18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 July 2009
http://mssanz.org.au/modsim09

Modeling Fuel Consumption in Wheat Production Using Neural Networks

Majeed SAFA[1,2], **Sandhya Samarasinghe**[1],**Magdy Mohssen**[1]

[1] *Department of Environment Management, Lincoln University, New Zealand*

[2] *Agricultural Collages, Islamic Azad University (Saveh Branch), Iran*

Email: safam2@lincoln.ac.nz

Abstract: An artificial neural network (ANN) approach was used to model the fuel consumption of wheat production. This study was conducted over 18,316 hectares of irrigated and dry land wheat fields in Canterbury in the 2007-2008 harvest year. The data was collected from three different sources: questionnaire, literature review, and field measurements.

The developed model is capable of predicting fuel consumption in wheat production under different conditions. It can help farmers find the best practice to reduce their expenditure with minimum income reduction. This study investigates the potential for using neural networks to forecast fuel consumption, as compared to traditional regression models. This study examines more than 15 different technical, social and geographical inputs in wheat production in Canterbury to find the most important factors for model development. Finally, 8 variables: distance from nearest town, size of farm, farmer education, average size of paddocks, farmer experience, farmer age, proportion of wheat area(ha)/total area(ha), total tractor power(hp)/ farm(ha) were selected. The final model can predict fuel consumption to ±5.6 L/ha accuracy in wheat production by using different social, geographical, and technical factors.

Keywords: *Modeling. Neural Network, Fuel consumption, Wheat*

1. INTRODUCTION

Wheat is one of the eight food sources (wheat, rice, corn, sugar, cattle, sorghum, millet and cassava) which provides 70-90% of all calories and 66-90% of the protein consumed in developing countries. Globaly, wheat provides nearly 55% of the carbohydrate and 20% of the calories consumed globally (Breiman and Graur, 1995). Also, more than 40% of world grain is being fed as livestock. The wheat is cultivated under a wide range of climatic conditions. Most people consume wheat more than any other cereal grain (H Singh et al., 2007).Global production of bread wheat in 2003 was 557 million tons (Mt), with average yield of 2.68 t/ha (htp://apps.fao.org/). The world's major bread wheat producing areas are in northern China, northern India, northern USA and adjoining areas in Canada, Europe, Russia, Latin America and Africa (Kole, 2006). It covers around 25% of the total global area devoted to cereal crops. It is the staple food of nearly 35%of the world population and demand for wheat grows faster than that for any other major crop. In the last few decades, developed seed varieties have increased the yield, but in many areas using old growing systems, yields have stayed at less than desired levels. The forecasted global demand for wheat in 2020 varies between 840 and 1050 Mt (Kronstad, 1998; Rosegrant et al, 1995). To reach this target, global production will need to increase by 1.6% to 2.6% annually from the present production level of 560 Mt. Increases in realized grain yield have provided about 90% of the growth in cereal production since 1950 (Mitchell and Onco, 1997) and by the end of the first decade of the 21st century, most of the increase needed in world food production must come from higher absolute yields (Ruttan, 1993). For wheat, the global average grain yield must increase from the current 2.7 to 3.8 t/ha (Kole, 2006). This means that the average yield of wheat should increase about 40% in the short term.

Due to rising fuel price in the recent years, the price of production of crops that are dependent more on fuel has increased faster than that of other crops. On the other hand, farmers will select agricultural production with minimum fuel share. Moreover, during recent years production of ethanol from wheat have seen increased. The ethanol production from wheat is highly competitive and this new demand has raised the wheat price in global market exponentially.

Artificial neural networks (ANN) are becoming a common tool for modeling complex input- output dependencies (Maren et al, 1990; Samarasinghe, 2007). ANN is a structure composed of a number of interconnected units (neurons). Each neuron in the network is able to receive input signals, to process them, and to send an output signal. Each neuron is connected with at least one other neuron, and each connection is represented by a real number, called weight. the weights are adjusted so that the network attempts to produce the desired output. The main advantage of neural networks is that they are able to use some prior unknown information hidden in the data (but they are not able to extract it explicitly). The process of capturing the unknown information is called the learning of the neural network. A typical ANN is structured in three neural layers, an input layer, a hidden layer (some time more than one layer is necessary), and output layer. Information flows from the input layer to the output layer through the hidden layers. The learning process is improved over several iterations. As there are many free parameters (hidden layers and neurons, learning parameters genetic algorithms have bean use to optimise these parameters (Samarasinghe, 2007).

According to Melesse and Henley (2005), "Neural networks use machine learning based on the concept of self-adjustment of internal control parameters". Neural networks learn the relationship between the input parameters, the output parameters and the controlled and uncontrolled variables by studying previously recorded data (Soteris 2000). For this reason, the size of the data sample is so important, because without enough examples, neural networks cannot create correct relationships; this size can vary from few to sometimes thousands.

Some published papers (Bowers (1989), Riethmuller (1989) Pellizzi et al., 1988, and Serrano(2007)) are available for the determination of fuel consumption in agricultural products. However, no paper was found on modelling fuel consumption in agricultural production (especially by using social factors); therefore, the objective of this research is to determine and model fuel consumption in wheat production based on field operations by using a variety of factors. 41 Irrigated farms were selected randomly and information from each farm was collected through face to face interviews.

2. METHODOLOGY

The data was collected from three different sources: questionnaire, literature review, and field measurement. Different technical and social factors in wheat production operation such as tillage machinery, planters, fertilizer broadcasters, sprayers, irrigation, transportation, harvesting, farmer age, relevant experience, education, number of paddocks were determined. The number and duration of operations were investigated by questionnaire and personal interviews with farmers. Randomly selected farm owners completed the questionnaire. For use in the model, some of the parameters were converted from qualitative data to quantitative data, for example farmer's education was divided into five categories: primary school, high school, Diploma, undergraduate, and postgraduate. Also, after finishing the survey for better analysis, some indices were defined. The primary fuel input in crop farming is diesel. In this study, for estimating the amount of diesel consumption, the financial budget manual book (2008) was used. Undoubtedly, fuel consumption in farms is related to many factors. For this reason, farmers consume different amount of fuel under different condition in the same operation. But it was impossible to estimate fuel consumption for all of the operations in each farm. Also, most farmers did not have any idea (even estimation) of fuel consumption in their operations.

For predicting fuel consumption in wheat production, regression and ANNs methodologies were tested. Regression modelling was tested first for predicting fuel consumption. In the first step, the relationship between fuel consumption and variables were tested with linear and quadratic simple regression using the r^2 as a decision criterion. Then, a multiple regression model was developed.

$$Y=a_0+a_1V_1+a_2V_2+\cdots+a_nV_n+\epsilon \qquad Eq.1$$

where a0-an are the regression coefficient , V0-Vn are the independent variables, and ϵ is Error .The model was a linear form to represent linear relationships of dependent variable on the independent variables and interaction term between independent variables (Eq.1). This model has extensively been used in agricultural experiment evaluation with positive expected linear effects and negative quadratic effects (Colwell, 1994). A simple model with the highest r^2 is designed by using a combination of forward, backward and stepwise regression adjustments. Terms were maintained in the final model if they were significant at p=0.05 by the F test (Alvarez, 2009).

Neural networks can be successfully trained to describe separately the influence of energy sources, agricultural operation and indirect properties on energy consumption in wheat production. 8 variables were selected as inputs. 25% of samples were randomly selected for verification and 75% of samples were used for training. In this study, various aspects of ANNs were examined to find the best model. Preliminary trials indicated that two hidden layer networks yielded better results than one hidden layer networks. Based on this result, Fig.2 illustrates the NNs model that was developed to relate output to the input factors. In order to construct NNs, first the number of inputs and outputs should be determined. The next step is to define how to present the related data to the network. For this purpose, two different data formats are used, actual and incremental data. The reason for using such data is to provide the most relevant information to the network and then let the network do pattern matching among inputs and outputs. Sometimes, not only the straight data is provided, but also the differences between the present and previous status of the data, which is called incremental rate data (Kermanshahi, 2001). For simple nonlinear problems, one or few hidden neurons may be sufficient. However, for highly nonlinear problems involving many input variables, a large number of neurons may be necessary to correctly approximate the desired input-output relationship. Selecting number of neurons is an art more than science. When the number of hidden neurons is less than required, errors increase and correlation between inputs and outputs become weak and when the number of hidden neuron is more than required, problem of over learning sets in (Kermanshahi and Iwamiya, 2001). In this study, after several trails using genetic algorithm based optimisation, two hidden layers and 4 hidden neurons in each layer were used. The data were divided into training and validation. 75% of the data were used for training and one-fourth of the data was used for validation.

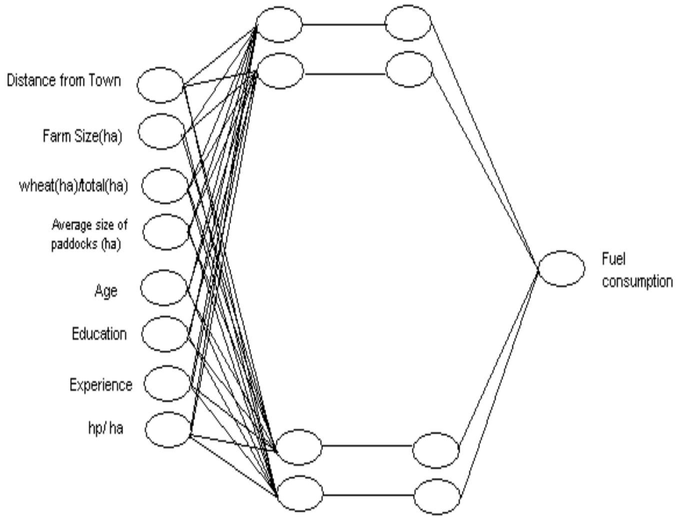

Fig. 1. Topology of feed forward neural network for Predicting Fuel consumption

After the network is initialized, it was trained for function approximation. During training, input pattern were randomly drawn and presented to NN, and weights were adjusted after each pattern. The learning rate controls the size of weight change in each epoch. A larger learning rate may lead to faster training answer but may oscillate around the minimum and never reach it. Consequently, a low learning rate of 0.01 was used in this model. The training was performed in batch form: calculating the error on all training patterns (epoch) with the current weights. The learning method that is used in study is Quick Prop. Quick Prop implicitly uses second order error to adjust smallest weights. Different functions such as sigmoid, Logistic and, Sine were tested and the choice of function was sigmoid of the general form:

$$Eq.2\ \sigma = [1 + e^{-u}]^{-1}$$

Weights were adjusted to minimize the mean square error between the outputs and the targets. The MSE, the most commonly used error indicator, of the prediction over all the training patterns for a network with one output neuron can be written as:

$$MSE = \frac{1}{2N}\sum_{i}^{N}(t_i - z_i)^2 \qquad Eq.3$$

Where t_i and z_i are the target and the predicted output for the *i*th training pattern, N is the total number of training patterns (Samarasinghe, 2007).

3. RESULTS AND DISCUSSION

ANNs showed that age and total tractor power per hectare (hp/ha) are the most important factors. Age had a strong impact on fuel consumption in wheat production due its direct and indirect effects through education and experience. Young farmers and more powerful tractors leading to reduced fuel consumption. Young farmers being more educated, it perfectly important of result; however, their limited experience, can dampens the effect.

High variability was observed in different inputs leading to a 3 fold difference in fuel consumption throughout the region (Table 1). The highest variability was in wheat area (ha)/total farm (ha), with an 18-fold difference. The distance from town was from 2 to 14 km. The range of total farm area was from 68 to 880 ha and the average size of paddocks was 11ha.

Table 1. Range of variability independent variables and fuel consumption

	Distance from Town(km)	Total Farm (ha)	Wheat area(ha) / total farm (ha)	Average size of paddocks(ha)	Age	Education	Experience	Total tractors power/total farm (hp/ha)	Fuel consumption
Mean	9	283	0.21	11	52	2	33	1.25	68.3
Max	14	880	0.53	27	65	4	52	3.53	96
Min	2	68	0.03	4	32	1	14	0.35	36
Std. Deviation	2.8	184.5	0.1	5.3	8.7	0.9	10.5	0.7	11.2
CV%	31	65	53	49	17	39	32	56	16

The average age of farmers was around 52 years and 57% of them studied at high school level. There were not any farmers with tertiary education. It was difficult to estimate farmers' experience, because majority of them grew up a farm and they have been involved in agricultural activities since childhood. The variability of the total tractor power/ total farm area (hp/ha) is from 3.53 to 0.35 hp/ha. It depends on the properties of tractors and equipment and also how frequently farmers use contractors. Data also show that more powerful tractors may reduce fuel consumption.

The highest standard deviation is for farm size (184.5) and lowest one is for wheat area/ total farm (0.1). Because of different scale, unit, and range of different variables, the coefficient of variation (CV= Standard deviation/mean) was used for better comparison. The CV is a useful statistic for comparing the degree of variation from one data series to another, even if the means are drastically different from each other. Table.1 shows the variables that are related to farm size have higher coefficient of variation. Age has the lowest CV (17%).

After using linear and quadratic models (p=0.05) it was observed that the relationship between independent variables and dependent variable was low. Positive significant association was observed between age and fuel consumption in both linear and quadratic model with r=0.32 for linear model and r= 0.41 for quadratic model. It means the farmers who are older; consume more fuel in their farms. Also there were negative significant correlation between education and fuel consumption; r was 0.36 and 0.39 for linear and quadratic models, respectively. It indicated that farmers with better education (young generation farmers) consume less fuel. The above results, also indicates evidence for a nonlinear relationship between fuel consumption and inputs. Other factors were not significantly correlated with fuel consumption. Surprisingly, there was no significant correlation between experience and fuel consumption but its correlation was higher than that of the rest of the factors. Multiple regression model could be fitted to fuel consumption data that accounted for around 31% of the variance (Figure 2A).

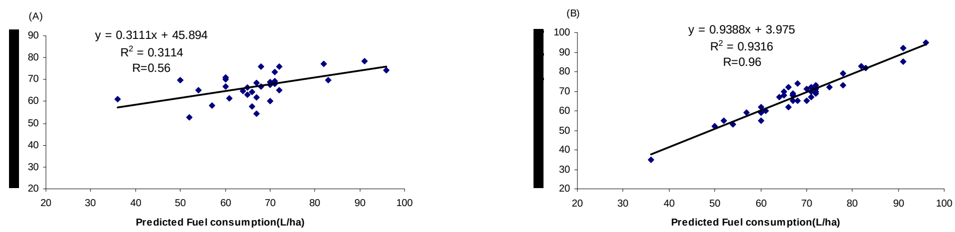

Fig.2. Relationships between observed and predicted fuel consumption (Training) (A) using lineal surface regression as estimation methodology (B) using an artificial neural network as estimation methodology.

In contrast, the ANN can predict fuel consumption in wheat production better than regression model and accounted for around 92% of the variance (Figure 2B). Figure 3 illustrates validation of multiple linear regression model and ANN model between observed and predicted values for fuel consumption. The r were observed as 0.34and 0.61 for multiple regression model and ANN, respectively.

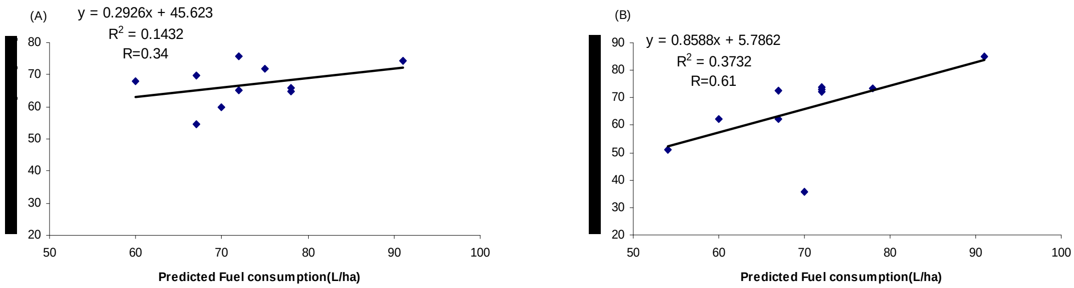

Fig.3. Relationships between observed and predicted fuel consumption (Validation) (A) using linear surface regression as estimation methodology (B) using an artificial neural network as estimation methodology.

Several models were developed and tested for fuel consumption in wheat production based on the eight input variables with genetic algorithm. The best set contained 8 factors: farmer age, relevant experience, farmer education, size of farm, distance of farm from town, total tractors power per hectare (hp/ha), average size of paddocks, and proportion of wheat area. A sensitivity study of model showed that total tractors power per hectare (0.0082) is the most important factor in this model. This was followed by farmer age (0.0079), experience (0.0039), size of farm (0.0033), education (0.0028), and size of paddocks (0.0021). The final model can predict fuel consumption to ±5.6 L/ha as shown in Figure 4.

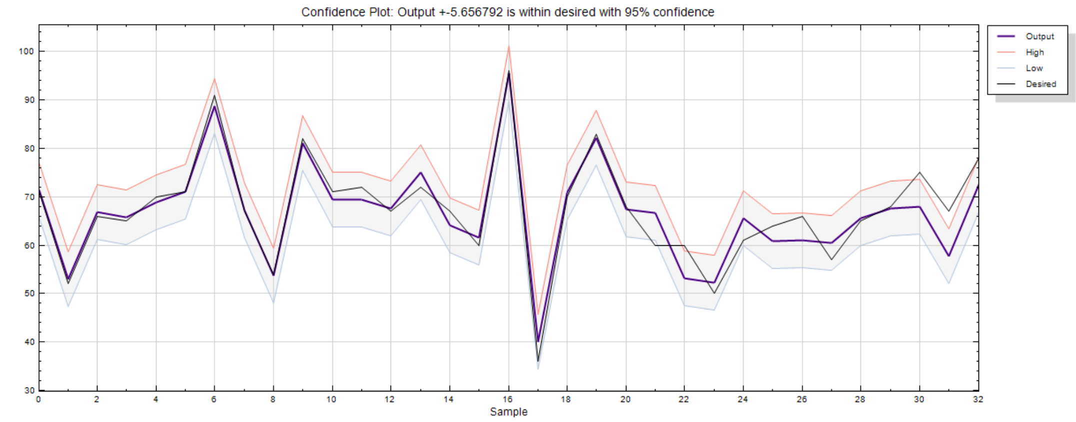

Fig.4. Observed and predicted fuel consumption using an artificial neural network with 95% confidence.

The MSE between predicted and measured For ANN model was around 0.0093. Ideally, the MSE values should be close to zero, indicating that, on average, there is no difference between the simulated and observed values.

4. CONCLUSION

The ANN model can help predict fuel consumption in different farms. If the model covers all production and activities on different farms, it can help farmers and government find out how to reduce fuel consumption on farms. Increasing the number of samples and testing more variables for a period of time, at least 5 years, can help design a model to predict the trend of fuel consumption in agricultural production under different circumstances. The ANN approach allowed a better predicting of fuel consumption than multiple regression when applied at a regional level. These results may be considered as a first step in developing methods suitable for prediction fuel consumption for the whole Canterbury Region by using social, technical, and geographical factors together. The methodology may be applied to other cropping areas of the World and to different crops.

REFERENCES

Alvarez, R. (2009). Predicting average regional yield and production of wheat in the Argentine Pampas by an artificial neural network approach. *European Journal of Agronomy*, 30(2), 70-77.

Bowers, C. G. J. (1985). "Southeastern tillage energy data and recommended reporting." Transactions of the ASAE 28(3): 731-737.

Breiman, A. Graur, D. 1995. Wheat Evaluation. Israel J Plant Sci. 43:58-95.

Colwell, 1994 J.D. Colwell, Estimating Fertilizer Requirements. A Quantitative Approach, CAB International, UK (1994) p. 259.

Kermanshahi, B. and Iwamiya,H.2001, Up to Year 2020 Load Forecasting Using Neural Nets. *Electrical and Energy Systems* 24 (2002), 789- 797.

Kole, C. 2006. Cereals and millets. Berlin ; New York, Springer.

Kronstad, W. E. 1998. Agricultural Development and Wheat Breeding in the 20th century. Proc. Of the 5th Int. Wheat Conf. Ankara. Turkey.

Lincoln University. (2008). Financial budget manual. Lincoln, N.Z.: Dept. of Farm Management and Rural Valuation, Lincoln College.

Maren, A .J. Harston. C. T and Pap, R .M. 1990. Handbook of Neural Computing Applications. Academic Press, San Diego.

Melesse, A. M., and Hanley, R. S. (2005). Artificial neural network application for multi-ecosystem carbon flux simulation. *Ecological Modelling*, 189(3-4), 305-314.

Mitchell, D. O.and Onco, M.D. 1998. The world Food Outlook. Cambridge University Press.

Murphy, J. D. and N. M. Power (2008). "How can we improve the energy balance of ethanol production from wheat?" *Fuel* 87(10-11): 1799-1806.

Ozkan, B., Akcaoz, H., & Fert, C. (2004). Energy input-output analysis in Turkish agriculture. *Renewable Energy*, 29(1), 39-51.

Pellizzi, G., Guidobono Cavalchini, A., Lazzari, M., & Commission of the European Communities. (1988). Energy savings in agricultural machinery and mechanization. London ; New York: Elsevier Applied Science.

Riethmuller, G. (1989). Draft Requirements of Tillage Equipment in the Western Australian Wheat belt. Conference on Agricultural Engineering 1988: An Australasian Conference to Celebrate the Australian Bicentennial.

Ruttan,V.W. 1993. Research to Meet Crop Production Needs: Into the 20th Century . International Crop Science Congress . CCSA. Madison , Wisconsin, USA. Pp 3-10.

Samarasinghe, S. (2007). Neural networks for applied sciences and engineering : from fundamentals to complex pattern recognition. Boca Raton, FL: Auerbach.

Serrano, J. M., J. O. Peça, et al. (2007). "Tractor energy requirements in disc harrow systems." *Biosystems Engineering* 98(3): 286-296.

Singh, H., A. K. Singh, et al. (2007). "Energy consumption pattern of wheat production in India." *energy* 32: 1848-1854.

Soteris, A. (2000). Applications of artificial neural-networks for energy systems. *Applied Energy* 67 (2000) 17-35.