Modelling Energy Use and Fuel Consumption in Wheat Production Using Indirect Factors and Artificial Neural Networks

Majeed Safa[1,*] and Sandhya Samarasinghe[2]

[1] Department of Agricultural Management and Property Studies, Lincoln University, New Zealand
[2] Department of Environmental Management, Lincoln University, New Zealand
Majeed.Safa@lincoln.ac.nz

Abstract. This study was conducted over 35,300 hectares of irrigated and dry land wheat fields in Canterbury, New Zealand in the 2007-2008 harvest year. The Artificial Neural Network models (ANNs), after examining more than 140 several direct and indirect parameters, can predict energy use and fuel consumption based on farm conditions, farmers' social considerations, farm operation, machinery condition and farm inputs, arable farms in Canterbury with an error margin of ±12% (± 2900 MJ/ha) and ±8% (± 5.6 l/ha), respectively.

Keywords: Modelling, Energy consumption, Fuel consumption, Neural Networks, Wheat.

1 Introduction

Energy is one of the important elements in modern agriculture. Energy consumption in agriculture has been increasing in response to the limited supply of arable land, increasing population, technological changes, and a desire for higher standards of living [1-3]. Some studies show that there is a positive relationship between energy usage and productivity [2, 4-9]. Also, there is a significant relationship between energy output and weather, price, yield, and technology [10].

In agriculture, fuel is used for transportation, running tractors and machinery, and irrigation. Due to rising fuel prices in recent years, the price of producing crops that are more dependent on fuel has increased faster than that of other crops. Under these circumstances, farmers will select agricultural production with a minimum fuel share [11]. In general, New Zealand farmers practise a form of 'industrialized' agriculture that relies on relatively high inputs of fossil fuels [12].

Energy modelling is an interesting subject among engineers and scientists who are concerned with energy production and consumption and environmental impacts [13, 14]. In the past, regression analysis was the common modelling technique used in energy studies. However, recently, neural networks (NN) have been increasingly used

* Corresponding author.

T. Huang et al. (Eds.): ICONIP 2012, Part II, LNCS 7664, pp. 25–32, 2012.

in energy studies [15]. Due to the ability of neural networks to model complex nonlinear systems in a flexible and adaptive manner, NN are increasingly being used for solving various problems in science and engineering [16]. NN have developed into a powerful approach that can approximate any nonlinear input-output mapping function to any degree of accuracy in an iterative manner. NN have many attractive properties for modelling complex production systems, and some of these are: universal function approximation capability, resistance to noisy or missing data, accommodation of multiple non-linear variables with unknown interactions and good generalization ability [17].

In the processing of inputs by the network, each neuron in the first layer (hidden layer) processes the weighted inputs through a transfer function to produce its output. The transfer functions may be a linear or a nonlinear function. There are several transfer functions, such as Logistic, Hyperbolic-tangent, Gussian, and Sine. The output depends on the particular transfer function used. This output is then sent to the neurons in the next layer through weighted corrections and these neurons complete their outputs by processing the sum of weighted inputs through their transfer functions. When this layer is the output layer, neuron output is the predicted output.

Several methods of error estimation have been proposed. The Mean square error (MSE) over all training patterns (Eq. 1) is the most commonly used error indicator. MSE is very useful to compare different models; it shows the networks ability to predict the correct output. The MSE can be written as:

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

where t_i and $z_i$are the actual and the predicted output for the i^{th} training pattern, and N is the total number of training patterns [18]. Root mean square error (RMSE) is another error estimation, which shows the error in the units of actual and predicted data.

When the number of variables is notably high, especially when there are limited number of samples, data reduction is useful. Also, when some input variables correlate with one another, another problem that is called multicollinearity, will appear. Correlation between inputs reduces the chance of having a unique solution [18]. Principle component analysis (PCA) is one of the bestmethods to select the most important uncorrelated variables. PCA uses the mean and variance of each input variable and the covariance between variables to create a covariance matrix (COV) [18] and transforms the COV to obtain independent components that are linear summations of original inputs.

2 Research Methods

2.1 Research Condition

This study was conducted over 35,300 hectares of irrigated and dry land wheat fields in Canterbury, New Zealand, in the 2007-2008 harvest year, which reported 87% of the wheat area and 66% of arable area harvested in New Zealand [19, 20]. 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 and 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.

2.2 Energy Use

Energy consumption is defined as the energy used for the production of wheat until it leaves the farm. The energy inputs estimated in this study are those that go into on-farm production systems before the post-harvest processes. In this study, energy consumption in wheat production was analyzed based on direct and indirect energy sources including human, fuel, fertilizers, pesticides, machinery, and seed.

In this study, energy conversion coefficients were investigated and selected carefully. In addition to direct and indirect energy inputs, there are other factors, such as technical, social, geographical, and financial factors, which may influence energy consumption indirectly. A wide range of factors, around 140, including farmers' social status, age of tractors and equipment, power of tractors, number and size of paddocks, and yield were studied. Moreover, these indirect factors and energy inputs were examined to design the model to predict energy consumption in wheat production. Involving indirect factors in energy prediction may help reduce energy consumption with minimum cost and minimum reduction to farmers' income.

On Average, energy consumption in wheat production in Canterbury was estimated about 22,566 MJ/ha. Of this, 36% was direct energy in the form of diesel at 3,121 MJ/ha and electricity at 4,870 MJ/ha in wheat production according to energy sources. Fertilizer is ranked first with 47% (10,651 MJ/ha) and electricity is ranked second with 22% (4,870 MJ/ha) of total energy.

2.3 Fuel Consumption

The best way to measure fuel consumption was by field measurements; filling the tractor tank twice, before and after each operation; however, for some operations, such as harvesting this method is too difficult. Due to different soil and machinery conditions and the large number of farms in this study, fuel estimations were derived from the Financial Budget Manual [21] of Lincoln University.

On average, 65.3 l/ha of fuel was consumed in wheat production. For better understanding, farm operations were classified into five clusters: tillage, drilling, fertilizer distributing, spraying and harvesting. Tillage was ranked highest in the systems, with 45% of total fuel consumption. Between the different tillage operations, ploughing and other primary and heavy operations used fuel more than secondary tillage operations.

2.4 Data Selection

Finding appropriate variables is the first step of model creation. For use in the model, it was necessary to select a limited number of variables. In this study, for variable redaction, principle component analysis (PCA) was used. There were approximately 140 variables that had the potential to be an input in the final model. Variables were selected that had no significant correlation and the highest link with energy use and fuel consumption separately.

After processing original data and input reduction, five variables from the PCs with the threshold cumulative variance of around 72% were selected to use as variables in the energy ANN model, which are crop area (ha), farmer's education, nitrogen consumption (kg), phosphate consumption (kg), and irrigation frequency.

After the PCA process, eight variables with the threshold cumulative range of around 77.6 % were selected to use as variables in the fuel ANN model. They comprised the size of wheat area (ha), farmer's education, number of sheep, herbicide and insecticide consumption (kg), phosphate consumption (kg), number of passes of plough and age of sprayer.

2.5 Modelling

NNs can be successfully trained to describe the influence of energy sources, agricultural operations, and indirect factors on energy consumption in wheat production. The sample size used in this study was 40 farms. Initially, a sample of 30 farms was randomly selected for training, and the remaining sample of 10 farms was used for validation.

After several trials by using Peltarion Synapse software, a modular neural network with two hidden layers was selected. In the modular network structure, the model is characterized by a series of independent neural networks after the input layer, which operates on the inputs to achieve some subtasks of the task the network expects to perform. These subtasks are trained separately with different examples from the sample and their outputs are summed in the output layer. The structure of the model prepares the network to use simultaneously different model functions for the data.The Peltarion Synapse softwareprovides a genetic algorithm-based optimization to examineand find the best various aspects of ANNs.

After the network was initialized, it was trained for function approximation. During training, input patterns were drawn randomly and presented to the NN; weights were adjusted after each epoch. The learning rate controls the size of weight change in each epoch. A higher learning rate may lead to faster training answer; but the weights may oscillate around the minimum and never reach it [18]. Consequently, a low learning rate of 0.01 was used in this model. The Quick Prop was used as the learning method; because, it was fast in reducing the error and finding the best model. Quick Prop implicitly uses the second derivative of error to adjust weights.

Different functions, such as hyperbolic tangent (tanh), logistic sigmoid, Gaussian Bell, linear and Sine, were tested to find best combination of functions and model structures. In the final energy model, linear function was applied for the input layer, hyperbolic tangent (tanh) function was selected for first hidden layer and logistic function was applied for second hidden layer.

In the final fuel model, logistic function was selected for the input layer and the first hidden layer; the hyperbolic tangent function was applied for the second hidden layer; and linear function was applied for the output layer.

3 Results

3.1 Energy Model

After testing different learning algorithms, neuron activation functions, and network structures, a modular network with two hidden layers were developed. As shown in Figure 1, after input layer the modular network is separated into two parts. The number of neurons in the first and second layers of the top part was optimized using a genetic algorithm optimizer that indicated 2 and 17 neurons for the first and second hidden layers, respectively. But, in second part, the number of neurons was optimized to be 18 and 17 for first and second hidden layers, respectively. The results are combined at the output layer to produce the final output, the energy consumption.

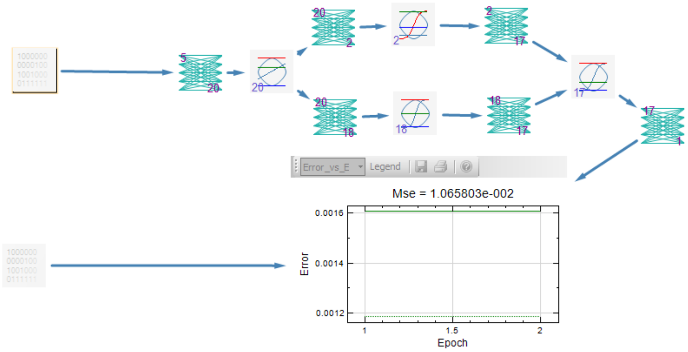

Fig. 1. Structure of modular network and number of neurons in each layer

The NN model after 100 iteration achieved the best results with scaled MSE= 1.06 E-2 (inputs and outputs were scaled between -1 and +1 for the neural networks model). The actual RMSE of the final NN model was estimated to be 1230 MJ/ha on validation data. It was the lowest RMSE between several NN models examined in this study. As shown in Figures 2, energy consumption estimated by the NN accounted for 90% and 96% of the actual variability in energy in training and validation data, respectively. Correlation between observed and predicted energy consumption is very high with r^2 =0.81 and r= 0.90 (training).

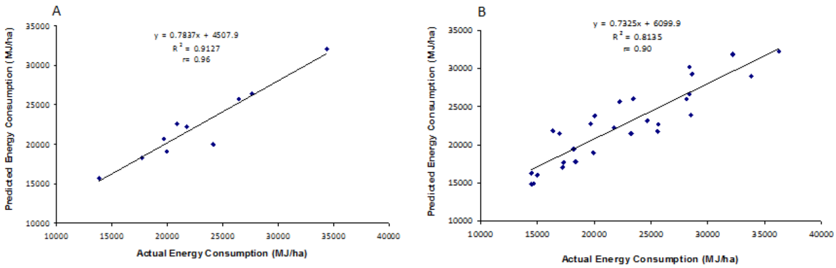

Fig. 2. Relationships between the observed and NN model predicted energy consumption, A) validation data. And B) Training data

The final model predicted energy consumption with an error margin of around ±6000 MJ/ha (training data) and an error margin of around ±2900 MJ/ha (validation data).

3.2 Fuel Model

As shown in Figure 3, a modular network with two hidden layers, was developed. Surprisingly, the numbers of neurons in the first and second layers of the both parts were similar, which were then optimized using a genetic algorithm optimizer that indicated 19 and 9 neurons, respectively. The results were combined in the output layer to produce the final output, fuel consumption.

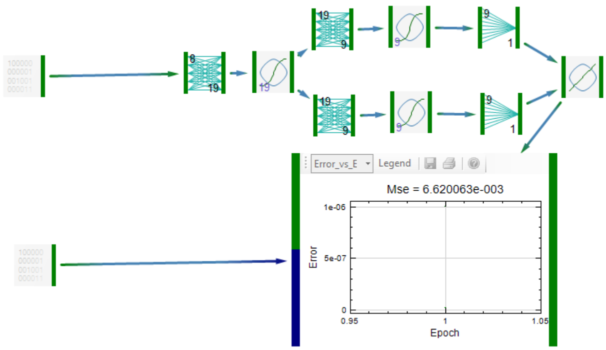

Fig. 3. Structure of modular network and number of neurons in each layer in fuel model

After 2000 iterations the ANN model achieved the best results with a scaled MSE of 6.6 E-3. Figure 4 illustrate the training and validation of the ANN model for observed and predicted values for fuel consumption. The r^2 was 0.78 and 0.54 for training and validation of the ANN model, respectively.The final model can predict fuel consumption to ±5.6 l/ha.

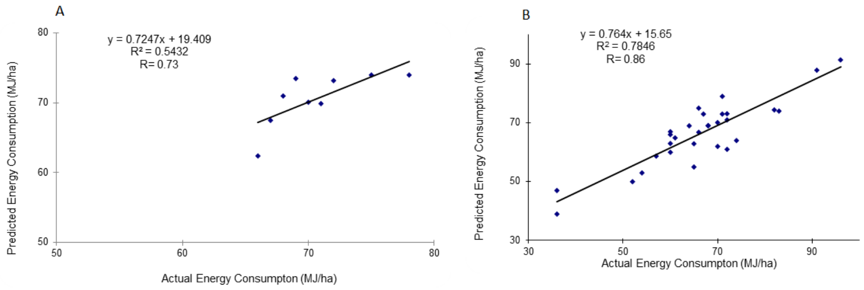

Fig. 4. Relationships between observed and predicted fuel consumption, A) validation data, and B) training data

4 Conclusion

There are several uncontrolled factors which could influence energy and fuel consumption in crop production. The final model can predict energy fuel consumption in wheat production with acceptably small error. It is noticeable that some of the variables in the final models are fixed and cannot be changed, or affect energy consumption indirectly. Therefore, the next step should be to explore in detail the links between input variables and energy and fuel consumption.

The result of this study shows the ability of NN model to predict energy use and fuel consumption in wheat production by using heterogeneous data. Using dissimilar variables, such as farm conditions and social factors would improve the ability of decision makers to look at the problem from different perspectives. These results may be considered as a first step in developing methods suitable for predicting technical factors on farms using social, technical and geographical factors together. Given the findings of this study, the most significant areas for improving overall modelling and conserving energy and fuel are as follows:

- Investigating on indirect factors, especially sociocultural parameters such as farmers' age and farmers' education, is the important step in farm management studies. As shown in this study, some of indirect factors can be used to predict technical factors on agricultural production.
- Increasing the number of samples and testing more variables for a longer period of time, at least five years, can help analyze trends in energy consumption in agricultural production in different regions under different conditions. Continuing this study over a period of time would help compare oil prices, wheat and other crop prices, and their effects on energy consumption and technology use on farms.
- Developing an NN model to estimate energy use of all products of each farm to find the most energy efficient combination of different agricultural production (rotation) and agricultural operation under different conditions. To develop this complex model, several farms must be involved and their production and operation must be investigated carefully.

References

1. Kizilaslan, H.: Input-Output Energy Analysis of Cherries Production in Tokat Province of Turkey. Appl. Energy 86, 1354–1358 (2008)
2. Hatirli, S.A., Ozkan, B., Fert, C.: Energy Inputs and Crop Yield Relationship in Greenhouse Tomato Production. Renew. Energy 31(4), 427–438 (2006)
3. Manaloor, V., Sen, C.: Energy Input Use and CO2 Emissions in the Major Wheat Growing Regions of India. In: International Association of Agricultural Economists Conference 2009, Beijing, China (2009)
4. Karkacier, O., Gokalp Goktolga, Z., Cicek, A.: A Regression Analysis of the Effect of Energy Use in Agriculture. Energy Policy 34(18), 3796–3800 (2006)
5. Karkacier, O., Gokalp Goktolga, Z.: Input-Output Analysis of Energy Use in Agriculture. Energy Conv. Manag. 46(9-10), 1513–1521 (2005)
6. Baruah, D.C., Bora, G.C.: Energy demand forecast for mechanized agriculture in rural India. Energy Policy 36, 2628–2636 (2008)
7. Singh, H., Mishra, D., Nahar, N.M.: Energy Use Pattern in Production Agriculture of a Typical Village in Arid Zone-Part III. Energy Conv. Manag. 45, 2453–2472 (2004)
8. Outlaw, J.L., Collins, K.J., Duffield, J.A.: Agriculture as a producer and consumer of energy, xv, 345 pages. CABI Pub., Wallingford (2005)
9. Smil, V.: Energy in Nature and Society: General Energetics of Complex Systems, xi, 480 pages. The MIT Press, Cambridge (2008)
10. Ozkan, B., Akcaoz, H., Fert, C.: Energy input-output analysis in Turkish agriculture. Renew. Energy 29(1), 39–51 (2004)
11. Safa, M., Samarasinghe, S., Mohssen, M.: Determination of Fuel Consumption and Indirect Factors Affecting It in Wheat Production in Canterbury. Energy 35(12), 5400–5405 (2010)
12. Wells, C.: Total Energy Indicators of Agricultural Sustainability: Dairy Farming Case Study, vii, 81 pages. Ministry of Agriculture and Forestry, Wellington (2001)
13. Al-Ghandoor, A., et al.: Residential Past and Future Energy Consumption: Potential Savings and Environmental Impact. Renew. Sust. Energy Rev. 13(6-7), 1262–1274 (2009)
14. Tester, J.W.: Sustainable Energy: Choosing Among Options, xxiii, 846 pages. MIT Press, Cambridge (2005)
15. Sözen, A.: Future Projection of the Energy Dependency of Turkey Using Artificial Neural Network. Energy Policy 37(11), 4827–4833 (2009)
16. Jebaraj, S., Iniyan, S.: A Review of Energy Models. Renew. Sust. Energy Rev. 10(4), 281–311 (2006)
17. Hagan, M., Demuth, H., Beale, M.: Neural Network Design. PWS Publishing Company, Boston (2002)
18. Samarasinghe, S.: Neural Networks for Applied Sciences and Engineering: from Fundamentals to Complex Pattern Recognition, xx, 570 pages. Auerbach, Boca Raton (2007)
19. Statistics New Zealand, Agricultural Production Statistics (Final): June 2007. Wellington, Statistics New Zealand (2008)
20. Statistics New Zealand, Agricultural Production Statistics: June 2007 (Provisional) Grain crops - Media Release, Increased Yields from Wheat and Barley Harvests. Wellington, Statistics New Zealand (2008)
21. Lincoln University, Financial Budget Manual. Dept. of Farm Management and Rural Valuation, Lincoln University, Lincoln (2008)