RICE YIELD CLASSIFICATION USING BACKPROPAGATION NETWORK

*P. Saad[a], ***N. K. Jamaludin, **S. S. Kamarudin
***A. Bakri, and *N. Rusli

**Northern Malaysia University College of Engineering (KUKUM)*
Taman JKKK, Kubang Gajah, 02600 Arau, Perlis, Malaysia.
*** Faculty of Information Technology*
Universiti Utara Malaysia, Sintok, Kedah, Malaysia
****Faculty of Computer Science and Information System*
University Technology Malaysia, Skudai, Johor, Malaysia

puteh@kukum.edu.my, khairah jamal@yahoo.com
sakira@uum.edu.my, aryati@fsksm.utm.my
nursalasawati@kukum.edu.my

ABSTRACT

Among factors that affect rice yield are diseases, pests and weeds. It is intractable to model the correlation between plant diseases, pests and weeds on the amount of rice yield statistically and mathematically. In this study, a backpropagation network (BPN) is developed to classify rice yield based on the aforementioned factors in MUDA irrigation area Malaysia. The result of this study shows that BPN is able to classify the rice yield to a deviation of 0.03.

Key words: Backpropagation Network, classification, rice yield, pests, diseases and weeds.

[a]Corresponding Author Address: Puteh Saad, PKK&P, KUKUM, Taman JKKK, 02600 Arau, Perlis, Malaysia. Telephone : 604-9798285 Fax : 604-9798292 Email : puteh@kukum.edu.my

1.0 INTRODUCTION

Accurate early estimates of the final yield are essential for the determination of management information and to evaluate the results of management strategies in agriculture. Thus it is important to consider the accuracy of the classification methods. Statistical methods are generally used to explain the spatial variations and classify crop yield at field scale (Sudduth *et al.*, 1996). However statistical methods require a normal distribution of input variables, which is not always the case (Atkinson and Tatnall, 1997). Many complex mathematical models have been developed for modeling crop yield (Williams, 1989). Although these models produce excellent results, they require numerous input variables that are time consuming and expensive to obtain in the field. BPN is proven to alleviate the limitations mentioned by the following observations. It has successfully classified corn yield based on soil texture, topography, pH and some soil nutrient element (Sudduth *et al.*, 1988). The classification results were better than results produced using non-parametric statistical benchmark methods. Another contribution also to corn yield classification is the development of 4 BPN models using topographic features, vegetation indices and textural indices. When compared to the multiple regression method (MLR), commonly used as a benchmark, the BPN outperforms the MLR in terms of classification accuracy (Serele *et al.*, 2000). Another application of BPN in crop yield classification is to classify wheat yield using climatic observation data (Safa *et al.*, 2002). The yield was classified with a maximum error of 45-60 kg/ha two months before crop ripening.

The objective of our study is to classify rice yield based on diseases, pests and weeds as parameters using BPN. The data were obtained from MUDA rice growing area in Kedah, Malaysia. In the following section, we describe the backpropagation algorithm. Section 3.0 describes the methodology adopted to perform the study. In section 4.0 we presented the results and discussion. Section 5.0 provides the conclusion.

2.0 BACKPROPAGATION ALGORITHM

An Artificial Neural Network consists of a large number of processing elements called neurons. Each neuron is connected to other neurons by weighted links. Each neuron has a set of input links from other neurons, a set of output links to other neurons, a current activation level and an activation function. The learning process involves updating connection weights so that a network can efficiently perform a specific classification task. Classification performance is dependent on the linearity of feature space. Linear feature space eases classification task, however non-linear feature space makes classification difficult.

The most commonly used family of neural networks for pattern classification tasks is the feed-forward network of multi-layer perceptron (MLP) with backpropagation (BP) learning algorithm and Radial Basis Function Network (Haykin, 1994). MLP with BP has been adopted my many to solve classification tasks due to its effectiveness, hence we adopted it in this study.

The classification using BPN involves three stages: the feed forward of the input classification pattern, the calculation and backpropagation of the associated error and the adjustment of the connection weights (Haykin, 1999).

During feed forward computation, each input neuron (x_i) receives an input signal and broadcasts the signal to each of the hidden neurons $Z_1, \ldots, Z_p$. Each hidden unit then computes its activation and sends its signal (z_j) to each output neuron. Each output neuron (Y_k) computes its activation (y_k) to form the response of the net for the given input pattern.

Next section is the description of the data preparation and methodology adopted to perform the rice yield classification.

3.0 METHODOLOGY

The data for this study was taken from MUDA area in Kedah, Malaysia. There are 4 areas with 27 localities. With the introduction of MUDA irrigation scheme, farmers can plant paddy twice a year. The samples obtained from Muda Agricultural Development Authority (MADA) range from 1995 to 2001. There are 2 planting seasons for each year thus generating a total of 14 seasons represented by *S1* to *S14*. There are 3 parameters affecting the rice yield namely; diseases, pests and weeds. Diseases and pests, each consists of 12 types. However, for weeds there are 11 kinds, making a total of 35 input parameters. We reduced the parameters into 3 by taking the sum of effects produced by different kinds of diseases, pests and weeds. The final table consists of 3 columns representing factors affecting yield and the last column representing the yield obtained. Sample data are taken from an area that consists of five localities; Locality *A1*, *B1*, *C1*, *D1* and *E1*. Each table belongs to one locality, consisting of 14 rows, representing 14 seasons. An example of the sample is depicted in Table 1.

Table 1: Effects of Parameters on Rice Yield for Locality *A1*

Seasons	*Pests*	*Diseases*	*Weeds*	*Yield*
1/95 (S1)	63.63	189.11	91.85	3530
2/95 (S2)	73.48	283.49	120.72	4601
1/96 (S3)	139.93	148.76	72.86	3732
2/96 (S4)	36.60	331.93	53.45	3766
1/97 (S5)	54.07	25.57	122.75	2764
2/97 (S6)	74.70	21.80	161.34	3817
1/98 (S7)	101.47	55.02	93.60	3760
2/98 (S8)	125.53	229.23	136.16	4267
1/99 (S9)	285.11	96.42	307.18	4113
2/99 (S10)	127.81	71.76	161.66	4110
1/00 (S11)	165.96	28.61	232.77	4193
2/00 (S12)	101.47	114.08	176.10	3618
1/01 (S13)	184.86	41.89	323.79	4117
2/01 (S14)	88.93	73.59	428.19	3650

The following are the steps taken to classify the rice yield.

1. Data Preprocessing
2. Determination of BPN Parameters
    - Determination of learning rate
    - Determination of momentum
    - Determination of number of nodes in the hidden layer
3. Classification using the BPN

1. Data Preprocessing

Since the BPN is a data-driven model that abides the "garbage in – garbage out" principle, data of insufficient quality have a tendency to bring failure to the application. BPN is only as good as the input data used to train it, thus preprocessing is of paramount importance especially when analyzing real-life data to deliver a successful application (Waseem, 2002). In this case, the improved unit range technique is chosen as the normalisation technique to normalise the data into (0.1-0.9) range shown in Equation 1 (Saad, 2003).

$$x' = 0.9 * \left(\frac{x - x_{\min}}{x_{\max} - x_{\min}} \right) + 0.1 \tag{1}$$

where

x'	= normalised data
x	= raw data
x_{max}	= a maximum data value
x_{min}	= a minimum data value

Data are normalized into (0.1-0.9) range since the activation function for the backpropagation algorithm used in this study is of unipolar sigmoid, signified by Equation (2).

$$y_i = \frac{1}{1 + e^{-\lambda x_i}} \tag{2}$$

where

y_i	=	the output
λ	=	a constant chosen to be 1- indicating the slope of the gradient descent
x_i	=	the input parameter

2. Determination of BPN Parameters

The values for Neural Network parameters such as learning rate (α), momentum (β) values and number of hidden layers are problem dependent (Haykin, 1999). Thus in study, the values are determined empirically as reported in the Results and Discussion section.

3. Classification using the BPN

For each locality, data for 10 seasons are utilised for classification with the model parameters obtained in step 2.

4.0 RESULTS AND DISCUSSION

When data in Table 1 are normalised using the formula given in Equation (1), the following output is obtained as shown in Table 2.

Table 2: Normalized Data

Seasons	*Pests*	*Diseases*	*Weeds*	*Yield*
1/95 (S1)	0.108768	0.539483	0.102471	0.416984
2/95 (S2)	0.148404	0.843807	0.179511	0.900000
1/96 (S3)	0.415798	0.409377	0.051796	0.526946
2/96 (S4)	0.100000	0.900000	0.900000	0.545455
1/97 (S5)	0.070299	0.012156	0.184928	0.100000
2/97 (S6)	0.153314	0.100000	0.287906	0.573217
1/98 (S7)	0.261036	0.107116	0.107141	0.542188
2/98 (S8)	0.357853	0.668849	0.220713	0.818182
1/99 (S9)	0.900000	0.240609	0.677083	0.734349
2/99 (S10)	0.367027	0.161094	0.028876	0.732716
1/00 (S11)	0.520542	0.021959	0.478518	0.777899
2/00 (S12)	0.261036	0.297553	0.327294	0.464888
1/01 (S13)	0.596596	0.064779	0.721407	0.736527
2/01 (S14)	0.210575	0.166994	0.900000	0.482308

10 samples are then used to train the BPN with different α and β values. The deviation is used as the metric to determine the suitable α and β values.

$$deviation = \frac{\sum_{i=1}^{10} out_n - out_t}{10} \quad \text{for } out_n > out_t, \tag{3}$$

$$deviation = \frac{\sum_{i=1}^{10} out_t - out_n}{10} \quad \text{for } out_t > out_n, \tag{4}$$

where

out_n - the network output
out_t - the target output

The resulting deviation for different learning rate (α) and momentum (β) values are shown in Table 3.

Table 3: Deviation for Various α and β Values

β \ α	*0.1*	*0.2*	*0.3*	*0.4*	*0.5*
0.6	0.110107	0.109659	0.109186	0.108782	0.108294
0.7	0.010906	0.011114	0.110805	0.110393	0.110029
0.8	0.108046	0.107499	0.106927	0.110344	0.110099
0.9	0.107086	0.109286	0.108862	0.109484	0.109367

From Table 3, it is observed that the combination that produces the lowest deviation is when $\alpha = 0.3$ and $\beta = 0.8$. The corresponding α and β values are then used in training to determine the optimal number of nodes in the hidden layer.

Similarly, a deviation is computed using Equation (3) and Equation (4) in the determination of optimal number of nodes in the hidden layer, producing results shown in Table 4.

Table 4: Deviations for Different Number of Nodes in Hidden Layer

Architecture	*Deviation*
3-2-1	0.037355
3-3-1	0.031384
3-4-1	0.032050
3-5-1	0.031011
3-6-1	0.032162

From Table 4, the optimal number of nodes in the hidden layer is 5, since it produces the lowest deviation.

Table 5 illustrates the NN model parameters used for classification of the rice yield.

Table 5: Neural Network Parameters

Parameters	*Values*
Learning Rate	0.3
Momentum	0.8
Number Nodes in the Input Layer	3.0
Number of Nodes in the Hidden Layer	5.0
Number of Nodes in the Output Layer	1.0

Hence the above parameters are used to classify the rice yield given effects of diseases, pests and weeds. The results for an area with 5 locations are reported here. Table 6 shows the number of iterations and error produced when the algorithm has converged during the learning session. Table 7 to Table 11 show the network output and target output during classification. Graph of Rice Yield is plotted against seasons for 5 locations as depicted in Figure 1 to Figure 5.

Table 6: Iteration and Error reached During Convergence for Each Locality

Locality	*Number of Iteration*	*Error*
A1	2404	0.000999
B1	7646	0.000998
C1	1418	0.001000
D1	7631	0.001000
E1	4553	0.001000

For all the localities, it is observed that the network output tally with the actual output with the average mean deviation of 0.03. The mean deviation that was obtained is better than the similar study conducted on wheat (Safa *et al.* , 2002) and corn (Soele *et al.* , 2000) yield prediction. However, the samples data that we acquired from MADA is limited, only 14 seasons per locality. We hope in future research to be able to collect more data or look at other strategies to format the data so that the sample size increases.

Table 7: Classification Results for Locality *A1*

Network Output	*Actual Output*	*Deviation*
0.419159	0.416984	0.002175
0.944004	1.000000	0.055996
0.562095	0.526946	0.035149
0.550675	0.545455	0.005220
0.096919	0.000000	0.096919
0.569010	0.573217	0.004207
0.486479	0.542188	0.055709
0.829625	0.818182	0.011443
0.724650	0.734349	0.009699
0.766308	0.732716	0.033592
	Mean Deviation	0.031011

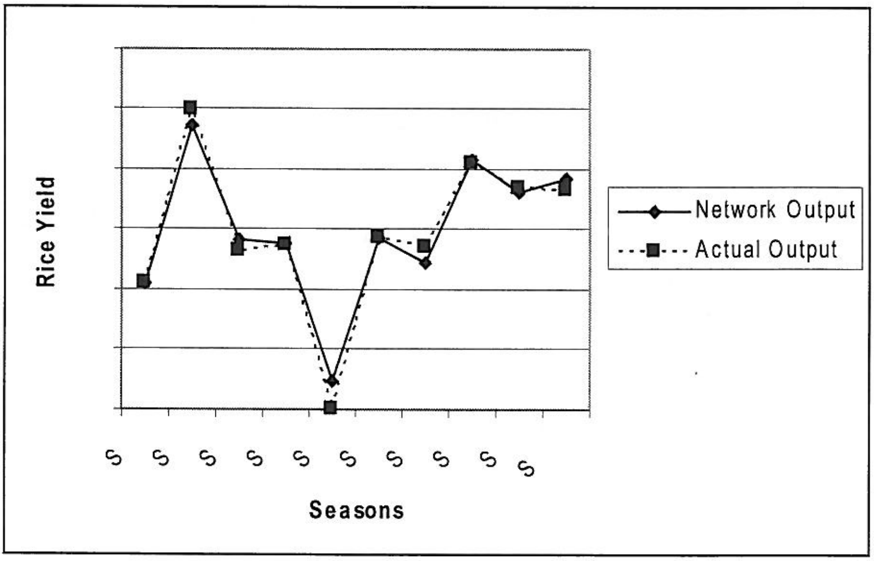

Fig. 1: Graph of Rice Yield vs. Seasons for Locality *A1*

Table 8: Classification Results for Locality *B1*

Network Output	*Actual Output*	*Deviation*
0.107411	0.100000	0.007411
0.896320	0.900000	0.003680
0.286910	0.252066	0.034844
0.806020	0.798347	0.007673
0.000017	0.012397	0.012380
0.993023	0.991736	0.001287
0.139611	0.130579	0.009032
0.976576	0.974380	0.002196
0.402718	0.292562	0.110156
0.396624	0.508264	0.111640
	Mean Deviation	0.030030

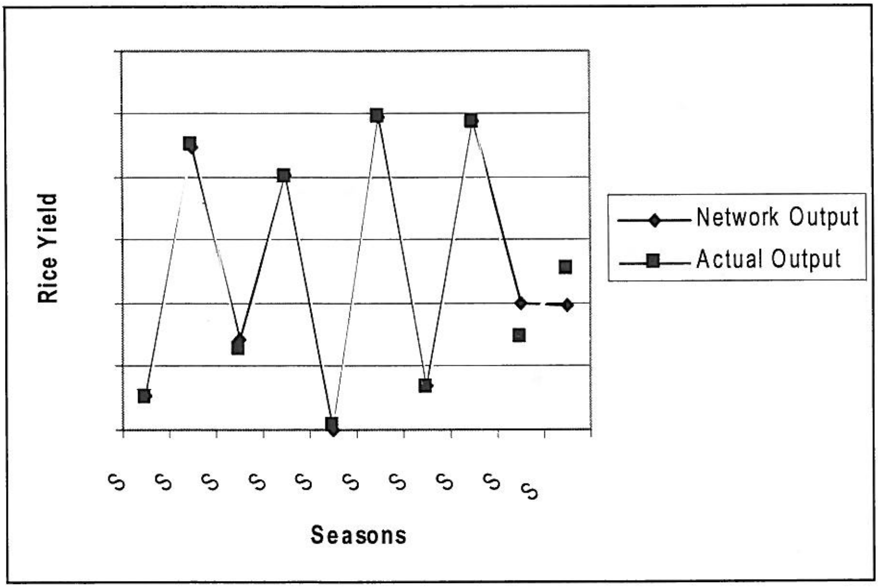

Fig. 2: Rice Yield vs. Seasons for Locality *B1*

Table 9: Classification Results for Locality *C1*

Network Output	*Actual Output*	*Deviation*
0.058718	0.102548	0.043830
0.899658	1.000000	0.100342
0.417551	0.384711	0.032840
0.581067	0.575513	0.005554
0.061388	0.000000	0.061388
0.850606	0.848975	0.001631
0.263302	0.302672	0.039370
0.940772	0.932256	0.008516
0.588916	0.596022	0.007106
0.692335	0.679925	0.012410
	Mean Deviation	0.031299

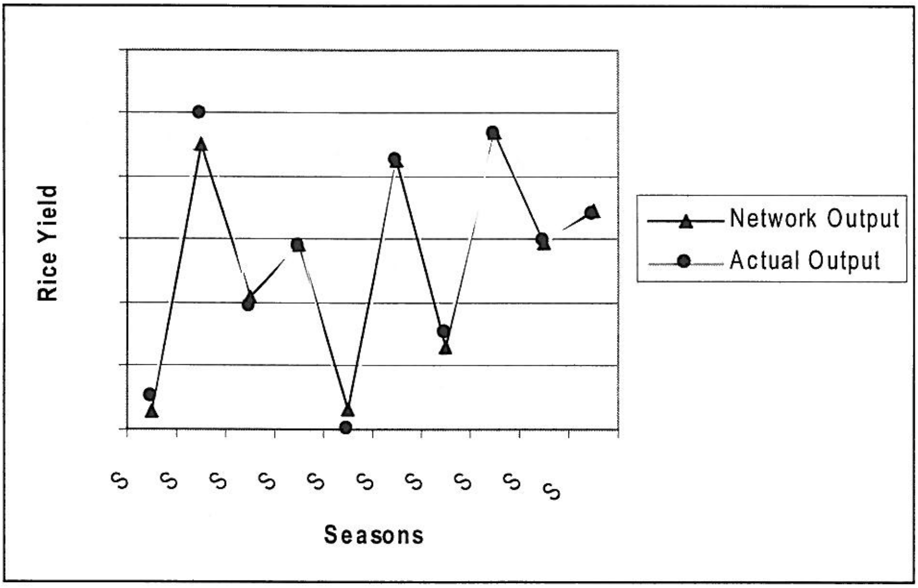

Fig. 3: Rice Yield vs. Seasons for Locality *C1*

Table 10: Classification Results for Locality *D1*

Network Output	*Actual Output*	*Deviation*
0.418405	0.433198	0.014793
0.725356	0.745518	0.020162
0.023335	0.000000	0.023335
0.614180	0.623482	0.009302
0.493217	0.503181	0.009964
0.720751	0.736264	0.015513
0.148133	0.060150	0.087983
0.435083	0.437247	0.002164
0.604854	0.550029	0.054825
0.443680	0.521689	0.078009
	Mean Deviation	0.031605

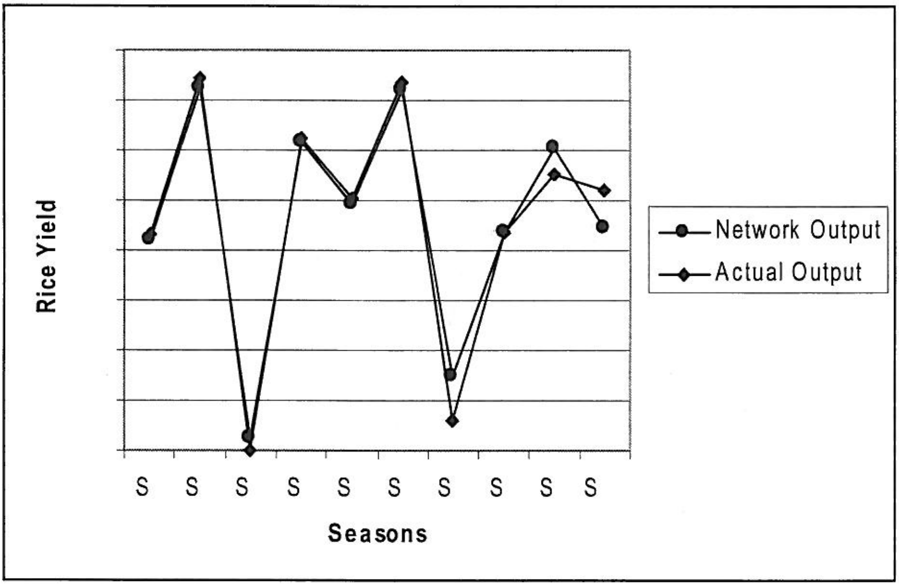

Fig. 4: Rice Yield vs. Seasons for Locality *D1*

Table 11: Classification Results for Locality *E1*

Network Output	*Actual Output*	*Deviation*
0.081368	0.000000	0.081368
0.428650	0.420411	0.008239
0.430374	0.444783	0.014409
0.737409	0.737243	0.000166
0.317190	0.320640	0.003450
0.894358	1.000000	0.105642
0.452117	0.461538	0.009421
0.803562	0.790556	0.013006
0.462060	0.440213	0.021847
0.612565	0.606245	0.006320
	Mean Deviation	0.026387

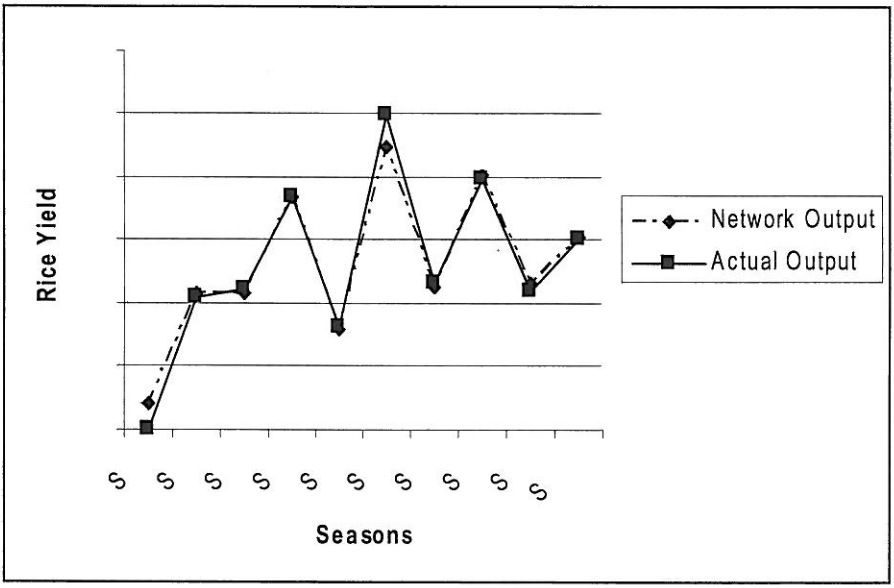

Fig. 5: Rice Yield vs. Seasons for Locality *E1*

Table 12: Average Mean Deviation

Locality	*Average Deviation*
A1	0.031011
B1	0.030030
C1	0.031299
D1	0.031605
E1	0.026387
Average Deviation for All Localities	0.030066

5.0 CONCLUSION

In this study, a backpropogation network (BPN) is developed to classify rice yield based on diseases, pests and weeds as parameters for MUDA irrigation areas Malaysia. The result of this study shows that BPN is able to classify the rice yield to a deviation of 0.03.

ACKNOWLEDGEMENT

We are indebted to Ministry of Science Technology and Environment (MOSTE), Malaysia for providing funds to carry out this research project under the IRPA Grant No: 04-02-06-0066 EA001 with the title of "Development of Intelligent Decision Support System for Rice Yield Prediction in Precision Farming". The duration of the grant is from 1/1/2003 to 31/12/2005.

REFERENCES

Atkinson, P., & Tatnall, A. (1997). Neural networks in remote sensing. *International Journal of Remote Sensing, 18*(4), 699-709.

Haykin, S. (1994). *Neural networks – A comprehensive foundation.* Toronto, Canada: Maxwell Macmillan.

Haykin S. (1999). *Neural Networks: A comprehensive foundation.* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Saad, P. (2003). T*rademark image classification approach using neural network and rough set theory*. Ph.D. Dissertation, University Technology Malaysia: In submission.

Safa, B., Khalili, A., Teshnehlab, M, & Liaghat, A. (2002). Prediction of wheat yield using artificial neural networks. In *15th Conference on Biometeorology/erobiology and 16th International Congress of Biometeorology*: Kansas City, MO. [Extended Abstract]. Retrieved July, 2003 from http://ams.confex.com/ms/15BioAero/15BioAero/abstracts/48422.htm

Serele, C. Z., Gwyn, Q. H. J., Boisvert, J. B., Pattey, E., McLaughlin, N., & Daoust, G. (2000). Corn yield prediction with artificial neural network trained using airborne remote sensing and topographic data. *In Proceeding of the International Conference Geoscience and Remote Sensing* Symposium,*1*, 24-28 July 2000. (pp. 384 –386). Honolulu, HI, USA:IEEE.

Sudduth, K., Fraisse, C., Drummond, S., & Kitchen, N. (1988). Integrating spatial data collection, modeling and analysis for precision agriculture. *First International Conference on Geospatial Information in Agriculture and Forestry*, *2*. June, 1998. (pp. 166-173). Lake Buena Vista, Florida: Altarum.

Sudduth, K., Fraisse, C., Drummond, S., & Kitchen, N. (1996). Analysis of spatial factors influencing crop yield. In P.C. Robert, R.H. Rust & W.E. Larson (Ed.). In *Proceedings of 3rd International Conference on Precision Agriculture*. June, 23 - 26, 1996. (pp. 129-140). Madison:ASA Misc. Publ.

Waseem, A. A. (2002). Data preprocessing for neural Nets. In *Proceeding of the International Conference on Artificial Intelligence in Engineering & Technology (ICAiET-2002)*, (pp. 683 – 686), Kota Kinabalu, Sabah, Malaysia.

Williams, J., Jones, C., Kiniry, J., & Spanel D. (1989). The EPIC crop growth model *Transactions of the American Society of Agricultural Engineer*, *32*, 497-511.