Journal of Theoretical and Applied Information Technology
10th August 2013. Vol. 54 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

THE CONSTRUCTION OF A VISITOR NUMBER FORECASTING MODEL FOR THE SHIHMEN RESERVOIR RECREATION AREA USING ARTIFICIAL INTELLIGENCE TECHNOLOGY

HAN-CHEN HUANG
Department of Leisure Management, Yu Da University, Taiwan
E-mail: hchuang@ydu.edu.tw

ABSTRACT

Visitor number is an important reference for recreation area managers when developing tourism and marketing plans, and recreation area managers must be able to precisely forecast visitor numbers to possess sufficient information for establishing management strategies and plans. Developing an effective and accurate visitor number forecasting model can allow managers to respond in advance and adjust operating and development directions accordingly. Consequently, this study used an artificial neural network (ANN) to establish a visitor number forecasting model. The results of verification using actual data indicated that the absolute relative error (ARE) of the model developed in this study was 17.6565% and the correlation was 0.773. These results show that the model is highly accurate for forecasting visitor number.

Keywords: *Recreation Area*, *Tourism*, *Forecasting*, *Artificial Neural Network*, *Visitor Number*

1. INTRODUCTION

Recreation area managers require the ability to precisely estimate or forecast visitor number to possess sufficient information for developing management strategies and plans. An understanding of visitor number data can be employed to[1,2]:

- Define recreational and service facility design standards.
- Evaluate recreational service performance.
- Evaluate the relationship between degree of use and impact on the social and physical environment.
- Reduce conflict between different customer groups.
- Identify the popular attractions in the park.
- Determine and confirm potential problems for the popular attractions in the park.
- Identify demand trends and forecasts.
- Organize recreational service facilities and service staff.
- Schedule maintenance work and staff assignments.
- Allocate and distribute public and service facilities in the park.
- Monitor whether tourists are adhering to use regulations.
- Evaluate the importance that using the natural area for recreation has for socioeconomics and politics.

Visitor number is not merely a quantified value; instead, it is an important basic datum. How to improve the precision of visitor number estimations or forecasts is a significant topic for various recreation management units [1-3]. Therefore, this study employed an artificial neural network (ANN) to develop a visitor number forecasting model. The results can provide recreation area managers with more reliable and efficient reference data.

2. LITERATURE REVIEW

2.1 Visitor Number

Visitor number is the most direct performance evaluation indicator for a recreation area. The higher the visitor number, the more popular the destination is. Visitor number data is differentiated based on varying meaning or significance as follows: visitor number, individual visitor number, and recreation visitor-day [1-3]. Of these, the visitor number reflects the net visitor number of a recreation area within a specific time period. Previous estimations of total visitor numbers to large recreation areas typically employed the accumulated number of visitors to various

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

attractions. This type of accumulated total visitor number data is typically higher than the number of actual visitors.

2.2 Forecasting

Stynes [3,4] identified and classified the following four types of visitor number forecasting methods: (1) the Delphi technique; (2) time series or trend extension models; (3) structural models; and (4) system or simulation models. Other commonly employed forecasting methods proposed in relevant research [5-16] include judgment analysis, survey analysis, trend analysis, cause analysis, and artificial intelligence. Furthermore, in recent years, artificial intelligence forecasting models have gradually replaced conventional forecasting models.

An ANN simulates the process of a biological brain after experiencing external stimulus and learning. Through continuous learning and error corrections, an ANN model adapts to the environment. According to the learning style or type, ANNs can be classified as supervised learning, unsupervised learning, associated learning networks, and optimization application networks [16-20].

3. BPNN

Back-propagation neural networks (BPNNs) have been shown to be effective tools for a wide range of fields [21-25]. BPNNs are a type of supervised learning algorithm ANN, and are formed by combining multilayer perceptrons and error back-propagation [17]. The architecture or framework of a typical BPNN is shown in Figure 1.

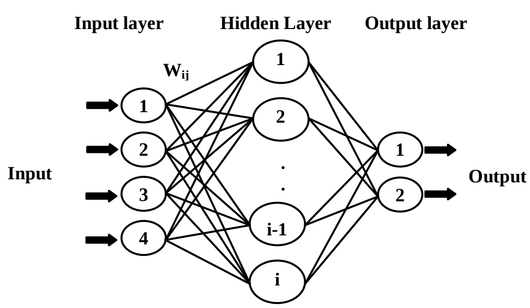

Figure 1:BPNN architecture

The phases of BPNN computing processes can be divided into learning processes and recall processes [26].

3.1 Learning Process

Step 1.Establish the architecture parameters and learning parameters.

Steps 2.Randomly generate weight matrix and bias vector initial values.

Step 3.Input training examples, including input values (X_1, X_2, X_3) and target or desired output values (t_1, t_2, t_3).

Step 4.Calculate the inferred output values (y_1,y_2,y_3)

(1) Hidden layer (h_1, h_2, h_3) (Eqs. 1 and 2)

$$net_k = \sum_i W_{ik} x_i - \theta_k \quad (1)$$

$$h_k = \frac{1}{1+\exp(-net_k)} \quad (2)$$

(2) Output layer (y_1, y_2, y_3) (Eqs. 3 and 4)

$$net_j = \sum_{ki} W_{kj} h_k - \theta_j \quad (3)$$

$$y_j = \frac{1}{1+\exp(-net_j)} \quad (4)$$

Step 5.Calculate the difference or gap amount δ (Eqs. 5 and 6).

(1) Output layer

$$\delta_j = (t_j - y_j)\cdot y_j \cdot (1 - y_j) \quad (5)$$

(2) Hidden layer

$$\delta_k = (\sum_j \delta_j \cdot W_{kj})\cdot h_k \cdot (1 - h_j) \quad (6)$$

Step 6.Calculate the weight correction and bias correction amounts (Eqs. 7 to 10)

(1) Output layer

$$\Delta W_{kj}(n) = \eta \delta_j h_k + \alpha \cdot \Delta W_{kj}(n-1) \quad (7)$$

$$\Delta \theta_j(n) = -\eta \delta_j + \alpha \cdot \Delta \theta_j(n-1) \quad (8)$$

(2) Hidden layer

$$\Delta W_{jk}(n) = \eta \delta_k x_i + \alpha \cdot \Delta W_{ik}(n-1) \quad (9)$$

$$\Delta \theta_k(n) = -\eta \delta_k + \alpha \cdot \Delta \theta_{kj}(n-1) \quad (10)$$

Step 7.Update the weight and bias values (Eqs. 11 to 14)

(1) Output layer

$$W_{kj} = W_{kj} + \Delta W_{kj} \quad (11)$$

$$\theta_j = \theta_j + \Delta \theta_j \quad (12)$$

(2)Hidden layer

$$W_{jk} = W_{jk} + \Delta W_{jk} \quad (13)$$

$$\theta_k = \theta_k + \Delta \theta_k \quad (14)$$

Journal of Theoretical and Applied Information Technology
10th August 2013. Vol. 54 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Step 8.Repeat Steps 3 to 7 until convergence (i.e., no significant changes in error or after executing a certain number of learning cycles)

3.2 Recall Process

Step 1.Establish network parameters.
Step 2.Read the weight matrix and bias vector.
Step 3.Input the input vector for unknown data (X_1, X_2, X_3).
Step 4.Calculate the inferred output vector.
(1)Hidden layer output values (h_1, h_2, h_3) (Eqs. 15 and 16)

$$net_k = \sum_i W_{ik} x_i - \theta_k \quad (15)$$

$$h_k = \frac{1}{1+\exp(-net_k)} \quad (16)$$

(2) Calculate the inferred output values (y_1, y_2, y_3) (Eqs. 17 and 18)

$$net_j = \sum_k W_{kj} h_k - \theta_j \quad (17)$$

$$y_j = \frac{1}{1+\exp(-net_j)} \quad (18)$$

4. MODEL CONSTRUCTION

4.1 Input and Output Variables

This study used data of the monthly visitor numbers to Taiwan's Shihmen Reservoir from 1993 to 2012 (T_{t-1}: t-1 month visitor number to T_{t-12}: t-12 month visitor number) as the input variables to forecast the t month visitor number T_t (Table 1). Overall, 70% of the data were used as training data, 15% were used as cross-validation data, and 15% were used as testing data (Figure 2).

Absolute relative error (ARE; Eq. 19) and correlation (r) were adopted as indicators for evaluating the model.

- ARE: The smaller the value, the smaller the error between the forecast value and the target value.

$$ARE = \frac{1}{n}\sum_{1}^{n}\left|\frac{y_i - y'_i}{y_i}\right| \cdot 100\% \quad (19)$$

where n is number of the forecasting periods, y_i is the actual value for the i period , and y'_i is the forecast value for the i period.

- Correlation (r): As r approaches 1, the ANN forecasting results improve.

Table1: Forecasting model variables

Input					Output
T_{t-12}	T_{t-11}	....	T_{t-2}	T_{t-1}	T_t
T_{t-11}	T_{t-10}	....	T_{t-1}	T_t	T_{t+1}
T_{t-10}	T_{t-9}	....	T_t	T_{t+1}	T_{t+2}
.	.	....	.	.	.
.	.	....	.	.	.
.	.	....	.	.	.
T_{t+n-10}	T_{t+n-9}	....	T_{t+n-2}	T_{t+n-1}	T_{t+n-2}
T_{t+n-11}	T_{t+n-10}	....	T_{t+n-1}	T_{t+n}	T_{t+n-1}
T_{t+n-12}	T_{t+n-11}	....	T_{t+n-2}	T_{t+n-1}	T_{t+n}

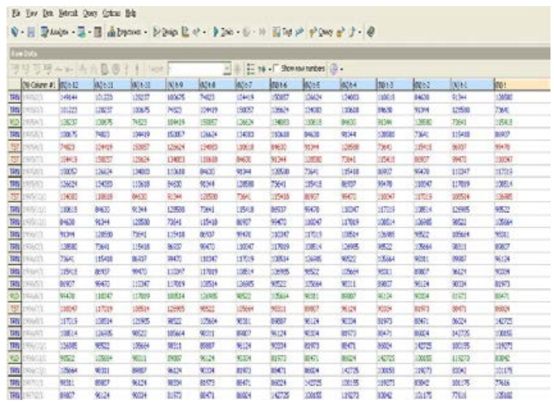

Figure 2: Input and output variable data

4.2 Architecture Design

The network input activation function used was the hyperbolic tangent, the output error function used was the sum-of-squares, and the output activation function used was a logistic model. The architecture design employed the test error as the determination standard to identify the optimal network architecture, which was 12-16-1 (Figure 3 and Table 2). The input layer comprised 12 neurons. The architecture also contained one hidden layer with 16 neurons and an output layer with 1 neuron.

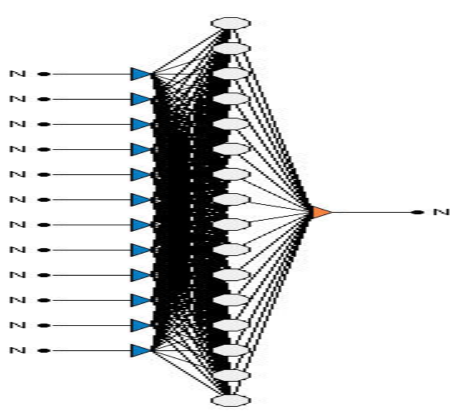

Figure 3: Network architecture

Journal of Theoretical and Applied Information Technology
10th August 2013. Vol. 54 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 2: The optimal network architecture search results

ID	Architecture	# of Weights	Fitness	* Test error	Akaike's criterion	R-Squared	Correlation	Train Error
6	[12-16-1]	225	0.000063	15775.57421	90.000874	0.492737	0.727311	13956.94043
4	[12-12-1]	169	0.00006	16533.007091	0.000974	0.520434	0.716225	13523.902344
7	[12-14-1]	197	0.000059	16850.623047	0.000921	0.506806	0.717428	13734.813477
2	[12-30-1]	421	0.000059	16992.535156	0.000651	0.521029	0.724984	13870.926758
5	[12-8-1]	113	0.000059	17091.146484	0.00109	0.484408	0.708484	13691.851563
8	[12-17-1]	239	0.000058	17266.970703	0.000855	0.501792	0.686888	13680.399414
3	[12-19-1]	267	0.000057	17397.966797	0.000014	0.400526	0.60450	13900.449219
1	[12-2-1]	29	0.000057	17474.771484	0.001331	0.521257	0.737823	13895.50293
9	[12-15-1]	211	0.000057	17499.861328	0.000897	0.474514	0.719267	13844.567383

4.3 Train and Test Results

For BPNN training, the quick propagation training algorithm was used. The training algorithm parameters were a quick propagation coefficient = 1.75 and a learning rate = 0.1. To prevent network overfitting, an overtraining control method was adopted, as shown in Figure 4. Figure 5 shows the model training and validation results, which indicated an excellent convergence regarding error. The distribution of the actual values and the model output values is shown in Figure 6. As the results in the figure indicate, the model output values are primarily distributed on both sides of the diagonal (Output/Target= 1), denoting a good forecast ability. The forecasting model developed in this study can reflect fluctuations and change (Figure 7).

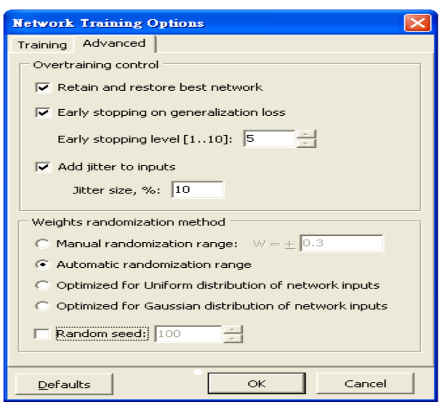

Figure 4: Overtraining control method

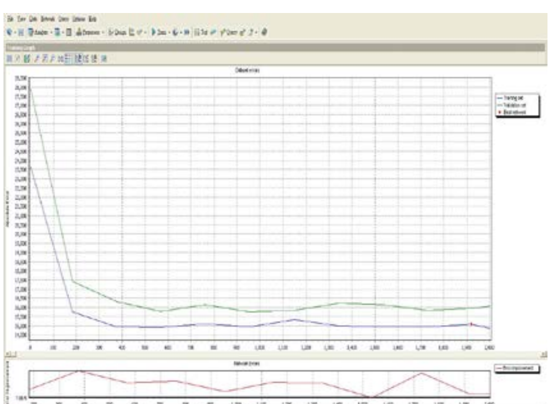

Figure 5: Model training and verification results

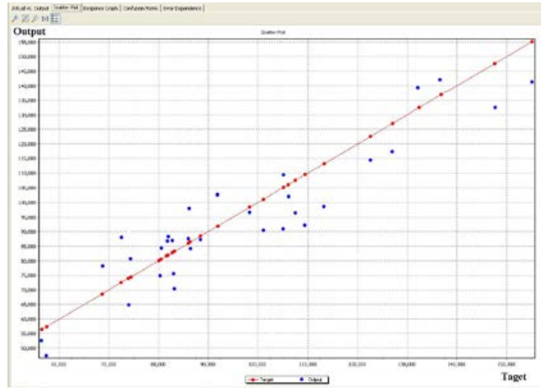

Figure 6: Scatter plot of the actual values and the model output values

The model training results are shown in Table 3, where ARE was 16.593%, and correlation was 0.802. This indicates that the training model was able to learn the training data with less than a 20% error. The model test results are shown in Table 4, where ARE was 17.6565%, and correlation was 0.773, indicating that the forecasting model developed in this study possesses considerable ability to forecast the actual visitor number.

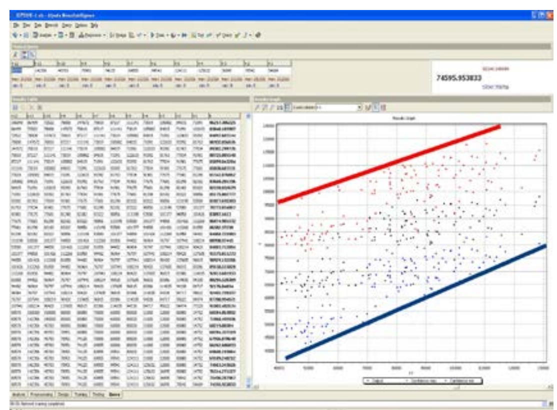

Figure 7: Model and output interface following construction

Table 3: Model training results

Summary	Target	Output	AE	ARE
Mean:	94898.971429	83427.024385	16822.623143	0.16593
Std Dev:	22325.82449	15741.41773	15052.540376	0.120331
Min:	56377	47241.784867	1301.162169	0.014707
Max:	155274	131534.754587	73468.729403	0.497513

Table 4: Model forecasting results

Summary	Target	Output	AE	ARE
Mean:	80816.246753	81326.127422	13420.969595	0.176565
Std Dev:	24905.261464	17740.400155	11057.104121	0.148087
Min:	27562	50513.179292	4.838274	0.000079
Max:	150057	122886.164922	63073.516272	0.832711

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

5. CONCLUSION

This study used a BPNN to develop a forecasting model. After employing 12 input variables to forecast future changes in visitor number, empirical analysis indicated that the model possessed significant forecasting ability; ARE was 17.6565% and the correlation coefficient was 0.773. These values confirm that the forecasting model can forecast actual visitor numbers. Thus, the model developed in this study can provide managers with forecasts of changes in future visitor numbers and enable them to respond to future operating problems before they occur.

REFRENCES:

[1] G. Cessford and A. Muhar, "Monitoring Options for Visitor Numbers in National Parks and Natural Areas", *Journal for Nature Conservation*, Vol. 11, No. 4, 2003, pp. 240-250.

[2] B.S. Lin and Y.J. Lin, "Estimation and Forecast of Recreation Use in Island-Recreation Site: A Case Study of Kinmen National Park", *Journal of National Park*, Vol. 20, No. 1, 2010, pp. 1-14.

[3] D.J. Synes, "An Introduction to Recreation Forecasting, in S. Lieber and D. Fesenmaier (Eds), Recreation planning and management", 1983, pp. 87-95, Pennsylvania: Venture Press.

[4] D.J. Stynes, "Time Series and Structural Recreation Forecasting Models, in S. Lieber and D. Fesenmaier (Eds), Recreation planning and management", 1983, pp. 105-119, Pennsylvania: Venture Press.

[5] Y.H. Lin and L.C. Chung, "Positioning Analysis of Taiwan Travel Market-Perspectives of Travel Related Practitioners from China and Taiwan", *Journal of Island Tourism Research*, Vol. 4, No. 3, 2011, pp. 1-23.

[6] Leon Oblak, Lidija Zadnik Stirn, Maja Moro, Jasna Hrovatin, Samo Mole, and Manja Kitek Kuzman, "Choice of Quantitative Method for Forecasting of Parquet Sales", *Drvna Industrija*, Vol. 63, No. 4, 2012, pp. 249-254.

[7] H.L. Wong, S.C. Ku, and H.W. Chen, "ARIMA and Neural Network Model for Passenger Flow Forecasting- A Case of Hong Kong Airport", *Ming-Hsin Journal*, Vol. 38, No. 2, 2012, pp. 69-83.

[8] S.Y. Huang, A.A. Chiu, and B.C. Wang, "Applying Intellectual Capital on Financial Distress Prediction Model in Taiwan Information Technology and Electronic Industry", *International Journal of Advancements in Computing Technology*, Vol. 4, No. 8, 2012, pp. 270-280.

[9] S. Pant and A. K. Wadhwani, "A Real Time GRS Detection System used Error Back Propagation Neural Network", *International Journal of Engineering Science and Technology*, Vol. 3, No. 9, 2011, pp. 7205-7210.

[10] Y.M. Chan, "Forecasting Tourism: A Sine Wave Time Series Regression Approach", *Journal of Travel Research*, Vol. 32, No. 1, 1993, pp. 58-60.

[11] C.H. Wang, "Predicting Tourism Demand using Fuzzy Time Series and Hybrid Grey Theory", *Tourism Management*, Vol. 25, No. 1, 2004, pp. 367-374.

[12] F.L. Chu, "Forecasting Tourism: A Combined Approach", *Tourism Management*, Vol. 19, No. 6, 1998, pp. 515-520.

[13] Loganathan Nanthakumar, Thirunaukarasu Subramaniam, and Mori Kogid, "Is 'Malaysia Truly Asia'? Forecasting Tourism Demand from Asean Using SARIMA Approach, Tourismos: An International Multidisciplinary", *Journal of Tourism*, Vol. 7, No. 1, 2012, pp. 356-381.

[14] Massimiliano Castellani, Maurizio Mussoni, and Pierpaolo Pattitoni, "Air Passenger Flows: Evidence from Sicily and Sardinia", *Almatourism*, Vol. 1, No. 2, 2011, pp. 16-28.

[15] C.J. Lin, H.F. Chen, and T.S. Lee, "Forecasting Tourism Demand using Time Series, Artificial Neural Networks and Multivariate Adaptive Regression Splines: Evidence from Taiwan", *International Journal of Business Administration*, Vol. 2, No. 2, 2011, pp. 14-24.

[16] Ruhaidah Samsudin, Puteh Saad, and Ani Shabri, "Hybridizing GMDH and Least Squares SVM Support Vector Machine for Forecasting Tourism Demand", *International Journal of Research and Reviews in Applied Sciences*, Vol. 3, No. 3, 2010, pp. 274-279.

[17] H.C. Huang, A.Y. Chang, and C.C. Ho, "Using Artificial Neural Networks to Establish a Customer-cancellation Prediction Model", *Przeglad Elektrotechniczny*, Vol. 89, No. 1b, 2013, pp. 178-180.

[18] H.C. Huang and C.C. Ho, "Back-Propagation Neural Network Combined With a Particle Swarm Optimization Algorithm for Travel Package Demand Forecasting", *International*

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Journal of Digital Content Technology and its Applications, Vol. 4, No. 17, 2012, pp. 194-203.

[19] T.S. Lan, P.C. Chen, C.H. Lan, and K.C. Chuang, "A Study of Using Back-Propagation Neural Network for the Sales Forecasting of the Thin Film Sputtering Process Material", *International Journal of Advancements in Computing Technology*, Vol. 4, No. 8, 2012, pp. 118-127.

[20] P.C. Chen, C.Y. Lo, H.T. Chang, and L.C. Yu, "A Study of Applying Artificial Neural Network and Genetic Algorithm in Sales Forecasting Model", *Journal of Convergence Information Technology*, Vol. 6, No. 9, 2011, pp. 352-362.

[21] D.S. Li and Q. He, "A PSO-Based Neural Network Ensemble with the Application of Flame Combustion Diagnosis", *Journal of Theoretical and Applied Information Technology*, Vol. 48, No. 1, 2013, pp. 60-66.

[22] G. Liu, H. Wang, and B.L. Duan, "Controller of PH Process based on Fuzzy Neural Network", *Journal of Theoretical and Applied Information Technology*, Vol. 48, No. 1, 2013, pp. 168-172.

[23] P.L. Zhang and L.F Cheng, "Application of Fuzzy Clustering Neural Network in Conjunction Speech Recognition", *Journal of Theoretical and Applied Information Technology*, Vol. 48, No. 1, 2013, pp. 319-324.

[24] L. Gao, S.L. Pan, and J. Shen, "A Robot Path Planning Scheme based on Neural Network", *Journal of Theoretical and Applied Information Technology*, Vol. 46, No. 2, 2012, pp. 654-658.

[25] L.J. Wang and D.F. Wang, "Application of Clonal Selection Algorithm on Ignition Advanced Angle of Hydrogen-Fueled Engine", *Journal of Theoretical and Applied Information Technology*, Vol. 46, No. 2, 2012, pp. 719-724.

[26] M.T. Hagan, H.B. Demuth, and M.H. Beale, "Neural Network Design, 1996, Boston, MA: PWS Publishing.