Materials and Design 112 (2016) 410–418

Contents lists available at ScienceDirect

Materials and Design

journal homepage: www.elsevier.com/locate/matdes

# Correlation between molecular features and electrochemical properties using an artificial neural network

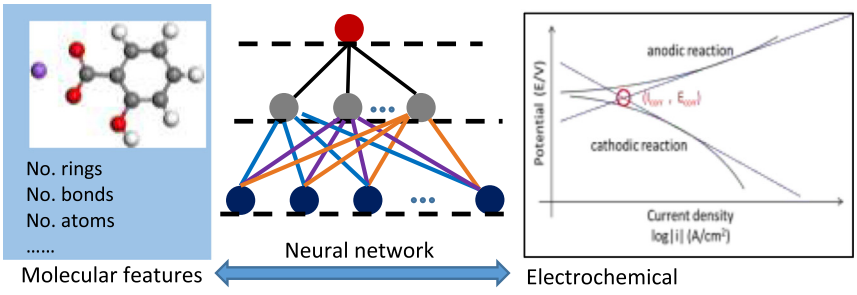
Fiona Fang Chen [a,*], Michael Breedon [a], Paul White [a], Clement Chu [a], Dwaipayan Mallick [b], Sebastian Thomas [c], Erik Sapper [d], Ivan Cole [a]

[a] CSIRO Manufacturing, Private Bag 10, Clayton South, Victoria 3169, Australia
[b] INSA Lyon, 20 Avenue Albert Einstein, 69100 Villeurbanne, France
[c] Department of Materials Science and Engineering, Monash University, Australia
[d] Boeing Research & Technology, PO Box 516, M/C S102-2152, St. Louis, MO 63166-0516, USA

## HIGHLIGHTS

- A combined experimental and modelling approach to elucidate key molecular properties of corrosion inhibiting molecules.
- Electrochemical properties are correlated with molecular features using a neural network model for inhibitor design.
- Robust predictions of electrochemical properties are achieved via an automatically trained network from measurements.
- Impact of molecular features on the effectiveness of corrosion inhibitor on an aluminium alloy is assessed and ranked.

## GRAPHICAL ABSTRACT




## ARTICLE INFO

*Article history:*
Received 10 June 2016
Received in revised form 26 August 2016
Accepted 23 September 2016
Available online 26 September 2016

*Keywords:*
Electrochemical property
Molecular structure
Corrosion inhibitor
Artificial neural network
Molecular modelling

## ABSTRACT

The increasing demand for environmentally-friendly and non-toxic coating systems from the aerospace and heavy industry sectors is driving innovation in corrosion inhibitor design and functional coating development. A fundamental understanding of how molecular structure and functionality influences the electrochemical responses of inhibited coatings is crucial for the design of effective functional coatings to replace stalwart, yet highly toxic industrial solutions. In this paper, an artificial neural network approach is presented to quantitatively study the relationship between the structural/molecular features of inhibitor compounds and their experimentally measured electrochemical properties. The presented method is applied to correlate molecular features of corrosion inhibitors with experimentally obtained corrosion potential (Ecorr), corrosion current (Icorr) and anodic/cathodic Tafel slopes. The neural network model, trained through an automatic optimization process, was able to predict the electrochemical performance for a given inhibitor molecule candidate. We will demonstrate how it can be utilised to assess the impact of molecular structure on the final effectiveness of the candidate corrosion inhibitor molecule. The presented neural network learning method could be applied to other areas in materials science for accelerating general materials discovery and functional coating design.



* Corresponding author.
*E-mail address:* Fiona.Chen@csiro.au (F.F. Chen).

http://dx.doi.org/10.1016/j.matdes.2016.09.084
0264-1275/

## 1. Introduction

Quantitative studies of structure-property relationship (QSPR) have been carried out in a wide range of applications from chemical and biological science and engineering [1–5], to the design and development of novel materials [6–8]. In corrosion science, electrochemical responses of inhibited metals may be correlated to the molecular structure of corrosion inhibitors included in the paint system [9–11]. Modelling this relationship is of particular importance in enabling and augmenting next-generation environmentally-friendly and sustainable inhibitors for corrosion protection of manufactured products, vehicle platforms and civil infrastructure.

Regression analysis is a common approach for estimating relationships among variables and has been extensively studied in literature, from the simplest form of linear regression to more sophisticated machine learning techniques [12]. As one particular type of machine learning algorithm, the artificial neural network (ANN) approach is inspired by the structure and functional aspects of biological neural networks, and is used to perform tasks such as nonlinear approximation, prediction, game/control theory, pattern recognition and classification, etc., by employing varying combinations of supervised learning, unsupervised learning and reinforcement learning [13–16]. In recent years, regression analysis based on neural networks has been extensively used in QSPR and quantitative structure-activity relationship (QSAR) studies for modelling physico-chemical properties or biological activity with respect to chemical structures [17–23]. The fundamental basis of the neural network model is to build up learning rules by training an artificial network from a given set of molecular features and then to predict the associated properties by interpreting the output of the trained network. Bayesian regularization in the network training has been shown to overcome problems of overfitting and overtraining in back-propagation neural networks and increase the robustness of the network performance [24–26].

In this work, we focus on modelling and correlating the electrochemical properties of inhibited metals, including corrosion potential, corrosion current and anodic/cathodic Tafel slopes, with structural descriptors of inhibitor molecules using the Bayesian regularization neural network approach. The proposed method is based on a three-layer neural network architecture with an automatic training process to minimise the correlation error. This optimization process effectively selects the most representative molecule subsets from the entire sample space. The remaining un-trained datasets can be used to validate the model prediction performance. The prediction capability is of great interest, and enables the effective and efficient design of new corrosion inhibitors. Given a large starting set of inhibitor molecule candidates with known structure characteristics, a robust prediction of the electrochemical response of a given inhibitor molecule can help select better candidates for experimental validation, rather than testing all possible molecules of a given chemistry. Such an approach will greatly reduce the time and effort required to identify a suitable candidate, accelerating the development section of the R&D cycle. In the experimental section, we will demonstrate that training optimization is an important process for building a robust network prediction model.

Additionally, the proposed neural network model is able to assess the importance of an individual structural descriptor to the resulting electrochemical properties, providing invaluable insights to guide molecular design of next-generation inhibitors. The assessment is based on a fraction of the network connecting weights on a particular structural descriptor over the total weights of the whole network. The higher the fraction on a molecular feature, the higher the impact of this feature to the modelled property.

Details of the artificial neural network model are presented in the next section, with an emphasis on Bayesian regularization training and optimization. In Section 3, the network model is used to correlate the corrosion potential (Ecorr), corrosion current (Icorr), anodic Tafel slope, and cathodic Tafel slope separately using a set of molecular descriptors. A number of important molecular features are identified and the simplified network performances are compared with the original model. The training optimization is demonstrated using 80% of input data for training and 20% for validation. The prediction accuracy on the validation set is presented for model justification.

## 2. Method

### 2.1. Neural network architecture

The basic architecture of an artificial neural network consists of at least three layers: an input layer, an output layer and one or more hidden layers. Neurons in the input layer represent measured molecular descriptors, while neurons in the output layer predict the engineering properties that relate to molecular characteristics. Input neurons are distributed to subsequent hidden layers and finally to the output layer via weighted connections. Each node in the network operates by taking the sum of its weighted inputs and then passes the result through a nonlinear activation function.

Consider a simple one hidden layer network illustrated in Fig. 1, let $f$ be a non-linear activation function. $X$, $A$ and $\theta$ be input neurons, the linear operator and the bias term on $X$. Similarly, let $H$, $B$ and $\phi$ denote the hidden neurons, the linear operator and the bias term on $H$, respectively. Then the output layer neurons $Y$ can be derived as

$$H = f(A * X + \theta) \tag{1}$$

$$Y = f(B * H + \varnothing) \tag{2}$$

The most frequently used non-linear activation function is a sigmoid function, which is modified from a binary step function and is in a form of:

$$f(u) = \frac{1}{1 + e^{-u/\tau}} \tag{3}$$

where $\tau$ is a tuneable number and referred as the temperature of the neuron. The higher the temperature the more gently the sigmoid changes. At very low temperature it approaches a step function. The sigmoid function introduces non-linearity in the network and is computationally attractive in its derivative calculation: $f'(u) = f(u) * (1 - f(u))/\tau$.

In general, a neural network can be reviewed as a nonlinear multivariable function of input vector $X$ and weight vector $\boldsymbol{w}$: $Y \sim (X, \boldsymbol{w})$. The weight vector is composed of weights in each layer for each neuron, including bias terms. The neural network error function, associated with given output $Y^*$, is defined as:

$$E_d = \frac{1}{2}(Y - Y^*)^2 \tag{4}$$

Substituting $H$ and $Y$ by Eqs. (1) and (2), $E_d$ is a multivariable function of $X$ and $\boldsymbol{w}$.

The neural network approach contains two phases: a training phase and a classification phase. The training phase aims to determine the best network weights by minimising the network error $E_d$. Once a network has been trained based on the experimental observation $Y^*$, the classification phase will predict the network output $Y$ for any given input $X$ and trained weight $\boldsymbol{w}$. With regarding to the computation strategy of multiple layer neural network, the classification phase is relatively simple and referred to as forward propagation, which is to feed input data in hidden layer to determine hidden neurons by Eq. (1) and then forward propagate to the output layer to calculate output neurons by Eq. (2).

The training process, on the other hand, is more complicated and performed by back propagation [27]. Back propagation is one of the most popular network training algorithms, using gradient descent or steepest descent method. It minimises the network total error by adjusting the weights in a way that the negative gradient of the error function with respect to the weights, pointing in the direction that will most quickly reduce the error.