NUEJ

NARESUAN UNIVERSITY ENGINEERING JOURNAL

July-December 2022 Vol.17, No.2
ISSN : 2651-1568

NUEJ
Naresuan University Engineering Journal

Editorial Team

NUEJ
Naresuan University
Engineering Journal

Editorial Board

NUEJ
Naresuan University
Engineering Journal

Editorial Board

NUEJ
Naresuan University
Engineering Journal

Research Articles

NUEJ
Naresuan University
Engineering Journal

Aims and Objectives

The primary objective of the *Naresuan University Engineering Journal (NUEJ)* is to publish high quality research articles presenting contemporary developments in theory, design, and applications in all areas of Engineering, Science and Technology, including research in Civil and Environmental Engineering, Mechanical Engineering, Electrical and Computer Engineering, Industrial, Chemical, and Material Engineering. NUEJ covers all multidisciplinary research in associated areas**,** such as Mechatronics, Energy, Industrial and Engineering Design, Manufacturing Technology, Engineering Management and Medical Engineering.

Journal Policies

Naresuan University Engineering Journal (NUEJ) is a peer reviewed journal, regularly published with 2 issues per year (January – June, and July – December). Submissions must be original, unpublished works, and not currently under review by other journals. NUEJ will consider only submitted works which respect research ethics, including confidentiality, consent, and the special requirements for human and animal research. All research articles dealing with human or animal subjects must attach an approval certificate from the appropriate Ethics Committee. Additionally, the research articles dealing with human subjects must provide evidence of informed consent.

Editorial board of NUEJ reserves the right to decide whether the submitted manuscript should be accepted for publication. The final decision of the editorial board cannot be appealed.

The submitted manuscript has to be written in English only, and can be in Microsoft Word (doc or docx) or PDF file format. The corresponding author is required to register and submit the manuscript at https://ph01.tci-thaijo.org/index.php/nuej

Editorial board of Naresuan University Engineering Journal

Faculty of Engineering, Naresuan University,

Phitsanulok, 65000, Thailand

Tel. +66 (0) 55 964092

Fax +66 (0) 55 964000

Email: nuej@nu.ac.th

NUEJ
Naresuan University
Engineering Journal

Editor's Note

Naresuan University Engineering Journal (NUEJ) provides a resource and platform for sharing new knowledge and information in all areas of engineering. The journal has the following procedures. The journal receives articles, then reviews the articles using a strict peer review process to ensure the publication maintains high standards. Articles are accepted for publication only if at least two out of three independent reviewers provide approval, followed by Editors permission; the Editor-in-Chief makes the final decision in any case. NUEJ is approved to be in the first tier of the Thai-Journal Citation Index (TCI 1) and the Asean Citation Index (ACI).

As part of the journal's efforts to continue to grow and improve, and aligned with the goal of becoming a worldwide reference, all manuscripts must be submitted in English. This issue, Vol. 17 No. 2 (2022), is the second English publication and there are eight fascinating engineering articles, three of which (the 4th, 5th, and 8th) are extensions of work first presented at the 5tn International Conference on Water Resources Engineering (ICWRE 5, 2021). As always, sincere appreciation is extended to the authors and reviewers for their contributions.

Assoc. Prof. Dr. Panus Nattharith

Editor

Naresuan University Engineering Journal

NUEJ
Naresuan University
Engineering Journal

Electricity Theft Detection in Electrical Distribution System Using Long Short-Term Memory

Chintana Xayalath[1], Sutthichai Premrudeepreechacharn [2]* and Kanchit Ngamsanroaj[3]

[1] Department of Mechanical, Faculty of Engineering, Chiang Mai University, Chiang Mai 50200, Thailand
[2] Department of Electrical Engineering, Faculty of Engineering, Chiangmai University, Chiang Mai 50200, Thailand
[3] Electricity Generating Authority of Thailand (EGAT), Nonthaburi, Thailand

* Corresponding author e-mail: suttic@eng.cmu.ac.th

(Received: 26 November 2021, Revised: 25 August 2022, Accepted: 05 September 2022)

Abstract

This paper presents the method of detection of power theft in the electrical distribution system by using the real data from industry consumers that use meter AMR. The real data were feature extracted by labeling event types by the number demonstrated to normal, voltage theft, and current theft. Then, the data were fed to Long Short-Term Memory (LSTM) for the created model by training and testing. The accuracy results were shown the model can be classified accurate archive to 99%.

Keywords: Non-technical loss, Power theft, Automatic Meter Reading, Long Short-Term Memory classification.

1. INTRODUCTION

Electricite Du Laos (EDL) is a state own enterprise responsible for distribution and transmission systems in the country. The municipal is one of the branches in Vientiane and has a large distribution system with a total of 164,126 customers. The Automatic Meter Reading (AMR) is installed in 135 customers as reported by (EDL, 2020). This area is a centralized country that has power use variety so that power consumption has large.

Along with power distribution, power losses have appeared in the power system. It is contained technical and non-technical losses. Which indicated to health index of the electrical distribution system and related revenue of the Organization.

Table 1 Summarize power consumption and power loss in Municipal Vientiane

No	Description	Energy (kWh)	Percentage
1	Energy demand sent out substation	1,130,046,682	100%
2	Energy from the billing system	981,961,993	86.9%
3	Technical loss	36,781,606	3.25%
4	Non- Technical loss	111,304,083	9.85%

As shown in Table 1, has demonstrated the ratio of non-technical losses (NTL) is coverage of the power sent from the substation more than technical losses. it is happened from power theft, meter fault, human error, unregistered meter, meter tampering, etc., as reported in the (CIRED, 2017) report.

The major problem of EDL is power theft which occurred by a fraudulent customer by going to adjust the meter. It made the power consumption in the billed less than actual. The most issue has appeared in commercial customer as shown in Table 2.

Table 2 Summarize non-technical loss in Municipal Vientiane

No.	Description	Non–technical losses (kWh)	Percentage
1	Household	988,215	0.88%
2	Commercial	109,625,838	98.49%
3	Other	690,030	0.61%

In previous years ago many methods related to power theft have been conducted. The traditional method of detecting power theft inspects the abnormal power consumption in the electrical bill and upgraded the measuring equipment in power distribution, such as the smart meter. Monitoring the power usage of customers in the smart grid, when using the smart meter for a long time have very large data in the data analytics (Bula et al., 2016). Due to the large amount of historical data being stored in the database. The data on power consumption is applied to the algorithms of machine learning. To detect NTL, Principal Component Analysis (PCA) can help to improve support vector machine performance and predictive accuracy (Toma et al., 2019). (Glauner et al., 2016) used the neighborhoods feature and master data of customers from successive consumption to apply logistic regression, k-nearest neighbors, linear support vector machine, and random forest to apply with logistic regression, k-nearest neighbors, linear support vector machine, and random forest. The result was compared to the accurate detection of NTL. (Micheli et al., 2019)

investigated NTL detection using a variant of Peer-to-Peer (P2P) computing to detect fraud in the case of unreliable smart meters. This research applied data collected from smart meters and data in the same neighborhood area for verification by NTL. This data has been implemented in the application of multi-linear regression. (Nabil et al., 2018) presented a Gradient Boosting Theft Detector (GBTD) based on several Gradient Boosting classifiers to improve the detection rate and False Positive Rate (FPR) of Gradient Boosting classifiers (GBCs). This research focuses on feature engineering-based improvements to the classifiers' ability to improve. Furthermore, a stochastic feature was used to generate the standard deviation, mean, minimum, and maximum values of daily electricity usage. There are 6 theft cases sampled from real-world theft patterns and applied with the dataset for validation of the proposed algorithm. (Buzau et al., 2019) used a hybrid deep neural network by combining the architecture of a long-short term memory with a multi-layer perceptron. It used feature detection anomalies and fraud in smart meters to train and test the algorithm. The performance classifiers of hybrid neural networks significantly outperform the state-of-the-art. (Long et al., 2020) used the method of data-driven combined algorithm to systematically identify anomalies of power loss in the distribution network. It consisted of abnormal feeder, time, and position detection.

In addition, the research paper by (Ghori et al., 2019) used real datasets from the company that supplies power in Pakistan. The study evaluated 15 algorithm classifiers of ML. It included the recent algorithm developed to compare the performance classifier in non-technical loss detection and the author, (Veerasamy et al., 2021) used a model recurrent neural network based on LSTM to classifier high impedance fault detection in solar PV.

This research aims to create a model to detect power theft in power distribution systems. The real data of case power theft has features extracted and applied with LSTM algorithms to establish a model.

This paper is organized as follows. The introduction of non-technical losses in the electrical distribution system was described in Section 1. Then, the power theft in AMR is described in Section 2. After that, the methodology for the detection of power theft is explained in Section 3. The simulation results are discussed in Section 4. Finally, the conclusion is described in Section

2. POWER THEFT IN AUTOMATIC METER READING (AMR)

2.1 Structure AMR system

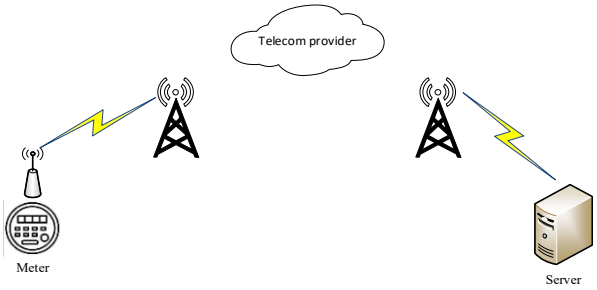


Figure 1 Automatic Meter Reading system.

The Automatic Meter Reading (AMR) is a similar meter measurement to the power parameter. It can be read, record values, and send the data to a server. The values include voltage, current, power consumption, energy import/export, etc. and the communication between meter and server uses the GPRS module. The data granularity is recorded every 15 minutes and the architecture system is shown in Figure 1.

2.2 Method AMR stealing electricity

The commercial customer is the customer category that uses power at high load capacity. While the customers consume large amounts of electricity, the electricity meter cannot support the power directly. It should have the equipment to scale the voltage and current appropriate with meter measurement range such as potential transformer (VT) and current transformer (CT) and has the meter diagram connection as shown in Figures 2 and 3.

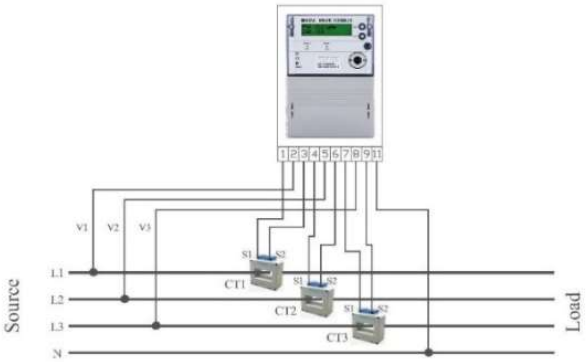


Figure 2 Single line diagram of meter 0.4 kV in EDL's standard.

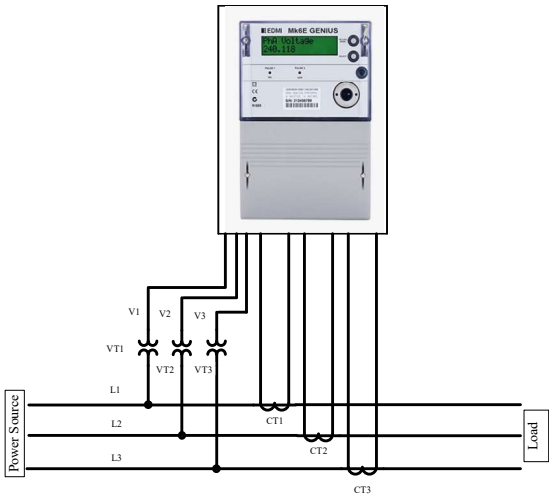


Figure 3 Single line diagram of meter 22kV in EDL's standard.

As Figures 2 and 3 are demonstrated the meter is connected with the power supplied. It has consisted of 2 different types connecting in the power distribution system. It also uses these types of meters in power distribution at low voltage and medium voltage. For the low voltage level meter, it should be used the current transformer (CT) to scale the range of current for measuring the electricity as shown in Figure 2. In the same way, in Figure 3 they also apply CT and installed voltage transformer (VT) to decrease voltage level from medium voltage level to applied at the range of meter requirement.

The power thefts on AMR all most is occurred from disturbing the measuring wire of voltage and current connect to the meter on site by cutting the measurement cable voltage and current 1 or 2 phases, short circuit the secondary winding of CT each other to make the measuring value is less than actual power utility.

3. METHODOLOGY

3.1 Long short-term memory

Long Short-Term Memory (LSTM) is a specialized Recurrent Neural Network. Which has architecture shown in Figure 4. It has the capability of learning long-term dependencies. Traditional RNNs have the problem of vanishing and exploding gradients. In vanishing gradients, weights in the early layers of the network get updated with very low values and, hence, train very slowly. This is because weight updates are proportional to the gradient of the error function, and in the initial layers, during backpropagation, gradients multiply with small activation values repeatedly, which further leads to even smaller values. Similarly, in exploding gradients, the gradients become too large and update the weight parameters by a large value and hence become harder to converge to an optimum value.

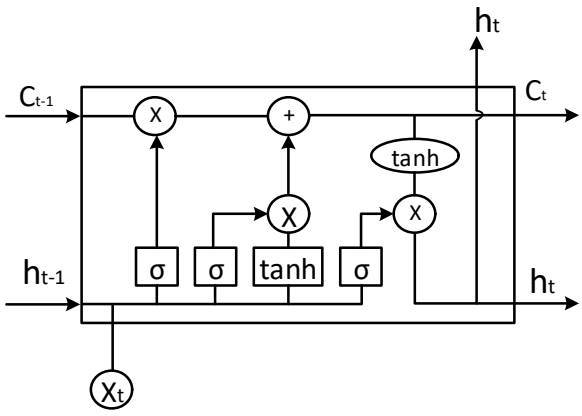


Figure 4 Long Short-Term Memory Architecture

The LSTM has three gates: the input gate i_t, the forget gate f_t and the output gate O_t. The forget gate takes the previously hidden state information (h_{t-1}) and current input x_t through a pointwise multiplication operation and decides what is memorized or forgotten the information from the cell state. This gate uses the sigmoid activation function to approve the predicts an output of either 0 or 1. A value of 1 shows that the relevant information should be kept in the cell state, while 0 represents irrelevant information, which is discarded from the cell state. The forget gate, input gate, and output gate are described in Equations (1) – (5).

$$f_t = \sigma\left(W_f[\, h_{t-1}, x_t \,] + b_f\right) \quad (1)$$

where W_f represents the weight and b_f is the bias of the forget gate f_t. The σ is applied as the activation function on the forget gate.

The input gate decides what information is going to be stored in the cell state information C_t^i of input gate. It takes the input x_t and previous hidden state h_{t-1} and applies and tanh activation functions through a pointwise multiplication operation as follows:

$$i_t = \sigma\left(W_i[\, h_{t-1}, x_t \,] + b_i\right) \quad (2)$$
$$C_t^i = \acute{C}_t * i_t \quad (3)$$
$$\acute{C}_t = \tanh\left(W_C[\, h_{t-1}, x_t \,] + b_C\right) \quad (4)$$

where W_i and W_C represent the weights of the input gate i_t and new cell state, respectively. The b_i and b_C are the biases of the network, and $\acute{C}_t$ is the previously hidden cell state information. To update the information on the current cell state C_t, Equations (1) multiply by the previous cell state C_{t-1} are summed with The Equations (2) multiplied by Equations (3):

$$C_t = f_t * C_{t-1} + i_t * \acute{C}_t \quad (5)$$

Finally, in Equation (5), the output gate is determined. The output gate takes the current input x_t and previously hidden state h_{t-1} with the implication of the activation

function σ. The b_o is added as a bias to the output network.

$$O_t=\sigma\,(W_o[\,h_{t-1}\,,\,x_t\,]+b_o\,) \quad (6)$$

The updated output gate O_t and the information from the cell state C_t are used to perform the pointwise multiplication operation to get the next hidden state h_t, given by Equation (7):

$$h_t= O_t * \tanh\,(C_t) \quad (7)$$

The input of model LSTM can be applied with the data of research x_t is the input at time step data and a vector of 6 points that corresponds to the value of voltage and current and O_t is the output of the LSTM classifier to event-type (normal/theft by voltage/theft by current).

3.2 Data Collection

The simulation of this study use data from the meter AMR low voltage shown in Figure 2. which load profile dataset has been collected from the AMR database system at EDL. The data record during (01/03/2014 - 30/12/2019). it had been recorded every 15 minutes. The data consist of the value of voltage phase 1, 2, 3, current phase 1, 2, and 3 to show attribute behavior power usage of customers and have 28,382,400 data set.

3.3 Feature Extraction of Power Theft

The study used data from Vientiane municipal customers to create a feature extract by selecting Voltage L1(V_1), Voltage L2 (V_2), Voltage L3 (V_3), and Current L1 (I_1), Current L2 (I_2), and Current L3 (I_3) to demonstrate a characteristic pattern to describe power theft behaviour and the type of power theft below:

3.3.1 Theft Detection from Voltage Measurement (V_{theft})

This is present V_{theft} detection, the research applies the parameter of voltage recorded the database of meter AMR. This problem rarely appears in the pattern voltage value such as meter failure, voltage drop, external disturbance, power outage, etc. So, this study implements the voltage parameter to detect the abnormality caused by the fraud meter. according to Figure 5, there are 6 parameters including 3-phase voltage and 3-phase currents. From observation the pattern characteristics of the data which occur 2 periods time abnormal from 20 - 35 and 44 - 55, it causes of power theft on voltage phase 1.

As more as Vtheft 1 phase is directly impacted to NTL, Moreover the stealing power can be done more than 1 phase. So, Figure 6 demonstrates the feature of Vtheft 2 phases. It is affected by the power consumption record of the meter. This leads to an increase in power losses of NTL.

3.3.2 Theft Detection from Current Measurement (I_{theft})

The current is a variable to illustrate the attribute of load. It has the feature very complex for identifiers of the event that occurred in the data record. That the problem uses the current 3 phases is based on the detection of power theft. In addition, the study uses the pattern of voltage to help approved the event. Both parameters are the relation shown in Figures 7 and 8.

As Figure 7 can be found that the profile of voltage and current in periods time 1 – 16 demonstrates that voltage 3 phases are normal and current 3 phases have the value close up zero. Enough periods time 21 the has power use the current 3 phases gradually change go along with behaviors of load. But the pattern of phase 3 is more different between the 2 phases in periods time 17 – 21. that may illustrate the abnormal load.

In addition, Itheft1 phase is directly impacted by NTL, Furthermore, the stealing power can be done in more than 1 phase. So, Figure 8 demonstrates the feature of Itheft 2 phases. It is affected by the power consumption record of the meter. This leads to an increase in power losses of NTL.

From detection voltage theft and current theft to observations. The data of parameters is very difficult if using human inspection, the abnormal of meter as well as power theft. It is greater than when we have the tool to help inspection by the application of Artificial Intelligence. which uses data from above to help classifier issue that problem for more efficiency in detection.

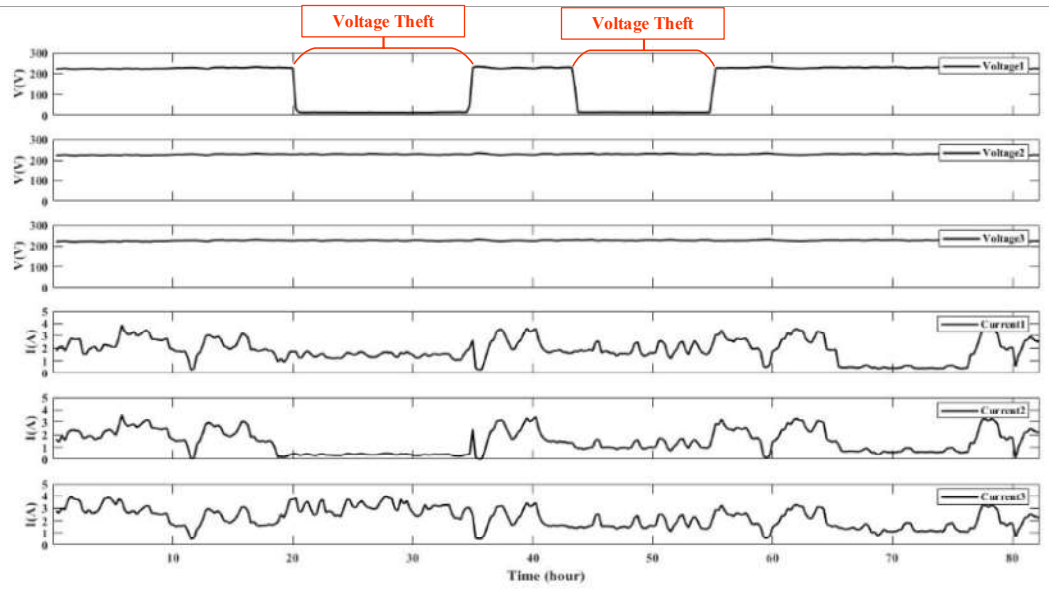


Figure 5 Voltage and current characteristics of power theft by voltage phase 1

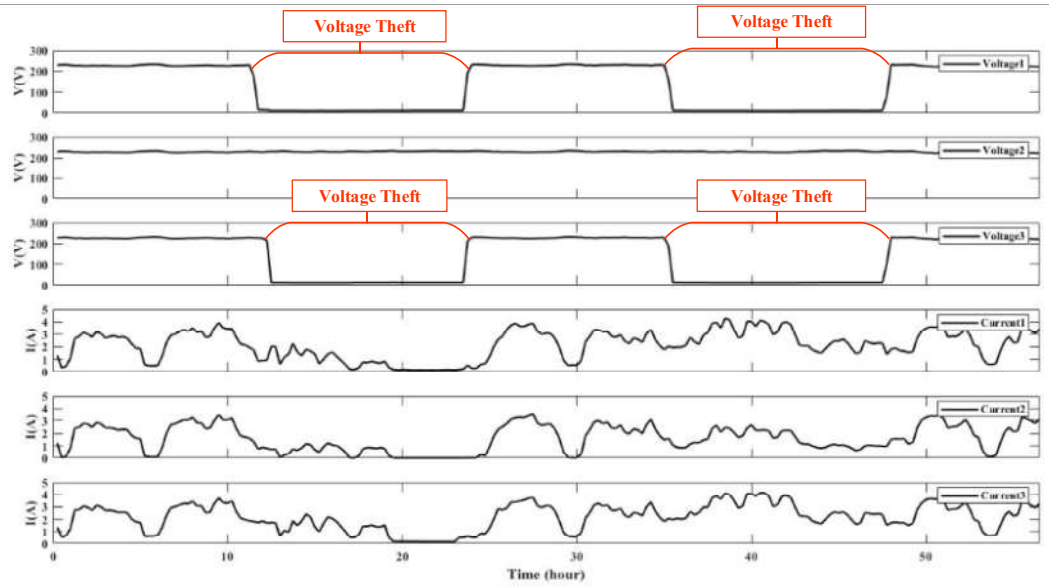


Figure 6 Voltage and current characteristics of power theft voltage phase 1 and 3

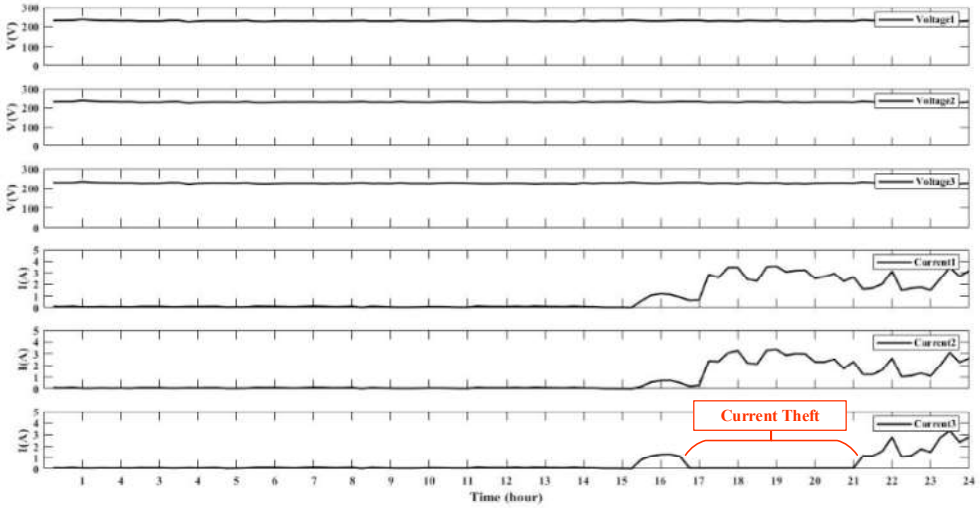


Figure 7 Voltage and current characteristics of power theft from current phase 3

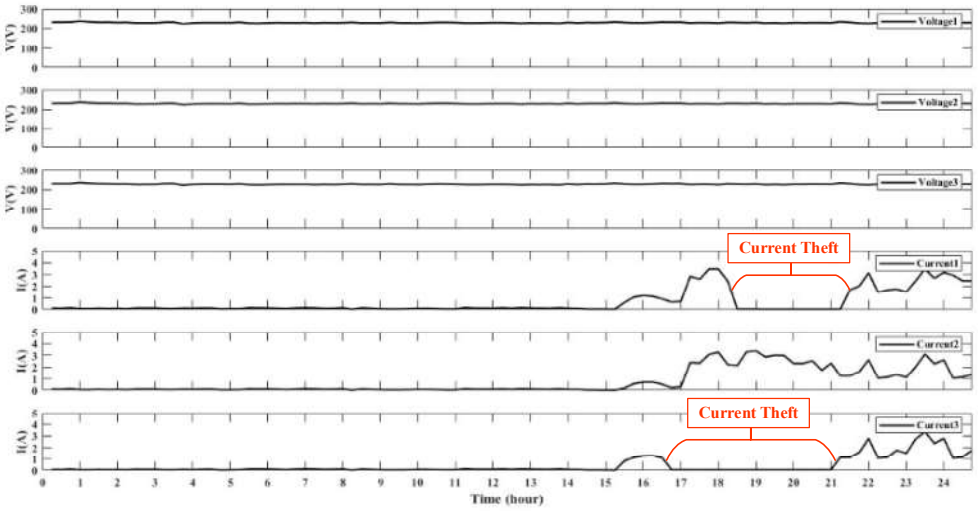


Figure 8 Voltage and current characteristics of power theft from current phase 1 and 3

3.4 Data Preprocessing

This study is using the real data of power theft from the AMR to feature extract. The dataset has included Voltage L1 (V1), Voltage L2 (V2), Voltage L3 (V3), Current L1 (I1), Current L2 (I2), and Current L3 (I3) to define the event type each time step of data. The event typically consists of normal, Vtheft, and Itheft, represented by numbers 1, 2, and 3, respectively. as shown in Table 3. The total dataset feature extracted contained the attribute of voltage and current to demonstrate normal, Vtheft, and Itheft in the numbers 28.237.278, 9,277.605, 52.791.

3.5 Experimental

The detection method of power theft is manipulating the feature extracted data to be applied with an LSTM algorithm which has a process as shown in Figure 9.

In model development to detect power theft. The study is using the feature extracted data in Table 3 to apply the LSTM algorithm. by defining the parameters of V1, V2, V3, I1, I2, I3 as input and event type output of the model. Then, the data has divided between 80% and 20% following the number of each class for the train and test model. The data were fed to LSTM and computed to the accuracy when receipts high accurate. the model will be utilized the next time.

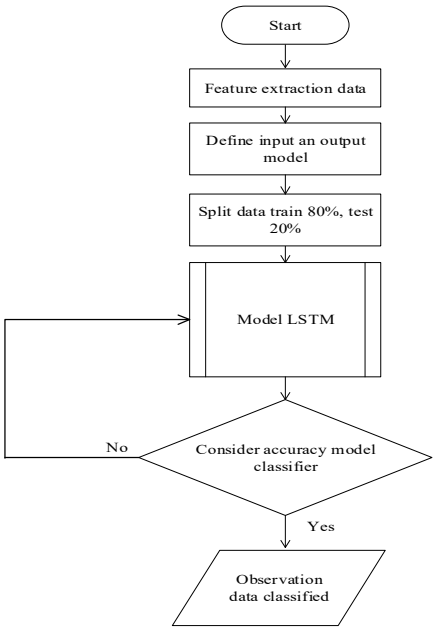


Figure 9 Flowchart model of LSTM classifier

Table 3 Data feature extracted

V1	V2	V3	I1	I2	I3	Event type
9	224.2	11	1.468	0.95	1.497	2
10.6	224.2	11.9	1.665	0.946	1.663	2
12	224.2	12.7	1.31	0.568	1.24	2
11.9	222.8	12.6	1.388	0.868	1.388	2
11.7	220.2	12.4	1.641	1.129	1.467	2
208.1	222.3	206.1	1.042	0.572	1.031	1
223.2	223.6	222.5	0.764	0.315	0.657	1
223.6	224	223	0.579	0.168	0.557	1
224.6	224.5	223.9	0.453	0.026	0.293	1
230.7	230.6	229	0.09	0.079	0.089	1
228.3	227.4	226.6	0.086	0.079	0.086	1
230.4	229.5	228.7	0.042	0.033	0.05	1
230	229.3	228.4	0.038	0.024	0.045	1
230.7	229.8	229.1	0.035	0.021	0.041	1
233.6	232.8	232	0.025	0.007	0.03	1
230.6	229.3	228.8	0.587	0.19	0.827	1
228.7	227.7	227.2	1.057	0.569	1.14	1
229.1	228.2	227.5	1.185	0.705	1.27	1
230.7	229.7	229	1.121	0.715	1.293	1
232.8	231.8	230.8	0.902	0.523	1.107	1
232	231.2	230.2	0.631	0.211	0.211	1
232.4	231.1	230.6	0.684	0.299	0.670	1
228.2	227	226.3	2.839	2.393	0.067	3
229.8	228.6	227.9	2.588	2.343	0.067	3
229.5	228.3	227.8	3.457	3.084	0.067	3
227.8	226.5	225.7	3.457	3.277	0.067	3
231.4	230.4	229.6	2.425	2.218	0.067	3
230.5	229.9	228.5	2.27	2.131	0.067	3
229.4	228.7	227.6	3.499	3.306	0.067	3
231.2	230.6	229.3	3.563	3.386	0.067	3

4. RESULT

This section is describing the simulation results of model power theft detection in a power distribution system on AMR. The study used real data to extract features by labeling event data as normal, V_{theft} and I_{theft}, and then applied an LSTM classified model for training and testing the event. In this work, we used data to train 80% of the classifier algorithm and 20% to test it. The results of model classification are shown in the pattern of the confusion matrix and accuracy.

Table 4 Experimental results of classifier model by using training data

No.	Type of Data set	Data set	Detection Result			Detection Accuracy
			Normal	V_{theft}	I_{theft}	
1	Normal	**15,237,942**	**15,237,942**	0	0	**100%**
2	V_{theft}	**7,422,565**	0	**7,422,565**	0	
3	I_{theft}	**43,141**	0	0	**43,141**	

As Table 4 to illustrated that the number actual of each class can be learned the feature of event type accurately. It demonstrated the feature unconfigure. the accuracy of the model is achieved at 100%. That shows a model has high accuracy.

Table 5 Experimental results of classifier model by using test data

No.	Type of Data set	Data set	Detection Result			Detection Accuracy
			Normal	V_{theft}	I_{theft}	
1	Normal	**3,807,480**	**3,807,480**	0	0	**99.95%**
2	V_{theft}	**1,854,500**	**3,384**	**1,851,116**	0	
3	I_{theft}	**9,650**	0	0	**9,650**	

The result in Table 5 demonstrates that is some missed prediction from the actual power theft by voltage prediction to a normal 3,384 data set. Due to this, the attribute data between both is similar. When compare the total data is less. However, the overall accuracy of detection of the event is achieved at 99.95%. Which the accuracy detection is higher than previous research on these issues.

5. CONCLUSION

The paper presented the detection method for electricity theft in the power distribution system by using the data feature extracted from industry customers to be applied with the LSTM algorithm. The results illustrate that the model can classify the event type to appear in the data accurately. The accuracy of a model is achieved at 100% of training and testing, which demonstrates the model can distinguish between the pattern of power theft and normal activity accurately. However, the model also helps reduce the meter inspection onsite and reduce power losses for a long time.

This research can be a guideline for applying the model to detect the problem the same as this paper in the existing system. Before applying this model to detect and achieve higher efficiency the data must relate to a large dataset. The preparation data for the train and test model should be feature extracted and accurate. In divided data following each class to avoid model overfitting and the selection activation function of the model should be appropriated with the number of classes output.

6. References

Bula, I., Hoxha, V., Shala, M., & Hajrizi, E. (2016). Minimizing non-technical losses with point-to-point measurement of voltage drop between "SMART"meters. *IFAC PapersOnLine*, *49*(29),206-211.

Buzau, M. M., Tejedor-Aguilera, J., Cruz-Romero, P., & Gómez-Expósito, A. (2019). Hybrid deep neural networks for detection of non-technical losses in electricity smart meters. *IEEE Transactions on Power Systems*, *35*(2), 1254-1263.

EDL (2020). Annual Report. EDL Vientiane.

Ghori, K. M., Abbasi, R. A., Awais, M., Imran, M., Ullah, A., & Szathmary, L. (2019). Performance analysis of different types of machine learning classifiers for non-technical loss detection. *IEEE Access*, *8*, 16033-16048.

Glauner, P., Meira, J. A., Dolberg, L., State, R., Bettinger, F., & Rangoni, Y. (2016, December). Neighborhood features help detecting non-technical losses in big data sets. In *Proceedings of the 3rd IEEE/ACM International Conference on Big Data Computing, Applications and Technologies* (pp. 253-261).

Long, H., Chen, C., Gu, W., Xie, J., Wang, Z., & Li, G. (2020). A Data-Driven Combined Algorithm for Abnormal Power Loss Detection in the Distribution Network. *IEEE Access*, *8*, 24675-24686.

Micheli, G., Soda, E., Vespucci, M. T., Gobbi, M., & Bertani, A. (2019). Big data analytics: an aid to detection of non-technical losses in power utilities. *Computational Management Science*, *16*(1), 329-343.

Nabil, M., Ismail, M., Mahmoud, M., Shahin, M., Qaraqe, K., & Serpedin, E. (2018, August). Deep recurrent electricity theft detection in AMI networks with random tuning of hyper-parameters. In *2018 24th International Conference on Pattern Recognition (ICPR)* (pp. 740-745). IEEE.

Toma, R. N., Hasan, M. N., Nahid, A. A., & Li, B. (2019, May). Electricity theft detection to reduce non-technical loss using support vector machine in smart grid. In *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)* (pp. 1-6). IEEE.

Veerasamy, V., Wahab, N. I. A., Othman, M. L., Padmanaban, S., Sekar, K., Ramachandran, R., ... & Islam, M. Z. (2021). LSTM recurrent neural network classifier for high impedance fault detection in solar PV integrated power system. *IEEE Access*, *9*, 32672-32687.

Working Group on Losses Reduction CIRED WG CC-2015-2 (2017). Reduction of Technical and Non-Technical Losses in Distribution Networks. CIRED http://www.cired.net.

NUEJ
Naresuan University
Engineering Journal

Analysis of Aeration Rate and Bubble Characterizations with Different Diffusers

Prattakorn Sittisom

Department of Civil Engineering, Faculty of Engineering, Naresuan University, Phisanulok, Thailand

Corresponding author e-mail: prattakorns@nu.ac.th

(Received: 22 January 2022, Revised: 2 June 2022, Accepted: 10 June 2022)

Abstract

This study addressed the increasing of dissolved oxygen (DO), using DO meter, bubble vertical velocity and size between different types of air diffuser, using Tracker software. Based on the findings of this study, it was found that bubble vertical velocity and size are affected by several factors such as flow rate, type of air diffuser, and flow turbulence. Among these factors, it was noticed that the flow rate has the greatest effect on the bubble vertical velocity and size followed by type of air diffuser. Obtained results show that a large sphere diffuser (D_l) provided the greatest increasing rate of DO and finer bubbles which easily dissolved into the water. But a small sphere (D_s) and cylinder (D_c) diffuser provided bigger bubbles with higher bubble vertical velocity which led to less of contact time and lower DO increasing rate.

Keywords: Dissolved Oxygen, Aeration, Bubble Vertical Velocity, Air Diffuser

1. INTRODUCTION

One of the most common methods for treating domestic and industrial wastewaters is activated sludge. Diffused aeration or mechanical aeration are used to provide oxygen to the wastewater in an activated sludge process. The diffuser is used to introduce air bubbles near the bottom of the tank in diffusing aeration. For optimal oxygen transfer to water (Atta et al., 2011).

Aeration of wastewater improves the removal of Biological Oxygen Demand (BOD), which is the amount of oxygen consumed by microorganisms in the oxidation of pollutants in wastewater, Nitrogen removal, and aerobic sludge treatment. And the cost of aeration comes to more than 50% of the total energy consumption of the entire treatment process and that could be reduced substantially by using intermittent aeration (Garcia-Ochoa et al., 2000; Calik et al., 2004; Liu et al., 2006).

The type, size, and shape of the diffuser; the air flow rate; the depth of submersion; tank design; and wastewater characteristics all influence the effectiveness of oxygen transfer. The oxygen concentration inside and outside the air bubbles is the most critical component. The second most significant aspect is surface area as higher surface area enhances oxygen transfer. A longer bubble residency time for air bubbles inside water allows oxygen to dissolve in water. A small bubble size improves oxygen transfer for two reasons. First, if a large bubble divides into smaller bubbles, the surface area of these smaller bubbles increases. Second, a smaller bubble will have less buoyancy force, resulting in a longer residency time and more travel time to reach the water's surface (Sommerfeld, M., 2009).

Levitsky et al. (2005, p. 242) studied the water oxygenation in an experimental aerator with different air–water interaction patterns. A device for water saturation by gas using enhanced air–water interaction was studied experimentally. The flow of gas and water in the device was organized in a way providing efficient gas dispersion into fine bubbles at relatively low gas and liquid supply pressures.

This permits water oxygenation to be improved and the aeration expenses to be reduced as compared with existing aerators. The setup for experimental study of the device and the measurement procedure were described. The data obtained confirm an efficiency of the proposed aerator for water oxygenation.

Alkhalidi & Amano (2011, p. 397) presented the factors affecting air bubble size when air is injected through a perforated membrane into a water pool. Critical factors that govern the size of air bubbles are the air pressure, the flow rate and the hole size of the diffuser membrane. In order to have better understanding of how bubble size can be affected and what the most affecting conditions are, the study was conducted in a computational fluid dynamic investigation that was validated by the experimental results.

This study presents aeration rate from different types of air diffuser and factors affecting bubble formation through water tank, flow rate effect and bubble vertical velocity and bubble size using a high-speed camera and software to validate the experimental study.

2. MATERIALS AND METHODS

2.1 Experimental Set-up

This study investigates the increasing of dissolved oxygen level, in aeration tank as shown in Figure 1.

The aeration model is the experimental model which is used to show the increasing level of dissolved oxygen. The water tank consists of glass basin and top-open with dimension 30cmx30cmx120cm and water level was set at 100 cm height in 25°C water tank.

There were 4 types of air diffuser, Figure 2, 2-cm-diameter small sphere (D_s), 4-cm-diameter big sphere (D_b), 3-cm-diameter and 5-cm-high cylinder (D_c) and 6-cm-diameter large sphere (D_l) diffuser, with average 0.1 mm porosity, were used in this experiment by placed at bottom of the water tank. The aeration was set at 20 L/min for every diffuser. DO was measured every 20 min in first hour, every 30 min in second hour and every 60 min in third and fourth hour.

To adjust DO level from 0 mg/L for every diffuser, Sodium Sulfite (Na_2SO_3) was used in this experiment as oxygen scavenger.

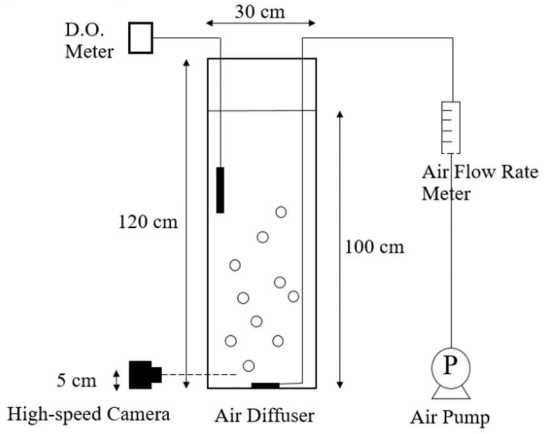


Figure 1 Schematic of Experimental Set-up

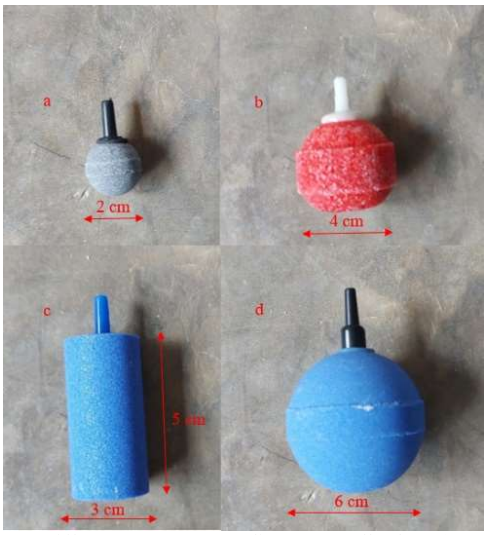


Figure 2 Diffuser types (a) 2-cm-diameter small sphere (D_s) (b) 4-cm-diameter big sphere (D_b), (c) 3-cm-diameter and 5-cm-high cylinder (D_c), (d) 6-cm-diameter large sphere (D_l)

2.2 Image Acquisition

The images of the two-phase flow were collected by a high-speed camera, SONY RX100V, with a resolution of 1980 x 1080 pixel which was set for 3.7 second video recording at 1,000 frames per second (fps) for bubble vertical velocity and size analysis.

2.3 Equivalent Bubble Diameter

Since for the present operational regime the bubble shape was in the ellipsoidal regime, also the major (A) and minor (B) axes as well as the bubble vertical velocity were determined. The equivalent bubble diameter, Figure 3, is determined from the major and minor axis as: $D_A = (AB)^{1/2}$.

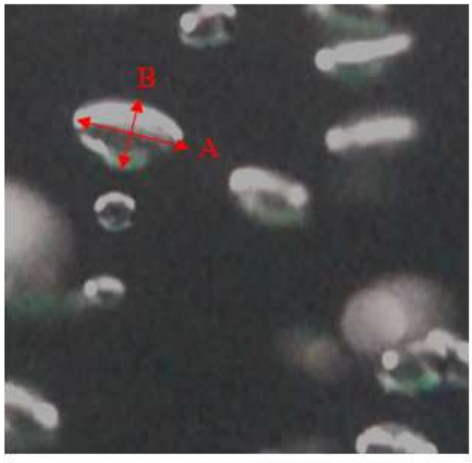


Figure 3 Major (A) and Minor (B) axes of ellipsoidal bubble

2.4 Bubble displacement measurement

Displacement was determined from the still images collected from the digital video camera. Bubble displacement was computed from the still frames obtained from the video image. The still frames were opened in commercial software, "Tracker", which was capable of showing location on an image.

The pixel coordinates (X and Y) of the bubbles center were marked and recorded. X coordinate corresponds to the distance from the left edge and Y coordinate corresponds to the distance from the top edge respectively. The pixel line running through the center of the bubble release point was known. The deviation of the bubble center from the release point was computed by subtracting the X of the bubble center from the X of the bubble release point.

2.5 Tracker Software

Tracker is a video analysis and modeling software developed by Open Source Physics (OSP) with a framework using Java. Tracker is a free application used to analyze videos. Trackers are used to detect an object, the position of the object, the object's velocity or the object's acceleration. The use of this Tracker software media can make it easier to create and represent data in graphical form. Therefore, Tracker is widely used to

detect motion material. The purpose of this study was to determine the bubbles' displacement and vertical velocity of using Tracker software.

The data analysis technique is called a Particle Model. The Particle Model is a mathematical model of a point mass. The step positions of the particle are determined by the parameters of the model rather than being marked. Particle Model has a start and end frame that define the frames of the video in which they are drawn.

An object can be said to be moving if every time its position changes with a certain reference. Some of the quantities related to straight motion include distance and space, speed and velocity, and acceleration. Based on the speed, straight motion can be divided into two, namely regular straight motion and straight motion that changes regularly. Displacement is a vector quantity that states the change in the position of the object at the reference point. Whereas distance is a scalar quantity which states the length of the path that an object passes. Speed is the result of a comparison between the distance traveled and the time an event occurs. If velocity is defined as the distance traveled per unit time, velocity is defined as displacement per unit time.

Regular straight motion is the motion of an object in a straight line at a fixed speed. Fixed speed means that the acceleration experienced by an object that is moving in a regular straight is zero. Therefore, the formula that applies to this motion is the formula for speed and speed. Speed can be formulated as follows:

$$v = \frac{s}{t} \quad \text{..... (1)}$$

Where:
v: speed or velocity (m/s)
s: distance (m)
t: time (s)

Acceleration is the change in speed per unit of time. If an object is moving at an increasing speed, it is called positive acceleration, whereas if an object is moving at a decreasing speed, it is called negative acceleration.

$$a = \frac{v}{t} \quad \text{..... (2)}$$

Where:
a: acceleration (m/s2)
v: velocity (m/s)
t: time (s)

Straight motion with regular changes is the motion of an object in a straight line at a constant acceleration. In addition, objects can also experience negative acceleration or it can be called deceleration. Straight motion changes regularly have an almost constant acceleration or deceleration so that the graph showing the acceleration against time is a horizontal line, while the speed changes so that the graph of velocity against time is a straight line with a slope that intersects the vertical axis.

3.RESULTS AND DISCUSSIONs

The result shows from Figure 4 that using 4 types of air diffuser enhanced the oxygen transfer differently. The variation of air diffuser affects the oxygen mass transfer level. Figure 4 shows that at air flow rate 20 L/min, the dissolved oxygen level in the case of using D_s diffuser increased from 0 mg/L to reach the level of 6.01 mg/L, D_b diffuser increased from 0 mg/L to reach the level of 6.44 mg/L, D_c diffuser increased from 0 mg/L to reach the level of 6.11 mg/L and D_l diffuser increased from 0 mg/L to reach the level of 6.76 mg/L.

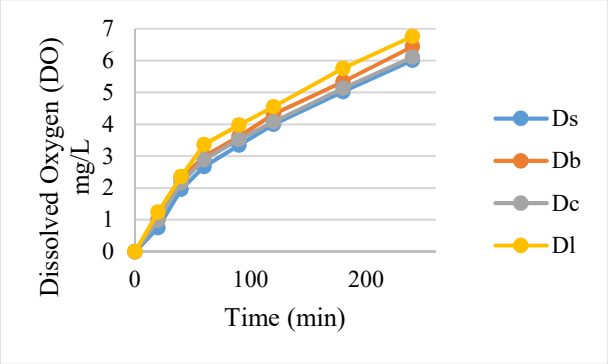


Figure 4 Dissolved oxygen (DO) level at air flow rate 20 L/min

Figure 5 shows that bubbles being captured by the high-speed camera. Figure 6 shows the comparison between bubble size and air diffuser type. Bubble diameter size from D_s was found to be around 2.5 ± 0.5 mm major diameter and 2.0 ± 0.5 mm minor diameter, bubble diameter size from D_b was found to be around 2.0 ± 0.3 mm major diameter and 1.5 ± 0.3 mm minor diameter, bubble diameter size from D_c was found to be around 2.5 ± 0.5 mm major diameter and 2.0 ± 0.5 mm minor diameter and bubble diameter size from D_l was found to be around 1.0 ± 0.2 mm major diameter and 1.0 ± 0.2 mm minor diameter by averaging a random sample of bubbles in a 3.7s video captured at 1,000 fps.

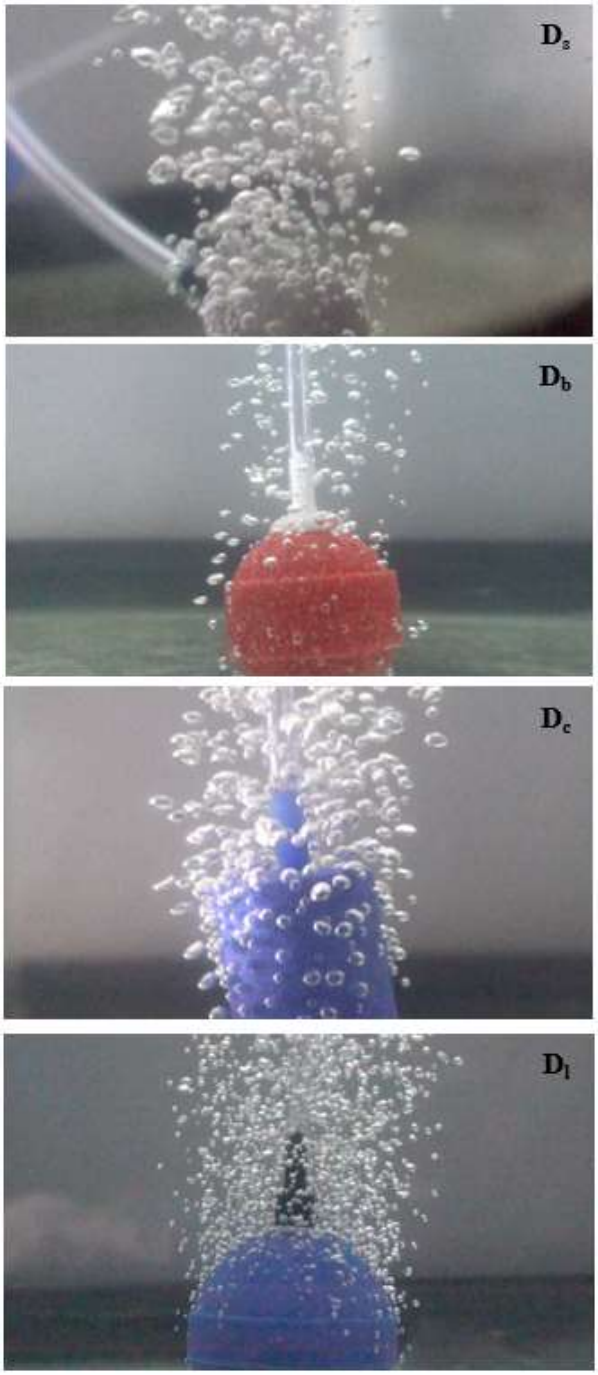


Figure 5 Air diffuser during the experiment

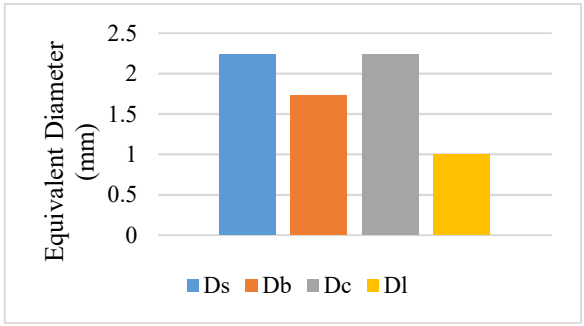


Figure 6 Bubble size from each diffuser

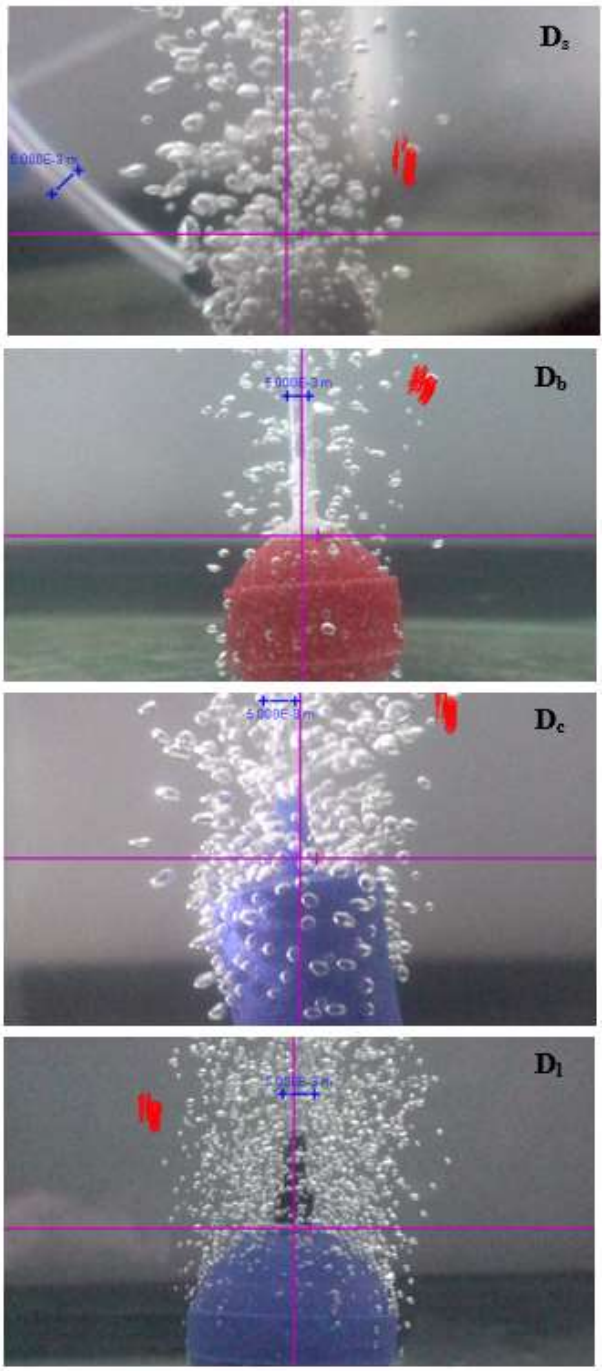
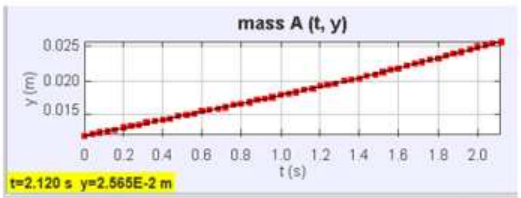


Figure 7 Screenshot of Tracker software

Figure 7 shows the screenshot from Tracker software which used for bubble vertical velocity analysis. Bubbles were selected randomly in a super slow-motion record, 40 times slow-motion.

Figure 8 shows extracted motion data of example bubble from different air diffuser. For bubble vertical velocity measurement, it was seen from Figure 5 and 7, Tracker uses the frame and frame rate to calculate t, and it uses the scale and coordinates of the marks to calculate x and y coordinates for the bubbles. It uses numerical differentiation to calculate y-velocity, vertical velocity.

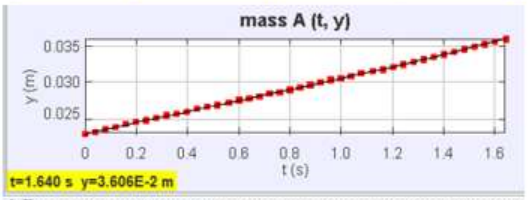
D_s



Columns ▾ mass A ▾

t (s)	x (m)	y (m)	v_y (m/s)
1.680	-2.329E-2	2.240E-2	7.468E-3
1.720	-2.338E-2	2.270E-2	7.261E-3
1.760	-2.347E-2	2.298E-2	7.198E-3
1.800	-2.356E-2	2.327E-2	7.425E-3
1.840	-2.366E-2	2.357E-2	7.540E-3
1.880	-2.375E-2	2.388E-2	7.542E-3
1.920	-2.384E-2	2.418E-2	7.398E-3
1.960	-2.395E-2	2.447E-2	7.477E-3
2.000	-2.403E-2	2.478E-2	7.499E-3
2.040	-2.414E-2	2.507E-2	7.351E-3
2.080	-2.423E-2	2.536E-2	7.251E-3
2.120	-2.432E-2	2.565E-2	

D_b

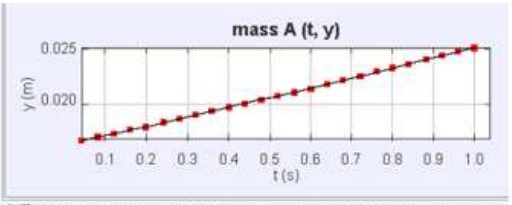


Columns ▾ mass A ▾

t (s)	x (m)	y (m)	v_y (m/s)
0.000	2.380E-2	2.306E-2	
0.040	2.388E-2	2.339E-2	8.184E-3
0.080	2.401E-2	2.371E-2	7.775E-3
0.120	2.410E-2	2.402E-2	7.379E-3
0.160	2.424E-2	2.430E-2	7.668E-3
0.200	2.437E-2	2.463E-2	7.840E-3
0.240	2.452E-2	2.493E-2	7.444E-3
0.280	2.467E-2	2.523E-2	7.375E-3
0.320	2.489E-2	2.552E-2	7.338E-3
0.360	2.510E-2	2.581E-2	7.323E-3
0.400	2.531E-2	2.611E-2	7.836E-3
0.440	2.546E-2	2.644E-2	7.964E-3

Figure 8(1) Bubble displacement, y-axis and time graph

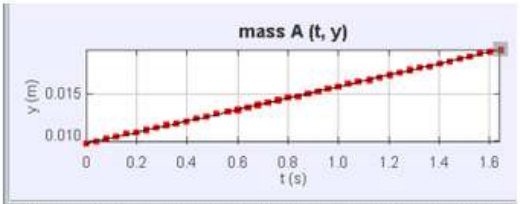
D_c



Columns ▾ mass A ▾

t (s)	x (m)	y (m)	v_y (m/s)
0.040	2.379E-2	1.681E-2	
0.080	2.383E-2	1.713E-2	7.621E-3
0.120	2.389E-2	1.742E-2	7.366E-3
0.160	2.394E-2	1.772E-2	7.610E-3
0.200	2.399E-2	1.803E-2	8.199E-3
0.240	2.404E-2	1.838E-2	8.705E-3
0.280	2.409E-2	1.873E-2	8.796E-3
0.320	2.410E-2	1.908E-2	8.832E-3
0.360	2.410E-2	1.943E-2	8.778E-3
0.400	2.410E-2	1.978E-2	8.639E-3
0.440	2.410E-2	2.012E-2	8.576E-3
0.480	2.410E-2	2.047E-2	8.301E-3

D_l



Columns ▾ mass A ▾

t (s)	x (m)	y (m)	v_y (m/s)
1.200	-2.241E-2	1.694E-2	6.101E-3
1.240	-2.242E-2	1.719E-2	6.059E-3
1.280	-2.242E-2	1.742E-2	5.979E-3
1.320	-2.242E-2	1.766E-2	5.953E-3
1.360	-2.242E-2	1.790E-2	6.030E-3
1.400	-2.242E-2	1.815E-2	6.308E-3
1.440	-2.242E-2	1.840E-2	6.210E-3
1.480	-2.242E-2	1.864E-2	6.112E-3
1.520	-2.242E-2	1.889E-2	5.919E-3
1.560	-2.241E-2	1.912E-2	6.035E-3
1.600	-2.241E-2	1.938E-2	6.128E-3
1.640	-2.241E-2	1.961E-2	

Figure 8(2) Bubble displacement, y-axis and time graph

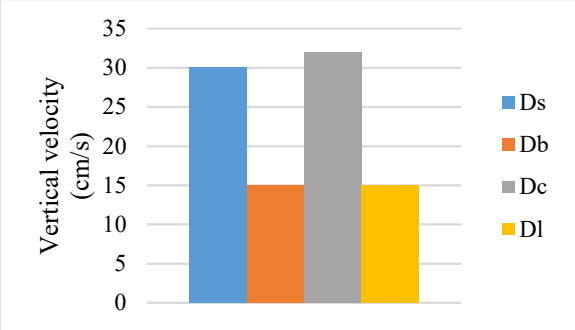


Figure 9 Bubble vertical velocity from each diffuser

The computational results presented in Figure 9 show the bubble vertical velocity from autotracker function. Bubble vertical velocity from D_s was calculated 30 cm/s, Bubble vertical velocity from D_b was calculated 15 cm/s, Bubble vertical velocity from D_c was calculated 32 cm/s and Bubble vertical velocity from D_l was calculated 10 cm/s. Figure 8 shows the screenshot of Tracker software which used for bubble vertical velocity analysis.

The results from the experiment presented that at aeration rate 20 L/min after 4 hours D_l provided the best DO at 6.76 mg/L with the finest bubble and slowest vertical velocity which let the bubbles have more total surface area and contact time in water tank. A similar trend was observed by Amma et al. (2015) who showed factors affecting fine bubble creation and bubble size for activated sludge.

On the other hand, D_c shape let bubble more easily merge each other which bigger bubbles were formed. The bigger ones have less surface area which increase vertical velocity. Therefore, bubbles provided from Dc have less contact time and DO in water tank.

4. Conclusion

The experiment was carried out to investigate the effect of different air diffuser on dissolved oxygen (DO) level, bubble vertical velocity and size. The following are the results:

1. The optimum air diffuser is the large sphere diffuser (D_l), finest porosity, achieving the highest DO level.
2. The bubble equivalent diameter relate significantly with vertical velocity.
3. The DO level depends on bubble vertical velocity and size. The smaller bubbles relate to slower vertical velocity. More contact time is greater DO level obtain.
4. The smaller diffuser increases bubble merging opportunity with lower DO level.

5. Acknowledgment

The Faculty of Engineering, Naresuan University provided financial support for this study, which is gratefully acknowledged.

6. References

Alkhalidi, A. and Amano, R. (2011) Factor Affecting Bubble Creation and Bubble Size. Proceedings of 2011ASME IMECE. ASME International Mechanical Engineering Congress & Exposition, vol. 1, pp. 397–401.

Atta, N. N., Elbaz, A. A. and Sakr, A. H., (2011). Effect of Pressure, Water Depth and Water Flow Rate on Oxygen Saturation Level in Activated Sludge Process. American Journal of Engineering and Applied Sciences, 4 (4): 435-439.

Calik, P., Yilgora, P., Ayhanb, P. and Demir, A. S., (2004). Oxygen transfer effects on recombinant benzaldehyde lyase production. Chemical Engineering and Science, 59: 5075-5083.

Garcia-Ochoa, F., Castro, E.G. and Santos, V.E., (2000). Oxygen transfer and uptake rates during xanthan gum production. Enzyme Microbiological Technology, 27: 680-690.

Levitsky, S., Grinis, L., Haddad, J. and Levitsky, M. (2005) Water Oxygenation in an Experimental Aerator with Different Air/ Water Interaction Patterns. HAIT Journal of Science and Engineering B, 2, 242–253.

Liu, Y.S., Wu, J.Y. and Ho, K.P. (2006). Characterization of oxygen transfer conditions and their effects on Phaffia rhodozyma growth and carotenoid production in shake- flask cultures. Biochemical Engineering Journal, 27: 331-335.

Sommerfeld, M., (2009). Analysis of Hydrodynamics and Microstructure in a Bubble Column by Planar Shadow Image Velocimetry. Industrial & Engineering Chemistry Research, 48, 330–340.

7. BIOGRAPHIES

Dr. Prattakorn Sittisom received the B.Eng. from Chiang Mai University in Environmental Engineering, Thailand in 2010, the M.Sc. from Chulalongkorn University in Environmental Management, Thailand in 2013 and the M.Eng. in Advanced Engineering and PhD. in Science and Technology in Engineering from the Nagasaki University, Japan in 2016 and 2019, respectively. He is currently an Instructor of Environmental Engineering at Chiang Mai University, in Chiang Mai province, Thailand. His research interests include Membrane Technology, Automation System Technology, Sensor and IoT.

NUEJ
Naresuan University
Engineering Journal

Numerical heat transfer study of an impinging jet of nanofluid of TiO_2 on a chip surface

Koonlaya Kanokjaruvijit*, Kamolthip Tuamjan, Pimwipa Srirat, Taechinee Rattanasangsri, Suttinon Panyadibwong and Pongpun Othaganont

Department of Mechanical Engineering, Faculty of Engineering, Naresuan University, Phitsanulok, Thailand

* Corresponding author e-mail: koonlayak@nu.ac.th, koonlaya@gmail.com

(Received: 18 April 2022, revised: 29 August 2022, Accepted: 5 September 2022)

Abstract

A numerical investigation of 2D axisymmetric heat transfer of an impinging jet of water-based titanium dioxide (TiO_2) nanofluid on a CPU chip surface is performed by using the finite element method with k-ε turbulence model with the wall treatment. The flat plate impingement is also studied to compare the heat transfer and flow characteristics with those of the chip plate and the average heat transfer results agree well with the experimental results obtained from literature. Parametric effects such as nanofluid concentration (ϕ), Reynolds number (Re_{Dj}) and jet-to-plate spacing (H/D_j) are examined. The nanofluid concentration is in the range of 0-6% by volume. The tested Re_{Dj} is between 2000 and 8000. The jet-to-plate spacing is between 2-4. The maximum heat transfer enhancement in terms of average Nusselt number of the TiO_2 nanofluid compared to that of the water is 18.24% for the chip plate impingement at Re_{Dj} = 2000 and ϕ = 6%; however, this maximum enhancement is 47.13% in terms of the average heat transfer coefficient. After the multiple linear regression analysis, the nondimensional heat transfer correlations are obtained. Finally, the ratios of pumping power between nanofluid and base fluid are plotted and found the penalty of 1.5 to nearly 4 times.

Keywords: Jet Impingement, Nanofluid, Titanium Dioxide, CPU Chip

1. NOMENCLATURE

c_p	specific heat, J/kg-K
D_j	jet diameter, m
h	heat transfer coefficient, W/m^2-K
H	distance between jet exit and stagnation point, m
k	turbulent kinetic energy
Nu	Nusselt number
PP	pumping power, W
q_s''	surface heat flux, W/m^2
r	radial coordinate, m
R_{chip}	radius of chip, m
R_s	radius, m
Re_{Dj}	Reynolds number based on jet diameter
p	pressure, Pa
T	temperature, K
u	fluid velocity, m/s
z	axial coordinate, m

Greek symbols

ε	dissipation rate of turbulent kinetic energy
λ	thermal conductivity
μ	fluid dynamic viscosity, Pa-s
ϕ	nanoparticles volume fraction
ρ	density

Subscripts

avg	average property
bf	with respect to base fluid
chip	with respect to chip
j	with respect to jet
nf	with respect to nanofluid
p	with respect to nanoparticles
s	with respect to surface

2. INTRODUCTION

Due to the increasing need for shrinking the size of electronic components while enhancing their capacity, the heat fluxes or power density is then substantially increased. Conventional heat sink used as a passive cooling method has probably become unfit to the packaging of the devices. Direct heat transfer from each component has been brought to a few researchers' interests and proved that it is a promising cooling scheme design for the next-generation higher performance devices. van Erp et al. (2020) has proposed a new design of electronic cooling by directly embedding multiple microfluidic channels inside the chip substrate using water as working fluid and reported the heat removal of 17 kW/cm^2 with only 0.57 W/cm^2 pumping power. Direct multiple impinging jets on a high-power device was proposed by Wei et al. (2019) reporting the highest heat transfer coefficient of 6.25×10^4 W/m^2-K with the pumping power as low as 0.3 W.

Nanofluid jet impingement has recently attracted an increasing number of researchers since Choi and Eastman have discovered the enhancement of base fluid thermal conductivity by using nanoparticles in 1995. Thus, the nanofluid could become a game changer in the heat transfer augmentation for components that are exposed to a high heat generation such as heat exchangers, electronic components, automobile components and thermal treatment of metals. In addition to increasing the thermal conductivity of the base fluid, Lv et al. (2017a,b) explained how the nanofluid augmented heat transfer in a single free impinging jet by the fact that the nanoparticles vigorously bombarded at the impingement surface. This caused intense turbulence and thinner boundary layer thickness near the surface leading to the heat transfer improvement.

Generally, the base fluids of the nanofluids are water, engine oil and ethylene glycol. However, most studies have used water or deionized water. The nanoparticles are usually metallic compounds such as Al_2O_3, SiO_2, CuO, ZnO and TiO_2. The research on nanofluid jet impingement has been carried out through experiments and numerical method in either a free-surface jet and a confined jet mostly to investigate the main parametric effects on heat transfer such as Reynolds number, jet-to-plate spacing and nanofluid concentration. Then, the correlations were presented. Zeitoun & Ali (2012), Modak et al. (2015) and Lv et al. (2017b) experimentally exhibited the heat transfer enhancement of the water-based Al_2O_3 nanofluid by approximately 60% at moderate concentration ($\phi \leq 6\%$ by volume) compared to the base fluid. Lv et al. (2017a) investigated the SiO_2 nanofluid impingement leading to the maximum 40% enhancement at Re = 8000 and ϕ = 3%. Barewar et al. (2019) studied Zuo nanofluid impingement at low concentration ($0.02 \leq \phi \leq 0.1\%$) and reported the maximum heat transfer augmentation by 55% at ϕ = 0.1% and H/D_j = 3.5. Naphon & Nakharintr (2012) used TiO_2 nanofluid at a fixed ϕ = 0.2% impinging on a mini fin heat sink to test the influence of the jet inlet temperature (T_j) and found that at higher T_j, i.e., 30°C the nanofluid did not help enhance the heat transfer. However, at lower T_j, i.e., 20°C the nanofluid became enhancing the heat transfer. The pressure drop data were also reported.

There have been a growing number of numerical analyses on the nanofluid impinging jet. Manca et al. (2011) investigated a slot jet of Al_2O_3 nanofluid using a single phase and k-ε turbulence model to study the effects of Reynolds number, jet-to-plate spacing (H/b) and concentration. They found the maximum heat transfer enhancement by 18% at ϕ = 6% and H/b = 10. Furthermore, they explained the decrease in heat transfer at the small H/b that was because the jet counter-rotating vortices were forced to leave before touching the stagnation region causing the secondary peak. Two-phase model for the Al_2O_3 nanofluid impingement was investigated by Huang & Jang (2013) and Abdelrehim et al. (2019). Abdelrehim et al. (2019) studied the laminar jet impingement and reported that the two-phase model gave better heat transfer results than the single-phase model when compared to the experimental ones. Alabdaly et al. (2019) numerically investigated a 2D planar laminar impinging jet of SiO_2 nanofluid on different sizes of semicircle bumps installed at some distance after the stagnation region. The higher Nusselt number was reported at the higher bump.

TiO_2 is nontoxic and inexpensive material used in variety of our everyday life including as a dye for off-white color in soap making process. However, not much research on nanofluid heat transfer include TiO_2 as nanoparticles, even though the heat transfer augmentation among TiO_2, Al_2O_3 and CuO nanofluids were found insignificantly different by Roy et al. (2012). Furthermore, few research work has reported pumping power or pressure drop (mostly carried out via numerical analysis research such as Abdelrehim et al. (2019), Alabdaly & Ahmed (2019), Roy et al. (2012) and Manca et al. (2011), but the experimental work by Naphon & Nakharintr (2012) showed the data of pressure drop) and the heat transfer correlations (carried out through either numerical investigation or experiments such as Lv et al, (2017a, b), Zeitoun & Ali (2012), Barewar et al. (2019), Modak et al. (2015) and Abdelrehim et al. (2019)) altogether, which are essential to heat transfer-related machine design.

The objectives of this research are to propose an alternative way of onsite cooling a CPU chip and to find the heat transfer characteristics of a circular impinging jet of water-based TiO_2 nanofluid on to a constant heat flux surface for both flat and chip surfaces. Parametric effects such as nanofluid concentration (ϕ), Reynolds number (Re) and jet-to-plate spacing (H/D_j) are considered. The heat transfer results are reported in terms of local and area average Nusselt numbers. With the use of the multiple linear regression, the heat transfer correlations are presented. Finally, pumping power ratios of the nanofluid and the base fluid are also discussed.

3. RESEARCH METHODOLOGY

Throughout this study, a steady state, incompressible single-phase turbulent flow is assumed. In addition, at the impingement surface, the constant heat flux is assigned with assumption of no heat radiation. The jet impingement system is semi-confined with no end effect assumption.

3.1 Computational Domain

Two schematic setups of the 2D axi-symmetric computational domains were tested in this study as shown in Fig 1. A circular jet nozzle of radius 2 mm is attached to a flat surface to form an upper boundary for the computational domain. A CPU chip of 9.15 mm radius (R_{chip}) and 5 mm height exhibited in Fig.1(a) is placed under the nozzle. The constant heat flux surface is

assigned to the chip surface only, and the rest of the impingement plate is insulated. In Fig.1(b), the entire flat plate is assigned constant heat flux. All surfaces within the domain are at no-slip condition with the radius (R_s) of 25 mm. After impinging, the spent fluid is to leave the exit, at which, the pressure is set to be equal to the ambient pressure.

Nanofluid of TiO_2 nanoparticles in water as a base fluid is the working fluid in this study. Pure water is also used for baseline cases. Hence, in this study, the heat transfer of the free-surface liquid jet impingement is considered.

3.2 Governing Equations

Under assumptions of steady state, incompressible, turbulent, single-phase flow and negligible heat radiation, The Reynolds time averaged Navier-Stokes (RANS) equations are shown in as follows

Continuity:

$$\frac{\partial \bar{u}_i}{\partial x_i} = 0 \tag{1}$$

Momentum:

$$\rho \bar{u}_i \frac{\partial \bar{u}_i}{\partial x_j} = -\frac{\partial \bar{p}}{\partial x_i} + \frac{\partial}{\partial x_j}\left[\mu\left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i}\right) - \rho\overline{u_i' u_j'}\right] \tag{2}$$

Energy:

$$\rho \bar{u}_i \frac{\partial \bar{T}}{\partial x_j} = \frac{\partial}{\partial x_j}\left[\frac{\mu}{Pr}\frac{\partial \bar{T}}{\partial x_j} - \rho\overline{T' u_j'}\right] \tag{3}$$

where $\bar{u}_i$, $\bar{u}_j$, $\bar{T}$ and $\bar{p}$ are time-averaged velocity components, temperature and pressure, respectively.

u_i', u_j' and T' are fluctuating quantities.

Noting that the Reynolds stress term, $-\rho\overline{u_i' u_j'}$, is based upon the Boussineq hypothesis show as follows

$$-\rho\overline{u_i' u_j'} = \nu_t\left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i}\right) - \frac{2}{3}\delta_{ij} \tag{4}$$

and the turbulent heat flux term

$$\rho\overline{T' u_j'} = \frac{\mu_t}{Pr_t}\left(\frac{\partial \bar{T}}{\partial x_j}\right) \tag{5}$$

The k-ε turbulence model with wall treatment for the steady state impinging jet with the assumption of fully developed turbulent flow is exhibited in the following equations. These equations combined with the above RANS equations thus govern the turbulent flow and heat transfer of the domain.

Energy:

$$\frac{\partial(\rho k u_i)}{\partial x_j} = \frac{\partial}{\partial x_j}\left(\mu + \frac{\mu_t}{\sigma_k}\right)\frac{\partial k}{\partial x_j} + (G_k + G_b) - \rho\varepsilon - Y_M + S_\varepsilon \tag{6}$$

Dissipation:

$$\frac{\partial(\rho\varepsilon u_i)}{\partial x_j} = \frac{\partial}{\partial x_j}\left(\mu + \frac{\mu_t}{\sigma_\varepsilon}\frac{\partial \varepsilon}{\partial x_j}\right) + C_1\frac{\varepsilon}{k}(G_k + G_{3\varepsilon}G_b) - C_2\rho\frac{\varepsilon^2}{k} + S_\varepsilon \tag{7}$$

where G_k and G_b denote the turbulent kinetic energy (tke) production terms due to the velocity gradients and the buoyancy, respectively. Y_M is the fluctuation rate associated with the tke dissipation. σ_k and σ_ε represent effective Prandtl numbers, which associates to eddy diffusion of k and ε to the momentum eddy viscosity as shown below

$$\sigma_k = \frac{\nu_t}{\nu_k} \quad \text{and} \quad \sigma_\varepsilon = \frac{\nu_t}{\nu_\varepsilon}$$

and the eddy viscosity is

$$\nu_t = \frac{C_\mu k^2}{\varepsilon}$$

The empirical constant values for jet impingement are C_μ = 0.09, C_1 = 1.44, C_2 = 1.92, σ_k = 1.0, and σ_ε = 1.3. The buoyancy dependence rate of ε is denoted $C_{3\varepsilon}$ which is equal to $\tanh|v/u|$. S_k and S_ε represent the generation terms.

3.3 Finite Element Method

A commercial program of finite element method, COMSOL, was used throughout the study. The program uses free meshing techniques to generate unstructured meshes covering the computational domain. Each mesh element is a triangle of 3 vertex node with linear shape function for all unknown variables: velocity components, pressure and temperature.

Note on the wall treatment with the use of k-ε turbulence model in COMSOL that the wall life-off from the physical surface was defined as $\delta_W = \frac{\delta_W^+ \mu}{\rho u_\tau}$, where the term δ_W^+ is determined from $\delta_W^+ = \max\left(\frac{h_w}{2}\frac{\rho C_\mu^{1/4}\sqrt{k}}{\mu}, 11.06\right)$ whether which one is higher. The first term is derived from the law of the wall, and the second one refers to the distance measured from the wall, where the logarithmic layer encounters with the laminar sublayer. Here h_w is denoted by the height of the mesh adjacent to the wall.

The boundary conditions shown in Fig.1. All surfaces are set no-slip conditions. The thermal boundary conditions at surfaces are set insulated or dT/dt = 0 except at the impingement surface, i.e., the entire plate for the flat plate case (Fig.1(a)) and the chip surface for the chip plate (Fig.1(b)). The inlet toward the impingement plate is at the nozzle exit. Finally, after impinging the fluid leaves the surface toward the exit to ambient.

3.4 Data Reduction

In this study, 3 parametric effects are considered: Reynolds number (Re_{Dj}), jet-to-plate spacing (H/D_j) and nanofluid concentration (ϕ), and the values are shown in

Table 1. The heat transfer results are expressed in terms of heat transfer coefficient (h) and Nusselt number (Nu_{Dj}). The results are presented as the dimensionless heat transfer correlations for both flat and chip surfaces.

Single phase fluid is assumed throughout the study. The nanoparticles of TiO_2 change the thermal properties of the base fluid, such as density (ρ), dynamic viscosity (μ), thermal conductivity (λ) and specific heat (c_p) according to the empirical formulae summarized by Said et al. (2014).

$$\rho_{nf} = (1-\phi)\rho_{bf} + \phi\rho_P \quad (8)$$

$$\mu_{nf} = \mu_{bf}(123\phi^2 + 7.3\phi + 1) \quad (9)$$

$$\lambda_{nf} = \lambda_{bf}\left(\frac{\lambda_P + 2\lambda_{bf} + 2(\lambda_P - \lambda_{bf})\phi}{\lambda_P + 2\lambda_{bf} - (\lambda_P - \lambda_{bf})\phi}\right) \quad (10)$$

$$c_{p,nf} = (1-\phi)c_{p,bf} + \phi c_{p,P} \quad (11)$$

The subscripts nf, bf and P denote nanofluid, base fluid and nanoparticle, respectively.

The Reynolds number is defined based on jet diameter (D_j) as follows

$$Re_{D_j} = \frac{\rho u_j D_j}{\mu} \quad (12)$$

After obtaining the surface temperature (T_s) from the computation, Newton's law of cooling gives

$$h = \frac{q_s''}{(T_s - T_j)} \quad (13)$$

where the surface heat flux is set to be 500 kW/m^2. This value was obtained from the specification of the CPU chip of Intel Core i7-920 (Kirsch (2008)). Then, the area-averaged heat transfer coefficient is calculated.

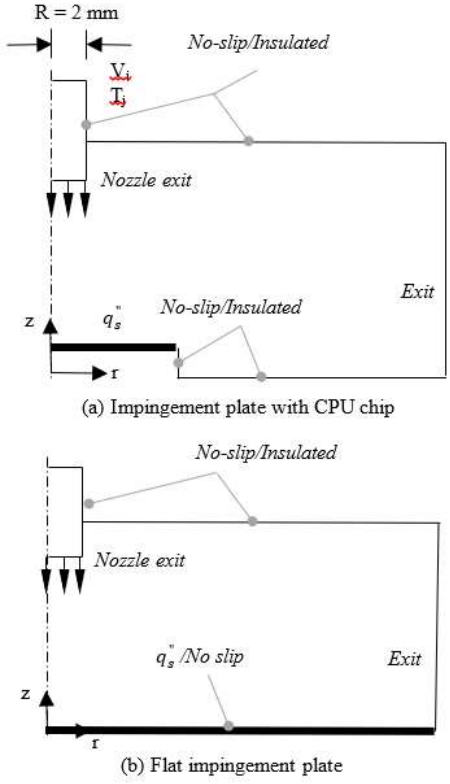


Figure 1 Computational domains and boundary conditions (Not to scale)

Table 1 Parametric effects and their values

Parametric effects	Values
Re_{Dj}	2000, 4000, 6000, 8000
H/D_j	2, 3, 4
ϕ (% Volume)	0, 2, 4, 6

3.5 Mesh Density Dependence Check

Four different models of number of meshes are tested as shown in Fig.2: 4800, 9480, 49628 and 100677 meshes, namely models A, B, C and D, respectively, in the flat impingement surface domain for Re = 2000, 4000, 6000 and 8000. The surface mesh sizes are mainly refined: 0.1, 0.05, 0.01 and 0.005 mm for models A, B, C and D, respectively. The results in terms of average Nusselt numbers are compared as well as computational time and RAM usage. Moreover, the mesh quality is also taken into consideration as the finer mesh costs lower quality, despite less error [Holzbecher & Si]. Model D is thus assumed to give the most accurate solution in terms of average Nusselt number.

The maximum difference from model D is found at model B for 7% at Re_{Dj} = 8000 while model A 6% at Re_{Dj} = 6000. However, when comparing models C and D, the difference is close to 0%. Considering RAM usage and computing times between models C and D, model C uses up to 0.5 GB less than model D while the computing time is nearly half of that of model D. The average mesh quality of model C is 11% higher. Therefore, model C is chosen to fit the accuracy and the specification of the computer used in this study is CPU Intel Corei5-8300H 4.0 GHz and 8 GB RAM.

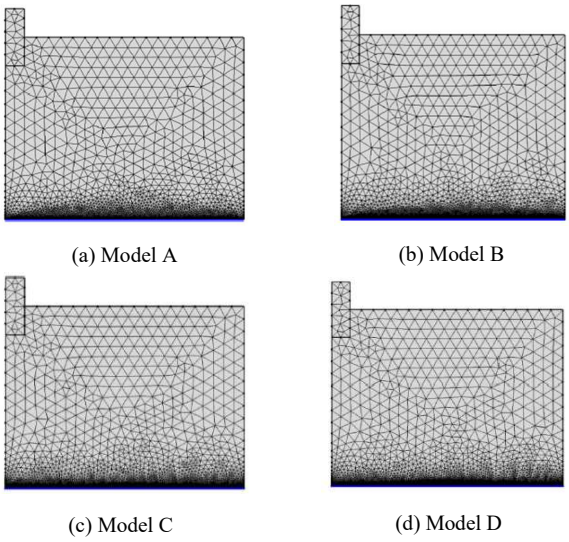


Figure 2 Mesh refinement especially near wall region

3.6 Validation

Fig.3 shows the comparison of the average Nusselt number obtained from the present study to the

experimental results of Sorour et al. (2019) at $H/D_j = 4$ and Re = 8000-14000. The differences are found 5.71%, 4.88%, 6.25% and 9.09% at 8000, 10000, 12000 and 14000, respectively. The differences were probably because the experimental setup of Sorour et al. (2019) was carried out through the heat flux of 14,920 W/m^2, D_j = 6 mm and D_s = 320 mm while the present computational domain was set to the heat flux of 500 kW/m^2, D_j = 4 mm and D_s = 50 mm. Sorour et al. (2019) showed that H/D_j did not significantly affect the Nusselt number.

The comparison of area averaged heat transfer coefficients of the current study to those of Lv et al. (2017) are shown in Fig.4. The tested Reynolds number in this study is from 2000 to 8000, while Lv at al. (2017) from 6114 to 12228, D_j = 2 mm, D_s = 100 mm, but the surface thermal condition was not reported. The similar trends of the data are found. Furthermore, both Lv et al. (2017) and the current study have also found that H/D_j slightly influences the heat transfer for the free-surface water impingement. However, for the clear illustration purpose, the heat transfer results for each H/D_j value are shown separately.

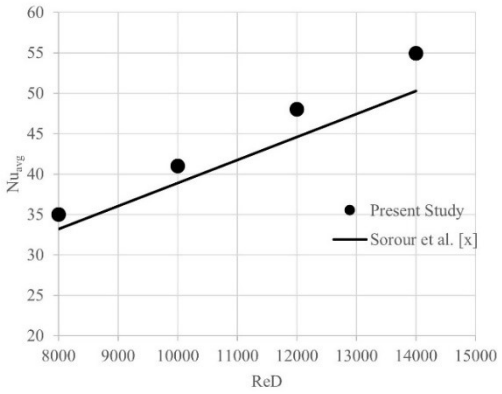
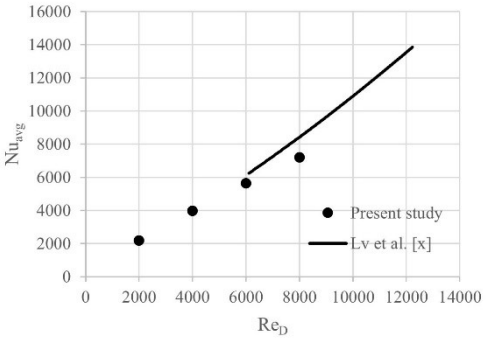
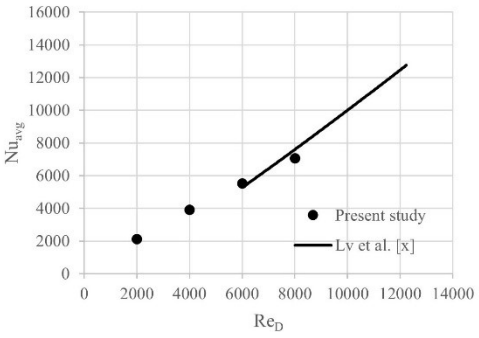


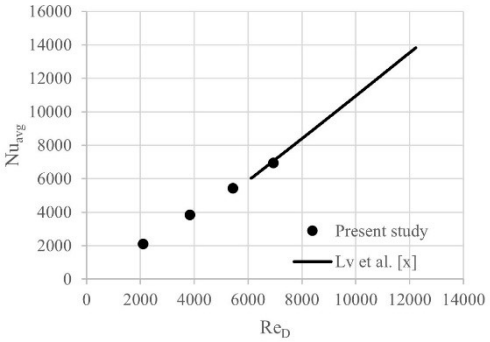
Figure 3 Comparison of the average Nusselt number of the present study and Sorour et al. (2019)



(a) $H/D_j = 2$



(b) $H/D_j = 3$



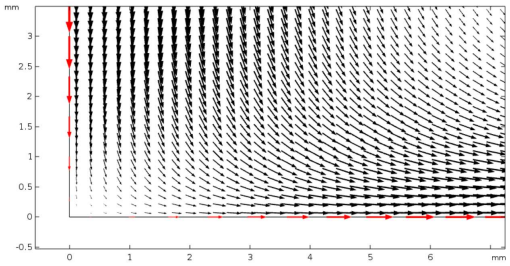
(c) $H/D_j = 4$

Figure 4 Comparison of heat transfer coefficients of the present study to those of Lv et al. (2017)

4. RESULTS AND DISCUSSION

4.1 Comparison of flow fields of flat plate impingement and chip plate impingement

Flat plate impingement as a baseline case is carried out along with the chip plate impingement in order to better understand the heat transfer results on the chip plate later on. Fig.5(a) shows the oncoming jet impinging on a flat surface and changing the flow direction along the impingement surface to form a wall jet then leave at the exit. In Fig.5(b), after impinging on a chip, the flow cannot leave towards the exit immediately, but circulates near the vicinity of the height of the chip prior to changing the direction along the flat area. This could be explained by the fact that the edge of the chip disrupts the boundary layer of the flow. Furthermore, according to Zhao et al. (2002), the chip surface covers the stagnation zone, while in Fig.5 (a) the viscous boundary layer starts right next to the stagnation zone. Note that the constant heat flux is assigned to the entire length of the flat plate case, while only to the chip surface and the flat area is insulated.



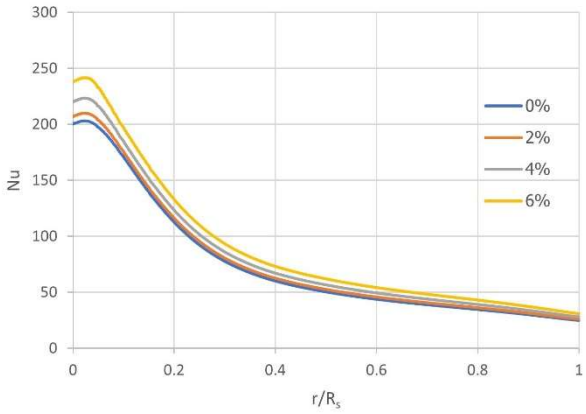
(a) Flat plate impingement

(b) Chip plate impingement

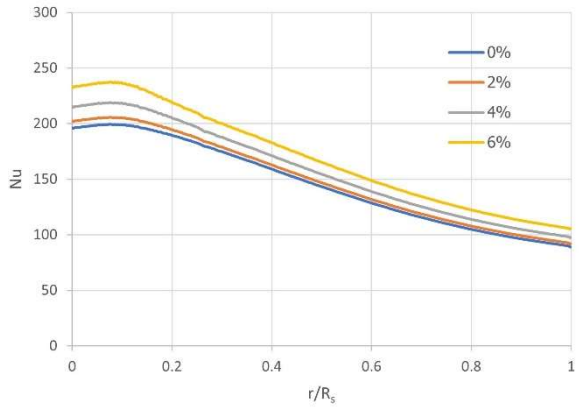
Figure 5 Vector plots of flow near the impingement vicinity

4.2 Effect of Nanofluid Concentration

Fig.6 shows the local Nusselt number distributions at Re = 8000 and H/D_j = 2 for different nanofluid concentrations. Note that only the chip surface is assigned constant heat flux and the heat transfer results on the chip surface are taken into account. Nusselt number increases with the nanofluid concentrations for both cases of flat plate and chip plate impingement. In Fig.6(a) for the flat plate impingement, the heat transfer enhancement of nanofluids compared to the base fluid at the stagnation point for Φ = 2, 4 and 6% are 3.38, 9.97 and 18.97%, respectively. However, the maximum value is shifted slightly away from the centerline and becomes the secondary peak for all cases. The differences in the secondary peak between the nanofluids of Φ = 2, 4 and 6% and the base fluid are 3.37, 9.96 and 18.97%, respectively. For the chip plate impingement, the differences in the stagnation points are 3.08, 9.80 and 18.89%, respectively. The differences in the secondary peaks are 9.08, 9.81 and 18.95%, respectively. According to Manca et al. (2011), the counter-rotating vortices of the oncoming jet towards the target plate are responsible for the secondary peaks as the vortices do not yet touch the stagnation point but are pushed away parallel to the surface towards the exit.



(a) Flat Plate



(b) Chip Plate

Figure 6 Local Nusselt number for different concentrations of Nanofluids at Re = 8000 and H/D_j = 2

However, the flow characteristics of nanofluids are not different from that of the base fluid (water), because the assumption of single phase nanofluid only changes the numeric values of the thermal properties of the working fluid. Nonetheless, Lv et al. (2017) confirmed with their experimental results that addition of the nanoparticles into a base fluid did not affect the flow characteristic of the base fluid explained by the individual nanoparticle bombardment effect on the surface. Thus, this helps enhance the heat transfer of the base fluid.

Fig.6(a) shows the comparison of area averaged Nusselt numbers at different concentrations for both flat

plate and chip plate impingement. The results are in accordance with the local Nusselt number variation that the heat transfer enhancement increases with the nanofluid concentration. The chip plate impingement gives higher average Nu than the flat plate due to the fact that only the chip surface was considered in the calculation, while the entire surface of flat plate impingement was brought into the calculation. The maximum heat transfer enhancement in terms of Nu_{avg} is 24.57% found at Re_{Dj} = 2000 and ϕ = 6% for the flat plate impingement and 18.24% at Re_{Dj} = 2000 and ϕ = 6% for the chip plate impingement. However, when considering in terms of area average heat transfer coefficients as exhibited in Figs.6(b) and (c) for the flat and the chip impingement, respectively, the maximum enhancement is also found at Re_{Dj} = 2000 and ϕ = 6% by 47.13% and 39.65%. Nevertheless, noting that the enhancement found at other values of Re_{Dj} (ϕ = 6%) is only 1-2% lower than at Re_{Dj} = 2000.

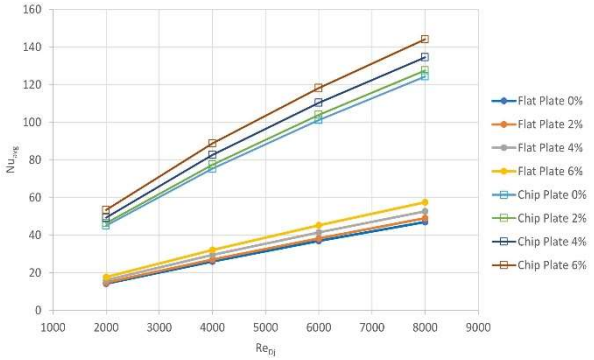
4.3 Effect of Reynolds Number

Generally, the impingement heat transfer results are significantly influenced by Reynolds number. Higher Re leads to higher heat transfer results as illustrated in Fig.7 for both flat and chip plate impingement as the jet velocity at the higher Re_{Dj} becomes higher and impinges more vigorously causing the boundary layer thinner. This also allows the jet with higher Re_{Dj} to exchange more momentum leading to higher heat transfer enhancement. The area average Nusselt numbers are plotted against Re_{Dj} at different values of jet-to-plate spacing, which will be discussed in the following topic. Re_{Dj} = 8000 gives the highest followed by Re_{Dj} = 6000, 4000 and 2000, respectively.

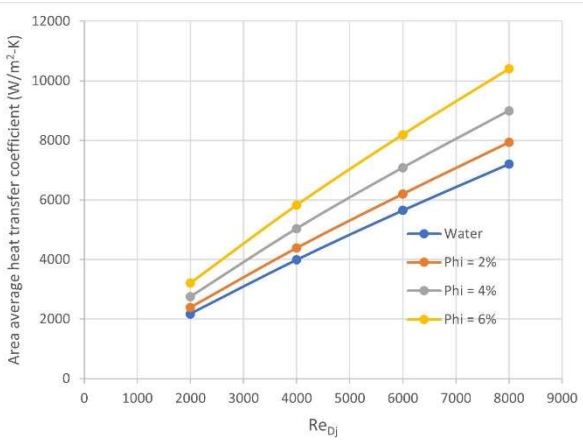
The area averaged Nusselt numbers at different Re and H/D_j are plotted in Fig.8 for ϕ = 4%. Consider Fig.8(a), H/D_j does not play important role to Nu_{avg} values. Re_{Dj} = 8000 gives the area average Nu approximately 70% higher than Re = 2000, and 44% and 21 higher than Re_{Dj} = 4000 and 6000, respectively, at all values of H/D_j. In Fig.8(b), the chip plate impingement shows the influence of H/D_j; however, the differences are similar. At Re_{Dj} = 8000, Nu_{avg} is approximately 64%, 39% and 18% higher than those at Re_{Dj} = 2000, 4000 and 6000, respectively.

4.4 Effect of Jet-to-Plate Spacing

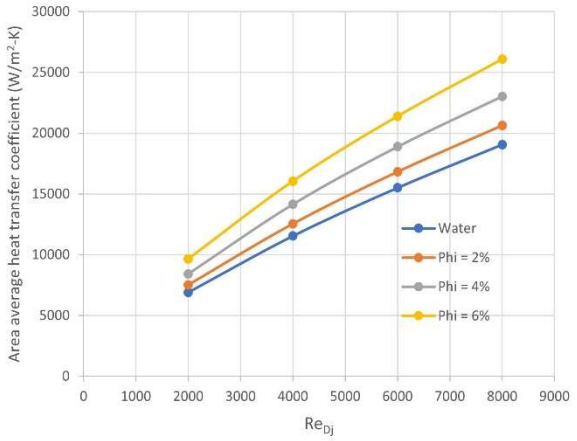
Fig.10 shows the distribution of local Nusselt number for different jet-to-plate spacings at Re_{Dj} = 6000 and ϕ = 4%. Secondary peaks are found in all cases and prominently shown in the case of H/D_j = 2. This has been explained by some researchers (Cornaro et al. (1999), Manca et al.(2011) and Lv et al. (2017)) by the fact that at low H/D_j a jet cannot establish fully developed flow before approaching the target plate, thus the counter-rotating vortices are pushed parallel to the impingement surface towards the exit. Notwithstanding that, H/D_j = 2 leads to the highest local heat transfer results compared to H/D_j = 3 and 4.



(a) Average Nusselt number



(b) Area average heat transfer coefficient for flat plate impingement

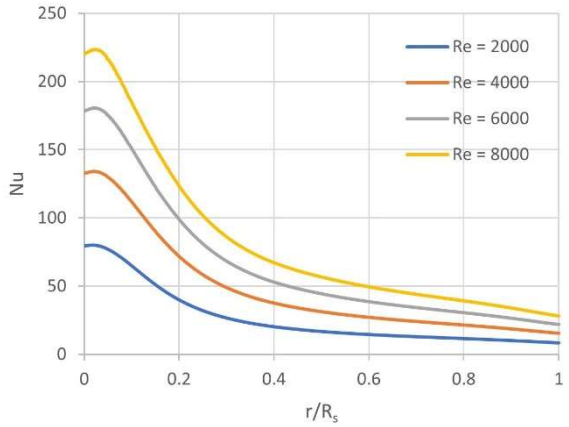


(c) Area average heat transfer coefficient for chip plate impingement

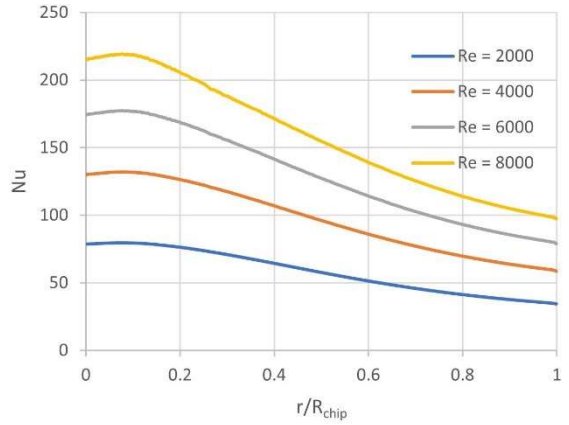
Figure 7 Average heat transfer results at H/D_j = 2.

Referring to the area averaged Nusselt numbers in Fig.9, H/D_j does not significantly affect the area average Nusselt number. This can be shown in the local distribution in Fig.10(a) that the local Nu graphs collapse after the stagnation zone. The area averaged Nu for the chip plate impingement is higher than the flat plate

impingement, because only the chip surface is included in the area average. The differences in percentage at each Re_{Dj} are quite similar such as at $H/D_j = 2$, Nu_{avg} is approximately 9% and 18% higher than those at $H/D_j = 3$ and 4, respectively, while at $H/D_j = 3$, Nu_{avg} is approximately 8% higher than that of $H/D_j = 4$.

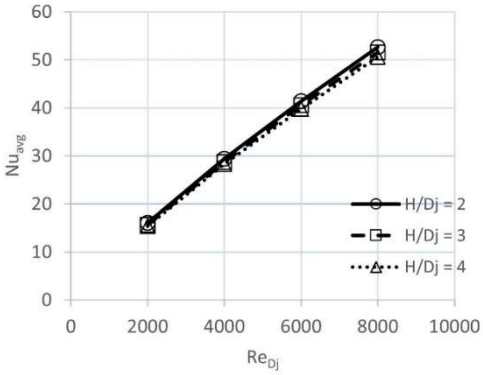


(a) Flat Plate

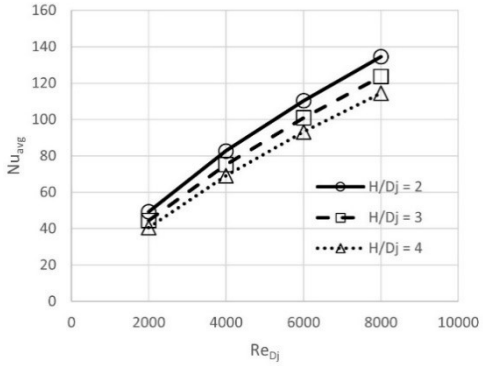


(b) Chip Plate

Figure 8 Comparison of local Nusselt number at different Reynolds number at $\Phi = 4\%$ and $H/D_i = 2$.



(a) Flat Plate



(b) Chip Plate

Figure 9 Comparison of area averaged Nusselt numbers for $\phi = 4\%$.

4.5 Heat Transfer Correlations

Nondimensional heat transfer correlations were obtained from the multiple linear regression analysis through the analysis of variance (ANOVA). The area averaged Nusselt number becomes the dependent variable or the response variable, and the independent variables or explanatory variables are Re_{Dj}, H/D_j and Φ. Then, all variables are log-transformed, then the multiple linear regression is performed with the R^2 value greater than 0.99 for all cases. In addition, 2-sided p-values for testing the coefficients are less than 0.05. Finally, after anti-logging, 4 heat transfer correlations are obtained: water impinging on a flat plate, nanofluid impinging on a flat plate, water impinging on a chip plate and nanofluid impinging on a chip plate.

For flat plate impingement, the correlation of Nusselt number to Reynolds number and jet-to-plate spacing is shown as follows

$$Nu_{avg} = 0.0209 Re_{D_j}^{0.8638} \left(H/D_j\right)^{-0.0557} \quad (16)$$

For the nanofluid impingement on a flat plate, the correlation is

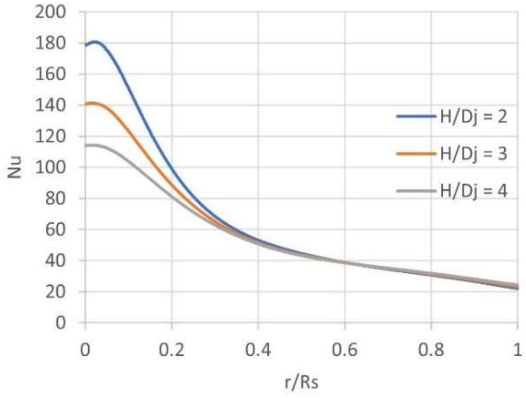
$$Nu_{avg} = 0.0557 Re_{D_j}^{0.8587} \left(H/D_j\right)^{-0.056} \phi^{0.1458} \quad (17)$$

For impinging on the chip surface, only the heat transfer of the chip top surface was considered. In the case of water jet, the correlation is then

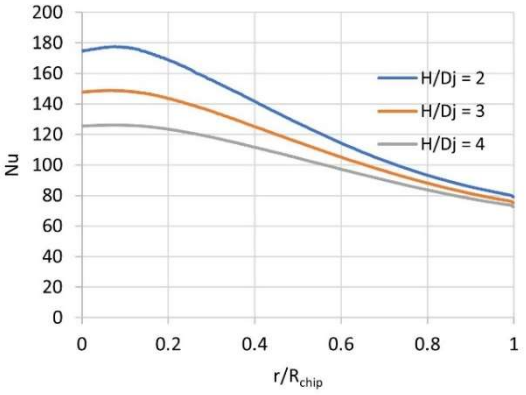
$$Nu_{avg} = 0.1922 Re_{D_j}^{0.7414} \left(H/D_j\right)^{-0.2622} \quad (18)$$

Finally, in the case of the nanofluid jet impinging on a chip, the correlation becomes

$$Nu_{avg} = 0.3247 Re_{D_j}^{0.7368} \left(H/D_j\right)^{-0.2522} \phi^{0.1206} \quad (19)$$



(a) Flat Plate



(b) Chip Plate

Figure 10 Distribution of local Nusselt numbers at different jet-to-plate spacing, Re_{Dj} = 6000, Φ = 4%

4.6 Pumping Power

Fig.11 shows the plots of pumping power ratio between the pumping power of nanofluid (PP_{nf}) and the pumping power of base fluid (PP_{bf}) against Reynolds number at H/D_j = 2. Since nanoparticles cause the viscosity of the base fluid to increase, the pumping power of the nanofluid is obviously higher than that of the base fluid. When ϕ = 2%, the pumping power becomes almost 50% higher than that of the base fluid at all values of Re_{Dj}. The pumping power is increased to be more than twice and almost 4 times of the base fluid, when the concentrations are increased to 4% and 6%, respectively. Therefore, the proper selection of the nanofluid concentration is important for nanofluid jet impingement, especially for experiments.

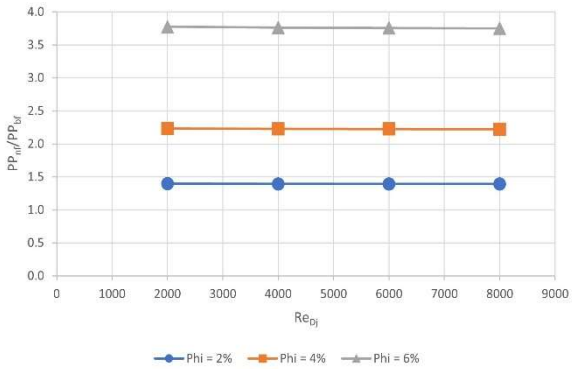


Figure 11 Pumping power ratios of nanofluid and water for chip plate impingement at H/D_j = 2

5. CONCLUSIONS

A numerical heat transfer study on water-based TiO_2 nanofluid impingement on a chip surface and on a flat surface was carried out to investigate the parametric effects such as nanofluid concentration, Reynolds number and jet-to-plate spacing as follows

1. The heat transfer increased with the increasing concentration due to the increasing in thermal conductivity. The maximum heat transfer enhancement in terms of Nusselt number was 24.57% and 18.24% for the flat plate and the chip plate impingement cases, respectively, at ϕ = 6% and Re_{Dj} = 2000.
2. As Reynolds number increased, the heat transfer also increased. This was explained by the fact that at higher Reynolds number the impinging jet thinned the boundary layer and allowed the greater momentum transfer.
3. Due to the counter-rotating vortices being pushed (according to Cornaro et al. (1999) and Manca et al. (2011)), the secondary peaks at the stagnation region were shown at all H/D_j cases. The heat transfer increased with decreasing H/D_j. However, the local Nu distribution showed that the local Nu collapsed after impingement region.

The heat transfer correlations for cases of both flat plate and chip plate impingement for the water and the nanofluid were presented by using the multiple linear regression analysis of log-log scale. After anti-logging, each correlation is summarized as follows

Water-flat plate impingement:

$$Nu_{avg} = 0.0209 Re_{D_j}^{0.8638} \left(H/D_j\right)^{-0.0557}$$

Nanofluid-flat plate impingement:

$$Nu_{avg} = 0.0557 Re_{D_j}^{0.8587} \left(H/D_j\right)^{-0.056} \phi^{0.1458}$$

Water-chip plate impingement:

$$Nu_{avg} = 0.1922 Re_{D_j}^{0.7414} \left(H/D_j\right)^{-0.2622}$$

Nanofluid-flat plate impingement:

$$Nu_{avg} = 0.3247 Re_{D_j}^{0.7368} \left(H/D_j\right)^{-0.2522} \phi^{0.1206}$$

Even though the nanofluid jet impingement has shown a promising heat transfer enhancement, the

penalty in pressure drop in terms of pumping power was also considered. At ϕ = 6%, the pumping power was found almost 4 times of that of water, while at ϕ = 2% the pumping power became 50% of that of water. These data are crucial for heat transfer component designers.

6. REFERENCES

Abdelrehim, O., Khater, A., Mohamad, A.A. & Radwan, A. (2019). Two-phase simulation of nanofluid in a confined single impinging jet. *Case Studies in Thermal Engineering*, 14, 100423. https://doi.org/10.1016/j.csite.2019.100423

Alabdaly, I.K. & Ahmed, M.A. (2019). Numerical investigation on the heat transfer enhancement using a confined slot impinging jet with nanofluid. *Propulsion and Power Research*, 8(4), 351-361. https://doi.org/10.1016/j.jppr.2019.06.004

Barewar, S.D., Tawri, S. & Chougule, S.S. (2019). Heat transfer characteristics of free nanofluid impinging jet on flat surface with different jet to plate distance: An experimental investigation. *Chemical Engineering & Processing: Process Intensification*, 136, 1–10. https://doi.org/10.1016/j.cep.2018.12.001

Choi, S.U.S. & Eastman, J.A. (1995). Enhancing thermal conductivity of fluids with nanoparticles. *ASME International Mechanical Engineering Congress & Exposition, November, 12-17, 1995, San Francisco, CA, USA.* https://www.osti.gov/biblio/196525-enhancing-thermal-conductivity-fluids-nanoparticles

Cornaro, C., Fleischer, A.S. & Goldstein, R.J. (1999). Flow visualization of a round jet impinging on cylindrical surfaces. *Experimental Thermal and Fluid Science*, 20(2), 66-78. https://doi.org/10.1016/S0894-1777(99)00032-1

Huang, J.B. & Jang, J.Y. (2013). Numerical Study of a Confined Axisymmetric Jet Impingement Heat Transfer with Nanofluids. *Engineering*, 5, 69-74. doi:10.4236/eng.2013.51b013

Holzbecher, E., & Hand, S. (2008). Accuracy test for COMSOL – and Delaunay meshes. COMSOL. *Excerpt from the Proceedings of the COMSOL Conference 2008 Hannover.* https://www.comsol.com/paper/accuracy-tests-for-comsol-and-delaunay-meshes-5436

Kirsch, N. (2008, Nov 03). Intel Core i7 920, 940 and 965 Processor Review. Retrieved from http://www.legitreviews.com/intel-core-i7-920-940-and-965-processor-review_824#vkxPL2aZHwzDJ1zl.99

[a]Lv, J., Chang, S., Hu, C., Bai, M., Wang, P. & Zeng, K. (2017). Experimental investigation of free single jet impingement using Al_2O_3-water nanofluid. *International Communications in Heat and Mass Transfer*, 88, 126–135. http://dx.doi.org/10.1016/j.icheatmasstransfer.2017.08.017

[b]Lv, J., Hu, C., Bai, M., Zeng, K, Chang, S. & Gao, D. (2017). Experimental investigation of free single jet impingement using SiO_2-water nanofluid. *Experimental Thermal & Fluid Science, 84*, 39-46. http://dx.doi.org/10.1016/j.expthermflusci.2017.01.010

Manca, O., Mesollela, P., Nardini, S. & Ricci, D. (2011). Numerical study of a confined slot impinging jet with nanofluids. *Nanoscale Research Letters,* 6, 188. http://www.nanoscalereslett.com/content/6/1/188

Modak, M., Srinivasan, S., Garg, K., Chougule, S.S., Agarwal, M.K. & Sahu, S.K. (2015). Experimental investigation of heat transfer characteristics of the surface using Al_2O_3–water nanofluids. *Chemical Engineering & Processing: Process Intensification*, 91, 104–113. http://dx.doi.org/10.1016/j.cep.2015.03.006

Naphon, P. & Nakharintr, L. (2012). Nanofluid jet impingement heat transfer characteristics in the rectangular mini-fin heat sink. *Journal of Engineering Physics and Thermophysics*, 85(6). https://link.springer.com/article/10.1007/s10891-012-0793-8

Roy, G., Gherasim, I., Nadeau, F., Poitras, G. & Nguyen, C.T. (2012). Heat transfer performance and hydrodynamic behavior of turbulent nanofluid radial flows. International Journal of Thermal Sciences 58, 120-129. doi:10.1016/j.ijthermalsci.2012.03.009

Said, Z., Saidur, R., Hepbasli, A. & Rahim, N.A. (2014). New thermophysical properties of water based TiO_2 nanofluid – the hysteresis phenomenon revisited. *International Communications in Heat and Mass Transfer*, 58, 85-95. http://dx.doi.org/10.1016/j.icheatmasstransfer.2014.08.034

Sorour, M.M., El-Maghlany, W.M., Alnakeeb, M.A. & Abbass, A.M. (2019). Experimental study of free single jet impingement utilizing high concentration SiO_2 nanoparticles water base nanofluid. *Applied Thermal Engineering*, *160*, 114019. https://doi.org/10.1016/j.applthermaleng.2019.114019

van Erp, R., Soleimanzadeh, R., Nela, L, Kampitsis, K. & Matioli, E. (2020). Co-designing electronics with microfluidics for more sustainable cooling. *Nature, 585*, 211-216. https://www.nature.com/articles/s41586-020-2666-1

Wei, T., Oprins, H., Cherman, V., Qian, J., Wolf, I.D., Beyne, E.& Baelmans, M. (2019). High-efficiency polymer-based direct multi-jet impingement cooling solution for high-power devices. *IEEE Transaction on Power Electronics*, 34(7), 6601-6612.

Zeitoun, O. & Ali, M. (2012). Nanofluid impingement jet heat transfer. *Nanoscale Research Letters*, 7:139. http://www.nanoscalereslett.com/content/7/1/139

Zhao, Y., Masuoka, T., Tsuruta, T. & Ma, C.F. (2002) Conjugate heat transfer on a horizontal surface impinged by circular free-surface liquid jet. *JSME International Journal series B,* 45(2), 307-314. https://doi.org/10.1299/jsmeb.45.307

NUEJ
Naresuan University Engineering Journal

Tracing Crop Water Requirement in the Pumping, Gravitational and Inundation Irrigation Schemes Using Cloud–Based IrriSAT Application

Paphanin Phutonglom[1], Areeya Rittima[2,*], Yutthana Phankamolsil[3], Allan Sriratana Tabucanon[4], Wudhichart Sawangphol[5], Jidapa Kraisangka[6], Yutthana Talaluxmana[7], and Varawoot Vudhivanich[8]

[1, 2] Graduate Program in Environmental and Water Resources Engineering, Department of Civil and Environmental Engineering, Faculty of Engineering, Mahidol University, Thailand
[3] Environmental Engineering and Disaster Management Program, Mahidol University, Kanchanaburi Campus, Thailand
[4] Faculty of Environment and Resource Studies, Mahidol University, Thailand
[5, 6] Faculty of Information and Communication Technology, Mahidol University, Thailand
[7] Department of Water Resources Engineering, Faculty of Engineering, Kasetsart University, Thailand
[8] Department of Irrigation Engineering, Faculty of Engineering, Kamphaeng Saen, Kasetsart University, Thailand

* Corresponding author e-mail: areeya.rit@mahidol.ac.th

(Received: 26 Nov 2021, Revised: 31 Jul 2022, Accepted: 25 October 2022)

Abstract

Tracing crop coefficient (K_c) at all the stages of crop growth is commonly essential for an accurate estimation of crop water use. This study applied the cloud–based IrriSAT application to trace the dynamic values of crop coefficient in three different sorts of irrigation schemes; pumping, gravitational and inundation irrigation for estimating crop water requirement in the Chao Phraya River Basin (CPYRB), Thailand. Three selected irrigation schemes; Bang Bal (BB), Thabua (TB), and Yom–Nan (YN) representing pumping, gravitational, and inundation irrigation schemes were selected to trace crop coefficient values of in–season and off–season crops and to estimate long–term crop water requirement (ET_c) from 2015–2020. The results of dynamic values of K_c–IrriSAT were verified and adjusted with average K_c established by the Royal Irrigation Department (K_c–RID) which were calculated as a function of K_c from field observation for the different types of crops and accumulated area size monitored by the Geo–Informatics and Space Technology Development Agency (GISTDA). The results revealed the similar patterns of average K_c generated by IrriSAT corresponding to the average K_c–RID. After the calibration procedure was successfully done, the correlations between K_c–IrriSAT adjusted and average K_c–RID for BB, TB, and YN irrigation schemes are relatively higher with R^2 of 0.8304, 0.8466, and 0.8314, respectively. In addition, it shows the explicit variability on monthly and yearly crop water demands of these three sorts of irrigation schemes when the adjusted K_c–IrriSAT was employed. It would be concluded that cloud–based IrriSAT application can be a very supportive tool in estimating the actual crop water requirement particularly for irrigators to evaluate the current status of irrigation water use and to improve the irrigation efficiency at the field scale.

Keywords: Crop Coefficient, Crop Water Requirement, Reference Crop Evapotranspiration, Cloud–Based IrriSAT Application, Normalized Difference Vegetation Index.

1. INTRODUCTION

Thailand is one of the top leaders for agricultural production in the Southeast Asia (BOI, 2021). It is stated that the economic development of Thailand has been predominantly driven by the agricultural sector (Singhapreecha, 2014). Therefore, enhancing agricultural productivity in the large–scale irrigation schemes play an important role to raise livelihood of the local people and to drive the economic growth of the country (Ministry of Agriculture and Cooperatives of the Kingdom of Thailand, 2016). Importantly, water supply facilities and sufficient irrigation water should be provided to farmers in association with the agricultural water demand to increase crop yields.

Increasing water demand for multiple uses in the Chao Phraya River Basin (CPYRB) which is located in the central region of Thailand, has embraced the risk in water resource management specially to satisfy agricultural water demand in the Greater Chao Phraya Irrigation Schemes (GCPYIS). Due to uncertainty of water supply from the headwater of the Chao Phraya River, the extreme events of flood and drought have been frequently occurred in the past few decades (Thanadachophol et al., 2020). Therefore, tracing crop water requirement (ET_c) at all the stages of crop growth in the various sorts of irrigation schemes is necessarily essential to estimate the right amount of crop water demand in GCPYIS.

Crop Water Requirement (CWR) is principally the precise amount of water consumed through evapotranspiration and to meet crop water needs during the specified time periods. In other words, CWR, also known as crop evapotranspiration (ET_c), is described as the depth of water (millimeters) needed to compensate for the water losses through crop evapotranspiration. The main factors affecting crop water requirement are climate factors, crop types, and growth stage of crops. Crop water requirement can be derived based upon the reference crop evapotranspiration (ET_o) and crop coefficient (K_c). The reference evapotranspiration (ET_o) is the rate of evapotranspiration from a hypothetical reference crop which is relatively subject to climate conditions (Pereira & Alves, 2005). The crop coefficient (K_c) varies accordingly with crop types and development stages of the crops. The values of crop coefficient for a given crop have represented the dynamics of crop evapotranspiration (Pandey, 2021).

Since most of the irrigated land area in Thailand are continually cultivated throughout the year, it is rarely possible to clearly determine the beginning of cultivation and to find the dynamic changes of K_c and crop evapotranspiration (ET_c) values over a year. As the satellite–based crop monitoring platform is well proven its capability to monitor crop growth status and vegetation indices, this study aims at tracking the dynamic values of crop water requirement (ET_c) using cloud–based IrriSAT application in the different sorts of irrigation schemes to provide useful information and tools for the analysis of agricultural water requirement in the Chao Phraya River Basin. The cloud–based IrriSAT application is satellite–based irrigation scheduling service developed in 2005. It was designed to help farmers with irrigation management at a wide range of irrigation scales (Hornbuckle et al., 2016). The reference crop evapotranspiration (ET_o) is estimated by FAO Penmen Monteith equation using observations from weather stations. IrriSAT can also anticipate crop water requirement by referring to the strong relations between the Normalized Difference Vegetation Index (NDVI) from the cultivated land and crop coefficient (K_c) (Hornbuckle et al., 2016). Moreover, water balance deficit in the root zone of crops based upon the water balance approach can be traced to indicate the levels of irrigation water requirements.

2. METHODOLOGY

2.1 Study Area

The study area is in the Greater Chao Phraya Irrigation Scheme (GCPYIS) occupying an irrigation service area of more than 19,654 km^2 (12 million rai). GCPYIS lies in the Lower Ping (LPRB), Lower Nan (LNRB) and Chao Phraya–Thachin (CPY–TCRB) River Basins which are major parts of river basin cluster of CPYRB as can be seen in Fig.1–Fig.3. The general characteristics of 35 irrigation schemes in CPYRB are summarized in Table 1. However, only three different sorts of irrigation schemes in GCPYIS representing pumping, gravitational, and inundation irrigation were selected to trace the dynamic values of crop water requirement in this study; Bang Bal (BB), Thabua (TB), and Yom–Nan (YN). Pumping irrigation is powered by pumping system installed at the site to supply the irrigation water to the fields. For gravitational irrigation, the flow irrigation water is directly supplied to the fields through the canals off taking from the headworks. Inundation irrigation, also known as river–canal irrigation, is a type of direct irrigation without construction of hydraulic structures to control the water level in the river.

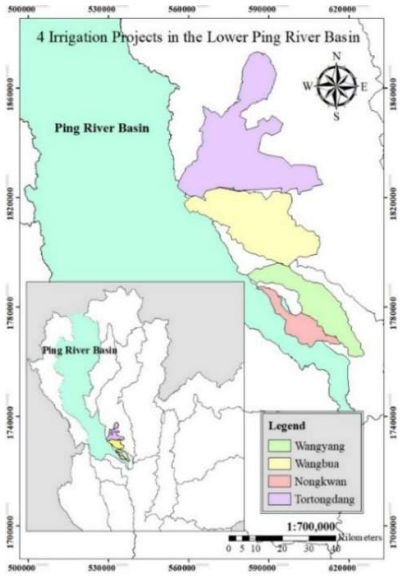


Figure 1 Four irrigation schemes in LPRB

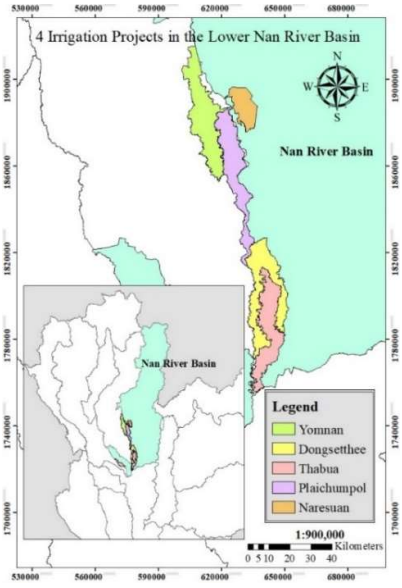


Figure 2 Five irrigation schemes in LNRB

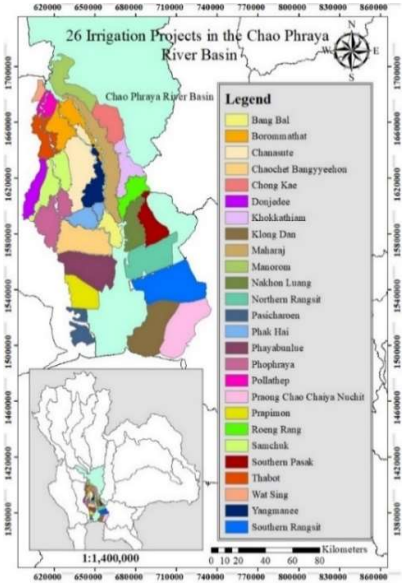


Figure 3 Twenty-six irrigation schemes in CPY–TCRB

Table 1 General characteristics of irrigation schemes in CPYRB

No.	Name of Irrigation Scheme	Type of Irrigation Scheme	Irrigation Area (km^2)
Lower Ping River Basin			
1	Tortongdang	Inundation	992
2	Wangyang	Inundation	1,336
3	Wangbua	Inundation	1,129
4	Nongkwan	Inundation	
Lower Nan River Basin			
1	Dongsetthee	Gravitation	449
2	Thabua*	Gravitation	398
3	Plaichumpol	Gravitation	433
4	Naresuan	Gravitation	172
5	Yom–Nan*	Inundation	515
Chao Phraya–Thachin River Basin			
1	Wat Sing	Pumping	123
2	Bang Bal*	Pumping	268
3	Pollathep	Gravitation	200
4	Thabot	Gravitation	338
5	Samchuk	Gravitation	609
6	Donjedee	Gravitation	270
7	Phophraya	Gravitation	659
8	Borommathat	Gravitation	677
9	Chanasute	Gravitation	880
10	Yangmanee	Gravitation	403
11	Phak Hai	Gravitation	338
12	Maharaj	Gravitation	853
13	Manorom	Gravitation	529
14	Chong Kae	Gravitation	449
15	Khokkathiam	Gravitation	421
16	Roeng Rang	Gravitation	320
17	Southern Pasak	Gravitation	404
18	Nakhon Luang	Gravitation	476
19	Northern Rangsit	Gravitation	773
20	Southern Rangsit	Gravitation	1,112
21	Chaochet Bangyeehon	Gravitation	754
22	Phayabunlue	Gravitation	814
23	Prapimon	Gravitation	426
24	Pasicharoen	Gravitation	337
25	Klong Dan	Gravitation	819
26	Praong Chao Chaiya Nuchit	Gravitation	978
		Total	19,654

As CPYRB is in the tropical climate influenced by northeast and southwest monsoons, rainfall data exhibits the temporal and spatial variation particularly at the river basin level. The mean annual rainfall in the Central Thailand from 2003 to 2007 ranges between 899 and 1,136 millimeters (mm). The mean monthly temperature data evaluated from the long–term record steadily varies between 22^{o}–31^{o}Celsius. Dry season which is a period of low rainfall and off–season crops cultivation is undertaken, runs from November to April. Wet season generally begins in May and lasts in October when in–season crops cultivation is sparsely implemented during time periods. In addition, high values of monthly evaporation loss are obviously found in this region ranging between 120–130 mm for the period 1991–2000 of climatological data reported by the Thai Meteorological Department (TMD).

2.2 Data Collection

Data required for this study includes (1) GIS shapefiles of the study area collected from the Land Development Department (LDD) of Thailand, (2) crop coefficient values provided by the Royal Irrigation Department (RID) of Thailand, (3) planting area of four main types of crops including rice, maize, sugarcane, and cassava monitored by the GISTDA from 2018 to 2019, and (4) climate data collected from the Thai Meteorological Department (TMD) from 2000 to 2020 from the nearest climate stations. The simplified overview of data collection process is presented in Fig.4.

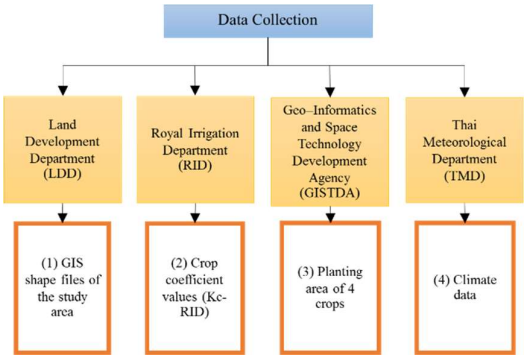


Figure 4 Simplified overview of data collection process

The GIS shapefiles were used to delineate the location of the irrigation area by IrriSAT application. Crop coefficient values (K_c) corresponding to Penman–Monteith equation of major crops, which are publicly provided by the Royal Irrigation Department (RID) from field measurement, was used to validate the results of K_c estimation from IrriSAT. The planting area of four main types of crops monitored by GISTDA using remote sensing technique was used to verify the area size of each crop. In addition, the long–term monthly climate data including atmospheric pressure, temperature, relative humidity, wind speed as well as sunshine duration from 2000 to 2020 from the nearest climate stations in the study area was used to calculate the reference crop evapotranspiration (ET_o).

2.3 Estimating Crop Coefficient Using Cloud–Based IrriSAT Application

Estimating the dynamic values of crop coefficient over the growth stages can be commonly implemented by cloud–based IrriSAT application which is the satellite–based decision support tool for irrigators (Hornbuckle et al., 2016). In fact, cloud–based IrriSAT application has been developed to estimate K_c, ET_o and to predict daily and seven–day crop water use (ET_c) (Hornbuckle et al., 2016). These are presented as aggregated values at various spatial and temporal scales. Due to the limits of global climate data at ground stations in Thailand available on the cloud platform, consequently, ET_o and ET_c cannot be generated and presented. However, the aggregated values of crop coefficient can be only evaluated by IrriSAT. To identify the crop growing area as input data of cloud–based IrriSAT application, the GIS shape files of three irrigation schemes must be converted into Keyhole Markup Language (KML) files as typically illustrated in Fig.5.

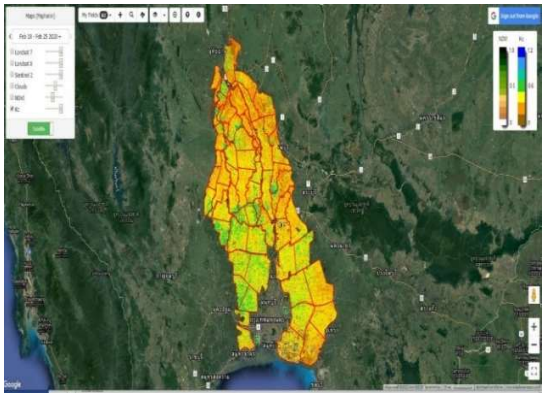


Figure 5 Display of the study area imported in cloud–based IrriSAT application

The maximum, average and minimum crop coefficient (K_c) values were then achieved. The results were automatically displayed in the form of the time series of the crop coefficient according to the specified duration of planting and harvesting dates of crops. However, crop water demand was not directly calculated by IrriSAT in this study. Therefore, the crop water demand (ET_c) for each irrigation area was computed by referring to the calculated reference crop evapotranspiration (ET_o) and crop coefficient performed by IrriSAT (K_c–IrriSAT). ET_o calculator (Raes, 2012) was used as the analytical tool to calculate reference crop evapotranspiration. The chart of crop coefficient generated from IrriSAT is illustrated in Fig.6.

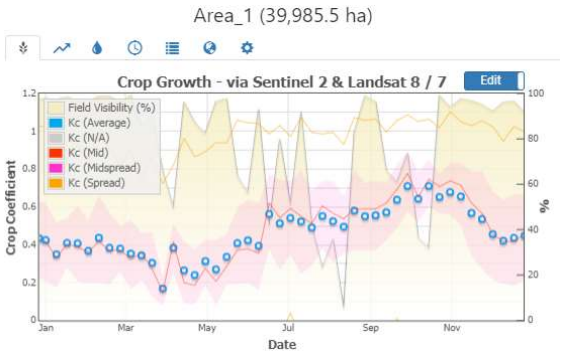


Figure 6 Typical chart of crop coefficient generated by cloud–based IrriSAT application

Cloud–based IrriSAT application was brought to estimate crop coefficient for three main irrigation schemes; BB, TB, and YN from 2015 to 2020. Various forms of crop efficient namely; K_c(average), K_c(observed), K_c(override), K_c(stddev), K_c(min), K_c(Q1), K_c(median), K_c(Q3), and K_c(max) as well as field visibility (%) were accordingly generated. However, only K_c(average) was used to compare with those values of crop coefficient (K_c–RID) performed by using observation data from RID and GISTDA.

2.4 Estimating Average Crop Coefficient (Average Kc–RID) Using Observation Data

The results of dynamic values of K_c–IrriSAT from 2015 to 2020 were verified and adjusted with average K_c–RID which were calculated as a function of K_c from field observation for the different types of crop and accumulated area size of crops monitored using the remote sensing technique (GISagro 4.0) by GISTDA. Due to the limit of GISagro 4.0, the average K_c–RID on the weekly scale can be computed based upon four main types of crop namely; (1) rice, (2) sugarcane, (3) maize, and (4) cassava by using the Eq. (1).

$$\text{Average } K_c\text{–RID} = \frac{(Kcri \times Areari) + (Kcsu \times Areasu) + (Kcmi \times Areami) + (Kcca \times Areaca)}{\text{Total Area}} \quad (1)$$

where, Kcri, Kcsu, Kcmi, Kcca are crop coefficients of rice, sugarcane, maize, and cassava, respectively. Areari, Areasu, Areacmi, Areaca are the accumulated

planting areas of rice, sugarcane, maize, and cassava, respectively.

Therefore, the calculation of average K_c–RID of BB, TB, and YN irrigation schemes was accordingly estimated based upon these four major crops planted from 2018 to 2019 which occupied only 58%, 66%, and 64% of the total cultivated areas, respectively.

2.5 Calibrating Crop Coefficient Values Done by IrriSAT

Calibrating K_c values performed by IrriSAT was conducted using least square criterion to envisage the degree of agreement between K_c–IrriSAT and average K_c–RID and to find the adjusted factors for the specified time periods (Kyaw et al., 2020). The method of least squares is a standard approach in regression analysis to approximate the solution of overdetermined systems by minimizing the sum of the squares of the residuals made in the results (Demaison & Vogt, 2020). In this study, two different periods of planting in–season and off–season crops in the area were identified to compute the adjusted factors of K_c–IrriSAT. These adjusted factors were solved using optimization solver based upon the long–term data sets of K_c–IrriSAT and average K_c–RID.

2.6 Estimating Reference Crop Evapotranspiration (ET_o)

The monthly calculations of reference crop evapotranspiration (ET_o) was implemented based upon the FAO Penman–Monteith equation using ET_o calculator (Saha, 2020). The Penman–Monteith equation requires air temperature, humidity, solar radiation, and wind speed data as key inputs as expressed in Eq. (2).

$$ET_o = \frac{0.408\Delta(R_n - G) + \gamma\frac{900}{T+273}u_2(e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (2)$$

where, ET_o is reference evapotranspiration (MJ m^{-2} day^{-1}), R_n is net radiation at the crop surface (MJ m^{-2} day^{-1}), G is soil heat flux density (MJ m^{-2} day^{-1}), T is the mean air temperature at 2 m height (°C), u_2 is the wind speed at 2 m height (m s^{-1}), e_s is saturation vapor pressure (kPa), e_a is actual vapor pressure (kPa), $e_s - e_a$ is the saturation vapor pressure deficit (kPa), Δ is the slope of the vapor pressure curve (kPa $°C^{-1}$), and γ is the psychometric constant (kPa $°C^{-1}$).

2.7 Calculating Long–Term Crop Water Requirement (ET_c)

The final step is to calculate long–term crop water requirement (ET_c) from 2015–2020 using Eq. (3) (Allen et al., 1998) after the calibration procedure for K_c adjustment and ET_o calculations were successfully done.

$$ET_c = K_c \times ET_o \quad (3)$$

where, ET_c is crop water requirement (mm/period), K_c is crop coefficient done by could–based irrisat application (K_c–irriSAT adjusted) and average K_c–RID, and ET_o is reference crop evapotranspiration (mm/period).

3. RESULTS AND DISCUSSIONS

3.1 Crop Coefficient (K_c) Generated from Cloud–Based IrriSAT Application

The dynamic values of K_c–IrriSAT of three irrigation schemes were generated in many forms from 2015 to 2020 and were displayed in almost one week timeframe. However, only maximum K_c(average) was presented and used to compare the results with average K_c–RID as summarized in Table 2. It is found that the maximum values of K_c(average)–IrriSAT are 0.7019, 0.7997, and 0.7763 for BB, TB, and YN irrigation schemes, respectively which are relatively lower than those received from K_c(average)–RID with 1.4638, 1.4402, and 1.5042, respectively. In addition, the average values of K_c–IrriSAT among these different types of irrigation schemes are in the same range.

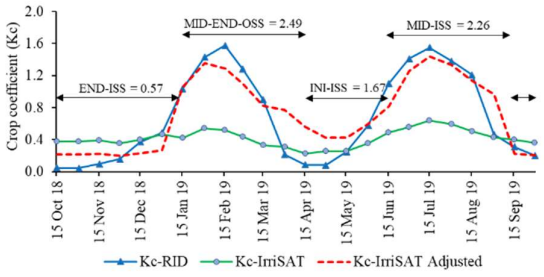
Table 2 Characteristics of crop coefficient values obtained from cloud–based IrriSAT application

Name of Irrigation Scheme	Type of Irrigation Scheme	Max. K_c(avg.)–IrriSAT	Max. K_c(avg.)–RID
BB	Pumping	0.7019	1.4638
TB	Gravitation	0.7997	1.4402
YN	Inundation	0.7763	1.5042

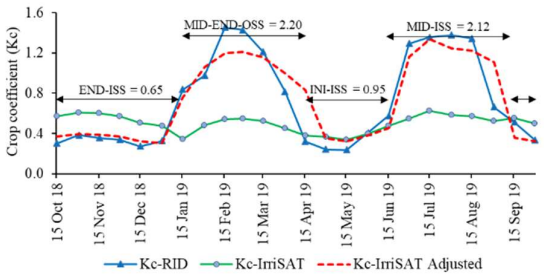
3.2 Relationship between K_c–IrriSAT and average K_c–RID

The relationships between K_c–IrriSAT and average K_c–RID before and after the calibration procedure corresponding to specific growing periods of three main irrigation schemes are presented in Fig.7 and Fig.8. Before calibrating, the patterns of K_c–IrriSAT and average K_c–RID over the growth stages of crops are likely similar. However, the K_c values calculated by IrriSAT for three irrigation schemes are highly deviated from average K_c–RID values in some growing periods in both in–season and off–season crops. It is found that the values of K_c–IrriSAT are higher than average K_c–RID from filed observation in initial and late stages of crop growth in dry and wet seasons for these three irrigation schemes as can be seen in Fig.7. Meanwhile, the lower values of K_c–IrriSAT are found in the mid–stages of dry and wet seasons. The reason might be that evaluating K_c by IrriSAT on cloud–based platform entails the entire planting area. However, calculating average K_c–RID is manipulated based upon some specific types of crop in a given area. Therefore, calibrating K_c values performed by IrriSAT was then conducted using least square criterion to envisage the degree of agreement (R^2) to the average K_c–RID. After calibrating, the adjusted factors

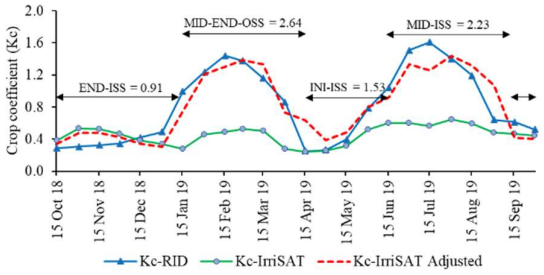
corresponding to time periods identified can be made as presented in Table 3. The comparison of K_c–IrriSAT before and after adjustments with the average K_c–RID are shown in Fig.7. Correlations between K_c–IrriSAT after adjusted and average K_c–RID for BB, TB, and YN irrigation schemes are relatively higher with R^2 of 0.8304, 0.8466, and 0.8314, respectively as can be seen in Fig.8.



(a) Bang Bal



(b) Thabua

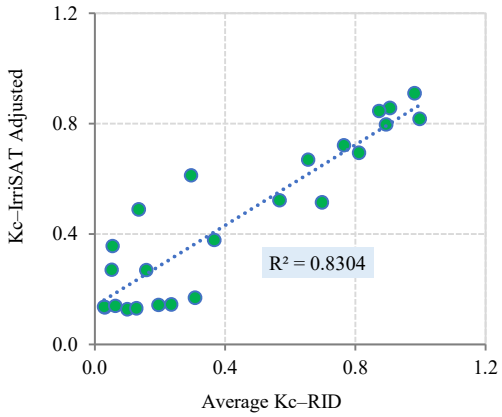


(c) Yom–Nan

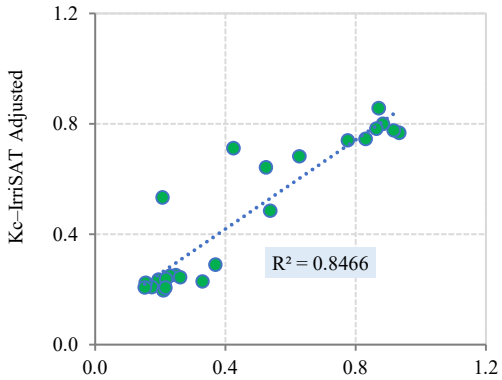
Figure 7 Pattern of K_c value over the growth stages of crops

Table 3 Characteristics of crop coefficient values obtained from cloud–based IrriSAT application

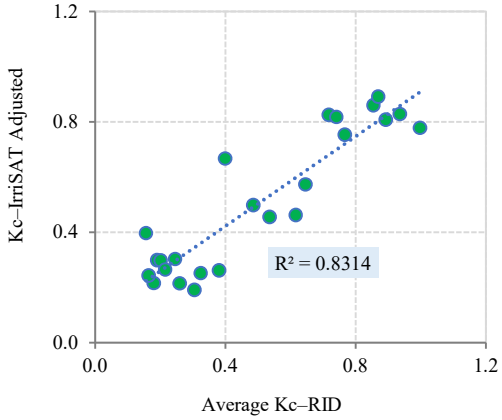
	Adjusted Factors			
Crop	**Off Season Crop (OSS)**	**In Season Crop (ISS)**		
Stage	Mid–End	Initial	Mid	End
Period	1st Jan to 15th Apr	15th Apr to 15th Jun	15th Jun to 1st Sep	1st Sep to 1st Jan
BB	2.49	1.67	2.26	0.57
TB	2.20	0.95	2.12	0.65
YN	2.64	1.53	2.23	0.91



(a) Bang Bal



(b) Thabua



(c) Yom–Nan

Figure 8 Correlation between K_c–IrriSAT adjusted and K_c–RID

3.3 Results of Reference Crop Evapotranspiration (ET_o)

The calculation of the reference crop evapotranspiration (ET_o) for BB, TB, and YN irrigation schemes was accomplished using observed climate data at the nearest weather stations located in Ayutthaya, Phichit–Nakhon Sawan, and Phitsanulok–Sukothai Provinces, respectively.

The results of ET_o calculation are illustrated in Table 4. It is found that the average values of ET_o in the Pichit and Nakhon Sawan Provinces are relatively closer to the Phitsanulok and Sukhothai Provinces due to the similar physical circumstances in the Lower Nan Basin. On the other hand, the values of ET_o calculated using climate data in the Ayutthaya Province in the Chao Phraya–Thachin River Basin seem to be bigger over the year. However, the ranges of ET_o in this region vary between 4.30–6.10, 2.61–5.45, and 2.78–4.37 mm/day for BB, TB, and YN, respectively which is in a similar range comparing with those previously reported by several studies (HII, 2012a; 2012b, 2012c, 2012d; NRCT, 2022).

Table 4 Reference crop evapotranspiration values estimated by the FAO–Penman Monteith formula

Month	ET_o (mm/day)		
	FAO–Penman Monteith Formula		
	BB	TB	YN
Jan	4.69	2.95	2.87
Feb	4.47	3.88	3.37
Mar	5.35	5.00	3.88
Apr	5.39	5.45	4.28
May	5.23	5.08	4.37
Jun	5.29	4.84	4.26
Jul	5.23	4.61	4.19
Aug	4.98	4.33	4.14
Sep	4.31	3.85	3.94
Oct	4.30	3.32	3.54
Nov	4.79	2.84	3.03
Dec	5.29	2.61	2.78

3.4 Comparison of Crop Water Requirement between ET_c–IrriSAT and ET_c–RID

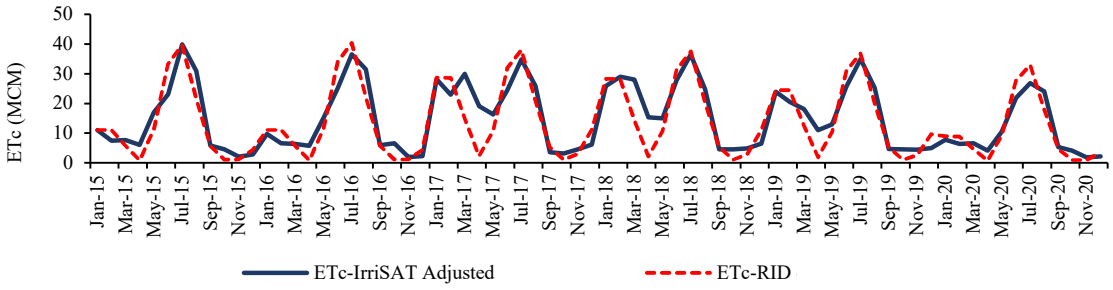
Fig.9 and Table 5 show the calculated values of monthly and yearly crop water demands (ET_c) for three different sorts of irrigation schemes using two kinds of K_c namely, (1) K_c–IrriSAT adjusted and (2) average K_c–RID. It is illustrated that applying the average K_c–RID and adjusted K_c–IrriSAT values provide the similar patterns of the monthly crop water demands from 2015 to 2020 for BB, TB, and YN irrigation schemes. In addition, it shows the equality of mean of the yearly crop water demands of two datasets for these three irrigation schemes. However, when the adjusted K_c–IrriSAT was adopted under the same circumstances of cultivated area size used, the explicit variability on yearly crop water demands of BB, TB, and YN irrigation schemes was found. This is because the cloud–based IrriSAT application can provide the dynamic values of K_c in accordance with the changes in planting area size and NDVI values (Hornbuckle et al., 2016). Among the different sorts of irrigation schemes selected in GCPYIS, the mean values of monthly and yearly crop water demand performed by adjusted K_c–IrriSAT are likely close to those obtained by K_c–RID for all sorts of irrigation. Moreover, the greater variability in the values of monthly and yearly crop water demands made by adjusted K_c–IrriSAT are predominantly found. However, the further study on the performances of IrriSAT in estimating K_c and ET_c values and relation between K_c and Normalized Difference Vegetation Index (NDVI) in the different sorts of irrigation is highly encouraged for the achievement of satellite–based crop water requirement estimation.

4. CONCLUSION

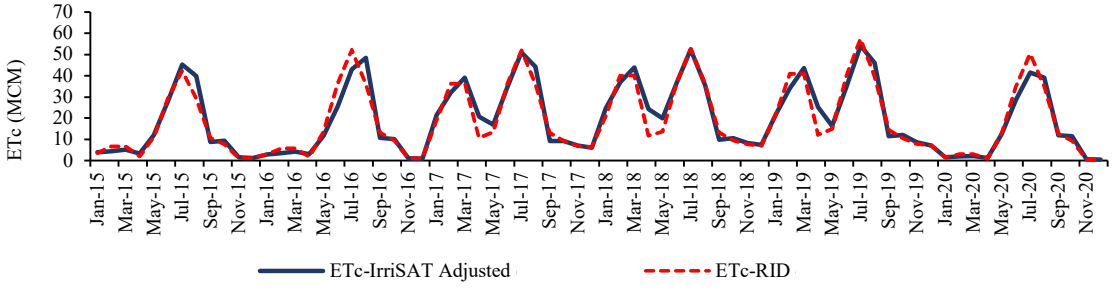
The cloud–based IrriSAT application can be a very supportive tool for tracking the dynamics of crop coefficient which is a key parameter for accurate estimation of crop water use. This study revealed the application of IrriSAT in estimating actual crop water demand promptly on cloud–based platform in various types of irrigation schemes in Thailand. In addition, the calibration procedures to find the adjusted factors of dynamic crop coefficients estimated by IrriSAT were also envisaged. The results show that the cloud–based IrriSAT application can deliver the explicit variability on monthly and yearly crop water demands in these three sorts of irrigation schemes which represent the pumping, gravitational and inundated irrigation in GCPYIS. In addition, it can be used for crop water demand estimation particularly in small to large scale irrigation areas. Importantly, it is very helpful for the water resources planners to identify affordable water delivery and to improve the irrigation efficiency at the field scale corresponding to the dynamic values of estimated crop water demand and water supply status.

5. ACKNOWLEDGMENT

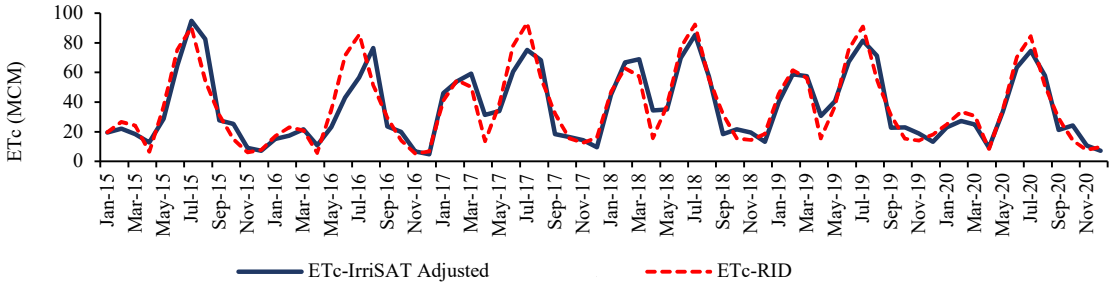
Authors would like to acknowledge the Thailand Science Research and Innovation (TSRI) for providing financial support. We are cordially grateful to the Royal Irrigation Department (RID) of Thailand and the Thai Meteorological Department (TMD) for providing research data.



(a) Bang Bal



(b) Thabua



(c) Yom–Nan

Figure 9 Patterns of monthly ET_c values from 2015 to 2020 in each irrigation project

Table 5 Yearly crop water requirement values from 2015 to 2020

Year	Crop Water Requirement, ETc (MCM)					
	Bang Bal (BB)–Pumping		Thabua (TB)–Gravitation		Yom–Nan (YN)–Inundation	
	Kc–IrriSAT Adjusted	Average K_c–RID	Kc–IrriSAT Adjusted	Average K_c–RID	Kc–IrriSAT Adjusted	Average K_c–RID
2015	157.83	146.23	163.34	151.90	413.24	394.47
2016	153.20	148.46	165.41	178.92	319.60	366.20
2017	218.61	195.91	291.79	275.60	487.18	501.78
2018	222.21	193.22	310.28	289.91	535.16	526.37
2019	190.35	179.70	315.26	307.44	525.43	516.58
2020	121.00	120.62	153.39	165.03	376.51	398.45
Avg.	177.20	164.02	233.25	228.13	422.85	450.64
SD.	40.06	30.19	79.94	70.11	86.92	71.70
Var.	1,604.65	911.53	6,390.70	4,914.90	7,554.81	5,141.48

6. REFERENCES

Allen, R., Wright, J.L., Pruitt, W.O., Pereira, L., & Jensen, M. (1998). Water Requirements. Design and Operation of Farm Irrigation Systems 1998, 29, DOI: 10.1007/978–94–011–5131–3_2.

Board of Investment of Thailand (BOI). (20201). Thailand Advantages: Trade Spotlight. Retrieved from https://www.boi.go.th/index.php?page=thailand_advantages.

Demaison, J., & Vogt, N. (2020). Least–Squares Method. Accurate Structure Determination of Free Molecules, 233–263. DOI: 10.1007/978–3–030–60492–9_9.

Hornbuckle, J., Vleeshouwer, J., Ballester, C., Montgomery, J., Hoogers, R., & Bridgart, R. (2016). Estimating the crop coefficient. IrriSAT Technical Reference, 1–17.

Hydro–Informatics Institute (HII). (2012a). 25 Basins Database Development Project: Ping Basin. Bangkok: HII.

Hydro–Informatics Institute (HII). (2012b). 25 Basins Database Development Project: Nan Basin. Bangkok: HII.

Hydro–Informatics Institute (HII). (2012c). 25 Basins Database Development Project: Chao Phraya Basin. Bangkok: HII.

Hydro–Informatics Institute (HII). (2012d). 25 Basins Database Development Project: Thachin Basin. Bangkok: HII.

Kyaw, K. M., Rittima, A., Phankamolsil, Y., Tabucanon, A. S., Sawangphol, W., Kraisangka, J., Talaluxmana, Y., & Vudhivanich, V. (2020). Tracing Crop Water Demand in the Lower Ping River Basin, Thailand using Cloud–Based IrriSAT Application. Proceedings of the 22nd IAHR–APD Congress 2020, September 14–17, Sapporo, Japan.

Ministry of Agriculture and Cooperatives of the Kingdom of Thailand (MOAC). (2016). Economic characteristics of Thailand. Agricultural Development Plan No. 12, Thailand, 12–81.

National Research Council of Thailand (NRCT). (2022). Multiple Reservoir Re–Operation System for Long–Term Water Supply Management in the Greater Chao Phraya River Basin Using Artificial Intelligence Technique (Phase 2). Bangkok: NRCT.

Pandey, V. (2021). Estimation of Crop Water Requirement. Fertigation Technologies for Micro Irrigated Crops, Apple Academic Press, 109–136. DOI: 10.1201/9781003084136–10.

Pereira, L.S., & Alves, I. (2005). Crop Water Requirement. Encyclopedia of Soils in the Environment, 322–334. DOI: 10.1007/s12205–013–1638–5.

Raes, D. (2012). The ETo Calculator: Reference Manual Version 3.2. Food and Agriculture Organization of the United Nations, Italy.

Saha, A. (2020). How to use FAO developed ETo Calculator to estimate evapotranspiration from weather data. Determination of Hydro–geological Properties for Upper Soil Layer Regarding Irrigation design and Groundwater Recharge in Sylhet region, DOI: 10.13140/RG.2.2.27723.34080.

Singhapreecha, C. (2014). Economy and Agriculture in Thailand. Faculty of Economics, Chiang Mai University.

Thanadachophol, T., Teamsuwan, V., Udomsap, L., Wongsamut, W., & Chamnankaew, U. (2020). An Analysis of Water Management in 2020 Drought Conditions, Thailand, 287–288.

7. BIOGRAPHIES

Paphanin Phutonglom received her bachelor degree in Civil Engineering in 2019 and master degree in Environmental and Water Resources Engineering in 2022 from Faculty of Engineering, Mahidol University, Thailand. She is currently working at the Office of the National Water Resources.

Areeya Rittima received her doctoral degree in Irrigation Engineering from Kasetsart University, Thailand in 2006. She is currently an associate professor at the Department of Civil and Environmental Engineering, Faculty of Engineering, Mahidol University, Thailand. Her research interests include application of Artificial Intelligence (AI) for dam–reservoir operation system, water allocation concept, and conjunctive water use.

Yutthana Phankamolsil received the doctoral degree in Irrigation Engineering from Kasetsart University, Thailand in 2008. Currently, he is an assistant professor in the Environmental and Disaster Management Program, Kanchanaburi Campus, Mahidol University, Thailand. His research interests are hydro–informatic, socio–hydrology, agent–based modeling, micro–irrigation control.

Allan Sriratana Tabucanon received his Ph.D. in Urban Engineering from the University of Tokyo, Japan. Currently, he is an assistant professor at the Faculty of Environment and Resource Studies, Mahidol University, Thailand. His research interests include climate change impact prediction, 2D urban flood simulation, flood damage modelling and formulation of flood alleviation plan in tangible and intangible.

Wudhichart Sawangphol received the Master of Information Technology (MIT) with honors in Data Management, Software Engineering, and Knowledge Engineering and Ph.D. in Ontology Reasoning and Optimization from Monash University, Australia in 2013 and 2017. Currently, he is a lecturer at the Faculty of Information and Communication Technology, Mahidol University, Thailand. His research interests are automated reasoning, ontology, ontology reasoning, optimization, machine learning, deep learning, and data visualization.

Jidapa Kraisangka received the Ph.D. in Information Science from the University of Pittsburgh, USA in 2019. Currently, she is an instructor at the Faculty of Information and Communication Technology, Mahidol University, Thailand. Her research interests are probabilistic and decision–theoretic methods in decision support systems, clinical decision support system machine learning, and data visualization.

Wudhichart Yutthana Talaluxmana received his doctoral degree in Irrigation Engineering from Kasetsart University, Thailand in 2013. Currently, he is an assistant professor at the Department of Water Resources Engineering, Faculty of Engineering, Kasetsart University, Thailand. His research interests include water resources planning and management, hydraulics and water resources development.

Varawoot Vudhivanich received his doctoral degree in Civil Engineering (Water Resources Planning and Management) from Colorado State University, USA in 1986. He is currently an associate professor and senior expert at the Department of Irrigation Engineering, Faculty of Engineering at Kamphaeng Saen, Kasetsart University, Thailand. His research interests include hydrology, hydraulic and irrigation and water management.

NUEJ
Naresuan University
Engineering Journal

Evaluating Hydroelectricity Production Re-operating with Adapted Rule Curve Under Climate Change Scenarios: Case Study of Bhumibol Dam in Thailand

Khin Muyar Kyaw[1], Areeya Rittima[2]*, Yutthana Phankamolsil[3], Allan Sriratana Tabucanon[4], Wudhichart Sawangphol[5], Jidapa Kraisangka[6], Yutthana Talaluxmana[7], and Varawoot Vudhivanich[8]

[1, 2*] Graduate Program in Environmental and Water Resources Engineering, Department of Civil and Environmental Engineering, Faculty of Engineering, Mahidol University, Thailand
[3] Environmental Engineering and Disaster Management Program, Mahidol University, Kanchanaburi Campus, Thailand
[4] Faculty of Environment and Resource Studies, Mahidol University, Thailand
[5, 6] Faculty of Information and Communication Technology, Mahidol University, Thailand
[7] Department of Water Resources Engineering, Faculty of Engineering, Kasetsart University, Thailand
[8] Department of Irrigation Engineering, Faculty of Engineering at Kamphaeng Saen, Kasetsart University, Thailand

*Corresponding author's e–mail: areeya.rit@mahidol.ac.th

(Received: 2 August 2022, revised: 12 October 2022, Accepted: 25 October 2022)

Abstract

Hydroelectricity production is being impacted by climate change due to the considerable changes in water availability of reservoir system and dam release. This study aims at evaluating the response of energy production of the Bhumibol dam through the reservoir re–operation system with non–engineering adaptation measures due to climate change. Re–operating the Bhumibol (BB) dam with adapted rule curve and modelling exercise with MIKE11 to predict series of reservoir inflow were conducted under RCP4.5 and RCP8.5 climate change scenarios. The adapted rule curves of BB dam were established by either increasing or lowering the upper and lower rule levels of 0.5 meters from the rule curves which were developed by EGAT in 2012. In addition, the standard operation policy was applied to specify the amount of water release corresponding to the adapted rule curves established. The water balance–based reservoir re–operation model was developed using MATLAB Simulink Toolbox for short–term simulation from 2012 to 2018. Influence of climate change on the seasonal and yearly reservoir inflows were considerably investigated. In addition, the relation of current and projected inflows, and the response of dam release and hydroelectric production of BB dam were then evaluated. The results of the short–term simulation from 2012 to 2018 show that dam release is likely to be increased corresponding to the high variability of projected inflows. Therefore, the seasonal and yearly hydroelectricity production are accordingly increased when re–operating dam under RCP4.5 and RCP8.5 inflows. It is found that the yearly hydroelectricity production with RCP4.5 and RCP8 inflows are about 52% and 30% respectively higher than the current inflow. It is also revealed that re–operating dam with the different types of adapted rule curves does not alter the volume of released water and energy production generated from the reservoir radically because the standard operating rules were adopted for all adapted rule curves. Importantly, the study on the adaptation measures to climate change would help increase understanding of necessity of new operational rules for dam and reservoir re–operation to cope well with instability of reservoir water supply for sustainable hydropower production in future.

Keywords: Adapted Rule Curve, Climate Change, Hydroelectricity, Water Balance–based Simulation Model.

1. INTRODUCTION

Great attention has been paid to climate change for few decades as it has caused serious impact on water resources management and responses to the natural disasters. The climate change has drastically altered the frequency and intensity of extreme rainfall creating large uncertainty of water availability in hydrologic cycle and occurrences of unprecedented huge floods and prolonged droughts (Kundzewicz et al., 2014). Moreover, the changes in rainfall patterns and streamflow discharges may adversely impact the efficiency of hydropower generation (ASEAN Development Bank, 2012; Goyal & Surampalli, 2018). Therefore, the dam and reservoir operation practices have to be re–examined and efficiently improved through adaptation measures based upon the rational and up–to–date information. For hydroelectric dam, the non–engineering adaptation measures such as re–operating reservoir with the new operating rules, improving the predictability of hydrological data, and developing the localized regional climate model to suggest the operational changes in reservoir management, can be very useful and need to be more explored for applications (Jia, 2016).

Most of the reservoirs in Thailand have been designed to serve multiple water uses such as irrigation, municipality, ecology, and hydropower generation. As the demand for electricity has become more competitive due to a massive increase in population and human being, a large number of small to large hydropower plant development projects has been established by aiming to effectively support the electricity requirement and to manage the renewable energy sources (Aroonrat & Wongwises, 2015; Punarai et al., 2015). Therefore, the attempt to re–operate the dam and reservoir system under climate change and to evaluate the response to energy production was made in this study by selecting Bhumibol dam as the study area.

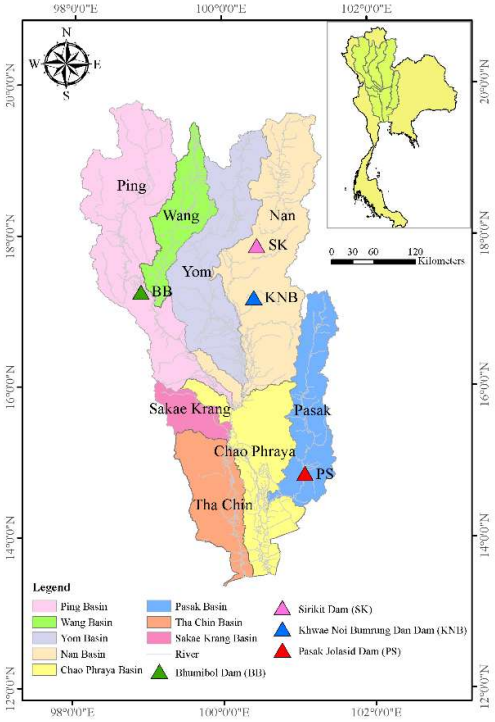


Figure 1 Bhumibol dam in the Greater Chao Phraya River Basin (GCPYRB)

The Bhumibol (BB) dam is a multipurpose dam built across the Ping River in the northern elevated plain of Thailand as shown in Fig.1. It is the major source of water supply in the Greater Chao Phraya River Basin (GCPYRB). GCPYRB is termed for this study to describe the principal river basin cluster for cooperative management of water resources in the central region of Thailand covering Ping, Wang, Yom, Nan, Pasak, Sakae Krang, and Tha Chin River basins with the total area of 138,977 km^2. The Bhumibol dam has been jointly operated with Sirikit (SK) and Khwae Noi Bumrung Dan (KNB) dams not only to supply water for national domestic, agricultural, industrial uses, as well as the ecological needs downstream, but also for hydropower production to supply electric power in the nearby region. The hydropower plant of Bhumibol dam has been built since 1961 and enlarged to the total installed capacity of 779.2 MW in 1991 with the average energy of 1,037 GWHr/yr (Thongthamchart & Raphitphan, 2016). The Electricity Generating Authority of Thailand (EGAT) is a key institutional operator responsible for dam and reservoir operation and hydroelectricity production.

In this study, the relation of reservoir inflow, dam release and hydroelectric production of BB dam which is altered by the climate change impact and the consequences of reservoir operation policy undertaken, are considerably investigated through the development of reservoir re–operation model using the adapted rule curves. The adapted rule curves were generated by increasing and lowering the levels of upper rule curve (URC) and lower rule curve (LRC) of ±0.5 m from existing rule curve established by EGAT in 2012 (Kyaw et al., 2022). The long–term response of energy production among the current and future climate change scenarios were examined and tested using statistical procedure to exhibit the statistical differences of obtained energy as the results of reservoir re–operations and climate change.

2. METHODOLOGY

In this study, the projection of reservoir inflow of BB dam was carried out through the platform of MIKE11 Zero. MIKE11 RR NAM Model and MIKE11 HD, the physically–lumped model, were adopted for rainfall–runoff simulation into Ping River Basin. Prediction of rainfall and evaporation under climate change scenarios was implemented based on the simulation of EC–EARTH under RCP4.5 and RCP8.5 scenarios regionally downscaled by RegCM4 with 25 km x 25 km grid size over the study area (Tabucanon et al., 2021). The projected inflow of BB dam was then generated for five periods namely, 2000–2020 (baseline), 2021–2040, 2041–2060, 2061–2080 and 2081–2099. However, the baseline period was only used for this study due to time duration of the available historical record. The water balance–based reservoir re–operation model was then developed using MATLAB Simulink Toolbox, as typically shown in Fig.2.

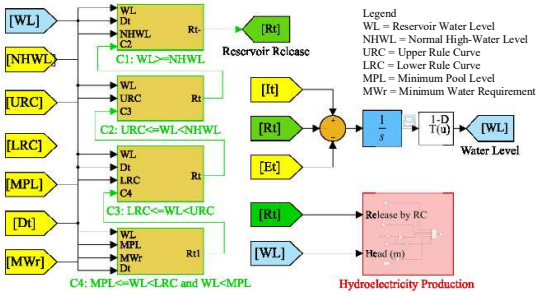


Figure 2 Water balance–based reservoir re–operation model using adapted rule curve

To emphasize the response of energy production when altering the reservoir operating policy under impact of climate changes, the adapted rule curves of BB dam which were modified from the rule curve established by EGAT in 2012, were used. The daily short–term simulation based upon the associated reservoir data were then performed from 2012 to 2018 under current and projected inflows. Identifying the water demand for BB dam was referred to yearly water allocation plan to reach all the water use sectors in GCPYRB. Water supplied to the target demand nodes was shared by BB and SK dams in the proportion of 0.44:0.56 (Tabucanon et al., 2021).

In the final step, the potential hydropower energy under current and future RCP 4.5 and RCP8.5 inflows were investigated to envisage the statistical difference of obtained energy by using ANOVA statistical procedure (St & Wold, 1989).

2.1 Adapted Rule Curve–Based Reservoir Re–operation Model

The reservoir rule curve is regarded as the common tool to provide useful guidance in the decision–making process of dam release. In this study, the adapted rule curve–based reservoir re–operation model for BB dam was established based on the existing rule curve developed in 2012 by EGAT. To accomplish the modelling practice for reservoir re–operation, four main scenarios of adapted rule curve were generated by increasing and lowering the water storage levels of upper rule curves (URC) and lower rule curves (LRC) of ±0.5 meters from the existing rule curve. The reservoir operating policies were set up in accordance with the standard operating policy (SOP) (Neelakantan & Sasireka, 2013; Kangrang et al., 2018) aiming to reduce the risk of unmanageability of reservoir operation system to handle with water deficit and flood while maximizing potential hydroelectricity production over the specific time periods. The water release from the reservoir is specified as the same amount of target demand when the reservoir storage is between LRC and URC. However, the water can only be released with the minimum water requirement of 5 million cubic meter per day (MCM/day) when the reservoir storage is lower than LRC. In case of reservoir storage is higher than URC, the reservoir water release is accordingly based on the conditions of excessive water and maximum turbine discharge of the hydropower system. The total amount of the water storage above normal high–water level (NHWL) is specified as a spilled water when the water level is above NHWL. Moreover, the reservoir water could not be released in case of the reservoir storage is lower than minimum pool level (MPL) as shown in Fig.3.

2.2 Estimation of Hydroelectricity Production

To simulate hydroelectricity production as a result of water release when re–operating with adapted rule curve, the water balance–based approach (Carvajal et al., 2017). was applied in accordance with the current and projected inflows as shown in Eq. (1).

$$S_{t+1} = S_t + I_t - E_t - R_t \quad (1)$$

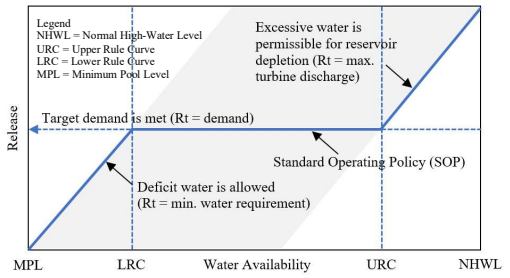


Figure 3 Application of standard reservoir operating policy with rule curves

where S_{t+1} represents the water storage of the reservoir at time step t+1; S_t is the initial storage of the reservoir at time step t; I_t is the reservoir inflow volume at time step t; E_t is the evaporation loss from the reservoir at time step t; and R_t is the water release volume or the reservoir outflow discharging into the hydropower turbines. S_t is constrained by the maximum water storage (S_{max}) and minimum water storage (S_{min}), which dynamically changes due to the associated reservoir data and amount of water release specified by adapted rule curves. The relationship between water storage and water level can be found using reservoir water level–area–capacity curve. R_t is determined by the water balance–based reservoir operation model developed using MATLAB Simulink Toolbox. Finally, daily hydropower electricity output of Bhumibol dam can be computed using Eq. (2).

$$E = \eta \rho g Q H t \quad (2)$$

where, E is the daily electricity output in kilowatts hour (KWHr). η is the overall efficiency of the hydropower plant in percentage, ρ is the water density (1,000 kg/m^3), g is the acceleration of gravity which is 9.81 m/s^2. Q represents the reservoir outflow discharging through the hydropower turbines (m^3/s). H is the hydraulic head (m) which can be calculated by the difference in height between the headwater level in the reservoir and tail water level downstream of the dam. t is the number of working hours for power generation over the specified time periods.

2.3 Statistical Analysis of Hydroelectricity Production Using ANOVA–Test

The energy production performed by the existing and adapted rule curves under the current and projected inflows were tested using one–way analysis of variance (ANOVA) test to determine whether there is a significant difference between the means of two energy dataset (Davis & Mukamal, 2006; Raftery et al., 1995; Labovitz,

1970). In the other words, ANOVA–test, which is a parametric test of difference, was used to describe whether impact of climate change has an effect on the potential energy production at level of significance (α) of 0.05.

3. Results and Discussion

3.1 Influence of Climate Change on the Seasonal and Yearly Reservoir Inflows

To emphasize on the influence of climate change on the seasonal and yearly inflows of Bhumibol dam, the long–term recorded inflow and projected inflow performed by MIKE11 were investigated and compared. It is found from recorded inflow that the average yearly inflow of the BB reservoir from 1969 to 2018 is approximately 5,694 MCM/yr. More than 80% of the total inflow is contributed to BB dam in wet season (May–Oct, WS) and the remaining is occurred in dry season (Nov–Apr, DS). Table 1 and Fig.4 also indicate that the long–term projected yearly inflow into BB dam under RCP4.5 and RCP8.5 tends to be decreased predominantly for all the specific time periods except in 2041–2060. In comparison with the baseline period (2000–2020), RCP4.5 scenario exhibits the increase in inflow in dry season (Nov–Apr) by +0.07%, +10.00%, +15.42% and +6.25% in 2021–2040, 2041–2060, 2061–2080 and 2081–2099, respectively. However, the opposite results are obviously found in wet season (May–Oct) as the change are expected to be –10.44%, +9.60%, –13.01% and –2.63%. For RCP8.5 case, the potential increase in inflow in dry season (Nov–Apr) is +8.14%, +8.15% and +22.71% in 2041–2060, 2061–2080 and 2081–2099, respectively except in 2021–2040 period showing the decrease in inflow by –5.03% deviated from the baseline period. Similarly, the percentage change in projected inflow in wet season (May–Oct) is fluctuated around the baseline by –4.68%, +20.17%, –10.13% and +18.04% in 2041–2060, 2061–2080 and 2081–2099, respectively. The results indicate considerable variations in the seasonal and yearly patterns of reservoir inflow which are key factor influencing the complexity and effectiveness of reservoir management in both the current and future operations.

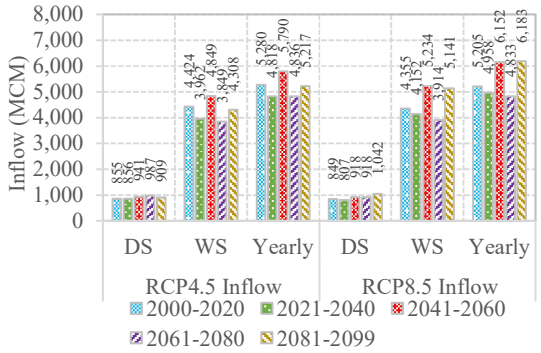


Figure 4 Projected changes in average seasonal and yearly reservoir inflows of Bhumibol dam

3.2 Response of Reservoir Water Release When Re–operating with Adapted Rule Curves under Climate Change Scenarios

As aforementioned, the average long–term yearly inflow of the Bhumibol dam is projected to significantly decrease under climate change scenarios (RCP4.5 and RCP8.5). However, in this study the simulation run of reservoir re–operation model was implemented using the short–term dataset from 2012 to 2018 since the rule curves has been established in 2012 by EGAT. It is noticeable that the average values of projected inflows from 2012 to 2018 are 6,358 and 5,168 MCM/yr for RCP4.5 and RCP8.5, respectively which are much higher than the current inflow of 4,269 MCM/yr as shown in Fig.5. These average values of projected inflows from 2012 to 2018 appear largely inconsistent with the average values of 5,280 and 5,205 MCM for RCP4.5 and RCP8.5 from 2000 to 2020, and 5,694 MCM from historical record from 1969 to 2018 due to the occurrences of extreme flow events in two consecutive years in this region creating severe flood in 2011 and handling flood risk in 2012. This reflects that climate change has significantly created the serious impact on the large variation of reservoir inflow for both short– and long–term data. The increase in projected inflows from 2012 to 2018 could lead to the increase in the seasonal and yearly release volumes from reservoir as expressed in Table 2 since determination of amount of reservoir release is substantially subject to the total inflow. The Fig.5 and Fig. 6 show strong correlation between the reservoir inflow and water releases when existing rule curve and 4 scenarios of adapted rule curves were employed. It exhibits that releasing water from BB dam is not only based on the conditions of available water storage and target water demand but also significantly subject to the extent of incoming inflows. Therefore, for the RCP4.5 and RCP8.5 inflow cases, the seasonal and yearly releases are relatively higher than water release for the current inflow case. It is also found that the released water in dry season is higher than in wet season for all scenarios. However, altering the operating rule curves by increasing and lowering the levels of ±0.5 meters while the same

standard rules were applied, shows the slight differences in terms of seasonal and yearly release volumes from the reservoir system. Moreover, monthly water shortage is not existed for the current and projected inflows scenarios when re–operating with adapted rule curves from 2012 to 2018. Importantly, even the huge amount of water release in wet season obtained from RCP4.5 and RCP8.5 inflow scenarios is profoundly noticed, however, the maximum values of dam releases in wet season are 3,018 and 2,234 MCM for RCP4.5 and RCP8.5 inflow scenarios, respectively which are definitely associated with the water allocation plan specified in normal operation periods.

Table 1 The recorded and projected reservoir inflows of Bhumibol dam for the specified time periods

Month	Recorded Inflow (MCM)	Projected Inflow under RCP4.5 (MCM)					Projected Inflow under RCP8.5 (MCM)				
Periods	1969–2018	2000–2020[1/]	2021–2040	2041–2060	2061–2080	2081–2099	2000–2020[1/]	2021–2040	2041–2060	2061–2080	2081–2099
Jan	132.20	117.25	116.49	125.56	127.85	130.11	118.11	112.31	132.20	136.68	133.59
Feb	59.08	71.31	69.96	75.19	72.42	73.30	71.08	69.09	80.57	77.40	82.20
Mar	38.02	60.30	55.07	65.30	140.95	55.70	57.73	52.50	61.79	58.17	62.97
Apr	48.12	56.24	81.80	72.12	76.07	65.92	65.53	46.95	54.80	47.89	105.81
May	247.46	227.58	328.21	403.26	281.83	264.96	243.73	240.15	264.25	106.10	339.54
Jun	317.08	425.75	288.30	657.90	356.29	465.86	403.38	355.65	429.98	251.00	444.47
Jul	379.35	428.09	378.96	414.31	342.98	428.75	440.60	278.99	455.13	299.64	548.93
Aug	923.49	726.06	584.56	707.69	672.31	651.17	563.78	576.64	1,302.83	662.55	994.77
Sep	1,504.04	1,351.93	1,185.17	1,420.61	1,061.18	1,342.31	1,418.73	1,207.00	1,461.26	1,243.92	1,438.88
Oct	1,209.42	1,264.98	1,197.26	1,245.40	1,134.31	1,154.85	1,285.25	1,493.23	1,320.32	1,351.16	1,374.54
Nov	590.14	352.97	348.38	398.26	359.56	385.30	345.57	332.92	382.38	397.67	420.61
Dec	245.24	197.31	184.27	204.51	210.40	198.50	191.29	192.82	206.67	200.65	236.97
Yearly	5,693.64	5,279.77	4,818.43	5,790.10	4,836.15	5,216.72	5,204.77	4,958.25	6,152.17	4,832.84	6,183.27
Δ%			(–8.74)	(+9.67)	(–8.40)	(–1.19)		(–4.74)	(+18.20)	(–7.15)	(+18.80)
DS	1,112.80	855.37	855.99	940.93	987.26	908.83	849.30	806.59	918.40	918.47	1,042.14
Δ%			(+0.07)	(+10.00)	(+15.42)	(+6.25)		(–5.03)	(+8.14)	(+8.15)	(+22.71)
WS	4,580.84	4,424.40	3,962.44	4,849.17	3,848.89	4,307.89	4,355.47	4,151.66	5,233.77	3,914.36	5,141.13
Δ%			(–10.44)	(+9.60)	(–13.01)	(–2.63)		(–4.68)	(+20.17)	(–10.13)	(+18.04)

Remark: [1/] baseline period

Table 2 Seasonal and yearly water release when re–operating with adapted rule curves from 2012 to 2018

Types of Inflow	Reservoir Water Release (MCM)														
	Existing Rule Curve			(+0.5 m) URC			(–0.5 m) URC			(+0.5 m) LRC			(+0.5 m) LRC		
	DS	WS	Yearly	DS	WS	Yearly	DS	WS	Yearly	DS	WS	Yearly	DS	WS	Yearly
Current	2,632	1,702	4,334	2,632	1,702	4,334	2,632	1,700	4,332	2,632	1,700	4,332	2,632	1,703	4,335
RCP 4.5	3,008	3,019	6,027	3,018	2,985	6,003	3,000	3,049	6,048	3,009	3,018	6,027	3,021	3,006	6,027
Δ%	(+14)	(+77)	(+39)	(+15)	(+75)	(+39)	(+14)	(+79)	(+40)	(+14)	(+77)	(+39)	(+15)	(+77)	(+39)
RCP 8.5	2,869	2,224	5,093	2,859	2,222	5,082	2,851	2,261	5,111	2,861	2,234	5,094	2,861	2,234	5,094
Δ%	(+9)	(+31)	(+18)	(+9)	(+31)	(+17)	(+8)	(+33)	(+18)	(+9)	(+31)	(+18)	(+9)	(+31)	(+18)

Remark: [1/] the different values compared with the current inflow case

Table 3 Seasonal and yearly hydroelectricity production when re–operating with adapted rule curves from 2012 to 2018

Types of Inflow	Hydroelectricity Production (GWhr)														
	Existing Rule Curve			(+0.5 m) URC			(–0.5 m) URC			(+0.5 m) LRC			(+0.5 m) LRC		
	DS	WS	Yearly	DS	WS	Yearly	DS	WS	Yearly	DS	WS	Yearly	DS	WS	Yearly
Current	660	377	1,037	660	377	1,037	659	376	1,035	660	377	1,036	660	377	1,037
RCP 4.5	788	789	1,577	793	782	1,575	784	794	1,578	788	789	1,577	791	786	1,577
Δ%	(+19)	(+109)	(+52)	(+20)	(+108)	(+52)	(+19)	(+111)	(+52)	(+19)	(+109)	(+52)	(+20)	(+109)	(+52)
RCP 8.5	781	568	1,349	780	569	1,349	775	576	1,351	779	571	1,349	778	570	1,348
Δ%	(+18)	(+51)	(+30)	(+18)	(+51)	(+30)	(+17)	(+53)	(+30)	(+18)	(+52)	(+30)	(+18)	(+51)	(+30)

Remark: [1/] the different values compared with the current inflow case

Table 4 Results of ANOVA test for the daily energy performed various reservoir re–operation models

Re–operation	Diff.	Source of Variation	Sum of Square	Degree of Freedom	Mean Squared	F	p–value	F critical
Existing Rule Curve	$\Delta E^{1/}$	Between Groups	2798714108	1	2.799E+09	151.8326	2.11E–34	3.8433
		Within Groups	94228936748	5112	18432891			
		Total	97027650856	5113				
	$\Delta E^{2/}$	Between Groups	936810177.7	1	936810178	79.3998	6.94E–19	3.8433
		Within Groups	60314713786	5112	11798653			
		Total	61251523963	5113				
(+0.5 m) URC	$\Delta E^{1/}$	Between Groups	2776119189	1	2.776E+09	150.2820	4.51E–34	3.8433
		Within Groups	94432634805	5112	18472738			
		Total	97208753994	5113				
	$\Delta E^{2/}$	Between Groups	935875162.3	1	935875162	79.3168	7.23E–19	3.8433
		Within Groups	60317542813	5112	11799206			
		Total	61253417975	5113				
(–0.5 m) URC	$\Delta E^{1/}$	Between Groups	2824975555	1	2.825E+09	153.5799	9E–35	3.8433
		Within Groups	94031007217	5112	18394172			
		Total	96855982772	5113				
	$\Delta E^{2/}$	Between Groups	954884620.2	1	954884620	80.6869	3.65E–19	3.8433
		Within Groups	60497651904	5112	11834439			
		Total	61452536524	5113				
(+0.5 m) LRC	$\Delta E^{1/}$	Between Groups	2800368148	1	2.8E+09	151.9171	2.03E–34	3.8433
		Within Groups	94232199964	5112	18433529			
		Total	97032568113	5113				
	$\Delta E^{2/}$	Between Groups	938147770.1	1	938147770	79.5401	6.47E–19	3.8433
		Within Groups	60294268970	5112	11794654			
		Total	61232416740	5113				
(–0.5 m) LRC	$\Delta E^{1/}$	Between Groups	2799519087	1	2.8E+09	151.9133	2.03E–34	3.8433
		Within Groups	94205960025	5112	18428396			
		Total	97005479113	5113				
	$\Delta E^{2/}$	Between Groups	927441951.2	1	927441951	78.9799	8.55E–19	3.8433
		Within Groups	60029006055	5112	11742763			
		Total	60956448006	5113				

Remark: $\Delta E^{1/}$ = the difference between current inflow and projected RCP4.5 case
$\Delta E^{2/}$ = the difference between current inflow and projected RCP8.5 case
F = F–statistic value representing how much the variability among the means exceeds that expected one
F critical = value of the F–statistic at the threshold probability α of mistakenly rejecting a true null hypothesis

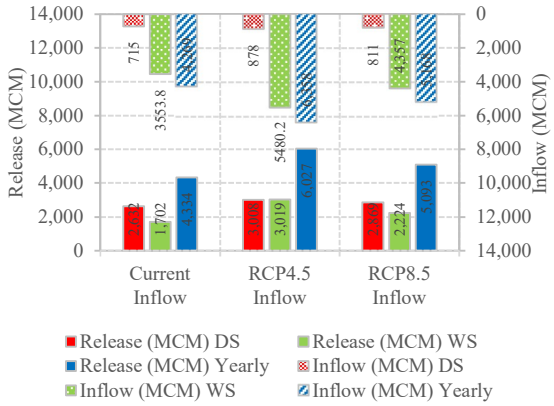


Figure 5 Relation of average reservoir water release with the current and projected inflows from 2012 to 2018 when existing rule curve was employed

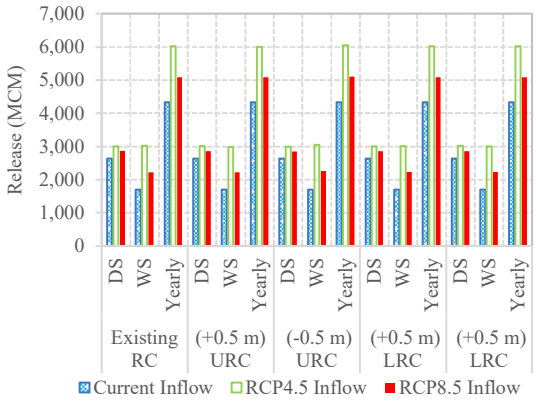


Figure 6 Average seasonal and yearly water releases when re–operating with adapted rule curve under RCP4.5 and RCP8.5 inflows from 2012 to 2018

3.3 Response of Hydroelectricity Production When Re–operating with Adapted Rule Curves under Climate Change Scenarios

In principle, the potential hydroelectricity production is relatively subject to the volumetric released water, hydraulic heads and overall efficiency of hydropower plant. As the dam release from 2012 to 2018 is likely to be increased due to the high variability of projected inflow, therefore, the seasonal and yearly hydroelectricity production are obviously increased when re–operating dam under RCP4.5 and RCP8.5 inflows. The yearly energy amounts to 1,037 GWhr for the current inflow when re–operating reservoir with the existing rule curve. By comparing with the current inflow case, the yearly energy is expected to increase of 52% and 30% under RCP4.5 and RCP8.5 scenarios when the existing rule curve is employed. However, there is no significant differences when changing the operating policy by the adapted rule curve as expressed in Table 3 and Fig. 7. The response of hydraulic head for hydropower generation was also investigated in this study as shown the results in Fig.8. It is revealed that the average values of monthly hydraulic head under RCP4.5 and RCP8.5 scenarios are higher than the current inflow case, leading to the increasing in energy production over the simulation time periods.

3.4 ANOVA Test for Hydroelectricity Production under Climate Change Scenario

The impact of climate change on the potential energy production of BB dam was re–diagnosed through the analysis of ANOVA–test to compare the means of energy when re–operating reservoir under the current and projected inflows. The results in Table 4 obviously show a significant difference between the means of two energy dataset at level of significance (α) of 0.05 when existing and adapted rule curves were employed. This implies that the changes in the extent of reservoir inflow and volume of released water due to impact of climate change would influence to the potential energy production in future.

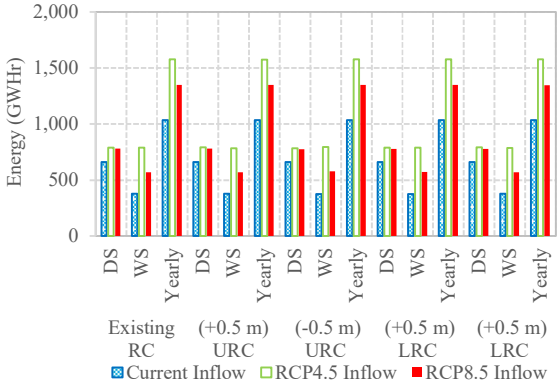


Figure 7 Average seasonal and yearly hydropower generation when re–operating with adapted rule curve under climate change scenarios

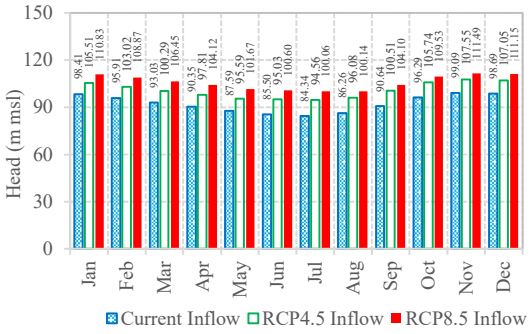


Figure 8 Response of average hydraulic head for hydropower generation under climate change scenarios

4. CONCLUSION

Hydroelectric generation is sensitively affected by the changes in water availability of the reservoir system and dam release. Based on the short–term simulation of BB dam from 2012 to 2018, dam release is likely to be increased corresponding to the high variability of projected inflow. Therefore, the seasonal and yearly hydroelectricity production are significantly increased when re–operating dam under RCP4.5 and RCP8.5 inflows. It is also revealed that re–operating dam with the different types of adapted rule curves does not alter the volume of released water and energy production generated from the reservoir radically because the standard operating rules were adopted for all adapted rule curves. However, the further study on the adaptation measures to climate change would help increase understanding of necessity of new operational rules for dam and reservoir re–operation to cope well with instability of reservoir water supply in future.

5. ACKNOWLEDGMENT

Authors would like to acknowledge the Thailand Science Research and Innovation (TSRI) for providing financial support. We are thankful to the Electricity Generating Authority of Thailand (EGAT), Thailand Meteorological Department (TMD), and the Royal Irrigation Department (RID) for providing research data.

6. REFERENCES

Aroonrat, K., & Wongwises, S. (2015). Current status and potential of hydro energy in Thailand: A review. *Renewable and Sustainable Energy Reviews, 46*, 70–78. https://doi.org/10.1016/j.rser.2015.02.010

ASEAN Development Bank. (2012). Climate risk and adaptation in the electric power sector. Retrieved from http://hdl.handle.net/11540/887.

Carvajal, P. E., Anandarajah, G., Mulugetta, Y., & Dessens, O. (2017). Assessing uncertainty of climate change impacts on long–term hydropower generation using the CMIP5 ensemble–the case of Ecuador. *Climatic Change, 144*(4), 611–624.

Davis, R. B., & Mukamal, K. J. (2006). Hypothesis testing: means. *Circulation, 114*(10), 1078–1082. https://doi.org/10.1161/CIRCULATIONAHA.105.586461

Goyal, M. K., & Surampalli, R. Y. (2018). Impact of climate change on water resources in India. *Journal of Environmental Engineering, 144*(7), 04018054. http://doi.org/10.1061/(ASCE)EE.1943–7870.0001394

Jia, J. (2016). A technical review of hydro–project development in China. *Engineering, 2*(3), 302–312. https://doi.org/10.1016/J.ENG.2016.03.008

Kangrang, A., Prasanchum, H., & Hormwichian, R. (2018). Development of future rule curves for multipurpose reservoir operation using conditional genetic and tabu search algorithms. *Advances in Civil Engineering*, Article ID 6474870, 1–10. http://dx.doi.org/10.1155/2018/6474870

Kundzewicz, Z. W., Kanae, S., Seneviratne, S. I., Handmer, J., Nicholls, N., Peduzzi, P., & Sherstyukov, B. (2014). Flood risk and climate change: global and regional perspectives. *Hydrological Sciences Journal, 59*(1), 1–28. https://doi.org/10.1080/02626667.2013.857411

Kyaw, K. M., Rittima, A., Phankamolsil, Y., Tabucanon, A. S., Sawangphol, W., Kraisangka, J., Talaluxmana, Y & Vudhivanich, V. (2022). Assessing Reservoir Reoperation Performances through Adapted Rule Curve and Hedging Policies under Climate Change Scenarios: In–depth Investigation of Case Study of Bhumibol Dam in Thailand. *Engineering Access, 8*(2), 179–185. http://doi: 10.14456/mijet.2022.23

Labovitz, S. (1970). The nonutility of significance tests: The significance of tests of significance reconsidered. *Pacific Sociological Review*, 13*(3), 141–148.* https://doi.org/10.2307%2F1388411

Neelakantan, T. R, & Sasireka, K. (2013). Hydropower reservoir operation using standard operating and standard hedging policies. *International Journal of Engineering and Technology, 5*(2), 1191–1196. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.412.1094&rep=rep1&type=pdf

Punarai, N., Khumkaew, S., & Kluabwang, J. (2015). Application of differential evolution to parameter identification of Bhumibol Hydro Power Plant in Thailand. *Proceedings of the International MultiConference of Engineers and Computer Scientists.*

Raftery, A. E., Gilks, W. R., Richardson, S., & Spiegelhalter, D. (1995). Hypothesis testing and model. *Markov chain Monte Carlo in practice*, 165–187.

St, L., & Wold, S. (1989). Analysis of variance (ANOVA). Chemometrics and intelligent laboratory systems, 6(4), 259–272. https://doi.org/10.1016/0169–7439(89)80095–4

Tabucanon, A. S., Rittima, A., Raveephinit, D., Phankamolsil, Y., Sawangphol, W., Kraisangka, J., Talaluxmana, Y., Vudhivanich, V., & Xue, W. (2021). Impact of climate change on reservoir reliability: a case of Bhumibol Dam in Ping River Basin, Thailand. *Environment and Natural Resources Journal*, *19*(4), 266–281. http://dx.doi.org/10.32526/ennrj/19/2021012

Thongthamchart, C., & Raphitphan, N. (2016). Performance of Bhumibol concrete arch dam after 50 years of operation. EGAT, Thailand. *In: Proceedings for ASIA*. https://www.academia.edu/25900993/Performance_of_Bhumibol_concrete_arch_dam_after_50_years_of_operation

7. BIOGRAPHIES

Khin Muyar Kyaw was born in Bago city, Myanmar in 1993. She received the B.E degree in Civil Engineering from Technological University (Toungoo), Myanmar in 2014 and M.E degree in Civil Water Resources Engineering from Yangon Technological University, Myanamar in 2019. She is currently a Ph.D. candidate at the Faculty of Engineering, Mahidol University, Thailand. Her current research interests include application of Artificial Intelligence (AI) for dam–reservoir operation system and impact of climate changes on dam–reservoir operation.

Areeya Rittima received her doctoral degree in Irrigation Engineering from Kasetsart University, Thailand in 2006. She is currently an associate professor at the Department of Civil and Environmental Engineering, Faculty of Engineering, Mahidol University, Thailand. Her research interests include application of Artificial Intelligence (AI) for dam–reservoir operation system, water allocation concept, and conjunctive water use.

Yutthana Phankamolsil received the doctoral degree in Irrigation Engineering from Kasetsart University, Thailand in 2008. Currently, he is an assistant professor in the Environmental and Disaster Management Program, Kanchanaburi Campus, Mahidol University, Thailand. His research interests are hydro–informatic, socio–hydrology, agent–based modeling, micro–irrigation control.

Allan Sriratana Tabucanon received his Ph.D. in Urban Engineering from the University of Tokyo, Japan. Currently, he is an assistant professor at the Faculty of Environment and Resource Studies, Mahidol University, Thailand. His research interests include climate change impact prediction, 2D urban flood simulation, flood damage modelling and formulation of flood alleviation plan in tangible and intangible.

Wudhichart Sawangphol received the Master of Information Technology (MIT) with honors in Data Management, Software Engineering, and Knowledge Engineering and Ph.D. in Ontology Reasoning and Optimization from Monash University, Australia in 2013 and 2017. Currently, he is a lecturer at the Faculty of Information and Communication Technology, Mahidol University, Thailand. His research interests are automated reasoning, ontology, ontology reasoning, optimization, machine learning, deep learning, and data visualization.

Jidapa Kraisangka received the Ph.D. in Information Science from the University of Pittsburgh, USA in 2019. Currently, she is an instructor at the Faculty of Information and Communication Technology, Mahidol University, Thailand. Her research interests are probabilistic and decision–theoretic methods in decision support systems, clinical decision support system machine learning, and data visualization.

Wudhichart Yutthana Talaluxmana received his doctoral degree in Irrigation Engineering from Kasetsart University, Thailand in 2013. Currently, he is an assistant professor at the Department of Water Resources Engineering, Faculty of Engineering, Kasetsart University, Thailand. His research interests include water resources planning and management, hydraulics and water resources development.

Varawoot Vudhivanich received his doctoral degree in Civil Engineering (Water Resources Planning and Management) from Colorado State University, USA in 1986. He is currently an associate professor and senior expert at the Department of Irrigation Engineering, Faculty of Engineering at Kamphaeng Saen, Kasetsart University, Thailand. His research interests include hydrology, hydraulic and irrigation and water management.

NUEJ
Naresuan University
Engineering Journal

Drought Analysis in the Eastern Economic Corridor by using the Standardized Precipitation Index (SPI)

Polpech Samanmit, Jutithep Vongphet* and Bancha Kwanyuen

Department of Irrigation Engineering, Faculty of Engineering at Kamphaeng saen, Kasetsart University, Nakhon Pathom, Thailand

* Corresponding author e-mail: fengjtv@ku.ac.th

(Received: 26 July 2022, Revised: 27 October 2022, Accepted: 03 November 2022)

Abstract

This research aimed to analyze the severity level of drought by using the Standardized Precipitation Index (SPI) in the Eastern Economic Corridor (EEC). The six meteorological stations, Bangkok Metropolis, Chanthaburi, Chonburi/Sattahip, Koh Sichang, Pattaya, and Khao Ito, over 30 years data from 1988 to 2017, were used in this research. The three patterns of moving cumulative rainfall of 3, 6, and 12 months were used to estimate the indices of SPI3, SPI6, and SPI12, respectively. As a result, the averaged values of 3 patterns of SPI ranged from 0.035 to 0.180, which was interpreted as a level of near normal. The maximum and minimum values of the SPI could experience drought conditions in both severe to extreme drought and wet. However, the results demonstrated that more than 62 % of the SPI ranged near normal. The value of SPIs approximated moderate to extreme droughts below 17 %. The results of this research could be a part of the information to support water management in the study area to develop a suitable plan to mitigate the effect of drought in the future.

Keywords: Drought, Eastern Economic Corridor, Rainfall, Standardized Precipitation Index.

1. INTRODUCTION

Drought was a natural disaster in Thailand that hugely impacted and damaged various sectors. During the past 20 – 30 years (1989 – 2013), many provinces suffered in drought throughout the country (Wichitarapongsakun et al.,2016). The drought process affected water shortage in many activities, such as water supply, agriculture, and industry, for a long time (Damberg and AghaKouchak, 2014). Drought could be divided into four categories: meteorological, hydrological, agricultural, and socioeconomic droughts; each drought was related to and affected each other (Wilhite and Glantz, 1985). Thailand was prone to droughts in two periods: (1) winter to summer in the second half of October onwards in a drought manner and (2) in the middle of the rainy season (late June to July) in a precipitation deficit (Hydro-Informatics Institute, 2012, Thai Meteorological Department, 2016). The drought resulted from the northeast and southwest seasonal monsoons. In addition, the Enzo phenomenon caused climate variability in precipitation deficit that affected water supply problems (Thanapakpawin et al., 2011). The severity level of drought can be assessed from several drought indices, which were widely used in foreign countries and Thailand, such as Effective Drought Index (EDI), Generalized Monsoon Index (GMI), Standardized Precipitation Index (SPI), Normalized Difference Vegetation Index (NDVI), Palmer Drought Severity Index (PDSI), Etc. (Thai Meteorological Department, 2012, Pinthong and Kwanyuen, 2016, Adisa et al., 2021). The Standardized Precipitation Index (SPI) was one of the most widespread drought indices used to analyze the severity level of drought abroad, such as the analysis of drought levels in South Africa with SPI and EDI indices (Adisa et al., 2021), comparison of drought indices between SPI and SPEI in Mongolia china (Pei et al., 2020), analysis of future drought indices in RCP8.5 cases with SPI and EDI in the Langat basin, Malaysia (Huang et al., 2016), Etc. Including the analysis of drought indices in Thailand, such as the analysis of drought levels with the SPI in the Sakae Krang river basin (Wichitarapongsakun et al., 2016,Wichitarapongsakun et al.,2017), the lower Mae Klong river basin (Madusanka and Venkataramana, 2017), the Huai Ko Kaeo sub-basin of the Pasak river basin (Wichitarapongsakun, 2015), Ping River Basin (Chaito and Khamkong, 2020), 15 areas of the north (Kaingam and Chotamonsak, 2019), comparison of SPI, SPEI and SPAEI analysis results in the Chi river basin (Homdee et al., 2016), Etc. The SPI could be analyzed over several periods from 1 to 72 months. The meaning of SPI3 and SPI6 are used for the analysis of drought that occurs during the crop season, and SPI12 is used for the study of change in drought during the year (Mckee et al., 1993, Thai Meteorological Department, 2012, Sriwongsitanon, 2015).

The Eastern Economic Corridor (EEC) is an essential area for the economic development of Thailand. This study aims to support water resource management between surface water and groundwater (Conjunctive use) integrated into the water management system. To be suitable for water supply and demand according to the drought conditions in the study area That is established to enable Thailand to develop economically into a high-income country and the developed countries according to the 20-year National Strategic by focusing on enhancing the competitiveness of the manufacturing and service sectors in the industrial sector and increasing the capacity of reservoirs and allocating water to support the increasing trend of water demand (The Eastern Economic Corridor Office of Thailand, 2018).

The research objective was to analyze the severity level of drought using the Standardized Precipitation Index (SPI) based on six meteorological stations covering the area of the Eastern Economic Corridor in 3 patterns, including SPI3, SPI6, and SPI12. The purpose of this research was to study the nature, severity, and frequency of droughts in Thailand in the study area for 30 years (1988 – 2017), which will be one of the components of information supporting water management such as rainfall, runoff, water demand, water balance, Etc.

2. DATA AND METHOD

2.1 Study Area

Eastern Economic Corridor (EEC) is located in eastern Thailand, which is established to enable Thailand to develop economically into a high-income country and the developed countries according to the 20-year National Strategic by focusing on enhancing the competitiveness of the manufacturing and service sectors in the industrial sector and increasing the capacity of reservoirs and allocating water to support the increasing trend of water demand (The Eastern Economic Corridor Office of Thailand, 2018). The EEC had an area of approximately 13,266 sq. km and covered three provinces: Chachoengsao, Chonburi, and Rayong. (Figure 1)

2.2 Rainfall data

Observed rainfall data in this research were collected from six meteorological stations of the Thailand Meteorological Department (TMD). That was the neighborhood of the three provinces in the area of the EEC to cover the context of the drought situation covering the study area, consisting of Bangkok Metropolis, Chanthaburi, Chonburi/Sattahip, Koh Sichang, Pattaya, and Khao Ito (Table 1 and Figure 1), during the years 1998 – 2017 (30 years). The Double-mass curve method was used to verify the consistency of cumulative annual rainfall at six stations (Wichitarapongsakun et al., 2016, Taesombat, 2012).

The average annual rainfall of the six meteorological stations ranged between 1,110.17 – 2,959.62 mm/year. The maximum value was 4,028.06 mm/year in 2006 at the Chanthaburi station. The minimum value was 708.23 mm/year in 1993 at the station of Pattaya.

2.3 Standardized Precipitation Index (SPI)

The standardized Precipitation Index (SPI) was developed to analyze the severity level of drought based on cumulative rainfall for interest (Mckee et al., 1993). The SPI could be studied over several periods from 1 month to 72 months, allowing water supply estimates from short to long periods. The SPI period was selected with sufficient rainfall to affect the five types of water resources: soil moisture, Groundwater level, stream, reservoir storage, and snow. A short period of 3 and 6 months was recommended for the growing season, and a more extended period of 12 and 24 months for the changing of drought during the year.

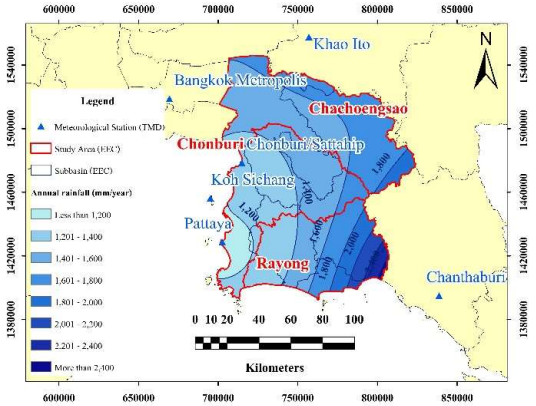


Figure 1 Study area and Average annual rainfall

Generally, precipitation had a Gamma Distribution function (Mckee et al., 1993). However, since the determination of the SPI requires cumulative rainfall, it was determined using the Cumulative probability Density Function (CDF), then transformed to a standard normal (Z) with a mean of 0 and variance of 1. The SPI can analyze the severity level of drought or humidity of rainfall in each area. For equations (1) to (6) show in this paper and more detail, refer to reference reports and research articles (Thai Meteorological Department, 2012, Sriwongsitanon, 2015).

The cumulative probability Density Function (CDF) is Shown as Equation **(**1).

$$G(x) = \frac{1}{\Gamma(\hat{\alpha})}\int_0^x t^{\hat{\alpha}-1}e^{-t}dt \qquad (1)$$

In a month where no rain was reported When x = 0, the cumulative density function must be converted to Equation (2).

$$H(x) = q + (1-q)G(x) \qquad (2)$$

Where q is the number of non-rainy days (m) divided by the number of days studied (n), and G(x) is the Cumulative probability Density Function (CDF).

The SPI index can be calculated from Equation **(3)** and Equation (4) as follows:

$$Z = SPI = -\left(t - \frac{c_0+c_1t+c_2t^2}{1+d_1t+d_2t^2+d_3t^3}\right) \qquad (3)$$

Where $0 < H(X) \leq 0.5$

$$Z = SPI = +\left(t - \frac{c_0+c_1t+c_2t^2}{1+d_1t+d_2t^2+d_3t^3}\right) \qquad (4)$$

Where $0.5 > H(X) < 1$

where (t) can be obtained from Equation **(5)** and Equation (6).

$$t = \sqrt{\ln\left(\frac{1}{(H(x))^2}\right)} \qquad (5)$$

Where $0 < H(X) \leq 0.5$

$$t = \sqrt{\ln\left(\frac{1}{(1.0-H(x))^2}\right)} \qquad (6)$$

Where $0.5 > H(X) < 1$

Where $C_0 = 2.515517$, $C_1 = 0.802853$, $C_2 = 0.010328$, $d_1 = 1.432788$, $d_2 = 0.189269$ and $d_3 = 0.001308$

Table 1 Meteorological stations data and Average annual rainfall for 30 years (1988 – 2017)

Stations ID	Stations name	Latitude	Longitude	Mean	Maximum	Minimum
484550	Bangkok Metro.	13.73333	100.56667	1,730.50	2,783.18	985.77
484800	Chanthaburi	12.60000	102.11667	2,959.62	4,028.06	1,509.01
484590	Chonburi/Sattahip	13.36667	100.98333	1,367.00	2,110.24	836.17
484600	Koh Sichang	13.16667	100.80000	1,162.46	1,682.17	865.28
484610	Pattaya	12.91667	100.86667	1,110.17	1,655.96	708.23
484300	Khao Ito	14.07776	101.37988	1,735.55	2,581.20	1,189.48

The percentage of drought occurrences of SPI equations (3) and (4) can be classified as the drought classification level criterion, as shown in Table 2.

Table 2 Drought classification according to the SPI

Category	SPI
Extreme drought	< -2.00
Severe drought	-1.50 to -1.99
Moderate drought	-1.00 to -1.49
Near normal	-0.99 to 0.99
Moderate wet	1.00 to 1.49
Severe wet	1.50 to 1.99
Extreme wet	> 2.00

Source: Mckee et al. (1993)

This research analyzed the SPI with 3 patterns based on the cumulative monthly rainfall of 3, 6, and 12 months to be SPI3, SPI6, and SPI12, respectively, with 6 meteorological stations from the years 1988 to 2017 (30 years).

3. RESULTS AND DISCUSSION

Drought analysis of SPI with three different periods was illustrated by Table 3 The mean SPI values of all meteorological stations ranged from 0.035 to 0.180, interpreted as near normal. At the same time, the maximum and minimum values ranged from 1.920 to 3.127 and -4.182 to -1.529, indicated as severe to extreme wet and drought, respectively. For example, the maximum values of SPI3 were 2.916 at Koh Sichang, and SPI6 and SPI12 were 2.648 and 3.127, respectively, at the station of Chanthaburi. The minimum values of SPI3, SPI6 and SPI12 were -3.744, -4.182 and -2.613 at the stations of Khao Ito, Bangkok Metropolis and Pattaya, respectively.

Table 4 to Table 6 demonstrated the possibility of drought severity in the percentage of the whole simulation time, and Figure 2 to Figure 4 showed drought incidence in terms of both wet (> 0.99) and drought (< -0.99) over 30 years (1988 – 2017), with a nonspecific pattern. The SPI3 significantly indicated the possibility of drought to near normal as the percentage of 68.99 – 71.51 (%) at all meteorological stations. In contrast, moderate, severe and extreme droughts ranged from 5.31 – 11.17 (%), 3.07 – 4.19 (%), 0.84 – 3.35 (%), respectively. As same as, SPI 6 also interpreted the possibility of drought to near normal with the percentage 67.89 -72.39 (%), while moderate, severe and extreme droughts ranged from 4.51 - 10.99 (%), 2.25 – 4.79 (%), 0.56 - 3.10 (%), respectively. Finally, the SPI12 also interpreted almost

Table 3 Classification of SPI by meteorological stations

Stations	SPI3			SPI6			SPI12		
	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min
Bangkok Metropolis	0.058	2.137	-3.528	0.068	2.279	-4.182	0.081	2.014	-2.011
Chanthaburi	0.035	2.637	-2.778	0.070	2.648	-2.300	0.067	3.127	-1.553
Chonburi/Sattahip	0.067	2.385	-2.661	0.081	1.920	-2.676	0.090	2.058	-2.026
Koh Sichang	0.091	2.916	-3.449	0.130	2.621	-2.543	0.180	2.211	-1.529
Pattaya	0.065	2.419	-3.074	0.085	2.408	-2.976	0.121	1.963	-2.611
Khao ITo	0.059	2.559	-3.744	0.079	2.291	-2.030	0.089	2.215	-2.163

Table 4 Percentage of SPI3 by meteorological stations

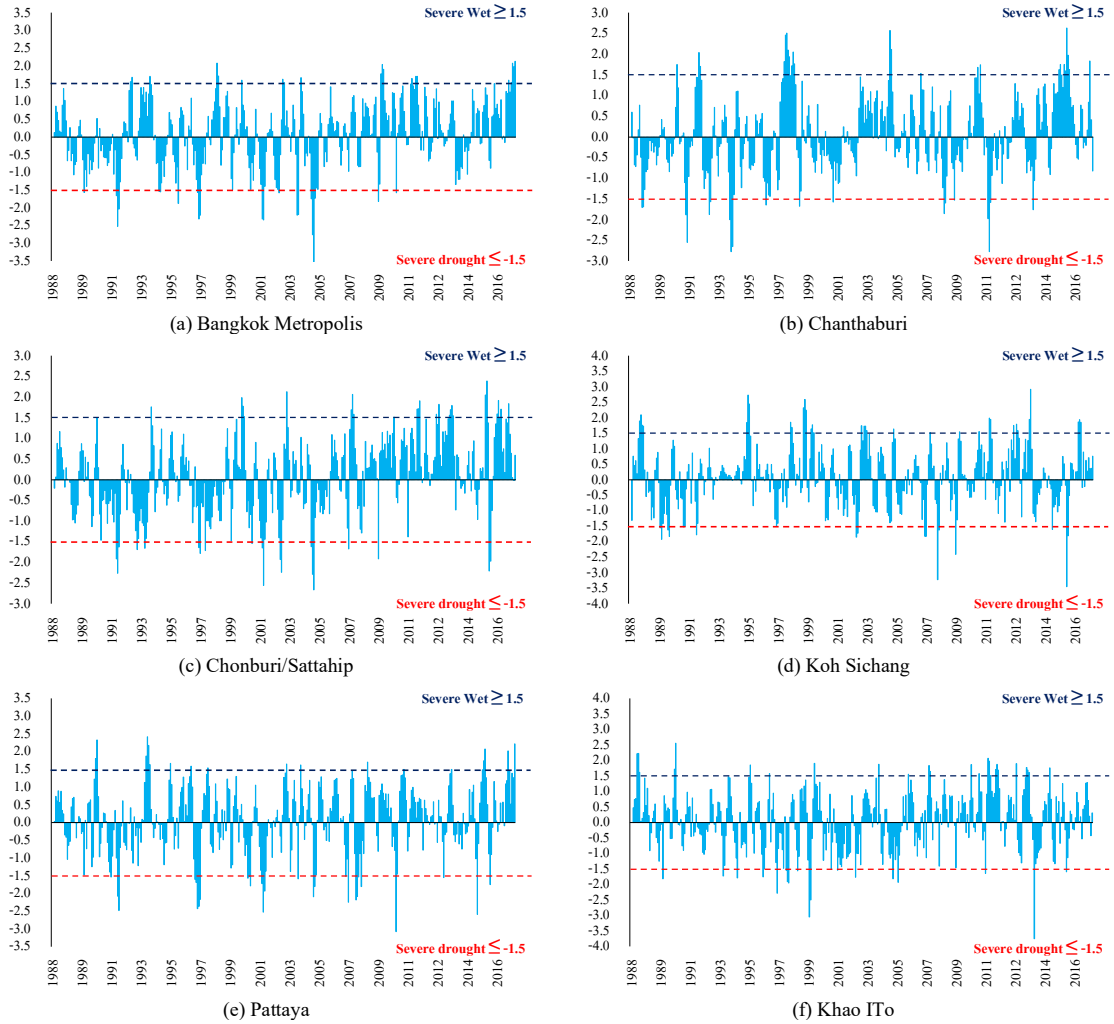
Category	Bangkok Metropolis	Chanthaburi	Chonburi/Sattahip	Koh Sichang	Pattaya	Khao ITo
Extreme drought	2.79	1.40	1.68	0.84	3.35	1.12
Severe drought	3.07	3.91	3.91	3.91	4.19	4.19
Moderate drought	7.26	5.31	8.10	9.50	5.31	11.17
Near normal	68.99	71.51	69.83	69.27	69.55	69.27
Moderate wet	12.01	11.17	9.22	7.26	13.13	7.26
Severe wet	4.75	4.47	6.15	7.26	2.79	5.87
Extreme wet	1.12	2.23	1.12	1.96	1.68	1.12

Table 5 Percentage of SPI6 by meteorological stations

Category	Bangkok Metropolis	Chanthaburi	Chonburi/Sattahip	Koh Sichang	Pattaya	Khao ITo
Extreme drought	1.69	0.56	1.41	0.56	3.10	0.56
Severe drought	4.79	3.66	3.38	2.25	4.51	2.54
Moderate drought	6.20	7.89	9.01	9.58	4.51	10.99
Near normal	72.39	70.42	70.14	67.89	71.27	72.11
Moderate wet	9.30	8.45	9.58	10.70	11.83	6.48
Severe wet	5.07	4.79	6.48	5.07	3.94	6.48
Extreme wet	0.56	4.23	0.00	3.94	0.85	0.85

Table 6 Percentage of SPI12 by meteorological stations

Category	Bangkok Metropolis	Chanthaburi	Chonburi/Sattahip	Koh Sichang	Pattaya	Khao ITo
Extreme drought	0.29	0.00	0.29	0.00	1.43	0.29
Severe drought	6.02	0.29	4.01	0.29	5.44	3.15
Moderate drought	6.02	10.60	9.74	7.74	5.44	7.74
Near normal	71.63	73.64	69.63	70.77	73.07	77.36
Moderate wet	8.60	8.88	11.46	9.46	12.61	5.44
Severe wet	7.16	2.58	4.01	7.16	2.01	4.58
Extreme wet	0.29	4.01	0.86	4.58	0.00	1.43



(a) Bangkok Metropolis (b) Chanthaburi

(c) Chonburi/Sattahip (d) Koh Sichang

(e) Pattaya (f) Khao ITo

Figure 2 Changes in the SPI3 Index Classified by Meteorological Stations

simulation time to near normal with 69.63 - 77.36 (%), while moderate, severe and extreme droughts ranged from 5.44 – 10.60 (%), 0.29 – 6.02 (%) and 0.00 – 1.43 (%), respectively.

Shows consecutive periods of drought for the SPI12, as it was an indicator of year-to-year drought changing by considering the SPI ≤ -1.5 (severe to extreme drought). Bangkok Metropolis had five periods: May – June 1992, November 1997 – April 1998, July – October 2001, January – July 2002, and February – April 2005. Chanthaburi had a maximum annual rainfall of 1 month (April 1995). Chonburi/Sattahip had four periods: November 1993 – April 1994, September – October 2001, February – April 2002, and July – August 2005. Koh Sichang had only one month (July 2001). Pattaya had four periods: April – September 1992, October – December 2000, April – December 2001, and January – July 2002. Khao Ito had two periods: August – September 1999 and October 2001 – April 2002.

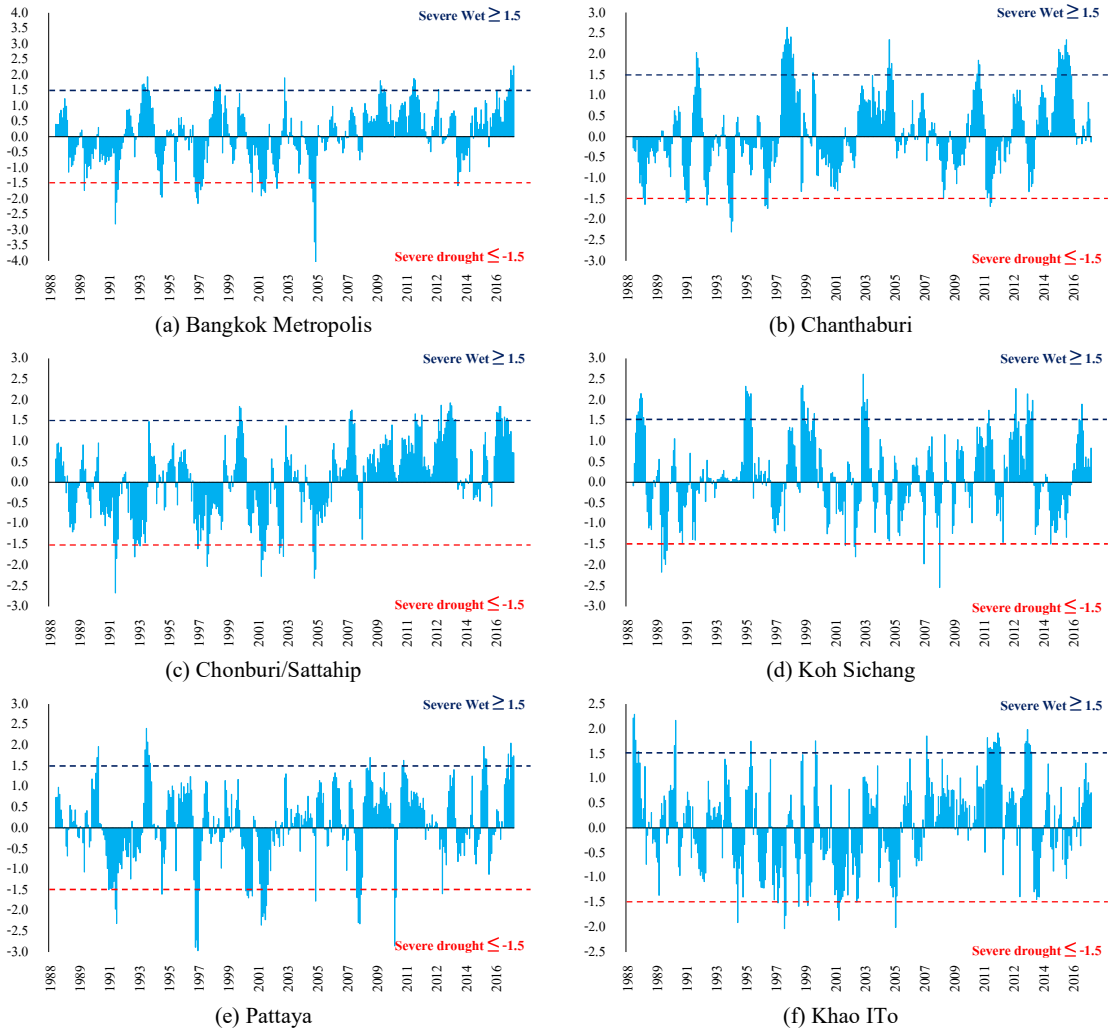


Figure 3 Changes in the SPI6 Index Classified by Meteorological Stations

The results of the analysis of all SPI showed that most of the study areas had near-normal rainfall. From the percentage of the SPI was close to near normal at 62%. However, there were also different levels of drought likelihood.

4. CONCLUSION

The drought severity levels were analyzed using the Standardized Precipitation Index (SPI) in the Eastern Economic Corridor (EEC) area. The six meteorological stations were analyzed in three patterns, including SPI3,

SPI6, and SPI12. The result showed that the mean of three SPI patterns ranged from 0.035 to 0.180, which was interpreted as the range of near normal. Considering the maximum and minimum values of the SPI in all periods, it was found that the value was greater than 1.5 and less than -1.5, respectively. Therefore, they indicated drought occurrence in both severe to extreme drought and severe to extreme wet. However, the likelihood of drought was at different levels in percentage. Most of the percentage of drought was more than 62 (%) in the study area, which experienced drought as near normal. The likelihood of drought (< -0.99) since moderate to extreme drought in the 30 years (1988 – 2017). There was a range below 17 % of the likelihood of drought and wet (100 %) in all SPI patterns, which was a slight likelihood of drought. The possibility of wet (> 0.99) since moderate to extremely wet in the 30 years (1988 – 2017) ranged below 21 % of the likelihood of drought and wet (100 %) in all SPI patterns. However, it was found that drought in the study area had uncertain patterns of drought and wet by considering the drought incidence curve. Hence, it is necessary to consider optimizing the water management plan according to the drought conditions.

However, in the analysis of drought levels, other indices consider other relevant climate variables, such as Temperature, Soil moisture, and satellite images, such as the Normalized Difference Vegetation Index (NDVI), which can be used to analyze with the SPI for more excellent reliability. Finally, the SPI was a convenient index that can be used to analyze the severity level of drought because it uses only rainfall data to represent the drought situation, with the data over 30 years. That took a long time to explore the accuracy of the results. Moreover, this research will be a starting point for further analysis of the SPI in the form of future projections of rainfall data under climate change in the study area.

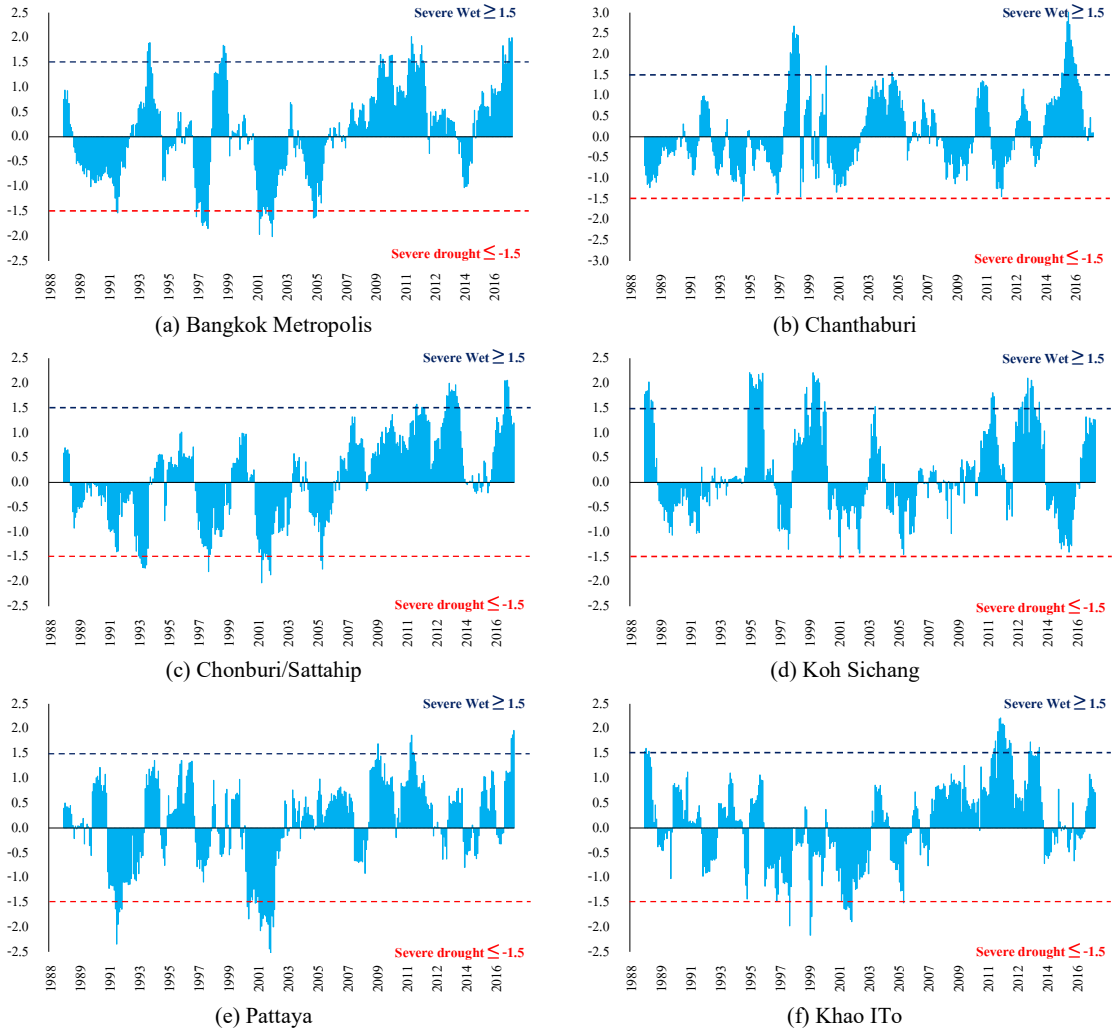


Figure 4 Changes in the SPI12 Index Classified by Meteorological Stations

5. ACKNOWLEDGMENT

I want to thank the advisors and members of the Department of Irrigation Engineering for providing knowledge and advice for completing the research.

6. REFERENCES

Adisa O. M., Masinde M. and Botai J. O. (2021). Assessment of the Dissimilarities of EDI and SPI Measures for Drought Determination in South Africa. Water 2021 13(82), 1-17. DOI: 10.3390/w13010082

Chaito, T. and Khamkong, M. (2020). Time Series Model for Standardized Precipitation Index in the Ping River Basin of Chiang Mai Province. The Journal of KMUTNB 2020, 1-14. DOI: 10.14416/j.kmutnb.2020.12.012

Damberg L.and AghaKouchak A. (2014). Global trends and patterns of drought from space. Theoretical and applied climatology 2014; 117(3-4), 441-448.

Homdee, T., Pongbut, K. and Kanae, S. (2016). A Comparative performance analysis of three standardized climatic drought indices in the Chi River basin, Thailand. Agriculture and Natural Resources 50, 211-219. DOI: 10.1016/j.anres.2016.02.002

Huang Y. F., Ang J. T., Tionga Y. J., Mirzaei M. and Mat M. Z. (2016). Drought Forecasting using SPI and EDI under RCP-8.5 Climate Change Scenarios for Langat River Basin, Malaysia. Procedia Engineering, 154, 710-717.

Hydro-Informatics Institute. (2012). Data collection and data analysis for the 25 rivers basin and model for flood and drought in the Ping river basin (in Thai). Available Source: http://www.thaiwater.net/web/

Kaingam, V. and Chotamonsak, C. (2019). Analysis of Drought in Northern Thailand Using Standardized Precipitation Index. J Sci Technol MSU 39(3), 313-322.

Madusanka T. and Venkataramana S. (2017). Characterization of future drought conditions in the Lower Mekong River Basin. Weather and Climate Extremes 17, 47-58. DOI: 10.1016/j.wace.2017.07.004

Mckee T. B., Doesken N. J. and Kleist J. (1993). The Relationship of Drought frequency and duration to time scales. Eighth Conference on Applie Climatology, American Meteorological Society, Jan 17-23, Anaheim CA, 179-186.

Pei Z., Fang S., Wang L. and Yang W. (2020). Comparative Analysis of Drought Indicated by the SPI and SPEI at Various Timescales in Inner Mongolia, China. Water 2020 12(1925), 1-20. DOI: 10.3390/w12071925

Pinthong, A. and Kwanyuen, B. (2016). Drought Monitoring by Composite Drought Index. Rajamangala University of Technology Thanyaburi 15(2), 45-53. DOI: 304080-1-10-20180403

Sriwongsitanon, N. (2015). Drought Monitoring and Early Warning for Thailand. Department of Water Resources Engineering, Kasetsart University, 17-27.

Taesombat, W. (2012). Engineering Hydrology. Department of Irrigation Engineering, Kasetsart University Kamphaeng saen campus, 96-100.

Thai Meteorological Department. (2016). Climate Change. Available Source: https://www.tmd.go.th/info/info.php?FileID=86/

Thai Meteorological Department. (2012). Study on Drought Index in Thailand 2012, Agro - meteorological Division, Meteorological Development Bureau, 33-60.

Thanapakpawin P., Boonya-aroonnet S., Chankarn A., Chitradon R. and Snidvongs A. (2011). Chapter 7 Thailand drought risk management: macro and micro strategies. Droughts in Asian Monsoon Region (Community, Environment and Disaster Risk Management, Volume 8) Emerald Group Publishing Limited 2011; 8, 121-140.

The Eastern Economic Corridor Office of Thailand. (2018). Combined diagram for the development of the Eastern Economic Corridor 2017 - 2022. The Eastern Economic Corridor Office of Thailand (EECO), 71-80.

Wichitarapongsakun, P. (2015). Drought Analysis in the Rain-fed Agriculture Area Using the Standardized Precipitation Index (SPI) in the HuaiKoKaeo sub-basin of the Pasak River Basin. 12th THAICID National Symposium, 150-163.

Wichitarapongsakun, P., Sarin, C., Klomjek, P. and Chuenchooklin, S. (2016). Meteorological drought in the Sakea Krang River basin using the Standardized Precipitation Index (SPI) and the Meteorological Drought Index (D). Naresuan University Journal: Science and Technology 2016 ; 24(3), 123-135. DOI: 1549--4903-1-10-20161011_2

Wichitarapongsakun, P., Sarin, C., Klomjek, P. and Chuenchooklin, S. (2017). Rainfall prediction and meteorological drought analysis in the Sakae Krang River basin of Thailand. Agriculture and Natural Resources 50, 490-498. DOI: 10.1016/ j.anres.2016.05.003

Wilhite D. A. and Glantz M. H. (1985). Understanding: the drought phenomenon: the role of definitions. Water international 1985l ; 10(3), 111-120.

NUEJ
Naresuan University
Engineering Journal

LCS-based Thai Trending Keyword Extraction from Online News

Kietikul Jearanaitanakij[1,*], Nattapong Kueakool[1,2], Puwadol Limwanichsin[1,2], Tiwat Kullawan[2] and Chankit Yongpiyakul[2]

[1] Department of Computer Engineering, School of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand
[2] InfoQuest Limited, Bangkok, Thailand

* Corresponding author e-mail: kietikul.je@kmitl.ac.th

(Received: 23 August 2022, Revised: 4 November 2022, Accepted: 15 November 2022)

Abstract

A trending keyword is a common word or a phrase that is most frequently mentioned in the current period. Extracting trending keywords from Thai online news is not trivial. A too-short keyword may not have a specific meaning because it may be just a common word that does not have any significance to the interpretation. On the other hand, a long common keyword conveys a better meaning. However, the running time to extract the long keyword from a collection of documents may not be bounded within a reasonable time. A problem statement of this research is to find a varying-length trending keyword from Thai online news within a reasonable running time. We propose a novel method to extract trending keywords by applying the longest common substring (LCS) algorithm. The common keywords having high occurrence frequency are selected as the trending keywords. The proposed method inherits the advantage of the reasonable running time from the dynamic programming technique of the LCS algorithm. The experimental results on various sources of Thai online news agencies indicate a superior precision of the proposed method over char-N-gram and word-N-gram strategies.

Keywords: Longest common substring, Natural language processing, Online news, Thai trending keyword, Varying-length keyword.

1. INTRODUCTION

Among all sources of news in Thailand, it is undeniable that online news is the most popular platform. There is a lot of online news issued per day making the current trend extraction difficult. To find the trending keyword, two mechanisms are needed to implement. The first is the process of finding the common keyword from a collection of documents. The second is the ability to determine that the selected common keyword is the current news trend. There are many methods for extracting a common keyword from documents. Lee and Kim (2008) apply TF-IDF (Robertson, 2004) variants and filter keywords across domains to extract keywords from a huge pile of documents. The experimental results show that their approach can efficiently remove meaningless Korean words. Ma et al. (2008) identify the query and the topic-related features for each word that cooccurs in the same window length and then calculate the importance of the word from the combined features. Their experimental results are competitive with other candidates. Zhang et al. (2008) introduce the keyword extraction based on Conditional Random Fields (CRF) that use most of the features in the document. The CRF model outperforms other off-the-shelf machine learning algorithms such as SVM and multiple linear regression models in terms of F1 score. However, a computational cost of training CRF model is expensive.

Once the common keyword is detected, the next step is to confirm that the keyword represents the actual news trend. Various approaches are proposed to identify the trending keyword. Shimizu et al. (2005) extract the trending keyword from the press releases to analyze the competitors' strategies. Words that are matched with a description pattern, i.e., a product common phrase observed from press releases, are considered to be a product trending keywords. However, their method requires a user to specify a description pattern which may not be appropriate for an automatic system. Sutheebanjard and Premchaiswadi (2010) applied four string matching techniques, e.g., Guth, Levenshtein, Damerau-Levenshtein, longest common substring and longest common subsequence to disambiguate Thai personal names from a collection of articles. They found that the longest common subsequence was the best Thai personal name matching with 94.43% of F-Score. Aiello et al. (2013) identify the trending topics on Twitter by combing N-grams occurrence with a topic-ranking score representing the rapidly emerging topic. Their method outperforms other techniques on Twitter data. Madani et al. (2015) discover trending topics from Twitter in real-time by using topic clustering. Tweets thesaurus created

in the form of a bag of words can semantically represent different terms corresponding to the same trending topic. Ousirimaneechai and Sinthupinyo (2018) detect the trending keywords using the simple character N-grams to tokenize document and find keywords. They identify the trending keyword by filtering a set of stop words generated by their algorithm. Their method requires neither word tokenization tools nor external stop words list. According to the experimental results, their method frequently takes too-short words as the keywords and depends mainly on the generated stop words. Indra et al. (2019) compare two methods, namely document pivot, and BN-grams, for detecting trending topics on Indonesian tweets. Their experimental results on ten topics indicate that BN-grams have better topic recall than the document pivot. Alzubi et al. (2020) extract trending scientific topics by scoring topics according to two factors; the number of citations and the number of accepted papers on that topic. In addition, their method can recommend topics customized to each user profile. However, they do not experiment on the real dataset but on the synthesized 3000 papers instead. Tanantong et al. (2020) extract trending keywords from Thai Twitter by using the N-gram-based word-combination technique. Words that are located in an adjacent position are considered to combine into one candidate keyword. The candidate keyword which has the highest rate of appearances within three consecutive days on Twitter is extracted as the trending keyword.

According to the above works, none of them extracts long and varying-length common keywords which tend to convey better meaning than short keywords. In this research, we propose Thai trending keyword extraction from online news by using the longest common substring (LCS) algorithm and the rescaled keyword frequency. LCS is a powerful tool that can find the longest common keyword from the collection of documents. It can be customized to search for varying-length keywords. While LCS helps us find the longest common keywords, the rescaled frequency of keyword occurrences identifies whether the common keyword is the current trend. The experimental results on the online Thai news collected from different sources show a superior performance of the proposed method compared to other algorithms.

2. Related Works

Two fundamental concepts, e.g., the longest common substring, Thai word segmentation, related to this research are described in the first two subsections. In addition, two previous works are also briefly explained since we compare their results with the proposed method. Let us abbreviate techniques in (Ousirimaneechai and Sinthupinyo, 2018) and (Tanantong et al., 2020) as Char-N-Gram and Word-N-Gram, respectively.

2.1 The Longest Common Substring Problem

The longest common substring (LCS) is a classical problem in computer science (Gusfield, 1997). The problem is about finding the longest substring of two or more strings. For instance, the LCS of "AABCDEFG" and "ABABCDEGH" is "ABCDE". By using a dynamic programming technique, we can locate LCS between two strings of lengths m and n within O(mn) time complexity. Fig. 1 illustrates pseudocode for finding LCS between two strings (S1 and S2) by a dynamic programming technique. The algorithm processes the LCS matrix from left to right and top to bottom. For each element indexed by row i and column j, if S1[i] is equal to S2[j] then the value of LCS_Mat[i, j] equals the value from the previous upper left element plus one. Otherwise LCS_Mat[i, j] equals to zero. After all elements in the LCS matrix have been filled, the common substring which has the longest length will be chosen as the solution. The example for extracting the longest substring from the above two strings can be demonstrated in Fig. 2. Denote that the green diagonal elements represent the longest common substring which has the length of five alphabets, i.e., ABCDE. More details of the LCS algorithm including its variations (Mousavi et al., 2012; Beal et al., 2016; Charalampopoulos et al., 2021; Akmal et al., 2021) can be found in the reference section.

```
Function LCS (S1, S2)
  Let S1 and S2 are strings of length m and n, respectively.
      LCS_Mat is a matrix of dimension m x n.
      max_length holds the length of LCS found so far, initial value = 0.
      lcs is the longest common substring, initial value = empty.
      S1[a:b] represents the inclusive substring from S1[a] to S1[b].

  for i ← 1 to m
    for j ← 1 to m
      if S1[i] == S2[j]
        if i == 1 or j == 1
            LCS_Mat[i, j] ← 1
        else
            LCS_Mat[i, j] ← LCS_Mat[i − 1, j − 1] + 1
        if LCS_Mat[i, j] > max_length
            max_length ← LCS_Mat[i, j]
            lcs ← {S1[i − max_length + 1: i − max_length + i]}
        else if LCS_Mat[i, j] == max_length
            lcs ← lcs ∪ { S1[i − max_length + 1: i − max_length + i] }
      else
        LCS_Mat[i, j] ← 0
  return lcs
```

Figure 1 Pseudocode of LCS algorithm

	A	B	A	B	C	D	E	G	H
A	1	0	1	0	0	0	0	0	0
A	1	0	1	0	0	0	0	0	0
B	0	2	0	2	0	0	0	0	0
C	0	1	0	0	3	0	0	0	0
D	0	0	0	0	0	4	0	0	0
E	0	0	0	0	0	0	5	0	0
F	0	0	0	0	0	0	0	0	0
G	0	0	0	0	0	0	0	1	0

Figure 2 LCS example for strings AABCDEFG and ABABCDEGH

2.2 Thai Word Segmentation

One of the hardest parts of Thai natural language processing is word segmentation. Due to the lack of space between Thai words, we need to apply subtle techniques to separate them into individual words. There are several well-known algorithms for Thai word segmentation, e.g., Newmm, Deepcut, and AttaCut. Newmm, implemented in PyThaiNLP (Phatthiyaphaibun et al., 2016), uses a maximal matching algorithm, based on words in a dictionary, and the Thai character cluster. It always breaks the same longest word no matter the position of the word in the context. For example, a string "ตากลม" is always segmented into "ตาก | ลม" (to gain air). On the other hand, DeepCut (Kittinaradorn et al., 2019) and its variance AttaCut (Chormai et al., 2019) apply a more advanced technique, i.e., a convolutional neural network, to segment a sentence into a list of words depending on the word surrounding context. From the previous example, the string "ตากลม" can be either segmented into "ตาก | ลม" (to gain air) or "ตา | กลม" (round eye) depending on its surrounding words. To conform with the LCS algorithm, we employ Newmm algorithm to consistently tokenize sentences into a list of the longest words.

2.3 Char-N-Gram

Char-N-Gram (Ousirimaneechai and Sinthupinyo, 2018) begins the process by tokenizing text into five-character grams and counting the number of occurrences of them. Afterward, the K-Means clustering algorithm is applied to the counted grams to group those grams into K clusters by using the elbow method. The lowest rank cluster is discarded since it has a low chance of containing keywords. The keyword clusters from other days are combined and clustered with the K-Means algorithm again. The cluster with the highest mean of occurrences will be assigned as the global keyword grams. For each news post, replace characters that are not in any part of elements in the global keyword grams with a blank symbol and split the resulting text by using a blank symbol. The global keywords are words that remained in the news post after removing stop words. Finally, to find the trending keywords, pick the global keywords which have a high frequency of occurrences within a day.

2.4 Word-N-Gram

Word-N-Gram (Tanantong et al., 2020) applies a simpler idea. It replaces stop words from the original post with a blank symbol and combines the adjacent non-blank words. These adjacent words are candidates for the trending keywords. The candidate which has the highest frequency of occurrences in all posts within a period is counted as a trending keyword.

3. Proposed Method

The following steps describe the proposed method along with the diagram in Fig. 3.

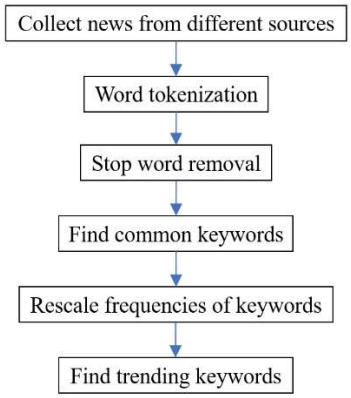


Figure 3 Proposed method

Step 1: Collect the current news posts from different sources and sort them by a chronological order.
Step 2: Tokenize words in news posts by using Newmm algorithm.
Step 3: Remove stop words from the tokenized news.
Step 4: Find the common keywords of all news posts by using the LCS algorithm.
Step 5: Rescale occurrence frequencies of all common keywords.
Step 6: Pick the rescaled LCS keywords which have high occurrence frequency as the current trends.

After breaking news posts into a list of words by Newmm technique, we follow the process in (Tanantong et al., 2020) to remove stop words in step 3. Stop words tend to be insignificant to the main idea and frequently appear in news. To avoid selecting common keywords which are not necessary to be trending keywords, we monitor words that occur every day for one month and remove them from a list of tokenized words. It is worth noting that the row and column elements of the LCS matrix in step 4 are represented as the tokenized words (not alphabets). To avoid too-short and too-long keywords, we select only common phrases whose lengths are within three to eight words. As a result, the proposed method can produce the varying-length trending keywords. The example of LCS matrix calculation in step 4 can be illustrated in Fig. 4. The purpose of the rescaling

process in step 5 is to ensure that all common keywords returned from LCS every day are on the same scale. The rescaled frequency of each LCS keyword is the ratio between the number of its occurrences and the total number of occurrences of all LCS common keywords within a certain period.

	รัฐบาล	มี	นโยบาย	คน	ละ	ครึ่ง	เหมือน	ให้	ข้าว
โครงการ	0	0	0	0	0	0	0	0	0
คน	0	0	0	1	0	0	0	0	0
ละ	0	0	0	0	2	0	0	0	0
ครึ่ง	0	0	0	0	0	3	0	0	0
ถือ	0	0	0	0	0	0	0	0	0
เท็โน	0	0	0	0	0	0	0	0	0
นโยบาย	0	0	1	0	0	0	0	0	0
ที่	0	0	0	0	0	0	0	0	0

Figure 4 Example of the LCS matrix for a common keyword "คนละครึ่ง"

4. EXPERIMENTAL RESULTS

We collect 155,661 online posts from eight different Thai news agencies; Thairath, Thai PBS, Spring News, Siamrath, Sanook, Post Today, Komchadluek, and Amarin, between 1st September 2021 and 31st December 2021. Table 1 shows the number of news instances from each agency. The ground truth of each day, i.e., the actual trending keyword, is manually assigned as a set of 1 – 3 keywords by our news experts. Each instance of news in the dataset composes of date, time, and news content.

To have a fair comparison, we conduct the experiments of all methods on the standard Google Colab (cpu: 2.2 GHz/single core/cache: 56320 KB, ram: 13.3 GB, disk: 108 GB). We begin the experiment with the quest for a suitable Thai word tokenization. Later in the section, the experimental results for top–10 precision, running time, and word cloud are provided.

Table 1 The number of posts collected from news agencies

News Agency	Number of news posts
Thairath	27544
Post Today	16988
Sanook	10743
Thai PBS	5895
Komchadluek	25450
Amarin	10999
Siamrath	51211
Spring	6831

The objective of the first experiment is to confirm that a suitable word tokenization algorithm for the proposed method is Newmm. We randomly pick 10000 news instances from the dataset and apply both Newmm and AttaCut to break news into words. The number of segmented words that appear in the LEXiTRON dictionary (Trakultaweekoon et al., 2007) and the running time of each algorithm are reported in Table 2. Newmm has better performance than AttaCut in terms of both percentages of words that appear in the LEXiTRON dictionary and running time. Moreover, words segmented by Newmm are consistent and do not depend on the surrounding context which is suitable for finding the trending keywords. As a result, we employ Newmm as the Thai word segmentation. To break English words which may embed in Thai news, we simply split text by using a blank character as a delimiter. As we will see in the experimental section, the Thai trending keyword tends to be a phrase that does not depend on the surrounding context. Therefore, the percentage of words that appear in the LEXiTRON dictionary is a suitable indicator to measure the consistency of word tokenization of the two algorithms.

Table 2 Word tokenization comparison between Newmm and AttaCut.

Measure	Newmm	AttaCut
Percentage of words appear in LEXiTRON dictionary (%)	82.60	44.47
Running time (second)	9.18	88.27

The experiments of different methods are conducted on the dataset described in the previous section. The ground truth for each day is manually assigned as a set of 1 – 3 trending keywords by our volunteers. News instances are fed to the algorithms in chronological order starting from the oldest to the latest. Each algorithm predicts ten trending keywords for each day. If the volunteers found that at least one of ten predictions semantically matches the ground truth of that day, the prediction is counted as a hit, i.e., top-10 precision. The task of precision judging is not trivial to execute without human intervention since the predicted trending keywords may not exactly match the ground truth keywords. In other words, they implicitly match the meaning of the ground truths. For example, the keyword "แอนชิลี สก็อต-เคมมิส" (Anchilee Scott-Kemmis) implicitly matches the ground truth "มิสยูนิเวิร์สไทยแลนด์" (Miss Universe Thailand), "แอพเป๋าตัง" (Paotang application) implicitly matches the ground truth "โครงการคนละครึ่ง" (The half-half co-payment project). Therefore, we need human to judge the top-10 precision. The top-10 precision percentages of the proposed method (LCS-based) and the other two techniques; Char-N-Gram (Ousirimaneechai and Sinthupinyo, 2018) and Word-N-Gram (Tanantong et al., 2020), are shown in Table 3.

LCS-based method gains the best top-10 precision of all three months. Since the number of words in a common keyword can be varied between three and eight, the LCS-based method can produce meaningful keywords. In contrast, Char-N-Gram limits the length of the keyword to a gram of five characters and keeps only the cluster that has the highest mean of occurrences. Although this cluster usually contains keywords that have the highest mean of occurrences, those keywords tend to be short and less meaning, e.g., "การ" (task), "เพราะ" (because), "จำนวน"

(amount). Therefore, it frequently misses long keywords like "น้อม รำลึก ถึง พระ มหา กรุณา ธิ คุณ" (bow in remembrance of His Majesty the King), "เจ้า พนักงาน ละเว้น โดย มิ ชอบ" (staff wrongful refrain), "ฮู แต ยือ ลอ เระ บา เจาะ" (Hutaeyelor village, Bacho district), etc. On the other hand, Word-N-Gram does not suffer from short keywords. After stop word removals, it concatenates the remaining adjacent words resulting in keywords with different lengths. However, Word-N-Gram does not try to produce the longest keyword like the LCS-based method. The keywords from Word-N-Gram which have the highest rate of occurrences may be long, e.g., "มหาวิทยาลัย" (university), "อาจารย์เจ้าหน้าที่" (staff and faculty), but do not have specific meanings. As a result, more than half of trending keywords from Word-N-Gram do not semantically match any of the three ground-truth keywords.

The next experiment is to measure the running time among three methods. Table 4 compares the running time (in seconds) of trending keyword extraction for all news instances within three months. Word-N-Gram spends the least running time because of its simple procedure. In contrast, Char-N-Gram takes so much time in finding all possible 3-gram, 5-gram, and 8-gram words and running two rounds of K-Means clustering. The LCS-based method inherits a moderate running time of O(mn) from the dynamic programming technique of the LCS algorithm (Gusfield, 1997); where m and n are news lengths. Although we can further optimize the running time of LCS by creating a generalized suffix tree for strings, there is a tradeoff in memory space for storing a large tree.

Table 3 Top-10 Precisions of three methods

Measure	Char-N-Gram (N = 3 / 5 / 8)	Word-N-Gram	LCS-based
October	11.45/13.04/12.94	26.09	63.04
November	11.98/14.13/14.03	39.13	57.60
December	16.54/19.56/18.99	45.65	60.86
Average	13.32/15.58/15.32	36.96	60.50

Table 4 Running time in second of three methods

Measure	Char-N-Gram (N = 3 / 5 / 8)	Word-N-Gram	LCS-based
October	5211/6318/7410	320	955
November	5345/6655/7833	322	885
December	5572/6845/7942	330	856
Average	5376/6606/7728	324	898.67

To have a quick and visual insight of the trending keywords from different methods, we generate word clouds to highlight popular keywords based on their frequencies of occurrences. The larger font of the word, the more frequently it occurs in news posts. Fig. 5 illustrates word clouds on 25th October 2021 which has two ground truths; "พิธีอัญเชิญพระเกี้ยว" (Phra Kiew coronet parade) and "มิสยูนิเวิร์สไทยแลนด์" (Miss Universe Thailand). Keywords in red ellipses are trends that semantically match the trending ground truths. Trending keywords from Char-N-Gram and Word-N-Gram do not partially match any keywords in the ground truths. Keywords from Char-N-Gram are too short while those from Word-N-Gram are not related to the ground truths. In contrast, two top trending keywords, i.e., "นิสิตอัญเชิญเกี้ยว" (students who participate in Phra Kiew coronet parade), and "ชิลีสก๊อตมิส" (Anchilee Scott), from LCS-based method semantically match the target trending keywords. Notice that many long keywords from Word-N-Gram have a small frequency of occurrences, e.g., "มันสำปะหลังข้าวโพดเลี้ยงสัตว์" (cassava, corn, animal husbandry), "หลอมรวมความสมัครสมาน" (gather unity), "อนึ่งหย่อมความกดอากาศ" (by the way, the air pressure patch) In addition, the air pressure patch. After combining words leftover from the stop word removals, those long keywords of Word-N-Gram do not reflect any trend in the current news posts. On the other hand, the LCS-based method intends to find the longest common keywords to help promote the collocation degree of keywords; making them cover more meanings in the ground truths.

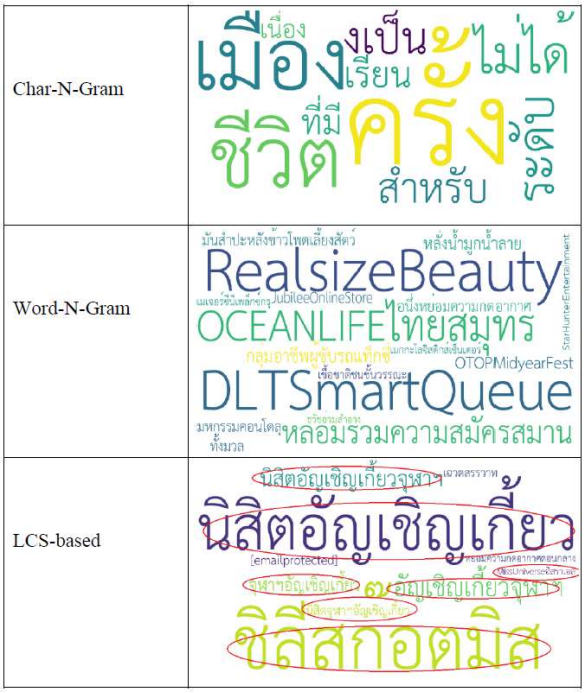


Figure 5 Word clouds for trending keywords on 25 October 2021

Fig. 6 expresses a similar phenomenon as Fig. 5. It demonstrates the word cloud on 23rd November 2021 which has the ground truths "พัคชินฮเย ชเวแทจุน" (Park Shin-hye & Choi Tae-joon), "ดราม่า แอนชิลี สก็อต" (Anchilee Scott accused of Thailand flag abuse). Neither Char-N-Gram nor Word-N-Gram captures any trend in the ground truths whereas the LCS-based method covers all current trends. Fig.7 shows another word cloud on 24th December 2021 which has the ground truths "เลือกตั้งซ่อม กทม." (Bangkok

election), "ฟุตบอลไทย-เวียดนาม AFF" (Thailand-Vietnam AFF football), "โอมิครอน" (Omicron). Both Char-N-Gram and Word-N-Gram implicitly capture only one keyword "โอมิครอน" (Omicron) while the LCS-based method semantically hits all trending ground truths.

Interestingly, Word-N-Gram contains more English words than other methods in Fig. 5-7. These are situations where Word-N-Gram lost important Thai stop words while replacing them with blank symbols. After combining non-blank words, its keywords may look strange and have a low frequency of occurrences. As a result, it misses the ground truth trending keywords and continues extracting lower-frequency keywords which may include English keywords.

Figure 6 Word clouds for trending keywords on 23 November 2021

Figure 7 Word clouds for trending keywords on 24 December 2021

5. CONCLUSION

We propose a novel method to extract Thai trending keywords from online news by using the longest common substring (LCS) algorithm to find a set of the longest common keywords. The longest common keywords that have a high frequency of occurrences will be selected as the trending keywords. The experimental results on 155,661 online news posts collected from eight Thai news agencies between September 2021 and December 2021 indicate the superior performance of the LCS-based method over other techniques. The contribution of the proposed method is the effort to find the longest trending keyword from the current collection of news. The LCS-based method can find the keywords with varying ranges, e.g., 3 to 8 words, depending on the context of the news. In practice, we recommend a user to find a suitable range from the sample news posts. We use the range between 3 and 8 words because of two factors. First, a glance from our dataset does not show a keyword with more than eight words. Second, we need to control the running time of the LCS algorithm within a reasonable bound. The advantage of long keywords is the ability to express trends more meaningfully. The LCS-based method possesses the highest top-10 precision, on average, 60.50% and the moderate running time at 898.67 seconds compared to other methods. In addition, it is possible to apply the LCS-based method to extract trending keywords from other languages by simply changing a word tokenization module. One of the possible future works of the proposed method is to apply the cosine similarity to group similar trending keywords together and express them as one keyword. This improvement can help limit the trending keywords to a smaller set of keywords and make it easy

to interpret news insight. Despite its moderate running time, the LCS-based method may suffer from the O(NxN) operation for coupling a pair of news to find LCS. It is recommended to preliminary run the LCS-based method to find a suitable value of N that produces an acceptable total running time or apply multiple threads to gain the concurrency of LCS operations.

6. ACKNOWLEDGMENT

We gratefully acknowledge support from eight Thai news agencies, including Thairath, Post Today, Sanook, Thai PBS, Komchadluek, Amarin, Siamrath, and Spring, for their valuable online news posts.

7. REFERENCES

Aiello, L. M., Petkos, G., Martin, C., Corney, D., Papadopoulos, S., Skraba, R., G¨oker, A., Kompatsiaris, Y., & Jaimes, A. (2013). Sensing trending topics in Twitter. *IEEE Transaction on Multimedia*, *15(6)*, 1268–1282. https://doi.org/10.1109/TMM.2013.2265080

Akmal, S., & Williams, V. V. (2021). Improved approximation for longest common subsequence over small alphabets. *48th International Colloquium on Automata, Languages, and Programming* (pp. 1-19). ArXiv. https://doi.org/10.48550/arXiv.2105.03028

Alzubi, S., Hawashin, B., Mughaid, A., & Jararweh, Y. (2020). Whats trending? an efficient trending research topics extractor and recommender. *11th International Conference on Information and Communication Systems* (pp. 191-196). IEEE. https://doi.org/10.1109/ICICS49469.2020.239519

Beal, R., Afrin, T., Farheen, A., & Adjeroh, R. (2016). A new algorithm for "the LCS problem" with application in compressing genome resequencing data. *BMC Genomics*, *17(4)*, 369–381. https://doi.org/10.1186/s12864-016-2793-0

Charalampopoulos, P., Kociumaka, T., Pissis, S. P., & Radoszewski, J. (2021). Faster algorithms for longest common substring. *29th Annual European Symposium on Algorithms* (pp. 1-30). ArXiv. https://doi.org/10.48550/arXiv.2105.03106

Chormai, P., Prasertsom, P., & Rutherford, A. (2019). AttaCut: a fast and accurate neural Thai word segmenter. *ArXiv*, *1911(7056)*, 1–13. https://doi.org/10.48550/arXiv.1911.07056

Gusfield, D. (1997). *Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology*. Cambridge University Press. https://doi.org/10.1017/CBO9780511574931

Indra, S. K., Winarko, E., & Pulungan, R. (2019). Trending topics detection of Indonesian tweets using BN-grams and Doc-p. *Journal of King Saud University - Computer and Information Sciences*, *31(2)*, 266–274. https://doi.org/10.1016/j.jksuci.2018.01.005

Kittinaradorn, R., Chaovavanich, K., Achakulvisut, T., Srithaworn, K., Chormai, P., Kaewkasi, C., Ruangrong, T., & Oparad, K. (2019, September 23). DeepCut: A Thai word tokenization library using Deep Neural Network. Retrieved from https://doi.org/10.5281/zenodo.3457707

Lee, S., & Kim, H. (2008). News keyword extraction for topic tracking. *Fourth International Conference on Networked Computing and Advanced Information Management* (pp. 554-559). IEEE. https://doi.org/10.1109/NCM.2008.199

Ma, L., He, T., Li, F., Guil, Z., & Chen, J. (2008). Query-focused multi-document summarization using keyword extraction. *International Conference on Computer Science and Software Engineering* (pp. 20-23). IEEE. https://doi.org/10.1109/CSSE.2008.1323

Madani, A., Boussaid, O., & Zegour, D. E. (2015). Real-time trending topics detection and description from Twitter content. *Social Network Analysis and Mining*, *5(59)*, 1–13. https://doi.org/10.1007/s13278-015-0298-5

Mousavi, S. R., & Tabataba, F. (2012). An improved algorithm for the longest common subsequence problem. *Computers & Operations Research*, *39(3)*, 512–520. https://doi.org/10.1016/j.cor.2011.02.026

Ousirimaneechai, N., & Sinthupinyo, S. (2018). Extraction of trend keywords and stop words from Thai Facebook pages using character n-grams. *International Journal of Machine Learning and Computing*, *8(6)*, 589–594. http://www.ijmlc.org/vol8/750-ML0015.pdf

Phatthiyaphaibun, W., Chaovavanich, K., Polpanumas, C., Suriyawongkul, A., Lowphansirikul, L., & Chormai, P. (2016, June 27). PyThaiNLP: Thai Natural Language Processing in Python. Retrieved from http://doi.org/10.5281/zenodo.3519354

Robertson, S. (2004). Understanding inverse document frequency: on theoretical arguments for IDF. *Journal of Documentation*, *60*(5), 503–520. https://doi.org/10.1108/00220410410560582

Shimizu, Y., Akiyoshi, M., & Komoda, N. (2005). A method of extracting product trend keywords from press releases to analyze product strategy of competitors. *International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce* (pp. 631-635). IEEE. https://doi.org/10.1109/CIMCA.2005.1631539

Sutheebanjard, P., & Premchaiswadi, W. (2010). Disambiguation of Thai personal name from online news articles. *International Conference on Computer Engineering and Technology* (pp. V3-302-V3-306). IEEE. https://doi.org/10.1109/ICCET.2010.5485879

Tanantong, T., Kreangkriwanich, S., & Laosen, N. (2020). Extraction of trend keywords from Thai Twitters using n-gram word combination. *International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology* (pp. 320-323). IEEE. https://doi.org/10.1109/ECTI-CON49241.2020.9158061

Trakultaweekoon, K., Porkaew, P., & Supnithi, T. (2007). LEXiTRON vocabulary suggestion system with recommendation and vote mechanism. *Proceedings of Symposium of Natural Language Processing* (pp. 43-48). National Electronics and Computer Technology Center. lexitron.nectec.or.th/2009_1/paper/paper_3.pdf

Zhang, C., Wang, H., Liu, Y., Wu, D., Liao, Y., & Wang, B. (2008). Automatic keyword extraction from documents using conditional random fields. *Journal of Computer Information Systems*, *4(3)*, 1169–1180. https://core.ac.uk/download/pdf/11884499.pdf

NUEJ
Naresuan University
Engineering Journal

Establishment of Water Quality Indexes for Raw Water in the Water Supply

Akarayut Kraikriangsri*, Sitang Pilailar and Suwatana Chittaladakorn

Department of Water Resources Engineering, Faculty of Engineering at Bangkhen, Kasetsart University, Bangkok, Thailand

* Corresponding author e-mail: akarayut.kr@ku.th, tul4936@gmail.com

(Received: 29 July 2022, Revised: 3 November 2022, Accepted: 17 November 2022)

Abstract

Over two decades, climate variability has caused a significant change in human well-being, particularly in the availability of water resources. The water shortage and insufficient water supply especially affect the quality of life in many ways. The Pollution Control Department (PCD) of the Metropolitan Waterworks Authority (MWA) has selected thirteen parameters considered essential as indicators of raw water suitability. The water quality of the Chao Phraya River at 18 PCD stations along the 379-km stretch from Chainat to Samutprakarn has been recorded for analysis and considered for inclusion in a Water Quality Index (WQI) that includes color, turbidity, Biochemical Oxygen Demand (BOD), Dissolved oxygen (DO), pH, Nitrate-Nitrogen (NO_3-N), Fe, Mn, Total Dissolved Solids (TDS), Total Phosphorous (TP), Electrical Conductivity (EC), salinity, and hardness. The WQI provides a single value for water quality suitability, is useful for the MWA in water supply production, and reflects the constituents that may be harmful to the consumer and the efficiency of production. For example, if the salinity in raw water is greater than 0.25 mg/l, the salt content of the water cannot be removed by the MWA system, resulting in a water supply with a salty flavor. Therefore, if the deteriorated water quality of raw water is predicted in advance, measured at a point upriver from the Samlae intake station, appropriate solutions can be put in place. Using PCD water quality records, the average WQI score of the Chao Phraya River in the rainy season is 74.63 and in the dry season is 78.37 (from a full score of 100). These values indicate the deterioration of raw water quality during the low flow period, a situation that the MWA should acknowledge and make periodical inspections to identify and control.

Keywords: MWA, Rating Curves of Water Quality Parameters, Surface Water Quality Standard, WQI, Water Supply Production.

1. INTRODUCTION

Water is an essential resource for living organisms in various ways. However, several human activities such as the excretion of waste in daily life and wastewater discharge from agricultural and industrial activities lead to the deterioration of water quality. Contamination of water sources causes the loss of water use benefits Pilailar and Urantinon (2019). As a result, the quality of life of living beings worsens.

To better understand the water quality in an area of interest, this should be determined by a combination of physical, chemical, and biological characteristics, and the integrated water situation should be evaluated using an appropriate technique, such as a Water Quality Index (WQI), the numerical summation of multiple water quality parameters into a single value. The WQI was calculated as a single value from multiple test results of water quality in a water basin by Horton (1965). It has since been widely applied for generating trends, evaluating, and communicating the overall water quality for the public to understand, and allowing comparisons among different watercourses or different locations in the same watercourse.

The WQI concept integrates all water quality parameters such as Biochemical Oxygen Demand (BOD) and Dissolved oxygen (DO) into scores that can indicate water quality for multiple purposes. The 4 steps for developing WQI are (1) selecting a set of water quality parameters of interest, (2) developing sub-indices-transforming the different units and dimensions of water quality parameters to a common scale, (3) assigning weights to the water quality parameters based on their relative importance to overall water quality, and (4) aggregating sub-indices to produce an overall index Boyacıoğlu (2009).

In Thailand, the WQI has been used to research and develop water quality standards for the Chao Phraya River at the raw water intake station of the MWA in Pathum Thani Province. The parameters were measured in physical, chemical, and biological properties. It was found that the water quality was satisfactory since the raw water was self-purified along the canal that transports the water to the production plant Thamkasem et al., (1986). However, at the production plant, turbidity, smell, color, and bacteria remain to be removed for making safe-drinking water, following World Health Organization (WHO) water quality standards guidelines World Health Organization (2011).

In 1995, the water quality management division of the Pollution Control Department (PCD) of the Metropolitan Waterworks Authority (MWA) modified the WQI of Brown Brown et al., (1970) to consider the basis of research in this field for evaluating the overall water quality in Thailand's rivers Simachaya (2000). Following their suggestions, unweighted WQI was applied for assessing the overall water quality by the PCD, in which all water quality parameters of interest were assumed to have the same level of importance. Before determining the WQI in each area, the values of each water quality parameter included in the WQI model are converted into sub-index scores between 0 and 100 using the rating curve technique developed for Thailand's rivers Landwehr (1979).

In 2007, a WQI was developed for the Chao Phraya River by selecting the data from seven water quality monitoring stations. Thirteen parameters were included in cluster type analysis (CA), factor analysis (FA), multiple regression, and statistical processes to analyze the data to select the parameters suitable for the setting of the WQI Prakirake et al., (2009).

Although WQI has been widely used in evaluating water quality in several countries, the usefulness of the WQI depends on its purpose. In 2015, an assessment of the water quality of the Ganges River in India using the WQI tool was undertaken, and the WQI values enabled an easy understanding of the trends of the water quality. The results pointed towards the requirement for urgent plans to prevent river water pollution. In 2018, Ahmed Hamdan created a WQI map based on Geographic Information System (GIS) for the Shatt Al-Arab River and its branches in Iraq to inform local authorities and convince them to take action Meher et al., (2015).

In Thailand, a WQI to measure the water quality situation of the six river basins was created by the PCD, referred to as the PCD-WQI which included eleven water quality parameters. A sampling station in Phra Nakhon Si Ayutthaya was the sampling point on the Chao Praya River. The PCD-WQI value was also able to be used to monitor the management of water resources for sustainability Tiyasha et al., (2020), and has been used to assess the chemical, physical, and microbiologic features of water bodies in temperate latitudes and was developed as an ecosystem-specific WQI, termed the ES-WQI Gradilla-Hernández et al., (2020).

In the current study, a set of water quality parameters was used to measure raw water quality in the Chao Phraya River at nine PVD stations along 180 kilometers of the river from Chainat to the Samlae raw water pumping station in Pathum Thani province. The variation in raw water quality due to the seasonal change has been a concern. To measure this seasonal variation, water quality assessment was divided into the rainy season, June to September, and the dry season, October to February. Based on the surface water quality standard of the PCD and the drinking water quality standards of the WHO water quality standards guidelines, 13 water quality parameters including color, turbidity, BOD, DO, pH, NO_3-N, Fe, Mn, TDS, TP, EC, salinity, and hardness were selected.

A time-series method was applied to analyze the water samples and measure the quality parameters. A time series method is an efficient and straightforward method for analyzing the past behavior of a time-series variable to predict its future behavior when independent causal variables influencing the time-series variable are unknown or cannot be determined. This technique as applied to the creation of the WQI requires uncomplicated data sets and less time for computation. In addition, other factors influencing water quality (e.g., changes in land use and population growth) were assumed to indirectly reflect the changing patterns of the 13 water quality parameters. The modified WQI was used to evaluate the raw water quality at MWA water supply production facilities. In addition, the WQI indicates the tendency for the deterioration of raw water over time, informing MWA awareness and planning for future, periodical inspections of the raw water at their inspection points.

The purpose of this study was to determine the raw water quality before the MWA analysis starting at raw water receival locations, for raw water from the Chao Phraya River into the Eastern Water Supply Canal at Samlae, Pathum Thani Province. The raw water is pumped and transmitted to the Bangkhen Water Treatment Plant (WTP), Thon Buri WTP, and the Samsen WTP. If the MWA can ascertain the water quality of raw water, in the form of the WQI, before the pumping station, it would be able to better manage the water supply production system.

2. STUDY AREA AND WATER QUALITY MEASURING STATION

Raw water used in the MWA water treatment process is taken from the Chao Phraya River at the Samlae Pumping Station which is located at Samlae in Pathum Thani Province, about 18 km north of the Bangkhen Water Treatment Plant. The Samlae Pumping Station is the first point that takes water from the Chao Phraya River into the MWA's east canal and is located 41 kilometers away from Bangkok, and 90 kilometers away from the Gulf of Thailand (Figure 1). The capacity of the station is 3.8 million cubic meters per day.

The data used to analyze the WQI values were obtained from 9 PCD water quality sampling stations upstream from the Samlae raw water pumping station to the Chao Phraya dam in Chainat province, as depicted in Figure 2 with location details in Table 1. In ordinary conditions, the PCD measures the quality of raw water four times per year.

NUEJ
Naresuan University
Engineering Journal

Figure 1 Samlae raw water pumping station, Pathum Thani

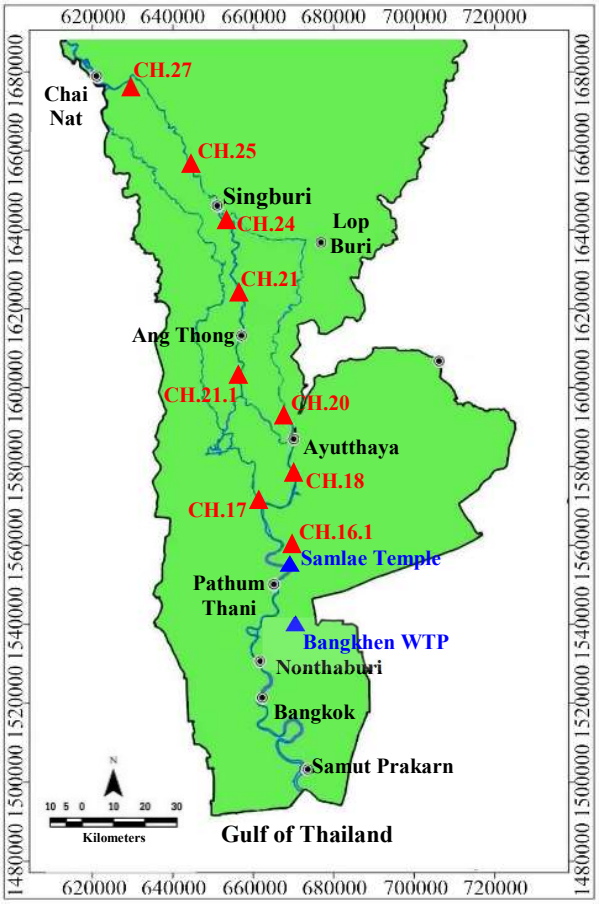


Figure 2 Location of the 9 PCD water quality sampling stations

The PCD in the Bureau of Water Quality Management monitored the water quality at 18 stations along the Chao Phraya River twice in the rainy season and twice in the dry season, and 28 water quality parameters were measured. In 2016, the PCD evaluated that the Chao Phraya River receives approximately 4.8 million cubic meters of wastewater per day from communities (10%), industrial plants (35%), aquaculture farms (10%), and swine farms (5%). As a result, the Biochemical Oxygen Demand Loading (BOD-loading) is approximately 369,421 kg/day. In the dry season, when there is no rainwater drainage of fertilizers and pesticides from the fields into the Chao Phraya River, the river receives the most waste from the community, industry, and agriculture, with the burden of the community waste the most. The potential discharging sources are residential areas, temples, government offices, schools, hospitals, and restaurants Pilailar and Urantinon (2019). These discharges cause high BOD with low DO, mainly as the river flows through Bangkok which is the most densely populated area along the Chao Praya River. Although The Chao Phraya River flowing through the Samlae raw water pumping station has water quality in the tier 3 criterion, it tends to deteriorate under the influence of rising tides allowing less freshwater to flow in the lower parts of the river.

Table 1 Location of the PCD's water quality sampling stations

Station	Distance from Samlae Pumping Station (KM.)	Water quality sampling station	District	Province
CH16.1	0	Samlae raw water pumping station	Mueang	Pathum Thani
CH17	5	Sam Khok Police Station	Sam Khok	Pathum Thani
CH18	28	Bang Pa-In Paper Factory	Bang Pa-In	Ayutthaya
CH20	47	Pom Phet	Mueang	Ayutthaya
CH21.1	69	Sa Kaeo temple	Pa Mok	Ang Thong
CH21	87	Ang Thong City Hall	Mueang	Ang Thong
CH24	131	The bridge crossing the Chao Phraya River	Mueang	Sing Buri
CH25	148	In-Buri Bridge	In-buri	Sing Buri
CH27	181	Chao Phraya Dam	Sapphaya	Chai Nat

The recorded data from the Samlae raw water quality sampling station on the Samlae canal in Pathum Thani Province also shows the trend of BOD and DO, as shown in Figure 3. Those data show that the water river tends not to meet the criterion of raw water for water supply production.

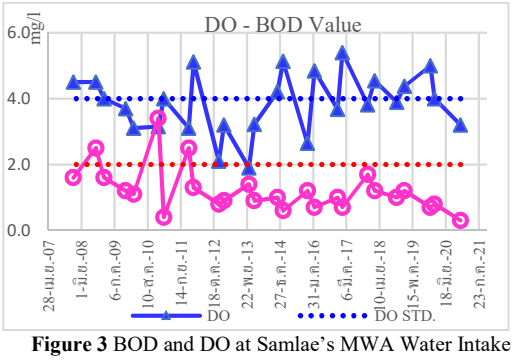


Figure 3 BOD and DO at Samlae's MWA Water Intake

3. METHODOLOGY

3.1 Data collection

The water quality data at nine stations along the Chao Phraya River, gathered by the PCD from 2008 to 2020, are shown in Table 1. These data include the water quality measures at the Samlae pumping station of the MWA which adopted these data based on the water quality standards guidelines that were recommended by several officials at the MWA, PCD, and WHO. The parameters related to the water supply products, previously discussed, were selected to consider the WQI for raw water.

3.2 Development of the WQI

The first step was the selection of the parameters that influence the water quality of raw water in the water supply production. This initial step started with examining data availability; the screening procedures and the multivariate statistical methods that were applied Prakirake et al., (2009). The rating curves of the selected parameter for each were then developed. The objective of this step was to transform the water quality parameter observations into a common scale (0 to 100).

In establishing the rating curves, expert judgment and statistical methods were used to develop the rating curves and define key points in agreement with the permissible limits from water quality standards. The actual parameter transformation was achieved through categorical scaling and linear interpolation. The weights corresponding to each parameter, ranging from 0 and 1 , were assigned according to their importance to the overall water quality. To assign weights to the parameters, participatory-based approaches that have been widely used in the Delphi method and the Analytical Hierarchy Process were adopted, and are admittedly subjective.

The fourth and final step was to select the weighted sub-index values' aggregation method and the formula WQI value. The standard formula for the WQI aggregation method, Horton (1965) is shown in Eq. (1)

$$WQI = \sum_{i=1}^{n} Wi \times Qi \quad (1)$$

where Qi is the water quality parameter s' index i
Wi is the associated weighing factor for each parameter
n is the number of parameters

This equation gives the result of a single value of WQI, which ranges from 0 to 100. It classifies the water body's quality into five different categories.

3.3 Procedure of WQI

As stated in the introduction section, there are four steps for the development of WQI. Figure 4 summarizes the methods used for each step in the WQI modification for raw water specific to MWA.

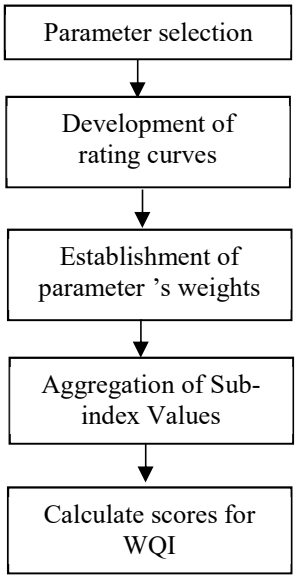


Figure 4 Procedure of Establishing WQI for Raw Water

3.4 Parameter selection

In this lateral study, the raw water quality data from the PCD was collected 4 times per year, twice in the rainy season, June to September, and twice in the dry season, October to February. These data can be classified for their significance in the parameters in the rainy season and dry season. Therefore, the weighting criteria scores of significant parameters will be different, water quality parameters and given scores (ranging from the poorest, 0, to the best, 100) to each range of values of their selected parameters. Other information influencing water quality (e.g., land use change, water consumption, and human activities along the riversides) were also aggregated from the water quality experts' opinions. The data obtained from the water quality experts and Brown et al., (1970) were used to create the rating curves for the 13 water quality parameters previously described The thirteen rating curves were tested for a year using the values of the associated water quality parameters observed at the nine sampling stations and with the recorded data from 2008 to 2020, and their suitability was adjusted for better determination of the water conditions in the Chao Phraya River.

In general, expert judgment can be incorporated into the selection of parameters through three approaches; individual interviews, interactive groups, and the Delphi method Meyer and Booker (1990). Of the three approaches, the Delphi method is the one that has been widely used for the selection of parameters Prakirake et al., (2009). The purpose of this method is to mine the view or opinions of experts without having the experts congregate at an agreed time and define the Delphi method Linstone and Turoff (1976).

3.5 Development of rating curves

The observed values of each water quality parameter were converted into sub-index scores for the parameter. To assign the sub-index values, the water quality parameters are presented as a two-dimensional (X-Y coordinate) plot where the sampling data in the X-Axis and Q-value within the Y-Axis, ranging on the scale of 0 to 100, where 0 and 100 represented the poorest and the highest of the water quality.

The construction of WQI was developing the rating curves for each of the selected parameters. The rating curves were extended to reach the complete concentration distribution ranges of the water quality parameters. Additionally, the rating curve for pH was horizontally displaced, so the maximum suitable pH, (value = 7) sub-index value was assigned to 100% of the Q-value in the Y-Axis.

Based on these criteria, the sub-index transformation curves for each selected parameter were assigned. Then, rating curve equations (sub-index functions) for each parameter were numerically formulated by curves fitting and approximation of function method using Microsoft Excel®. To aid the rating equation, scores for selected parameters were then approximated Uddin et al., (2021).

The techniques for creating similar sub-index transformation curves for water quality Prakirake et al., (2009).

3.6 Establishment of parameter's weights

The third step for constructing the WQI was to establish the weights of the 13 selected parameters to measure their relative importance to the final index value. Weights were assigned only to the desired parameters. A biplot was used to interpret these components and each of the selected water quality parameters' vectors. The length of the vector from the origin to the coordinates reflects the variance of that variable. To determine the weights of each selected parameter, the range scale was determined and normalized by the sum of all the range scales. Thus, the sum of all normalized values.

In this study, the weights of each selected parameter were determined through seasonal change, dry season, and rainy season. A Temporary Weight (TW) was developed by dividing the Significant Rating (SR) of each parameter. The weighting Scale (WS) for each parameter was defined as the ratio of the TP weight to the summation of the TW weights. The WS was obtained by approximating the ratio of the TW for each variable to the summation of the TW values Prakirake et al., (2009) TW and WS of the thirteen parameters according to Eq. (2) and Eq. (3)

$$TW = \frac{Highest\ of\ Significant\ Rating}{Each\ Significant\ Rating} \quad (2)$$

$$WS = \frac{Eac\ \ Temporary\ Weight}{\sum Temporary\ Weight} \quad (3)$$

In addition, the thirteen parameters were selected for the development of WS through seasonal consideration: dry season and rainy season.

The importance values and roles, for the dry season, were conductivity, salinity, and DO, and BOD with WS of 0.10 considered the most important variable for determining water supply quality while hardness and TDS with WS of 0.08 were the second most important, as shown in Table 2.

In the rainy season, TDS and turbidity with WS of 0.10 were considered the most important variables for determining water supply quality while color, hardness, Fe, NO_3-N, BOD, DO, and TP with WS of 0.08 were of lesser importance, as shown in Table 3.

Table 2 Selected Parameters Weights in Dry Season

Parameter	Unit	Weighting Scale		
		SR	TW	WS
pH	-	1.40	0.70	0.07
Color	Pt-Co	1.40	0.70	0.07
Hardness	mg/l	1.60	0.80	0.08
TDS	mg/l	1.60	0.80	0.08
Fe	mg/l	1.20	0.60	0.06
Mn	mg/l	1.20	0.60	0.06
NO_3-N	mg/l	1.20	0.60	0.06
BOD	mg/l	2.00	1.00	0.10
DO	mg/l	2.00	1.00	0.10
Total PO_4-P	mg/l	1.40	0.70	0.07
Conductivity	(µS)	2.00	1.00	0.10
Salinity	ppt	2.00	1.00	0.10
Turbidity	NTU	1.40	0.70	0.07
		∑	10.20	1.00

Table 3 Selected Parameters Weights in Rainy Season

Parameter	Unit	Weighing Scale		
		SR	TW	WS
pH	-	1.00	0.50	0.05
Color	Pt-Co	1.60	0.80	0.08
Hardness	mg/l	1.60	0.80	0.08
TDS	mg/l	2.00	1.00	0.10
Fe	mg/l	1.60	0.80	0.08
Mn	mg/l	1.40	0.70	0.07
NO_3-N	mg/l	1.60	0.80	0.08
BOD	mg/l	1.60	0.80	0.08
DO	mg/l	1.60	0.80	0.08
Total PO_4-P	mg/l	1.60	0.80	0.08
Conductivity	(µS)	1.40	0.70	0.07
Salinity	ppt	1.40	0.70	0.07
Turbidity	NTU	2.00	1.00	0.10
		∑	10.20	1.00

3.7 Classification for categories of water quality

The WQI scores for all stations on the river were used to categorize the overall water quality into classes following Notification No. 8: Surface water quality standard, the 1992 Thailand Enhancement and Conservation of National Environmental Quality Act. The five surface water quality classes with specific characteristics are shown in Table 4.

Table 4 Categories of Water Quality and WQI Scores

Class	Descriptor Categories	Score	Descriptive
1	Very good	85 - 100	Extra clean freshwater resources. Consumption requires only ordinary processing for pathogenic destruction.
2	Good	80 - 85	Very clean freshwater resources. Consumption requires an ordinary water treatment process which minor purification required before use.
3	Average	65 - 80	Medium clean freshwater resources. Consumption requires a conventional water treatment process before use.
4	Poor	40 - 65	Fairly clean freshwater resources. Consumption requires a specific or advanced water treatment process before use.
5	Very poor	< 40	Unacceptable

Considering the relationship between surface water quality standard categories and the range scales of the selected water quality parameter, the maximum-average-minimum of WS scores for the water collected was used as the baseline classification. The actual water quality is in the range of very poor to very good. Ranges of scale were adjusted using the average value of WS obtained from the MWA's Samlae raw-water intake, and the water quality data from the PCD. For the rating curve for each water quality classification, the scores of the very good class for each WQI were taken as the maximum data point of the "good" class. The maximum data point of the "good" class was 83, which may be approximated as the baseline of the very good class score (score over 85). Therefore, range scales for the good and average for the individual class were designated, respectively, as the average scores to upper limit scores and average scores of the "average" to average values of the "good" class. Therefore, the range of the "poor" class spanned between average scores of the "average" and "very poor" class is shown in Table 5.

Table 5 relationship of selected parameters and range scale

Parameter	Unit	Water quality Classification for water treatment plant				
		Very good	Good	Average	Poor	Very poor
pH	-	7	6.8 – 7.2	6.5 - 8.5	< 5.0 > 9.0	< 2.0 >12.0
Color	Pt-Co	≤ 10	20	100	200	> 200
Hardness	mg/l as $CaCo_3$	≤ 20	50	100	200	400
TDS	mg/l	≤ 200	1,000	1,200	1,500	> 1,500
Fe	mg/l	≤ 1	0.5	50	70	> 70
Mn	mg/l	≤0.5	1	5	10	> 10
NO_3-N	mg/l	≤ 2.5	5	10	15	> 15
BOD	mg/l	≤ 1	1.5	2.00	2.5	5
DO	mg/l	≥ 6.6	5.32	4.00	2.28	< 2.0
Total PO_4-P	mg/l	≤ 0.04	0.15	0.3	0.7	> 0.7
Conductivity	(µS)	50	200	500	1,000	1,500
Salinity	ppt	0.05	0.15	0.25	1.50	2.00
Turbidity	NTU	< 8	20	75	144	> 144

4. RESULT AND DISCUSSION

The relationship between the observed values from nine water quality sampling stations of the PCD and water quality sampling data from the MWA Samlae raw-water pumping station were converted into rating curves for each water quality parameter. These rating curves were modified from the relationship of the selected parameters and the range scales and categories of the surface water quality standards, presented as two-dimensional (X-Y coordinates), as illustrated in Figures 5-1, 5-2 up to 5-13. The associated equations for converting to a score were numerically formulated by curve fitting and the approximation of function method using Microsoft Excel®. With the aid of the rating equation, scores for each selected parameter were then approximated, as shown in Table 6.

The Q-value for all water quality parameters in this study evaluated the water quality, where input the water quality sampling data in the X-Axis and the Q-value within the Y-Axis, ranging on the scale of 0 to 100, where 0 and 100 represented the poorest and the highest of the water quality.

The values of the thirteen parameters obtained in the analysis of the average of the water quality sampling data are shown in Table 5. The highest are TDS, hardness, and BOD. It can be seen that water quality at that time is still good or very good. However, there is local pollution upstream, especially in the social and industrial expansion areas.

For long-term analysis, surface water quality data obtained between 2005 and 2020 from the natural water resources intended for water supply, from the Chao Phraya River at the upper MWA Samlae Raw-water Intake, Amphur Muang Pathum Thani province, were analyzed and evaluated using the WQI. The water qualities of the Chao Phraya River in general were average.

Table 6 Equations of sub-index rating curve for water quality parameters' index

Parameter	Equation
pH	$Qi = -21.429x^2 + 300x - 967.14$
Color	$Qi = 0.0006x^2 - 0.4706x + 92.424$
Hardness	$Qi = 0.0008x^2 - 0.5581x + 102.6$
TDS	$Qi = -9E\text{-}06x^2 - 0.0358x + 105.09$
Fe	$Qi = -0.0015x^2 - 0.9562x + 98.91$
Mn	$Qi = 0.2362x^2 - 10.786x + 100.73$
NO_3-N	$Qi = 0.1099x^2 - 8.4305x + 117.88$
BOD	$Qi = -0.1056x^2 - 10.732x + 102.66$
DO	$Qi = 0.1233x^2 + 16.848x - 15.933$
Total PO_4-P	$Qi = 255x^2 - 309.27x + 94.929$
Conductivity	$Qi = 3E\text{-}05x^2 - 0.1098x + 100.64$
Salinity	$Qi = 14.341x^2 - 69.274x + 86.549$
Turbidity	$Qi = -4E\text{-}06x^2 - 0.6169x + 96.173$

Where x is the water quality from the raw water quality data

In this study, the WQI was evaluated against that reported by the MWA at the Samlae raw-water intake. The results suggested that the WQI value, calculated from the selected parameters from the water sampling station during the dry season, (1st October to 30th April) is appropriate for the water supply. The WQI score was 78.37, and the range of the WQI in categories in class 3. In the rainy season, (1st May to 31st August) the calculated WQI score was 74.63, and was in the same range of WQI as in the dry season, as illustrated in Tables 7 and T8. The water quality was average to good because the upstream of Chao Phraya River was less contaminated by pollution point sources throughout the river, including communities, industrial plants, aquaculture, and various animal farms than downstream.

WQI Calculation and parameters analysis should be considered the peak or critical event of parameter values. For example, in February 2021, drought and the incursion of saltwater in river basins caused a crisis in Thailand, resulting in more discharge from upstream reservoirs to manage such problems. Upstream water from the Chao Phraya River basin was not enough to push seawater intrusion back at the MWA Samlae Raw-water Intake.

Thus, seawater contamination in raw water leads to increasing salinity of the water supply (MWA raw water standard salinity value $\leq$ 0.25 g/l), and water treatment plants cannot remove saltiness from the raw water. In this situation, water quality is not up to potable water standards which may affect consumers and cause damage to the country's industrial sector in the future.

The WQI score was applied in evaluating the trend of water quality. Results are illustrated in Figure 6. It can be seen that in the dry season, the WQI value tends to be better than the WQI values in the rainy season. The rainy season WQI may be influenced by soil erosion, chemicals, and industrial and community effluents that flow into the raw water sources. If considering the trend of annual average discharge value. It can be seen that the trend of 2018-2020 was a continuous decrease in runoff volume. This will affect the WQI values for both the dry and rainy seasons during 2018-2020. The WQI values in the dry season were 78.89, 81.79, 82.58, 81.76, 81.14, 78.70, and 67.94, the WQI values in the rainy season were 68.75, 74.07, 78.73, 79.59, 79.12, 77.20, 76.52 and 68.50 (forecast value), respectively which seems to indicate that the WQI value is relative to the decrease of annual discharge.

If the forecasting value is true, the water quality became poor water, fairly clean freshwater resources for consumption purposes require a conventional water treatment process or advanced water treatment process before being used since the water quality was under average level.

Typically, the climate variability effect or drought encountered, when evaluating water quality, yields obscure water quality. However, the index proposed shows the effect of climate variability on water quality prediction, in particular the raw water quality of the water treatment process.

5. CONCLUSIONS

The established WQI can be used instead of the traditional water quality evaluation since the developed WQI requires only 13 selected parameters rather than the usual 42 parameters generally used. This illustrates that the WQI provides a more convenient way to understand water quality than a long list of numerical values for a group of parameters. This enables the public and non-scientific communities to share and understand the data monitored. The WQI can assist those in charge of the water resource to make more efficient use of the Decision Support System (DSS) of the authorities to improve water quality to sustain water supply usage.

Overall, it may be inferred that the WQI may serve as a tool for describing water quality in terms of the specific water usage for the water supply. The WQI may also be used as a tool for measuring water quality trends in the water source. The water supply index developed in this study and the minimum operators may be used as an indicator reflecting the presence of pollutants in the river or water resource.

As a result, the MWA should be aware of the need to periodically inspect and sample the raw water, and to develop a tool and methods to solve the problem more efficiently and thereby reduce the impact of water pollution more successfully than before. Appropriate tools and strategies, such as those developed in this study, can solve the various problems, particularly saltwater intrusion, more effectively.

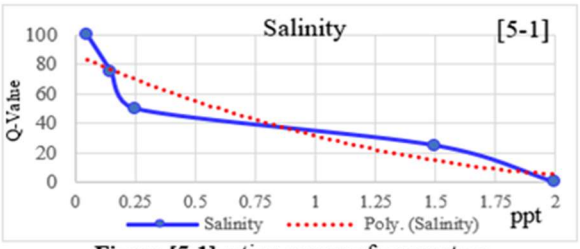


Figure [5-1] rating curves of parameters

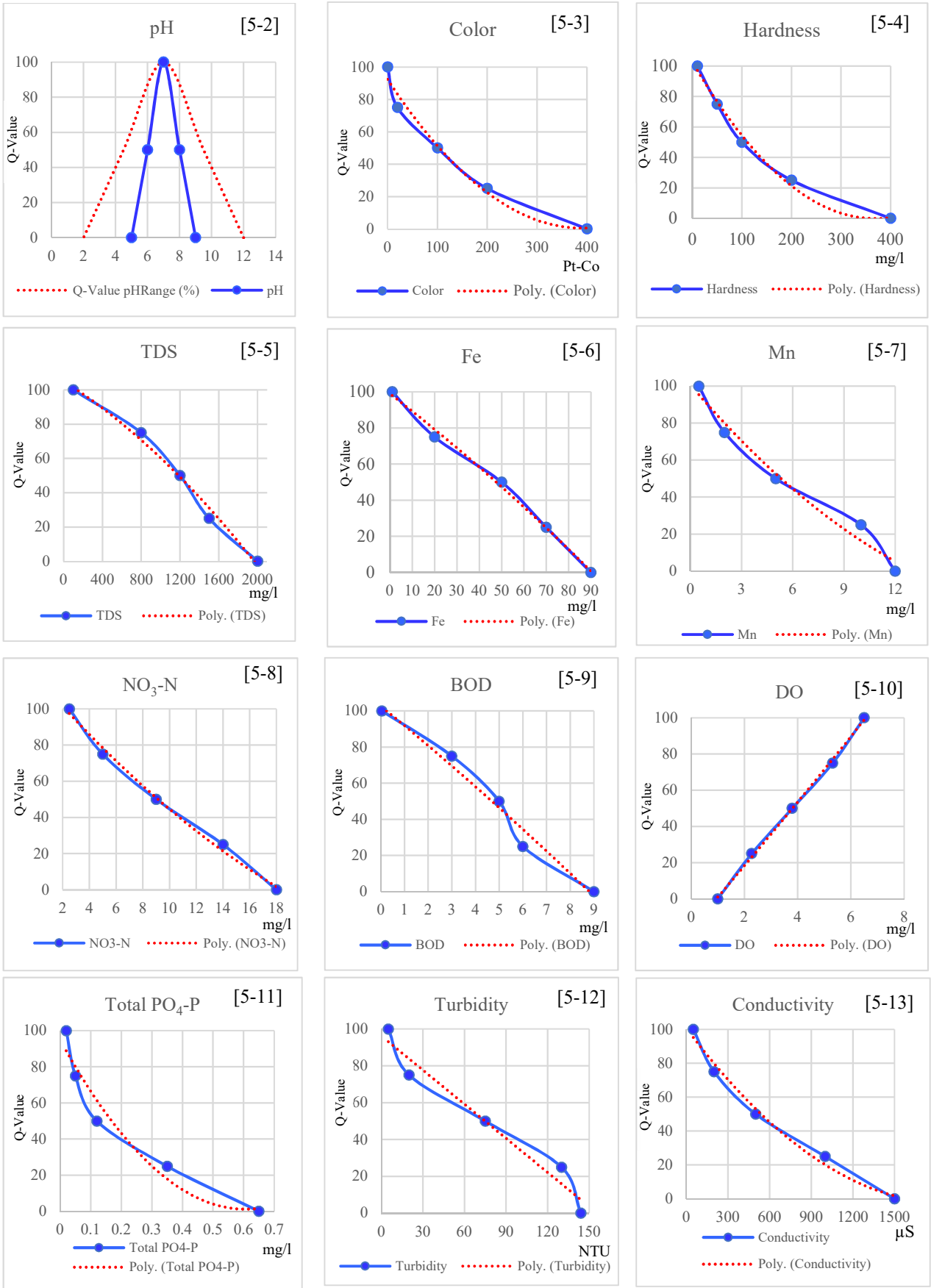


Figure [5-2] – [5-13] Rating curves of water quality parameters

Table 7 the average WQI value in the dry season (1st September – 30th April, 2010 - 2021)

Parameter	Unit	Weighting Scale			Sampling data	Q-Value	WS*Q-Value
		SR	TW	WS			
pH	-	1.40	0.70	0.07	7.35	80.270	5.51
Color	Pt-Co	1.40	0.70	0.07	1.00	91.954	6.31
Hardness	mg/l	1.60	0.80	0.08	101.98	54.006	4.24
TDS	mg/l	1.60	0.80	0.08	278.61	94.417	7.41
Fe	mg/l	1.20	0.60	0.06	0.89	98.056	5.77
Mn	mg/l	1.20	0.60	0.06	0.09	99.804	5.87
NO_3-N	mg/l	1.20	0.60	0.06	1.30	100.000	5.88
BOD	mg/l	2.00	1.00	0.10	1.85	82.444	8.08
DO	mg/l	2.00	1.00	0.10	3.86	50.917	4.99
Total PO_4-P	mg/l	1.40	0.70	0.07	0.10	65.421	4.49
Conductivity	(µS)	2.00	1.00	0.10	280.96	72.159	7.07
Salinity	ppt	2.00	1.00	0.10	0.20	73.268	7.18
Turbidity	NTU	1.40	0.70	0.07	24.50	81.054	5.56
		∑	10.20	1.00		WQI =	78.37

Table 8 the average WQI value in the rainy season (1st May – 31st August, 2010 - 2021)

Parameter	Unit	Weighting Scale			Sampling data	Q-Value	WS*Q-Value
		SR	TW	WS			
pH	-	1.00	0.50	0.05	7.53	76.816	3.77
Color	Pt-Co	1.60	0.80	0.08	11.00	87.320	6.85
Hardness	mg/l	1.60	0.80	0.08	40.00	81.556	6.40
TDS	mg/l	2.00	1.00	0.10	1113.00	54.096	5.30
Fe	mg/l	1.60	0.80	0.08	0.47	98.460	7.72
Mn	mg/l	1.40	0.70	0.07	0.02	100.514	6.90
NO_3-N	mg/l	1.60	0.80	0.08	2.26	100.000	7.84
BOD	mg/l	1.60	0.80	0.08	2.00	80.774	6.34
DO	mg/l	1.60	0.80	0.08	4.10	55.216	4.33
Total PO_4-P	mg/l	1.60	0.80	0.08	0.01	92.777	7.28
Conductivity	(µS)	1.40	0.70	0.07	2232.00	5.021	0.34
Salinity	ppt	1.40	0.70	0.07	0.40	61.134	4.20
Turbidity	NTU	2.00	1.00	0.10	34.00	75.194	7.37
		∑	10.20	1.00		WQI =	74.63

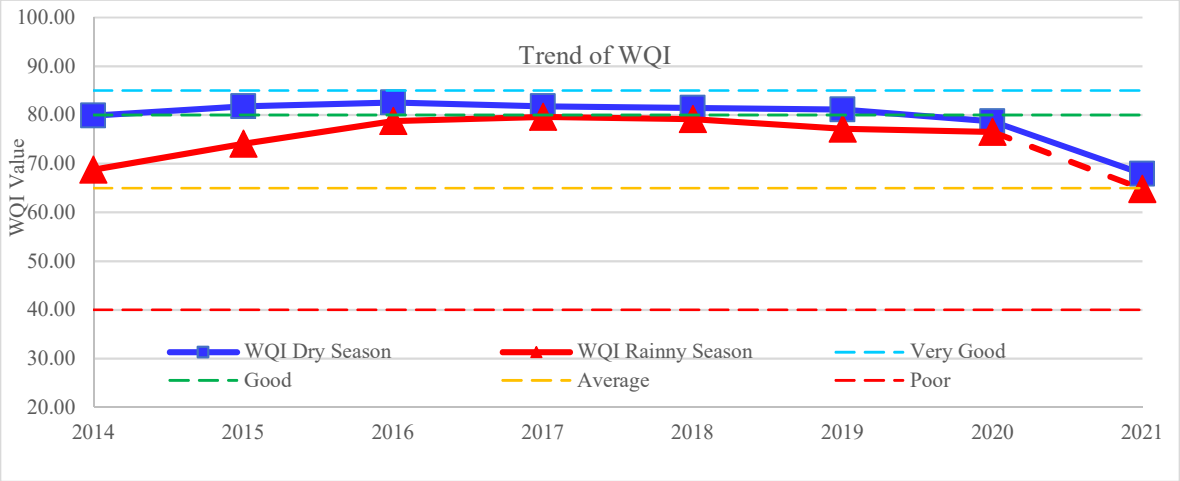


Figure 6 Trend in water quality of WQI scores for the dry season and rainy season between 2014 – 2021

6. ACKNOWLEDGMENTS

The author wishes to thank the Nationwide surface water quality database system (IWIS), Pollution Control Department (PCD), and Metropolitan Waterworks Authority (MWA) for providing water quality data for the years 2008-2020 and 2005-Feb 2021. In addition, the authors would like to acknowledge Assistant Professor Dr. Sitang Pilailar and Professor Dr. Suwatana Chittaladakorn for their valuable comments and constructive discussion. We also thank Mr. Roy I. Morien of the Naresuan University Graduate School for his editing of the grammar, syntax and general English expression in this paper.

7. REFERENCES

Brown, R. M., McClelland, N.I., Deininger, R.A., & Tozer, R.Z. (1970). A water quality index-do we dare?, *Water& Sewage Works. 117*, 339-343.

Boyacıoğlu, H. (2009). Development of a water quality index based on a European classification scheme. *Water SA*, *33*, 101-106. doi:https://doi.org/10.4314/WSA.V33I1.47882

Gradilla-Hernández, M. S., de Anda, J., Garcia-Gonzalez, A., Montes, C. Y., Barrios-Piña, H., Ruiz-Palomino, P., & Díaz-Vázquez, D. (2020). Assessment of the water quality of a subtropical lake using the NSF-WQI and a newly proposed ecosystem specific water quality index. *Environmental Monitoring and Assessment, 192(5)*,296.doi:https://www.doi.org/10.1007/s10661-020-08265-7

Horton, R. K. (1965). An index-number system for rating water quality. *Journal of Water Pollution Control Federation, 37(3)*, 300-306.

Landwehr, J. M. (1979). A statistical view of a class of water quality indices. *Water Resources Research, 15*(2), 460-468. doi:https://doi.org/10.1029/WR015i002p00460

Linstone, H., & Turoff, M. (1976). The Delphi Method: Techniques and Applications. *Journal of Marketing Research*, *13*(3), 317–318. doi:https://www.doi.org/10.2307/3150755

Meher, P., Sharma, P., Gautam, Y. P., Kumar, A., & Mishra, K. P. (2015). Evaluation of Water Quality of Ganges River Using Water Quality Index Tool. *Environment Asia, 8*, 124-132. doi:https://www.doi.org/10.14456/ea.2015.15

Meyer, M. A., & Booker, J. M. (1990). *Eliciting and analyzing expert judgment: A practical guide*. United States. doi: https://doi.org/10.2172/5088782

Pilailar, S., & Urantinon, A. (2019). Chao Phraya River contamination sources investigation and water quality capacity in dry season evaluation. *Journal of Environmental Management, 15*(1), 62-83. doi:https://www.doi.org/10.14456/jem.2019.4

Prakirake, C., Chaiprasert, P., & Tripetchkul, S. (2009). Development of specific water quality index for water supply in Thailand. *Songklanakarin Journal of Science and Technology, 31*, 91-104. https://www.scopus.com/inward/record.uri?eid=2-s2.067650922100&partnerID=40&md5=80437ee843645c58e8a67fad184c1ee3

Simachaya, W. (2000). *Water quality criteria and standards of water quality in Thailand*. Paper for the Workshop on "Environmentally Sound Technology on Water Quality Management" UNEP, Mekong River Commission (MRC), Bangkok. https://www.pcd.go.th/waters/%E0%B9%80%E0%B8%AD%E0%B8%81%E0%B8%AA%E0%B8%B2%E0%B8%A3-water-quality-management-in-thailand

Thamkasem, T., Chaotanont, P. & Naenna, P. (1986, January 27-29). Water quality of The Chao Phra Ya River at the MWA raw water pump station. Tambol Samlae, Amphor Muang, Pathumthani. *24th Academic Conference of Kasetsart University in Environment, Thiamkomkris Building, Faculty of Forestry*, *Kasetsart University, Bangkok,* Part 8, 1-15.https://kukrdb.lib.ku.ac.th/proceedings/KUCON/search_detail/dowload_digital_file/3944/7096

Tiyasha, Tung, T. M., & Yaseen, Z. M. (2020). A survey on river water quality modelling using artificial intelligence models: 2000–2020. *Journal of Hydrology, 585*, 124670. doi:https://doi.org/10.1016/j.jhydrol.2020.124670

Uddin, M. G., Nash, S., & Olbert, A. I. (2021). A review of water quality index models and their use for assessing surface water quality. *Ecological Indicators, 122*, *107218*, 1-21. doi:https://doi.org/10.1016/j.ecolind.2020.107218

World Health Organization. (2011). *Guidelines for Drinking-water Quality, Fourth Edition.* WHO https://apps.who.int/iris/bitstream/handle/10665/44584/9789241548151_eng.pdf