Application of Artificial Neural Networks on Predicting Corrosion Rates of Carbon Steel in Taiwan Industrial Zones

Hsiang-Teng Lin, Chien-Ming Lo and Min-Der Lin[*]

Department of Environmental Engineering, National Chung Hsing University, Taichung, Taiwan, R.O.C.

Abstract—**This study employed artificial neural network (ANN) to develop a regional forecasting model to predict atmospheric corrosion rates of carbon steels within general industrial zones and coastal industrial zones in Taiwan. Analyzed data are based on the results of metal atmospheric corrosion rates monitoring project executed by The Institute of Harbor & Marine Technology Center in Taiwan. The result shows that sulfur dioxide deposition is the most significant factor to impact carbon steel corrosion rate in general industrial zones. However, for coastal industrial zones both sulfur dioxide deposition and chloride deposition are significant factors. The results reveal that the corrosion rates predicted by ANN have the most accurate performance. Furthermore, duplicating the extreme values of training set data of ANN can reduce the errors of corrosion rates' prediction. As for corrosion classification category predictions, the results show that ANN can accurately predict the cases for coastal industrial zones, but there are up to 24% of misjudgments for cases of general industrial zones.**

Keywords-atmospheric corrosion; artificial neural network; carbon steel

I. INTRODUCTION

Due to the fact that the metal atmospheric corrosion phenomenon is affected by many influencing parameters and the prediction about corrosion rates is very difficult, many studies have focused on either the effect of environmental variables on corrosion kinetics or the developments of metal corrosion rate prediction models. Feliu, Morcillo and Feliu Jr. collected worldwide atmospheric corrosion and environmental data from a comprehensive literature survey to establish general corrosion damage functions for mild steel, zinc, copper and aluminum, in terms of simple meteorological and pollution parameters [1, 2]. The product of the chloride level and the temperature was found to be a very important determinant of mild steel, zinc and copper corrosion. Besides, the product of the SO_2 level and the wetness time was highly relevant to mild steel corrosion. Tidblad and Kucera collected and analyzed atmospheric corrosion data of carbon steel, zinc, copper, limestone and paint coated steel from 16 test sites in Asia and Africa [3]. Three exposure periods include one, two and four years were presented. It is found that SO_2 and acid rain are the most important factors for all materials. On the other hand, HNO_3 shows correlation to corrosion of zinc and limestone. Sabah, Charles, Siddiqui, and Saleh employed regression analysis and additive models using median polish to investigated the atmospheric corrosions of various metals

caused by the degrading effects of air pollutions in Oman [4]. Aluminum, brass, copper, epoxy, galvanized, mild steel and stainless steel were used for investigation. The results showed that copper and mild steel were the most corrosive metals while stainless steel and epoxy were the least corrosive, and carbonates were the main component of corrosion, followed by chlorides and sulphates.

Mendoza and Corvo used a stepwise multiple regression technique to develop an atmospheric corrosion model for steel in Cuba, including outdoor-indoor exposure based on the influence of time of wetness and pollutants[5]. Conditions of different atmospheres, such as rural, urban-industrial, and coastal, are investigated. The results indicated that the classification of corrosivity based on environmental data according to ISO 9223 is not in agreement with that based on the corrosion rates. The influence of time and quantity of rain is also found to be very important for characterizing differences between indoor and outdoor corrosion.

Although there are many studies using traditional regression models to predict atmospheric corrosion rates of metals, and investigate the relation between corrosion factors and corrosion rates, but the results are not quite satisfied. Instead, artificial neural network (ANN) is a potential candidate to establish a metal corrosion prediction model, without regarding the complex mechanisms about physical diffusions and chemical reactions.

Cai, Cottis, and Lyon used world-wide data compiled from 42 references for the atmospheric corrosion of steel and zinc to develop an ANN model for corrosion rates prediction [6]. Input variables include temperature, time of wetness, exposure time, sulphur dioxide concentration and chloride concentration. The results showed that the ANN model is promising and accounts for about 70% of the variance of the atmospheric corrosion of steel and zinc.

To predict the corrosion rate of carbon steel in the context of the Iberoamerican Corrosion Map (MICAT) Project, which includes seventy-two test sites in fourteen countries throughout Iberoamerica, Pintos, Queipo, Troconis de Rincón, Rincón, and Morcillo developed ANN and regression analysis models as functions of time of wetness, chloride deposition rate, sulfate deposition rate, relative humidity, precipitation, and temperature [7]. The results indicated that the ANN model exhibited superior performance than the regression model, and

is able to provide reasonable corrosion rates for a variety of climatological and pollution conditions.

Díaz, López, and Rivero used an ANN model and experimental corrosion data of low alloy steel subtracts in three test sites in Uruguay to develop the damage function of carbon steel [8, 9]. The ANN numerical model showed attractive results regarding goodness of fit and residual distributions.

Kenny, Paredes, Lacerda, Sica, Souza, and Lázaris applied ANN to predict low-carbon steel, copper and aluminum corrosion rates according to environmental parameters along the transmission electric energy lines in the area of Sao Luis-Maranhao, Brazil [10]. Exposure time, relative humidity, time of wetness, precipitation, wind velocity, solar radiation, chloride ions, sulfur compounds, dustfall, and temperature were employed as the inputs of ANN model. The obtained results indicate that ANN can be used as a good corrosion estimator.

Jančíková, Kreislová, Koštial, Ružiak, and Bogar utilized ANN to predict atmospheric corrosion of steel in Czech [11]. Temperature, relative humidity, precipitation, pH of rainfall, air pollution by sulphur dioxide and exposition time were used as an input vector. Corrosion weight loss of structural carbon steel represented an output vector. It is found that ANN is advantageous and the relative error of steel corrosion loss prediction is only 6 %.

Halama, Kreislova, and Lysebettens used data from long term exposure (1968 to 1995) of carbon steel samples from three different locations in the Czech Republic to develop an ANN model to predict atmospheric corrosion of carbon steel [12]. The results showed that the ANN model is robust and relatively fast and precise, with errors below 20%.

This study employs ANN and the data of atmosphere corrosion investigation of Carbon steel in Taiwan to develop a corrosion prediction model. The data were monitored in different industrial zones and collected by The Institute of Harbor and Marine Technology (IHMT) from July, 2009 to March, 2011. The ANN model aims to predict the atmospheric corrosion rates of carbon steel in Taiwan's industrial area.

II. METHODOLOGY

A. Preparation of Data Sets

This study used the ISO corrosion environment category to classify the original corrosion rates of carbon steel in Taiwan. Since there are not C1 and C2 categories monitored in Taiwan, these two categories are not considered in this study. However, category C5+ is further divided into C6, C7 and C8 categories, as shown in TABLE I.

TABLE I. CATEGORY OF CARBON STEEL CORROSION RATES

Category	Carbon Steel Corrosion Rates, γ_{corr}
C3	$25 < \gamma_{corr} \le 50$ μm/yr
C4	$50 < \gamma_{corr} \le 80$ μm/yr
C5	$80 < \gamma_{corr} \le 200$ μm/yr
C6	$200 < \gamma_{corr} \le 300$ μm/yr
C7	$300 < \gamma_{corr} \le 400$ μm/yr
C8	$400 < \gamma_{corr}$ μm/yr

Due to the corrosion rates of carbon steel is closely related to the concentrations of environmental pollutants and meteorology variables, this study summarized variables from previous studies, and conduct correlation analysis of these variables and carbon steel corrosion rates. The results are listed in TABLE II.

Since small amounts of extreme values contained in the training data set of ANN will affect the training results and prediction performance. This study duplicated the extreme values (DEV) in the training data set to let ANN get more familiar with them, and attempt to increase the prediction ability and accuracy of ANN. Extreme values in this study are those corrosion rates belong to C3 and C+ (i.e., C6~C8) categories.

B. Artificial Neural Network

An ANN is a highly connected array of elementary processors called neurons. The multi-layered perceptron (MLP) ANN is the most widely used model. It consists of one input layer of neurons, one or more hidden layers of neurons, and a final layer of output neurons. Each neuron in a layer is connected to the neurons in the adjacent layer with different weights. Signals pass from the input layer to the output layer. In the hidden and the output layers, each neuron receives signals from the neurons of the previous layer. The passing signals are weighted by the activation (or transfer) functions between neurons.

The back-propagation network (BPN) ANN is one of the widely used MLP ANN models. The principle of BPN is using steepest gradient descent method to minimize the deviations between observed and predicted data. Such a deviation minimization is achieved by the learning process. The learning process take one training sample for each time, and one "learning epoch" is completed when all training samples are all finished. A BPN can undergo sample training for several learning epoch until it is converged. Figure 1 illustrates a typical architecture of BPN ANN, and the important elements of ANN include input layer, hidden layer, and output layer, as described below [13]:.

TABLE II. THE VARIABLES USED BY ANN TO PREDICT ATMOSPHERIC CORROSION RATES OF CARBON STEEL

Variables	Unit
SO₂ Deposition Rate	mg/m²day
Chloride Deposition Rate	mg/m²day
Exposure Time	hour
Time of Wetness	hour
Rain Duration Time	hour
Solar Radiation	hour
Rainfall	mm
Average Temperature	℃
Average Wind Velocity	m/s
Average Wind Direction	360 degree

1. Input Layer: Input layer is used to input variables to ANN. The number of neurons is problem-depend, and the input variables are based on the selection of the independent variables of the problem under investigated.

2. Hidden Layer: Hidden layer is used to transform the signals received from the input layer to suitable forms for the usage of the output layer. Originally, there were no hidden layers in the earliest ANN models, and the outputs were merely simple functions of the inputs. The setup of hidden layers between the input and output layers could significantly enhance the capabilities of ANN models.

3. Output Layer: The output layer is used to show the results of output variables of the network. The number of neurons of output layer depends on the complexity of the problem.

4. Activation function: The activation function defines the relation of input variables (X_i) and output variables (Y_i), $Y_j = f(X_j)$. Two types of activation functions are commonly used. Discrete activation functions include linear function and step function. Continuous activation functions include sigmoid function and hyperbolic tangent function.
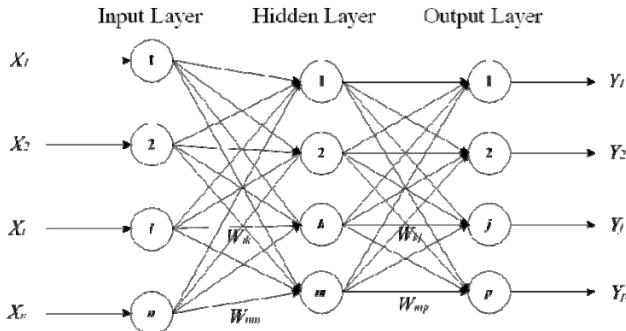

FIGURE I. BPN FRAMEWORK.

The symbols in FIGURE I are summarized in the following: X_i: variables of input layer ($i = 1,2,3,....n$); W_{ik}: weighted values between input layer and hidden layer ($k = 1, 2, 3, \cdots, m$); W_{kj}: weighted value between hidden layer and output layer ($j = 1, 2, 3, \cdots, p$); Y_j: The variables of output layer.

Root of mean square of error (RMSE), mean absolute error (MAE) and determination coefficient (R^2) were employed to evaluate the prediction performance of the ANN models, and are described in the following equations:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(T_i - Y_i)^2}{n}}$$

(1)

$$MAE = \frac{\sum_{i=1}^{n}|T_i - Y_i|}{n}$$

(2)

$$R^2 = \frac{SSR}{SST}$$

(3)

$$SSR = \sum_{i=1}^{n}(Y_i - \overline{T})^2$$

(4)

$$SST = \sum_{i=1}^{n}(T_i - \overline{T})^2$$

(5)

where T_i and Y_i are the observation and prediction values of sample i, respectively; n is the number of samples, SSR is the regression sum of squares; SST is the total sum of squares; $\overline{T}$ is the average of observation values of the samples.

Different types of ANN models were developed to respectively predict the corrosion rates and corrosion categories of carbon steel in Taiwan's general and coastal industrial areas. The models are summarized as:

1. CRG model: predict the corrosion rates of carbon steel in general industrial area.

2. CRC model: predict the corrosion rates of carbon steel in coastal industrial area.

3. CCG model: predict the corrosion categories of carbon steel in general industrial area.

4. CCC model: predict the corrosion categories of carbon steel in coastal industrial area.

The input variables for ANN models include SO_2 deposition rate, exposure time, time of wetness, rain duration time, solar radiation, rainfall, average temperature, average wind velocity, and average wind direction. However, for corrosion predictions of coastal industrial areas, one more input variable, chloride deposition rate, is included. On the other hand, the output variables for corrosion rates prediction (CRG & CRC) and corrosion categories prediction (CCG & CCC) are carbon steel corrosion rate (μm/year) and corrosion category (C3 ~ C8), respectively.

Since the collected corrosion data with low and high corrosion rates (regarding as extreme values) are relatively few, this work duplicates the extreme values in the ANN training data sets, namely Duplicate Extreme Value ANN (DEV-ANN), to improve the ANN prediction accuracy of extreme values. The results and performance of DEV-ANN will be investigated and compared with traditional ANN.

III. RESULTS AND DISCUSSION

The comparisons of the *RMSE* and R^2 of the ANN and DEV-ANN applied to corrosion rate prediction for general (CRG model) and coastal (CRC model) industrial zones are summarized in TABLE III and TABLE IV, respectively. It illustrates that both of the training and testing *RMSE*s of ANN can be reduced, and the R^2 increased, by duplicating the extreme values.

TABLE III. PERFORMANCE COMPARISON OF ANN AND DEV-ANN FOR GENERAL INDUSTRIAL ZONES (CRG MODEL).

Model	Training RMSE	Testing RMSE	R^2
ANN	0.0520	0.0436	0.68
DEV-ANN	0.0288	0.0285	0.90

TABLE IV. PERFORMANCE COMPARISON OF ANN AND DEV-ANN FOR COASTAL INDUSTRIAL ZONES (CRC MODEL).

Model	Training RMSE	Testing RMSE	R^2
ANN	0.0823	0.0658	0.61
DEV-ANN	0.0479	0.0473	0.88

The observed and predicted carbon steel corrosion rates by ANN and DEV-ANN models for different corrosion categories are listed in TABLE V and TABLE VI. The mean absolute errors are also included. It can be found that the prediction errors (*MAE*) of DEV-ANN for categories C3, C4, and C5 are significantly less than ANN. However, ANN predictions are more accurately than DEV-ANN for category C5+. Generally, The DEV-ANN can successfully and accurately predict the carbon steel corrosion rates, especially for coastal industrial zones.

TABLE V. COMPARISON OF DIFFERENT CATEGARY PREDICTIONS OF ANN AND DEV-ANN FOR GENERAL INDUSTRIAL ZONES(CRG MODEL).

Model	Items	C3	C4	C5	C5+
ANN	Observed	38.71	65.71	117.93	293.34
	Predicted	60.05	79.11	103.92	278.20
	MAE	21.34	13.40	14.01	15.14
DEV-ANN	Observed	38.71	65.71	117.93	293.34
	Predicted	45.49	71.70	110.90	239.00
	MAE	6.78	5.99	7.03	54.34

Unit: μm/y

TABLE VI. COMPARISON OF DIFFERENT CATEGARY PREDICTIONS OF ANN AND DEV-ANN FOR COASTAL INDUSTRIAL ZONES(CRC MODEL).

Model	Items	C3	C4	C5	C5+
ANN	Observed	39.94	63.97	127.26	285.09
	Predicted	175.69	59.85	118.63	289.21
	MAE	135.75	4.12	8.63	4.12
DEV-ANN	Observed	39.94	63.97	127.26	285.09
	Predicted	42.93	64.93	126.14	280.73
	MAE	2.99	0.96	1.12	4.36

Unit: μm/y

Besides metal corrosion rates, it will be very practical and convenient if the model can predict in which corrosion category a certain environmental and meteorological circumstance will result. Therefore, this study also applies ANN in corrosion category prediction. The statistical results shown in TABLE VII and TABLE VIII indicate that ANN can provide accurate predictions about carbon steel corrosion categories. Although there are up to 24% misjudgments for

general industrial zone predictions, the coastal industrial zone predictions are almost 100% accurate.

TABLE VII. STATISTICAL RESULTS FOR CORROSION CATEGORY PREDICTION BY ANN.

Model	Training RMSE	Testing RMSE	R^2
For General Industrial Zones	0.1033	0.1802	0.76
For Coastal Industrial Zones	0.0345	0.0727	0.98

TABLE VIII. ACCURACY OF CORROSION CATEGORY PREDICTION

Model	C3	C4	C5	C6	C7	C8
For General Industrial Zones	89%	78%	91%	76%	83%	100%
For Coastal Industrial Zones	100%	100%	96%	100%	100%	50%

IV. CONCLUSION

Several ANN models for prediction of carbon steel atmospheric corrosion in Taiwan industrial zones are described. The methodology involves preprocessing the monitored data set, setting the ANN model architectures, and training and testing the ANN models. The results indicate that duplicating the extreme values of training set data of ANN can significantly reduce the prediction errors. Generally, ANN can provide accurate predictions about carbon steel corrosion categories, especially for cases of coastal industrial zones.

ACKNOWLEDGMENT

REFERENCES

[1] Feliu, S., Morcillo, M., & Feliu, S. JR. (1993). The prediction of atmospheric corrosion from meterological and pollution parameters- Ⅰ. Annual corrosion. *Corrosion Science, 34*, 403-414.

[2] Feliu, S., Morcillo, M., & Feliu, S. JR. (1993). The prediction of atmospheric corrosion from meterological and pollution parameters- Ⅱ. Long-term forecasts. *Corrosion Science, 34*, 415-422.

[3] Tidblad, J. & Kucera, V. (2008). Atmospheric corrosion effects of air pollution on materials and cultural property in Asia and Africa. *17th International Corrosion Congress 2008: Corrosion Control in the Service of Society, 2*, 703-716.

[4] Sabah, A. A., Charles, S. B., Siddiqui, R. A., & Saleh, M. A. (2003). Atmospheric corrosion of metals. *JCSE, 5*, 1-22.

[5] Mendoza, A. R., & Corvo, F. (1999). Outdoor and indoor atmospheric corrosion of carbon steel. *Corrosion Science, 41*, 75-86.

[6] Cai, J., Cottis, R. A., & Lyon, S. B. (1999). Phenomenological modeling of atmospheric corrosion using an artificial neural network. *Corrosion Science, 41*, 2001-2030.

[7] Pintos, S, Queipo, N. V., Troconis de Rincón, O., Rincón, A., & Morcillo, M. (2000). Artificial neural network modeling of atmospheric corrosion in the MICAT project. *Corrosion Science, 42*(1), 35-52.

[8] Díaz, V., López, C., & Rivero, S. (2003). Low carbon steel corrosion damage prediction in rural and urban environments. *Revista de Metalurgia, 39*, 188-193.

[9] Díaz, V., & López, C. (2007). Discovering key meteorological variable in atmospheric corrosion through an artificial neural network model. *Corrosion Science, 49*, 949-962.

[10] Kenny, E. D., Paredes, R. S. C., Lacerda, L. A., Sica, Y. C., Souza, G. P., & Lázaris, J. (2009). Artificial neural network corrosion modeling for metals in an equatorial climate. *Corrosion Science, 51*, 2266-2278.

[11] Jančíková, Z., Kreislová, K., Koštial, P., Ružiak, I., & Bogar,M. (2011). Prediction of steel atmospheric corrosion by means of artificial intelligence methods. *METAL,18.-20. 5*, 1201-1205.

[12] Halama, M., Kreislova, K., & Lysebettens, J. V. (2011). Prediction of atmospheric corrosion of carbon steel using artificial neural network model in local geographical regions. *Corrosion, 67(6)*, 065004-1-065004-6.

[13] Yeh, Y.C. (1997). *Applied Artificial Neural Network*. Schalors Books.