2022

Parameters optimization in Chlor-Alkali ion-exchange membrane electrolysis process with artificial neural network

Kittapas Sukantowong
Faculty of Engineering

Parameters optimization in Chlor-Alkali ion-exchange membrane electrolysis process with artificial neural network

Mr. Kittapas Sukantowong

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering in Chemical Engineering
Department of Chemical Engineering
FACULTY OF ENGINEERING
Chulalongkorn University
Academic Year 2022

การวิเคราะห์ค่าควบคุมเหมาะสมในกระบวนการคลอร์-อัลคาไลน์แบบเมมเบรนอิเล็กโทรไลซิสโดยใช้โครงข่ายประสาทเทียม

นายกฤตภาส สุคันโธวงศ์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมเคมี ภาควิชาวิศวกรรมเคมี
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2565

Thesis Title	Parameters optimization in Chlor-Alkali ion-exchange membrane electrolysis process with artificial neural network
By	Mr. Kittapas Sukantowong
Field of Study	Chemical Engineering
Thesis Advisor	Associate Professor SUPAREAK PRASERTHDAM

Accepted by the FACULTY OF ENGINEERING, Chulalongkorn University in Partial Fulfillment of the Requirement for the Master of Engineering

..... Dean of the FACULTY OF ENGINEERING

(Professor SUPOT TEACHAVORASINSKUN)

THESIS COMMITTEE

..... Chairman

(Professor Dr. PIYASAN PRASERTHDAM)

..... Thesis Advisor

(Associate Professor SUPAREAK PRASERTHDAM)

..... Examiner

(KRITCHART WONGWAILIKHIT)

..... External Examiner

(Jakkrit Suriboot)

กฤตภาส สุคันโธวงศ์ : การวิเคราะห์ค่าควบคุมเหมาะสมในกระบวนการคลอร์-อัลคาไลน์แบบเมมเบรนอิเล็กโทรไลซิสโดยใช้โครงข่ายประสาทเทียม. (Parameters optimization in Chlor-Alkali ion-exchange membrane electrolysis process with artificial neural network) อ.ที่ปรึกษาหลัก : รศ. ดร.ศุภฤกษ์ ประเสริฐธรรม

ตั้งแต่เทคโนโลยีเมมเบรนอิเล็กโทรไลซิสถูกนำมาใช้ในกระบวนการคลอร์-อัลคาไลน์ ผู้ผลิตหลายแห่งได้ทำการเสนอการปรับปรุงเพื่อพัฒนาประสิทธิภาพ และ ลดการใช้พลังงานมาอย่างต่อเนื่อง ซึ่งทำให้ความสัมพันธ์ของตัวแปลควบคุมในกระบวนการผลิตมีค่าแตกต่างไปจากเดิม เพื่อวิเคราะห์ความสัมพันธ์ และ ค่าควบคุมเหมาะสมแบบเฉพาะ เทคโนโลยีการเรียนรู้ของเครื่องถูกมาใช้ในการวิดคราะห์ความสัมพันธ์ซับซ้อนของตัวแปล ผ่านโมดูลการสร้างชุดข้อมูลทำนายด้วยโครงข่ายประสาทเทียม (nftool) เพื่อใช้ในการวิเคราะห์หาความสัมพันธ์ระหว่าง 5 ค่าควบคุม ประกอบด้วย ความหนาแน่นกระแสไฟฟ้าต่อพื้นที่ (CD, KA/m2) ระยะเวลาการเดินเครื่อง (DOL, day) อัตราการไหลของน้ำเกลือป้อนเข้า (QFB, m3/h) อัตราการไหลของโซดาไฟป้อนเข้า (QHD, m3/h) อุณหภูมิเซลล์ (T, degC) และ 1 ผลลัพท์ คือ แรงดันไฟฟ้าต่อเซลล์ (CV, V) โดยข้อมูลถูกรวบรวมมากจากฐานข้อมูลประวัติค่าควบคุมกระบวนการผลิต (exaquantum) ผลการวิจัยแสดงให้เห็นว่าในการสร้างความสัมพันธ์ระหว่าง CD และ CV ให้ค่า RMSE 0.0167 V ซึ่งให้ผลลัพธ์ดีกว่าการทดถอยสมการเส้นตรง กรณี 2 ตัวแปร DOL ที่เป็นตัวแปรที่ 2 ช่วยส่งเสริมให้การทำนายแม่นยำขึ้นที่ RMSE 0.0065 V จากความต้านทางของเมมเบรนสูงขึ้นตามระบะเวลาการใช้งาน ในกรณีที่เพิ่มค่าควมคุมเป็น 3 ตัวแปร T ที่เป็นตัวแปรที่ 3 ให้ความแม่นยำ และ เป็นรูปแบบที่สามารถทำนาย CV ได้ดีที่สุดในงานวิจัยนี้ โดยให้ค่า RMSE ที่ 0.0043 V ซึ่งเกิดจากการปรับอุณหภูมิตามสภาะวแวดล้อมการผลิต ในส่วนของ 4 ตัวแปร ไม่พบความสัมพันธ์ที่นำไปสู่การทำนายที่แม่นยำขึ้น สามารถสรุปได้ว่าการนำโครงข่ายประสาทเทียมมาใช้ในการหาค่าควบคุมเหมาะสมมีความสามารถเพียงพอต่อการทำนายค่า CV ที่แม่นยำ หลังจากการเดินเครื่องเป็นเวลานาน และ สามารถใช้เพื่อทดถอยหาค่า CV ที่แท้จริง เพื่อเปรียบเทียบคุณภาพของเซลล์ไฟฟ้าเคมีในแต่ละสภาวะได้

สาขาวิชา วิศวกรรมเคมี ลายมือชื่อนิสิต .....

ปีการศึกษา 2565 ลายมือชื่อ อ.ที่ปรึกษาหลัก .....

6370010221 : MAJOR CHEMICAL ENGINEERING

KEYWORD: chlor-alkali, Electrochemical, data analytics, machine learning, Artificial Neural Networks, digital trandformation, commercial scale chemical plant

Kittapas Sukantowong : Parameters optimization in Chlor-Alkali ion-exchange membrane electrolysis process with artificial neural network. Advisor: Assoc. Prof. Dr. SUPAREAK PRASERTHDAM

Since ion-exchange membrane electrolysis cell has developed for producing Chlor-Alkali products. New higher efficiency and lower consumption technology are released from licensors yearly, which made the process correlation deviate from the original design. The machine learning is used with “Neural Network Fitting Tool (nftool)” in MATLAB. To find a correlation between 5 inputs consisting of current density (CD, KA/m2), operation day (DOL, day), feed brine flow rate (QFB, m3/h), feed caustic flow rate (QHD, m3/h), cell temperature (T, degC) and one output which is cell voltage (CV, V). Datasets were collected from the plant information management system “exaquantum” historian database. The result is shown only on CD as the predictor gives RMSE at 0.0167 V. In 2 predictors, DOL as the second gave RMSE at 0.0065 V, which can conclude that DOL (or clogging factor) has the most impact on CV increasing. In 3 predictors, T as the third gave RMSE at 0.0043 V, from controlled temperature set point change. Developed ANN optimization model can be used to optimize controlled parameters to predict suitable CV after a long run (high DOL) or to compare electrolysis effectiveness by regressing CV for comparing at the same condition.

Field of Study: Chemical Engineering Student's Signature .....

Academic Year: 2022 Advisor's Signature .....

ACKNOWLEDGEMENTS

I would like to express my special thanks to my thesis advisor "Associate Professor Dr. Suppareak Praserthdam" of my master's degree study, "Dr. Meena Rittiruam" who was like my thesis co-advisor for their continuous guidance and support in completing my project. Besides my advisor, I would also like to extend my gratitude to all of the thesis committee members "Professor Piyasan Praserthdam," "Dr. Kritchart Wongwailikhit" and "Dr. Jakkrit Suriboot" for their encouragement, insightful comments, and excellent suggestions. My sincere thanks also go to "Autchara Pratumtheng" department administration for providing me with advice and information, the Dean of the Faculty of Engineering, Chulalongkorn University, for offering me the flexible and online study program, and also all people at AGC for giving me great opportunities in production engineer experiences and allow me to attend the class. I want to thank Professor Piyasan Praserthdam again for giving me a perspective on a master's degree in chemical engineering. It helps me a lot in decision-making to join the master's degree program. Last but not least, I would like to thank my family: my parent "Porameth Sukantowong" and "Somsook Sae-tan" for supporting me spiritually throughout my life.

Kittapas Sukantowong

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1 : INTRODUCTION

1.1 INTRODUCTION

Chlor-Alkali process (CA) is an industrial process used to produce Chlorine (Cl_2), Caustic Soda (NaOH), and other valued byproducts such as Sodium Hypochlorite (NaOCl) and Hydrochloric acid (HCl). All products from the CA process are essential raw materials in various industries. Cl_2 is crucial in disinfection and petrochemical processes such as PVC and PC polymer. NaOH in commodity products and pH neutralization. The high demand for CA products from the rapid growth of the infrastructure industry in Southeast Asia led to the development of scale with a 2021 annual capacity of 78 million metric tons (EMR report) NaOH worldwide, and in Thailand, will increase to more than 1 million metric tons NaOH in 2025.

Since 1972 [1], they are development of membrane electrolysis cell. The primary process of CA production is the electrolysis of brine (NaCl) aqueous solution, which is more environmentally compatible than mercury cells in the past. The electrolysis process has a high electrical energy consumption, which latest technology in 2022 claim power consumption at 2,000 kWh per metric ton NaOH (information from licensor handbook). Thus, electrical power consumption is the direct operating cost of the CA process and a key factor of production profit.

Currently, CA electrolysis technology has continued to develop year by year both in catalyst terms and internal cell structure for improving product quality, cell efficiency, durability, and power consumption. All improvements in the CA electrolysis cell have a slight change in reaction kinetics and hydrodynamics inside (some improvement has a significant change in a component inside the cell for uniform internal brine circulation). The cell objective is to reduce the gap between the anode and cathode electrode from gas forming inside and causing electrical resistance. The reducing gap improves electrical conduction through electrolytes resulting in lower power consumption.

According to industrial plant production, many technologies operate together that make the conventional method (customarily calculated based on the effect of mass transfer flux) might be not compatible with all types of operating electrolysis cells. This thesis is to find the solution for predicting cell efficiency and optimum correlation by artificial neuron network (ANNs) – a machine learning algorithm - between controlled parameters, e.g., capacity current load, feed brine concentration, feed caustic concentration, cell temperature, and output parameters, e.g., depleted brine pH, product quality with unknown hidden causation individually [1]. The proposal is to find the optimum parameters and impact of each operation-controlled parameter for the electrolysis cell to operate at the lowest operation cost (mainly electrical power consumption) to lead to the highest production profit.

1.2 OBJECTIVE

This research aims to find the optimum parameters and impact of each operation-controlled parameter in CA membrane electrolysis process cells by ANNs.

1.3 RESEARCH SCOPE

1.3.1. ANNs algorithm based on scale conjugate gradient method in MATLAB neural network fitting tool (nftool) will be chosen to develop predictive models.

1.3.2. The six parameters in March 2022 and February 2023 will be collected and pre-processed for training, validation, and testing dataset. An output parameter is cell voltage (CV); input parameters are five operating parameters. To predict cell voltage in any case of operating parameters adjustment.

1.3.3. The optimum number of hidden nodes and hidden layers in ANNs that provides the highest testing performance - determined by minimum root mean square error (RMSE) - from varying the number of hidden nodes from 5 to 10 and hidden layers from 1 to 5 by trial and error will be chosen to train and compare the impact of each input parameter.

1.3.4 RMSE of the train, validate, and test dataset will determine the impact of each input parameter.

1.4 RESEARCH METHODOLOGY

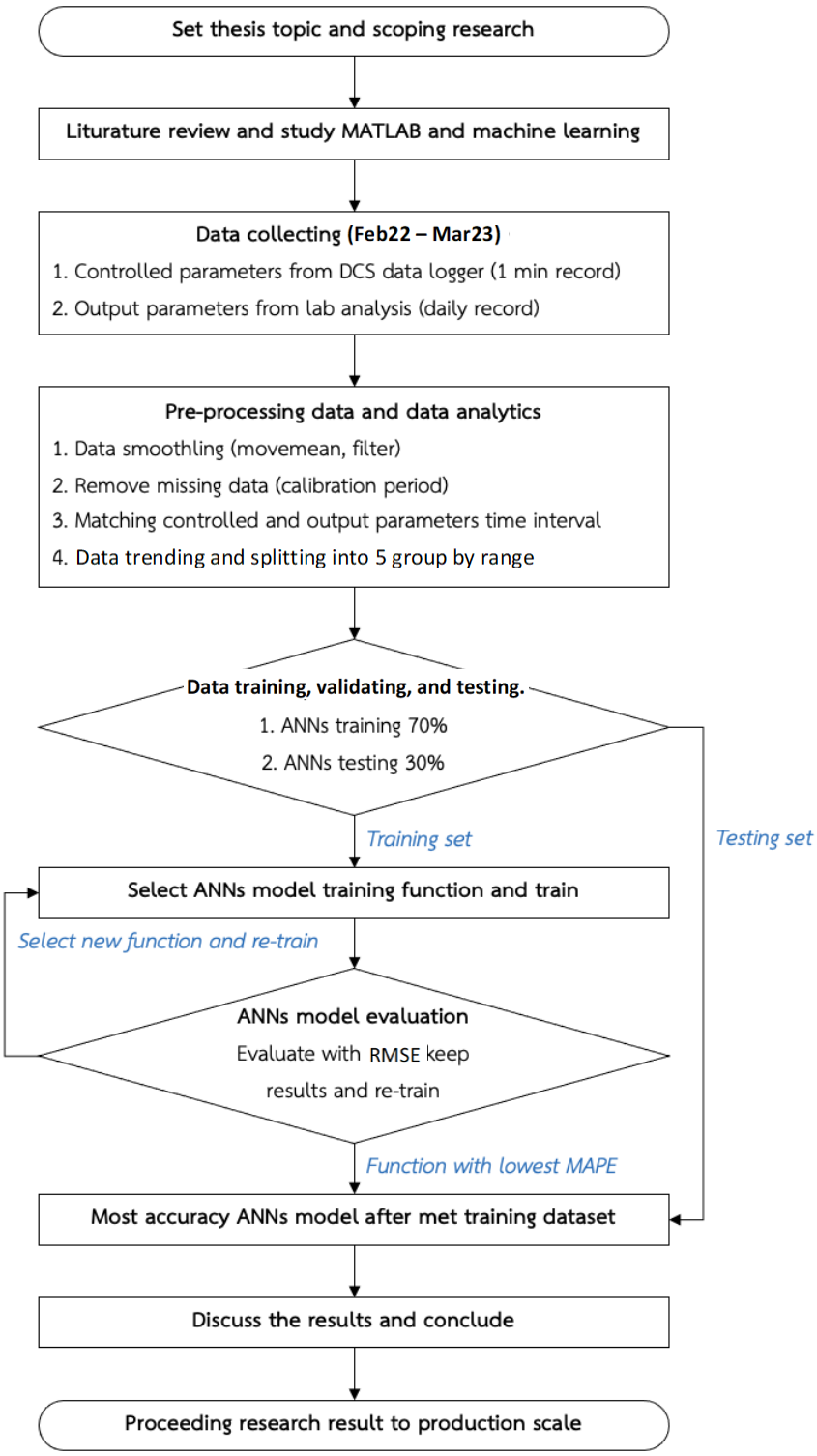


Figure 1 Thesis methodology flow chart

CHAPTER 2 : LITERATURE REVIEW

2.1 CHLOR-ALKALI PROCESS

Chlor-Alkali process (CA) is an industrial process used to produce Chlorine (Cl_2), Caustic Soda (NaOH), and other valued byproducts such as Sodium Hypochlorite (NaOCl) and Hydrochloric acid (HCl). All products from the CA process are essential raw materials in various industries. Cl_2 is crucial in disinfection and petrochemical processes such as PVC and PC polymerization. NaOH in commodity products and pH neutralization. The high demand for CA products in all industries leads to scale growth with a 2022 annual capacity of 78 million metric tons of NaOH worldwide [2].

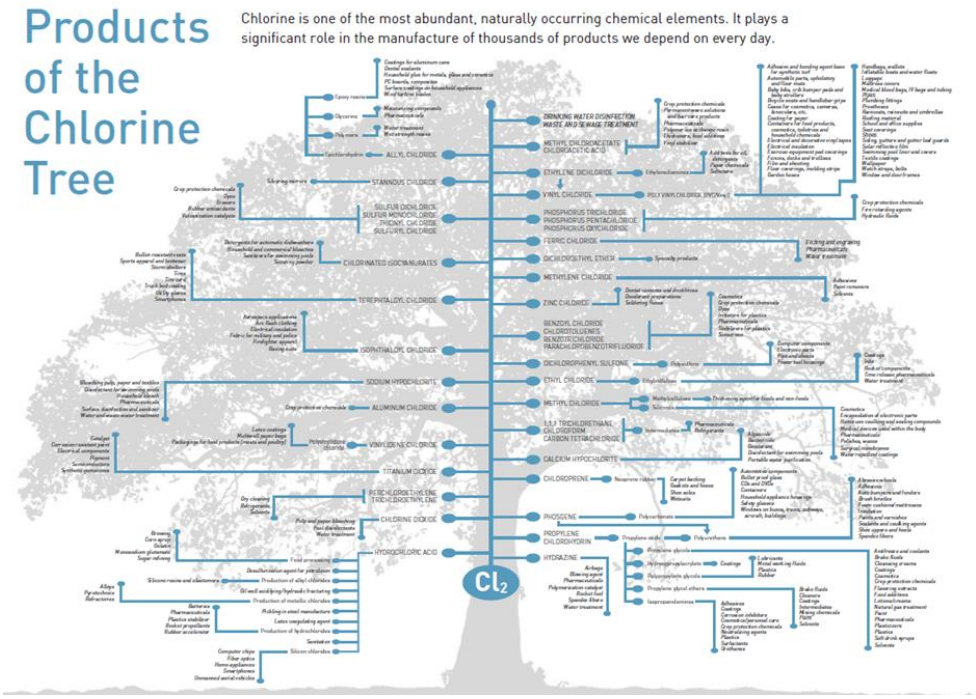


Figure 2 Products of Chlorine tree [3]

Since 1972 [1], It is the development of membrane electrolysis cell. The primary process of CA production is the electrolysis of brine (NaCl) aqueous solution, which is more environmentally compatible than mercury cell in the past. The electrolysis process has a high electrical energy consumption, and the latest technology in 2022 claim power consumption at 2,000 kWh per metric ton NaOH [4]. Thus, electrical power consumption is the direct operating cost of the CA process and a key factor of production profit.

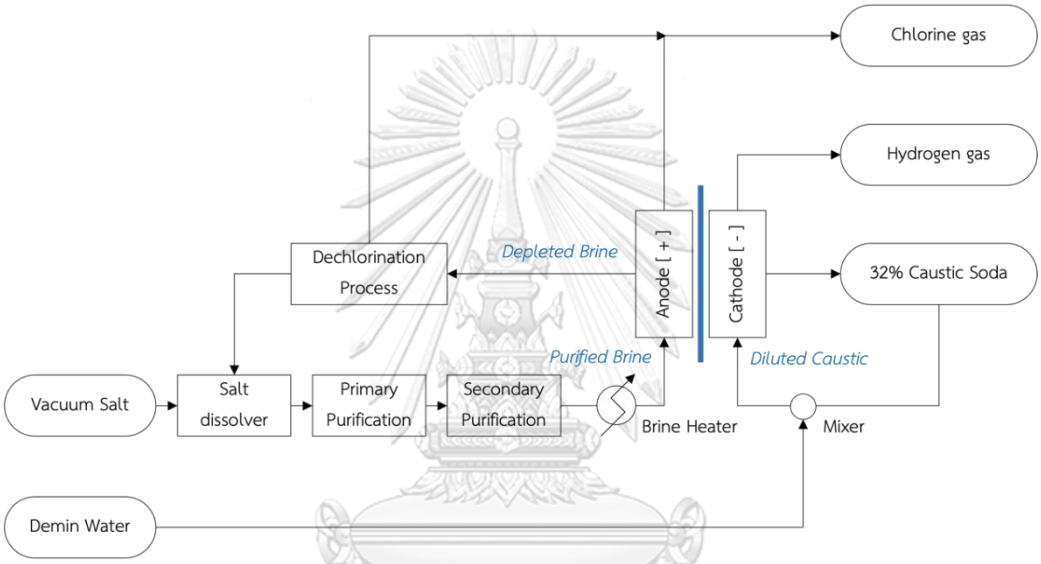


Figure 3 Chlor-Alkali process overview

CA brine membrane electrolysis process [5] starts with two raw materials – NaCl salt (use vacuum salt according to low Ca^{2+}, Mg^{2+}, I^- and Br^- in salt, which is less effect on membrane operation) as anode feeding and demineralized water as cathode feeding. NaCl vacuum salt as solid salt from the supplier must pass four raw material feeding treatments before feeding to the electrolysis cell.

1. Salt dissolver for dissolving solid salt into demineralized water to be a brine solution.
2. Primary purification removes suspended solid (SS) from brine by anthracite bed filter.
3. Secondary brine purification removes undesired ions, e.g., Ca^{2+}, Mg^{2+}, and Al^{3+}, from brine by chelating resin bed towers.

4. Brine heater for heating purified brine at suitable operating cell temperature by heat balancing with heat generation from electrolysis, raw salt, demineralized water, and recycled output brine and controlling cell temperature by temperature-indicated control (TIC).

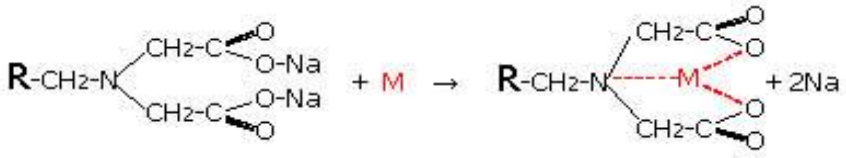


Figure 4 Chelating resin mechanism on alkali earth ion [6]

After purified brine reacted in the electrolysis cell, some NaCl solution remained from electrochemical conversion (depleted brine). Usually, input NaCl at a concentration of 300 g/L will remain in the output at 230 g/L (reaction consumed 70 g/L of purified brine). This remaining will return to the salt dissolver as recycling of raw material in the process. The remaining brine output has dissolved free-Chlorines in the form of Cl_2, $HClO^-$, ClO^- and ClO_3^- from electrolysis reaction, which affects the recycling process because free-Chlorine in recycled brine can damage chelating resin in the secondary purification. Thus, depleted brine must remove dissolved free-Chlorine before recycling to salt dissolver in dechlorination. The dechlorination process removes dissolved free-Chlorine with reduced depleted brine pH by HCl to pH below 2 for transforming all free-Chlorine species in solution into Cl_2. Then, remove the dissolved Cl_2 in the brine solution with a vacuum hydro cyclone. Cl_2 gas in solution separates in cyclone to the gas phase and sent to combine with Cl_2 gas product. Liquid phased from cyclone remains only depleted brine to recycle process.

Reaction undergoes in the anode and cathode side separately. First, Fed purified brine NaCl into the anode side (called anolyte). An aqueous solution of NaCl separated into Na^+ and Cl^- ion in demineralized water. Cl^- ion from electricity become Cl_2 gas and floats out separate from the liquid. Lonely ion Na^+ in the anode side permeates through an ion-exchange membrane to the cathode side and meets OH^- from the catholyte reaction resulting in the NaOH product.

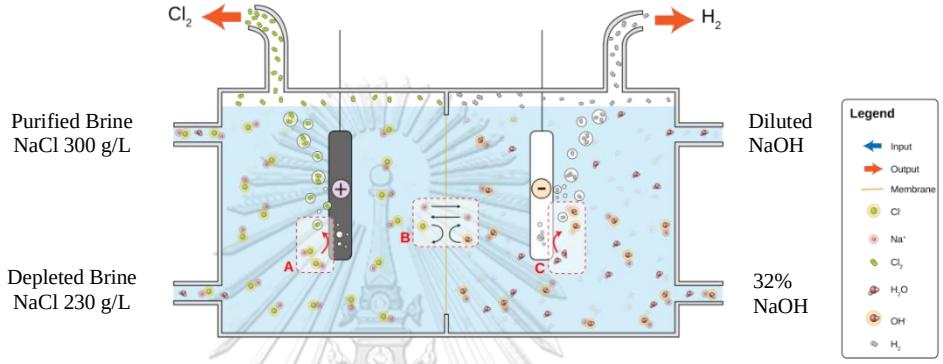


Figure 5 Chlor-Alkali electrolysis cell diagram [7]

Anode side reaction

$NaCl \longrightarrow Na^+ + Cl^-$ -------- equation 1

$2Cl^- \longrightarrow Cl_2 + 2e^-$ (E_0 1.36 V) -------- equation 2

Cathode side

$H_2O \longrightarrow H^+ + OH^-$ -------- equation 3

$2H^+ + 2e^- \longrightarrow H_2$ (E_0 -0.83 V) -------- equation 4

$Na^+ + OH^- \longrightarrow NaOH$ -------- equation 5

Total reaction

$2NaCl + 2H_2O \longrightarrow Cl_2 + H_2 + 2NaOH$ -------- equation 6

2.2 MEMBRANE THEORY

The ion-exchange membrane's main component is fluoropolymer due to excellent chemical compatibility with base, acid, and chlorinated conditions [8]. The membrane consists of 2 main layers. First, the sulfonate layer on the anode side is a highly conductive and sacrificial fiber for reinforcement and conducts electrical conduction through the membrane. Second, the carboxylate layer on the cathode side is highly ion-selective for selecting Na^+ to permeate ions. The surface of the membrane is applying an anti-gas stagnation coating.

The critical performance of polymer-based ion-exchange membranes is membrane duration and ion-selection channel size, which control impurity, permeated rate, uncontrolled ion selection, and reverse diffusion. The wrong channel size results in high power consumption from high resistance, permeation, and product contamination. It occurs from reverse diffusion of product on the cathode side to the anode or leak of minus ion from the anode to the cathode side.

1. Current density (CD) - current input for reaction divided by effective membrane area - is the driving force of ion permeation and leads reaction by transferring an electron to the reactant.
2. Membrane service life (determined by the day of life, DOL) related to the accumulation of impurities in raw material, such as solid particles or other ion species in brine, can plug or penetrate inside the membrane channel leading to low membrane efficiency.
3. Water content is related to brine and caustic concentration, and the flow rate of raw material consists of feed brine flow rate (Q_{FB}) and feed caustic flow rate (Q_{HD}) on the membrane surface. Water content led to polymer water absorption and affected channel size.
4. Cell temperature (T) directly affects the channel size of a membrane according to polymer properties. The membrane sheet and the membrane's channel will expand at high temperatures and shrink at low temperatures.

Reaction developed Chlorine to free-Chlorine.

Chlorite generation from leaked OH^-

$$Cl_2 + OH^- \longrightarrow HOCl^- + Cl^- \quad \text{-------- equation 7}$$

Evolution of free-Chlorine (Hypochlorite and Chlorate)

$$HClO^- \longrightarrow ClO^- + H^+ \quad \text{-------- equation 8}$$

$$2HClO^- + ClO^- \longrightarrow ClO_3^- + 2H^+ + 2Cl^- \quad \text{-------- equation 9}$$

O_2 evolution from decomposed Hypochlorite

$$2ClO^- \longrightarrow O_2 + 2Cl^- \quad \text{-------- equation 10}$$

O_2 evolution from oxidation of OH^-

$$OH^- \longrightarrow O_2 + 2H_2O + 4e^- \ (E_0\ 0.40\ V) \quad \text{-------- equation 11}$$

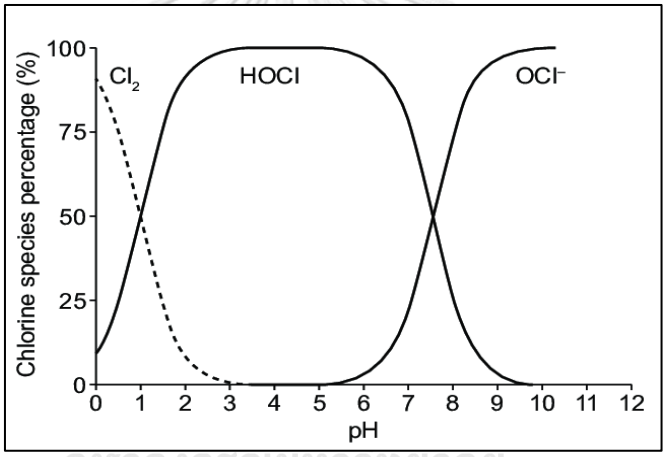


Figure 6 Development of free-Chlorine in various pH [9]

2.3 CELL VOLTAGE (CV)

Cell voltage is directly related to power consumption (PC) by an equation of electrical power (P = IV in direct current) which from the rectifier supplies stable current (I) as the production rate required to produce desired products. The voltage or cell voltage appears consequently from Ohm's law (V = IR while R is assumed constant). Thus, the relation between CV and CD could be a linear correlation, while resistance (R) is a slope of the function. There are several reasons to control cell resistance in commercial-scale CA plants, for example.

1. Catalytic reaction rate from different electrode coating.
2. Membrane channel size by cell condition (explained in membrane theory) and membrane material or technology.
3. Electrolysis cell design of fluid flow or recycling inside the cell.
4. Resistance of cell material.
5. Brine and caustic raw material quality.

CA electrolysis reaction rate is calculated by 'Faraday's law of electrolysis' with the importance number 96,450 coulomb per 1 electron mole. It means that to produce 1 mole of Cl_2 as reaction mass balance, NaCl 2 mole will be consumed and need electricity 96,450x2 = 192,900 coulomb (1 coulomb = 1-ampere sec) or 53.58 ampere in 1 hour.

Current density (CD) means electrical current intake to electrolysis reaction ratio to 1 m^2 of membrane effective area. Usually, the membrane cell was limited membrane effective area by technology provider design which is optimum hydrodynamic circulation inside electrolysis cell maintenance and manufacturer. In recent technology, the membrane can operate at the highest current density, up to 8 KA/m^2.

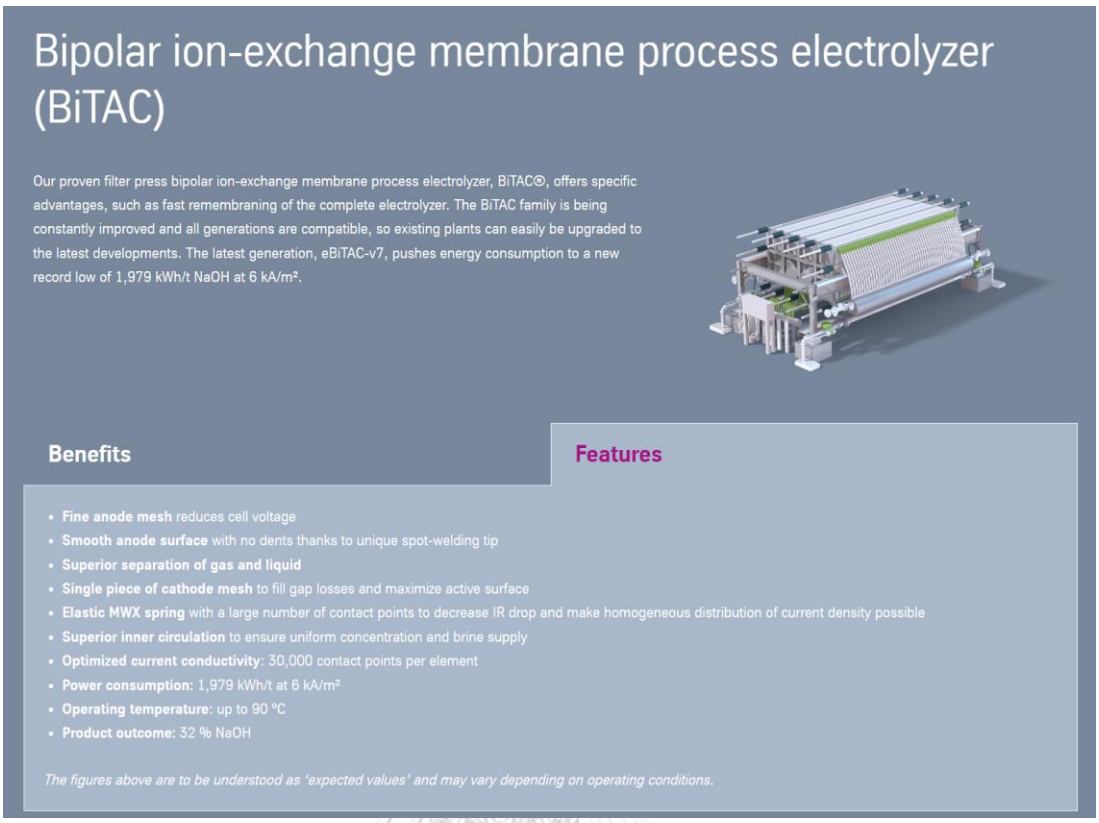


Figure 7 Internal circulation improvement design to improve capacity [4]

One electrolysis cell consists of many membrane cells connected in a series of electrical circuits to reduce loss from electrical resistance (busbar loss). The maximum series number is related to DC supply (in commercial use, thyristor rectifier to convert AC to DC). According to a higher cell number with a higher total voltage, I/A should match the electricity supplied on the AC side. In conclusion, nowadays, traditional industrial Chlor-alkali electrolysis cell operation. Bipolar design circuit in cell series to reduce busbar loss. The maximum CD limit was an efficient design from a technology provider. Maximum capacity per electrolysis cell (no. of membrane cells in a circuit) limited by I/A ratio. High production capacity was from the high number of parallel electrolysis cell operations.

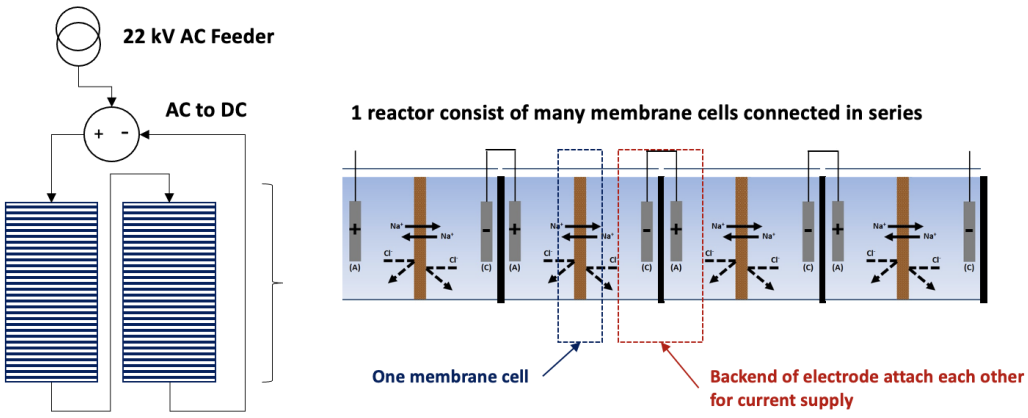


Figure 8 Chlor-Alkali electrolysis cell configuration in commercial plant [5]

According to Faraday's law of electrolysis and electrolysis cells configuration, the production mass balance of Cl_2 and NaOH was calculated by the following equation when

1. DC current (I) Current input to the cell (related to 96,450 Faraday's number)
2. IEM no. the number of the membrane in an electrolysis cell.
3. 0.945 basis efficiency of current to each product.

Chlorine production

$$\frac{\text{DC Current [kAs]} \times 1000 \left[\frac{\text{A}}{\text{kA}}\right]}{96450 \left[\frac{\text{coulomb}}{\text{mole e}^-}\right]} \times 0.5 \left[\frac{\text{mol e}^-}{\text{mol Cl}_2}\right] \times 71 \left[\frac{\text{g}}{\text{mol Cl}_2}\right] \times \text{IEM} \times 0.945 \text{ [std eff.]} = \textbf{Chlorine Production [g/s]}$$

-------- equation 12

Caustic soda production

$$\frac{\text{DC Current [kAs]} \times 1000 \left[\frac{\text{A}}{\text{kA}}\right]}{96450 \left[\frac{\text{coulomb}}{\text{mole e}^-}\right]} \times 1 \left[\frac{\text{mol e}^-}{\text{mol NaOH}_s}\right] \times 40 \left[\frac{\text{g}}{\text{mol NaOH}_s}\right] \times \text{IEM} \times 0.945 \text{ [std eff.]} = \textbf{Caustic soda Production [g/s]}$$

-------- equation 13

2.4 MACHINE LEARNING (ML)

Machine learning is a modern way to find the complex correlation between input and output data which does not have to be deeply in the numerical model [7]. The principle of almost ML models is to cluster input data to several groups as output class by trialing the most precise mathematics model.

There are three types of machine learning model approaches based on the learning method: supervised, unsupervised, and reinforcement learning. Supervised learning requires labeled input and output data for the training phase. When the model has identified the relationship between the input and output data, it can use to classify new and unseen datasets and predict outcomes.

2.5 ARTIFICIAL NEURAL NETWORK (ANNs)

Neural networks were beginning to be the modeling of complex manufacturing processes [1]. These layers have a certain number of input nodes, hidden nodes, and output nodes. A simple ANNs workflow in *Figure 9* shows a network with one node in the input layer, two hidden nodes, and one node in the output layer. ANNs will learn to find a correlation between input and output data through manipulation in the hidden layer.

As a greater number of inputs, such as parameters in the chemical plant, make ANNs more complex and results in more complex hidden layer correlation, the Primary hidden node is not enough to predict precise correlation and need to be calculated in the more hidden layer to find a secondary, tertiary relative. However, too much-hidden layer correlation can make ANN's model overfitted, and too small might not find any correlation in the network.

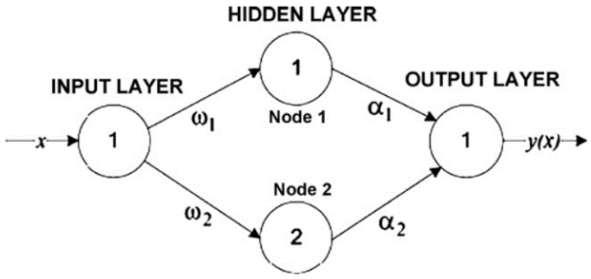


Figure 9 Schematics of simple ANN [10]

2.6 MODEL CONFIRMATION METHOD

2.6.1 Data splitting

To evaluation of ANN, trained model data much be split into 2 or 3 groups. In the case of 2 groups, the first is the training group for performing backpropagation training and creating the network regression model. Second, is the test group for testing the accuracy of the generated network model with the unseen dataset. The problem of the two groups is when the model is optimized by repeat generating model in the training dataset and accuracy checking in the testing dataset by random train-test selection. The testing dataset is not totally unseen, then become the model bias.

For this reason, the general data splitting is usually divided into 3 groups. By adding a validation dataset to use as model testing in repeating optimization models, before checked with the testing dataset, which is unseen.

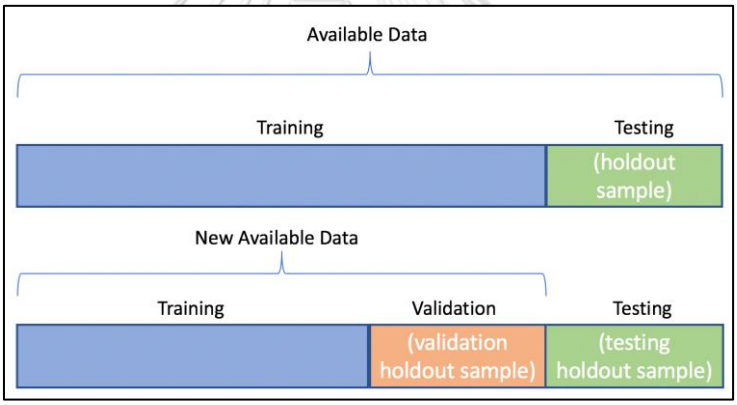


Figure 10 Data splitting; train, validate and test [11]

2.6.2 Train-to-test ratio

Comparing results of varying train-to-test ratios is a way to evaluate the performance of any dataset distribution. The different train-to-test ratios can give different model performance results according to randomized training, validating, and testing dataset. The commonly train-to-test ratio is 70:30, 80:20, 90:10, or even 50:50. It depends on regression practice or the characteristics of each dataset. Several train-to-test ratios should be compared to confirm the model performance of the new dataset in any distributed data and control the confidence interval of the training set to generate a network.

2.6.3 Cross validation

Cross-validation or k-fold cross-validation (k-fold cv) is one method to validate model performance by finding the average error between each iteration of the dataset (usually 5 or 10 iterations). In each iteration, groups of train and test datasets will be varied following the number of iterations to compare data training bias in each dataset and give a model more confidence.

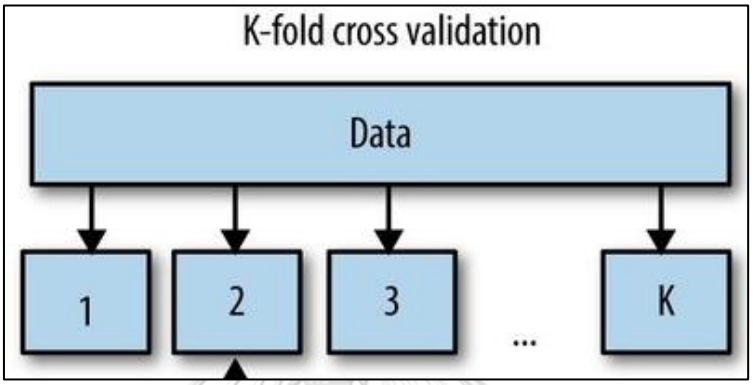


Figure 11 Schematic of k-fold cross validation [12]

2.6.4 Hold-out validation

Hold-out validation is another method to validate model performance by holding validation data for testing the model to be unseen data. This method seems like a blind test of a model to use the generated model to predict output in a different dataset. It might be a disadvantage in the small dataset, which can occur overfitting data due to training and validation datasets having different data distributions.

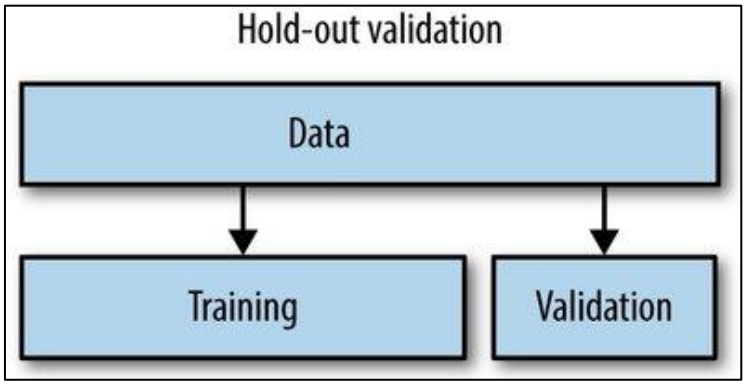


Figure 12 Schematic of hold-out validation [12]

2.6.5 Under-fitting, over-fitting, and best-fitting

Typically, regression or model-fitting has three categories: under-fitting, best-fitting, and over-fitting. The under-fitting model cannot regress the training dataset to meet accuracy criteria, and the over-fitting regression model works well in the training dataset. However, in the testing dataset, results become worse from too much fitting or noise generation. Meanwhile, the best-fitting is the optimum one that gives the highest accuracy in any dataset.

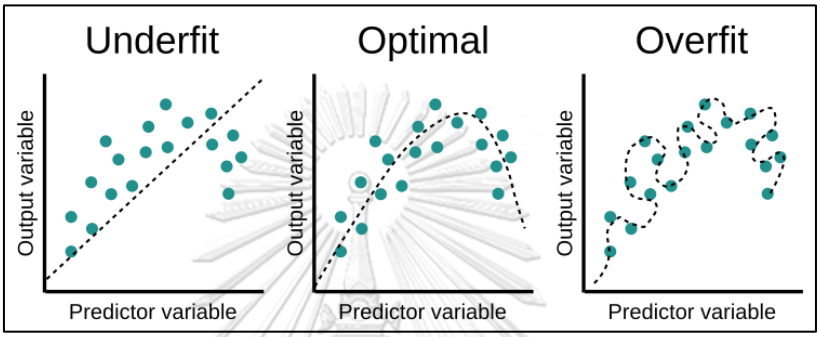


Figure 13 Visualization of example underfit, optimal (best fit) and overfit [13]

2.6.6 Early stopper

Early stopper [14] is method to find optimum value of prediction error by early stopping model fitting in test group before runner to terminal epoch as train group. That is selecting only converge model which gave most fitted value before model fitting become overfitted by diverge error.

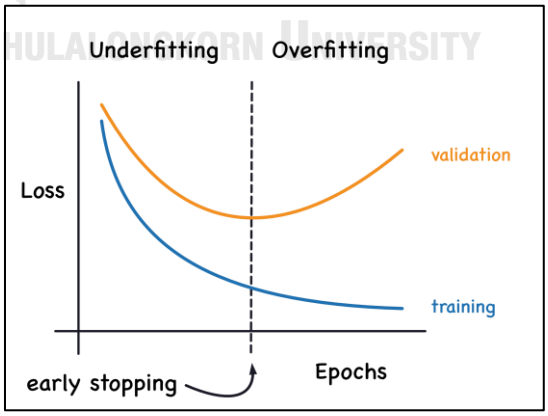


Figure 14 Example plot of early stopper [14]

2.7 PERFORMANCE EVALUATION METHOD

There are several ways to measure a predictive model's performance to evaluate the prediction's accuracy [15]. The most practical ways are measuring fitting errors by different values between the actual and predicted data.

2.7.1 Mean Absolute Error (MAE)

MAE is the most straightforward error measurement, calculated average magnitude of all regressed and actual data, the perfect fitted shown MAE at 0. However, MAE can easily deviate from outliers to make significant one error dominate overall performance, and the absolute make this evaluation method cannot indicate between under and overperformance.

$$MAE = \frac{1}{n}\sum |y-y_i| \qquad \text{-------- equation 14}$$

While; y is actual data, y_i is regressed data at the same input, n is number of data

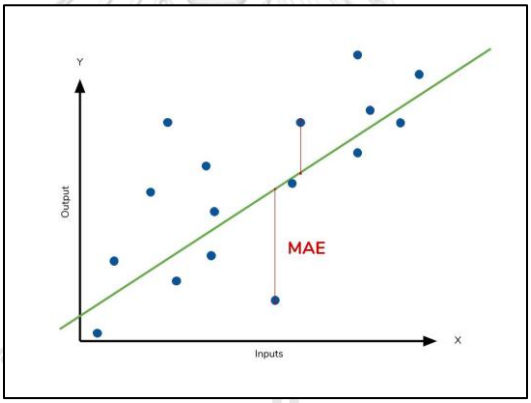


Figure 15 Mean Absolute Error (MAE) [15]

2.7.2 Mean Absolute Percentage Error (MAPE)

MAPE is the percentage form of MAE. The error calculation is the same as MAE, only converted to percentage form. MAPE also has a high effect from outliers, the same as MAE. The percentage is easier to understand and generalized results. However, MAPE has a limitation at an undefined value when data is at 0 (become 0/0), and the percentage value is less symetric error compared to the original.

$$MAE = \frac{100\%}{n}\sum \left|\frac{y-y_i}{y}\right| \qquad \text{-------- equation 15}$$

While; y is actual data, y_i is regressed data at the same input, n is number of data

2.7.3 Mean Square Error (MSE)

MSE is like MAE but uses a square to indicate magnitude substitute absolute. The MSE gives much more error due to the square value, and for this reason, MSE becomes non-linear (exponential growth of error) and gives a penalty to a value with a large error more than a small error. Similarly, MSE has an effect from outliers more than MAE also.

$$MAE = \frac{1}{n}\sum (y-y_i)^2 \qquad \text{-------- equation 16}$$

While; y is actual data, y_i is regressed data at the same input, n is number of data

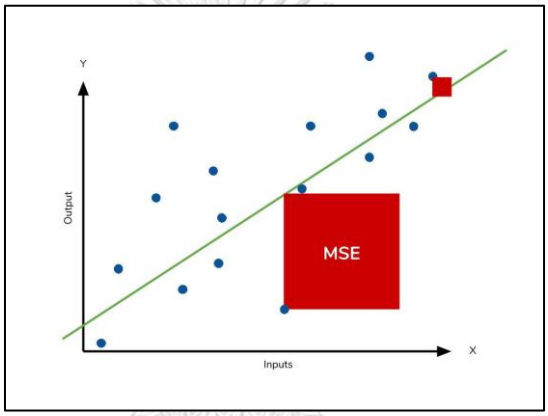


Figure 16 Mean Square Error (MSE) [15]

2.7.4 Root Mean Square Error (RMSE)

RMSE is a square root form of MSE that converts square error to the original scale. RMSE can measure error as an original unit of output and is easier to understand.

$$MAE = \frac{1}{n}\sum \sqrt{(y-y_i)^2} \qquad \text{-------- equation 17}$$

While; y is actual data, y_i is regressed data at the same input, n is number of data

2.7.5 Mean Percentage Error (MPE)

MPE is like MAPE but without magnitude indication by absolute. MPE is in percentage form, which is easier to understand, but also has limitation with undefined value when the data is at 0. MPE can show positive and negative error effects because it does not have absolute or square to eliminate under and overperformance. However, the effect of positive and negative errors can be deducted from each other when the model faces a uniform dataset. MPE can show the trend but cannot verify the regression accuracy.

$$MAE = \frac{100\%}{n} \sum \left(\frac{y - y_i}{y} \right) \quad \text{-------- equation 18}$$

While; y is actual data, y_i is regressed data at the same input, n is number of data

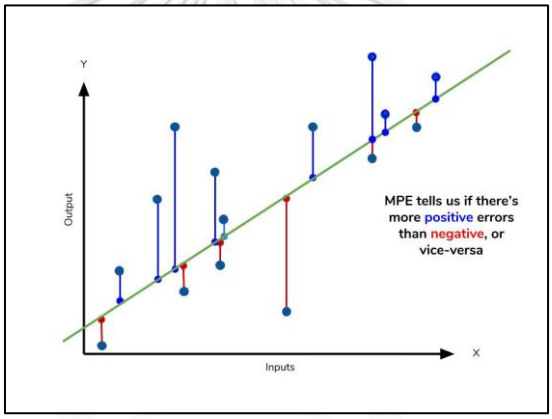


Figure 17 Mean Percentage Error (MPE) [15]

Table 1 Summary of error measurement methods [15]

No.	Method	Eliminate negative	Outlier effect	Easy to indicate	Original scale
1	MAE	Yes	Low	No	Yes
2	MAPE	Yes	Low	Yes	No
3	MSE	Yes	Much	No	No
4	RMSE	Yes	Much	No	Yes
5	MPE	No	Low	Yes	No

2.8 LITERATURE REVIEW

In 2006, A.A. Jalali et al. [16] studied the effect of process conditions on the cell voltage of a Chlor-Alkali membrane cell. Five process parameters were studied: anolyte pH, cell temperature, electrolyte velocity, brine concentration, and current density on laboratory-scale Chlor-Alkali membrane cell. The analysis goes with a statistical method, which is an analysis of variance (ANOVA) to evaluate the effectiveness of operating parameters. The result found that current density and cell temperature were the most striking parameters of cell voltage. Current density contributes 69.94% effect on cell voltage in a sensitivity analysis.

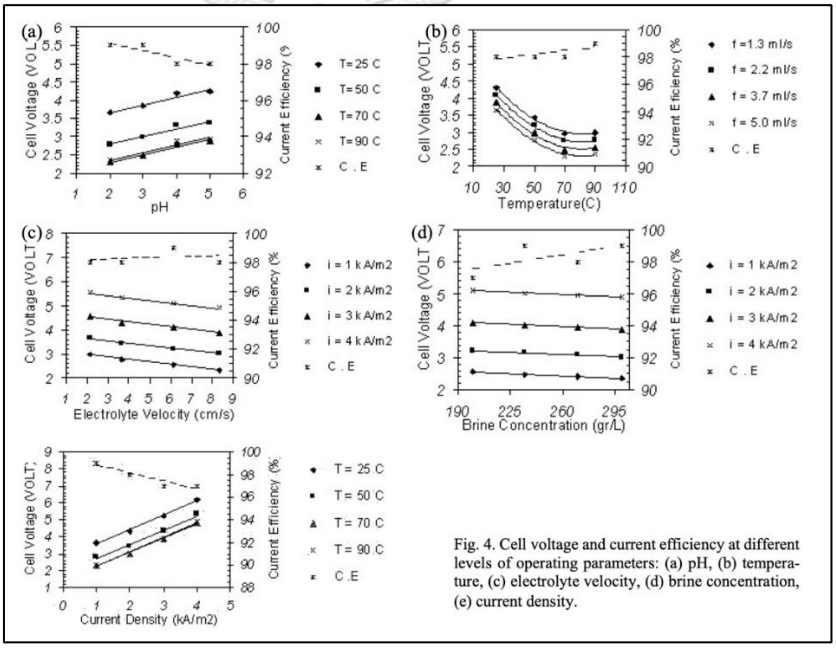


Figure 18 Plot of cell voltage with different operating parameters

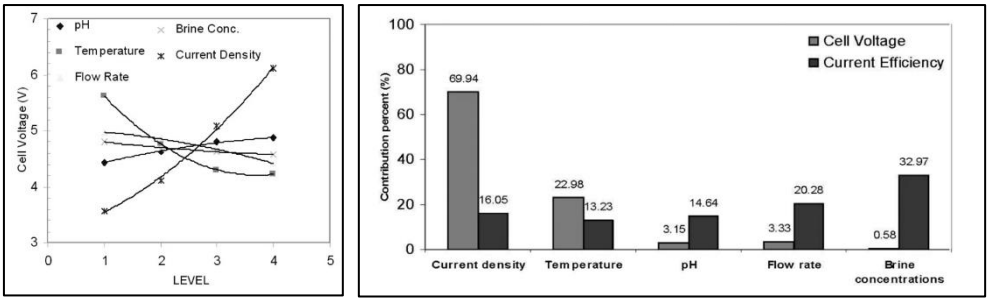


Figure 19 Sensitivity plot of each parameters impact level to CV,

And contribution percent chart of each operating parameters

In 2008, T. Mirzazadeh et al. [17] studied the effect of various parameters on caustic current efficiency (CCE) on zero-gap oxygen-depolarized cathode (ODC) Chlor-Alkali cell with artificial neural network (ANN) Lavenberg-Marquardt (LM) back propagation method. This research studies seven operating parameters: anolyte pH, temperature, brine flow rate, brine concentration, oxygen temperature, and oxygen flow rate. The study aims to find the optimum of each parameter condition with the best CCE. Results show that ANN with the LM method has better CCE prediction accuracy than actual data. The ANN-generated data can assist in trending and finding correlations to predict CCE.

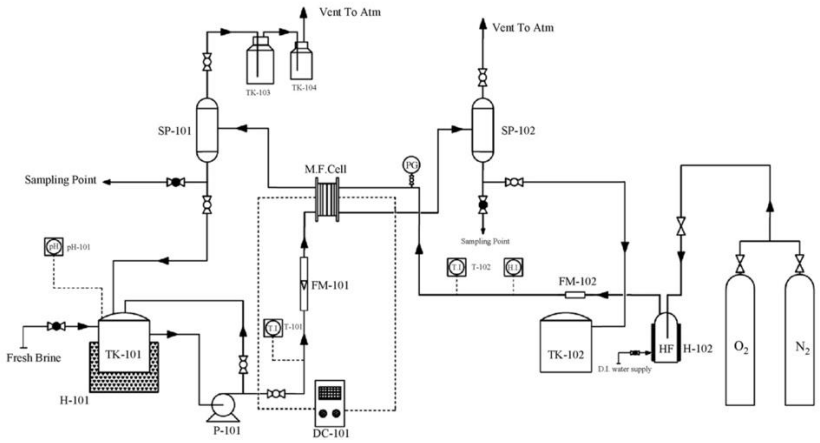


Figure 20 Schematic of set-up laboratory-scale Chlor-Alkali ODC membrane cell

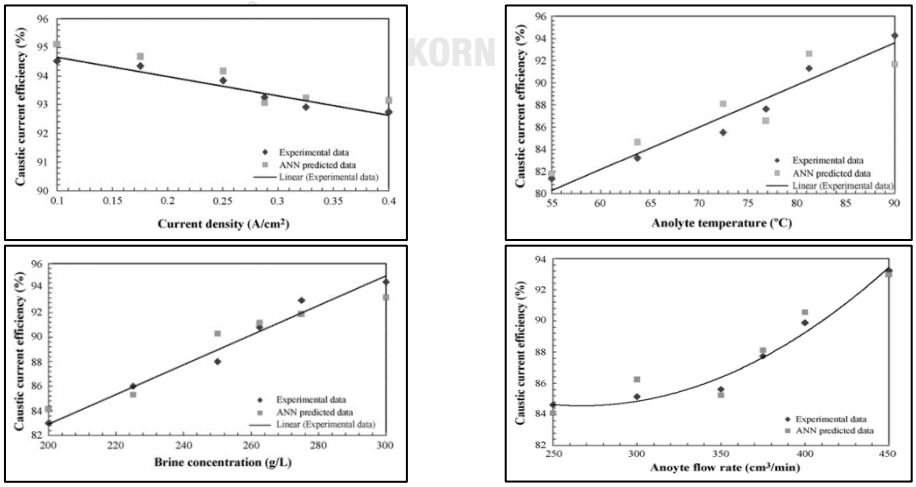


Figure 21 Analysis regression of each parameters to CCE

In 2008, N. Shojai Kaveh et al. [10] studied the development of a backpropagation (BP) algorithm, an artificial neural network (ANN) model, on various operating parameters to validate model prediction and study the effect of each parameter. The experiment undergoes six operating parameters: anolyte pH, temperature, brine concentration, current density, and run time in set-up lab-scale Chlor-Alkali electrolysis. The number of hidden nodes and hidden layers was selected at the optimum point by trial and error, which was 6-7-5-1 (6 input parameters, seven hidden nodes in the first hidden layer, five hidden nodes in the second layer, and one output parameter). The result of the performance model in this study showed an RMSE of 0.036 in the training group and 0.043 in the test group. Sensitivity was analyzed and plotted into the level graph to find the impact of each parameter. It found that current density impacts cell voltage most, with 59.26% on ANN prediction and 58.19% on experimental data.

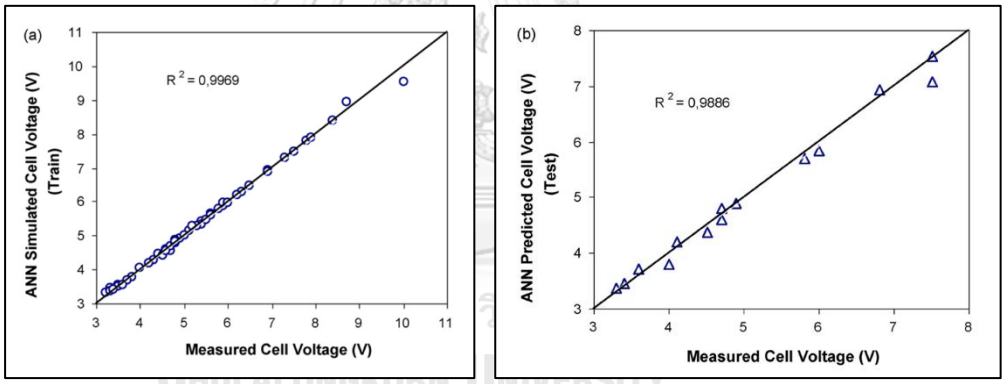


Figure 22 Regression plot of ANN result with actual cell voltage

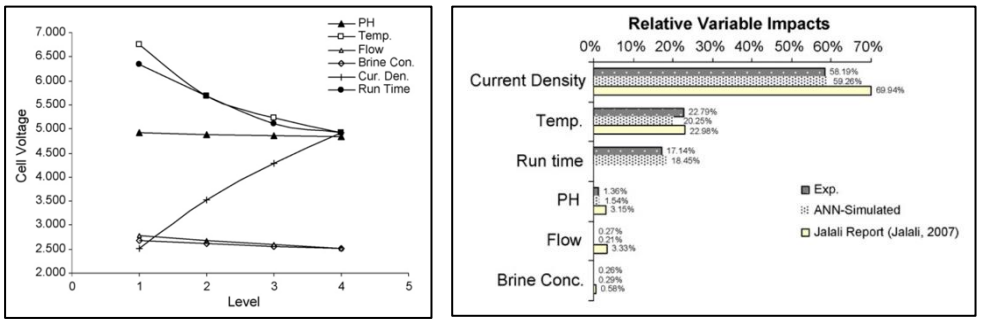


Figure 23 Sensitivity analysis of each parameters impact level to CV

In 2009, N. Shojai Kaveh et al. [18] studied to investigate the impact of operating parameters by the support vector machine (SVM) technique – a machine learning method -. The study tested with a set-up Chlor-Alkali electrolysis cell in the laboratory and varied operating conditions to obtain data. The research performs six operating parameters: anolyte pH, temperature, brine concentration, current density and run time, and one output as cell voltage. The result shows SVM model can be used to predict the impact of parameters and not only approximately, but SVM also can predict cell voltage with RMSE 0.161. From the sensitivity study, current density gave the highest impact among all operating parameters, with 54.28% on SVM simulation and 55.42% on experimental data.

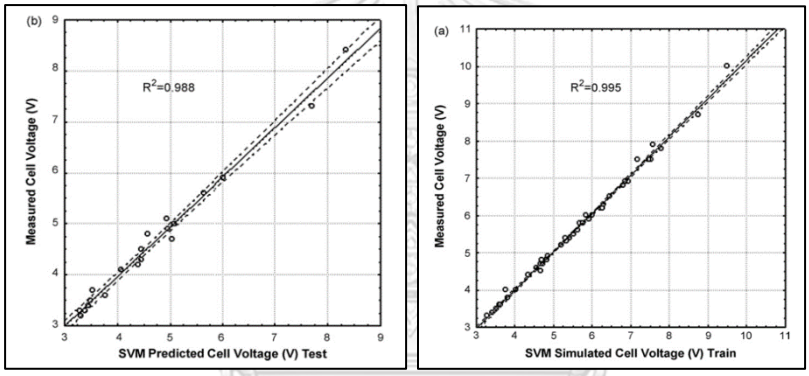
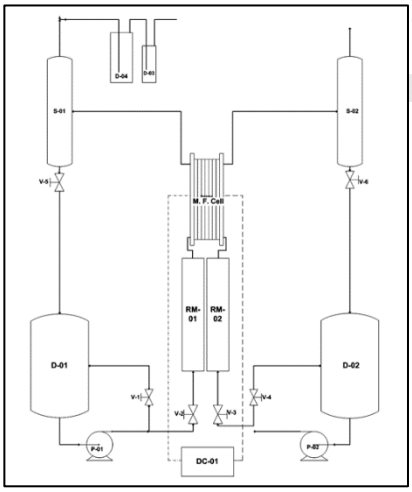


Figure 24 Regression plot of SVM result with actual cell voltage



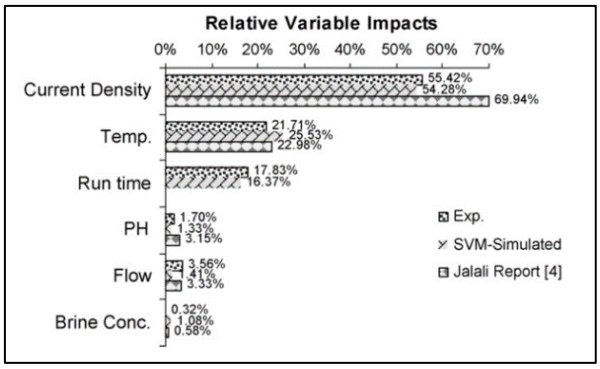


Figure 25 (Left) Schematic of set-up laboratory-scale Chlor-Alkali membrane cell

Figure 26 (Right) Sensitivity plot of each parameters impact level to CV

From the works literature review, several techniques were used for predicting cell voltage, giving satisfactory results. Chlor-Alkali voltage and current efficiency prediction are in focus because the process consumes much power in production; multi-operating parameters come with a highly complex relationship, and ordinary regression is challenging to show highly accurate prediction. Implementing machine learning is the new trend of correlation finding in recent years. From research machine learning, most of the research studies in set-up laboratory-scale Chlor-Alkali membrane cell, which can control environmental effects and vary operating parameters for testing. In commercial-scale size plant studies, there is more limit on testing validation but giving much more data. It is potential to study ANN regression with operating parameters to cell voltage in a commercial-scale Chlor-Alkali plant.

CHAPTER 3 : METHODOLOGY

3.1 DATASET AND PREPROCESSING

3.1.1 Data source

The dataset for analysis in this thesis was from the Yokogawa plant information management system (exaquantum). Parameters in the process were measured by an instrument transmitter and sent to a distributed computer system (DCS). DCS receives the signal of the transmitter and transforms the signal into measurable parameter data for reading, visualizing, and controlling. After the process, data are in a usable format and sent this information to a server (historian database). Plant information management systems (PIMS) have become an integral part of the tools to connect DCS historian databases to users in local networks or web servers (exaquantum is a trademark of Yokogawa's PIMS). Exaquantum acquires process information to users for data analytics, visualization, and decision-making.

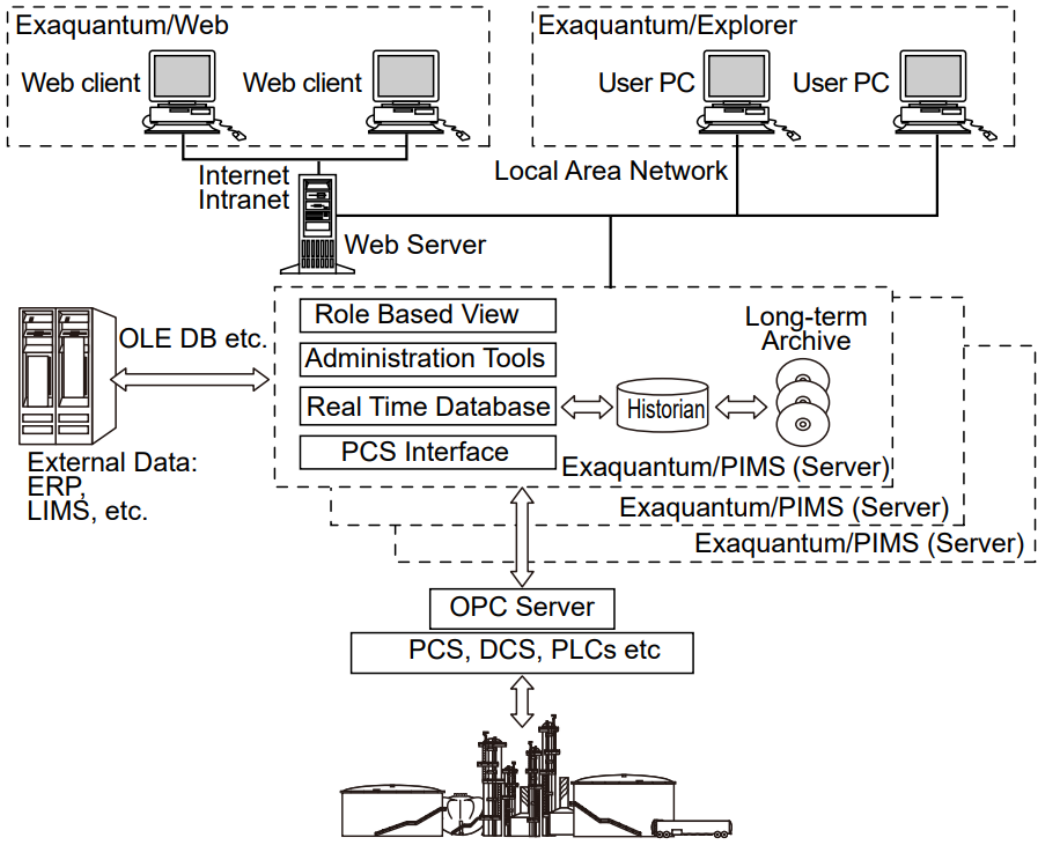


Figure 27 Yokogawa plant information management system (exaquantum) [19]

3.1.2 Dataset

This thesis collected operational raw data in 1 year period from March 2022 to February 2023 in hourly intervals (a total of 8,761 data). There are six features in the dataset shown in the table's column.

Input parameters (can be called predictors in the ANN model) consisting of

1. Current density (CD, V) is electrical current through a square-meter membrane. It was used for standardized operation capacity (refer to Faraday equation 12, 13 in terms of current, I), directly related to the permeating rate of substance through the membrane.
2. Operation day (DOL, day) refers to increasing membrane resistance due to clogging accumulation on the membrane surface and blocking the permeation of substances.
3. Feed brine flow rate (Q_{FB}, m^3/h) refers to water content on the membrane surface, leading to membrane channel expansion or shrinking.
4. Feed caustic flow rate (Q_{HD}, m^3/h) refers to water containing the same as Q_{FB} but on the cathode side.
5. Cell temperature (T, °C) is the temperature that occurs during operation. It can be from heat received from feed brine, feed caustic temperature, or heat generated from the electrolysis cell due to electrical resistance. Cell temperature

Output parameters (can be called response in ANN model) is

1. Cell voltage (V) in unit V shows the voltage across a membrane from the current applied and the resistance effect of operating parameters. Cell voltage led to power consumption and approximate membrane efficiency.

Table 2 List of parameters, types, ranges and units

Type	Abbr.	Description	Range	Unit
Output data	CV	Cell voltage	2.0-3.4	V
Input data	I	Current	6.0-19.7	KA
Input data 1	CD	Current density	1.8-6.0	KA/m^2
Input data 2	DOL	Operation day	0-365	Day
Input data 3	Q_{FB}	Feed brine flow rate	12.0-30.0	m^3/h
Input data 4	Q_{HD}	Feed brine flow rate	31 ± 1	m^3/h
Input data 5	T	Cell temperature	85-95	°C
Evaluation	RMSE	Root means square error	-	V

	A	B	C	D	E	F	G	H	I
1	DOL	I	CD	V	CV	FI3312A	FI3312B	TI3302A	TI3302B
2	1.000	19.686	6.011	311.829	3.057	28.951	31.324	88.364	86.681
3	1.042	19.685	6.011	311.797	3.057	29.264	31.376	88.354	86.677
4	1.083	19.683	6.010	311.793	3.057	29.071	31.471	88.328	86.650
5	1.125	19.686	6.011	311.795	3.057	28.901	31.488	88.278	86.537
6	1.167	19.684	6.010	311.796	3.057	29.470	31.255	88.282	86.590
7	1.208	19.683	6.010	311.774	3.057	28.899	31.488	88.299	86.621
8	1.250	19.680	6.009	311.778	3.057	29.830	31.422	88.306	86.553
9	1.292	19.684	6.010	311.801	3.057	29.386	31.237	88.346	86.650
10	1.375	19.688	6.012	311.806	3.057	28.743	31.218	88.393	86.707
11	1.417	19.685	6.011	312.384	3.063	29.800	31.945	88.373	86.709
12	1.458	19.687	6.011	311.707	3.056	29.270	31.809	88.357	86.724
13	1.500	19.690	6.012	311.803	3.057	29.961	31.762	88.421	86.814
14	1.542	19.689	6.012	311.876	3.058	29.879	31.659	88.439	86.897
15	1.583	19.690	6.012	311.886	3.058	29.725	31.858	88.431	86.798
16	1.625	19.692	6.013	311.857	3.057	29.813	31.651	88.396	86.856
17	1.667	19.693	6.013	311.848	3.057	29.557	31.742	88.361	86.770
18	1.708	19.689	6.012	311.827	3.057	29.290	31.712	88.294	86.696
19	1.750	19.689	6.012	311.838	3.057	29.925	31.818	88.271	86.620
20	1.792	19.684	6.010	311.851	3.057	29.241	31.740	88.275	86.513
21	1.833	19.683	6.010	311.858	3.057	29.681	31.765	88.288	86.589
22	1.875	19.684	6.010	311.843	3.057	29.142	31.715	88.255	86.638
23	1.958	19.682	6.010	311.835	3.057	29.298	31.897	88.294	86.612
24	2.000	19.685	6.011	311.848	3.057	29.059	31.589	88.228	86.624
25	2.042	19.686	6.011	311.893	3.058	29.117	31.553	88.225	86.559
26	2.083	19.691	6.013	311.898	3.058	29.427	31.693	88.252	86.585

Figure 28 Six features of 8,760 raw data in Mar 22 – Feb 23 in CSV file

3.1.3 Data preprocessing

According to data collected from “exaquantum” data historians, which record hourly raw data include non-operation, calibration, and unsteady-state periods or raw data that come with signal loss and noise. Pre-processing performs to cut outlier data, clean data, and transform all parameters to the same time interval. To visualize a group of data, linear regression between CV and CD will be plotted. Referring to Ohm’s law, CV and CD direct theoretical correlation will be shown in linear. The deviate RSQ > 0.9.

Ohm’s law

$$V = IR \quad \text{-------- equation 19}$$

$$CV = CD*R \quad \text{-------- equation 20}$$

While CV means voltage across 1 electrolysis cell.

CD means current through a unit area of the membrane.

In ANN model training, performance results in test group correlation might be biased by accidentally similar data in the train, validate, and test groups from random data selection. Reduce data training bias cross-checking must be performed to validate the test and train the group to confirm the performance of the ANN model.

3.1.4 Method to remove outlier data.

1. Cutting non-operation period by data filter method.
    a. Removing the dataset with current (I) < 6 KA. In typical operation practice, the minimum load was set at 6 KA because it is the lowest current supply for electrolysis reactions that can make on-spec products.
    b. Removing dataset in which cell voltage (CV) < 2 V. Referring to reaction equation 6, cell voltage less than 2 V is insufficient for a brine electrolysis reaction.

2. Cutting unsteady-state period by the data filter method.
    a. Removing dataset at 09:00 AM due to the load decreasing period for on-peak operation. During load decreasing input parameters might be performed lead or lag to output and effect to correlation fitting.
    b. Removing the dataset at 10:00 PM due to the load-increasing period for off-peak operation. During load decreasing input parameters might be performed lead or lag to output and effect to correlation fitting.
3. Splitting of data before and after a turnaround. In the long operation period (before turnaround), cell voltage (CV) increased due to clogging accumulation on the membrane. After long stopping, for example, turnaround or commercial shutdown, CV might drop slightly due to the flushing effect during drainage or non-operation liquid circulation. Some impurities clogged on the membrane surface will loosen at this time.

3.1.5 Method to reduce training data bias.

1. Raw data will be equally separated into 5 groups. To validate and compare the results of each dataset.
    a. Group A (GA): 1/3 of data after preprocessing, before turnaround
    b. Group B (GB): 1/3 of data after preprocessing, before turnaround
    c. Group C (GC): 1/3 of data after preprocessing, before turnaround
    d. Group D (GD): 1/2 of data after preprocessing, after turnaround
    e. Group E (GE): 1/2 of data after preprocessing, after turnaround
2. Raw data will be varied train-to-test ratio by
    a. Train 70%, validate 15% and test 15% (70:30)
    b. Train 80% validate 10% and test 10% (80:20)
    c. Train 90% validate 5% and test 5% (90:10)

Table 3 List of datasets

Abb.	Detail	Period	Estimate no.
RW	Original collected data	01 Mar 22 - 01 Mar 23	8,761 data
PS	Pre-processed, cut stop	01 Mar 22 - 01 Mar 23	< 8,761 data
PU	Pre-processed, cut unsteady	01 Mar 22 - 01 Mar 23	< 8,761 data
PBT	Pre-processed, before T/A	01 Mar 22 - 07 Nov 22	2 of 3 of PU
PAT	Pre-processed, after T/A	11 Nov 22 - 01 Mar 23	1 of 3 of PU
GA	Reducing bias, Group A (PBT)	01 Mar 22 - 24 Apr 22	1,186
GB	Reducing bias, Group B (PBT)	24 Apr 22 - 17 Jun 22	1,186
GC	Reducing bias, Group C (PBT)	17 Jun 22 - 10 Aug 22	1,186
GD	Reducing bias, Group D (PAT)	11 Nov 22 - 06 Jan 23	1,186
GE	Reducing bias, Group E (PAT)	06 Jan 23 - 28 Feb 23	1,186

3.2 DATA ANALYTICS

Based on numerical method. All parameters after preprocessing will manipulate with

1. Generate a correlation heatmap to see the relationship between each parameter, which is valid or invalid to each other when it becomes predictive.
2. Correlation plot between each input parameter to output. And regress by linear regression.
3. Trial second-order correlation plot by fixing the most accurate parameter in constant value to see the second relationship by heatmap of correlation.
4. Create regression equation from correlation.

3.3 NEURAL NETWORK FITTING

This step will be performed in MATLAB via nftool. When the pre-processed data is prepared in a CSV file, the data will be imported to the MATLAB workspace, and the neural network toolbox shall be executed.

3.3.1 Testing effect of number of hidden nodes

After splitting into five groups, the dataset will be trained and tested with various hidden node numbers to find an optimum number of hidden nodes. ANN will be trained with 5-10 hidden nodes and validated in 3 cases of train-to-test ratio.

1. Train 70%, validate 15% and test 15% (70:30)
2. Train 80% validate 10% and test 10% (80:20)
3. Train 90% validate 5% and test 5% (90:10)

3.3.2 Testing effect of number of hidden layers

Same as the previous section dataset will be trained and tested with various hidden layer numbers to find an optimum number of hidden layers. ANN will train with 1-5 hidden layers and validated in 3 cases of train-to-test ratio.

1. Train 70%, validate 15% and test 15% (70:30)
2. Train 80% validate 10% and test 10% (80:20)
3. Train 90% validate 5% and test 5% (90:10)

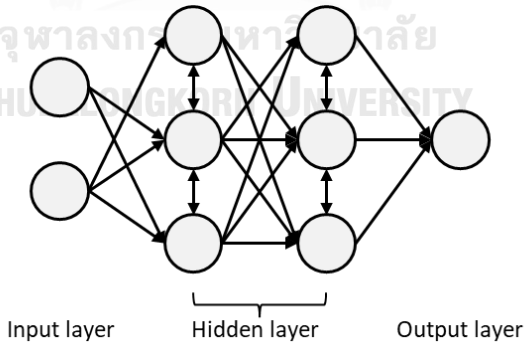


Figure 29 Schematic of simple ANN with 2 inputs, 3 hidden nodes, 2 hidden layers, and 1 output.

3.3.3 Comparing parameters impact

Operating parameters will be set as predicters, and train the ANN model with the Scale Conjugate Gradient algorithm to find regression with CV as a response. The predictors will pair up with each other in one predictor, two predictors, three predictors, and four predictors cases. RMSE will determine the impact of each case to find optimum predictors that give the most accurate CV. Predictors cases in this study are shown in table 3.

Table 4 Summary pairing of predictors cases

Dataset	Input				
	no.	1	2	3	4
PU	1	CD	-	-	-
PU	2	CD	DOL	-	-
PU	2	CD	Q_{FB}	-	-
PU	2	CD	Q_{HD}	-	-
PU	2	CD	T	-	-
PU	3	CD	DOL	Q_{FB}	-
PU	3	CD	DOL	Q_{HD}	-
PU	3	CD	DOL	T	-
PU	3	CD	Q_{FB}	Q_{HD}	-
PU	3	CD	Q_{FB}	T	-
PU	3	CD	Q_{HD}	T	-
PU	4	CD	DOL	Q_{FB}	Q_{HD}
PU	4	CD	DOL	Q_{FB}	T
PU	4	CD	DOL	Q_{HD}	T
PU	4	CD	Q_{FB}	Q_{HD}	T

3.4 PERFORMANCE EVALUATION

The performance of each model will be calculated based on its error as root mean square error (RMSE) for eliminate effect of subtraction error and determined in absolute form. The preliminary result shows mean square error in the range of 0.00001 to 0.001, which RMSE is better for comparing the slight difference in prediction error with the original scale.

Root Mean Square Error (RMSE)

$$MAE = \frac{1}{n}\sum \sqrt{(y-y_i)^2}$$

While; y is actual data, y_i is regressed data at the same input, n is number of data

3.5 BLIND TEST AND PARAMETERS OPTIMIZATION STUDY

After testing the number of hidden nodes, number of hidden layers, and impact of parameters, the optimum result with the lowest RMSE will implement to blind-test with actual data. In this thesis, training and validate datasets were puck at the dataset before the turnaround, and the test performance of the model was with actual data in another dataset, which is the dataset after the turnaround, to see the accuracy of the ANN model.

Table 5 Dataset for train, validate ANN model and for test model performance

Abb.	Detail	Period	ANN group
PBT	Pre-processed, before T/A	01 Mar 22 - 07 Nov 22	Train and Validate
PAT	Pre-processed, after T/A	11 Nov 22 - 01 Mar 23	Test

Study of parameters optimization, the optimum model will be used to predict cell voltage (CV) in multiple cases to find results after selecting the operating parameter in each mode. There is plotting the result of predicted cases compared with the actual of each parameter's upper and lower limit.

CHAPTER 4 : RESULTS AND DISCUSSION

4.1 DATASET AND PREPROCESSING

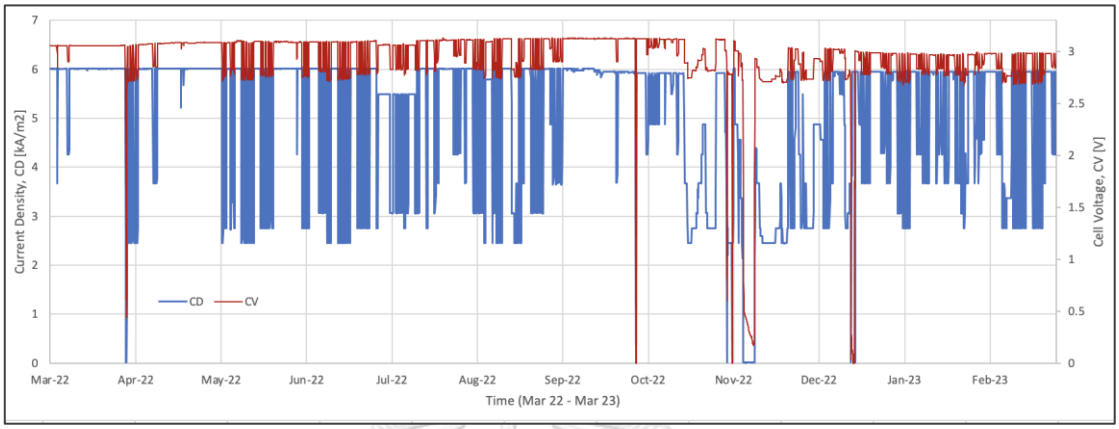


Figure 30 Raw data of CD and CV plot with time

Raw data of CD and CV appear in fluctuating trend according to practical daily production plans designed to operate high capacity (high CD) in off-peak periods to minimize power consumption in on-peak periods for operation cost optimization. These raw data will be preprocessed by removing outliers from stopping, removing unsteady state period, selected data range to avoid turnaround, respectively. To compare the readiness of the theoretical dataset base will be applied by referring to the linear correlation between CD and CV of Ohm's law as equation 15.

4.1.1 Result of removing outlier data

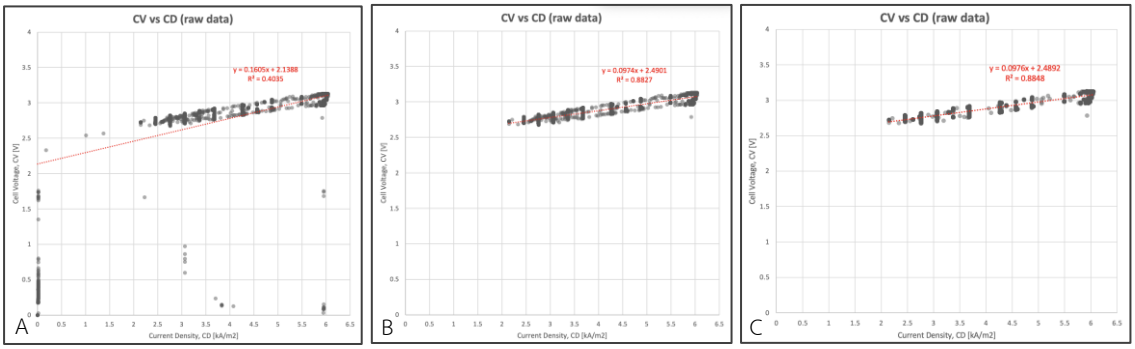


Figure 31 Linear regression of CD and CV in various dataset

A) Linear regression of CD and CV by raw data

B) Linear regression of CD and CV after remove outlier

C) Linear regression of CD and CV after removing unsteady state period

Correlation plot between CD and CV, graph show in linear form as Ohm's law. *Figure 31 A* raw data plot without preprocessing appears to scatter the outlier, which makes the linear regression trend deviate from result RSQ 0.4035 *Figure 31 B* After cutting the outlier by removing stopping period data result in RSQ 0.8829, which is much better and can be assumed to be linear. Conversely, this dataset will analyze with other parameters, e.g., DOL, Q_{FB}, Q_{HD}, and T, which may be found in lead or lag data changing during a heating-up or load-changing step. In eliminating the problem, data between 08.00 and 22.00 of all day will remove from all unsteady state data in the load-changing period between on-peak and off-peak. After cutting unsteady state data, the result shows RSQ 0.8848, which is like the previous step, but stabilizes other parameters; this step must perform.

4.1.2 Result of reducing training data bias

Table 6 Summary of dataset, period and number of data

Abb.	Detail	Period	Amount data	RSQ
RW	Original collected data	01 Mar 22 - 01 Mar 23	8,761	0.4035
PS	Pre-processed, cut stop	01 Mar 22 - 01 Mar 23	8,599	0.8829
PU	Pre-processed, cut unsteady	01 Mar 22 - 01 Mar 23	7,881	0.8848
PBT	Pre-processed, before T/A	01 Mar 22 - 07 Nov 22	5,508	0.9612
PAT	Pre-processed, after T/A	11 Nov 22 - 01 Mar 23	2,373	0.9787
GA	Reducing bias, Group A (PBT)	01 Mar 22 - 24 Apr 22	1,186	0.9983
GB	Reducing bias, Group B (PBT)	24 Apr 22 - 17 Jun 22	1,186	0.9955
GC	Reducing bias, Group C (PBT)	17 Jun 22 - 10 Aug 22	1,186	0.9955
GD	Reducing bias, Group D (PAT)	11 Nov 22 - 06 Jan 23	1,186	0.9728
GE	Reducing bias, Group E (PAT)	06 Jan 23 - 28 Feb 23	1,186	0.9955

Data after preprocess by remove outlier and unsteady state period following assumption of after long-term stop such as turnaround or cell maintenance, CV refer to *Figure 23 C* scatter of data seem split to 2 linear trends. To avoid dataset across this period data were split into 5 groups which GA, GB and GC were collected before turnaround and GD and GE were collected after turnaround.

Datasets GA, GB, GC, GD, and GE will be used to train models to compare each algorithm's performance, varying the number of hidden nodes and varying number of hidden layers. It also tests in various train-to-test ratios at 70:30, 80:20, and 90:10. The use of small datasets and many groups proposes to reduce data bias.

Datasets PBT and PAT will be used to train models to compare the impact of each parameter because of more data amount and cover more operation cases. PBT will be used in the train and validate model, and PAT will be used in the test-to-test model performance without bias.

4.2 DATA ANALYTICS

4.2.1 Correlation between each parameter to output

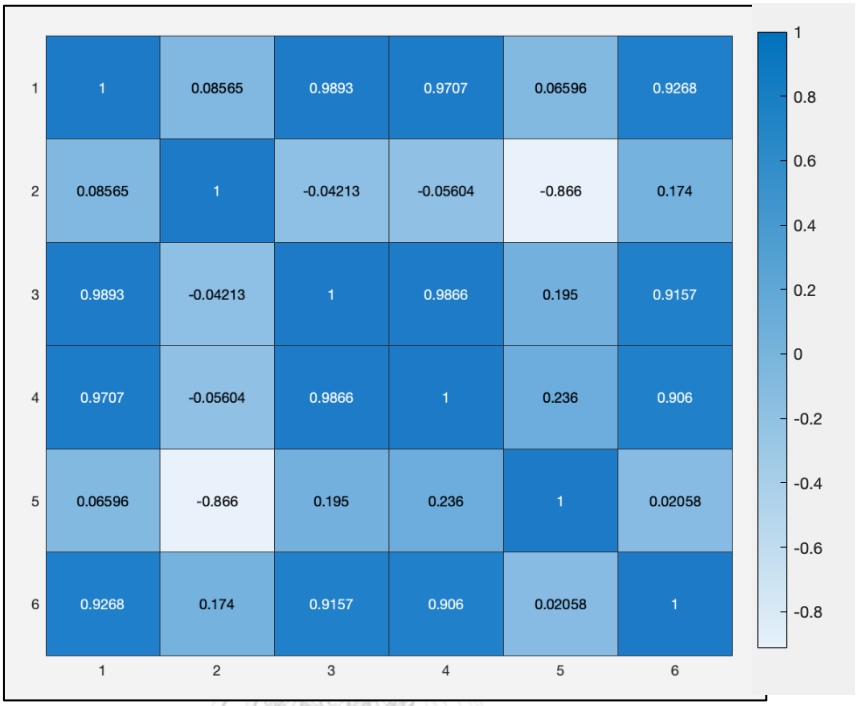


Figure 32 Heat map correlation between each parameter

Table 7 Definition of number and parameters in heat map Figure 24

No.	Type	Abbr.	Description	Range	Unit
1	Output data	CV	Cell voltage	2.0-3.4	V
2	Input data 1	DOL	Operation day	0-365	Day
3	Input data 2	CD	Current density	1.8-6.0	KA/m^2
4	Input data 3	Q_{FB}	Feed brine flow rate	12.0-30.0	m^3/h
5	Input data 4	Q_{HD}	Feed brine flow rate	31 ± 1	m^3/h
6	Input data 5	T	Cell temperature	85-95	°C

Result of correlation heat map shown

1. Correlation from 5 inputs to output
    a. The highest correlation is CD. It directly correlates to CV fitted in linear regression following Ohm's law in equation 15. Current density determines production capacity following Faraday's law in equations 12, 13. Furthermore, other operating controlled parameters will be adjusted to support CD in production.
    b. Q_{FB}, In production philosophy, the brine feed flow rate is adjusted to control the concentration on the membrane and feed NaCl as raw material in production. The significant correlations occur by manipulation.
    c. T, the temperature controls membrane channel sizing to achieve optimum voltage and impurities content. Both Q_{FB} and T are not direct parameters t control CV, but they are dependent variables that highly accurate regression from the effect of CD or production capacity.
2. Internal correlation between 5 inputs

    There are three correlations between input parameters: CD-Q_{FB}, CD-T, and T-Q_{FB}. All 3 cases show an invalid relationship which is the effect of dependent variables (Q_{FB} and T) on the primary variable (CD), as explained in the previous section.

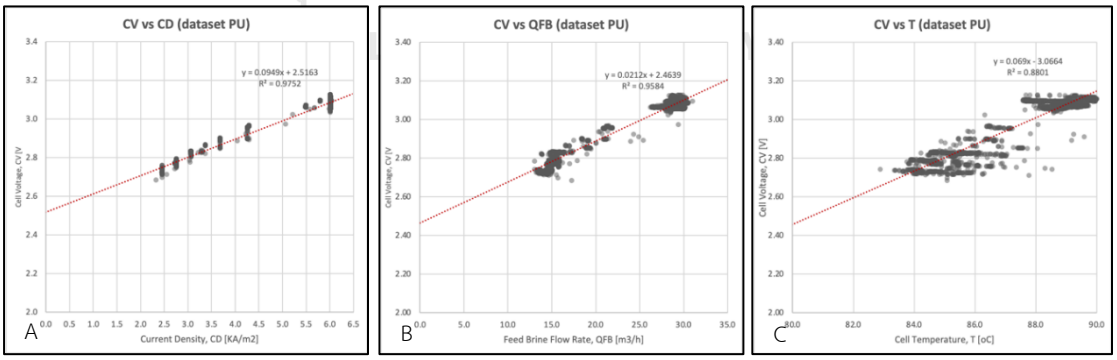


Figure 33 Linear regression plot of CD, Q_{FB} and Q_{HD} (significant parameters) to CV

A) Linear regression plot of CD to CV gives RSQ 0.9753

B) Linear regression plot of Q_{FB} to CV gives RSQ 0.9584

C) Linear regression plot of T to CV gives RSQ 0.8801

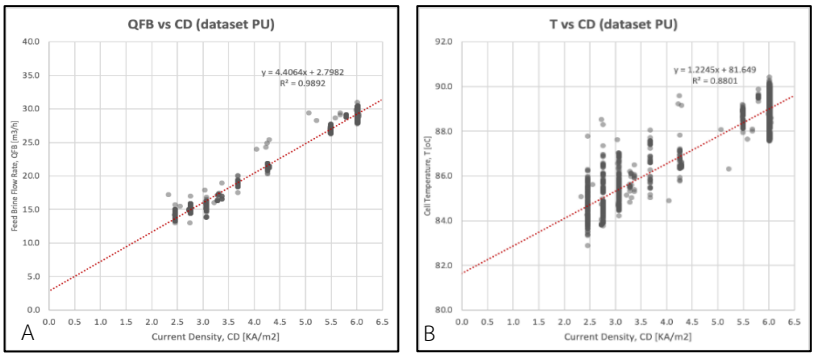


Figure 34 Linear regression plot of CD, Q_{FB} and Q_{HD} to CD (between inputs)

A) Linear regression plot of Q_{FB} to CD gives RSQ 0.9892

B) Linear regression plot of Q_{FB} to CV gives RSQ 0.8801

Linear regression plot of dependent variables, Q_{FB} and T, to a primary variable, CV, shows the high accuracy of a linear equation to data. That shows Q_{FB}, T, and CV are not independent. Both are controlled parameters to support each CD operation.

4.2.2 Correlation between each parameter to output, eliminate CD

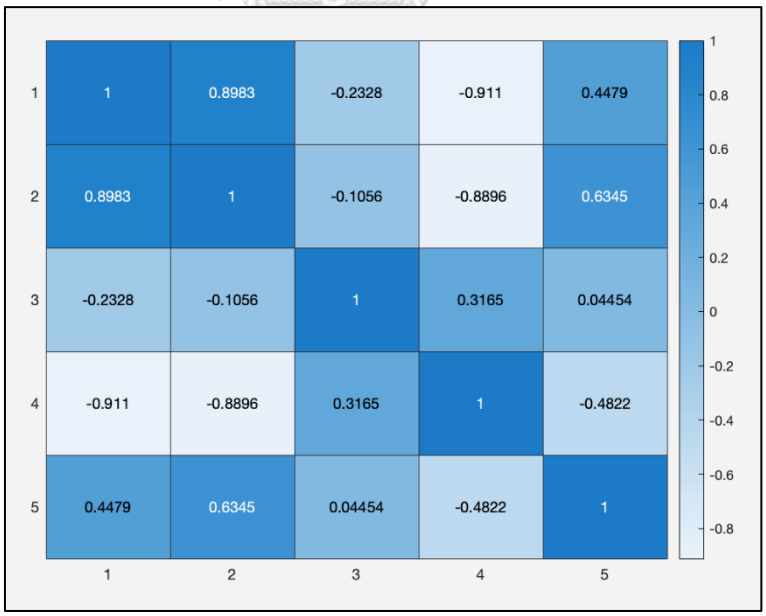


Figure 35 Heat map correlation between each parameter at constant CD (6 KA/m^2)

Table 8 Definition of number and parameters in heat map Figure 27

No.	Type	Abbr.	Description	Range	Unit
1	Output data	CV	Cell voltage	2.0-3.4	V
2	Input data 1	DOL	Operation day	0-365	Day
3	Input data 3	Q_{FB}	Feed brine flow rate	12.0-30.0	m^3/h
4	Input data 4	Q_{HD}	Feed brine flow rate	31±1	m^3/h
5	Input data 5	T	Cell temperature	85-95	°C

Second-order correlation plot at constant CD 6 KA/m^2. Found DOL has a relationship with CV in linear. Linear regression of the correlation gave RSQ 0.8829. While Q_{FB} and T, which correlated highly to CD and CV, do not have a significant relationship to CV. That means a high correlation to CV in heatmap mainly occurs by CD's effect.

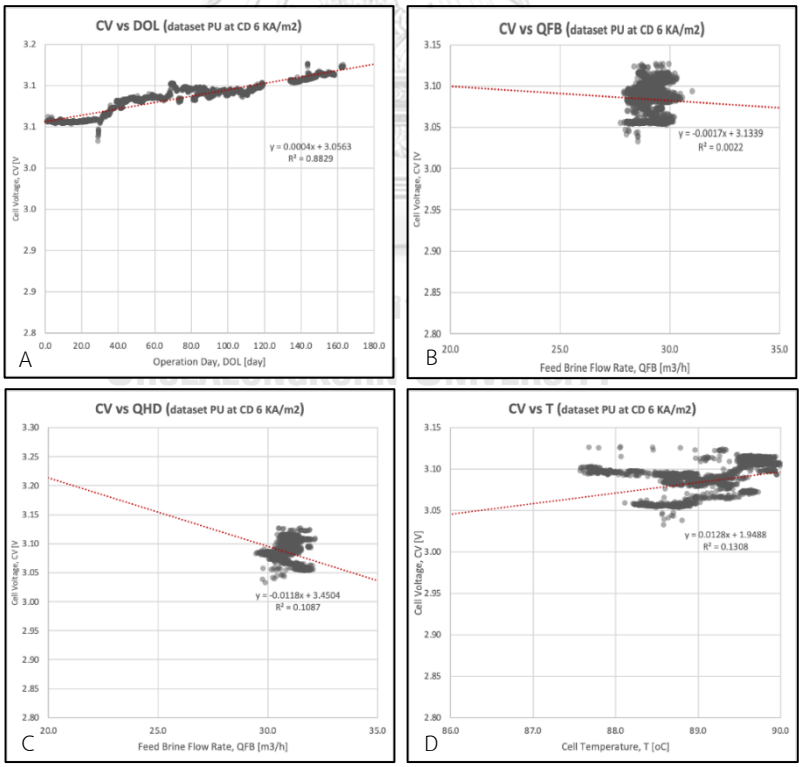


Figure 36 Linear regression plot of DOL, Q_{FB}, Q_{HD} and T to CV at constant CD

A) DOL and CV gives RSQ 0.8829 *B) Q_{FB} and CV gives RSQ 0.0022*

C) Q_{HD} and CV gives RSQ 0.1087 *D) T and CV gives RSQ 0.1308*

4.3 NEURAL NETWORK FITTING

4.3.1 Result of performance in comparing ANN algorithm

Table 9 Result of performance in comparing ANN algorithm

Dataset	Performance (RMSE, V) on test group							
	Scale Conjugate Gradient algorithm (SCG)			Levenberg Marquardt algorithm (LM)			Bayesian Regularization algorithm (BR)	
	Train	Validate	Test	Train	Validate	Test	Train	Test
GA	0.0092	0.0111	0.0192	0.0022	0.0021	0.0028	0.0010	0.0010
GB	0.0065	0.0067	0.0052	0.0013	0.0013	0.0017	0.0014	0.0014
GC	0.0059	0.0061	0.0072	0.0014	0.0016	0.0017	0.0013	0.0016
GD	0.0067	0.0077	0.0074	0.0028	0.0029	0.0056	0.0023	0.0039
GE	0.0023	0.0022	0.0022	0.0016	0.0014	0.0016	0.0014	0.0017
Average	0.0082 V			0.0027 V			0.0019 V	

The performance result in various ANN algorithms gave acceptable RMSE (accept at RMSE < 0.01 V due to the original data source being used to calculate at two decimal numbers). BR gave the best performance result with the lowest RMSE; the second was LM. Usually, BR deals with small datasets and is very good at handling noise which might not be suitable for implementation with actual commercial plants. LM is usually the fastest learning algorithm, but it takes more memory in the calculation. SCG uses less memory than other algorithms. For the dataset in this research, memory efficiency seems more practical to implement, and all gave acceptable results also. SCG will be used as an algorithm to train ANN in the next section.

4.3.2 Finding optimum number of hidden node

Table 10 Result of varying number of hidden node with train-to-test ratio 70:30

Dataset	Performance (RMSE, V) on test group					
	5 nodes	6 nodes	7 nodes	8 nodes	9 nodes	10 nodes
GA	0.0071	0.0113	0.0039	0.0056	0.0043	0.0091
GB	0.0108	0.0062	0.0037	0.0041	0.0058	0.0065
GC	0.0031	0.0081	0.0056	0.0060	0.0064	0.0051
GD	0.0058	0.0051	0.0054	0.0177	0.0064	0.0097
GE	0.0036	0.0060	0.0025	0.0030	0.0028	0.0058
Average	0.0061	0.0037	0.0042	0.0073	0.0051	0.0062
Accuracy rank	4	1	2	6	3	5

Table 11 Result of varying number of hidden node with train-to-test ratio 80:20

Dataset	Performance (RMSE, V) on test group					
	5 nodes	6 nodes	7 nodes	8 nodes	9 nodes	10 nodes
GA	0.0050	0.0054	0.0070	0.0051	0.0368	0.0062
GB	0.0170	0.0086	0.0036	0.0029	0.0085	0.0114
GC	0.0139	0.0157	0.0087	0.0073	0.0135	0.0090
GD	0.0070	0.0054	0.0057	0.0069	0.0069	0.0075
GE	0.0049	0.0059	0.0076	0.0055	0.0033	0.0046
Average	0.0096	0.0066	0.0065	0.0055	0.0122	0.0077
Accuracy rank	5	3	2	1	6	4

Table 12 Result of varying number of hidden node with train-to-test ratio 90:10

Dataset	Performance (RMSE, V) on test group					
	5 nodes	6 nodes	7 nodes	8 nodes	9 nodes	10 nodes
GA	0.0063	0.0117	0.0081	0.0096	0.0044	0.0063
GB	0.0066	0.0089	0.0042	0.0171	0.0086	0.0042
GC	0.0061	0.0030	0.0099	0.0100	0.0153	0.0074
GD	0.0083	0.0095	0.0144	0.0157	0.0167	0.0055
GE	0.0036	0.0054	0.0048	0.0082	0.0028	0.0034
Average	0.0062	0.0077	0.0083	0.0121	0.0076	0.0054
Accuracy rank	2	4	5	6	3	1

The results of varying hidden nodes 5 – 10 show

1. All node numbers gave acceptable results (accept at RMSE < 0.01 V due to the original data source was used to calculate at two decimal numbers) except seven nodes in 90:10 train-to-test-ratio gave RMSE higher than 0.121 V, which was from dataset GC and GE, it might be close to turning around period. It might appear less accurate from a small test ratio.
2. No significantly different between each number on nodes RMSE result. In the test-to-test ratio 70:30, 80:20, and 90:10, six hidden nodes, eight hidden nodes, and ten nodes gave the lowest RMSE in each case.
3. This thesis used six hidden nodes in further study to follow rule-of-thumb inputs+1 (n+1) hidden nodes, and this thesis will perform with a 70:30 train-to-test ratio to handle the amount of data.

4.3.3 Finding optimum number of hidden layer

Table 13 Result of varying number of hidden layer with train-to-test ratio 70:30

Dataset	Performance (RMSE, V) on test group				
	1 layer	2 layers	3 layers	4 layers	5 layers
GA	0.0154	0.0081	0.0043	0.0050	0.0075
GB	0.0063	0.0106	0.0106	0.0199	0.0107
GC	0.0071	0.0070	0.0183	0.0171	0.0071
GD	0.0078	0.0058	0.0136	0.0089	0.0147
GE	0.0035	0.0043	0.0050	0.0062	0.0033
Average	0.0062	0.0072	0.0104	0.0114	0.0087
Accuracy rank	1	2	4	5	3

Table 14 Result of varying number of hidden layer with train-to-test ratio 80:20

Dataset	Performance (RMSE, V) on test group				
	1 layer	2 layers	3 layers	4 layers	5 layers
GA	0.0095	0.0070	0.0076	0.0100	0.0090
GB	0.0062	0.0168	0.0089	0.0154	0.0061
GC	0.0037	0.0051	0.0054	0.0210	0.0196
GD	0.0073	0.0067	0.0068	0.0118	0.0090
GE	0.0042	0.0089	0.0072	0.0038	0.0087
Average	0.0062	0.0089	0.0072	0.0124	0.0105
Accuracy rank	1	3	2	5	4

Table 15 Result of varying number of hidden layer with train-to-test ratio 90:10

Dataset	Performance (RMSE, V) on test group				
	1 layer	2 layers	3 layers	4 layers	5 layers
GA	0.0073	0.0087	0.0082	0.0209	0.0121
GB	0.0053	0.0194	0.0200	0.0169	0.0189
GC	0.0076	0.0111	0.0050	0.0108	0.0217
GD	0.0113	0.0065	0.0060	0.0105	0.0310
GE	0.0070	0.0077	0.0049	0.0085	0.0063
Average	0.0077	0.0107	0.0088	0.0135	0.0180
Accuracy rank	1	3	2	4	5

The results of varying hidden layers 1 – 5 show

1. At 1-2, the hidden layer gave an acceptable result (accept at RMSE < 0.01 V due to the original data source used to calculate at two decimal numbers). At higher hidden layers, RMSE was also getting higher. 4-5 hidden layers have RMSE > 0.01 V; almost all cases that occurred from correlation in ANN overfit led to test results worsening RMSE.
2. This thesis used one layer in the further study because it gave the lowest RMSE in all cases, all in an acceptable range.

4.3.4 Result of comparing parameters impact

Table 16 Result of performance in comparing parameter impact

Dataset	Input					Performance (RMSE, V) on test group			Performance Ranking
	no.	1	2	3	4	Train	Validate	Test	
PU	1	CD	-	-	-	0.0170	0.0175	0.0167	12
PU	2	CD	DOL	-	-	0.0059	0.0052	0.0065	4
PU	2	CD	Q_{FB}	-	-	0.0174	0.0186	0.0171	13
PU	2	CD	Q_{HD}	-	-	0.0162	0.0175	0.0184	14
PU	2	CD	T	-	-	0.0141	0.0156	0.0148	9
PU	3	CD	DOL	Q_{FB}	-	0.0062	0.0063	0.0066	5
PU	3	CD	DOL	Q_{HD}	-	0.0068	0.0066	0.0070	‘6
PU	3	CD	DOL	T	-	0.0047	0.0044	0.0043	1
PU	3	CD	Q_{FB}	Q_{HD}	-	0.0142	0.0138	0.0143	8
PU	3	CD	Q_{FB}	T	-	0.0159	0.0161	0.0159	11
PU	3	CD	Q_{HD}	T	-	0.0156	0.0145	0.0149	10
PU	4	CD	DOL	Q_{FB}	Q_{HD}	0.0055	0.0057	0.0057	3
PU	4	CD	DOL	Q_{FB}	T	0.0073	0.0078	0.0081	7
PU	4	CD	DOL	Q_{HD}	T	0.0055	0.0051	0.0052	2
PU	4	CD	Q_{FB}	Q_{HD}	T	0.0227	0.0211	0.0221	15

The performance in comparing parameters impact showed a case of 3 predictors with CD, DOL, and T giving the highest accuracy in CV prediction.

Comparing direct correlation of CD as predictor and CV as response.

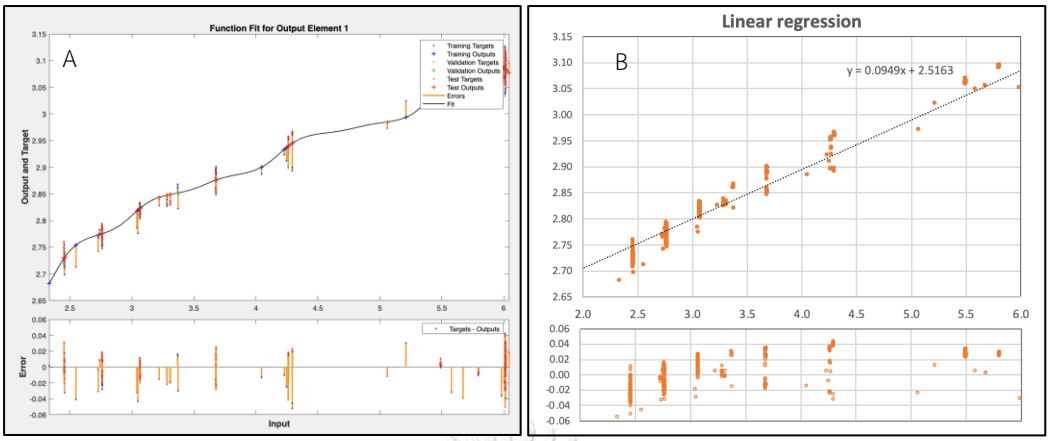


Figure 37 Correlation plot between CD and CV fitting

A) Fitting by scale conjugate gradient result in RMSE = 0.0167

B) Fitting by linear regression result in RMSE = 0.0197

The regression plot of CD and CV complies with the theory. Found a wide cluster of data in step intervals, which were from the operating point of current density. The actual operation is performed in daily capacity swing following product demand. There is a practical capacity adjustment.

The result of the correlation plot of 1 predictor between CD and CV by Scale Conjugate Gradient gave a lower RMSE, which is the result of non-linear fitting performing.

Comparing of selected 2 predictors to CV response

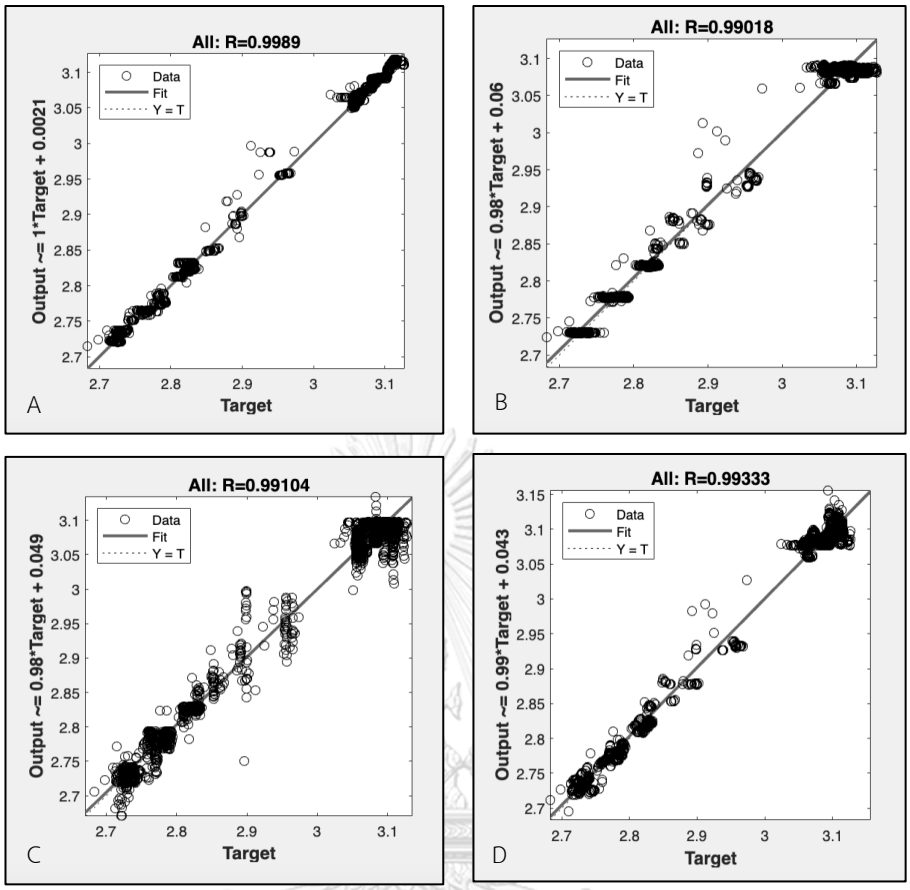


Figure 38 Regression plot of 2 input predictors

A) Regression plot 2 predictors CD, DOL by scale conjugate gradient

B) Regression plot 2 predictors CD, Q_{FB} by scale conjugate gradient

C) Regression plot 2 predictors CD, Q_{HD} by scale conjugate gradient

D) Regression plot 2 predictors CD, T by scale conjugate gradient

The plot of 2 predictors. The case of CD and DOL gave the lowest RMSE, which can conclude that the main deviation factor in each cluster deviation in *Figure 38 A* is DOL (operating time) which directly relates to clogging accumulation on the membrane. On the contrary, the feed flow rate of both Q_{FB} and Q_{HD} gave worse RMSE. Slight fluctuation of these parameters is not directly impacted cluster deviation in a steady state period. It might appear impaction in transient state data, which eliminate from the dataset in pre-processing.

Comparing of selected 4 predictors to CV response.

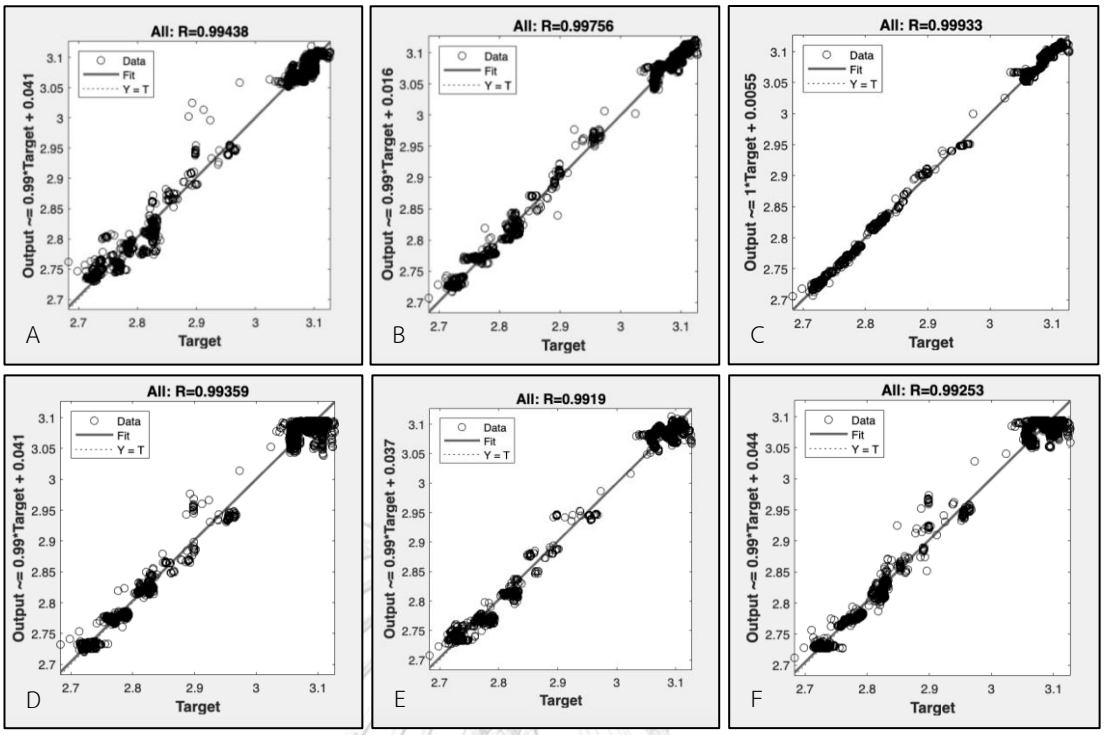


Figure 39 Regression plot of 3 input predictors

A) Regression plot CD, DOL, Q_{FB} by scale conjugate gradient

B) Regression plot CD, DOL, Q_{HD} by scale conjugate gradient

C) Regression plot CD, DOL, T by scale conjugate gradient

D) Regression plot CD, Q_{FB}, Q_{HD} by scale conjugate gradient

E) Regression plot CD, Q_{FB}, T by scale conjugate gradient

F) Regression plot CD, Q_{HD}, T by scale conjugate gradient

The plot of 3 predictors. The CD, DOL, and T cases in *Figure 31 C* gave the lowest RMSE. Referring to the plot of 2 predictors, DOL and T were in the set that gave the lowest RMSE. That from in some period that T was adjusted, for example, steam utility management or optimizing cell voltage after long operation by increasing temperature. Only DOL cannot regress the model in that period to fit the response (T can see in disturbance). To conclude, temperature helps to deduct the plot of predictors CD and DOL to the standard line leading to the model becoming more fitted.

Comparing of selected 4 predictors to CV response.

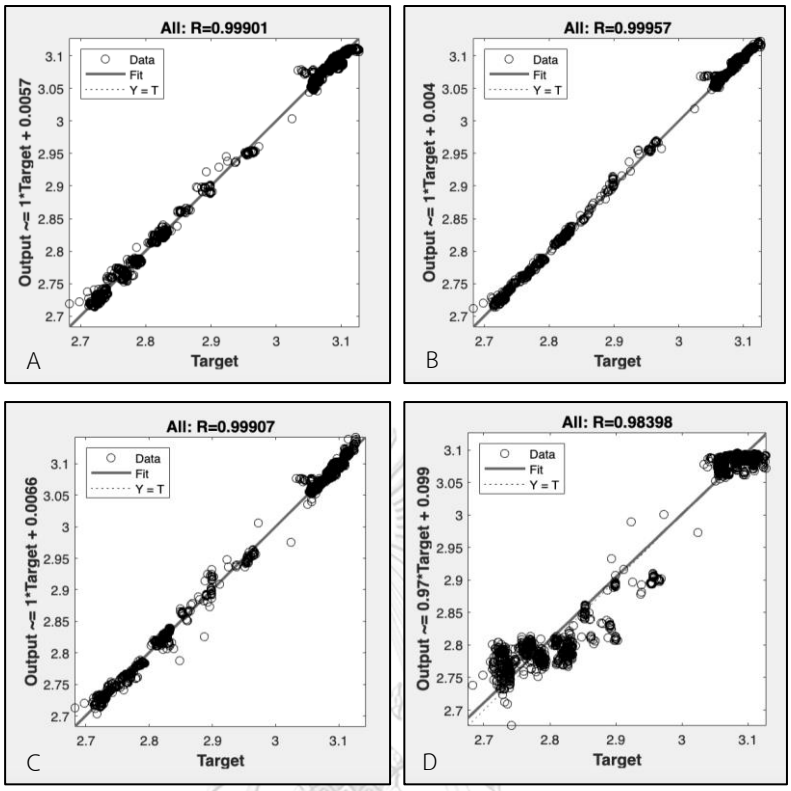


Figure 40 Regression plot of 4 input predictors

A) Regression plot CD, DOL, Q_{FB}, Q_{HD} by scale conjugate gradient,

B) Regression plot CD, DOL, Q_{FB}, T by scale conjugate gradient,

C) Regression plot CD, DOL, Q_{HD}, T by scale conjugate gradient,

D) Regression plot CD, Q_{FB}, Q_{HD}, T by scale conjugate gradient

The plot of 2 predictors. Not found significantly better correlation than the three predictors. In *Figures 40 A* and *Figure 40 C*, cases of CD, DOL, Q_{FB}, Q_{HD} and CD, DOL, Q_{HD}, T as predictors gave RMSE the lowest, but higher than three predictors cases of CD, DOL, T. In the case of *Figure 40 C* result becomes worse after adding Q_{HD} as the fourth parameter from no relationship and making the model overfit with noise. In the case of *Figure 40 A*, the lowest RMSE without T in a predictive model is the lowest. It might be from Q_{FB} and regress deviation as T as plot in section 4.2.

4.3.5 Summary of accuracy in each case.

Table 17 Summary of accuracy in each case

Dataset	Input					Performance (RMSE, V) on test group			Performance Ranking
	no.	1	2	3	4	Train	Validate	Test	
PU	3	CD	DOL	T	-	0.0047	0.0044	0.0043	1
PU	4	CD	DOL	Q_{HD}	T	0.0055	0.0051	0.0052	2
PU	4	CD	DOL	Q_{FB}	Q_{HD}	0.0055	0.0057	0.0057	3
PU	2	CD	DOL	-	-	0.0059	0.0052	0.0065	4
PU	3	CD	DOL	Q_{FB}	-	0.0062	0.0063	0.0066	5

DOL has the highest impact on cell voltage. The wide range of data of clusters in each step of operation capacity is mainly caused by membrane efficiency down by operation time or clogging by impurities. However, DOL cannot control; only feed purities can monitor to keep stable DOL (quality control was not measured and mentioned in this thesis).

T was the first runner-up for highest impact. It can synergy with DOL in model regression. In some cases of deviation from the regular operation, for example, adjust T for optimum cell voltage or steam management. DOL cannot regress the model to a standard line because T was disturbing operation data.

In the case of 4 predictors, the feed flow rate of both Q_{FB} and Q_{HD}, which mainly deviate by fluctuation, results in noise, not the adjustment of the main parameters. That gave worse RMSE, especially Q_{FB}. However, this research analyzes the steady-state dataset in which the feed flow rate was unaffected. If the dataset includes load changing period, it might be more impact.

4.4 RESULT OF BLIND TEST AND PARAMETERS OPTIMIZATION

4.4.1 Blind test with data before turnaround

After performing hold-on validation, the first step is to check the model accuracy at the same condition (before turnaround) by selecting data in March 2022 to June 2022 to train and validate the ANN model and selecting data in August 2022 to test for unseen testing-dataset in model evaluation. The results show that

1. By linear regression (equation created from CD and DOL) show RMSE in the train group at 0.0276 V and test group at 0.0206 V
2. By ANN prediction (model created from CD, DOL, and T) show RMSE in the train group at 0.0019 V and test group at 0.0181 V

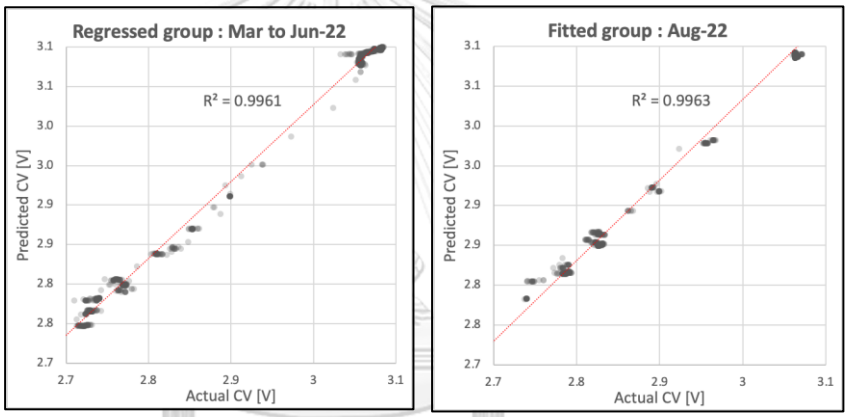


Figure 41 Linear regression fitting of test data in August 2022

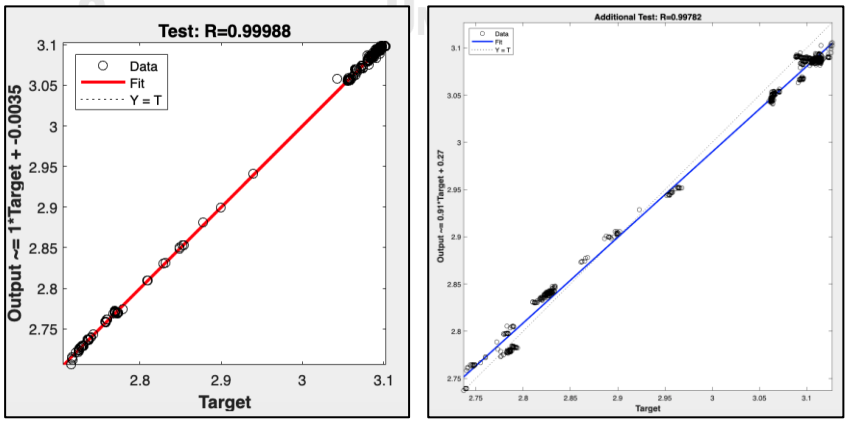


Figure 42 ANN model fitting of test data in August 2022

4.4.2 Blind test with data after turnaround

The Second step is to evaluate the model with different conditions by training the model before turnaround and testing with data before the turnaround period. They selected data from March 2022 to June 2022 to train and validate the ANN model (same as 4.4.1) and selected data in January 2023 to test. Cell voltage shows they have significantly different conditions from the flushing effect during the turnaround. To compare with the same ANN model, It has to reduce CV by 3% to represent the flushing effect. The results show that

1. By linear regression (equation created from CD and DOL) show RMSE in the training group at 0.0276 V and test group at 0.0137 V
2. By ANN prediction (model created from CD, DOL, and T) show RMSE in the training group at 0.0019 V and test group at 0.0126 V

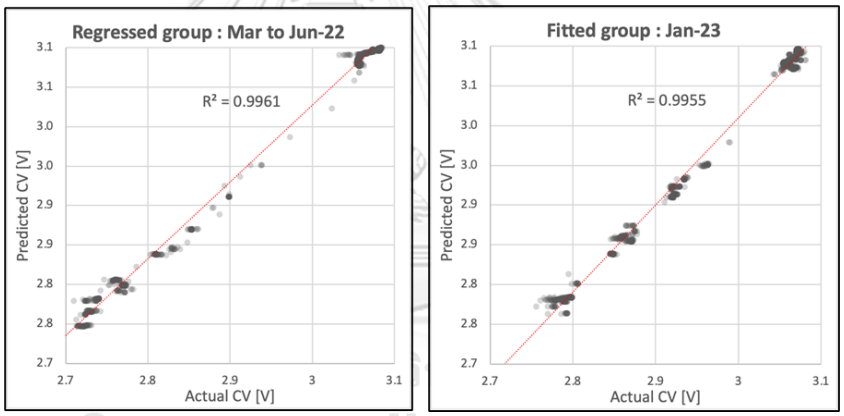


Figure 43 Linear regression fitting of test data in January 2023

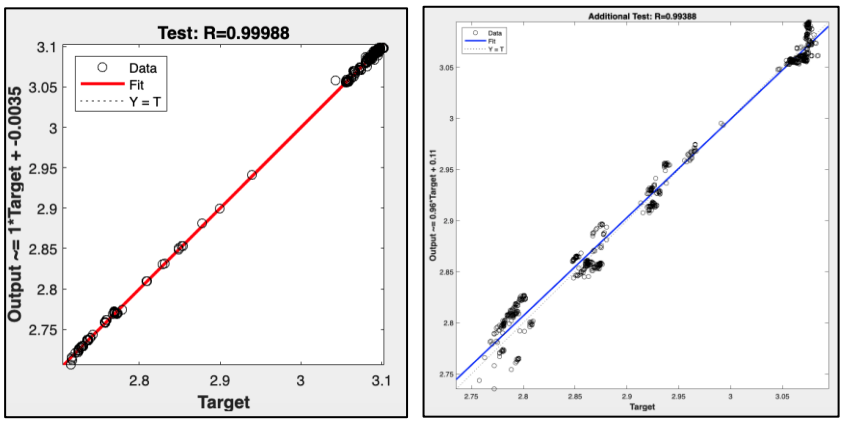


Figure 44 ANN model fitting of test data in January 2022

4.4.2 Parameters optimization study

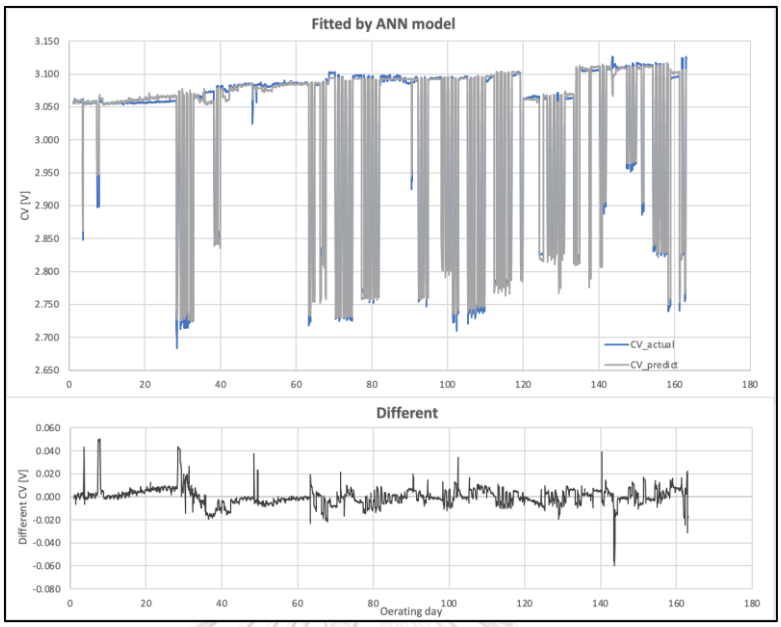


Figure 45 ANN model fitting in predictive CV (grey) comparing with actual data (blue)

The predictive model of ANN can handle noise from slightly fluctuating operating parameters, and the non-linear function seems more fitted. Some peaks could not be regressed at low production load after regressing CV to the standard line at CD 6 KA/m^2. Since we regressed CD to a high load, cell temperature was not regressed (T was a dependent variable that can affect as input parameter).

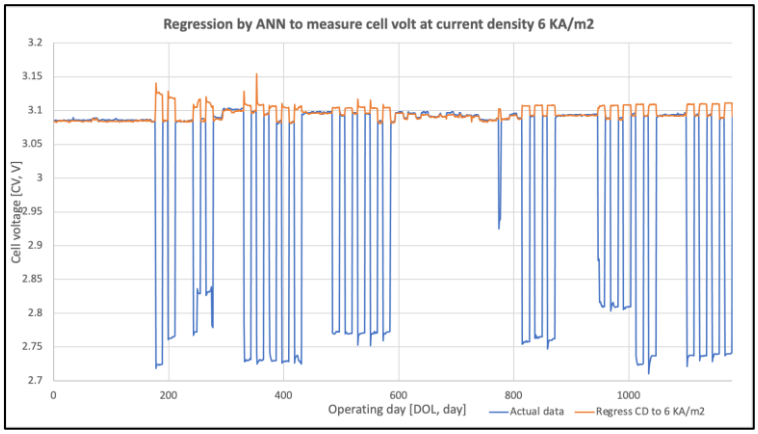


Figure 46 ANN model fitting in predictive CV after at CD 6 KA/m^2 (orange) comparing with actual data (blue)

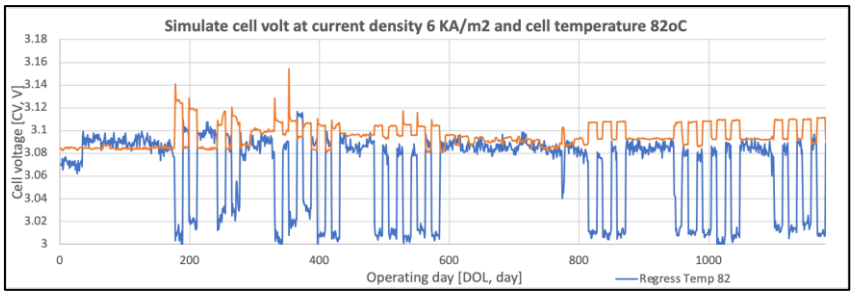


Figure 47 ANN model fitting in predictive CV at CD 6 KA/m^2 and 82°C (blue) Comparing with predictive CV at CD 6 KA/m^2 actual T (orange)

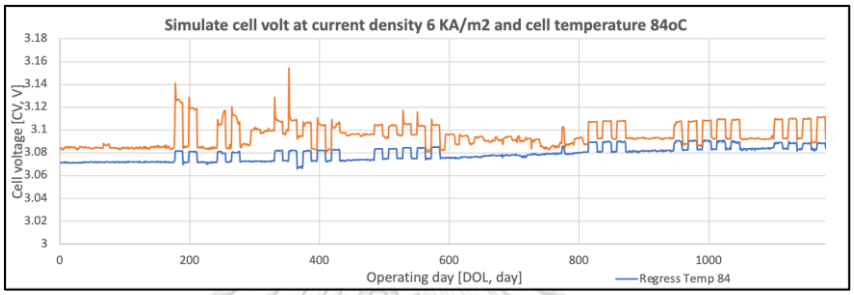


Figure 48 ANN model fitting in predictive CV at CD 6 KA/m^2 and 84°C (blue) Comparing with predictive CV at CD 6 KA/m^2 actual T (orange)

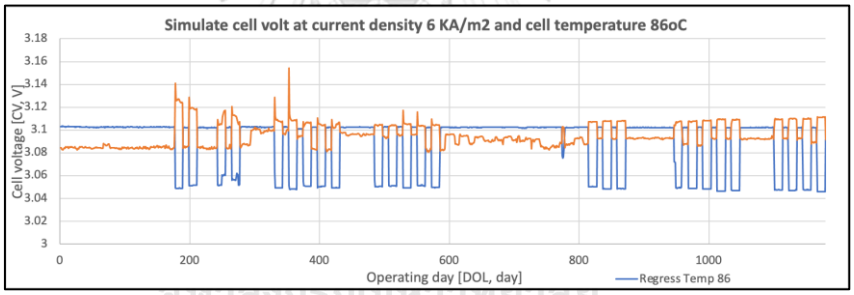


Figure 49 ANN model fitting in predictive CV at CD 6 KA/m^2 and 86°C (blue) Comparing with predictive CV at CD 6 KA/m^2 actual T (orange)

After regressing the temperature to see only the effect of DOL and finding CV growth due to operating time, it was found that temperatures below and over 82°C gave CV more deviation than CV at actual temperature. That means at 82°C and below, cell condition at various temperatures, as actual, is more stable in the wind range than at 82°C. However, cell temperature at 84°C gave the most stable CV in all ranges. Predicted values after regressing to 84°C deviate from the growth line by less than 0.02 V. At cell temperature, 84°C represented all CD at the same condition compared with other parameters.

CHAPTER 5 : CONCLUSION

Finding an optimum number of hidden nodes and layers with scale conjugate gradient algorithm found that 6-8 accuracy is not significantly different between hidden nodes. All gave acceptable RMSE; one hidden layer gave the most accuracy.

In comparing each parameter impact to cell voltage (CV) found that the case that gave minimum RMSE was three predictors consisting of current density (CD), operation date (DOL), and cell temperature (T) at 0.0043 in the test group. Exclude CD, which is theoretically directly related to CV; DOL has the highest impact on the cell voltage. The clusters in step of production capacity is caused by membrane efficiency down by operation time. The second impact parameter is T. T is directly related to ion-exchange efficiency by channel size to be extended or shrunk. Dataset some has appeared different T at the same capacity. It was to control plant stability.

Meanwhile, In the case of CD and DOL as predictors, RMSE was 0.0065. On the other hand, feed brine flow rate (Q_{FB}) and feed caustic flow rate (Q_{HD}) do not have a synergy to reduce RMSE after implementing additional parameters. In the case of implying Q_{FB} or Q_{HD} as the third or fourth predictor behind DOL and T, RMSE results become worse from non-related noise. It can assume that Q_{FB} and Q_{HD} control in range during steady-state operation. However, a dataset for training ANN removed the unsteady state period, which the impact of Q_{FB} and Q_{HD} might be more apparent.

The study can conclude that CD, DOL, and T as ANN 3 predictors have sufficient accuracy to predict CV and, after the predict CV in blind test (hold-out validation), found that in different operating condition models still give enough accuracy in prediction.

The study of optimum conditions by normalized CV. The result shown at 84°C of the test and train condition dataset gave the most stable CV that can represent all operation CD (production capacity) with an error less than 0.02 V. This thesis can apply to stabilize power consumption by predicting CV after extending membrane service life by controlling T in optimum value or predicting corrected CV value to comparing another uncontrol condition for measuring the performance of electrolysis cells.

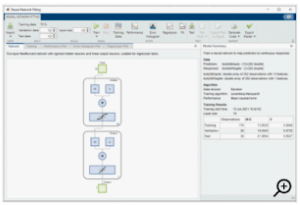
APPENDIX

Neural Net Fitting product description

Neural Net Fitting

Solve fitting problem using two-layer feed-forward networks

Description



The **Neural Net Fitting** app lets you create, visualize, and train a two-layer feed-forward network to solve data fitting problems.

Using this app, you can:

- Import data from file, the MATLAB® workspace, or use one of the example data sets.
- Split data into training, validation, and test sets.
- Define and train a neural network.
- Evaluate network performance using mean squared error and regression analysis.
- Analyze results using visualization plots, such as regression fit or histogram of errors.
- Generate MATLAB scripts to reproduce results and customize the training process.
- Generate functions suitable for deployment with MATLAB Compiler™ and MATLAB Coder™ tools, and export to Simulink® for use with Simulink Coder.

Note

To interactively build, visualize, and train deep learning neural networks, use the **Deep Network Designer** app. For more information, see Get Started with Deep Network Designer.

Open the Neural Net Fitting App

- MATLAB Toolstrip: On the **Apps** tab, under **Machine Learning and Deep Learning**, click the app icon.
- MATLAB command prompt: Enter `nftool`.

Examples

Fit Data with a Shallow Neural Network

Algorithms

The **Neural Net Fitting** app provides built-in training algorithms that you can use to train your neural network.

Training Algorithm	Description
Levenberg-Marquardt	Update weight and bias values according to Levenberg-Marquardt optimization. Levenberg-Marquardt training is often the fastest training algorithm, although it does require more memory than other techniques. To implement this algorithm, the **Neural Net Fitting** app uses the `trainlm` function.
Bayesian regularization	Bayesian regularization updates the weight and bias values according to Levenberg-Marquardt optimization. It then minimizes a combination of squared errors and weights, and determines the correct combination so as to produce a network that generalizes well. This algorithm typically takes longer but is good at generalizing to noisy or small data sets. To implement this algorithm, the **Neural Net Fitting** app uses the `trainbr` function.
Scaled conjugate gradient backpropagation	Scaled conjugate gradient backpropagation updates weight and bias values according to the scaled conjugate gradient method. For large problems, scaled conjugate gradient is recommended as it uses gradient calculations which are more memory efficient than the Jacobian calculations used by Levenberg-Marquardt or Bayesian regularization. To implement this algorithm, the **Neural Net Fitting** app uses the `trainscg` function.

See Also

Apps

Neural Net Time Series | Neural Net Clustering | Neural Net Pattern Recognition

Functions

`fitnet` | `feedforwardnet` | `trainlm` | `trainscg` | `trainbr`

Figure 50 Neural Net Fitting description [20]

MATLAB add-on Neural Net Fitting

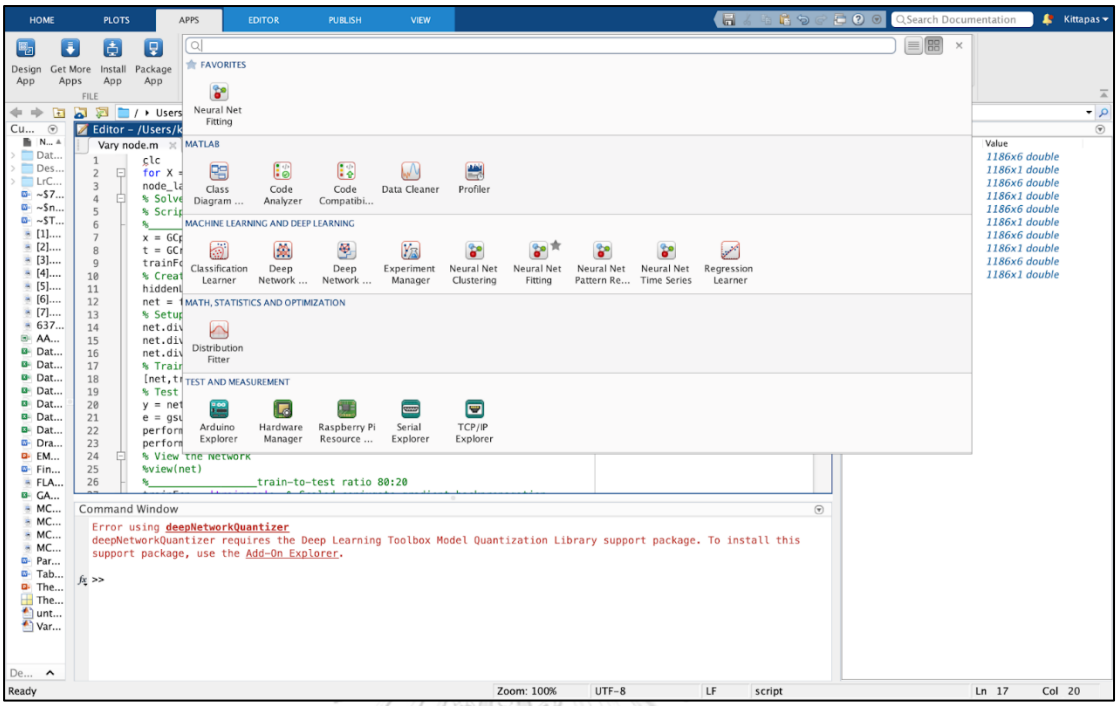


Figure 51 Neural Net Fitting application

Acting window of Neural Net Fitting “nftool”

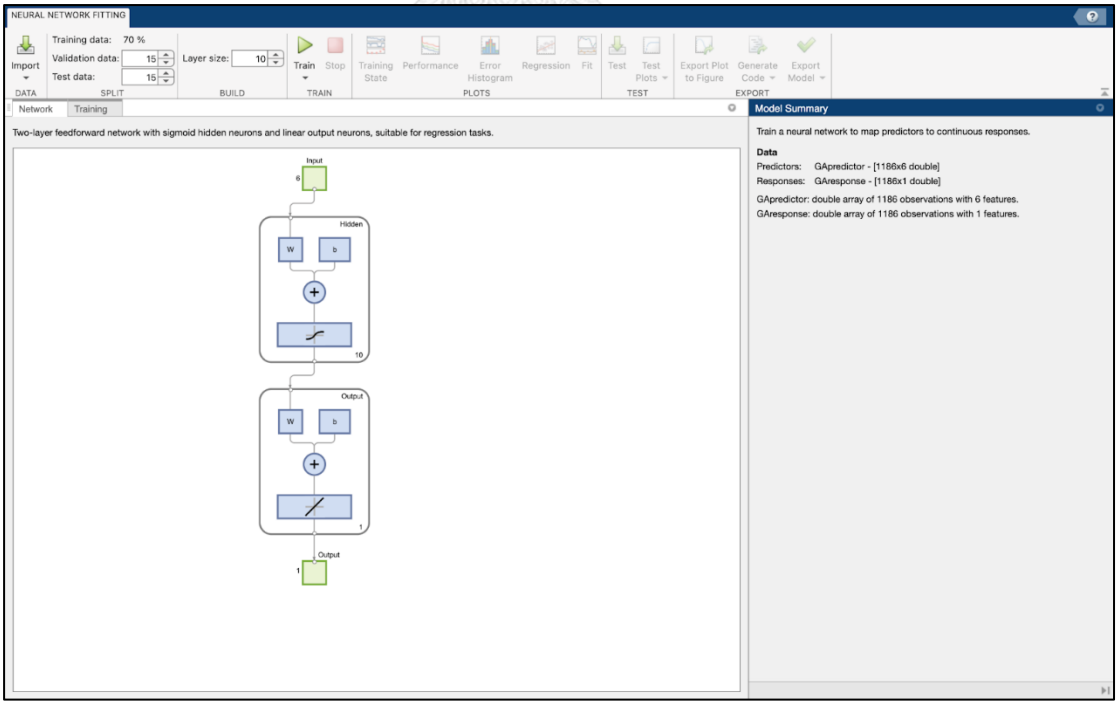


Figure 52 Neural Net Fitting network windown

Result window of Neural Net Fitting “nftool”

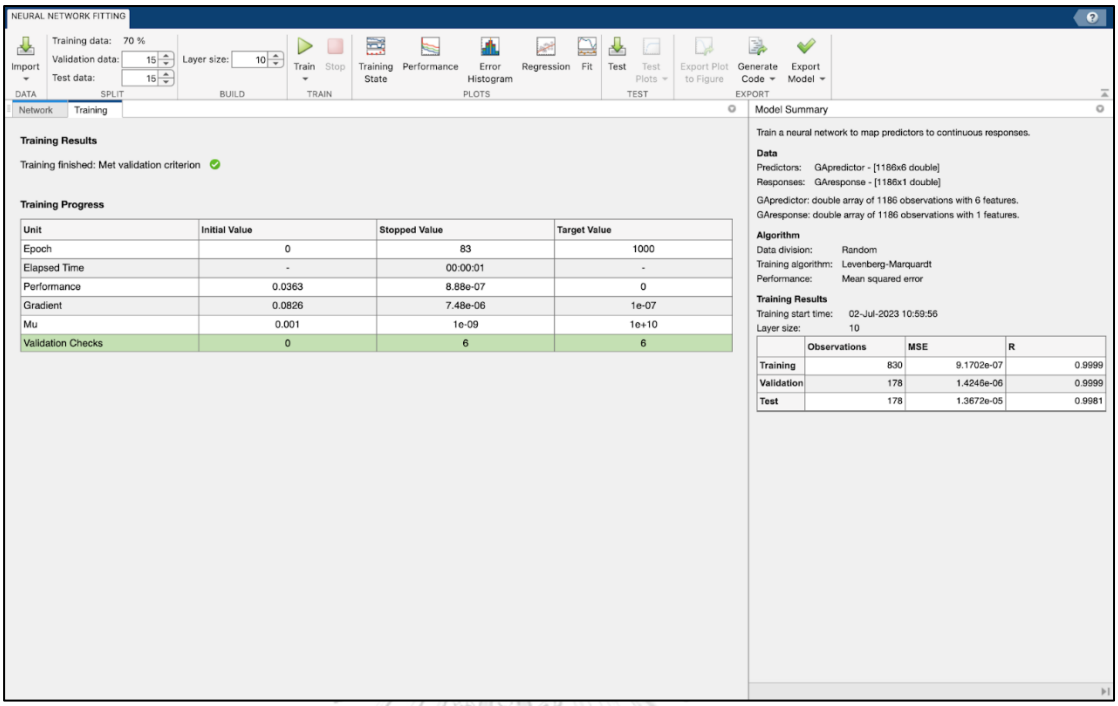


Figure 53 Neural Net Fitting result window

Dataset for ANN train and test (CSV file)

GITHUB repository clone link:

https://github.com/MACHINE11051/Thesis-ANN-Chlor-Alkali-Kittapas-Sukantowong.git

Abb.	Detail	Period	Estimate no.
RW	Original collected data	01 Mar 22 - 01 Mar 23	8,761 data
PS	Pre-processed, cut stop	01 Mar 22 - 01 Mar 23	< 8,761 data
PU	Pre-processed, cut unsteady	01 Mar 22 - 01 Mar 23	< 8,761 data
PBT	Pre-processed, before T/A	01 Mar 22 - 07 Nov 22	2 of 3 of PU
PAT	Pre-processed, after T/A	11 Nov 22 - 01 Mar 23	1 of 3 of PU
GA	Reducing bias, Group A (PBT)	01 Mar 22 - 24 Apr 22	1,186
GB	Reducing bias, Group B (PBT)	24 Apr 22 - 17 Jun 22	1,186
GC	Reducing bias, Group C (PBT)	17 Jun 22 - 10 Aug 22	1,186
GD	Reducing bias, Group D (PAT)	11 Nov 22 - 06 Jan 23	1,186
GE	Reducing bias, Group E (PAT)	06 Jan 23 - 28 Feb 23	1,186

Remark: AA in GITHUB refers to dataset PBT to train, validate and test ANN model.

Generated script for Neural Net Fitting with scale conjugate gradient algorithm

Dataset : GA

Hidden node : 10

Hidden layer : 1

Spitting : Train 70% Validate 15% Test 15%

```
% Solve an Input-Output Fitting problem with a Neural Network
% Script generated by Neural Fitting app
% This script assumes these variables are defined:
%   GApredictor - input data.
%   GAresponse - target data.
x = GApredictor';
t = GAresponse';
% Choose a Training Function
% For a list of all training functions type: help nntrain
% 'trainlm' is usually fastest.
% 'trainbr' takes longer but may be better for challenging problems.
% 'trainscg' uses less memory. Suitable in low memory situations.
trainFcn = 'trainscg';  % Scale Conjugate Gradient backpropagation.
% Create a Fitting Network
hiddenLayerSize = 10;
net = fitnet(hiddenLayerSize,trainFcn);
% Setup Division of Data for Training, Validation, Testing
net.divideParam.trainRatio = 70/100;
net.divideParam.valRatio = 15/100;
net.divideParam.testRatio = 15/100;
% Train the Network
[net,tr] = train(net,x,t);
% Test the Network
y = net(x);
e = gsubtract(t,y);
performance = perform(net,t,y)
```

```
% View the Network
view(net)
% Plots
% Uncomment these lines to enable various plots.
%figure, plotperform(tr)
%figure, plottrainstate(tr)
%figure, ploterrhist(e)
%figure, plotregression(t,y)
%figure, plotfit(net,x,t)
```

Generated script for Neural Net fitting with scale lavenberg marquardt

Dataset : GA

Hidden node : 10

Hidden layer : 1

Spitting : Train 70% Validate 15% Test 15%

```
% Solve an Input-Output Fitting problem with a Neural Network
% Script generated by Neural Fitting app
% This script assumes these variables are defined:
%   GApredictor - input data.
%   GAresponse - target data.
x = GApredictor';
t = GAresponse';
% Choose a Training Function
% For a list of all training functions type: help nntrain
% 'trainlm' is usually fastest.
% 'trainbr' takes longer but may be better for challenging problems.
% 'trainscg' uses less memory. Suitable in low memory situations.
trainFcn = 'trainlm';  % Levenberg-Marquardt backpropagation.
% Create a Fitting Network
hiddenLayerSize = 10;
net = fitnet(hiddenLayerSize,trainFcn);
% Setup Division of Data for Training, Validation, Testing
net.divideParam.trainRatio = 70/100;
net.divideParam.valRatio = 15/100;
net.divideParam.testRatio = 15/100;
% Train the Network
[net,tr] = train(net,x,t);
% Test the Network
y = net(x);
e = gsubtract(t,y);
performance = perform(net,t,y)
```

```
% View the Network
view(net)
% Plots
% Uncomment these lines to enable various plots.
%figure, plotperform(tr)
%figure, plottrainstate(tr)
%figure, ploterrhist(e)
%figure, plotregression(t,y)
%figure, plotfit(net,x,t)
```

Generated script for Neural Net fitting with scale Bayesian regularization

Dataset : GA

Hidden node : 10

Hidden layer : 1

Spitting : Train 70% Validate 15% Test 15%

```
% Solve an Input-Output Fitting problem with a Neural Network
% Script generated by Neural Fitting app
% This script assumes these variables are defined:
%   GApredictor - input data.
%   GAresponse - target data.
x = GApredictor';
t = GAresponse';
% Choose a Training Function
% For a list of all training functions type: help nntrain
% 'trainlm' is usually fastest.
% 'trainbr' takes longer but may be better for challenging problems.
% 'trainscg' uses less memory. Suitable in low memory situations.
trainFcn = 'trainbr';  % Bayesian Regualrization backpropagation.
% Create a Fitting Network
hiddenLayerSize = 10;
net = fitnet(hiddenLayerSize,trainFcn);
% Setup Division of Data for Training, Validation, Testing
net.divideParam.trainRatio = 70/100;
net.divideParam.valRatio = 15/100;
net.divideParam.testRatio = 15/100;
% Train the Network
[net,tr] = train(net,x,t);
% Test the Network
y = net(x);
e = gsubtract(t,y);
performance = perform(net,t,y)
```

```
% View the Network
view(net)
% Plots
% Uncomment these lines to enable various plots.
%figure, plotperform(tr)
%figure, plottrainstate(tr)
%figure, ploterrhist(e)
%figure, plotregression(t,y)
%figure, plotfit(net,x,t)
```

Script for testing optimum number of hidden node

Dataset : GA

Hidden node : 5 - 10

Hidden layer : 1

Spitting : Train 70% Validate 15% Test 15%

: Train 80% Validate 10% Test 10%

: Train 90% Validate 5% Test 5%

```
for X = 5:10              %run for loop 1 - 5
node_layer = [X]          %hidden node under for loop 1 - 5
%_____________________train-to-test ratio 90:10
x = GCpredictor';         %input data
t = GCresponse';          %traget data
trainFcn = 'trainscg';  % Scaled conjugate gradient backpropagation.
% Create a Fitting Network
hiddenLayerSize = node_layer;
net = fitnet(hiddenLayerSize,trainFcn);
% Setup Division of Data for Training, Validation, Testing
net.divideParam.trainRatio = 90/100;
net.divideParam.valRatio = 5/100;
net.divideParam.testRatio = 5/100;
% Train the Network
[net,tr] = train(net,x,t);
% Test the Network
y = net(x);
e = gsubtract(t,y);
performance_MSE_90_10 = perform(net,t,y);
performance_RMSE_90_10 = sqrt(performance_MSE_90_10);
%_____________________train-to-test ratio 80:20
trainFcn = 'trainscg';  % Scaled conjugate gradient backpropagation.
% Create a Fitting Network
hiddenLayerSize = node_layer;
```

```
net = fitnet(hiddenLayerSize,trainFcn);
% Setup Division of Data for Training, Validation, Testing
net.divideParam.trainRatio = 80/100;
net.divideParam.valRatio = 10/100;
net.divideParam.testRatio = 10/100;
% Train the Network
[net,tr] = train(net,x,t);
% Test the Network
y = net(x);
e = gsubtract(t,y);
performance_MSE_80_20 = perform(net,t,y);
%_____________________train-to-test ratio 70:30
trainFcn = 'trainscg';  % Scaled conjugate gradient backpropagation.
% Create a Fitting Network
hiddenLayerSize = node_layer;
net = fitnet(hiddenLayerSize,trainFcn);
% Setup Division of Data for Training, Validation, Testing
net.divideParam.trainRatio = 70/100;
net.divideParam.valRatio = 15/100;
net.divideParam.testRatio = 15/100;
% Train the Network
[net,tr] = train(net,x,t);
% Test the Network
y = net(x);
e = gsubtract(t,y);
performance_MSE_70_30 = perform(net,t,y);
performance_RMSE_70_30 = sqrt(performance_MSE_70_30);
%summary table
result_GA = [performance_RMSE_70_30 performance_RMSE_80_20
performance_RMSE_90_10]
end
```

Script for testing optimum number of hidden layer

Dataset : GA

Hidden node : 6

Hidden layer : 1 - 5

Spitting : Train 70% Validate 15% Test 15%

: Train 80% Validate 10% Test 10%

: Train 90% Validate 5% Test 5%

```
for X = 1:5                              %run for loop hidden layer 1 - 5
   switch X
   case 1  node_layer = [6]              %1 hidden layer
   case 2  node_layer = [6 6]            %2 hidden layers
   case 3  node_layer = [6 6 6]          %3 hidden layers
   case 4  node_layer = [6 6 6 6]        %4 hidden layers
   case 5  node_layer = [6 6 6 6 6]      %5 hidden layers
   end
%____________________train-to-test ratio 90:10
x = GEpredictor';        %input data
t = GEresponse';         %traget data
trainFcn = 'trainscg';  % Scaled conjugate gradient backpropagation.
% Create a Fitting Network
hiddenLayerSize = node_layer;
net = fitnet(hiddenLayerSize,trainFcn);
% Setup Division of Data for Training, Validation, Testing
net.divideParam.trainRatio = 90/100;
net.divideParam.valRatio = 5/100;
net.divideParam.testRatio = 5/100;
% Train the Network
[net,tr] = train(net,x,t);
% Test the Network
y = net(x);
e = gsubtract(t,y);
```

```
performance_MSE_90_10 = perform(net,t,y);
performance_RMSE_90_10 = sqrt(performance_MSE_90_10);
%_____________________train-to-test ratio 80:20
trainFcn = 'trainscg';  % Scaled conjugate gradient backpropagation.
% Create a Fitting Network
hiddenLayerSize = node_layer;
net = fitnet(hiddenLayerSize,trainFcn);
% Setup Division of Data for Training, Validation, Testing
net.divideParam.trainRatio = 80/100;
net.divideParam.valRatio = 10/100;
net.divideParam.testRatio = 10/100;
% Train the Network
[net,tr] = train(net,x,t);
% Test the Network
y = net(x);
e = gsubtract(t,y);
performance_MSE_80_20 = perform(net,t,y);
performance_RMSE_80_20 = sqrt(performance_MSE_80_20);
%_____________________train-to-test ratio 70:30
trainFcn = 'trainscg';  % Scaled conjugate gradient backpropagation.
% Create a Fitting Network
hiddenLayerSize = node_layer;
net = fitnet(hiddenLayerSize,trainFcn);
% Setup Division of Data for Training, Validation, Testing
net.divideParam.trainRatio = 70/100;
net.divideParam.valRatio = 15/100;
net.divideParam.testRatio = 15/100;
% Train the Network
[net,tr] = train(net,x,t);
% Test the Network
y = net(x);
```

```
e = gsubtract(t,y);
performance_MSE_70_30 = perform(net,t,y);
performance_RMSE_70_30 = sqrt(performance_MSE_70_30);
%summary table
result_GA = [performance_RMSE_70_30 performance_RMSE_80_20
performance_RMSE_90_10]
end
```

REFERENCES

1. Ito, H. and A. Manabe, *Chapter 8 - Chlor–alkali electrolysis*, in *Electrochemical Power Sources: Fundamentals, Systems, and Applications*, T. Smolinka and J. Garche, Editors. 2022, Elsevier. p. 281-304.
2. Research, E.M., *Global Chlor-Alkali Market Outlook*. 2022: https://www.expertmarketresearch.com/reports/chlor-alkali-market.
3. Chlor, E., *THE CHLORINE TREE - DOZENS OF USEFUL PRODUCTS IN WHICH CHLORINE CHEMISTRY PLAYS A KEY ROLE*, Products_of_the_Chlorine_Tree, Editor. 2019, Euro Chlor: https://www.eurochlor.org/publication/the-chlorine-tree-dozens-of-useful-products-in-which-chlorine-chemistry-plays-a-key-role/.
4. NUCERA, T., *Chlor-alkali solutions*, B.i.-e.m.p.e. (BiTAC), Editor. 2023, Thyssenkrupp NUCERA.
5. Lenntech, *Membrane cell process for chlor alkali production*, Chlorine-2, Editor. 2023, Lenntech.
6. Corporation, M.C., *DIAION™ CR11 Chemical Structure and Chelation Reaction*, 1189188_001, Editor. 2023, Mitsubishi Chemical Corporation.
7. Thomas F. O'Brien , T.V.B., Fumio Hine, *Handbook of Chlor-Alkali Technology*. Volume I: Fundamentals, ed. Chloralkali_membrane.svg. Vol. 1. 2005: Springer New York, NY.
8. Millet, P., *9 - Chlor-alkali technology: fundamentals, processes and materials for diaphragms and membranes*, in *Handbook of Membrane Reactors*, A. Basile, Editor. 2013, Woodhead Publishing. p. 384-415.
9. Mirza, N., *Minimizing Trihalomethane Formation through Source Water Monitoring and Optimizing Treatment Practices Monitoring and Optimizing Treatment Practices*. 2019.
10. Kaveh, N.S., S.N. Ashrafizadeh, and F. Mohammadi, *Development of an artificial neural network model for prediction of cell voltage and current efficiency in a chlor-alkali membrane cell.* Chemical Engineering Research and Design, 2008. **86**(5): p. 461-472.
11. algotrading101, *Train/Test Split and Cross Validation – A Python Tutorial*, training-validation-test-data-set, Editor. 2020, algotrading101.
12. More, V.D., *Machine learning Model Validation techniques*, emlm_0302-6a388b903f6e1e04c95e718940eff039, Editor. 2018, wordpress.
13. yngie-c.github, *Overfitting & Underfitting*, 5033807687188480, Editor. 2020, yngie-c.github.
14. Karagiannakos, S., *Regularization techniques for training deep neural networks*, early-stopping, Editor. 2021, AI SUMMER.
15. sefidian. *A guide on regression error metrics (MSE, RMSE, MAE, MAPE, sMAPE, MPE) with Python code*. 2022 [cited 2023 18/7/2023]; Available from: https://sefidian.com/2022/08/18/a-guide-on-regression-error-metrics-with-python-code/.
16. Jalali, A.A., F. Mohammadi, and S.N. Ashrafizadeh, *Effects of process conditions on cell voltage, current efficiency and voltage balance of a chlor-alkali membrane cell.* Desalination, 2009. **237**(1): p. 126-139.
17. Mirzazadeh, T., et al., *Optimization of caustic current efficiency in a zero-gap advanced chlor-alkali cell with application of genetic algorithm assisted by*

artificial neural networks. Chemical Engineering Journal, 2008. **140**(1): p. 157-164.
18. Shojai Kaveh, N., F. Mohammadi, and S.N. Ashrafizadeh, *Prediction of cell voltage and current efficiency in a lab scale chlor-alkali membrane cell based on support vector machines.* Chemical Engineering Journal, 2009. **147**(2): p. 161-172.
19. YOKAGAWA, *Plant Information Management System (Exaquantum)*, E. Overview, Editor. 2022, YOKAGAWA.
20. MathWorks, *Neural Net Fitting*, N.N. Fitting, Editor. 2023, MathWorks.

VITA

NAME	Kittapas Sukantowong
DATE OF BIRTH	18 October 1996
PLACE OF BIRTH	Bangkok, Thailand
INSTITUTIONS ATTENDED	Department of Chemical Engineering, Faculty of Engineering, Chulalongkorn University
HOME ADDRESS	11 Soi Suksawat 26 Yak 2-1, Bangpakok, Ratburana, Bangkok, Thailand 10140
PUBLICATION	The 32nd Thai Institute of Chemical Engineering and Applied Chemistry Conference