https://doi.org/10.65646/3rc20dmd3s1179

*Proceedings of the International Research Conference of the Open University of Sri Lanka (IRC-OUSL 2025)*

# FORECASTING THE WATER QUALITY INDEX USING THE SEASONAL AUTOREGRESSIVE MOVING AVERAGE MODEL AND THE PROPHET MODEL

***H. Dissanayake* and R. Punchi-Manage***

*Department of Statistics and Computer Science, Faculty of Science, University of Peradeniya, Sri Lanka*

Surface water quality (WQ) often exhibits strong seasonal and long-term trends influenced by both natural processes and human activity. Accurate WQ forecasting is vital for ecosystem protection, public health, and environmental management, but irregular data collection challenges conventional models like SARIMA, which assume evenly spaced observations. The objectives of this study are to compute the Weighted Arithmetic Water Quality Index (WAWQI) from irregularly collected physicochemical WQ parameters between March 2009 and September 2017 from 20 monitoring sites along the River Thames, to develop SARIMA models for WAWQI forecasting using regularly spaced time series constructed from the observed data, to implement the Prophet forecasting model to handle irregularly sampled WAWQI data while capturing trend and seasonality, to assess and compare the predictive performance of both models using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE), and to provide model selection recommendations for environmental monitoring applications with irregular sampling intervals. SARIMA models cannot be fitted to irregular data. Therefore, to fit the SARIMA model, missing values were linearly interpolated to create a regularly spaced time series. The Prophet model can be used for irregular data without interpolation. Time series data were split into 80% training and 20% testing sets. RMSE captures large errors, MAE shows average error, and MAPE allows relative comparison across sites. Results revealed consistent seasonality in WQ across most locations. Both models performed relatively well. However, Prophet outperformed SARIMA in terms of MAPE (0.02–0.3%) compared to SARIMA's 2–11%, indicating superior relative accuracy. However, SARIMA performed slightly better in terms of RMSE and MAE, particularly at sites with denser data. Therefore, comparisons should be interpreted with caution, as Prophet's higher RMSE and MAE at sparsely sampled sites highlight its reliance on raw, irregular data without interpolation. These findings underscore the importance of aligning model selection with specific forecasting objectives. SARIMA performs well with stable, regularly spaced data, whereas Prophet offers greater flexibility for handling nonlinear trends and irregular sampling. This study provides practical guidance for environmental analysts in selecting forecasting models that best align with their data characteristics and policy goals.

*Keywords*: prophet model, SARIMA model, Water Quality Index

**Corresponding Author: hashinid@sci.pdn.ac.lk*

# FORECASTING THE WATER QUALITY INDEX USING THE SEASONAL AUTOREGRESSIVE MOVING AVERAGE MODEL AND THE PROPHET MODEL

***H. Dissanayake* and R. Punchi-Manage***
*Department of Statistics and Computer Science, Faculty of Science, University of Peradeniya, Sri Lanka.*

## INTRODUCTION

Surface water quality (WQ) is influenced by both natural processes and anthropogenic activities, often exhibiting strong seasonal and long-term variations. Accurate forecasting of these changes is essential for ecosystem preservation, public health protection, and effective environmental management. However, ecological monitoring data are frequently collected at irregular intervals, posing a significant challenge to conventional time series forecasting models. The SARIMA model requires equally spaced time series and is well-suited for stationary data with clear seasonal structures (Box et al., 2015). In contrast, the Prophet model is an additive regression model designed to accommodate missing data and irregular sampling frequencies while capturing seasonality and trend components (Taylor & Letham, 2018). The objectives of this study are to compute the surface water quality across 20 monitoring sites along the River Thames, to develop SARIMA models for WAWQI forecasting using regularly spaced time series constructed from the observed data, to implement the Prophet forecasting model to handle irregularly sampled WAWQI data while capturing trend and seasonality, to assess and compare the predictive performance of both models using RMSE, MAE, and MAPE as evaluation metrics, and to provide model selection recommendations for environmental monitoring applications with irregular sampling intervals.

## METHODOLOGY

### Study Area

This study examines the River Thames in Southern England (51°38’N–51°83’N, 01°75’E–00°56’E) using data from 20 sites collected between March 2009 and September 2017. The original dataset includes 21 sites (Figure 1), but one site was excluded due to having too few observations. Physical and chemical parameters were obtained from the Environmental Information Data Centre (EIDC), managed by the UK Centre for Ecology & Hydrology. These parameters were not measured at equal time intervals, resulting in an irregularly spaced dataset that presents challenges for traditional time series analysis.

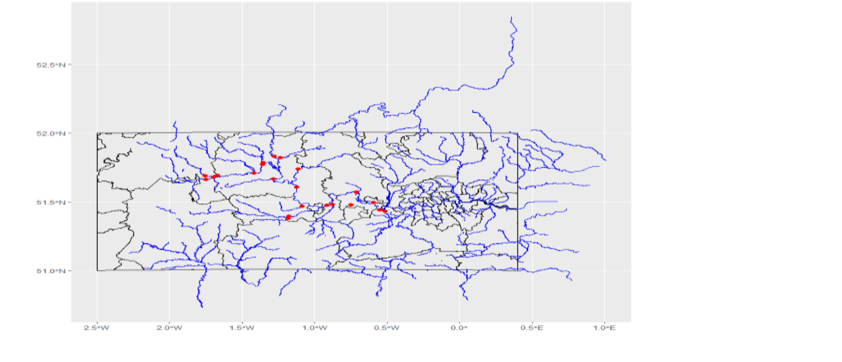


**Figure 1:** Map of the site's locations for the River *Thames*

## Weighted Arithmetic Water Quality Index

The Water Quality Index (WQI) simplifies complex chemical data into a single value to assess surface water quality. This study uses the WAWQI to classify water quality by purity level, following a multi-step calculation process. First, a unit weight ($W_n$) was assigned to each WQ parameter using the formula: $W_n = k/S_n$ where $S_n$ represents the standard permissible value for the $n^{\text{th}}$ WQ parameter and $k$ is the constant of proportionality, calculated as $k = 1/\Sigma[1/S_n]$. The standard values ($S_n$) were based on the WHO-recommended standards. Then, the quality rating scale ($Q_n$) for each parameter was calculated using the formula: $Q_n = [(V_n - V_{id})/(S_n - V_{id})] \times 100$ where $V_n$ is the estimated value of the n$^{\text{th}}$ WQ parameter and $V_{id}$ is the ideal value for $n^{\text{th}}$ parameter in pure water ($V_{id}$ for pH = 7 and zero for all other variables). Finally, the WAWQI for each observation was calculated using $WQI = \Sigma Q_n W_n / \Sigma W_n$.

## Seasonal Autoregressive Integrated Moving Average (SARIMA) Model

Time series diagnostics, including seasonal decomposition, Ljung–Box tests, and outlier detection, were conducted to identify trend, seasonality, and anomalies before fitting the SARIMA models. To model and forecast the Water Quality Index (WQI) over time for each monitoring site, Seasonal Autoregressive Integrated Moving Average (SARIMA) models were fitted. Seasonal models capture the seasonal variation of WQI. The data was arranged chronologically, and missing values were linearly interpolated. For each site, a time series object was created assuming monthly observations ($s$ = 12), and the data were split into training (80%) and testing (20%) subsets. A manual grid search was performed over a range of SARIMA hyperparameters *(p, d, q)* × *(P, D, Q)* to identify the optimal model for each site based on the lowest Root Mean Squared Error (RMSE) on the test set. The best-performing SARIMA model for each site was selected based on Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The SARIMA model was fitted using the *forecast* package in R (Hyndman et al., 2023).

## The Prophet Forecasting Model

The Prophet forecasting model is an additive regression model, which is used for forecasting time series data. It is designed to have intuitive parameters that can be adjusted without knowing the details of the underlying model. The Prophet model decomposes time series data into four main components: $y(t) = g(t) + s(t) + h(t) + \epsilon_t$, where $g(t)$ is the trend function, $s(t)$ is the seasonal effect, $h(t)$ is the effect of holidays, and $\varepsilon_t$ is the error term or idiosyncratic change. The holiday effect is not required and, therefore, is not considered in this study. In this study, the Prophet forecasting model was used to forecast the Water Quality Index (WQI) across the sampling sites. The data were transformed into a format compatible with Prophet, where the "sampling_date" variable was treated as the time index and the WQI as the target variable. For each site, a separate Prophet model was trained to understand site-specific WQI trends. The model was then used to generate a 90-day forecast for each site. To evaluate model performance, cross-validation was carried out by splitting the data into training and testing sets using an 80:20 ratio. Performance metrics, including RMSE, MAE, and MAPE were computed to assess prediction accuracy. Then, the forecasts were visualized with confidence intervals, allowing comparison across different sites. The Prophet model was implemented using the *prophet* package in R, developed by Facebook (Taylor & Letham, 2018). The performance of the SARIMA and Prophet models was assessed using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). RMSE emphasizes larger errors by squaring the residuals, MAE provides a linear measure of average error magnitude, and MAPE expresses errors as a percentage, facilitating interpretability across scales.

## RESULTS AND DISCUSSION

### Seasonal Autoregressive Integrated Moving Average (SARIMA) Model

The SARIMA models were fitted to the monthly WQI time series data for each monitoring site. before fitting the models, time series diagnostics revealed that most sites (e.g., TC1, TC10, TC4) exhibited clear seasonality, and TC20 showed an upward trend. Sites such as TC12, TC13, & TC17 remained relatively stable. The models captured both short-term trends and seasonal variations in WQI, providing insights into temporal fluctuations across different sites of the River *Thames*. Figure 2 indicates the SARIMA fitted models and forecasted WQI values for each monitoring site of the River *Thames.* The model demonstrated relatively stable performance across the majority of monitoring sites, especially those with denser and more regularly spaced data. RMSE values for SARIMA ranged from 1.41 (TC8) to 5.59 (TC2), while MAE ranged from 1.08 (TC8) to 4.66 (TC2). Sites such as TC10, TC12, and TC8 showed particularly low error values, indicating strong predictive accuracy where data density and interpolation quality were high. In terms of relative error, SARIMA's MAPE ranged from 1.61% to 11.05%. Notably, site TC15 recorded the highest MAPE (11.05%), along with relatively high RMSE and MAE, suggesting potential challenges in capturing more complex or irregular patterns at that location, even after interpolation. The performance of SARIMA was also influenced by the presence of outliers & changepoints in some sites. These irregularities led to higher RMSE & MAE values and reduced forecast stability, indicating SARIMA's limitations compared to Prophet's robustness with irregular data. However, SARIMA generally performed well at locations with smoother seasonal patterns and sufficient data continuity. For some sites, the Ljung–Box test indicated residual autocorrelation ($p < 0.05$), suggesting that SARIMA did not fully capture the underlying dependence. This limitation explains the need for comparison with Prophet.

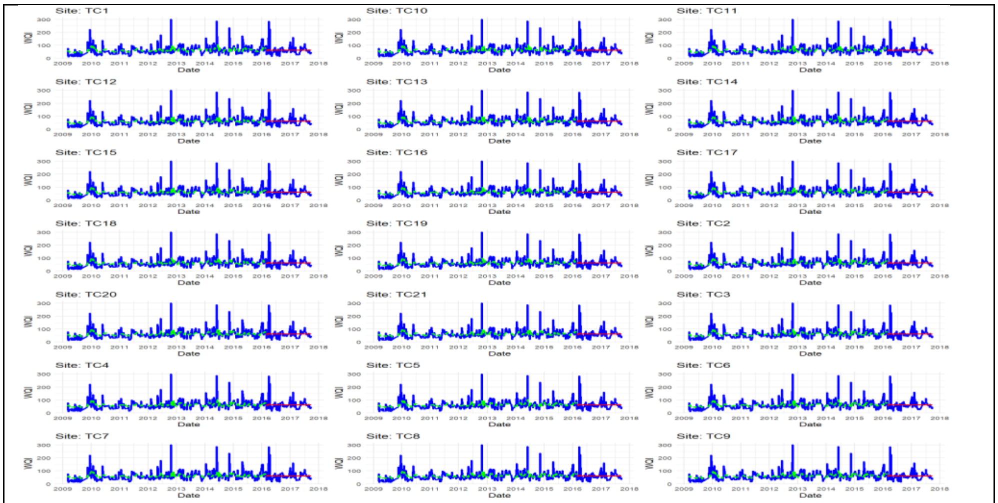


**Figure 2:** SARIMA fitted models and forecasted average monthly WQI values for each monitoring site of the River *Thames*. (The blue line indicates the actual values, and the green line represents the fitted values from the SARIMA model during the training period. The red line represents the

**The Prophet Forecasting Model**

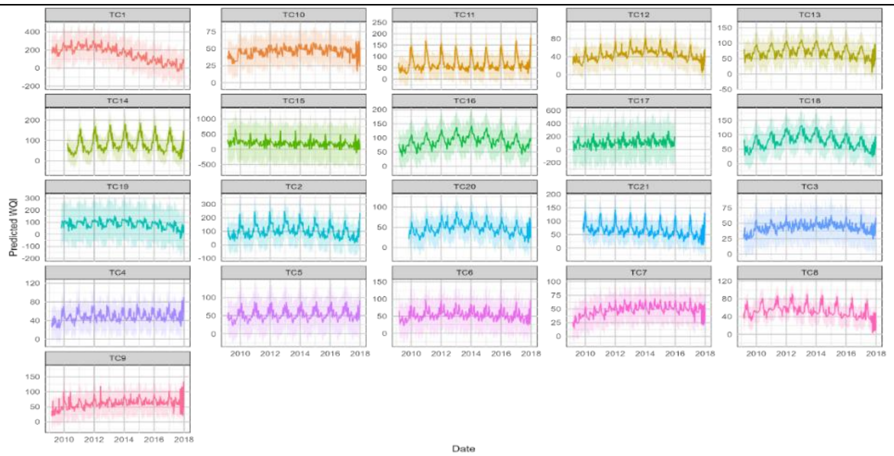


**Figure 3**: Prophet forecasting plots. The solid lines indicate the predicted WQI, while the shaded bands represent the 95% confidence intervals generated by the model.

**Table 1:** Performance matrices for each site.

| SITE | SARIMA | | | | Prophet | | |
|---|---|---|---|---|---|---|---|
| | SARIMA $(p,d,q)(P,D,Q)_{12}$ | RMSE | MAE | MAPE | RMSE | MAE | MAPE |
| TC1 | $(0,0,0)(0,1,1)_{12}$ | 4.14 | 3.26 | 6.54 | 6.94 | 5.05 | 0.093 |
| TC10 | $(0,0,2)(0,1,1)_{12}$ | 1.64 | 1.26 | 1.87 | 2.31 | 1.72 | 0.025 |
| TC11 | $(0,0,0)(0,1,1)_{12}$ | 2.49 | 1.91 | 3.50 | 4.35 | 3.26 | 0.054 |
| TC12 | $(0,1,2)(2,0,2)_{12}$ | 1.71 | 1.11 | 1.64 | 3.20 | 2.00 | 0.029 |
| TC13 | $(2,1,3)(2,0,1)_{12}$ | 2.46 | 2.02 | 3.02 | 3.77 | 2.84 | 0.041 |
| TC14 | $(2,1,2)(2,1,2)_{12}$ | 2.96 | 2.30 | 3.78 | 44.71 | 11.07 | 0.162 |
| TC15 | $(1,1,1)(0,1,1)_{12}$ | 5.54 | 4.49 | 11.05 | 6.90 | 4.95 | 0.109 |
| TC16 | $(0,1,0)(2,0,2)_{12}$ | 2.60 | 2.10 | 3.20 | 3.39 | 2.68 | 0.039 |
| TC17 | $(3,1,2)(1,0,2)_{12}$ | 4.56 | 3.49 | 5.65 | 90.19 | 21.36 | 0.321 |
| TC18 | $(3,1,3)(0,1,1)_{12}$ | 4.09 | 3.24 | 5.38 | 5.86 | 4.07 | 0.063 |
| TC19 | $(3,0,3)(1,1,0)_{12}$ | 5.14 | 4.38 | 7.64 | 12.51 | 6.80 | 0.112 |
| TC2 | $(0,0,1)(0,1,1)_{12}$ | 5.59 | 4.66 | 10.31 | 8.61 | 6.26 | 0.118 |
| TC20 | $(3,1,0)(2,1,1)_{12}$ | 2.06 | 1.39 | 1.96 | 11.39 | 3.54 | 0.047 |
| TC3 | $(0,0,2)(2,1,2)_{12}$ | 3.37 | 2.65 | 4.60 | 4.42 | 3.43 | 0.055 |
| TC4 | $(0,0,1)(2,1,2)_{12}$ | 3.30 | 2.80 | 4.58 | 4.23 | 3.19 | 0.049 |
| TC5 | $(0,0,0)(2,1,2)_{12}$ | 2.75 | 2.22 | 3.39 | 3.99 | 3.09 | 0.045 |
| TC6 | $(0,0,0)(0,1,1)_{12}$ | 3.29 | 2.65 | 4.20 | 4.85 | 3.75 | 0.056 |
| TC7 | $(0,0,1)(2,1,2)_{12}$ | 2.67 | 2.13 | 3.39 | 4.16 | 3.06 | 0.046 |
| TC8 | $(2,1,2)(1,0,1)_{12}$ | 1.41 | 1.08 | 1.61 | 2.49 | 1.64 | 0.024 |
| TC9 | $(0,1,0)(0,0,0)_{12}$ | 3.45 | 2.64 | 4.61 | 4.48 | 3.52 | 0.056 |

The performance of the Prophet model was evaluated using three error metrics: RMSE, MAE, and MAPE. The results showed that Prophet consistently achieved lower MAPE values than SARIMA at all sites, indicating better relative accuracy overall. MAPE values for Prophet ranged from 0.024%–0.321%, suggesting high relative accuracy across the sites. However, Prophet's RMSE and MAE values were occasionally larger than those of SARIMA, especially at sites characterized by sparse observations or high fluctuation. At TC17, for instance, Prophet yielded an RMSE of 90.19 and an MAE of 21.36, compared to SARIMA's much lower values of 4.56 and 3.49. Figure 3 indicates the Prophet forecasting plots with confidence intervals for each monitoring site of the River *Thames*. Several sites exhibit strong seasonality and wider uncertainty, suggesting variability in water quality trends, whereas sites like

TC10 and TC12 show more stable, consistent patterns with narrower prediction intervals. Overall, the model provided valuable insights into the temporal behavior of WQI across the study area.

## CONCLUSIONS/RECOMMENDATIONS

This study aimed to forecast the Water Quality Index (WQI) of the River Thames using two time series models: the SARIMA model and the Prophet model. Both models were effective in forecasting WQI, however, their performance varied depending on the data structure. For the sampled WAWQI data from the River Thames, the Prophet model outperformed SARIMA in terms of MAPE, indicating higher relative accuracy when handling the irregular data. SARIMA performed slightly better in absolute error metrics (RMSE and MAE) at sites with dense observations. Unlike SARIMA, Prophet does not require interpolation and is more sensitive to small datasets. This sensitivity may explain the higher RMSE and MAE observed at two sites with fewer data points for Prophet, a pattern not observed in SARIMA. The difference in accuracy metrics may result from SARIMA's ability to capture absolute errors under stable seasonal patterns, while Prophet performs better at handling relative errors under irregular seasonality due to its flexibility in modeling non-linear trends. These results highlight the importance of selecting forecasting models based on data structure and analytical objectives, as accurate WQI forecasting supports proactive water management and informed decision-making.

## REFERENCES

Box, G., Jenkins, G., & Reinsel, G. (2015). Time Series Analysis-Forecasting and Control. *John Wiley & Sons*.

Taylor, S. J., & Letham, B. (2018). Forecasting at Scale. *Taylor & Francis*, *72*(1), 37–45.