Predicting Changes in Construction Cost Indexes Using Neural Networks

By Trefor P. Williams,[1] Member, ASCE

Abstract: Construction cost indexes provide a comparison of cost changes from period to period for a fixed quantity of goods or services. Back-propagation neural-network models have been developed to predict the change in the *ENR* construction cost index for one month and six months ahead. A training set of macroeconomic data was developed for the period from 1967 to 1991. The neural-network models use inputs including recent trends in the index, the prime lending rate, housing starts, and the month of the year. Output from the neural-network models was compared with predictions made by exponential smoothing and simple linear regression. The prediction produced by the neural network gave a greater error than either exponential smoothing or linear regression. It can be concluded that the movement of the cost indexes is a complex problem that cannot be predicted accurately by a back-propagation neural-network model.

INTRODUCTION

Back-propagation neural networks are useful tools for pattern-recognition problems. Programs of this type have recently found application for predicting time series in the financial markets. Back-propagation networks are used to forecast technical bottoms and rallies in the Standard & Poor's 500. Zaremba (1990) has discussed the use of neural networks to predict trends in the future markets. The use of neural networks to predict trends in indexes like the Standard & Poor's 500 indicates that they have potential to predict variations and trends in construction cost indexes.

The level of construction prices is affected by many different factors. These factors include inflationary trends in the economy, the current level of construction activity, seasonal effects, and the cost of borrowing money. The purpose of this paper is to discuss how back-propagation neural networks can be used to develop predictive models of construction cost indexes. In construction, the ability to predict trends in prices, both short- and long-term, can result in more accurate bids. Models of this type can potentially allow contractors to incorporate expected price fluctuations into their bidding strategy. The ability to predict variations in prices can result in reduced construction costs by allowing material purchases to be better timed to take advantage of short-term fluctuations in material prices. Owner organizations may also benefit from the development of neural networks to predict changes in construction cost indexes. Improved budgeting decisions could be made if the trend in construction prices could be forecast accurately. This could allow improved analysis of the effects of delaying or accelerating large construction expenditures.

CONSTRUCTION COST INDEXES

A cost index provides a comparison of cost or price changes from period to period for a fixed quantity of goods or services (Ostwald 1984). Cost

[1]Asst. Prof., Dept. of Civ. Environ. Engrg., Rutgers Univ., P.O. Box 909, Piscataway, NJ 08855-0909.

Note. Discussion open until November 1, 1994. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on April 19, 1993. This paper is part of the ***Journal of Construction Engineering and Management***, Vol. 120, No. 2, June, 1994. ©ASCE, ISSN 0733-9364/94/0002-0306/$2.00 + $.25 per page. Paper No. 6007.

Downloaded from ascelibrary.org by San Jose State University on 10/16/25. Copyright ASCE. For personal use only; all rights reserved.

indexes permit an estimator to forecast construction costs from the present to future periods without going through detailed costing. Because construction costs vary with time due to changes in demand, economic conditions, and prices, indexes convert costs applicable at a past date to equivalent costs now or in the future.

This study centers on the behavior of the construction cost index published weekly in *ENR* magazine. The index was originally designed as a basic method to chart basic construction cost trends. The index is a weighted aggregate index of the prices of constant quantities of structural steel, portland cement, lumber, and common labor. The index consists of the cost of 200 h of common labor at the 20-city average labor rate, the cost of 1,135 kg (2,500 lb) of standard structural shapes, the cost of 1,024.22 kg (2,256 lb) of portland cement, and 331.62 board-m (1,088 board-ft) of 5.08 cm × 10.16 cm (2 in. × 4 in.) lumber at the 20-city price (Grogan (1992).

VARIATIONS IN CONSTRUCTION COST INDEX

Obviously, the *ENR* construction cost index is increasing in the long term. However, it is subject to considerable short-term variation. Fig. 1 illustrates the one-month percentage change in the construction cost index for the 186 consecutive months from August 1975 to January 1991. The one-month variation is characterized by having many small changes, with occasional significant jumps in construction prices.

Table 1 shows the observed monthly percentage change in the construction cost index for July 1967–December 1991. The monthly change is positive 88.17% of the time. The monthly change falls in the range between 0 and 1% for 67.38% of the time. However, there are extreme positive and negative variations observed. The six-month change in the construction cost index is shown in Table 2. These data are considerably smoother than the monthly variations. Still, there are several six-month periods where the construction cost index has decreased.

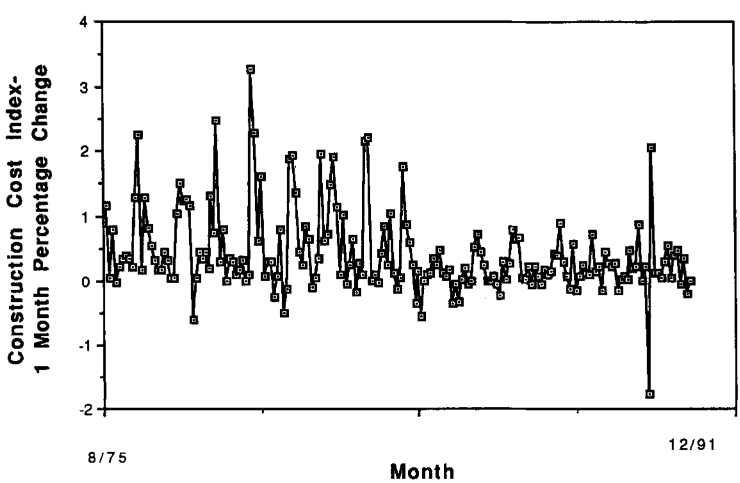

FIG. 1. Monthly Percentage Change in Construction Cost Index

J. Constr. Eng. Manage., 1994, 120(2): 306-320

Downloaded from ascelibrary.org by San Jose State University on 10/16/25. Copyright ASCE. For personal use only; all rights reserved.

TABLE 1. Observed One-Month Percentage Change in Construction Cost Index, 1967–91

One-month change in index (%) (1)	Number of observed cases (2)
Less than −0.5	5
−0.5 to −0.249	7
−0.25 to −0.001	21
0 to 0.249	78
0.25 to 0.449	61
0.5 to 0.749	26
0.75 to 0.999	23
1.0 to 1.5	31
Greater than 1.5	27

TABLE 2. Observed Six-Month Percentage Change in Construction Cost Index, 1967–91

Six-month change in index (%) (1)	Number of observed cases (2)
−1.0 to −0.001	9
0 to 0.999	40
1.0 to 1.999	62
2.0 to 2.999	36
3.0 to 3.999	39
4.0 to 4.999	39
5.0 to 5.999	27
6.0 to 6.999	17
7.0 to 7.999	10
8.0 to 8.999	2
9.0 to 9.999	4

ECONOMIC FACTORS INCLUDED IN MODEL

Two neural-network models were developed to predict changes in the *ENR* construction cost index. One predicts the one-month change, and the second predicts the six-month variation. The goal in developing the models was to find meaningful input factors that can be easily obtained from government reports or construction publications. This was done to make use of the model by construction practitioners practical.

Many different factors affect construction prices. The models use several input variables to attempt to capture the major factors that cause variations in construction prices. The variables selected as inputs to the models are:

- Percentage change in the construction cost index for one month
- The six-month percentage change in the construction cost index
- The prime lending rate
- The six-month percentage change in the prime lending rate
- The six-month change in the prime rate
- Number of housing starts for the month
- Percentage change in housing starts for one month

Downloaded from ascelibrary.org by San Jose State University on 10/16/25. Copyright ASCE. For personal use only; all rights reserved.

- Percentage change in housing starts for the preceding six-month period
- The month of the year

An initial analysis of the data collected indicated that a primary determinant of the construction cost index is its recent performance. Therefore the previous one-month and six-month variations in the index are included as inputs.

The prime rate is the interest rate charged by banks on short-term business loans. The relationship between construction prices and the prime lending rate is complex. An increase in interest rates raises the cost of capital projects and reduces industrial and commercial building. It also increases the cost and reduces the availability of funds for house purchase (Hillebrandt 1974). Interest rates also tend to mirror inflationary trends in the economy, with increasing prime rates observed during inflationary periods.

Housing starts are included in the model to reflect the level of activity in the construction industry. In periods of high housing starts, with high demand for construction materials, it would be expected that increases in the cost index would be higher. Seasonal variations in construction prices, and construction activities are also noted. An analysis of the data collected indicates that volatile changes in the construction cost index occur most frequently in the fall and spring.

The types of information required as inputs are widely disseminated and relatively easy to obtain. The construction cost index was obtained from the *ENR* annual cost roundup issue. Housing starts were obtained from *Current Construction Reports*, published monthly by the U.S. Department of Commerce. Prime rates were obtained from the *Federal Reserve Bulletin*, which is published monthly.

DEVELOPMENT OF NEURAL-NETWORK INPUT DATA

Data were obtained for the period from July 1967 to December 1991. For the one-month prediction, a training set of 215 cases and a test set of 63 cases were developed. For the six-month prediction, a training set of 207 cases and a test set of 66 cases were used. Table 3 shows a portion of the training set for the one period ahead prediction model. Columns 1 and 2 are the model outputs. The month is input as a binary variable. There is a different input for each month of the year. For example, if it is August, the August input is 1 and all other months are 0.

Relatively few examples exist of large variations in the construction cost index. To assist the neural network in identifying when a large variation will occur an additional input was provided in the one month prediction model. If the percentage change of the construction cost index was greater than 1.5% or less than −0.5%, then the actual percentage change of this data spike is input as a large variation. If there is no data spike, the large variation input is 0.

If was necessary to transform some of the data for input to the neural-network computer program. The construction cost index exhibits long-term increases. Therefore, the index will be a higher value in the future than it is now. The back-propagation algorithm does not extrapolate well. The training set must include data over the entire range of the input space (Freeman and Skapura 1991). A neural network that predicts the actual value of future construction cost indexes would not be possible. To overcome

Downloaded from ascelibrary.org by San Jose State University on 10/16/25. Copyright ASCE. For personal use only; all rights reserved.

Downloaded from ascelibrary.org by San Jose State University on 10/16/25. Copyright ASCE. For personal use only; all rights reserved.

TABLE 3. Portion of Training Set for One-Month Forecast

Predicted month (1)	Normalized index, one-month percentage change (2)	Normalized index, six-month percentage change (3)	Prime rate (4)	Prime rate, one-month percentage change (5)	Prime rate six-month percentage change (6)	Housing starts normalized (7)	Housing starts normalized, one-month percentage change (8)	Housing starts normalized, six-month percentage change (9)	Actual month (10)	Large variation (11)	Normalized next month, index percentage change (12)	Large variation, next month (13)
July 1975	0.740	0.710	7.15	−2.90	−40.56	0.239	0.757	0.801	June	0	0.579	0
August 1975	0.580	0.707	7.66	−1.30	−16.97	0.273	0.686	0.831	July	0	0.360	0
September 1975	0.360	0.712	7.88	−0.05	−0.635	0.296	0.665	0.860	September	0	0.510	0
November 1975	0.340	0.618	7.53	0.01	1.73	0.342	0.567	0.808	November	0	0.390	0
December 1975	0.390	0.452	7.26	0.19	2.62	0.284	0.508	0.780	December	0	0.420	0
January 1976	0.420	0.297	7.00	−0.15	−2.14	0.256	0.564	0.679	January	0	0.430	0
February 1976	0.430	0.224	6.75	−0.91	−13.48	0.363	0.914	0.809	February	0	0.420	0
March 1976	0.420	0.253	6.75	−1.13	−16.74	0.333	0.486	0.725	March	0	0.390	0
May 1976	0.610	0.326	6.75	−0.78	−11.56	0.366	0.676	0.683	May	0	0.790	2.249
June 1976	0.790	0.523	7.20	−0.06	−0.833	0.418	0.705	0.783	June	0	0.380	0
July 1976	0.380	0.505	7.25	0.25	3.45	0.360	0.520	0.762	July	0	0.604	0
August 1976	0.600	0.593	7.01	0.26	3.71	0.429	0.732	0.651	August	0	0.512	0
September 1976	0.512	0.641	7.00	0.25	3.57	0.570	0.817	0.822	September	0	0.455	0
October 1976	0.455	0.672	6.78	0.03	0.44	0.539	0.575	0.826	October	0	0.415	0
November 1976	0.415	0.576	6.50	−0.25	−3.85	0.533	0.612	0.796	November	0	0.382	0
December 1976	0.382	0.372	6.35	−0.85	−13.39	0.472	0.762	0.824	December	0	0.383	0
January 1977	0.383	0.371	6.25	−1.00	−16.00	0.343	0.086	0.643	January	0	0.438	0
February 1977	0.414	0.241	6.25	−0.75	−12.00	0.759	0.821	0.785	February	0	0.359	0

this problem, the construction-cost-index data were transformed to the percentage change of the index to make program inputs and outputs independent of the level of the construction cost index.

In addition, better results were obtained by normalizing the input data. Values of the percentage change in the index, percentage change in housing starts, and housing starts were transformed to fall in the range of 0–1. Back-propagation networks provide improved performance when the data are transformed in this manner. The following formula illustrates how the data were transformed. This formula applies when the maximum value is a positive number and the minimum is a negative number:

$$XT = \frac{X + X_{min}}{X_{max} + X_{min}} \tag{1}$$

where XT = the normalized value; X = the original value; X_{max} = the maximum positive value observed; and X_{min} = the absolute value of the minimum negative value observed.

MODEL OUTPUTS

The first model produces a prediction of the percentage change of the construction cost index one month in the future. The second model produces a prediction of the percentage change in the construction cost index six months in the future. The one-month prediction also produces an additional output. This model provides output indicating if a spike in the data is predicted. A data spike was defined as any monthly change greater than 1.5% or less than −0.5%. This output was included to help the neural network identify unusual variations in prices. Because spikes occur infrequently, there are few examples in the training set. The variable is 0 when there is no spike and is the actual percentage change in the construction cost index when a data spike exists.

NEUROSHELL PROGRAM

The neural-network model was constructed using the Neuroshell program (from Ward Systems Group, Inc., Frederick, Md.) on the IBM PC. This program implements a three-layer back-propagation model. Use of a shell program of this type is attractive for the construction environment. Model development does not require extensive knowledge of computer programming because the back-propagation model is predefined as part of the program. To use this program, a model developer is required only to define model inputs and outputs as well as to develop a suitable training set.

BACK-PROPAGATION ALGORITHM

The program employs an analog, three-layer, back-propagation network. Fig. 2 shows the basic configuration of the back-propagation network. During training the neural network processes patterns in a two-step procedure. In the first or forward phase of back-propagation learning, and input pattern is applied to the network, and the resulting activity is allowed to spread through the network to the output layer. The program compares the actual output pattern generated for the given input to the corresponding training-set output. This comparison results in an error for each neurode in the output layer. In the second or backward phase, the error from the output

Downloaded from ascelibrary.org by San Jose State University on 10/16/25. Copyright ASCE. For personal use only; all rights reserved.

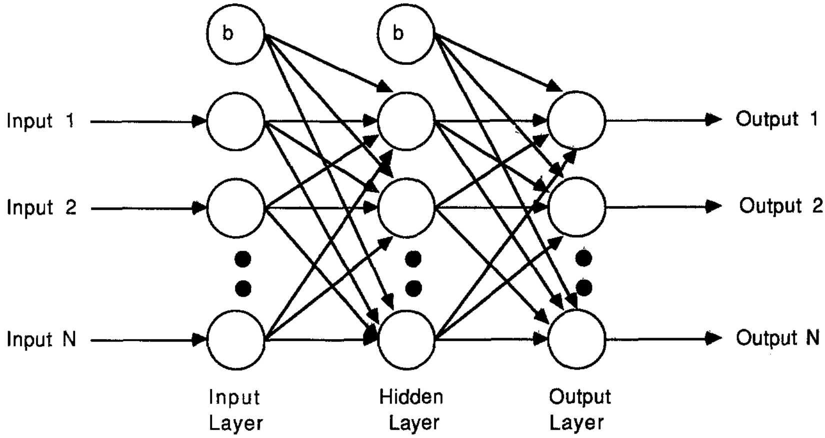

FIG. 2. Back-Propagation Network Structure

layer is propagated back through the network to adjust the interconnection weights between layers. This learning process is repeated until the error between the actual and desired output converges to a predefined threshold (Caudill and Butler 1990).

The relevant equations for the back propagation model can be summarized as follows (Freeman and Skapura 1991).

1. Apply the input vector to the input units:

$$\mathbf{x}_p = (x_{p1}, x_{p2}, \ldots, x_{pN})^t \tag{2}$$

These are in the input values for training pattern p. N is the number of nodes on the input layer. The pattern is fed forward.

2. Calculate the net-input values to the hidden layer units. The net input to the jth hidden node is:

$$\mathrm{net}_{pj}^h = \sum_{i=1}^{N} w_{ji}^h x_{pi} + \theta_j^h \tag{3}$$

where w_{ji}^h = the weight on the connection from the ith input unit; θ_j^h = a bias term that permits a more rapid convergence of the learning process; and the superscript h refers to quantities on the hidden layer.

3. Calculate the outputs from the hidden layer:

$$i_{pj} = f_j^h(\mathrm{net}_{pj}^h) = (1 + e^{-\mathrm{net}_{jk}^h})^{-1} \tag{4}$$

By applying a sigmoid threshold function, the output of the neuron is calculated. The function compresses the range of outputs to be between 0 and 1.

4. Move to the output layer and calculate the net-input values to each output-layer node:

$$\mathrm{net}_{pk}^o = \sum_{j=1}^{L} w_{jk}^o i_{pj} + \theta_k^o \tag{5}$$

J. Constr. Eng. Manage., 1994, 120(2): 306-320

Downloaded from ascelibrary.org by San Jose State University on 10/16/25. Copyright ASCE. For personal use only; all rights reserved.

where the superscript o refers to quantities on the output layer; and L = the number of nodes in the hidden layer.

5. Calculate the outputs on the output layer:

$$o_{pk} = f_k^o(\text{net}_{pk}^o) \tag{6}$$

6. Calculate the error for the output nodes:

$$\delta_{pk}^o = (y_{pk} - o_{pk})f_k^{o\prime}(\text{net}_{pk}^o) \tag{7}$$

where y_{pk} = the desired output value; and o_{pk} = the actual output value.

7. Calculated error terms for the hidden layer:

$$\delta_{pj}^h = f_j^{h\prime}(\text{net}_{pj}^h) \sum_k \delta_{pk}^o w_{kj}^o \tag{8}$$

Every weight update on the hidden layer depends on the error terms δ_{pk}^o on the output layer. The known output-layer errors are back-propagated to the hidden layer to determine the weight changes in that layer.

8. Update weights on the output layer:

$$w_{kj}^o(t + 1) = w_{kj}^o(t) + \eta\delta_{pk}^o i_{pj} \tag{9}$$

where η = the learning rate. This can be set to a value of 0–1 by the Neuroshell user. In this step, weights are modified slightly to produce a smaller error the next time the pattern is presented. The amount of weight modification is proportional to the amount of error. The larger the learning rate, the larger the weight change will be. For this problem the best results were with a learning rate of 0.6.

9. Update weights on the hidden layer:

$$w_{ji}^h(t + 1) = w_{ji}^o(t) + \eta\delta_{pj}^h x_i \tag{10}$$

10. Calculate the error term

$$E_p = \frac{1}{2} \sum_{k=1}^{M} \delta_{pk}^2 \tag{11}$$

The error term is a measure of how well the network is learning. When the error is acceptably small, training can be discontinued.

The Neuroshell program has several user defined variables. The primary variable is the number of hidden nodes. Generally, too few hidden nodes will result in two few distinctive problem characteristics being captured in the network. Defining too many hidden nodes causes much longer learning. At some point, increasing the number of hidden nodes does not greatly increase the capability of the network to classify. Experimentation was required to find the number of hidden nodes that produced optimum results. For the one-month model 20 hidden nodes produced the best results.

Overtraining is possible with neural networks. The Neuroshell program provides a feature that allows automatic comparison of the results obtained from a test set of inputs. A unique test set is developed that contains different cases from those used in the training set. The program calculates the mean square error obtained from the test set as training is taking place and automatically saves the network that produces the lowest error. Another factor controlled by the model developer is how training sets are presented to the

Downloaded from ascelibrary.org by San Jose State University on 10/16/25. Copyright ASCE. For personal use only; all rights reserved.

TABLE 4. One-Period Neural-Network Prediction Compared with Actual Results

Month (1)	Actual percentage change (2)	Predicted percentage change (3)	Difference (4)
September 1967	0.28	0.56	0.28
November 1967	0.09	0.48	0.39
January 1968	0.81	0.43	−0.38
March 1968	0.27	0.48	0.21
May 1968	1.58	0.66	−0.92
July 1968	0.35	0.9	0.55
September 1968	1.27	0.55	−0.72
November 1968	0.08	0.48	0.40
January 1969	1.23	0.53	−0.70
March 1969	0.73	0.52	−0.21
May 1969	0.72	0.76	0.04
July 1969	1.01	0.87	−0.14
November 1969	0.46	0.36	−0.10
January 1970	0.31	0.54	0.23
March 1970	0.23	0.48	0.25
June 1970	1.75	1.63	−0.12
August 1970	0.28	0.84	0.56
October 1970	0.91	0.55	−0.36
December 1970	0	0.40	0.40
February 1971	0.14	0.60	0.46
April 1971	1.12	0.93	−0.19
June 1971	2.39	2.36	−0.03
August 1971	0.68	0.84	0.16
October 1971	0.18	0.63	0.45
December 1971	0.42	0.41	−0.01
February 1972	0.30	0.63	0.33
April 1972	0.59	0.43	−0.16
June 1972	1.48	1.47	−0.01
August 1972	0.28	0.70	0.42
October 1972	0.45	0.47	0.02
December 1972	0.44	0.26	−0.18
February 1973	0.65	0.58	−0.07
April 1973	0.80	0.51	−0.29
June 1973	0.84	0.84	0
August 1973	0.05	0.60	0.55
October 1973	0.21	0.37	0.16
February 1974	0	0.47	0.47
April 1974	1.07	0.12	−0.95
May 1974	0	0.71	−0.71
November 1974	−0.29	0.49	0.78
January 1975	0.10	0.57	0.47
March 1975	0	0.60	−0.60
May 1975	1.34	2.03	0.69
July 1975	1.91	2.16	0.25
November 1975	−0.04	0.63	0.67
May 1976	1.29	0.47	−0.82
March 1977	0.32	0.59	0.27
February 1978	0.34	0.43	0.09

Downloaded from ascelibrary.org by San Jose State University on 10/16/25. Copyright ASCE. For personal use only; all rights reserved.

TABLE 4. *(Continued)*

(1)	(2)	(3)	(4)
November 1978	0.35	0.49	0.14
May 1979	0.10	0.44	0.34
November 1979	0.29	0.55	0.26
December 1979	0.29	0.40	0.11
September 1980	0.45	1.05	0.60
August 1981	1.92	0.95	−0.97
December 1982	0.84	0.31	0.53
April 1984	0.34	−0.01	−0.35
December 1984	−0.34	−0.33	0.58
July 1985	0.45	0.88	0.43
May 1986	0.78	0.58	−0.20
March 1987	0.16	0.34	0.18
April 1988	0.11	0.38	0.27
January 1989	0.26	0.45	0.19
September 1990	0.46	0.34	−0.12

network. Cases can be presented in order or at random. For this problem, random presentation of training cases produced the best results.

MODELING RESULTS

Tables 4 and 5 show a comparison between predictions made using the test sets and the actual change in the construction cost index. Table 4 shows the one-period-ahead prediction, and Table 5 shows the six-periods-ahead prediction. The model inputs for each test set are from data available from one month or six months earlier. For example, using the six-month model, the April 1976 prediction is based on input information from October 1975. The test sets were selected at random from the period 1967–90. The one-month prediction model has 63 cases, and the six-month test set has 66 cases. No test set case is included in the training set. The results are from the neural networks that minimized the mean square error of the test sets.

The neural networks produced poor predictions of the changes in the construction cost index. In the one-month model, the difference between the actual percentage change and the predicted percentage change was less than or equal to 0.25 in only 20 of 63 cases. For the six-month model, the difference between the actual percentage change and the predicted change was less than or equal to 0.5 for 20 of 66 cases.

COMPARISON WITH OTHER FORECASTING TECHNIQUES

To test the usefulness of the neural-network model further, the one-month prediction of the percentage change in the index was compared with the predictions made by two simple forecasting techniques. An exponential smoothing model and a linear regression model were also built. Forecasts were made using the three models for the period of March 1991–October 1992.

Exponential smoothing provides a forecasting technique that is widely used due to its simplicity, computational efficiency, and reasonable accuracy (Montgomery and Johnson 1976). Exponential smoothing is described by the following equations:

Downloaded from ascelibrary.org by San Jose State University on 10/16/25. Copyright ASCE. For personal use only; all rights reserved.

TABLE 5. Six-Period Neural-Network Prediction Compared with Actual Results

Month (1)	Actual percentage change (2)	Predicted percentage change (3)	Difference (4)
April 1976	1.49	3.31	1.82
October 1976	6.49	6.08	−0.41
August 1977	4.24	5.83	1.59
July 1978	5.58	5.19	−0.39
April 1979	1.24	2.31	1.07
October 1979	8.18	3.45	−4.73
April 1980	0.69	2.69	2.00
May 1980	0.27	2.80	2.53
February 1981	2.10	2.55	0.45
January 1982	1.25	3.27	2.02
May 1983	2.21	2.83	0.62
June 1984	1.25	1.72	0.47
September 1984	1.42	2.55	1.13
May 1985	0.34	−0.37	−0.71
December 1985	0.66	1.40	0.74
October 1986	2.42	3.57	1.15
August 1987	2.10	3.66	1.56
September 1988	1.16	4.24	3.08
June 1989	0.69	1.73	1.04
May 1990	0.85	0.96	0.11
February 1991	2.26	2.05	−0.21
April 1968	2.50	2.55	0.05
June 1968	4.86	3.24	−1.62
August 1968	4.87	4.44	−0.43
October 1968	5.56	4.42	−1.14
December 1968	3.93	3.12	−0.81
February 1969	4.73	2.01	−2.72
April 1969	4.74	2.67	−2.07
June 1969	5.44	3.45	−1.99
August 1969	4.88	4.96	0.08
October 1969	3.86	5.03	1.17
December 1969	2.69	3.31	0.62
February 1970	1.46	1.70	0.24
April 1970	2.27	1.38	−0.89
June 1970	5.09	2.88	−2.21
August 1970	7.55	4.08	−3.47
November 1970	6.50	4.35	−2.15
January 1971	3.49	3.04	−0.45
March 1971	5.02	2.71	−2.31
May 1971	6.84	3.28	−3.56
July 1971	9.45	5.32	−4.13
September 1971	9.55	7.58	−1.97
November 1971	6.85	6.62	−0.23
January 1972	4.04	3.54	−0.50
March 1972	2.55	3.22	0.67
May 1972	4.04	3.41	−0.63
July 1972	4.86	5.43	0.57
September 1972	4.99	6.58	1.59

J. Constr. Eng. Manage., 1994, 120(2): 306-320

Downloaded from ascelibrary.org by San Jose State University on 10/16/25. Copyright ASCE. For personal use only; all rights reserved.

TABLE 5. *(Continued)*

(1)	(2)	(3)	(4)
November 1972	4.05	5.01	0.96
January 1973	3.60	2.46	−1.14
March 1973	3.94	1.62	−2.32
May 1973	3.84	2.24	−1.60
July 1973	3.33	3.82	0.49
September 1973	3.64	5.25	1.61
November 1973	2.85	3.87	1.02
January 1974	2.02	1.46	−0.56
March 1974	0.59	0.60	0.01
July 1974	4.91	4.86	−0.05
September 1974	7.14	4.66	−2.48
October 1974	6.62	3.78	−2.84
February 1975	2.44	2.56	0.12
April 1975	1.64	2.01	0.37
June 1975	4.72	3.76	−0.96
August 1975	6.42	6.88	0.46
October 1975	6.89	8.53	1.64
December 1975	4.01	4.22	0.21

$$F(0) = A(1) \tag{12}$$

$$F(t) = \alpha A(t) + (1 - \alpha)F(t - 1) \tag{13}$$

$$f(t + t) = F(t) \tag{14}$$

where $A(t)$ = the actual data in period t; $F(t)$ = the smoothed value for period t; $f(t)$ = the forecast for period t; α = the smoothing constant; and τ = the time from t (Chang and Sullivan 1993). The selection of the smoothing constant α is important in determining the operating characteristics of the exponential smoothing model. The larger the value of α, the more rapid is the response to changes in $F(t)$.

Exponential smoothing was used to develop a one period ahead forecast for the monthly percentage change in the construction cost index. An α of 0.05 was used. This was the smoothing constant value that minimized the mean square error of the forecast for the period February 1967–February 1991. This low smoothing constant produces a forecast that does not respond rapidly to large variations in the index, although it produces predictions that have a lower overall error.

A simple regression model was also developed that provided a one-month ahead forecast. The model used the time period as the predictor variable and the percentage change in construction cost index as the response variable. For the first forecast, in March 1991, the model was calibrated using the historical data from February 1967–February 1991. For each following forecast, the model parameters were updated to include the newly available actual percentage change.

Fig. 3 shows the actual change, the neural network prediction, the exponential smoothing prediction, and the regression prediction. A sum of the squares of errors (SSE) was calculated for each type of prediction. The formula for this calculation is:

$$\text{SSE} = \sum_{i=1}^{n} (P - A)^2 \tag{15}$$

Downloaded from ascelibrary.org by San Jose State University on 10/16/25. Copyright ASCE. For personal use only; all rights reserved.

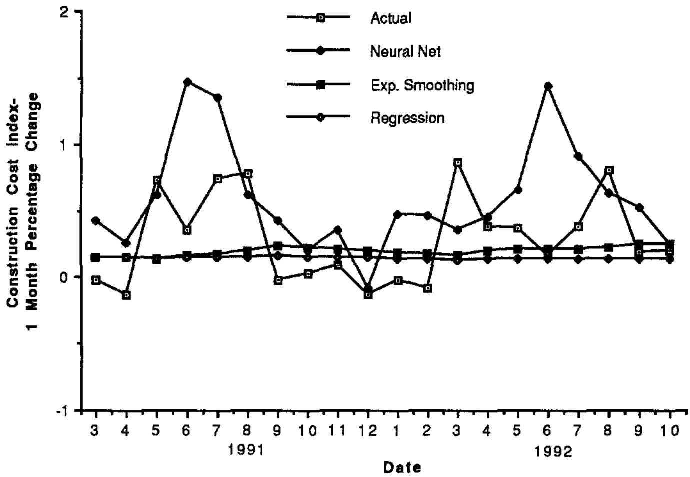

FIG. 3. Comparison of One-Period Predictions

where P = the prediction; A = the actual one-month percentage change; and n = the number of months predicted. The exponential smoothing model produced a SSE equal to 2.45. For regression the SSE was 2.65. The SSE for the neural network was 5.31. Although the exponential smoothing and regression techniques are unable to react to large variations in the index, they are overall more accurate predictors than the neural–network model.

CONCLUSIONS

This study has illustrated the difficulties in applying neural networks to predict changes in the construction cost index. Back-propagation neural networks have powerful pattern-recognition capabilities. However, the factors that influence construction prices are many and complex. It may be impossible to develop a simple neural-network model that can consistently make meaningful predictions of price changes. The model, as currently formulated, incorporates several major macroeconomic indicators of trends in the economy and the construction industry. Additional research is needed to identify variables better that can be used to develop predictive models. Possibly the neural-network model could be enhanced by adding more input parameters. However, this diminishes the attractiveness of the neural-network approach because it would require extensive data collection to make a single prediction.

The behavior of the economy is cyclical. To develop a meaningful training set required data from as far back as 1967. However, the economy has passed through several cycles in that time period. We are now experiencing a period of low interest rates and low inflation. Much of the data collected were from periods of high inflation, where variations in the index were more volatile. We now seem to be in a period where a different mechanism is affecting the change in the index. It is possible to conclude that the factors affecting prices are so complex that data collected from other business cycles have little relevance to predicting how prices behave now.

An additional problem with the model is the difficulty in predicting when extreme variations in prices will occur. Few training cases exist for extreme variations. Particularly, relatively few decreases in the construction cost

J. Constr. Eng. Manage., 1994, 120(2): 306-320

Downloaded from ascelibrary.org by San Jose State University on 10/16/25. Copyright ASCE. For personal use only; all rights reserved.

index have occurred. The network had trouble predicting price decreases because there were so few decreasing patterns in the test set.

APPENDIX I. REFERENCES

Caudill, M., and Butler, C. (1990). *Naturally intelligent systems*. The MIT Press, Cambridge, Mass.

Chang, Y.-L. and Sullivan, R. S. (1993). *Mac QS quantitative systems*. Prentice-Hall, Inc., Englewood Cliffs, NJ.

Freeman, J. A., and Skapura, D. M. (1991). *Neural networks: Algorithms, applications and programming techniques*. Addison-Wesley Publishing Co., Reading, Mass.

Grogan, T. (1992). "Cost history-keeping track of a moving target." *ENR*, 228(13), 42–47.

Hillebrandt, P. M. (1974). *Economic theory and the construction industry*. The Macmillan Press, Ltd., London, England.

Katz, J. O. (1992). "Developing neural network forecasters for trading." *Tech. Anal. of Stocks and Commodities*, Apr., 58–70.

Montgomery, D. C., and Johnson, L. A. (1976). *Forecasting and time series analysis*. McGraw-Hill Book Co., New York, N.Y.

Ostwald, P. E. (1984). *Cost estimating*. Prentice-Hall, Inc., Englewood Cliffs, N.J.

Zaremba, T. (1990). "Case study III: Technology in search of a buck." *Neural network PC tools, A practical guide*, R. C. Eberhart and R. W. Dobbins, eds., Academic Press, Inc., San Diego, Calif., 251–284.

APPENDIX II. NOTATION

The following symbols are used in this paper:

$A(t)$ = actual data in period t;
E_p = calculated error for training pattern p;
$F(t)$ = smoothed value for period t;
$f(t)$ = forecast for period t:
f_j^h = output function for jth unit of hidden layer;
$f_j^{h'}$ = derivative of output function for jth unit of hidden layer i;
f_k^o = output function for kth unit of output layer o;
$f_k^{o'}$ = derivative of output function for kth unit of output layer o;
i_{pj} = output of hidden layer node j for training pattern p;
L = number of nodes in hidden layer;
M = number of nodes on output layer;
N = number of nodes in hidden layer;
net_{pj}^h = net input to jth unit on hidden layer h for training pattern p;
= net input to kth unit on output layer o for training pattern p;
o_{pk} = output at the kth unit of output layer for training pattern p;
t = time period;
w_{ji}^h = weight on connection from ith input to jth unit on hidden layer;
w_{jk}^o = weight on connection from jth unit on hidden layer to kth unit on output layer;
X = original data value;
$X_{\max}$ = maximum observed value for data set;
$X_{\min}$ = absolute value of minimum negative observed value for data set;
XT = normalized value;
$\mathbf{x}_p$ = input vector for training pattern p;

J. Constr. Eng. Manage., 1994, 120(2): 306-320

Downloaded from ascelibrary.org by San Jose State University on 10/16/25. Copyright ASCE. For personal use only; all rights reserved.

x_{pi} = input for pattern p from node i on input layer;
y_{pk} = actual output at node k for pattern p;
δ^h_{pj} = error in hidden layer h at jth node for pattern p;
δ^o_{pk} = error at output node k on output layer o;
α = smoothing parameter for exponential smoothing;
η = neural-network learning rate;
θ^h_j = bias term to jth unit on hidden layer h; and
τ = time from t.

Downloaded from ascelibrary.org by San Jose State University on 10/16/25. Copyright ASCE. For personal use only; all rights reserved.