Int. J. Electrochem. Sci., 14 **(2019)** 1516 – 1528, doi: 10.20964/2019.02.72

International Journal of
ELECTROCHEMICAL
SCIENCE
www.electrochemsci.org

Determination of Corrosion Types from Electrochemical Noise by Gradient Boosting Decision Tree Method

Qiushi Li[1], *Jihui Wang*[1,*], *Ke Wang*[1,**], *Jing Yuan*[1,2], *Xuteng Xing*[1], *Xin Liu*[1], *Wenbin Hu*[1]

[1] State Key Laboratory of Hydraulic Engineering Simulation and Safety, Tianjin University, Tianjin 300350, P R China
[2] College of Physics Electronic Information Engineering, Qinghai University for Nationalities, Xining, Qinghai 810007, China
[*]E-mail: qsli@tju.edu.cn (Qiushi Li), jhwang@tju.edu.cn (Jihui Wang), kewang@tju.edu.cn (Ke Wang)

Received: 12 October 2018 / *Accepted:* 5 December 2018 / *Published:* 5 January 2019

The corrosion behavior of X65 steel and 304 stainless steel (SS) was investigated in typical passivation, uniform corrosion and pitting solution systems by electrochemical noise (EN) technique. Eleven characteristic parameters were extracted from EN data based on statistical analysis, shot noise theory, and wavelet analysis methods. Subsequently, the data samples composed by the extracted parameters were analyzed by gradient boosting decision tree (GBDT) model. The results indicated that the proposed GBDT model could efficiently and accurately discriminate the corrosion type for data samples containing X65 steel and 304SS. The discrimination results of GBDT for the corrosion type are consistent with their corroded morphology analysis. Among the eleven parameters extracted from EN measurements, noise resistance R_n, average frequency f_n and wavelet dimension of EPN (*WD_E*) have the greatest influence on GBDT model.

Keywords: Electrochemical noise, Gradient boosting decision tree, Corrosion type, Machine learning

1. INTRODUCTION

Electrochemical noise (EN) is one of the popular electrochemical testing methods introduced firstly by Iverson and Tyagai [1, 2]. The EN method has the advantages of simple device utilization and non-destructive testing ability. In particular, the EN method has the ability to obtain the information of both corrosion rate and corrosion type, which cannot be achieved by other electrochemical testing methods [3-5]. Due to these advantages, the EN method has become a widely used electrochemical testing method in the area of corrosion research.

In corrosion study, researchers have been interested in how to identify the corrosion type of a certain material. Many researches have focused on how to extract the characteristics for identifying the corrosion types from EN data. The methods for analyzing EN data can be divided into three major types, including time-domain analysis, frequency-domain analysis, and time-frequency joint analysis. The characteristics for identifying the corrosion types include the following aspects.

(1) Statistical characteristics such as noise resistance R_n, standard deviation σ, skewness, and kurtosis [6].

(2) Characteristic charge q and characteristic frequency f_n from shot noise theory [7-9].

(3) Wavelet energy distribution from wavelet analysis [10].

(4) Correlation dimension calculated from phase space reconstruction theory [11].

(5) Wavelet-based fractal dimension obtained from wavelet theory and fractal analysis [12].

As described above, corrosion types could be identified by different data analysis methods from EN data. Therefore, the EN method can efficiently obtain a large amount of data containing the corrosion type information. Due to the efficiency of using the EN method, the corrosion type discrimination based on EN data is suitable to be performed automatically. The automatic corrosion type discrimination using the EN method is a strong advantage over other testing methods. In recent years, thanks to the rapid development of technologies such as pattern recognition and machine learning, researchers have developed many new methods for EN data analysis to distinguish and discriminate corrosion types. For examples, the cluster analysis and linear discriminant analysis have been introduced by Huang [13-15] into EN data analysis to distinguish different pitting states of low carbon steel in $NaHCO_3$+NaCl solution. BP neural network has been applied by Li [16] to discriminate uniform corrosion, pitting and passivation of 304 stainless steel. The recurrence quantification analysis method has been used by Hou [17, 18] to analyze EN data and obtain characteristic parameters which are applied to linear discriminant analysis and random forest algorithm to discriminate the corrosion types of a carbon steel.

Although cluster analysis, linear discriminant analysis, BP neural network method and random forest method have been successfully applied in the discrimination of corrosion types, the corrosion data used are mainly on single material and relatively not too much data groups. The lack of data samples may lead to the weak representation of the corrosion type discrimination. Under the background of the vigorous development of big data nowadays, material corrosion databases are developing rapidly. Therefore, how to classify a large number of corrosion type data in the form of database more efficiently, accurately and intelligently is a new requirement for the corrosion type discrimination methods.

Gradient boosting decision tree (GBDT) is an iterative decision tree algorithm consisting of multiple decision trees. The results of all trees are added together to make the final result. The GBDT algorithm has been widely used in various fields since the introduction by Friedman [19]. Recently, the GBDT algorithm has attracted more attention because of its application as a machine learning algorithm for sorting and classification. Wang [20] has presented a new photovoltaic power prediction model based on the GBDT by using historical weather data and photovoltaic power output data. Li [21] has applied the GBDT for the detection of auxiliary through lanes at intersections. In addition to its strong predictive ability and good stability, GBDT also has the ability to effectively handle mixed features, which makes the GBDT algorithm very suitable for dealing with the corrosion types of various materials.

In this paper, the corrosion behavior of X65 steel and 304SS in passivation, uniform corrosion and pitting systems are collected by EN method. The obtained original EN data of these two materials are then mixed as the data samples. The gradient boosting decision tree (GBDT) model is used to discriminate the corrosion types of the mixed data sample. The results of corrosion type discrimination using this GBDT method are compared with the R_n-f_n scatter pattern method.

2. METHODOLOGY

2.1 Gradient boosting decision tree

The gradient boost decision tree (GBDT) is formed by combining the gradient boosting algorithm with the decision tree algorithm. GBDT uses decision trees as weak learners and builds the model in a stage-wise manner by optimizing the loss function [20]. After *N* times of iterations, a total of *N* decision trees (weak classifiers) will be obtained by GBDT. A final GBDT classifier model is constructed from these decision trees by weighting or voting on these weak classifiers.

For a *k*-class problem, the basic steps for training the GBDT classifier model are as follows.

Define *X* as the training sample, import *n* training samples *X* and set the relevant parameters. The number of iterations is set to *N*. The function $F_k(X)$ represents the estimated value of the sample *X* belonging to the *k*th class.

The function $p_k(X)$ is the probability of the sample *X* belonging to the *k*th class. $p_k(X)$ can be obtained from $F_k(X)$ according to equation (1).

$$p_k(X) = exp\big(F_{ki}(X)\big) \Big/ \sum_{l=1}^{K} exp\big(F_l(X)\big) \tag{1}$$

The loss function *L* is shown in equation (2), where y_k is the estimated value of the input data. If the sample *X* is in the *k*th class, y_k =1, otherwise y_k =0.

$$L = -\sum_{k=1}^{K} y_k \log p_k(X) \tag{2}$$

The gradient direction of residual reduction is calculated by using the loss function *L*, as shown in equation (3).

$$\tilde{y}_k = -[\partial L / \partial F_k(X)] = y_k - p_k(X) \tag{3}$$

The gain of each leaf node ω_{jk} is calculated according to equation (4). Let *J* represent the number of leaf nodes. By the split of the leaf node with the largest gain, a decision tree model containing *J* leaf nodes is obtained.

$$\omega_{jk} = (K-1)\sum_{i=1}^{n} \tilde{y}_{ik} \Big/ K\sum_{l=1}^{n} |\tilde{y}_{ik}|(1-|\tilde{y}_{ik}|) \ , \quad j = 1,\ 2,\ 3,\ \ldots,\ J \tag{4}$$

Then, the next iteration is performed and the estimated values are updated. Each cycle constitutes a new decision tree until the end of *N* iterations, outputting a GBDT classifier model consisting of *N* decision trees.

3. EXPERIMENTAL

3.1 Materials and electrolytes

The materials used in this work were X65 steel and 304SS. The chemical compositions of X65 steel and the 304SS are shown in Table 1. Coupons were cut to a dimension of 10 mm × 10 mm × 3 mm. The coupons for electrochemical measurements were encapsulated by epoxy resin after welded with copper wires. All coupons were ground by SiC paper up to 1500 grit, and then cleaned with ethanol.

Table 1. Chemical compositions (wt. %) of the X65 steel and the 304SS.

Alloy	C	Si	Mn	Ni	Cr	Mo	Cu	V	Nb
X65	0.09	0.26	1.30	0.15	0.04	0.17	0.13	0.04	0.03
304SS	0.035	0.66	2.00	9.27	18.65	-	0.13	-	-

Ten solution systems were selected to set up passivation, uniform corrosion and pitting for X65 steel and 304SS, as shown in Table 2. For X65 steel, the solution systems of 1 mol·L^{-1} $NaHCO_3$, different concentrations of dilute H_2SO_4 and 1 mol·L^{-1} $NaHCO_3$ with different concentrations of NaCl are corresponded to passivation, uniform corrosion and pitting respectively [22, 23]. For 304SS, the solution systems were 0.1 mol·L^{-1} NaOH, 0.5 mol·L^{-1} H_2SO_4 and 0.1 mol·L^{-1} $FeCl_3$, which corresponded to passivation, uniform corrosion and pitting [11, 16]. All the solutions were prepared with ultrapure water and analytical reagents.

Table 2. Solution systems selected for passivation, uniform corrosion and pitting of X65 steel and 304SS.

No.	Materials	Solution	Corrosion types	Labels*
1	X65	1 mol·L^{-1} $NaHCO_3$	Passivation	P1
2	X65	0.01 mol·L^{-1} H_2SO_4	Uniform corrosion	U1-1
3	X65	0.05 mol·L^{-1} H_2SO_4	Uniform corrosion	U1-2
4	X65	0.1 mol·L^{-1} H_2SO_4	Uniform corrosion	U1-3
5	X65	1 mol·L^{-1} $NaHCO_3$ + 0.1 mol·L^{-1} NaCl	Pitting	L1-1

6	X65	1 mol·L^{-1} $NaHCO_3$ + 0.2 mol·L^{-1} NaCl	Pitting	L1-2
7	X65	1 mol·L^{-1} $NaHCO_3$ + 0.5 mol·L^{-1} NaCl	Pitting	L1-3
8	304SS	0.1 mol·L^{-1} NaOH	Passivation	P2
9	304SS	0.5 mol·L^{-1} H_2SO_4	Uniform corrosion	U2
10	304SS	0.1 mol·L^{-1} $FeCl_3$	Pitting	L2

* For the labels, P represents passivation, U represents uniform corrosion, and L represents pitting.

3.2 Electrochemical noise measurements

The electrochemical noise spectrum X65 steel and 304SS in the above proposed solution systems were measured by electrochemical workstation (Autolab 302N, Metrohm, Switzerland) with a three-electrode system. The metal coupon (X65 steel and 304SS) was served as the working electrode (WE) with a working area of 1 cm^2. The same metal as WE was used as the counter electrode (CE), while an Ag/AgCl electrode (saturated KCl) was served as reference electrode (RE). The electrochemical potential noise (EPN) was recorded between WE and RE, and the electrochemical current noise (ECN) was recorded between WE and CE. The test frequency of EN was set to 2 Hz, and the duration of each EN measurement was set to 72 h.

Before the measurements of X65 steel in the pitting systems (No.5, No.6 and No.7 solution systems), X65 steel was firstly immersed in 1 mol·L^{-1} $NaHCO_3$ solution for 2 h for metal surface passivation, and then transferred to the testing solution systems.

3.3 Feature extraction

For the collected EN data, direct current (DC) component was firstly removed from the original EN data by the five-order polynomial fitting, and then EN data were cut into continuous segments. Each EN segment consisted of 1024 points (512 s recordings). Eleven characteristic parameters of EN were calculated from each segment, i.e. 1-noise resistance (R_n), 2-characteristic charge (q), 3-average frequency (f_n), 4-standard deviation of EPN (σ_E), 5-standard deviation of ECN (σ_I), 6-skew of EPN (V_{ske}), 7-skew of ECN (I_{ske}), 8-kurtosis of EPN (V_{kur}), 9-kurtosis of ECN (I_{kur}), 10-wavelet dimension of ECN (WD_I), and 11-wavelet dimension of EPN (WD_E) [6, 10, 16]. These characteristic parameters were used as the features for the corrosion type discrimination in this work. The importance of the features contributing to the accuracy of the GBDT model was estimated by calculating the average importance of the features in every single decision tree.

3.4 Morphology characterizations

After EN measurements, coupons were firstly taken out from the immersion solution. Coupons from passivation systems were then rinsed by ultrapure water and ethanol. Whereas coupons from uniform corrosion systems and pitting systems were then dipped in the rust removal solution to remove the corrosion products. The rust removal solution was composed of 500 mL/L hydrochloric acid, 500 mL/L ultrapure water and 3.6 g/L hexamethylenetetramine. After the removal of the corrosion products, the coupons were rinsed by ultrapure water and ethanol. The corroded morphologies of X65 steel and 304SS were observed by digital microscopy (VHX-2000, KEYENCE, Japan).

4. RESULTS AND DISCUSSION

4.1 Corroded morphology

Fig. 1 illustrates the corroded morphologies of X65 steel exposed to different solution for 72 hours. In 1 mol·L^{-1} $NaHCO_3$ solution, no corrosion is observed on X65 steel, indicating that X65 steel is in the passivation state (Fig.1a). In 0.01~0.1 mol·L^{-1} H_2SO_4 solution, an overall corrosion is observed on X65 steel, indicating that the uniform corrosion has occurred on X65 steel (Figs. 1b, 1c and 1d). In $NaHCO_3$ + NaCl solution, obvious corrosion pits are observed on X65 steel, indicating that pitting corrosion has occurred as shown in Figs. 1e, 1f and 1g.

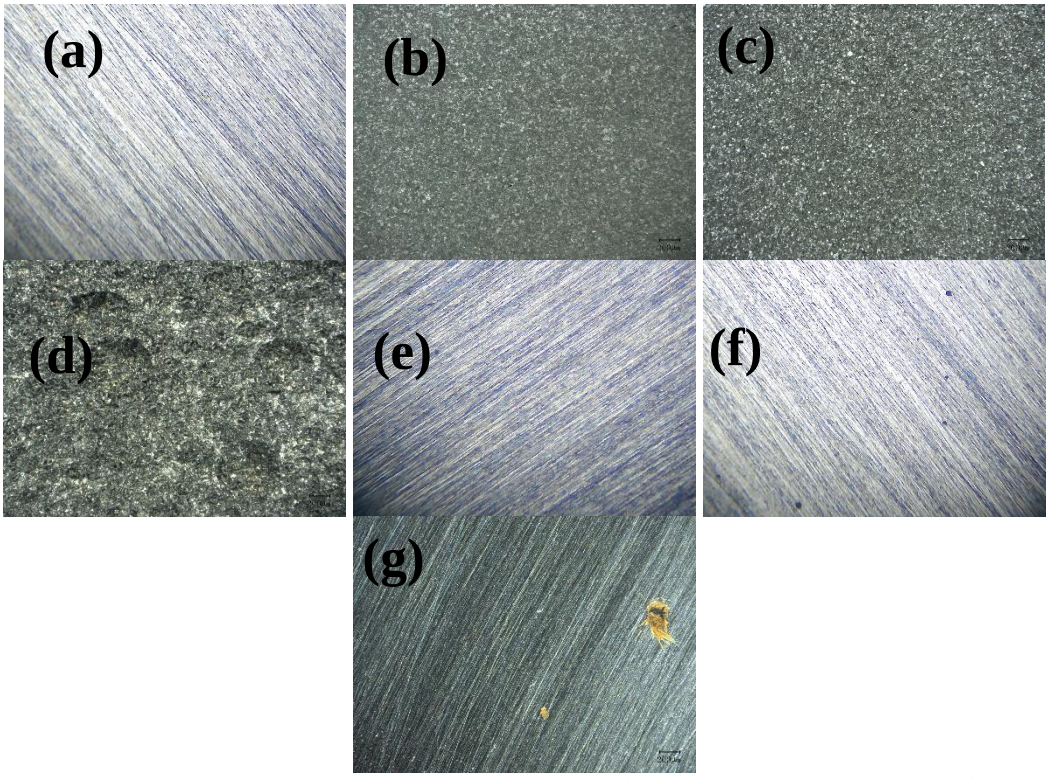


Figure 1. Corroded morphologies of X65 steel in 1 mol·L^{-1} $NaHCO_3$ (a), 0.01 mol·L^{-1} H_2SO_4 (b), 0.05 mol·L^{-1} H_2SO_4 (c), 0.1 mol·L^{-1} H_2SO_4 (d), 1 mol·L^{-1} $NaHCO_3$ + 0.1 mol·L^{-1} NaCl (e), 1 mol·L^{-1} $NaHCO_3$ + 0.2 mol·L^{-1} NaCl (f), and 1 mol·L^{-1} $NaHCO_3$ + 0.5 mol·L^{-1} NaCl (g) solutions.

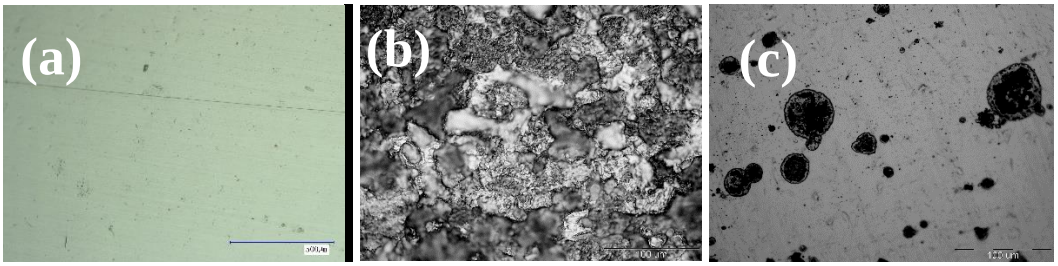
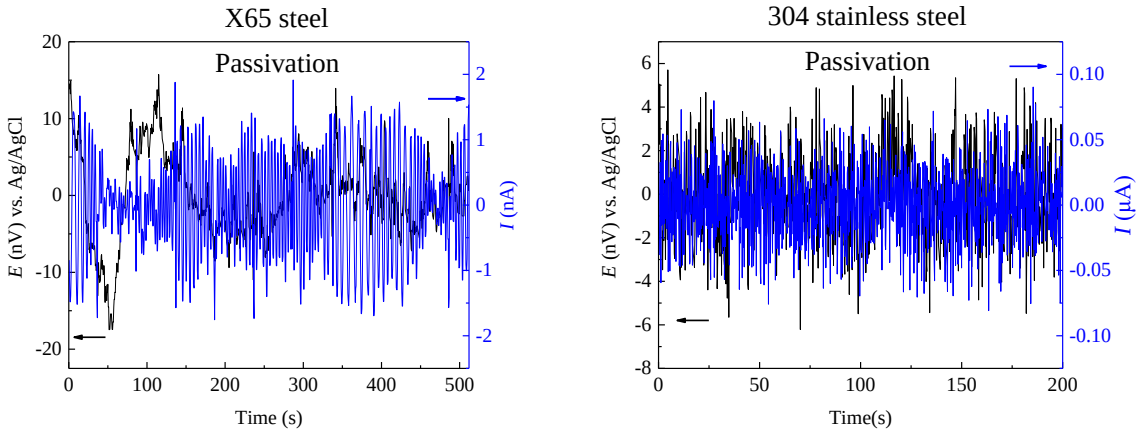


Figure 2. Corroded morphologies of 304SS in0.1 mol·L^{-1} NaOH (a), 0.5 mol·L^{-1} H_2SO_4 (b), and 0.1 mol·L^{-1} $FeCl_3$ (c) solutions.

Fig. 2 illustrates the corroded morphologies of 304SS exposed to different solution for 72 hours. In 0.1 mol·L^{-1} NaOH, 0.5 mol·L^{-1} H_2SO_4, and 0.1 mol·L^{-1} $FeCl_3$ solutions, typical passivation, uniform corrosion, and pitting corrosion are occurred on 304SS, respectively. All these results means that in the corresponding proposed solution systems the corrosion types of passivation, uniform corrosion, and pitting corrosion are occurred on X65 steel and 304SS.

4.2 EN characteristics

Fig. 3 illustrates the typical EPN and ECN spectra of X65 steel and 304SS exposed to different solution systems. As shown in Fig. 3, EN signals generated by passivation are similar to the white noise composed of frequent oscillations, whereas EN signals from uniform corrosion are very smooth with no sharp transients. The amplitudes of both EPN and ECN from uniform corrosion are larger than the amplitudes from passivation. Compared with passivation and uniform corrosion, EN signals generated by pitting corrosion show typical pitting transients, which are fast and sharp, and mostly occur simultaneously in EPN and ECN. A number of similar EN signals have been observed in various EN studies [11, 16, 18]. Therefore, these results indicate that the corrosion processes of X65 steel and 304SS in the corresponding solution systems are indeed passivation, uniform corrosion and pitting.



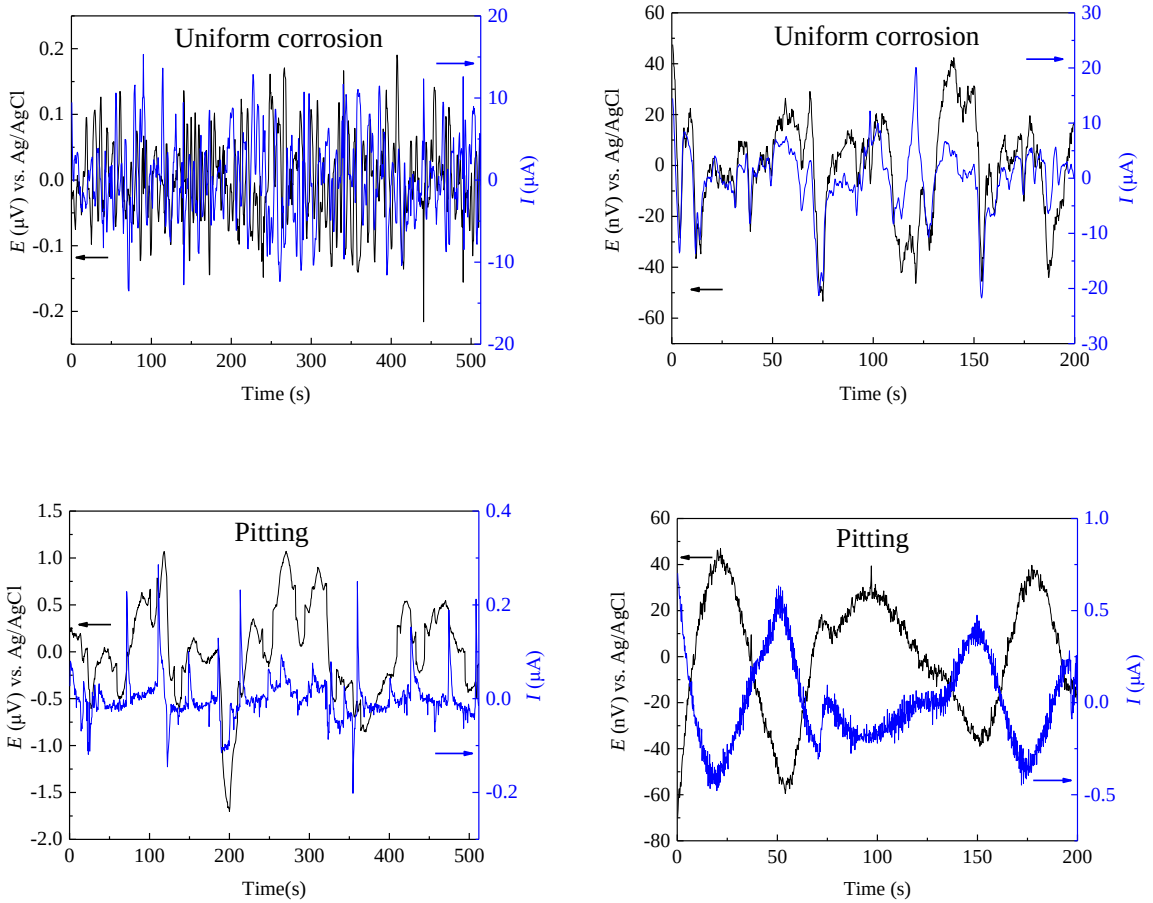


Figure 3. EPN and ECN spectra of X65 steel and 304SS in the typical passivation, uniform corrosion and pitting solution systems.

4.3 R_n-f_n discrimination analysis

Noise resistance R_n and average frequency f_n of X65 steel and 304SS in different solution systems are calculated by the corresponding EN data. Fig. 4 illustrates the scatter diagram of R_n versus f_n in different solution systems. As shown in Fig. 4, the data from X65 steel, represented by the solid points, can be divided into three regions by the solid lines (R_n=10^3-10^4 Ω·cm^2, f_n=10^3 Hz·cm^{-2}). The data from 304SS, represented by the open points, can be also divided into three regions by the dotted lines (R_n=10^5 Ω·cm^2, f_n=10^5 Hz·cm^{-2}). Cottis et al. have proposed using R_n and f_n as the indicators of corrosion types [8]. According to Cottis, for both X65 steel and 304SS, the divided three regions correspond to passivation processes (high f_n and large R_n), uniform corrosion processes (high f_n and small R_n), and pitting processes (low f_n and large R_n), respectively. These results indicate that the method of R_n-f_n scatter diagram can discriminate the corrosion type of data from X65 steel or 304SS.

However, if the data from X65 steel and 304SS are mixed, there are no lines that can divide the R_n-f_n scatter diagram into three regions. The method of R_n-f_n scatter diagram alone cannot discriminate the corrosion type of the data samples consisting of both X65 steel and 304SS.

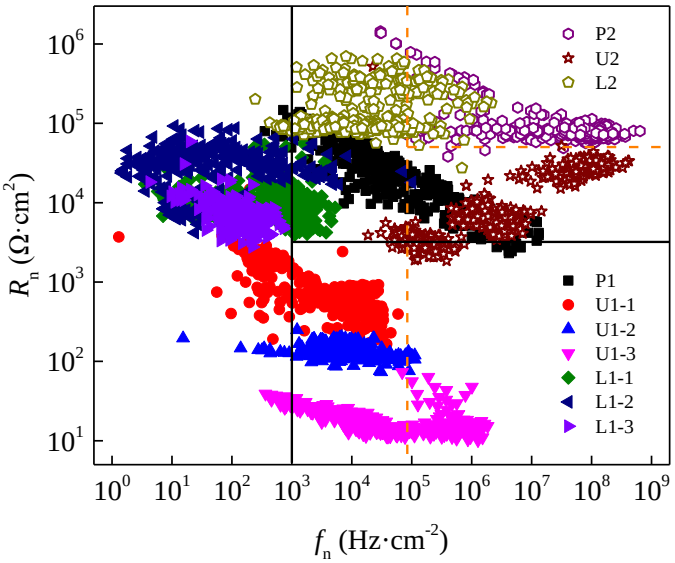


Figure 4. R_n–f_n scatter diagram of X65 steel and 304SS in typical passivation, uniform corrosion, and pitting solution systems.

4.4 GBDT discrimination analysis

After EN measurements in the different solution systems (Table 2), 3912 sets of data have been obtained, in which 2375 sets are for X65 steel and 1537 sets are for 304SS. Among the 2375 sets of X65 steel, 326 sets are passivation, 1056 sets are uniform corrosion, and 993 sets are pitting. Among the 1537 sets of 304SS, 404 sets are passivation, 749 sets are uniform corrosion, and 384 sets are pitting. All the EN data obtained from different materials in different solution systems at different time are mixed together in a random order. Eleven characteristic parameters in Table 3 are extracted from the EN data to form the data sample of the GBDT based on the statistical analysis, shot noise theory and wavelet analysis methods.

70% (2739) of all the data samples are selected as the training sets, and the rest 30% (1173) are selected as the test sets. To ensure the validity of the GBDT results, the ratios of the number of X65 steel and 304SS data in the training sets and the test sets are the same as the ratios in the original data samples. The eleven features of each sample are combined into a feature vector in a specific order. This feature vector is used as the input set of the GBDT. The corrosion type corresponding to each sample is used as the output set of the GBDT.

Table 3. Excerpted data samples for the GBDT. (Randomly taken from the mixed data samples)

Class No*	R_n	q	f_n	σ_E	σ_I	V_{ske}	I_{ske}	V_{kur}	I_{kur}	*WD_I*	*WD_E*
1	22675	1.39 E-11	27917	2.74 E-05	1.21 E-09	-2.77	0.68	19.04	4.77	1.98	1.50
1	65006	1.27 E-10	2141	8.39 E-05	1.29 E-09	0.75	0.03	4.19	4.13	1.97	1.34

1	80702	2.14 E-10	1306	1.08 E-04	1.35 E-09	-0.08	0.67	2.69	3.65	2.06	1.21
1	53090	5.18 E-11	7486	4.64 E-05	8.75 E-10	-0.61	0.38	4.98	6.36	2.10	1.20
…											
2	625.1	1.90 E-09	19861	8.27 E-05	1.32 E-07	0.63	0.83	3.30	4.24	1.78	2.01
2	115.7	1.72 E-08	16508	1.01 E-04	8.68 E-07	1.05	-0.96	5.07	4.81	2.05	1.97
2	13.61	3.75 E-09	45025	7.88 E-06	5.78 E-07	-0.075	0.029	2.72	3.05	2.16	2.19
2	11.71	4.06 E-09	59708	7.39 E-06	6.31 E-07	-0.206	0.381	3.175	3.732	1.90	1.71
…											
3	12740	1.87 E-09	304.6	2.58 E-04	2.02 E-08	0.190	9.085	4.069	112.3	1.90	1.34
3	36691	4.60 E-09	130.6	3.53 E-04	9.63 E-09	-0.629	1.045	3.197	5.423	1.94	1.36
3	6785	2.52 E-08	105.8	4.28 E-04	6.31 E-08	-0.428	0.948	2.800	6.977	1.82	1.60
3	7856	2.12 E-08	70.8	5.29 E-04	6.74 E-08	-0.224	1.571	2.461	15.43	1.84	1.50
…											

* The Class No. represents the corrosion type, 1-passivation, 2-uniform corrosion, and 3-pitting.

Fig. 5 illustrates the mean squared error (*mse*) of the GBDT model based on the eleven features. As shown in Fig. 5, *mse* value decreases as the number of iteration increases. At the beginning of the training, *mse* value decreases rapidly and then reaches the minimum value. The minimum value of test *mse* appears at the 20th iteration. After then, as the number of iteration increases, the test *mse* first increases and then decreases. When the number of iteration reaches 26, the test *mse* has decreased down to the minimum value, and then remains unchanged. Therefore, the training can be stopped at the 20th iteration, without increasing the number of iterations.

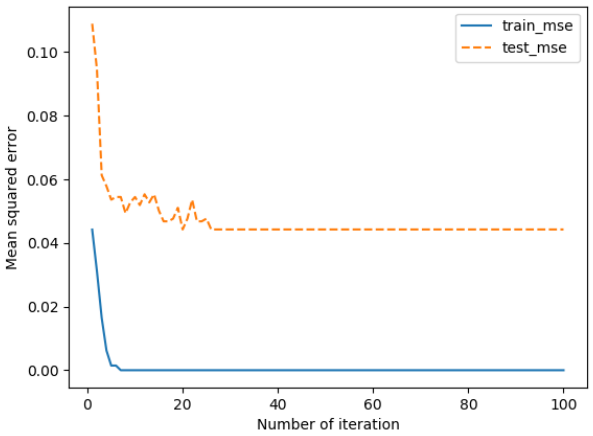


Figure 5. Mean squared error of the GDBT model based on the eleven features. (1-R_n, 2-q, 3-f_n, 4-σ_E, 5-σ_I, 6-V_{ske}, 7-I_{ske}, 8-V_{kur}, 9-I_{kur}, 10-*WD_I*, and 11-*WD_E*)

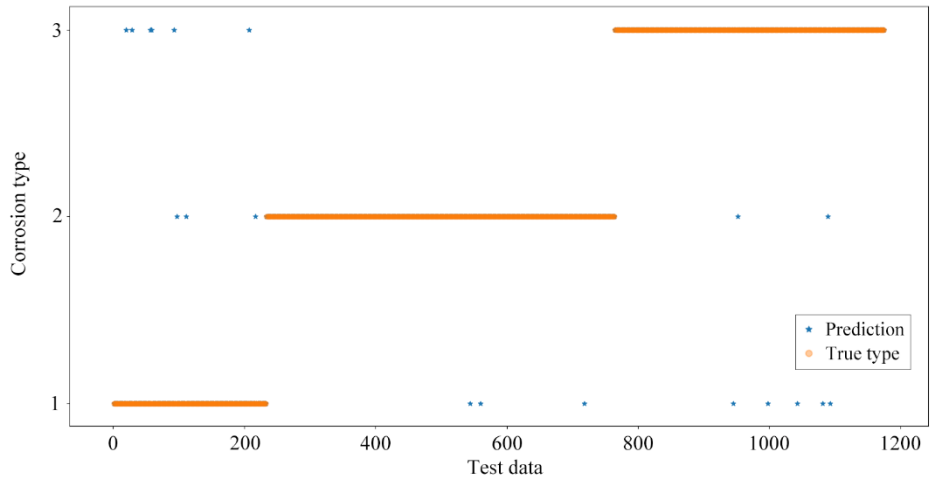


Figure 6. Comparison of the predicted corrosion types by the GBDT model with the actual corrosion types for test datasets based on the eleven features. (1-R_n, 2-q, 3-f_n, 4-σ_E, 5-σ_I, 6-V_{ske}, 7-I_{ske}, 8-V_{kur}, 9-I_{kur}, 10-*WD_I*, and 11-*WD_E*)

The test sets (1173) are presented to the GBDT model for corrosion type discrimination. Fig. 6 illustrates the accuracy of the GBDT model based on the eleven features, presenting by the comparison of the predicted corrosion types with the actual corrosion types for test datasets. The ordinate value represents the corrosion type. The values of 1, 2, and 3 represent passivation, uniform corrosion, and pitting, respectively. The abscissa represents the index of the test sets. The samples of No. 1-219 data are from the passivation solution system P1 and P2. No. 220-760 data belong to the uniform corrosion solution system U1-1, U1-2, U1-3, and U2; No. 761-1173 data are from the pitting solution system L1-1, L1- 2. L1-3, and L2. The coincidence of the predicted and true points indicates that the prediction result is accurate. As shown in Fig. 6, among the samples of No. 1-219 data that belong to the passivation, three samples are identified for the uniform corrosion, and six samples are identified for the pitting. Three samples that belong to the uniform corrosion are identified for the passivation. Two samples that belong to the pitting are identified for the uniform corrosion, and another five samples are identified for the passivation. Among the total of 1173 test sets, only 19 sets misclassified in the corrosion type prediction. The prediction accuracy of the GBDT is as high as 98.4%. Compared with the failure discrimination result of R_n-f_n scatter diagram for the two materials, GBDT is an excellent and developed model for identifying the corrosion type of data consisting of both X65 steel and 304SS.

Fig. 7 illustrates the importance of the EN variables contributing to the accuracy of the GBDT mode. In the order of importance from high to low, the eleven features are 1-R_n, 3-f_n, 11-*WD_E*, 7-I_{ske}, 4-σ_E, 10-*WD_I*, 9-I_{kur}, 6-V_{ske}, 2-q, and 5-σ_I. All these extracted features play a role on the training of GBDT. Among the eleven features, noise resistance R_n, average frequency f_n and wavelet dimension of EPN (*WD_E*) are the most significant. These three features have the strongest ability to reflect the characteristics of the corrosion type among the eleven features extracted in this paper.

For the discrimination of passivation, uniform corrosion, and pitting, the GBDT method obtains higher accuracy than the linear discriminant analysis method (88%) and the random forest method (93%)

[18]. Furthermore, the GBDT method is proved to be applicable to the data samples consisting of two materials. Hou has recently reported, based on the recurrence quantification analysis of the EN data, uniform corrosion and under deposit corrosion can be distinguished from each other by the random forest method [24]. The GBDT method deserves further studies on other corrosion types and data samples containing more materials.

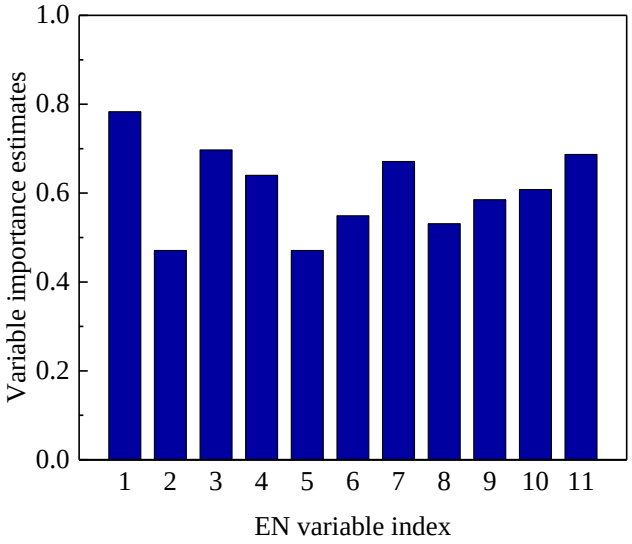


Figure 7. Importance of the EN variables contributing to the accuracy of the GBDT model. (1-R_n, 2-q, 3-f_n, 4-σ_E, 5-σ_I, 6-V_{ske}, 7-I_{ske}, 8-V_{kur}, 9-I_{kur}, 10-*WD_I*, and 11-*WD_E*)

5. CONCLUSIONS

Electrochemical noise (EN) data of X65 steel and 304SS in passivation, uniform corrosion, pitting solution systems have been analyzed by using the R_n-f_n scatter diagram method and GBDT model. By comparing these two methods, GBDT model is more suitable to discriminate the corrosion type for data samples containing different materials of X65 steel and 304SS.

The proposed GBDT model can efficiently and accurately discriminate the corrosion type of data samples consisting of X65 steel and 304SS. In typical passivation, uniform corrosion, pitting solution systems, the accuracy of the proposed GBDT model has reached 98.4%.

Among the eleven features extracted from EN measurements, noise resistance R_n, average frequency f_n and wavelet dimension of EPN (*WD_E*) have the greatest influence on the GBDT model. These three features have the strongest ability to reflect the characteristics of the corrosion type in this paper.

ACKNOWLEDGEMENTS
This work was supported by National Natural Science Foundation of China (No. 51771133 and No. 51471117) and National Key Basic Research Program of China (No.2014CB046801).

References

1. W. P. Iverson, *J. Electrochem. Soc.*, 115 (1968) 101.
2. V. A. Tyagai, *Electrochim. Acta*, 16 (1971) 1647.
3. F. Hass, A. C. T. G. Abrantes, A. N. Diógenes and H. A. Ponte, *Electrochim. Acta*, 124 (2014) 206.
4. E. C. Rios, A. M. Zimer, P. C. D. Mendes, M. B. J. Freitas, E. V. R. de Castro, L. H. Mascaro and E. C. Pereir, *Fuel*, 150 (2015) 325.
5. D. H. Xia, S. Z. Song and Y. Behnamian, *Corros. Eng. Sci. Techn.*, 51 (2016) 527.
6. C. A. Loto, *Alex. Eng. J.*, 57 (2018) 483.
7. H. A. A. Al-Mazeedi and R. A. Cottis, *Electrochim. Acta*, 49 (2004) 2787.
8. J. M. Sanchez-Amaya, R. A. Cottis and F. J. Botana, *Corros. Sci.*, 47 (2005) 3280.
9. M. G. Pujar, N. Parvathavarthini, R. K. Dayal and S. Thirunavukkarasu, *Corros. Sci.*, 51 (2009) 1707.
10. M. Shahidi, S. M. A. Hosseini and A. H. Jafari, *Electrochim. Acta*, 56 (2011) 9986.
11. D. Xia, S. Song, J. Wang, J. Shi, H. Bi and Z. Gao, *Electrochem. Commun.*, 15 (2012) 88.
12. X. Wang, J. Wang, C. Fu and Y. Gao, *Int. J. Electrochem. Sci.*, 8 (2013) 7211.
13. J. Y. Huang, X. P. Guo, Y. B. Qiu and Z. Y. Chen, *Electrochim. Acta*, 53 (2008) 680.
14. J. Y. Huang, Y. B. Qiu and X. P. Guo, *Electrochim. Acta*, 54 (2009) 2218.
15. J. Y. Huang, Y. B. Qiu and X. P. Guo, *Mater. Corros.*, 60 (2015) 527.
16. J. Li, W. Kong, J. Shi, K. Wang, W. Wang, W. Zhao and Z. Zeng, *Int. J. Electrochem. Sci.*, 8 (2013) 2365.
17. Y. Hou, C. Aldrich, K. Lepkova, L. L. Machuca and B. Kinsella, *Corros. Sci.*, 112 (2016) 63.
18. Y. Hou, C. Aldrich, K. Lepkova, L. L. Machuca and B. Kinsella, *Electrochim. Acta*, 256 (2017) 337.
19. J. H. Friedman, *Ann. Statist.*, 29 (2001) 1189.
20. J. Wang, P. Li, R. Ran, Y. Che and Y. Zhou, *Appl. Sci.*, 8 (2018) 689.
21. X. Li, Y. Wu, Y. Tan, P. Cheng, J. Wu and Y. Wang, *Int. J. Geo-Inf.*, 7 (2018) 317.
22. V. Deodeshmukh, A. Venugopal, D. Chandra, A. Yilmaz, J. Daemen, D. A. Jones, S. Lea and M. Engelhar, *Corros. Sci.*, 46 (2004) 2629.
23. L. Xu, H. Xiao, W. Shang, B. Wang and J. Zhu, *Corros. Sci.*, 109 (2016) 246.
24. Y. Hou, C. Aldrich, K. Lepkova and B. Kinsella, *Electrochim. Acta*, 274 (2018) 160.