Journal of Quality Measurement and Analysis *JQMA* 21(4) 2025, 77-92
e-ISSN: 2600-8602 *http://www.ukm.my/jqma*
https://doi.org/10.17576/jqma.2104.2025.05

A SYSTEM OF STOCHASTIC DIFFERENTIAL EQUATIONS FOR HANDLING UNCERTAINTIES IN WATER DEMAND

(*Sistem Persamaan Pembezaan Stokastik untuk Mengendalikan Ketidakpastian dalam Permintaan Air*)

ANIZA AKARAM, TAN CHEOW QING & ARIFAH BAHAR*

ABSTRACT

Managing uncertainties in water demand is a challenging task in modern water resource management, especially when factors like unpredictable rainfall patterns and sudden shifts in population growth influence consumption trends. Stochastic Differential Equations (SDEs) provide a powerful mathematical framework for modeling and analyzing the inherent uncertainties in water demand. In this study, SDEs are applied to examine water demand in Johor over the period 2005 to 2020, enabling the modeling and forecasting of demand within the context of these uncertainties. The water balance model, which tracks the inflow, outflow, and storage of water, is introduced to provide a more dynamic and realistic representation of water resource fluctuations. The study employs the Euler-Maruyama method for numerical solutions, offering a flexible and accurate approach for simulating the dynamic behavior of water demand. The findings highlight the importance of incorporating random variations and uncertainties into water demand forecasting, offering valuable insights for decision-makers in the water industry. This research combines stochastic models with water balance models to improve how we predict and manage water resources, helping to ensure a steady water supply and reduce long-term risks from water shortages.

Keywords: water demand uncertainty; Stochastic Differential Equations; water balance model; Euler-Maruyama; water demand forecasting

ABSTRAK

Menguruskan ketidaktentuan dalam permintaan air merupakan satu cabaran yang besar dalam pengurusan sumber air moden, terutamanya apabila faktor-faktor seperti pola hujan yang tidak menentu dan perubahan mendadak dalam pertumbuhan penduduk mempengaruhi trend penggunaan air. Persamaan Pembezaan Stokastik (SDE) menyediakan satu kerangka matematik yang berkuasa untuk memodelkan dan menganalisis ketidakpastian semula jadi dalam permintaan air. Dalam kajian ini, SDE digunakan untuk menganalisis permintaan air di Johor bagi tempoh 2005 hingga 2020, membolehkan pemodelan dan ramalan permintaan dalam kontejs ketidakpastian tersebut. Model imbangan air, yang menjejaki aliran masuk, aliran keluar, dan simpanan air, diperkenalkan untuk memberikan gambaran yang lebih dinamik dan realistik terhadap turun naik sumber air. Kajian ini menggunakan kaedah Euler-Maruyama untuk penyelesaian berangka, yang menawarkan pendekatan yang fleksibel dan tepat dalam mensimulasikan dinamik permintaan air. Hasil kajian menekankan kepentingan memasukkan variasi rawak dan ketidaktentuan ke dalam ramalan permintaan air, memberikan pandangan yang bernilai kepada pembuat keputusan dalam industri air. Penyelidikan ini menggabungkan model stokastik dengan model imbangan air untuk meningkatkan keupayaan meramal dan mengurus sumber air, bagi memastikan bekalan air yang stabil serta mengurangkan risiko jangka panjang akibat kekurangan air.

Kata kunci: ketidaktentuan permintaan air; Persamaan Pembezaan Stokastik; model imbangan air; Euler-Maruyama; ramalan permintaan air

1. Introduction

Water demand in Malaysia is a growing issue due to industrial expansion and population growth, both of which significantly impact water consumption patterns. As Malaysia continues to advance economically, the demand for water has risen steadily, creating increasing pressure on the country's existing water resources and infrastructure. This situation is further complicated by the diverse range of factors that influence water demand, including economic factors, particularly Gross Domestic Product (GDP), and variations in water consumption patterns. Economic factors, such as changes in GDP, can significantly impact water demand as economic growth drives higher industrial activity and increased water usage, while economic downturns may reduce consumption. These uncertainties make it challenging to accurately predict future water needs and plan for sustainable resource management.

Water demand forecasting is the process of predicting future water consumption patterns based on historical data, climate conditions and other relevant factors. One key purpose of water demand forecasting is to provide public water suppliers with future-oriented information (Billings & Jones 2011). As a result, extensive research has been conducted on water demand forecasting to prevent water resource shortages in various countries. With increasing population growth, consumption and the present threat of climate change, accurate water demand forecasting has emerged as a pivotal tool in water resource management and planning (Anang *et al.* 2019). Traditional water demand forecasting models use historical data and known factors to predict future needs. However, they often struggle with sudden weather changes, economic shifts, and rapid urban growth, which can make their predictions less reliable. These models may not accurately forecast water demand during extreme events or economic fluctuations. This shows the need for more advanced models that use stochastic techniques to better manage the unpredictability and variability in water demand.

This study proposes the use of Stochastic Differential Equations (SDEs) integrated with a water balance model. SDEs are particularly suited for modeling systems where uncertainty and random fluctuations play a significant role, as they allow for the inclusion of random variables that can capture the unpredictable nature of factors affecting water demand. By incorporating stochastic elements into the forecasting process, the proposed model aims to more accurately reflect the real-world variability in water demand, leading to more reliable predictions. This approach not only enhances the precision of water demand forecasts but also improves the model's ability to adapt to changing conditions, ultimately supporting more effective water resource management strategies.

This research aims to enhance water demand forecasting by utilizing a specialized approach in SDEs, specifically through the application of Geometric Brownian Motion (GBM). By integrating GBM, this study addresses uncertainties in water demand forecasting linked to fluctuations in economic indicators, particularly GDP, and varying patterns of water consumption. Through this approach, the research seeks to capture the dynamic nature of these variables, accounting for their randomness and variability, which are pivotal in accurately predicting water demand in regions such as Johor. This method provides a robust framework for analyzing how shifts in GDP and consumption trends influence future water requirements, helping to improve predictive accuracy in the face of economic and behavioral uncertainties.

2. Literature Review

Water authorities face challenges due to the unpredictable nature of water demand, which is influenced by various dynamic factors (Browne *et al.* 2013). Late detection of these uncertainties can result in inaccurate forecasts, potentially leading to water shortages. GDP, as one of the uncertainties, plays a crucial role in understanding and forecasting water demand, as it reflects the economic activities that drive the consumption of water resources. Research has consistently shown a strong correlation between economic growth and water usage, underscoring the importance of incorporating GDP as a variable in predictive models. For instance, the Organisation for Economic Co-operation and Development (OECD 2014) discussed how economic growth influences water demand, stressing the need to consider GDP when developing sustainable water management policies. Similarly, a study by the Russ *et al.* (2022) examines the interconnection between economic performance and water resources, demonstrating how fluctuations in GDP can affect both water quality and demand. These studies highlight the importance of integrating GDP into water demand forecasting to improve resource management and inform policy decisions.

Addressing uncertainty is crucial for accurate long-term water demand forecasting (Rinaudo 2015). These forecasts rely on historical data and future trend assumptions, but their accuracy can be compromised by missing or incorrect data and uncertain information quality. For an accurate and dependable forecast, it is essential to comprehend both the forecast itself and the associated uncertainties (Tiwari & Adamowski 2013). Water demand forecasting encompasses both short-term and long-term predictions. Short-term data, which covers periods from minutes to months, captures patterns, fluctuations, and variability over a brief time frame. This data is valuable for understanding immediate trends, making real-time decisions, and addressing current conditions. In contrast, long-term data, extending from months to years or even decades, helps identify broader patterns, trends, and cycles over an extended period. Various models have been used to forecast long-term water demand and these models can be categorized into probabilistic models and deterministic models (Almutaz *et al.* 2013).

Traditional time series methods for forecasting water demand, such as regression analysis, exponential smoothing, and Autoregressive Integrated Moving Average (ARIMA), are appreciated for their simplicity and ease of use. However, to make accurate predictions, it is crucial to consider factors like seasonality, external influences, and evolving patterns, which might require more advanced or combined methods. The ARIMA-M model has been introduced by combining ARIMA with Markov Chain error correction to forecast daily water consumption (Du *et al.* 2020). While ARIMA captures linear trends, the Markov Chain adjusts for random errors. However, ARIMA may struggle with non-stationary data and sudden changes, and the Markov Chain, though helpful, can be complex and may not fully address all errors, particularly in highly variable water demand. The model's success relies on high-quality data and careful calibration of both components to handle daily water usage complexities.

The studies by Herdiansyah *et al.* (2022) and Anang *et al.* (2019) both highlight challenges in accurately modeling water demand due to various methodological limitations. Herdiansyah *et al.*'s (2022) approach uses a simplified calculation, which, while easy to implement, fails to account for critical dynamic variables like seasonal changes, income variations, and shifts in consumer behavior. This simplification creates a gap, as it limits the model's ability to capture the complex, fluctuating nature of water demand. Similarly, Anang *et al.* (2019) identifies a strong correlation between income levels and water consumption, suggesting income changes significantly influence water usage. However, the use of Multiple Linear Regression in Anang *et al.*'s (2019) study presents an issue, as this method assumes minimal multicollinearity among variables. Since an assumption is violated, thus reducing the model's accuracy. In contrast, a stochastic approach, such as Stochastic Differential Equations (SDEs), could address these limitations. Unlike MLR, SDEs are designed to accommodate both the randomness of external factors and the presence of multicollinearity, making them better suited for handling the inherent variability and complex relationships within water demand data.

Stochastic time series methods are widely used for forecasting water demand due to the unpredictable nature of water usage patterns. These methods capture randomness and variability in consumption, making accurate predictions possible even amidst uncertainties. SDEs offer a framework that combines deterministic and stochastic elements, allowing researchers to model dynamic systems impacted by random fluctuations. Silva Santos *et al.* (2024) utilized a stochastic model based on Brazilian data to achieve a 40% reduction in residential water demand, demonstrating the model's relevance within Brazilian regions. However, adapting the model for different areas, which have unique factors like climate, demographics, and infrastructure, is essential for broader applicability. Similarly, Cominola *et al.* (2016) introduced a high-resolution stochastic model for residential water use, which, while effective for Monte Carlo simulations and scenario testing, does not include outdoor water data which is a key limitation in arid regions. Zubaidi *et al.* (2020) applied a stochastic model to forecast urban water demand in Baghdad, capturing variability through seasonality and short-term patterns, supporting accurate predictions for urban water management. Beyond water demand, SDEs are also valuable in other fields, such as modeling uncertainties in electrical power systems (Verdejo *et al.* 2019) and balancing real-time demand in wind power systems (Olsson *et al.* 2010). This versatility highlights the potential of SDE-based stochastic models for managing dynamic, uncertain systems across various domains.

Among Malaysian states, Johor stands out as a critical focus for water demand studies due to its rapid economic development, population growth, and its vital role in both domestic and transboundary water supply systems. Johor's economic landscape is undergoing rapid transformation, driven by substantial investments in data centers from global technology companies such as Amazon, Nvidia, Google, Microsoft, and ByteDance (Financial Times 2025). While these developments contribute significantly to economic growth, they also result in heightened water and energy demand, placing additional pressure on existing infrastructure. In contrast, Kelantan faces challenges in meeting its water demand due to limited infrastructure and over-reliance on groundwater sources, which often leads to supply disruptions (The Star 2024a). Similarly, Perlis, being primarily an agricultural state, struggles with water scarcity during dry spells, severely impacting paddy farmers and resulting in significant income losses (The Star 2024b). These inter-state differences underscore the necessity for localized, state-specific studies that account for economic activities, governance structures, and spatial variability in resource availability to support sustainable and efficient water resource management.

In conclusion, while SDEs have been extensively utilized in diverse fields for modeling complex systems, there is currently a gap in the application of SDEs specifically for forecasting water demand in Malaysia. This highlights an opportunity for further research to explore how SDEs could be applied to improve water demand forecasts in the Malaysian context, addressing the unique uncertainties and dynamics present in this region. Furthermore, while numerous studies on water demand have been conducted, regional differences in climate, infrastructure, and socio-economic factors highlight the need for region-specific research. The unique characteristics of the Johor region, including its climate, population growth, and urbanization trends, underscore the importance of examining water demand within this local context. These factors may influence water consumption patterns in ways that differ significantly from other areas, making targeted research in Johor essential to develop accurate forecasting models and effective water management strategies tailored to the region's specific needs.

3. Methodology

3.1. *Data source*

To conduct a comprehensive analysis of water demand, we relied on secondary data sourced from reputable institutions. The primary source of our data was the Department of Statistics Malaysia (2014), an official government agency responsible for collecting, analyzing, and disseminating statistical data across various sectors in Malaysia. For this study, we directly accessed DoSM records to obtain accurate and relevant data. Key variables used in our analysis included Gross Domestic Product (GDP), historical water consumption, and water demand. The dataset covers the period from 2005 to 2020, offering a robust timeline for detailed examination.

In addition, climate data necessary for the water balance model was sourced from Weather and Climate (n.d.). This data included critical variables such as temperature, precipitation, and humidity levels, which had been estimated based on 30 years of data. The data was obtained by navigating to the historical weather records section on the website and manually extracting the relevant information for the study.

3.2. *Stochastic differential equation*

In a study on factors influencing water demand, a correlation was found between income levels and water consumption, suggesting that changes in income might affect water usage. However, for Multiple Linear Regression to be valid, independent variables must not exhibit significant correlations with each other, a condition known as the absence of multicollinearity. When multicollinearity is present, the accuracy of hypothesis tests for individual regression coefficients can be compromised. This limitation underscores the need for alternative methods, such as Stochastic Differential Equations (SDEs), which can better handle complex relationships and uncertainties in water demand forecasting. SDEs are a class of differential equations used to model systems that are influenced by random factors. They extend traditional differential equations by incorporating stochastic processes and thus SDEs offer a realistic representation of complex systems, accounting for the inherent uncertainty and variability of real-world phenomena. This approach enhances both the analysis and simulation of systems with unpredictable behavior, leading to a better understanding and improved predictive capabilities of their dynamics. The general form of an SDE is:

$$dX(t) = \mu(t)\, X(t)\, dt + \sigma(t)\, X(t)\, dW(t). \tag{1}$$

where $X(t)$ is the state variable representing the system's state at time *t*, $\mu(t)X(t)$ is the drift term (deterministic part), $\sigma(t)X(t)$ is the diffusion term (stochastic part), and $W(t)$ is a Wiener process or Brownian motion, which introduces randomness into the system.

In the context of water demand, stochastic differential equations can be used to model uncertainties in water usage and forecast future water demand. The SDE for water demand can be represented as:

$$dD(t) = \mu(t)\, \mathrm{D}(t)\, dt + \sigma(t)\, \mathrm{D}(t)\, dW(t). \tag{2}$$

where $D(t)$ is water demand at time t, $\mu(t)D(t)$ is the drift term representing the average rate of change in water demand, influenced by factors such as seasonal variations and long-term trends, $\sigma(t)D(t)$ is the diffusion term representing the variability and uncertainty in water demand due to random factors such as sudden changes in climate or population, and $W(t)$ is a Wiener process that captures the stochastic nature of water demand.

3.2.1. *Geometric Brownian Motion*

Geometric Brownian Motion (GBM) is one of the most widely used models in stochastic processes, due to its ability to model random variables that evolve over time with both drift (trend) and volatility (random fluctuations). GBM is described by a SDE and is often applied to model where the future path of the variable is uncertain. The SDE for a GBM is written as in Eq. (2). The solution to this SDE provides the evolution of water demand over time:

$$D_t = D_0 e^{((\mu-(1/2)\sigma^2)t+\sigma W_t)} \quad (3)$$

This Eq. (3) shows that water demand follows a log-normal distribution, meaning it grows exponentially over time while subject to random fluctuations. $\left(\mu - \frac{1}{2}\sigma^2\right)$ reflects the adjusted drift rate, while σW_t introduces stochastic fluctuations.

3.3. *Water Balance Model*

A water balance model is an essential component in hydrology for depicting the dynamics of water flow and storage within a particular system, whether it be a watershed, lake, or groundwater aquifer. This model operates on the fundamental principle of mass conservation, which asserts that the variation in water storage within the system during a defined time frame is equivalent to the net difference between the inflow and outflow of water. In simpler terms, the water balance equation can be expressed as follows:

$$\Delta S = P - T - E - R. \quad (4)$$

where ΔS is the change in water storage, P is precipitation, E is evaporation, T is transpiration, and R is runoff. The accuracy and reliability of a water balance model's predictions can be greatly influenced by uncertainties in its components, especially precipitation and runoff.

4. Results and Discussion

4.1. *Descriptive statistics*

The uncertainties in water demand for this study are influenced by factors such as real income, measured by Gross Domestic Product (GDP) in RM Million, and domestic water consumption, recorded in Million Litres per Day (MLD).

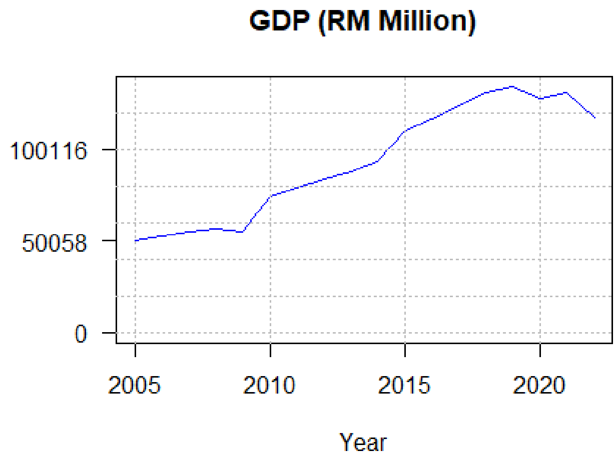

Figure 1: Time series plot for Gross Domestic Product (RM Million) from 2005 to 2022 in Johor

Figure 1 shows the trend of Gross Domestic Product (GDP) in Johor (in RM million) from the year 2005 to 2022. The plot demonstrates a general upward trajectory over the observed period, indicating economic growth. From 2005 to around 2014, the GDP shows a consistent but gradual increase, reflecting stable economic growth during this period. There is a noticeable jump in GDP starting around 2015, with a sharp rise peaking around 2018. This period might indicate stronger economic performance or an expansion in certain sectors contributing to GDP growth. After 2018, the GDP shows a period of stabilization or slower growth, followed by a slight decline around 2020, possibly influenced by external factors such as the global economic slowdown due to the COVID-19 pandemic. The GDP in 2021 and 2022 reflects a minor dip compared to the earlier peak but appears to recover slightly. The mean and variance for GDP are 93319.78 and 970608079 respectively. The mean suggests a strong average economic output. Meanwhile, the variance indicates significant fluctuations, reflecting a potentially volatile economic environment influenced by various external factors such as policy changes or global market shifts.

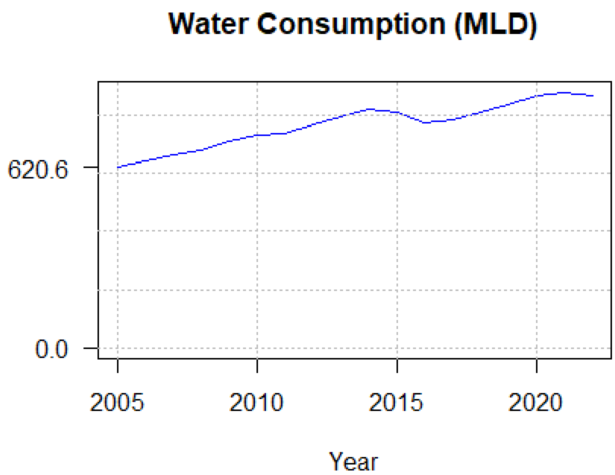

Figure 2: Time series plot for water consumption (MLD) for 2005 to 2022 in Johor

Figure 2 shows water consumption in Johor from 2005 to 2022. From 2005 to 2014, water consumption in Johor experienced a general upward trend, reflecting growing demand over this period. However, this trend shifted as water consumption decreased until 2016. Following this decline, water usage rose again until 2021, indicating a renewed increase in demand. In 2022, a slight decrease was observed, suggesting a potential stabilization or reduction in water consumption. The mean and variance for water consumption are 767.0333 and 6164.114 respectively. The mean provides a baseline for understanding typical usage patterns. Its variance indicates a more consistent consumption trend, suggesting that while there may be occasional spikes in usage, overall consumption remains relatively stable.

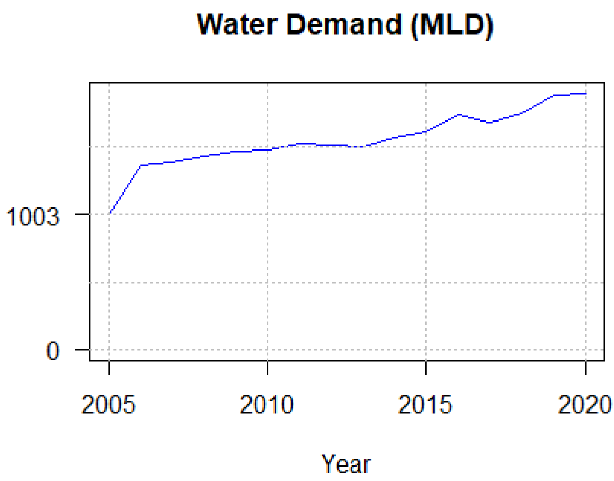

Figure 3: Time series plot for water demand (MLD) from 2005 to 2020 in Johor

The time series plot for water demand can be seen as in Figure 3. From 2005 to 2020, water demand in Johor exhibited a consistent upward trajectory. This indicates a sustained increase in water usage over the 15 years, reflecting either a rise in population, economic growth, or other factors driving higher consumption. The mean and variance for water demand are 1551.375 and 48525.85 respectively. The average water demand highlights the expected total demand for water resources and the variance implies variability in demand, possibly driven by seasonal changes or demographic shifts, indicating areas for further investigation or resource management.

4.2. *Normality test*

A normality test is a statistical procedure used to determine whether a dataset is well-modeled by a normal distribution, which is a key assumption in many statistical analyses. Conducting a normality test is essential because many statistical methods, such as t-tests, ANOVA, and regression analysis, rely on the assumption that the data is normally distributed. If the data does not meet this assumption, the results of these analyses may not be valid.

In this research, the Shapiro-Wilk test was used to assess the normality of real GDP and water consumption. The Shapiro-Wilk test is particularly powerful for detecting deviations from normality in small samples (typically $n < 50$) (Shapiro & Wilk 1965). It is more sensitive than other normality tests when sample sizes are limited. The Null Hypothesis (H_0)

is that the data follows a normal distribution, meanwhile the Alternative Hypothesis (H_A) refers to data is not normally distributed.

The results of the Shapiro-Wilk normality test indicate that the GDP data does not follow a normal distribution, as evidenced by a *p*-value of 0.03717, which is less than the 0.05 significance level. This suggests that the GDP data is significantly deviating from normality. In contrast, the water consumption data appears to follow a normal distribution, with a *p*-value of 0.5014. Since this *p*-value is greater than 0.05, we fail to reject the null hypothesis, indicating that the water consumption data is likely normally distributed.

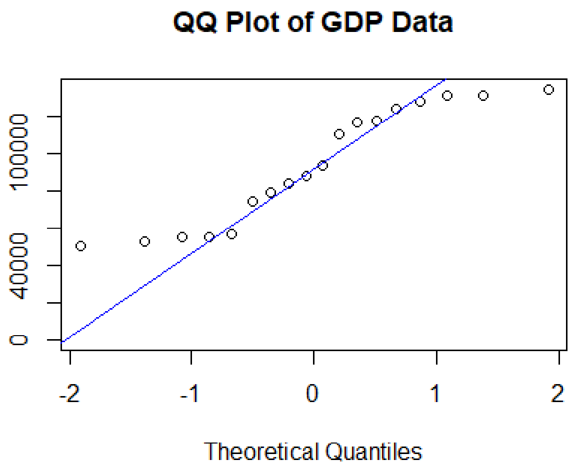

Figure 4: QQ Plot for Gross Domestic Product (RM Million)

A QQ (Quantile-Quantile) plot is a graphical tool used to assess whether a dataset follows a particular theoretical distribution, such as the normal distribution. Since our GDP has a smaller *p*-value, we use the QQ plot to show that the GDP data does not follow a normal distribution, particularly in the tails as in Figure 4. The data points generally follow a linear trend along the reference line, indicating that the GDP data approximately follows a normal distribution. At the extreme values (both high and low ends), there is some deviation from the line, suggesting slight departures from normality in the tails. This could indicate potential skewness or outliers in the GDP data distribution. Overall, the QQ plot suggests that the GDP data is reasonably close to a normal distribution, although minor deviations at the ends.

4.3. *Multicollinearity test*

Multicollinearity test is a method used to diagnose and address issues in regression analysis where predictor variables are highly correlated with one another. This correlation can complicate the estimation of regression coefficients, leading to unreliable or unstable results. Correlation coefficients close to +1 or -1 suggest strong correlations.

The results for the multicollinearity test for GDP and water consumption are 0.8904, and 0.8779 respectively. This shows there are highly strong correlation between water demand, GDP, and water consumption which results in the presence of multicollinearity. Multicollinearity can affect the overall performance of the regression model by reducing its predictive accuracy and making the interpretation of the coefficients less straightforward. Although the model might fit the data well, its ability to generalize to new data and provide reliable predictions might be compromised. This result implies that regression models are not suitable to model water demand.

4.4. *Geometric Brownian Motion*

Initially, we estimate the parameters for drift ($\hat{\mu}$) and volatility ($\hat{\sigma}$) using the Maximum Likelihood Estimation (MLE). MLE aims to find the parameter values that maximize the likelihood function, which measures the probability of observing the given data under the model. To analyze the time series data, we begin by calculating the log returns, which are the natural logarithms of the ratio of successive data points. We need to transform the data to log returns because log returns provide a way to stabilize the variance, as they normalize the effects of extreme values and make the data more homoscedastic. This transformation also enables additive modeling of returns over time, allowing for easier interpretation and forecasting. Additionally, log returns are more consistent with the assumption of continuous compounding, which aligns well with the mathematical framework of GBM. Next, we estimate the drift (μ) by computing the average of these log returns, representing the expected rate of change. Finally, determine the volatility (σ) by calculating the standard deviation of the log returns, which measures the variability or dispersion around the mean return.

Once the parameters are estimated, the GBM process is used to simulate the predicted values with standard normal random variables. To enhance the accuracy and reliability of these predictions, 100 random numbers are generated for each year, and the average of these numbers is used to forecast the values. Creating 100 random numbers of results in a solid sample size that is typically adequate for attaining statistical significance in various modeling contexts. This quantity effectively captures variability while keeping the analysis relatively straightforward. Conversely, generating an excessive number of random numbers can add unnecessary noise to the results, complicating the ability to identify significant patterns amid random variations. Utilizing 100 random numbers helps ensure a favorable signal-to-noise ratio. The Euler-Maruyama method has been utilized to estimate predicted values for GDP, water demand, and water consumption by applying GBM to model their stochastic behavior. The results for the calculated drift (μ) and the volatility (σ) for GDP are 0.05 and 0.2699 respectively. The calculation for predicted GDP follows the formula:

$$G_{t+1} = G_t e^{((\mu-(\sigma^2/2)\Delta t+\sigma\sqrt{\Delta t}Z_t))}. \tag{5}$$

where G is the predicted value of GDP, Δt is the time step, and Z_t is a standard normal random variable. The initial value is retained as it is in the original data, while the subsequent values are computed according to Eq. (5).

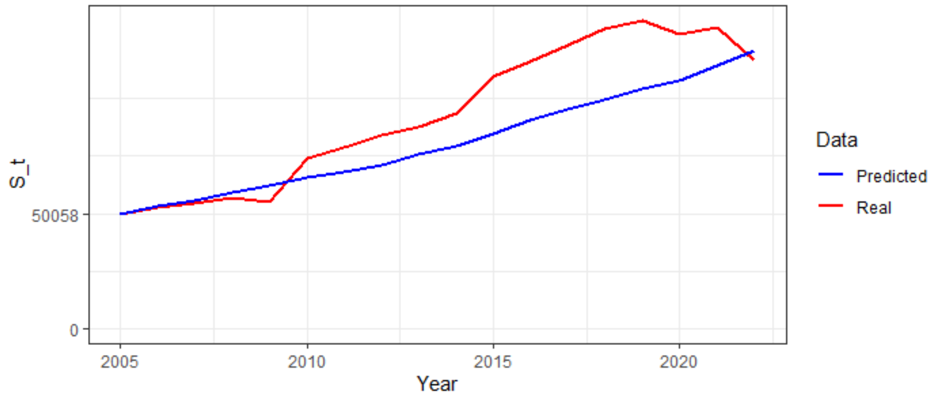

Figure 5: Comparison plot for real and predicted GDP

The same step is also used for both water consumption and water demand. As for water consumption, the calculated drift (μ) and the volatility (σ) are 0.1967 and 0.0773 respectively. The predicted water consumption is calculated using equation:

$$C_{t+1} = C_t e^{((\mu-(\sigma^2/2)\Delta t+\sigma\sqrt{\Delta t}Z_t))}. \tag{6}$$

where C is the predicted value of water consumption.

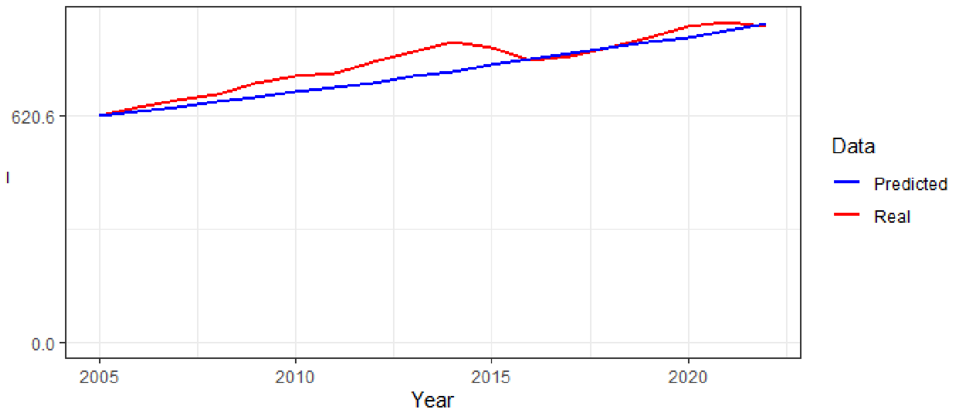

Figure 6: Comparison plot for real and predicted water consumption

The value for drift (μ) is 0.4259 and the volatility (σ) is 0.2477 for the parameter of water demand. The predicted water demand is based on equation:

$$D_{t+1} = D_t e^{((\mu-(\sigma^2/2)\Delta t+\sigma\sqrt{\Delta t}Z_t))}. \tag{7}$$

where D is the predicted water demand.

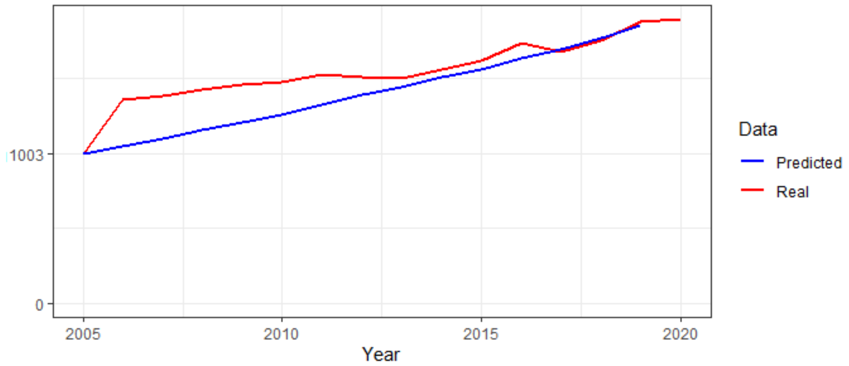

Figure 7: Comparison plot for real and predicted water demand

To check the accuracy of GBM model, we calculate using the Root Mean Square Error (RMSE) for GDP, water consumption and water demand. The RMSE for GDP is 17,254.27 RM million, indicating that, on average, the model's forecasts for GDP deviate from the actual figures by this amount. With an overall GDP of 1,679,756 RM million, this error accounts for approximately 1.03% of the total GDP. When put into context, this RMSE suggests that the predictive model is performing relatively well, as the error is just slightly above 1% of the total GDP. Such a level of accuracy is typically deemed acceptable in economic forecasting, particularly given the complexities and numerous variables that can influence GDP. A low RMSE in relation to overall GDP indicates that the model effectively captures general economic trends and fluctuations.

Similarly, the RMSE for water consumption is 35.9127 million liters per day reflects the average error in daily water consumption predictions. Considering the total daily consumption of 13,806.6 million liters, this RMSE constitutes about 0.26% of the overall consumption. This relatively small RMSE suggests that the model is fairly accurate, which is crucial for effective water resource management. Even minor inaccuracies can significantly impact resource allocation decisions. The low RMSE implies that the model likely considers various influencing factors, such as seasonal changes and population dynamics.

The RMSE for water demand is 165.4403 million liters per day, which indicates the average deviation in predicting daily water demand. When compared to the total demand of 24,822 million liters, this RMSE represents approximately 0.67% of total water demand. Although this RMSE reflects a reasonable level of error, it is higher than that for water consumption, suggesting that the model may encounter challenges in accurately predicting demand fluctuations. This discrepancy could be attributed to factors like changes in population growth, industrial activity, and climatic variations, which can all affect water demand.

4.5. *Water balance model*

In this study, the water balance model is calculated specifically for Johor Bahru due to the region-specific nature of atmospheric data, which can vary significantly across different locations. Johor Bahru is selected as a representative area for the state of Johor, as Malaysia's tropical climate generally exhibits minimal regional variability, with relatively small standard deviations in climatic parameters. This makes Johor Bahru an appropriate proxy for studying water balance trends within the broader context of Johor. Johor Bahru, being a highly urbanized region in Johor, experiences specific challenges related to water balance due to its growing population and industrial activities. The insights gained from this model are crucial for understanding the broader water resource dynamics in Johor, as Johor Bahru often serves as a representative area for the state. Johor's tropical climate, characterized by consistent rainfall patterns with minimal regional deviations, further reinforces the applicability of this model to the state's overall water management strategies. The water balance model equation used is:

$$\text{Water Balance} = \text{Precipitation} - \text{Evapotranspiration} - \text{Runoff} + \text{Inflow from Rivers} \quad (8)$$

where precipitation is the total amount of water entering the system in the form of rainfall, evapotranspiration is the amount of water that leaves the system through evaporation and plant transpiration, runoff is the amount of water that flows over the land surface and into rivers or storage reservoirs, and inflow from rivers is the additional water entering the system, expressed in million cubic (Chow et al. 1988; Thornthwaite and Mather 1955). Besides, Johor River Basin, which is central to the water resources of Johor Bahru is used as the catchment area which covers area approximately $2636km^2$.

A surface runoff coefficient of 0.5 is chosen for Johor Bahru due to its tropical climate, which is characterized by frequent and intense rainfall that generates significant runoff. This value reflects that half of the total precipitation contributes to the surface runoff, while the rest is either absorbed into the ground or lost through evaporation and transpiration. Urbanization in Johor Bahru further justifies this coefficient, as impervious surfaces like roads and buildings reduce water infiltration and amplify runoff. The high annual rainfall typical of the region supports the suitability of this value, ensuring the model accurately captures the hydrological dynamics of the area. Studies on tropical urban areas corroborate the use of a 0.5 runoff coefficient, highlighting its effectiveness in representing the combined influence of urban development and regional climatic conditions (Chow et al. 1988).

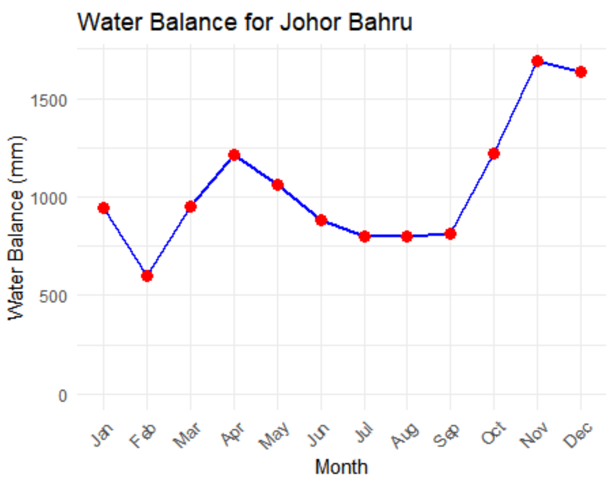

Figure 8: Water balance for Johor Bahru

The graph as in Figure 8 shows the water balance for Johor Bahru, derived from an analysis of data spanning 30 years. The months of November and December show a pronounced peak in water balance, exceeding 1500 mm, influenced by the effects of the Northeast Monsoon, which brings heavy rainfall to the area. The excess water during this period presents an opportunity to enhance water storage and replenish reservoirs, which are essential for addressing water demand during the drier months. However, in the absence of sufficient infrastructure to capture and store this surplus, a significant portion may escape as surface runoff, potentially causing flooding in urban and low-lying regions.

Conversely, February records the lowest water balance in Johor Bahru due to a combination of reduced rainfall and increased evapotranspiration. As it marks the tail-end of the Northeast Monsoon, the intensity and frequency of rainfall decline significantly compared to the peak months of November through January. This decrease in precipitation, coupled with rising temperatures and longer daylight hours, leads to higher evaporation and transpiration rates. Additionally, soil moisture begins to diminish following the heavy rains of previous months, resulting in less water retention. These factors collectively contribute to a lower net water gain, making February the driest month in terms of water balance.

The gradual increase in water balance from July to October reflects a transitional phase influenced by Malaysia's climatic patterns. According to the Malaysian Meteorological Department (MetMalaysia n.d.), the country experiences two main monsoon seasons: the Southwest Monsoon (May to September) and the Northeast Monsoon (November to March). Johor Bahru, located in southern Peninsular Malaysia, these transitions are particularly evident, with inter-monsoonal periods in April and October bringing variable winds and increased rainfall. These climatic shifts significantly impact the water balance in Johor Bahru during this time. On the other hand, the sharp declines after the rainy season highlight the rapid depletion of water reserves. From the water balance model, it can be seen that periods of high-water balance, especially in November and December, elevate the risk of flooding, whereas the steep drops during the drier months increase the potential for droughts, negatively impacting public health and economic activities. On top of that, agriculture, which depends on a stable water supply, experiences decreased crop yields during these dry spells.

5. Conclusion

Based on the results obtained, we can conclude that while the GBM model demonstrates high predictive accuracy based on the low RMSE value, it did not effectively capture or reflect certain underlying patterns or fluctuations in the data. This suggests that although the GBM model performs well overall in terms of prediction error, it is not sufficiently sensitive to variations in the data, which is important for understanding trends or changes in the system being studied. Therefore, further refinement or additional analysis are necessary to improve the model's ability to represent these fluctuations. The results for the water balance model highlight the importance of planning for water conservation and storage during periods of surplus, especially in November and December, to mitigate the impact of water scarcity during the dry months. By accounting for precipitation, evaporation, and runoff, the model helps identify potential surplus or deficits in water supply. This information is essential for planning water storage, distribution systems, and conservation measures, particularly in regions like Johor Bahru that experience both wet and dry seasons. Additionally, knowing the water balance supports sustainable water management, ensuring adequate supply during high-demand periods while minimizing the risk of shortages during dry spells.

Acknowledgments

The authors would like to thank Universiti Teknologi Malaysia, UTM Encouragement Research grant (Q.J130000.3809.31J69) and Ministry of Higher Education, Fundamental Research Grant Scheme (FRGS/1/2024/STG06/UTM/02/14) for their partial financial support.

References

Almutaz I., Ali E., Khalid Y. & Ajbar A.H. 2013. A long-term forecast of water demand for a desalinated dependent city: case of Riyadh City in Saudi Arabia. *Desalination and Water Treatment* **51**(31–33): 5934–5941.

Anang Z., Padli J., Rashid N.K.A., Alipiah R.M. & Musa H. (2019). Factors affecting water demand: Macro evidence in Malaysia. *Jurnal Ekonomi Malaysia* **53**(1): 17–25.

Billings R.B. & Jones C.V. 2011. *Forecasting Urban Water Demand*. 2nd Ed. Denver, CO: American Water Works Association.

Browne A.L., Medd W. & Anderson B. 2013. Developing novel approaches to tracking domestic water demand under uncertainty – A reflection on the "up scaling" of social science approaches in the United Kingdom. *Water Resources Management* **27**(4): 1013-1035.

Chow V.T., Maidment D.R. & Mays L.W. 1988. *Applied Hydrology*. New York, NY: McGraw-Hill Education.

Cominola A., Giuliani M., Castelleti A., Abdallah A.M. & Rosenberg D.E. 2016. Developing a stochastic simulation model for the generation of residential water end-use demand time series. *Proceedings of the 8th International Congress on Environmental Modelling and Software (iEMSs 2016)* in Toulouse, France.

Department of Statistics Malaysia. 2024. Data catalogue. https://open.dosm.gov.my/data-catalogue (1 October 2024).

Du H., Zhao Z. & Xue H. 2020. ARIMA-M: A new model for daily water consumption prediction based on the autoregressive integrated moving average model and the markov chain error correction. *Water* **12**(3): 760.

Financial Times. 2025. Tech groups to pay premium for energy for Malaysia data centres, says minister. https://www.ft.com/content/14d77d11-847e-4950-ab67-67a02e5324db (10 January 2025).

Herdiansyah A.R., Zahra R.A., Masjoyo Y.M., Muhammad A.F., Saputra M.R., Firdauzi L.B., Hafizha K.P. & Nurjani E. 2022. Analysis of meteorological water availability and water demand in Semarang Regency. *IOP Conference Series: Earth and Environmental Science* **1039**: 012011.

Malaysian Meteorological Department (MetMalaysia). n.d. Monsoon and climate of Malaysia. https://www.met.gov.my.

Organisation for Economic Co-operation and Development (OECD). 2014. Implications of water scarcity for economic growth. https://www.oecd-ilibrary.org/implications-of-water-scarcity-for-economic-growth_5jlssl611r32.pdf (1 October 2024).

Olsson M., Perninge M. & Söder L. 2010. Modeling real-time balancing power demands in wind power systems using stochastic differential equations. *Electric Power Systems Research* **80**(8): 966-974.

Rinaudo J.D. 2015. Long-term water demand forecasting. In Grafton Q., Daniell K., Nauges C., Rinaudo J.D. & Chan N. (eds.). *Understanding and Managing Urban Water in Transition*: 239–268. Dordrecht: Springer.

Russ J., Zaveri E., Desbureaux S., Damania R. & Rodella A.S. 2022. The impact of water quality on GDP growth: Evidence from around the world. *Water Security* **17**: 100130.

Shapiro S.S. & Wilk M.B. 1965. An analysis of variance test for normality (complete samples). *Biometrika* **52**(3-4): 591-611.

The Star. 2024a. Kelantan needs more water than operator can supply. https://www.thestar.com.my/news/nation/2024/03/08/kelantan-needs-more-water-than-operator-can-supply (1 July 2024).

The Star. 2024b. Padi farmers struggling amid water crisis in Perlis. https://www.thestar.com.my/news/nation/2024/07/19/padi-farmers-struggling-amid-water-crisis-in-perlis (1 August 2024).

Thornthwaite C.W. & Mather J.R. 1955. The water balance. *Publications in Climatology* **8**(1): 1–104. Centerton, NJ: Drexel Institute of Technology, Laboratory of Climatology.

Tiwari M.K. & Adamowski J. 2013. Urban water demand forecasting and uncertainty assessment using ensemble wavelet-bootsrap-neural network models. *Water Resources Research* **49**(10): 6486-6507.

Silva Santos V., Gadea A. & Cohim E. 2024. Stochastic model applied to water demand management in Brazil. *Water Supply* **24**(3): 692-706.

Verdejo H., Awerkin A., Kliemann W. & Becker C. 2019. Modelling uncertainties in electrical power systems with stochastic differential equations. *International Journal of Electrical Power and Energy Systems* **113**: 322-332.

Weather and Climate. n.d. *Johor Bahru climate by month*. https://www.weather-and-climate.com/.

Zubaidi S.L., Al-Bugharbee H., Muhsin Y.R., Hashim K. & Alkhaddar R. 2020. Forecasting of monthly stochastic signal of urban water demand: Baghdad as a case study. *IOP Conference Series: Materials Science and Engineering* **888**(1): 012018.

Department of Mathematical Sciences
Faculty of Science
Universiti Teknologi Malaysia Johor Bahru
81310 UTM Johor Bahru
Johor, MALAYSIA
E-mail: anizaakaram@graduate.utm.my, tanqing@graduate.utm.my, arifah@utm.my[*]

UTM-Centre or Industrial and Applied Mathematics,
Ibnu Sina Institute of Scientific and Industrial Research,
Universiti Teknologi Malaysia Johor Bahru
81310 UTM Johor Bahru
Johor, MALAYSIA
E-mail: arifah@utm.my[*]

Received: 3 January 2025
Accepted: 26 May 2025

[*]Corresponding author