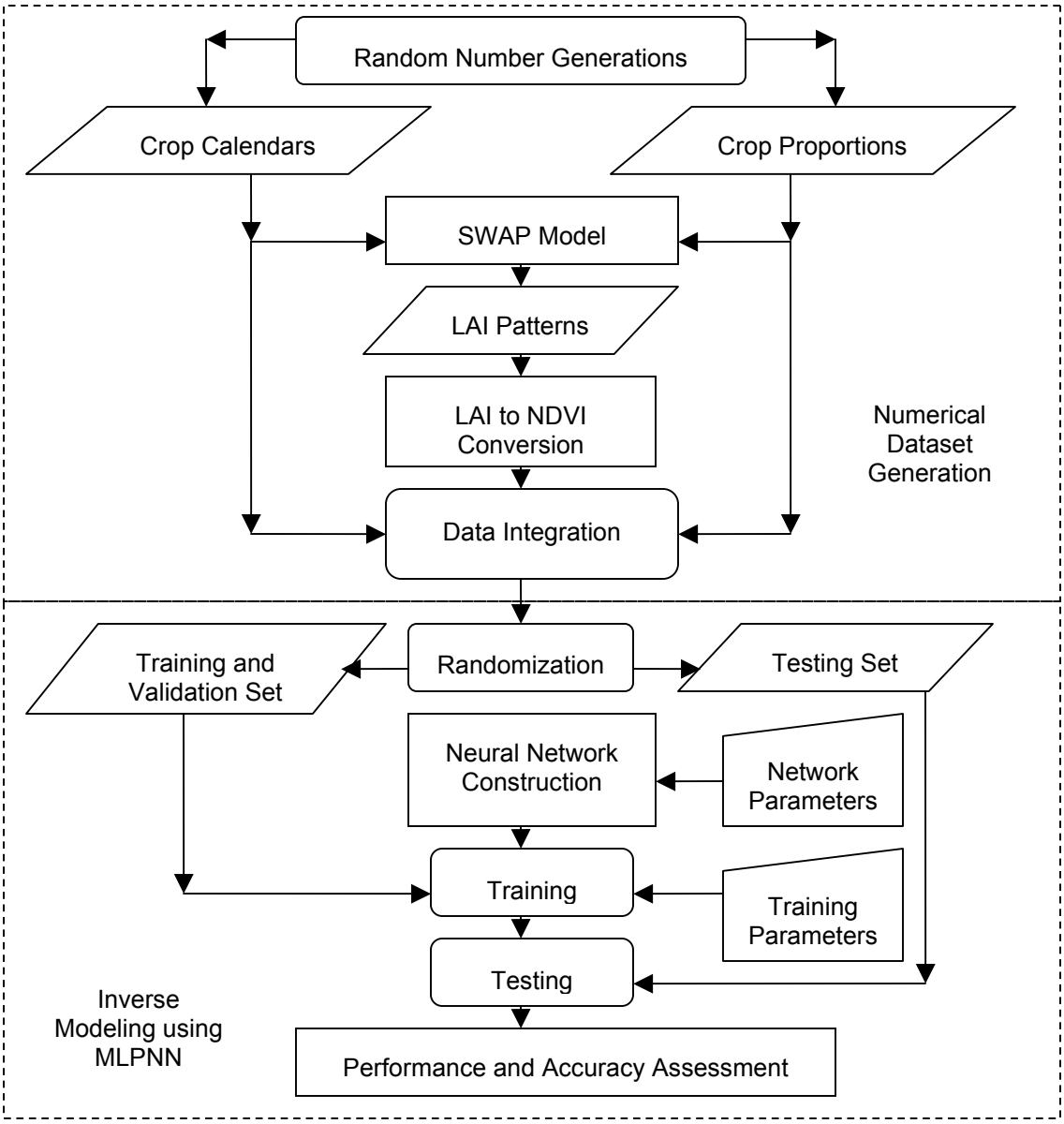

Fig. 2: Flow chart of the research. Upper half part of the flowchart shows the procedure to generate numerical data and the lower half part shows the use of the MLPNN for the inversion of the relationship between the input and output stakeholders perceived in the upper part.

Inversion of the Temporal NDVI Profiles:

The numerically-generated data were arranged in an input and output pairs and a neural network was trained to model the inverse relationship between agricultural parameters and the NDVI profiles. Out of 100 simulated NDVI patterns, 50 were assigned to the training set, 25 to the validation set, and the remaining 25 to the testing set.

Preprocessing and randomization:

Two of the typical preprocessing techniques applied to the dataset were normalization and Principle Component Analysis (PCA). Normalization transformed the dataset into a new

distribution having zero mean and unity standard deviation. PCA was performed to reduce the dimensionality of the dataset by removing those principal components which had a negligible variability associated with them. In this research, PCA preserved 99% variability of the dataset while reducing the dimensionality from 36 to 8. The preprocessing of dataset causes the equal learning rate and no synchronization of the weights in network training (Haykin, 1994). It also helps fit the available data into the limited numeric range available in the sigmiodal response function. The training, validation and testing sets were randomly selected and the training examples were presented to the network in a random order. This form of randomization is critical for improving the speed of convergence (Haykin, 1994).

Network Design:

A feedforward multilayer perceptron with error backpropagation neural net (MLPNN) was used as the model inverter. The inputs to the network were samples from the temporal NDVI profiles. Two different types of dataset were generated in the research; original 36 sampled NDVI patterns and PCA shortened 8 sampled NDVI patterns. Similarly, two different types of MLPNN were used. The first set of MLPNNs had 36 input nodes to accommodate 36 samples of a full-length NDVI pattern, and the second set of MLPNNs had 8 inputs to use 8 samples from the PCA shortened dataset. The output layer of the MLPNN consisted of six neurons to map an output pattern having e_{11}, e_{21}, e_{22}, a_1, a_2 and a_3. Target outputs were transformed into the unipolar range from +0.15 to +0.85 for the emergence dates and from 0 to 1 for the crop fractions. The range of crop fractions was not reduced to 15% to 85% as in emergence dates assuming that most of the randomly generated crop fractions will automatically be inside this range. Reduction of the unipolar range was necessary to cope with the dynamic range of sigmoid activation function, because the response of this activation function is flat at the limiting values.

Experiments on different architectures and training parameters:

Experiments were carried out with many different variations of one and two hidden layer networks. Comparative studies were carried out to see the differences in the network performance with full-length patterns and PCA reduced patterns. The network architecture that performed best was used to perform the experiments on the effects of varying learning rates and the momentum terms. The best combination of learning rate and momentum term was then assessed by evaluating the speed of convergence (number of epochs) and the training root mean squared errors (RMSE). Finally, the best network configuration and the best combination of learning rate and momentum term were used in the experiments with different activation functions.

Results and Discussion:

The NDVI patterns derived from the LAI patterns followed the emergence-growth-ripening cycle of the cropping practice very well (Fig. 3). It shows that the accuracy of the SWAP model and the NDVI-LAI relationship used in the research are sufficient for this study.

The experimental results with different network architectures (Table 1) showed that PCA could effectively be used to reduce the dimensionality of the dataset, which helps reduce the complexity of the network. As there were too many inputs in the full length patterns than the PCA shortened patterns, it needed many more connections to model the relationship. So, two hidden layers with a high number of neurons in each layer were found to be necessary in this case. The table clearly shows that the network performance (average of all three RMSEs) keeps decreasing with the decreasing numbers of hidden layers and the number of neurons in the

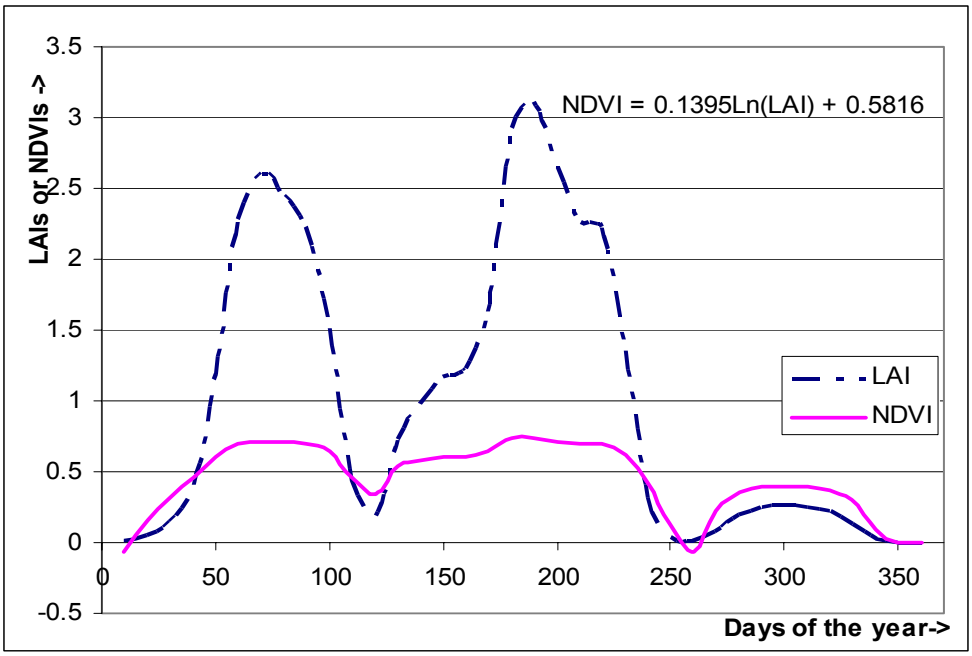

Fig 3: Sample LAI and NDVI patterns generated numerically using SWAP model

hidden layers for the network with full length inputs. However, only one hidden layer with fewer neurons was found to be enough to work with PCA shortened patterns. In this case, it is evident from the table that an increased number of layers and neurons did not really benefited to the network performance; rather decreased the performance in some cases. For instance, testing RMSE was 0.062 for the 08:14:06 network where as it was 0.047 for the 08:07:06 network.

Table 1: Comparative representation of the experiments with different network architecture.

Hidden Layers	Dataset*	Network Architecture**	Training Epochs	Training Time (ms)	RMSE Training	RMSE Validation	RMSE Testing
1	PCA	08:04:06	169	0063	0.036	0.045	0.040
		08:07:06	130	0076	0.035	0.050	0.047
		08:14:06	169	0203	0.030	0.054	0.062
	Full	36:05:06	091	0093	0.031	0.050	0.045
		36:10:06	065	0140	0.028	0.048	0.044
		36:21:06	156	0657	0.022	0.057	0.056
2	PCA	08:04:03:06	195	0094	0.034	0.048	0.049
		08:07:07:06	143	0140	0.034	0.053	0.047
		08:14:12:06	156	0344	0.030	0.050	0.049
	Full	36:05:04:06	182	0234	0.029	0.047	0.051
		36:15:10:06	091	0343	0.027	0.042	0.040
		36:26:16:06	143	1000	0.020	0.042	0.038

** Full length pattern has 36 samples and the PCA shortened pattern has 8 samples per pattern.*

*** Network Architecture representation is*

'Input_Neurons:Hidden_Layer1_Neurons:[Optional Hidden_Layer2_Neurons]:Output_Neurons"

The differences between validation and testing RMSEs of the 08:04:06 network and the 36:26:16:06 network (e.g. 0.040 – 0.038 = 0.002 for the testing set) were statistically insignificant at 95% confidence level. The network 36:26:16:06, however, was more complex than the 08:04:06 network. Consequently, the former architecture took only 63 milliseconds where as the later took 1000 milliseconds to converge. All other network architectures performed worse than these two. So, the best network chosen for the further experiments was the 08:04:06, which has only one hidden layer with four neurons in it and uses PCA reduced input patterns.

Experiments were carried out to see the effect of varying values of learning rate (Fig. 4) and momentum (Fig. 5) on network performance. We can think of the learning rate as the acceleration and think of momentum as the speed of the learning process. The smaller we make the learning rate, the smaller will the changes to the synaptic weights in the network be from one iteration to the next and the smoother will be the trajectory in weight space. Momentum term added in the delta rule helps avoid oscillation even with the higher value of learning rate (Haykin 94). Introduction of momentum term helps track the global minima, in case the training process could be trapped in local minima.

The higher learning rate generally meant faster convergence at the cost of degraded performance (Fig. 4). There is a breakeven point around the learning rate of 0.15, which is the optimal value of learning rate. However, particularly for this application, higher performance is more critical than the faster training time since the training session needs to be run only once for millions of feed-forward runs to be followed. So, the learning rate used was 0.1, a bit smaller than the breakeven value.

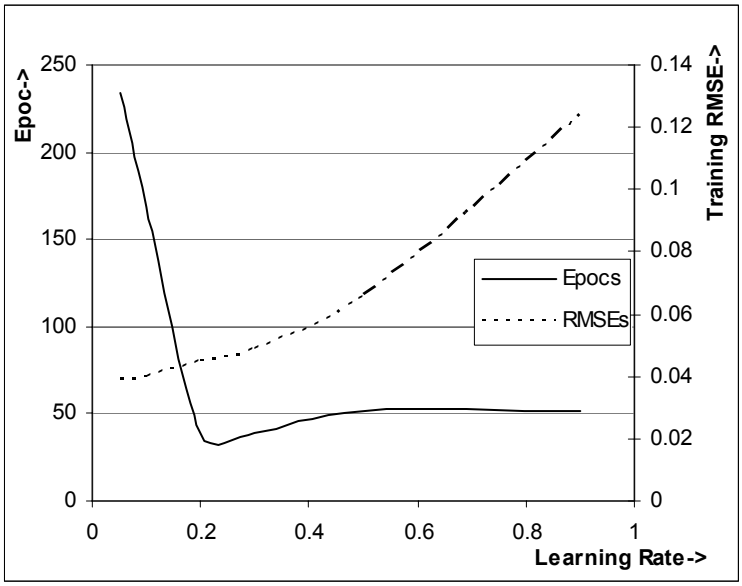

Fig. 4: Effect of learning rate on network performance

The network was not really able to learn at the momentum term of 0.9 (Fig. 5). For the range of 100 to 200 numbers of epochs, the network performed best with the momentum of 0.1, even though it was not significantly different than that with the momentum of 0.5 or 0.0.

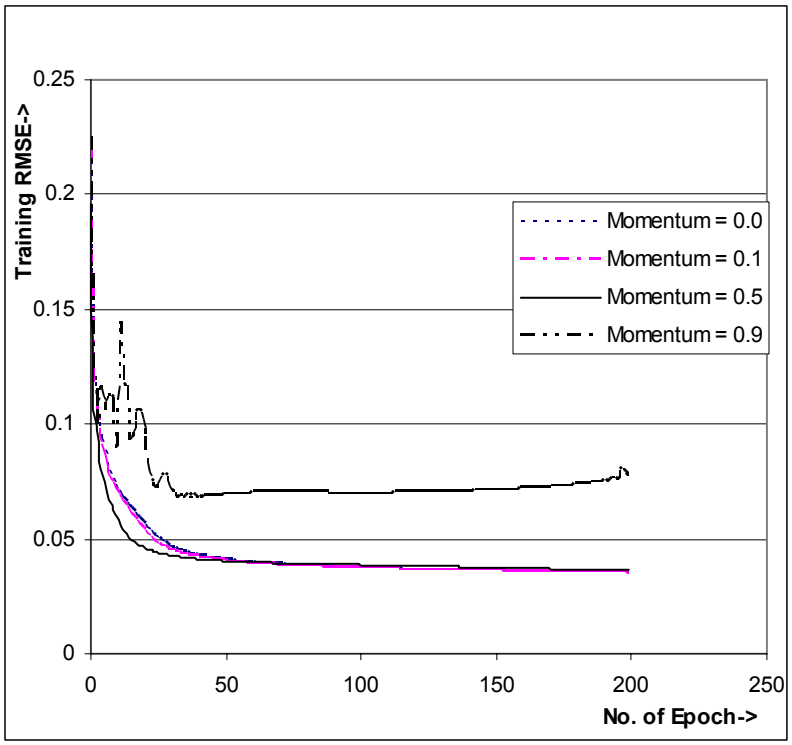

Fig.5: Learning curves for varying momentum term.

Finally, two types of hidden layer functions were tested. *'logistic sigmoid'* function performed better than *'hyperbolic sigmoid'* . It is because the unipolar response of *'logistic sigmoid'* matched to the output format used. Output layer activation function was *'pure linear'*.

The best network configuration of 08:04:06 was used to evaluate and summarize the test errors (Table 2). The learning rate used was 0.1, the momentum 0.1, the activation function 'logsig' with pure linear at the output layer and the training method was gradient decent with momentum. The table depicts the maximum and root mean squared error (RMSE) for all six output parameters from the MLPNN. The errors for the emergence dates given by the network are the fractions of a year (365 days) since they were converted into a scale of 0.0 to 1.0 while calculating the errors. To make the extent of these errors more meaningful, each such fraction has been converted into the number-of-days by multiplying it with 365 and is presented consecutively.

The RMSEs reported in the table were calculated using common RMSE formulae from statistics, which takes square root of the average of the squared errors for the outputs from all the test samples. Overall RMSE of the test set was 0.04. This means that, in general, there is a possibility of 14 days estimation error in the emergence dates and a 4% estimation error in the crop proportions. For the better understanding of the result, individual maximum errors and RMSEs were calculated for all the six outputs from the test results. The maximum error on network outputs for rice1 emergence dates is 0.078 (7.8%), which is almost 4 weeks. However, the RMSE for the same is 3.9%, which is equal to 14 days. The maximum error for the first emergence date of rice2 was only 0.043 or 16 days and the RMSE for the whole test outputs of this variable was 0.024, which is only 9 days. The maximum error for the second emergence date of rice2 was 0.083 or 30 days. However, the RMSE of 0.038, which is even smaller than that of the rice1, is evident of the fact that maximum error occurred in this case is only an outlier. Yet another outlier in the observations was the maximum value of the errors in the estimation of

the landuse proportion of rice2, which was 12.7%. The same for rice1 and rice3 were 0.076 and 0.106 with RMSE of 0.043 for both of them.

Table 2: Test errors produced by the best network configuration for the six output variables.

	Errors								
	Emergence Dates						Crop Proportions		
Error	rice1		1st, rice2		2nd, rice2				
Measures	Fraction	Days	Fraction	Days	Fraction	Days	rice1	rice2	rice3
Absolute Max Error	0.078	28	0.043	16	0.083	30	0.106	0.127	0.076
RMSE*	**0.039**	**14**	**0.024**	**09**	**0.038**	**14**	**0.043**	**0.065**	**0.043**

**Overall RMSE of the test set was 0.04.*

The technique achieved the RMSE of 14 days in the estimation of the emergence dates with 93.5% accuracy in the crop proportions estimation. This level of accuracy in sub-pixel agricultural practices quantification should be sufficient for the irrigation water management applications.

Computationally, the neural approach is much less costly relative to the GA approach in the study by Ines and Honda (2005). The approach here took 1.22 microseconds in a candidate experiment conducted in a laboratory PC with a Pentium 4 microprocessor running at 2.4 GHz and with 512 MB RAM to estimate the parameters for a single pixel. This was a tremendous improvement over the time taken by GA, which was about 10 minutes in similar hardware configuration. The training time of the candidate network was 63 ms. The computational speed achieved by the network is sufficient to apply the technique to the real satellite images having millions of pixels in a single scene.

Conclusions:

A number of different tests were carried out with different network architectures, learning rates, momentum terms and activation functions to see the performance on sub-pixel classification of a pixel with mixed vegetation signature. From this research, we can draw the following conclusions:

- The mean accuracy achieved for the rice proportion classification was varying from 93.5% to 95.7% and the RMSE of the estimated emergence dates was varying from 8 days to 14 days. These accuracies should be sufficient for the irrigation water management applications. The neural approach of classifying sub-pixel agricultural practices from a remote sensing imagery appears to be a practically applicable technique in water management practices.
- In a laboratory personal computer, it took only 1.22 ms to estimate the sub-pixel parameters. This computational speed should be enough to apply the approach to the real imageries having millions of pixels.
- In future, this study will be performed with some proportion of white noise added to the simulated data and finally with the real field data. Besides the computational benefits, this approach, in the implemented with real data, avoids dependency upon relationship between NDVI and LAI. This independence will remove some of the error sources associated in the experiments with numerical data.

A neural approach of problem solving has its own limitations. One of the major drawbacks is the generation of sufficient and well-distributed training and testing data sets. In this study also, this

problem has significantly attracted our attention since it has limited the immediate experiments with the real datasets. Further research is necessary to find the ways that can reduce the cost and time associated with field work data collection for the development of the training and testing datasets.

Acknowledgements

Executable programs, parameter files and related information for the SWAP model were provided by Dr. Kiyoshi Honda, Asian Institute of Technology, Thailand and Dr. Amor V.M. Ines, International Research Institute for Climate Prediction, The Earth Institute at Columbia University.

References

Brivio, P. A., M. Maggi and E. Binaghi. 2003. Mapping burned surfaces in Sub-Saharan Africa based on multi-temporal neural classification. *International Journal of Remote Sensing,* 24(20): 4003–4018.

Chen, Z., L. Tsang, Y. Kuga, and C. Chan. 1993. Active polarimetric microwave remote sensing of corn and multiparametric retrieval of corn parameters with an artificial neural network. *Microwave and Optical Technology Letters*, 6: 611-615.

Chen, K. S., S. K. Yen and D. W. Tsay. 1997. Neural classification of SPOT imagery through integration of intensity and fractal information. *International Journal of Remote Sensing,* 18(4): 763-783.

Defries, R.S., and J. R. G. Townshend. 1994. NDVI derived land cover classification at global scale. *International Journal of Remote Sensing,* 15(17): 3467-3486.

Delfrate F. and L.F. Wang. 2001. Sun flower biomass estimation using a scattering model and a neural network algorithm. *International Journal of Remote Sensing,* 22(7): 1235-1244.

Diane, M. M., E. J. Kaminsky and S. Rana. 1995. Neural Network Classification of Remote-Sensing Data. *Computers & Geosciences*, 21(3): 377-386.

Droogers, P. and W. Bastiaanssen. 2002. Irrigation Performance using Hydrological and Remote Sensing Modeling. *Journal of Irrigation and Drain Engineering,* 128(1): 11-18

Haykin S, 1994. Neural Networks: A Comprehensive Foundation. New York, N.Y.: Macmillan College Publishing Company.

Ines A. V. M. 2002. Improved crop production integrating GIS and Genetic Algorithm. PhD Dissertation. Bangkok, Thailand: Asian Institute of Technology.

Ines, A. V. M. and K. Honda. (2005). On quantifying agricultural and water management practices from low spatial resolution RS data using genetic algorithms: A numerical study for mixed-pixel environment. *Advances in Water Resources*, 28(8): 856-870.

Ito, Y. and S. Omatsu. 1997. Category classification method using a self-organizing neural network. *International Journal of Remote Sensing*. 18(4): 829-845.

Keiko, K., K. Pahari, M. Tamura and Y. Yasuoka. 1999. Feasibility Analysis for Vegetation Classification from Time Series NDVI data with “Green Census” data. *Asian Conference on Remote Sensing, 1999.*

Lacoul , M., L. Samarakkon and K. Honda. 2000. Vegetation Mapping in Ganges River Basin for Global Mapping Project. *Asian Conference in Remote Sensing, 2000.*

NASA. 2004. Relationship between NDVI and LAI: National Aeronautics and Space Administration. Available at:

http://ltpwww.gsfc.nasa.gov/bsb/hqvisit/16MAR2004/Deering.SibLAI_pt2.pdf. Accessed 15 Feb 2004.

Seckler, D., R. Baker and U. Amarasinghe. 1999. Water Security in the twenty first century. *Water Resources Development,* 15: 29-42.

Shimabukuro Y. E and J.A. Smith. 1991. The least-squares mixing models to generate fraction images derived from remote sensing multispectral data. *IEEE Transactions on Geoscience and Remote Sensing,* 29:16-20.

Sugita, M. and Y. Yasuoka. 1996. Land Cover Classification of East Asia Using Fourier Spectra of Monthly NOAA AVHRR NDVI Data. *IGARSS, 1996.*

SWAP 2006. Soil Water Atmosphere Plant. Available at http://www.swap.alterra.nl/. Accessed 12th May 2006.

Takeuchi, W., M. Tamura and Y. Yasuoka. 2003. Estimation of methane emission from West Siberian wetland by scaling technique between NOAA AVHRR and SPOT HRV. *Remote Sensing of Environment,* 85: 21-29.

Toung, T.P. and S.I Bhuian. 1999. Increasing water use efficiency in rice production: farm level perspectives. *Agricultural Water Management,* 40: 117-122.