Journal of Animal and Veterinary Advances 11 (14): 2503-2508, 2012
ISSN: 1680-5593

Comparison of Regression and Artificial Neural Network Models of Egg Production

[1,2]B.Y. Wang, [2]S.A. Chen and [2]S.W. Roan
[1]Livestock Management Division, Livestock Research Institute, Council of Agriculture, Mu-Chung Road, Hsinhua, 71246 Tainan, Taiwan
[2]Department of Animal Science, National Chung Hsing University, 250, Kuo Kuang Road, 40227 Taichung, Taiwan

Abstract: This study compared the relationship between egg production and the number of pullets, laying hens, culling birds and molting birds in Taiwan through Traditional Regression Methods and ANN (Artificial Neural Network) Models. Egg production data and the number of laying hens associated with each data set were gathered from the National Animal Industry Foundation for dates between January 2001 and March 2011, totalling 123 data sets. The final regression equations were: Traditional Regression Model: case = 2.77 + 0.696 Pmonth − 0.00621 $Pmonth^2$ − 0.00163 pullet + 0.0025 laying, R^2 = 0.699; ANN Model: case = 2.82+0.113 Pmonth − 0.00871 $Pmonth^2$ − 0.00157 pullet + 0.0024 laying, R^2 = 0.965. These results show that the ANN Method is more accurate than traditional Regression Models for predicting egg production in Taiwan.

Key words: Artificial Neural Network (ANN), egg production, regression, birds, Taiwan

INTRODUCTION

Taiwan produced between 80,000 and 90,000 cases of eggs per day (12 kg/case) in 2011. Over production and changes in egg consumption immediately affect egg prices because the egg is a basic ingredient in many dishes. For example, egg consumption fell approximately 10% in restaurants that provide meals for students during the Winter and Summer school vacations and egg prices were thus affected during those times. Such an impact is rare for other agricultural products.

The Artificial Neural Network (ANN) Model takes into account nonlinearities in the relationship between input and output information. The advantages of ANNs include knowledge plasticity of changing inputs and outputs, fault tolerance and interpolation capabilities (Zhang *et al*., 2007).

Researchers have compared ANN technologies and other simulation models of poultry growth. Roush *et al.* (2006) compared the Gompertz model and artificial neural network models of the performance of broiler growth using the Mean Square Error (MSE), Mean Absolute Deviation (MAD), Mean Absolute Percentage Error (MAPE) and bias as assessment criteria. The lowest MSE, MAD and bias were observed in the training data using the neural-developed neural network model while the lowest MSE, MAD and MAPE were observed in the validation data using the General Regression Neural Network (GRNN) Model. However, the lowest bias was observed using the neural-developed neural network model. When measured by bias, the Gompertz equation underestimated the performance of broiler growth whereas the neural-developed and genetic-developed neural network models produced little or no overestimation of the observed BW responses.

Ahmadi *et al.* (2008) assessed various feed compositions and broiler performance using the Group Method of Data Handling-type (GMDH) neural network models. The results indicated that the European Efficiency Factor (EFF) for training and validation during three periods, starting, growing and finishing, possessed good unity and the lowest MSE, MAD and MAPE values during the growing period with relatively low bias. The training data had lower bias than the validation data. Golian and Ahmadi (2008) used data to evaluate several egg production curves for a laying hen flock using the Neural Network (NN) Model. R^2, adjusted R^2, Mean Square Error (MSE), Residual Standard Error (RSE), Mean Absolute Percentage Error (MAPE) and bias were used to calculate the goodness of fit for the NN Model. The NN model adjusted R^2 value for the first and second egg production cycles were 0.999 and 0.998, respectively.

Ahmad (2009) compared the growth performance of broilers and guinea fowl by using an Artificial Neural

Corresponding Author: S.W. Roan, Department of Animal Science, National Chung Hsing University, 250, Kuo Kuang Road, 40227 Taichung, Taiwan

Network Model, the Gompertz Model and other Nonlinear Regression Models. The R^2 value of the BP3-NN Model was very close to 1.0 in broilers and the regression values were similar to real values. The Gompertz Model underpredicted growth in guinea fowl while the Logistic Model overpredicted guinea fowl growth in the 1st 3 weeks (0~2 weeks) and in the 5th week. The Ward-5 NN and GRNN Models produced similar R^2 values of approximately 0.96. The Ward-5 NN slightly overpredicted R^2 except in the 4th, 8th and 9th weeks. The GRNN slightly underpredicted R^2.

Savegnago *et al.* (2011) investigated the use of Neural Network Models on the egg production curves of White Leghorn hens by applying two different approaches: a Non-linear Logistic Model and two different Neural Network Models (Multilayer Perceptron (MLP) and radial basis function). Savegnago and his team used the mean absolute deviation, mean square error and R^2 to evaluate the fit of the models. The MLP model had the best fit in the test and validation phases, confirming that MLPs can be used as an alternative tool to fit egg production data.

This study compared the relationship between egg production and the number of pullets, laying hens, culling birds and moulting birds in Taiwan using Traditional Regression Methods and ANN Models. Researchers adopted above variables and simulated more than one variable combination to obtain the functional dependence of egg production which the government may use to guide and improve the poultry industry.

MATERIALS AND METHODS

Data collection (resources): Egg production data and the number of pullets, laying hens, culling birds and moulting birds in each data set were gathered from the National Animal Industry Foundation (2009) for dates between January 2001 and March 2011. In total, 123 data sets were collected.

Data selection

Time lag: Researchers introduced a time lag in the study because the pullets were too young to lay eggs. Pullets must mature to about 5 months before they can begin laying eggs which affects egg production. As such, researchers set 2 months of data in the study; 1 month represented pullets for the current study month and the other month represented laying hens 5 months ahead of the current study month. Researchers then used the total number of pullets for the current month to predict the number of laying hens and eggs produced 5 months in the future. Researchers found that the month data for pullets had a higher R^2 value than the month data for laying hens after comparing the two data sets with the number of laying hens and with egg production. Therefore, the month data mentioned later in this study refer to the month data for pullets.

Equation for number of laying hens: Researchers obtained the equation and the R^2 values for the month data and the number of laying hens by substituting these values into the statistical Software Minitab (1994) and gradually increasing the power of the exponents of both variables. Researchers then added the total number of pullets as another variable and increased the power of the exponent until the R^2 value ceased increasing. Finally, researchers added the data for the interaction between month and pullets to account for the relationship between these two variables.

Equation for egg production: Researchers obtained the equation for egg production in the same way we obtained the equation for the number of laying hens, the only difference being the addition of the laying hen data as another variable. Researchers who lack the real value for laying hens may use this equation to obtain predicted values and substitute them into the equation for the number of laying hens.

Steps for establishing a model

Traditional Regression Model-MiniTab (Input data): Input the required data including the month and total number of pullets as well as the known number of laying hens and egg production.

Select variables: Choose the known number of laying hens as a response for the laying hens equation and choose the month and the total number of pullets as predictors. Choose the known number of eggs produced as a response for the egg production equation and choose the month, total number of pullets and the number of laying hens as predictors.

Regression analysis: The laying hens and egg production equations were determined by regression analysis.

Artificial Neural Network Model-NeuroShell: Data were analysed using NeuroShell Software from Ward Systems Group (2010).

Step 1 (Source data import): Create the data set by inputting the month, the total number of pullets and the known number of laying hens and eggs produced into Microsoft Excel, 2007. Save the Excel spreadsheet as a CSV file and import it into the worksheet in NeuroShell.

Step 2 (Select the range of data): Choose the month with egg production as the training sample for the model. Choose the month a second time with unknown egg production as the validation data set for the model.

Step 3 (Select the Regression Model type): NeuroShell provides neural and genetic model types and the user may choose either of these two model types to perform a regression analysis.

Step 4 (Training the model): In this step, the software uses the data that has been chosen to train the model to build the model.

Step 5 (Applying the model): In this step, the software substitutes the data that has been chosen for validation into the model and calculates the predicted values.

Step 6 (Substituting the predicted values into the equation): Transfer the predicted values for month, number of total pullets and known number of laying hens into Minitab to obtain the regression equation.

RESULTS AND DISCUSSION

Illustration of the models: Table 1 shows the original variables. The definition of each variable is provided:

Table 1: Original data for egg production used in the study

Years	Month	Pmonth	Pullet[1]	Laying[1]	Case[1]	Years	Month	Pmonth	Pullet[1]	Laying[1]	Case[1]
2001	1	8	120.85	2,329.23	8.43	2005	1	8	122.55	2,450.08	9.13
	2	9	134.79	2,599.79	9.13		2	9	109.00	2,385.13	8.78
	3	10	169.43	2,702.80	8.92		3	10	146.26	2,481.62	9.10
	4	11	152.21	2,662.54	8.91		4	11	199.35	2,504.44	9.29
	5	12	157.18	2,584.43	8.89		5	12	159.92	2,389.78	8.73
	6	1	173.63	2,582.77	8.72		6	1	158.37	2,346.61	8.54
	7	2	141.56	2,566.74	8.64		7	2	179.92	2,311.97	8.09
	8	3	157.57	2,555.30	8.56		8	3	150.94	2,432.52	8.77
	9	4	186.02	2,554.01	9.00		9	4	155.36	2,299.69	8.34
	10	5	175.82	2,496.98	8.85		10	5	175.61	2,474.01	8.86
	11	6	131.44	2,478.04	8.89		11	6	117.72	2,385.74	8.81
	12	7	110.92	2,441.72	8.97		12	7	137.78	2,378.64	8.88
2002	1	8	90.34	2,440.24	8.90	2006	1	8	112.13	2,463.11	8.96
	2	9	103.71	2,424.79	8.67		2	9	125.20	2,386.27	8.54
	3	10	171.14	2,361.69	8.54		3	10	155.19	2,450.02	8.95
	4	11	122.78	2,426.13	8.85		4	11	171.42	2,438.23	8.88
	5	12	132.37	2,319.31	8.50		5	12	122.42	2,470.06	9.01
	6	1	169.28	2,226.78	8.16		6	1	127.28	2,312.31	8.45
	7	2	103.59	2,299.34	8.37		7	2	134.70	2,407.72	8.71
	8	3	137.47	2,319.58	8.46		8	3	114.19	2,337.34	8.52
	9	4	141.91	2,370.67	8.62		9	4	132.22	2,471.92	8.97
	10	5	141.85	2,391.43	8.72		10	5	99.37	2,418.77	8.98
	11	6	160.86	2,456.34	9.07		11	6	100.71	2,484.64	9.19
	12	7	134.20	2,487.82	9.04		12	7	121.74	2,512.87	9.26
2003	1	8	122.60	2,437.03	8.83	2007	1	8	103.92	2,503.38	9.21
	2	9	109.83	2,357.15	8.60		2	9	98.92	2,390.31	8.81
	3	10	159.00	2,382.58	8.65		3	10	153.34	2,373.31	8.72
	4	11	137.62	2,428.00	8.81		4	11	125.30	2,476.61	9.17
	5	12	159.94	2,371.19	8.56		5	12	125.15	2,425.26	8.84
	6	1	127.58	2,355.25	8.53		6	1	155.05	2,358.42	8.57
	7	2	130.79	2,259.79	7.87		7	2	112.56	2,367.11	8.44
	8	3	166.76	2,388.13	8.53		8	3	119.18	2,609.52	9.22
	9	4	135.35	2,380.20	8.55		9	4	144.45	2,504.69	8.96
	10	5	124.06	2,399.34	8.70		10	5	124.46	2,589.51	9.20
	11	6	155.61	2,454.86	8.98		11	6	110.46	2,565.85	9.33
	12	7	128.18	2,483.44	8.93		12	7	130.57	2,559.55	9.19
2004	1	8	124.36	2,337.73	8.47	2008	1	8	99.89	2,574.93	9.12
	2	9	126.96	2,331.33	8.24		2	9	86.53	2,457.98	8.58
	3	10	124.83	2,435.99	8.75		3	10	155.91	2,606.80	9.19
	4	11	115.35	2,434.48	8.81		4	11	129.60	2,504.01	8.98
	5	12	148.47	2,261.38	8.19		5	12	157.44	2,477.45	8.81
	6	1	116.79	2,237.32	8.13		6	1	130.34	2,434.06	8.59
	7	2	148.73	2,363.92	8.49		7	2	124.41	2,421.90	8.50
	8	3	104.37	2,161.68	8.09		8	3	169.31	2,474.64	8.69
	9	4	118.40	2,389.69	8.64		9	4	146.22	2,504.96	8.68
	10	5	132.52	2,415.43	8.83		10	5	131.85	2,518.52	8.90
	11	6	116.90	2,442.31	8.93		11	6	149.86	2,489.67	8.83
	12	7	115.71	2,463.53	9.15		12	7	130.82	2,554.74	9.08

Table 1: Continue

Years	Month	Pmonth	Pullet[1]	Laying[1]	Case[1]	Years	Month	Pmonth	Pullet[1]	Laying[1]	Case[1]
2009	1	8	138.10	2,463.59	8.65	2010	1	8	155.90	2,596.69	9.37
	2	9	128.24	2,428.98	8.44		2	9	124.27	2,467.34	8.93
	3	10	127.67	2,351.28	7.96		3	10	138.08	2,508.12	8.92
	4	11	134.91	2,559.79	8.72		4	11	131.67	2,538.41	9.24
	5	12	162.38	2,447.47	8.60		5	12	163.99	2,502.83	9.01
	6	1	133.63	2,392.45	8.45		6	1	127.02	2,506.51	9.04
	7	2	127.23	2,456.81	8.62		7	2	155.26	2,512.99	9.13
	8	3	158.55	2,529.62	8.78		8	3	120.39	2,543.34	9.07
	9	4	133.34	2,421.85	8.55		9	4	121.78	2,535.62	9.13
	10	5	157.48	2,540.41	9.10		10	5	139.62	2,488.14	8.99
	11	6	135.43	2,487.49	8.95		11	6	112.05	2,603.39	9.60
	12	7	131.41	2,571.49	9.35		12	7	91.97	2,522.21	9.33
						2011	1	8	108.66	2,503.61	9.01
							2	9	106.63	2,568.33	9.23
							3	10	128.11	2,539.45	9.52

[1]The units of Pullet and Laying are in tens of thousands of birds and case is tens of thousands of cases

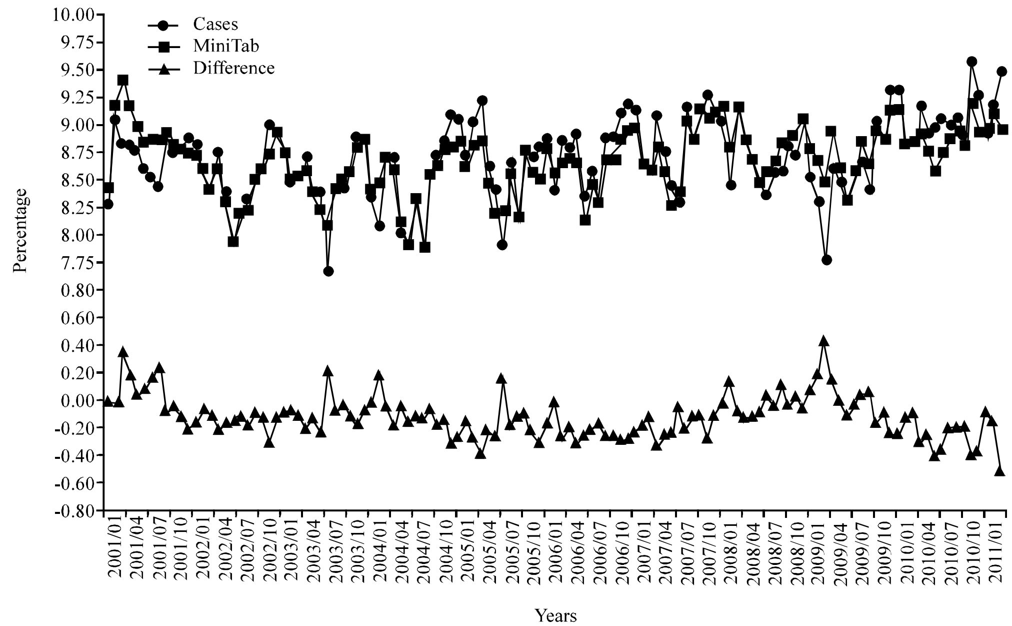

Fig. 1: Line chart of predicted egg production and the percent difference from the regression equation from MiniTab

- Year, month: The year and month (these variables are provided by the user)
- Pmonth: The month of pullets is the 5 months earlier than the month
- Pullet: The total number of pullets at Pmonth
- Laying: The number of laying hens
- Case: The number of cases of eggs produced

The regression equations:

- Traditional Regression Model-MiniTab: case = 2.77+ 0.696 Pmonth − 0.00621 Pmonth^2 − 0.00163 pullet + 0.0025 laying, $R^2 = 0.699$
- Artificial Neural Network Model-NeuroShell: case = 2.82 + 0.113 Pmonth − 0.00871 Pmonth^2 − 0.00157 pullet + 0.0024 laying, $R^2 = 0.965$

Data validation

Traditional Regression Model-MiniTab: The validation data were obtained by substituting the old data into the predictive equation and the resulting data were compared with the original data. The differences in the validation values and the original data were calculated (Fig. 1 and Table 2). The R^2, MSE and RMSE obtained through this method were 0.673, 0.031 and 0.177, respectively. The percentage errors were within 1% at 45.283 and within 3% at 89.623.

Artificial Neural Network Model-NeuroShell: The validation values and the original data were compared and the differences in the data sets were calculated (Fig. 2 and Table 3). The R^2, MSE and RMSE obtained through this method were 0.917, 0.006 and 0.075, respectively. The

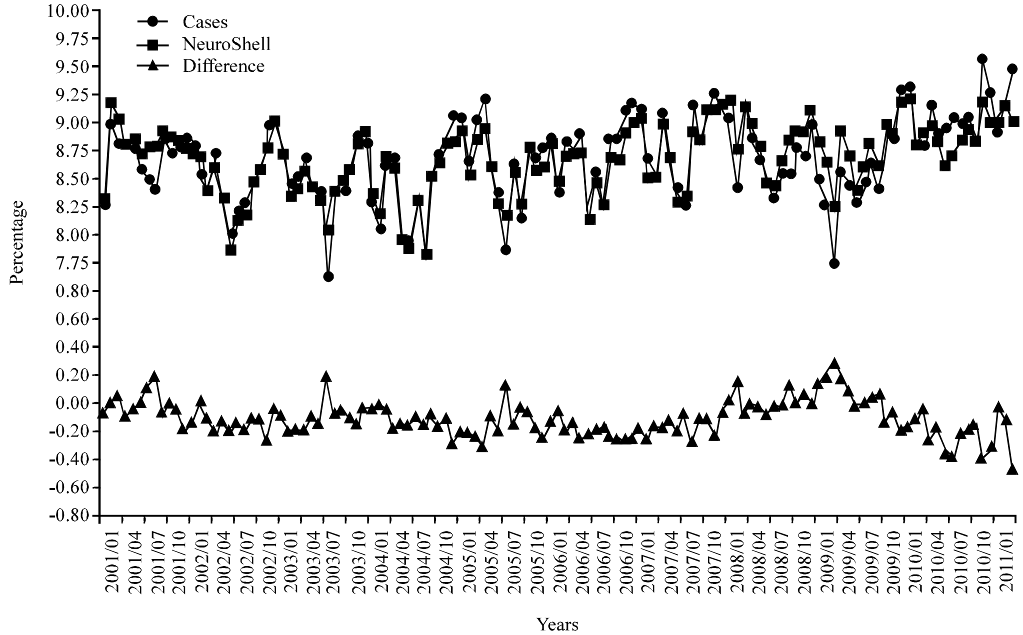

Fig. 2: Line chart of predicted egg production and percent difference from the regression equation from NeuroShell

Table 2: Regression results from MiniTab

MiniTab	Values
R^2	0.687
r	0.829
Mean Square Error (MSE)	0.001
Root Mean Square Error (RMSE)	0.023
Within 1%	42.276
Within 1-3%	43.089
Within 3-5%	13.008
>5%	1.626

Table 3: Regression results from NeuroShell

NeuroShell	Values
R^2	0.788
r	0.888
Mean Square Error (MSE)	0.022
Root Mean Square Error (RMSE)	0.149
Within 1%	45.528
Within 1-3%	43.902
Within 3-5%	9.756
>5%	0.813

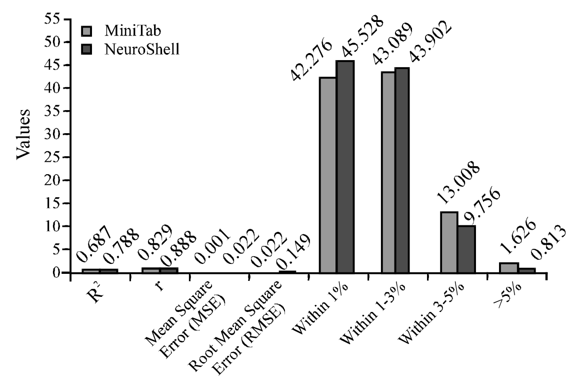

Fig. 3: Comparison of results from MiniTab and NeuroShell

percentage errors were within 1% at 59.434 and within 3% at 95.283. The comparison of the predicted values from MiniTab and NeuroShell (Fig. 3) demonstrated that NeuroShell is more accurate than MiniTab when regressing the data in this research.

Figure 3 shows the results of the comparison of MiniTab and NeuroShell. Researchers found NeuroShell to be more accurate than MiniTab with respect to the correlation, error and differences. Therefore, NeuroShell has better predictive ability than statistical regression methods. The results of the comparison are similar to the results of comparisons performed by Ahmadi *et al.* (2008), Golian and Ahmadi (2008), Ahmad (2009) and Savegnago *et al.* (2011) which all showed that neural network models provided the best fit when predicting poultry performance in terms of poultry growth and egg production.

Because this study utilised data from over a decade of market information, variations in the data may stem from variations in economic performance, national diet preferences, consumption habits and other factors. Although, the poultry industry and the egg production process have some impact on Taiwanese egg production, estimates of long-term changes in production can still function as a reference and as a warning. The results can successfully predict production and can be used to adjust production in the poultry industry if there is training in proper data treatment. Researchers can also forecast egg prices in Taiwan followed with this research method. Researchers hope this ability will contribute to egg production and egg price stability. Furthermore because other agricultural industries and production and regulation processes should have similar reference values to those of the poultry industry, researchers hope to also apply the Forecasting Method to these situations.

CONCLUSION

The results show that the ANN Method is more accurate than Traditional Regression Models for predicting egg production in Taiwan.

REFERENCES

Ahmad, H.A., 2009. Poultry growth modeling using neural networks and stimulated data. J. Applied Poult. Res., 18: 440-446.

Ahmadi, H., M. Mottaghitalab, N. Nariman-Zadeh and A. Golian, 2008. Predicting performance of broiler chickens from dietary nutrients using group method of data handling-type neural networks. British Poult. Sci., 49: 315-320.

Golian, A. and H. Ahmadi, 2008. Neural network model for egg production curve. J. Anim. Vet. Adv., 7: 1168-1170.

Minitab, 1994. Minitab statistic software release 10.1. Minitab Inc., USA.

National Animal Industry Foundation, 2009. The statistical data of egg production and laying hens. National Animal Industry Foundation, Taipei.

Roush, W.B., W.A. Dozier, 3rd and S.L. Branton, 2006. Comparison of Gompertz and neural network models of broiler growth. Poult. Sci., 85: 794-797.

Savegnago, R.P., B.N. Nunes, S.L. Caetano, A.S. Ferraudo, G.S. Schmidt, M.C. Ledur and D.P. Munari, 2011. Comparison of logistic and neural network models to fit to the egg production curve of White Leghorn hens. Poult. Sci., 90: 705-711.

Ward Systems Group, 2010. Neuroshell predictor tutorial. Ward Systems Group, UK.

Zhang, Z., Y. Wang, G. Fan and P.B. Harrington, 2007. A comparative study of multilayer perceptron neural networks for the identification of rhubarb samples. Phytochem. Anal., 18: 109-114.