



UNIVERSITI  
TEKNOLOGI  
MARA

Cawangan Kedah  
Kampus Sungai Petani



# e-PROCEEDINGS

of The 5<sup>th</sup> International Conference  
on Computing, Mathematics and  
Statistics (iCMS2021)

**4-5 August 2021**

**Driving Research Towards Excellence**



# **e-Proceedings of the 5<sup>th</sup> International Conference on Computing, Mathematics and Statistics (iCMS 2021)**

*Driving Research Towards Excellence*

Editor-in-Chief: Norin Rahayu Shamsuddin

Editorial team:

Dr. Afida Ahamad  
Dr. Norliana Mohd Najib  
Dr. Nor Athirah Mohd Zin  
Dr. Siti Nur Alwani Salleh  
Kartini Kasim  
Dr. Ida Normaya Mohd Nasir  
Kamarul Ariffin Mansor

e-ISBN: 978-967-2948-12-4

DOI

Library of Congress Control Number:

Copyright © 2021 Universiti Teknologi MARA Kedah Branch

All right reserved, except for educational purposes with no commercial interests. No part of this publication may be reproduced, copied, stored in any retrieval system or transmitted in any form or any means, electronic or mechanical including photocopying, recording or otherwise, without prior permission from the Rector, Universiti Teknologi MARA Kedah Branch, Merbok Campus. 08400 Merbok, Kedah, Malaysia.

The views and opinions and technical recommendations expressed by the contributors are entirely their own and do not necessarily reflect the views of the editors, the Faculty or the University.

Publication by  
Department of Mathematical Sciences  
Faculty of Computer & Mathematical Sciences  
UiTM Kedah

## TABLE OF CONTENT

### PART 1: MATHEMATICS

	Page
<b>STATISTICAL ANALYSIS ON THE EFFECTIVENESS OF SHORT-TERM PROGRAMS DURING COVID-19 PANDEMIC: IN THE CASE OF PROGRAM BIJAK SIFIR 2020</b> <i>Nazihah Safie, Syerrina Zakaria, Siti Madhahah Abdul Malik, Nur Bains Ismail, Azwani Alias Ruwaidiah Idris</i>	1
<b>RADIATIVE CASSON FLUID OVER A SLIPPERY VERTICAL RIGA PLATE WITH VISCOUS DISSIPATION AND BUOYANCY EFFECTS</b> <i>Siti Khuzaimah Soid, Khadijah Abdul Hamid, Ma Nuramalina Nasero, NurNajah Nabila Abdul Aziz</i>	10
<b>GAUSSIAN INTEGER SOLUTIONS OF THE DIOPHANTINE EQUATION <math>x^4 + y^4 = z^3</math> FOR <math>x \neq y</math></b> <i>Shahrina Ismail, Kamel Ariffin Mohd Atan and Diego Sejas Viscarra</i>	19
<b>A SEMI ANALYTICAL ITERATIVE METHOD FOR SOLVING THE EMDEN-FOWLER EQUATIONS</b> <i>Mat Salim Selamat, Mohd Najir Tokachil, Noor Aqila Burhanddin, Ika Suzieana Murad and Nur Farhana Razali</i>	28
<b>ROTATING FLOW OF A NANOFUID PAST A NONLINEARLY SHRINKING SURFACE WITH FLUID SUCTION</b> <i>Siti Nur Alwani Salleh, Norfifah Bachok and Nor Athirah Mohd Zin</i>	36
<b>MODELING THE EFFECTIVENESS OF TEACHING BASIC NUMBERS THROUGH MINI TENNIS TRAINING USING MARKOV CHAIN</b> <i>Rahela Abdul Rahim, Rahizam Abdul Rahim and Syahrul Ridhwan Morazuk</i>	46
<b>PERFORMANCE OF MORTALITY RATES USING DEEP LEARNING APPROACH</b> <i>Mohamad Hasif Azim and Saiful Izzuan Hussain</i>	53
<b>UNSTEADY MHD CASSON FLUID FLOW IN A VERTICAL CYLINDER WITH POROSITY AND SLIP VELOCITY EFFECTS</b> <i>Wan Faezah Wan Azmi, Ahmad Qushairi Mohamad, Lim Yeou Jiann and Sharidan Shafie</i>	60
<b>DISJUNCTIVE PROGRAMMING - TABU SEARCH FOR JOB SHOP SCHEDULING PROBLEM</b> <i>S. Z. Nordin, K.L. Wong, H.S. Pheng, H. F. S. Saipol and N.A.A. Husain</i>	68
<b>FUZZY AHP AND ITS APPLICATION TO SUSTAINABLE ENERGY PLANNING DECISION PROBLEM</b> <i>Liana Najib and Lazim Abdullah</i>	78
<b>A CONSISTENCY TEST OF FUZZY ANALYTIC HIERARCHY PROCESS</b> <i>Liana Najib and Lazim Abdullah</i>	89
<b>FREE CONVECTION FLOW OF BRINKMAN TYPE FLUID THROUGH AN COSINE OSCILLATING PLATE</b> <i>Siti Noramirah Ibrahim, Ahmad Qushairi Mohamad, Lim Yeou Jiann, Sharidan Shafie and Muhammad Najib Zakaria</i>	98

<b>RADIATION EFFECT ON MHD FERROFLUID FLOW WITH RAMPED WALL TEMPERATURE AND ARBITRARY WALL SHEAR STRESS</b>	<b>106</b>
<i>Nor Athirah Mohd Zin, Aaiza Gul, Siti Nur Alwani Salleh, Imran Ullah, Sharena Mohamad Isa, Lim Yeou Jiann and Sharidan Shafie</i>	

## **PART 2: STATISTICS**

<b>A REVIEW ON INDIVIDUAL RESERVING FOR NON-LIFE INSURANCE</b>	<b>117</b>
<i>Kelly Chuah Khai Shin and Ang Siew Ling</i>	
<b>STATISTICAL LEARNING OF AIR PASSENGER TRAFFIC AT THE MURTALA MUHAMMED INTERNATIONAL AIRPORT, NIGERIA</b>	<b>123</b>
<i>Christopher Godwin Udomboso and Gabriel Olugbenga Ojo</i>	
<b>ANALYSIS ON SMOKING CESSATION RATE AMONG PATIENTS IN HOSPITAL SULTAN ISMAIL, JOHOR</b>	<b>137</b>
<i>Siti Mariam Norrulashikin, Ruzaini Zulhusni Puslan, Nur Arina Bazilah Kamisan and Siti Rohani Mohd Nor</i>	
<b>EFFECT OF PARAMETERS ON THE COST OF MEMORY TYPE CHART</b>	<b>146</b>
<i>Sakthiseswari Ganasan, You Huay Woon and Zainol Mustafa</i>	
<b>EVALUATION OF PREDICTORS FOR THE DEVELOPMENT AND PROGRESSION OF DIABETIC RETINOPATHY AMONG DIABETES MELLITUS TYPE 2 PATIENTS</b>	<b>152</b>
<i>Syafawati Ab Saad, Maz Jamilah Masnan, Karniza Khalid and Safwati Ibrahim</i>	
<b>REGIONAL FREQUENCY ANALYSIS OF EXTREME PRECIPITATION IN PENINSULAR MALAYSIA</b>	<b>160</b>
<i>Iszuanie Syafidza Che Ilias, Wan Zawiah Wan Zin and Abdul Aziz Jemain</i>	
<b>EXPONENTIAL MODEL FOR SIMULATION DATA VIA MULTIPLE IMPUTATION IN THE PRESENT OF PARTLY INTERVAL-CENSORED DATA</b>	<b>173</b>
<i>Salman Umer and Faiz Elfaki</i>	
<b>THE FUTURE OF MALAYSIA'S AGRICULTURE SECTOR BY 2030</b>	<b>181</b>
<i>Thanusha Palmira Thangarajah and Suzilah Ismail</i>	
<b>MODELLING MALAYSIAN GOLD PRICES USING BOX-JENKINS APPROACH</b>	<b>186</b>
<i>Isnewati Ab Malek, Dewi Nur Farhani Radin Nor Azam, Dinie Syazwani Badrul Aidi and Nur Syafiqah Sharim</i>	
<b>WATER DEMAND PREDICTION USING MACHINE LEARNING: A REVIEW</b>	<b>192</b>
<i>Norashikin Nasaruddin, Shahida Farhan Zakaria, Afida Ahmad, Ahmad Zia Ul-Saufie and Norazian Mohamaed Noor</i>	
<b>DETECTION OF DIFFERENTIAL ITEM FUNCTIONING FOR THE NINE-QUESTIONS DEPRESSION RATING SCALE FOR THAI NORTH DIALECT</b>	<b>201</b>
<i>Suttipong Kawilapat, Benchlak Maneeton, Narong Maneeton, Sukon Prasitwattanaseree, Thoranin Kongsuk, Suwanna Arunpongpaisal, Jintana Leejongpermpool, Supattra Sukhawaha and Patrinee Traisathit</i>	

<b>ACCELERATED FAILURE TIME (AFT) MODEL FOR SIMULATION PARTLY INTERVAL-CENSORED DATA</b>	<b>210</b>
<i>Ibrahim El Feky and Faiz Elfaki</i>	
<b>MODELING OF INFLUENCE FACTORS PERCENTAGE OF GOVERNMENTS' RICE RECIPIENT FAMILIES BASED ON THE BEST FOURIER SERIES ESTIMATOR</b>	<b>217</b>
<i>Chaerobby Fakhri Fauzaan Purwoko, Ayuning Dwis Cahyasari, Netha Aliffia and M. Fariz Fadillah Mardianto</i>	
<b>CLUSTERING OF DISTRICTS AND CITIES IN INDONESIA BASED ON POVERTY INDICATORS USING THE K-MEANS METHOD</b>	<b>225</b>
<i>Khoirun Niswatin, Christopher Andreas, Putri Fardha Asa OktaviaHans and M. Fariz Fadilah Mardianto</i>	
<b>ANALYSIS OF THE EFFECT OF HOAX NEWS DEVELOPMENT IN INDONESIA USING STRUCTURAL EQUATION MODELING-PARTIAL LEAST SQUARE</b>	<b>233</b>
<i>Christopher Andreas, Sakinah Priandi, Antonio Nikolas Manuel Bonar Simamora and M. Fariz Fadillah Mardianto</i>	
<b>A COMPARATIVE STUDY OF MOVING AVERAGE AND ARIMA MODEL IN FORECASTING GOLD PRICE</b>	<b>241</b>
<i>Arif Luqman Bin Khairil Annuar, Hang See Pheng, Siti Rohani Binti Mohd Nor and Thoo Ai Chin</i>	
<b>CONFIDENCE INTERVAL ESTIMATION USING BOOTSTRAPPING METHODS AND MAXIMUM LIKELIHOOD ESTIMATE</b>	<b>249</b>
<i>Siti Fairus Mokhtar, Zahayu Md Yusof and Hasimah Sapiri</i>	
<b>DISTANCE-BASED FEATURE SELECTION FOR LOW-LEVEL DATA FUSION OF SENSOR DATA</b>	<b>256</b>
<i>M. J. Masnan, N. I. Maha3, A. Y. M. Shakaf, A. Zakaria, N. A. Rahim and N. Subari</i>	
<b>BANKRUPTCY MODEL OF UK PUBLIC SALES AND MAINTENANCE MOTOR VEHICLES FIRMS</b>	<b>264</b>
<i>Asmahani Nayan, Amirah Hazwani Abd Rahim, Siti Shuhada Ishak, Mohd Rijal Ilias and Abd Razak Ahmad</i>	
<b>INVESTIGATING THE EFFECT OF DIFFERENT SAMPLING METHODS ON IMBALANCED DATASETS USING BANKRUPTCY PREDICTION MODEL</b>	<b>271</b>
<i>Amirah Hazwani Abdul Rahim, Nurazlina Abdul Rashid, Abd-Razak Ahmad and Norin Rahayu Shamsuddin</i>	
<b>INVESTMENT IN MALAYSIA: FORECASTING STOCK MARKET USING TIME SERIES ANALYSIS</b>	<b>278</b>
<i>Nuzlinda Abdul Rahman, Chen Yi Kit, Kevin Pang, Fauhatuz Zahroh Shaik Abdullah and Nur Sofiah Izani</i>	

## **PART 3: COMPUTER SCIENCE & INFORMATION TECHNOLOGY**

- ANALYSIS OF THE PASSENGERS' LOYALTY AND SATISFACTION OF AIRASIA PASSENGERS USING CLASSIFICATION** 291  
*Ee Jian Pei, Chong Pui Lin and Nabilah Filzah Mohd Radzuan*
- HARMONY SEARCH HYPER-HEURISTIC WITH DIFFERENT PITCH ADJUSTMENT OPERATOR FOR SCHEDULING PROBLEMS** 299  
*Khairul Anwar, Mohammed A.Awadallah and Mohammed Azmi Al-Betar*
- A 1D EYE TISSUE MODEL TO MIMIC RETINAL BLOOD PERFUSION DURING RETINAL IMAGING PHOTOPLETHYSMOGRAPHY (IPPG) ASSESSMENT: A DIFFUSION APPROXIMATION – FINITE ELEMENT METHOD (FEM) APPROACH** 307  
*Harnani Hassan, Sukreen Hana Herman, Zulfakri Mohamad, Sijung Hu and Vincent M. Dwyer*
- INFORMATION SECURITY CULTURE: A QUALITATIVE APPROACH ON MANAGEMENT SUPPORT** 325  
*Qamarul Nazrin Harun, Mohamad Noorman Masrek, Muhamad Ismail Pahmi and Mohamad Mustaqim Junoh*
- APPLY MACHINE LEARNING TO PREDICT CARDIOVASCULAR RISK IN RURAL CLINICS FROM MEXICO** 335  
*Misael Zambrano-de la Torre, Maximiliano Guzmán-Fernández, Claudia Sifuentes-Gallardo, Hamurabi Gamboa-Rosales, Huizilopoztli Luna-García, Ernesto Sandoval-García, Ramiro Esquivel-Felix and Héctor Durán-Muñoz*
- ASSESSING THE RELATIONSHIP BETWEEN STUDENTS' LEARNING STYLES AND MATHEMATICS CRITICAL THINKING ABILITY IN A 'CLUSTER SCHOOL'** 343  
*Salimah Ahmad, Asyura Abd Nassir, Nor Habibah Tarmuji, Khairul Firhan Yusob and Nor Azizah Yacob*
- STUDENTS' LEISURE WEEKEND ACTIVITIES DURING MOVEMENT CONTROL ORDER: UİTM PAHANG SHARING EXPERIENCE** 351  
*Syafıza Saila Samsudin, Noor Izyan Mohamad Adnan, Nik Muhammad Farhan Hakim Nik Badrul Alam, Siti Rosiah Mohamed and Nazihah Ismail*
- DYNAMICS SIMULATION APPROACH IN MODEL DEVELOPMENT OF UNSOLD NEW RESIDENTIAL HOUSING IN JOHOR** 363  
*Lok Lee Wen and Hasimah Sapiri*
- WORD PROBLEM SOLVING SKILLS AS DETERMINANT OF MATHEMATICS PERFORMANCE FOR NON-MATH MAJOR STUDENTS** 371  
*Shahida Farhan Zakaria, Norashikin Nasaruddin, Mas Aida Abd Rahim, Fazillah Bosli and Kor Liew Kee*
- ANALYSIS REVIEW ON CHALLENGES AND SOLUTIONS TO COMPUTER PROGRAMMING TEACHING AND LEARNING** 378  
*Noor Hasnita Abdul Talib and Jasmin Ilyani Ahmad*

## **PART 4: OTHERS**

- ANALYSIS OF CLAIM RATIO, RISK-BASED CAPITAL AND VALUE-ADDED INTELLECTUAL CAPITAL: A COMPARISON BETWEEN FAMILY AND GENERAL TAKAFUL OPERATORS IN MALAYSIA** 387  
*Nur Amalina Syafiqa Kamaruddin, Norizarina Ishak, Siti Raihana Hamzah, Nurfadhlina Abdul Halim and Ahmad Fadhly Nurullah Rasade*
- THE IMPACT OF GEOMAGNETIC STORMS ON THE OCCURRENCES OF EARTHQUAKES FROM 1994 TO 2017 USING THE GENERALIZED LINEAR MIXED MODELS** 396  
*N. A. Mohamed, N. H. Ismail, N. S. Majid and N. Ahmad*
- BIBLIOMETRIC ANALYSIS ON BITCOIN 2015-2020** 405  
*Nurazlina Abdul Rashid, Fazillah Bosli, Amirah Hazwani Abdul Rahim, Kartini Kasim and Fathiyah Ahmad@Ahmad Jali*
- GENDER DIFFERENCE IN EATING AND DIETARY HABITS AMONG UNIVERSITY STUDENTS** 413  
*Fazillah Bosli, Siti Fairus Mokhtar, Noor Hafizah Zainal Aznam, Juaini Jamaludin and Wan Siti Esah Che Hussain*
- MATHEMATICS ANXIETY: A BIBLIOMETRIX ANALYSIS** 420  
*Kartini Kasim, Hamidah Muhd Irpan, Noorazilah Ibrahim, Nurazlina Abdul Rashid and Anis Mardiana Ahmad*
- PREDICTION OF BIOCHEMICAL OXYGEN DEMAND IN MEXICAN SURFACE WATERS USING MACHINE LEARNING** 428  
*Maximiliano Guzmán-Fernández, Misael Zambrano-de la Torre, Claudia Sifuentes-Gallardo, Oscar Cruz-Dominguez, Carlos Bautista-Capetillo, Juan Badillo-de Loera, Efrén González Ramírez and Héctor Durán-Muñoz*

# WATER DEMAND PREDICTION USING MACHINE LEARNING: A REVIEW

Norashikin Nasaruddin<sup>1</sup>, Shahida Farhan Zakaria<sup>2</sup>, Afida Ahmad<sup>3</sup>, Ahmad Zia Ul-Saufie<sup>4</sup>  
and Norazian Mohamaed Noor<sup>5</sup>

<sup>1,2,3</sup> Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (Kedah Branch), 08400 Merbok, Kedah, Malaysia. <sup>4</sup> Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450, Shah Alam, Selangor, Malaysia, <sup>5</sup> Sustainable Environment Research Group, Center of Excellence Geopolymer and Green Technology (CEGeoGTech), Universiti Malaysia Perlis, Kompleks Pengajian Jejawi 2, 02600 Arau, Perlis, Malaysia  
(<sup>1</sup> norashikin116@uitm.edu.my, <sup>2</sup> shahidafarhan@uitm.edu.my, <sup>3</sup> afidaahmad@uitm.edu.my, <sup>4</sup> ahmadzia101@uitm.edu.my, <sup>5</sup> norazian@unimap.edu.my)

Water is important and critical sources of life. Even though some countries enjoy tropical weather year-round with plenty of water resources like Malaysia, they are still facing scarcity issue. Water demand is influenced by various factors such as population, climate change and water utilization. This study reviews 45 Scopus articles from year 2015 to 2021 on predicting water demand using Machine Learning (ML) methods which include: neural network, random forest, decision tree, and hybrid optimisation models. The summary of ML methods on the evaluation of their performance in water demand prediction is identified by a comprehensive analysis of the literature. The narrative search of the most relevant literature is classified according to method, prediction type, prediction variables and accuracy rate. The review identified several machine learning methods that are commonly used which include decision tree, neural network, random forest and hybrid method. In conclusion, the study reports that the accuracy of the method varies according to types of prediction variables used.

**Keywords:** Water demand, Machine learning, Neural network, Decision tree

## 1. Introduction

Water is one of the important element and a critical source of life. With the growing of population as well as higher standard of living, the demand of water in daily use has also increased (Ali et al., 2014). However, some countries are still facing some problems with water scarcity issues even though they enjoy tropical weather year-round with plenty of water resources. Water demand is influenced by various factors such as population, climate change and water utilization (Ali et al., 2014).

Forecasting water demand with higher accuracy is useful to various parties for effective long-term planning in order to manage water systems efficiently and accommodating the increase in water demand. There are many uncertainties in predicting water demand and some of the challenges evolve from the relationship between human and natural systems in urban environments (House-Peters and Chang, 2011). Traditional methods such as time series and regression model were firstly used in forecasting water demand, however, the approaches have less accuracy (Zubaidi et al., 2020c). Recently, machine learning methods which include random forest, neural network, and support vector machine has been widely employed in the prediction of water demand.

There are limited studies done in Malaysia regarding water demand forecasting especially using machine learning method. Ali et al. (2014) and Mohd Firdaus and Talib (2016) investigated the pattern of water supply and demand in Langat catchment and Selangor area using Water Evaluation and Planning (WEAP) System. The Stockholm Environment Institute (SEI) created WEAP, a unique water resources and planning software that stimulates hydrologic patterns based on climatic input. This is a critical tool for informing society about climate change adaptation and policy making. WEAP also allows users to create scenarios using assumptions about water demand, infrastructure, and regulation, such as an increase in temperature or heavier rainfall. WEAP can combine all human activities in order to predict water shortages and water quality based on a model scenario. The latest study by Anang et al. (2019) used the pooled Ordinary Least Squares (OLS) to determine



the availability of water resources for 13 states in Malaysia. The pooling technique used compared changes in cross-sectional units to changes in individual units over time. It allows for more complex analysis than either cross-section or time series analysis alone. Furthermore, pooled regression uses pooled data sets to handle a larger number of data points.

This study reviews the current studies on the prediction of water demand. The review covers the application of ML methods (including optimization and hybrid models) and the discussions on the evaluation of their performance in water demand prediction. Narrative reviews of relevant literatures are classified according to method, prediction type, prediction variables and accuracy rate. The discussion and the conclusion of the reviews are presented in the last section of this paper.

## 2. Research Methodology

The SCOPUS online database has been assessed to identify the related documents on water demand forecasting using ML approaches. The database had been explored by using the advanced document search with keyword: (TITLE-ABS-KEY ( "water demand" ) AND TITLE-ABS-KEY ( "machine learning" OR "Deep learning" OR "ANN" OR "MLP" OR "ELM" OR "neural network" OR "ANFIS" OR "decision tree" ) ).

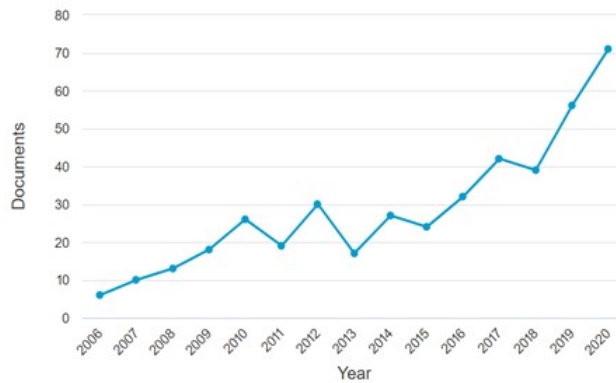


Figure 1: The number of SCOPUS publication in water demand prediction using Machine Learning approaches.

Figure 1 represents the number of SCOPUS publications in water demand prediction using ML approaches since 2006. The database shows an increment in the number of manuscripts published from year to year. In this paper, 46 relevant papers issued from 2015 to 2021 are reviewed. The revised papers are then categorised based on the ML model used which are Neural Network, Decision Tree, Random Forest, Hybrid, and Optimization models as illustrated in Section 3.

## 3. Machine Learning Models

The ML models presented in this paper include: 17 Artificial Neural Network, 9 Decision Tree, 4 Random Forest, 15 Hybrid and Optimisation. The review takes into consideration of the year the papers were published, the prediction variables and prediction type used in water demand analysis, the region in which the data were taken from, and lastly, the accuracy rate of the ML models chosen.

### 3.1 Artificial Neural-Network (ANN)

ANN is a computing system design that resembles the way human brain works where it takes the preceding examples into consideration to create a system of neuron hence makes a new decision. In modelling complex data pattern, ANN model is more practical and precise as the technique involved non-linear applications. ANN has been broadly used in various areas especially in prediction and classification. The extensive reviews of water demand forecasting using ANN is as in Table 1.

Table 1: Papers Review on Water Demand Analysis using ANN model.

References	Prediction Variable	Prediction Type	Region	Accuracy Rate
Carvalho et al. (2021)	Water demand	N/A	Fortaleza, Brazil	N/A
Xenochristou et al. (2020)	Water demand	1-7 days into the future	Southwest of England	MAPE = 3.2% – 17%, for a reduction in group size from 600 to 5 households.
Sidhu et al. (2020)	Crop water demand	Daily	Karnal in Haryana, India	Average accuracy rate of 71%
Xenochristou and Kapelan (2020)	Water demand	Daily	Southwest of England	$R^2 = 74.1\%$
Pesantez et al. (2020)	Water demand	Hourly	Cary, North Carolina, USA	Median RMSE = 9.5 – 16 gph of water
Ibrahim et al. (2020)	Water demand	Daily	Kuwait	Support Vector Linear Regression MAPE = 0.52, RMSE = 2.59, ARIMA MAPE = 1.8, RMSE = 9.4
Chen et al. (2020)	Evapotranspiration for cabbage farmland	Daily	Hunan Province	$R^2 = 0.78 - 0.81$
Smolak et al. (2020)	Short-term water demand	24hours forecast Weekly forecast	Wroclaw, Poland	MAPE = 90.4%
Rahim et al. (2020)	N/A	N/A	N/A	N/A
Lee and Derrible (2020)	Water Demand	Daily	USA	$R_{adj}^2 = 0.60$
Mouatadid et al. (2019)	Water irrigation flow	Daily	Palos de la Frontera	Wavelet-bootstrap using ANN: $R^2 = 0.9432$ , Bootstrap using Machine Learning: $R^2 = 0.9190$
Vonk et al. (2019)	Climate change and vacation absence	Yearly, daily	Netherlands and Belgium	The average demand increased between $-0.2\%$ and $+3.1\%$ , The peaking factor increased between $-2.9\%$ and $21.3\%$ .
Duerr et al. (2018)	Portable water billing records water usage data from TBW	Short term, Long term	Tampa Bay Water, Florida (Hillsborough, Pasco, and Pinellas)	N/A
Candelieri (2017)	Water demand	Hourly	Milan	N/A
An et al. (2015)	Building water demand	Area Precipitation per Demand Ratio (APDR)	Hong Kong	N/A
Safavi et al. (2015)	Baseline scenario	Annually	Iran	N/A
Liu et al. (2015)	Urban water use	Yearly	China	NSE = 0.87

### 3.2 Decision Tree

Decision tree uses supervised machine learning algorithms where the data analysis technique divides the data into many possible entities related to a given parameter. A decision tree has two main components which are nodes and leaves. The supervised learning approach analyses an item's observation in the branches to determine the item's goal value in the leaves. The leaves reflect the results, whereas the decision nodes represent the data splitting. One of the benefits of decision trees is that their outputs are simple to read and analyze without the need for statistical expertise. However, the main problem of the decision tree is generally leading to overfitting of the data which ultimately leads to wrong predictions. The summary of water demand studies using decision trees are as in Table 2.

Table 2: Papers Review on Water Demand Analysis using Decision Tree model.

References	Prediction Variable	Prediction Type	Region	Accuracy Rate
Ahcene and Saadia (2019)	hourly water request	Hourly	Rassauta, Algeria	$R = 0.95$
Oyebode (2019)	weather and socioeconomic variations	N/A	City of Ekurhuleni, South Africa	Pearson correlation-based ANN model produced the highest $R^2$ and NSE values at 0.9233 and 0.9001
Shah et al. (2018a)	temperature, rainfall, snow, snow depth, number of customers, median income, holidays and day of the year	Daily and monthly	Indiana, USA	mMLR = $-12.73\%$ , mFFNN = $-6.45\%$ , mMRNN = $-3.84\%$
Yin et al. (2018)	total population, urban population, and the primary, secondary, and tertiary industry gross domestic product, annual precipitation	Yearly	Wuxi City, China	Mean relative error: ANN = $-2.14\%$ , MLR = $-3.96\%$
Shah et al. (2018b)	Weather conditions	Daily	Indiana	Average error 2.31% and 1.90% with threshold values of 10% and 5%
Shabani et al. (2018)	Water demand	short term	Milan, Italy	$R^2 = 0.95$ and $0.93$ for training and testing data set
Farias et al. (2018)	RBF-ANN Drinking water demand	Hourly, daily, weekly	Barcelona	RMS = 0.0559
Loureiro et al. (2016)	Water demand	Daily	Portugal	Relative error, = $7\% - 21\%$
Ji et al. (2016)	Water supply operation	Monthly	China	83%

### 3.3 Random Forest

Random forest is made up of huge number of individual decision trees that work together as an ensemble. Each tree in the random forest produces a class prediction and the class with the most votes becomes the prediction model. Its popularity is due to its ease of use and adaptability as it can handle both classification and regression problems. Besides, it reduces overfitting in decision trees and helps to improve the accuracy. However, Random Forest creates a lot of trees and combines their output. In order to do so, this algorithm requires much more computational power and resources. The summary of water demand studies using random forest are presented in Table 3.

Table 3: Papers Review on Water Demand Analysis using Random Forest Model.

References	Prediction Variable	Prediction Type	Region	Accuracy Rate
Villarin and Rodriguez-Galiano (2019)	Domestic water consumption per capita	N/A	Seville, Andalusia	RMSE = 22.06 L/day/inhabitant, $R^2 = 0.46$
Park and Lee (2019)	daily usage of water per person, water price, population, and the information of the water source	Seasonal	Korea	N/A
Tulbure et al. (2016b)	Surface water dynamics	Long term trends	Australia	99.9%(0.02% standard error) with 87%(3%) and 96%(2%)
Tulbure et al. (2016a)	Surface water (SW)	Yearly	Australia	99.94%

### 3.4 Hybrid and Optimization

Hybrid model is a method to overcome the weakness of an original approach by combining two or more techniques. This model works significantly more efficient and has better predictive performance. Optimization approach is a process of modifying the hyperparameters. The aims of optimization algorithm are to identify the best parameter values under various conditions. This method has been used to solve problems in numerous applications domains. Table 4 presented the reviews on hybrid and optimization methodologies for prediction of water demand.

Table 4: Papers Review on Water Demand Analysis using Hybrid and Optimization models.

References	Prediction Variable	Prediction Type	Region	Accuracy Rate
Pandey et al. (2021)	Water demand	Hourly and monthly	Spanish and India	Spanish RMSE = 11.67% – 62.38%, Indian RMSE = 52.19% – 70.12%
Zubaidi et al. (2020a)	Urban water demand	Monthly	Southeast Water Utility in Melbourne	$R^2 = 0.9$
Zubaidi et al. (2020b)	Municipal water demand	Monthly	Melbourne City	ANFIS CE = 0.974, CSA-ANN CE = 0.971
Zubaidi et al. (2020d)	Long-term Municipal Water Demands	Monthly	Australia, Greater Melbourne, Victoria,	$R^2 = 0.96$ , RMSE = 0.0025
Zubaidi et al. (2020c)	Urban water demand	Monthly	Gauteng Province, South Africa	BSA – ANN: RMSE = 0.0099 M $\ell$ , CE = 0.979
Kazemi et al. (2020)	Water resources allocation	Monthly	North of Iran	$R^2 = 0.78 – 0.9$ , RMSE = 2.35 – 4.26
Peng et al. (2019)	growth period of crops, irrigation values (light intensity, soil moisture, soil electrical conductivity, and air temperature)	Long term	China	$R = 0.98963$ , Mean Square Error (MSE) = 0.00857724
Zhang et al. (2016)	Water demand	Annually	China	RMSE = 42.8%, MAE = 7.6%

Hou et al. (2016)	Water demand	Annually	China	N/A
Safavi and Enteshari (2016)	Water extraction quantities	Monthly	Iran	$R^2 = 0.86 - 0.99$
Abdullah et al. (2015)	Evapotranspiration (ET)	Daily	Iraq	$R^2 = 0.982$
Tiwari and Adamowski (2015)	Water demand	Weekly and monthly	Canada	$R^2 = 0.70 - 0.85$
Rahmani and Zarghami (2015)	Water resources management	Monthly time scale	Iran	$R^2 = 0.82 - 0.98$
Perea et al. (2015)	Irrigation water demand	Daily	Spain	$R^2 = 0.93$
Tiwari and Adamowski (2017)	Water demand	Daily	Canada	$R^2 = 0.78 - 0.92$

#### 4. Discussion

This study has reviewed papers on water demand using machine learning from 2015 to 2021. It is found that, most of the studies focus on short-term demand forecasting and very few studies address medium and long-term forecasting. An extensive literature review revealed that the most used machine learning method to predict water demand was Artificial Neural Network. Finding the best method that can provide the best results can be a difficult task because it depends on the predictive variables included and data that the studies used.

The accuracy of the predictive models was evaluated using variety of methods including accuracy rate,  $R^2$  and error rate such as RMSE and MAPE. The comparative assessment of the prediction models showed that the hybrid and optimized methods can be considered as the best assessment to predict future water demands.

#### 5. Conclusion

In the previous six years, the adoption of machine learning models has increased. The reviewed found that the most machine learning method used is Artificial Neural network, Support Vector Machine, Random Forest and Decision Tree. According to the study, the method's accuracy varies based on the type of prediction variables applied. In addition, the use of hybrids and optimization methods has risen considerably in recent years. The researchers have proven that hybrid and optimization models can improve accuracy and efficiency of the models. The research's next step is to create hybrid and optimization models with improved accuracy.

#### Acknowledgment

The research was funded by Geran Dalaman Negeri (GDN): 600-TNCPI 5/3/DDN (02) (004/2020) and thank you to Universiti Teknologi MARA for their support.

#### References

- Abdullah, S., Malek, M., Abdullah, N., and Mustapha, A. (2015). Feedforward backpropagation, genetic algorithm approaches for predicting reference evapotranspiration. *Sains Malaysiana*, 44(7):1053–1059. cited By 5.
- Ahcene, B. and Saadia, B. (2019). Energetic optimization and evaluation of a drinking water pumping system: Application at the rassauta station. *Water Science and Technology: Water Supply*, 19(2):472–481. cited By 2.
- Ali, M., Saadon, A., Abd Rahman, N. F., and Khalid, K. (2014). *An Assessment of Water Demand in Malaysia Using Water Evaluation and Planning System*.

- An, K., Lam, Y., Hao, S., Morakinyo, T., and Furumai, H. (2015). Multi-purpose rainwater harvesting for water resource recovery and the cooling effect. *Water Research*, 86:116–121. cited By 24.
- Anang, Z., Padli, J., Rashid, N. K. A., Alipiah, R. M., and Musa, H. (2019). Factors affecting water demand: macro evidence in malaysia.
- Candelieri, A. (2017). Clustering and support vector regression for water demand forecasting and anomaly detection. *Water (Switzerland)*, 9(3). cited By 55.
- Carvalho, T., De Souza Filho, F., and Porto, V. (2021). Urban water demand modeling using machine learning techniques: Case study of fortaleza, brazil. *Journal of Water Resources Planning and Management*, 147(1). cited By 0.
- Chen, H., Huang, J., and McBean, E. (2020). Partitioning of daily evapotranspiration using a modified shuttleworth-wallace model, random forest and support vector regression, for a cabbage farmland. *Agricultural Water Management*, 228. cited By 13.
- Duerr, I., Merrill, H., Wang, C., Bai, R., Boyer, M., Dukes, M., and Bliznyuk, N. (2018). Forecasting urban household water demand with statistical and machine learning methods using large space-time data: A comparative study. *Environmental Modelling and Software*, 102:29–38. cited By 16.
- Farias, R., Puig, V., Rangel, H., and Flores, J. (2018). Multi-model prediction for demand forecast in water distribution networks. *Energies*, 11(3). cited By 11.
- Hou, J., Fan, X., and Liu, R. (2016). Optimal spatial allocation of irrigation water under uncertainty using the bilayer nested optimisation algorithm and geospatial technology. *International Journal of Geographical Information Science*, 30(12):2462–2485. cited By 4.
- House-Peters, L. A. and Chang, H. (2011). Urban water demand modeling: Review of concepts, methods, and organizing principles. *Water Resources Research*, 47(5).
- Ibrahim, T., Omar, Y., and Maghraby, F. (2020). Water demand forecasting using machine learning and time series algorithms. pages 325–329. cited By 0.
- Ji, Y., Lei, X., Cai, S., and Wang, X. (2016). Application of a classifier based on data mining techniques in water supply operation. *Water (Switzerland)*, 8(12). cited By 2.
- Kazemi, M., Bozorg-Haddad, O., Fallah-Mehdipour, E., and Loáiciga, H. (2020). Inter-basin hydropolitics for optimal water resources allocation. *Environmental Monitoring and Assessment*, 192(7). cited By 1.
- Lee, D. and Derrible, S. (2020). Predicting residential water demand with machine-based statistical learning. *Journal of Water Resources Planning and Management*, 146(1). cited By 8.
- Liu, Y., Zhao, J., and Wang, Z. (2015). Identifying determinants of urban water use using data mining approach. *Urban Water Journal*, 12(8):618–630. cited By 6.
- Loureiro, D., Mamade, A., Cabral, M., Amado, C., and Covas, D. (2016). A comprehensive approach for spatial and temporal water demand profiling to improve management in network areas. *Water Resources Management*, 30(10):3443–3457. cited By 9.
- Mohd Firdaus, N. N. and Talib, S. (2016). Modeling water supply and demand for effective water management allocation in selangor. *Jurnal Teknologi*, 78.
- Mouatadid, S., Adamowski, J., Tiwari, M., and Quilty, J. (2019). Coupling the maximum overlap discrete wavelet transform and long short-term memory networks for irrigation flow forecasting. *Agricultural Water Management*, 219:72–85. cited By 15.
- Oyebode, O. (2019). Evolutionary modelling of municipal water demand with multiple feature selection techniques. *Journal of Water Supply: Research and Technology - AQUA*, 68(4):264–281. cited By 3.
- Pandey, P., Bokde, N., Dongre, S., and Gupta, R. (2021). Hybrid models for water demand forecasting. *Journal of Water Resources Planning and Management*, 147(2). cited By 1.
- Park, H. and Lee, D. (2019). Is water pricing policy adequate to reduce water demand for drought mitigation in korea? *Water (Switzerland)*, 11(6). cited By 1.

- Peng, Y., Xiao, Y., Fu, Z., Dong, Y., Zheng, Y., Yan, H., and Li, X. (2019). Precision irrigation perspectives on the sustainable water-saving of field crop production in china: Water demand prediction and irrigation scheme optimization. *Journal of Cleaner Production*, 230:365–377. cited By 14.
- Perea, R., Poyato, E., Montesinos, P., and Díaz, J. (2015). Irrigation demand forecasting using artificial neuro-genetic networks. *Water Resources Management*, 29(15):5551–5567. cited By 16.
- Pesantez, J., Berglund, E., and Kaza, N. (2020). Smart meters data for modeling and forecasting water demand at the user-level. *Environmental Modelling and Software*, 125. cited By 10.
- Rahim, M., Nguyen, K., Stewart, R., Giurco, D., and Blumenstein, M. (2020). Machine learning and data analytic techniques in digitalwater metering: A review. *Water (Switzerland)*, 12(1). cited By 9.
- Rahmani, M. and Zarghami, M. (2015). The use of statistical weather generator, hybrid data driven and system dynamics models for water resources management under climate change. *Journal of Environmental Informatics*, 25(1):23–35. cited By 32.
- Safavi, H. and Enteshari, S. (2016). Conjunctive use of surface and ground water resources using the ant system optimization. *Agricultural Water Management*, 173:23–34. cited By 16.
- Safavi, H., Golmohammadi, M., and Sandoval-Solis, S. (2015). Expert knowledge based modeling for integrated water resources planning and management in the zayandehrud river basin. *Journal of Hydrology*, 528:773–789. cited By 51.
- Shabani, S., Candelieri, A., Archetti, F., and Naser, G. (2018). Gene expression programming coupled with unsupervised learning: A two-stage learning process in multi-scale, short-termwater demand forecasts. *Water (Switzerland)*, 10(2). cited By 14.
- Shah, S., Hosseini, M., Miled, Z., Shafer, R., and Berube, S. (2018a). A water demand prediction model for central indiana. pages 7819–7824. cited By 0.
- Shah, S., Miled, Z., Schaefer, R., and Berube, S. (2018b). Differential learning for outliers: A case study of water demand prediction. *Applied Sciences (Switzerland)*, 8(11). cited By 2.
- Sidhu, R., Kumar, R., and Rana, P. (2020). Machine learning based crop water demand forecasting using minimum climatological data. *Multimedia Tools and Applications*, 79(19-20):13109–13124. cited By 1.
- Smolak, K., Kasieczka, B., Fialkiewicz, W., Rohm, W., Siła-Nowicka, K., and Kopańczyk, K. (2020). Applying human mobility and water consumption data for short-term water demand forecasting using classical and machine learning models. *Urban Water Journal*, 17(1):32–42. cited By 5.
- Tiwari, M. and Adamowski, J. (2015). Medium-term urban water demand forecasting with limited data using an ensemble wavelet-bootstrap machine-learning approach. *Journal of Water Resources Planning and Management*, 141(2). cited By 39.
- Tiwari, M. and Adamowski, J. (2017). An ensemble wavelet bootstrap machine learning approach to water demand forecasting: a case study in the city of calgary, canada. *Urban Water Journal*, 14(2):185–201. cited By 15.
- Tulbure, M., Broich, M., and Stehman, S. (2016a). Spatiotemporal dynamics of surface water extent from three decades of seasonally continuous landsat time series at subcontinental scale. volume 41, pages 403–404. cited By 3.
- Tulbure, M., Broich, M., Stehman, S., and Kommareddy, A. (2016b). Surface water extent dynamics from three decades of seasonally continuous landsat time series at subcontinental scale in a semi-arid region. *Remote Sensing of Environment*, 178:142–157. cited By 118.
- Villarin, M. and Rodriguez-Galiano, V. (2019). Machine learning for modeling water demand. *Journal of Water Resources Planning and Management*, 145(5). cited By 12.
- Vonk, E., Cirkel, D., and Blokker, M. (2019). Estimating peak daily water demand under different climate change and vacation scenarios. *Water (Switzerland)*, 11(9). cited By 1.
- Xenochristou, M., Hutton, C., Hofman, J., and Kapelan, Z. (2020). Water demand forecasting accuracy and influencing factors at different spatial scales using a gradient boosting machine. *Water Resources Research*, 56(8). cited By 4.
- Xenochristou, M. and Kapelan, Z. (2020). An ensemble stacked model with bias correction for improved water demand forecasting. *Urban Water Journal*, 17(3):212–223. cited By 5.

- Yin, Z., Jia, B., Wu, S., Dai, J., and Tang, D. (2018). Comprehensive forecast of urban water-energy demand based on a neural network model. *Water (Switzerland)*, 10(4). cited By 14.
- Zhang, H., Singh, V., Wang, B., and Yu, Y. (2016). Ceref: A hybrid data-driven model for forecasting annual streamflow from a socio-hydrological system. *Journal of Hydrology*, 540:246–256. cited By 29.
- Zubaidi, S., Abdulkareem, I., Hashim, K., Al-Bugharbee, H., Ridha, H., Gharghan, S., Al-Qaim, F., Muradov, M., Kot, P., and Al-Khaddar, R. (2020a). Hybridised artificial neural network model with slime mould algorithm: A novel methodology for prediction of urban stochastic water demand. *Water (Switzerland)*, 12(10). cited By 9.
- Zubaidi, S., Al-Bugharbee, H., Ortega-Martorell, S., Gharghan, S., Olier, I., Hashim, K., Al-bdairi, N., and Kot, P. (2020b). A novel methodology for prediction urban water demand by wavelet denoising and adaptive neuro-fuzzy inference system approach. *Water (Switzerland)*, 12(6). cited By 26.
- Zubaidi, S., Ortega-Martorell, S., Al-Bugharbee, H., Olier, I., Hashim, K., Gharghan, S., Kot, P., and Al-Khaddar, R. (2020c). Urban water demand prediction for a city that suffers from climate change and population growth: Gauteng province case study. *Water (Switzerland)*, 12(7). cited By 30.
- Zubaidi, S., Ortega-Martorell, S., Kot, P., Alkhaddar, R., Abdellatif, M., Gharghan, S., Ahmed, M., and Hashim, K. (2020d). A method for predicting long-term municipal water demands under climate change. *Water Resources Management*, 34(3):1265–1279. cited By 33.





**20  
21** **ICMS**  
INTERNATIONAL CONFERENCE ON COMPUTING,  
MATHEMATICS AND STATISTICS

e ISBN 978-967-2948-12-4



9 7 8 9 6 7 2 9 4 8 1 2 4