

Assessment and Forecasting of Water Quality Variables of the River Gomti, Lucknow, India

Amit Krishan¹; Ankita Srivastava^{2*}

¹Environmental Specialist, Altinok Consulting Engineering Pvt. Ltd., New Delhi-110077, Delhi, India

²President, Aarush Welfare Society, Lucknow-226017, Uttar Pradesh, India

Corresponding Author: Ankita Srivastava^{2*}

Publication Date: 2026/06/25

Abstract: This paper represents the anticipated monthly values of seven water quality indicators for the selected seven monitoring stations along the Gomti River in Lucknow, India. Time series forecasting technique was used to predict the water quality parameters. The outcomes showed that the suggested model was operating successfully. The Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) and Mean Absolute Error (MAE) values, for pH 1.56-2.18, 0.15-0.22 and 0.12-0.17; for TDS 42.04-79.10, 14.47-19.96 and 30.63-58.99; for TH 12.28-14.85, 4.17-5.29 and 8.70-10.32; for chlorides 1.41-2.60, 7.84-8.50 and 1.01-2.01; for DO 0.63-1.46, 7.46-45.07 and 0.46-1.06; for BOD 0.17-1.09, 3.65-18.83 and 0.13-0.76 and for TC 510-21100, 10.13-42.68 and 421-16220, ranged respectively. The predicted water quality parameter data for the study period (2013–2017) and the anticipated period (2018–2022) revealed a seasonal trend pattern, indicating that the seasonal inflows significantly affect water quality. For all sampling stations, the observed water quality parameter value increased from upstream to downstream, demonstrating the effects of rapid industrialization and urbanization and showing the same behavior for the predicted period. The study results indicate the suitability of the time series technique for predicting river water quality.

Keywords: River Gomti; Water Quality; Statistical Analysis; Time-Series Analysis; ARIMA.

How to Cite: Amit Krishan; Ankita Srivastava (2026) Assessment and Forecasting of Water Quality Variables of the River Gomti, Lucknow, India. *International Journal of Innovative Science and Research Technology*, 11(6), 1226-1253. <https://doi.org/10.38124/ijisrt/26jun378>

I. INTRODUCTION

Human well-being is always related to the availability of fresh water. Water from rivers, lakes, and other surface sources has long been the most prevalent option for municipal uses due to its easy accessibility [1, 2]. Consequently, most human advancements and settlements have occurred along river banks [3, 4]. Water is such a precious resource that even it cannot be replaced, and the benefits to humanity that come from its proper management, even the devastating consequences of its abuse, are undeniable [5, 6, 7].

Transportation, restitution, hydropower generation, and numerous household, industrial, and commercial operations all use river water in various ways [8, 9]. There is substantial pollution in the water from these sources, making the water unusable for beneficial uses and causing water-related diseases [10, 11]. The population suffering from water scarcity has increased sixteen-fold during the 1900s. As of 2005, more than 35% of the world's total population belongs to countries with water stress ranging from moderate to high. By 2050, this number is expected to rise to roughly 52%. Population expansion has sped up in emerging countries,

while environmental management systems have failed to meet the demand [10, 12, 13, 14].

There are two important rivers in the Indo-Gangetic Plain: The Ganga River and the Indus River, and their tributaries are the Gomti, Beas, Yamuna, Sutlej, Ravi Chambal and Chenab rivers. A significant source of drinking water for the Gomti River Basin (GRB) cities, including Lucknow, the Gomti River is a monsoon and groundwater-fed river [15]. Concern has been raised about the growing pollution load in rivers due to wastewater discharges from both residential and industrial sources. The deterioration in river Gomti's water quality due to the discharge of pollutants is well documented in research publications [15, 16, 17, 18, 19].

To keep the pollution levels within acceptable ranges, the prediction of river water quality is essential for effective basin management [11, 12]. The water quality can be modelled utilizing hydrochemical data and linear regression and structural equations, as well as predictability, trend and time-series analysis [20, 21, 22, 23]. Water quality managers employ analytical techniques like regression equations to

compare present water quality levels with predetermined requirements [22, 23, 24].

Gomti River is one of the most contaminated bodies of water in the region because of its extensive role in transporting urban and industrial wastes and agricultural runoff [1, 2, 18]. Therefore, comprehensive water quality management research is required to maintain water quality requirements. There have been estimations of water quality parameters' namely; Normalized Bayesian Information Criterion (BIC), Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and Stationary R-Squared (RMSE) using statistical analysis, time series, Auto-Regressive Integrated Moving Average (ARIMA), Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF).

II. MATERIALS AND METHODS

➤ Sampling Stations and Data Collection

The critical parameters that were considered for this study include pH, Total dissolved solids (TDS), Total

hardness (TH), Chloride, Dissolved oxygen (DO), Biochemical oxygen demand (BOD), and Total coliform (TC). Sampling has been done on monthly basis to obtain data on the aforementioned parameters of water quality at all selected seven sampling stations i.e., Station-1 Manjhi Ghat, Station-2 Up Stream water intake, Station-3 Kuriyaghat, Station-4 Downstream Mohan Meakins, Station-5 Nishat Ganj Bridge, Station-6 Upstream Barrage and Station 7 met STP Nala Bharwara of the Gomti river downstream named S1, S2, S3, S4, S5, S6, S7 respectively by Uttar Pradesh Pollution Control Board (UPPCB), Lucknow from the years 2013 to 2017. The first sample site, S1, was located on the outskirts of Lucknow city, in a low-pollution zone. The remaining five stations, S2-S6, were located in a high-pollution area in the middle of the stretch inside the city. The final sampling site, S7, was downstream of the city zone and suffered from moderate pollution. Fig. 1 shows the features of the sample stations.

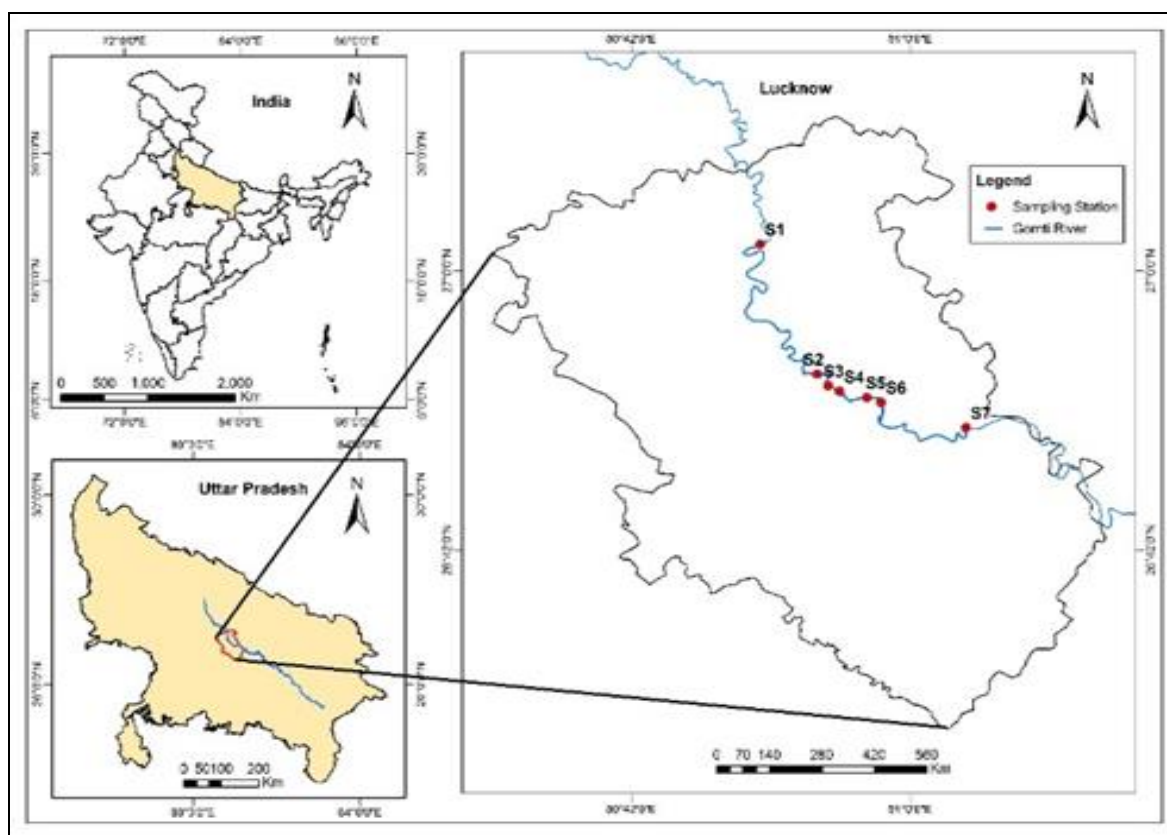


Fig 1 Sampling Stations in the River Gomti, Lucknow (Uttar Pradesh) India

➤ Study Area

The Indo-Ganga-Brahmaputra Plain, where the GRB is located, is the world's greatest alluvial tract. GRB is a noteworthy geographic feature of the Indian sub-continent because of its low altitude, low gradient, high rains, exceptionally productive soil, and large population density. The GRB experiences a dry hot summer from March to May, very heavy monsoon from June to September and a very cool winter from December to February like a humid subtropical

climate. Approximately 75%–80% of the year's total precipitation falls during the monsoon season, making it the wettest time of year. The post-monsoon season is marked by occasional rainfall and temperatures as high as 47°C and as low as 2°C. The GRB receives around 75% of its rainfall in the form of groundwater, while the rest evaporates or evapotranspires [1, 2, 18].

Lucknow is situated at an altitude of 123 meters above sea level, between 26.30° and 27.10°N and 80.30° and 81.13°E in the heart of the Indo-Gangetic alluvial plain. December–February sees a chilly dry winter and April–June sees a scorching dry summer. River Gomti covers a distance of 980 km from Madho- Tanda to the Himalayan foothills, a distance of around 50 km. It is common for the river to release over 600 m³/s of water during the monsoon, but this drops to 30 m³/s during the summer [15]. Lucknow's principal source of water, the Gomti River, has been under severe strain due to the city's growing population, which grew by 38% between 2001 and 2011 [19]. Drinking, industrial, and irrigating water depend on this river's water. As a result, groundwater production is affected by the over-extraction of subsurface water through tube wells positioned near riverbeds; the river water is a significant source of recharge [1, 2].

➤ *Statistical Analysis*

Different statistical tests (mean, median, mode, standard deviation, kurtosis, skewness, and coefficient of variation) have been performed to analyses the collected data.

- *R-Squared and Stationary R-Square*

R-squared is the proportion of the variance for a dependent variable that measures how much the model can describe the total variance in a series when it is stationary. When there is a trend or seasonal pattern, the stationary R-squared is favored over the ordinary R-squared. This is because the stationary section of the model is compared to a basic mean model using the stationary R-squared. Positive, negative, or zero stationary R-squared values are possible. Positive values mean the present model outperforms the baseline model. In the current model, positive numbers indicate better than the baseline model [22, 23, 25].

- *Root Mean Square Error (RMSE)*

The RMSE is a commonly used measure of value differences. It determines how far a dependent series differs from the predicted level [22, 23, 25]. The RMSE of an estimator in terms of the estimator parameter is calculated using equation 1

$$RMSE(\theta) = \sqrt{E[(\theta - \hat{\theta})^2]} \tag{1}$$

- *Mean Absolute Percentage Error (MAPE)*

MAPE is a metric for how far a dependent series deviates from its modelled level. It doesn't matter what units are used; therefore, it may be used to relate series come along with different units. Accuracy is given as a percentage as;

$$M = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{A_i - F_i}{A_i} \right| \tag{2}$$

Where

A_i - the actual value and

F_i - the predicted value.

MAPE = zero (for perfect fit),

But it has no restrictions at the upper level [22, 23, 25].

- *Mean Absolute Error (MAE)*

The MAE is a metric that measures how far a series differs from the level that its model anticipated. It is expressed in the same units as the original series. In addition, the MAE is a metric for determining how far forecasts or projections differ from the final result. It is provided by;

$$MAE = \frac{1}{n} \sum_{i=1}^n |F_i - A_t| = \frac{1}{n} \sum_{i=1}^n |e_i| \tag{3}$$

The absolute error is e_i, the prediction is F_i, and the determined value is A_i. In time-series analysis, it is a typical measure of forecast error [22, 23, 25].

- *Maximum Absolute Percentage Error (MaxAPE)*

The highest anticipated error is measured by the MaxAPE. It is used to imagine a forecasted worst-case scenario [22, 23, 25].

- *Maximum Absolute Error (MaxAE)*

The MaxAE, is a metric that estimates the biggest anticipated error in the same units as the series that it is reliant on. It is possible to utilise it to imagine the worst-case scenario while making forecasts. It's possible that MaxAE and MaxAPE will occur at separate times throughout the series. When the absolute error for a large series value is somewhat greater than the absolute error for a small series value, the maximum absolute error as well as maximum absolute percentage error occur at the smaller series value [22, 23, 25].

- *Normalized Bayesian Information Criterion (BIC)*

It is an overall measure of a model's fit that accounts for its complexity. It's a score based on mean square error that considers the model's number of parameters and the length of the series [22, 23, 25].

$$BIC = \chi^2 + k \cdot \ln(n) \tag{4}$$

It is used to assess the constant (k). It considers the performance of a parameterized model in terms of its ability to forecast data and penalizes the model's complexity, which is defined as the total number of parameters it contains.

- *Time Series*

It collects data points taken at regular intervals and often at successive times. Future projection methods are based on known previous events to predict data points before they are observed and included in the realm of time-series analysis. The time-series model considers that observations closer in time to one another are more tightly connected than those further apart in Time-series models. On the other hand, they are typically defined by a natural one-way ordering of time, which means that the values for a current period are derived from prior values instead of future values [26].

• *Auto-Regressive Integrated Moving Average (ARIMA)*

Three terms characterize the ARIMA model of a time series (p, d, q). The technique of detecting integer, generally extremely small values as 0, 1 or 2 of model patterns p, d, and q in data is known as time series identification. When the value is 0, the element in the model is no longer necessary. The middle element, d, is studied before p and q. The goal is to determine if the process is stationary and, if not, to make it stationary before deciding the values of p and q. A stationary process has a consistent mean and variation across the research period. AR (p) represents the time series of an autoregressive model [22, 23, 25], is defined as;

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + \varepsilon_t \quad (5)$$

Where the term ε_t is source of randomness which is referred as white noise, α_i denotes constants. It is supposed that it possesses the following characteristics:

$$E[\varepsilon_t] = 0,$$

$$E[\varepsilon_t^2] = \sigma^2,$$

$$E[\varepsilon_t \varepsilon_s] = 0 \text{ for all } t \neq s$$

Because autoregressive and moving average components might coexist in a series, both forms of correlations are necessary to model the patterns. To grasp this, if both elements are present only at lag 1, consider the linear equation;

$$y_t = x_t \beta + \varepsilon_t \quad (6)$$

$$\varepsilon_t = \rho \varepsilon_{t-1} + v_t \quad (7)$$

Where

$$1 < \rho < 1.$$

Where v_t is dependent and identically distributed (iid) and from expectation values;

$$E(v_t, v_{t-1}) = 0 \quad (8)$$

In this approach, the disturbance follows a first-order autoregressive model (AR1) process. As a result, the present error mixes the previous error plus some shock. As a result, the equation may be rewritten as follows:

$$y_t = x_t \beta + \rho \varepsilon_{t-1} + v_t \quad (9)$$

Also, we know that

$$y_{t-1} = x_{t-1} \beta + \varepsilon_{t-1} \quad (10)$$

$$\varepsilon_{t-1} = y_{t-1} - x_{t-1} \beta \quad (11)$$

From Eq. (9)

$$y_t = x_t \beta + \rho(y_{t-1} - x_{t-1} \beta + v_t)$$

$$y_t = x_t \beta + \rho y_{t-1} - \rho x_{t-1} \beta + v_t \quad (12)$$

• *Auto Correlation Functions (ACF) and Partial Auto Correlation Functions (PACF)*

ACF and PACF analysis are necessary to find an appropriate model for a given time series data. The link between the data in a time series is reflected in these statistical metrics. Plotting the ACF and PACF across consecutive time lags is typically beneficial for modelling and forecasting [23]. The sequence of AR and MA phrases can be determined using these plots. This may be represented as:

For a time series $\{x(t), t = 0, 1, 2, \dots\}$ at lag k the autocovariance is defined as;

$$\gamma_k = \text{Cov}(x_t, x_{t+k}) = E[(x_t - \mu)(x_{t+k} - \mu)] \quad (13)$$

At lag k the autocorrelation coefficient is defined as;

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad (14)$$

The mean of the time series is μ equal to $E[x_t]$. The time-series variance is the autocovariance at lag zero, i.e. γ_0 . The autocorrelation coefficient k is dimensionless; hence it is unaffected by the measurement scale, $-1 \leq \rho_k \leq 1$. Statistics Box and Jenkins termed γ_k as the theoretical autocovariance function (ACVF) and ρ_k as the theoretical autocorrelation function (ACF).

After allowing for data at intermediate delays, the partial autocorrelation function (PACF) is used to assess the correlation between an observation k periods ago and the current observation (i.e., at lags k). PACF (1) and ACF (1) are similar at lag 1.

The real or theoretical ACF and PACF values cannot be established since the stochastic process underlying a time series is unknown. Instead, these values should be derived from the training data, which is the most recent time series available. Sample ACF and PACF are the estimated ACF and PACF values from the training data, respectively. At lag k, the most credible sample estimate for the ACVF is;

$$c_k = \frac{1}{n} \sum_{t=1}^{n-k} (x_t - \mu)(x_{t+k} - \mu) \quad (15)$$

Then the estimate for the sample ACF at lag k is given by;

$$r_k = \frac{c_k}{c_0} \quad (16)$$

Here, $\{x(t), t = 0, 1, 2, \dots\}$ is the training series of size n with mean μ .

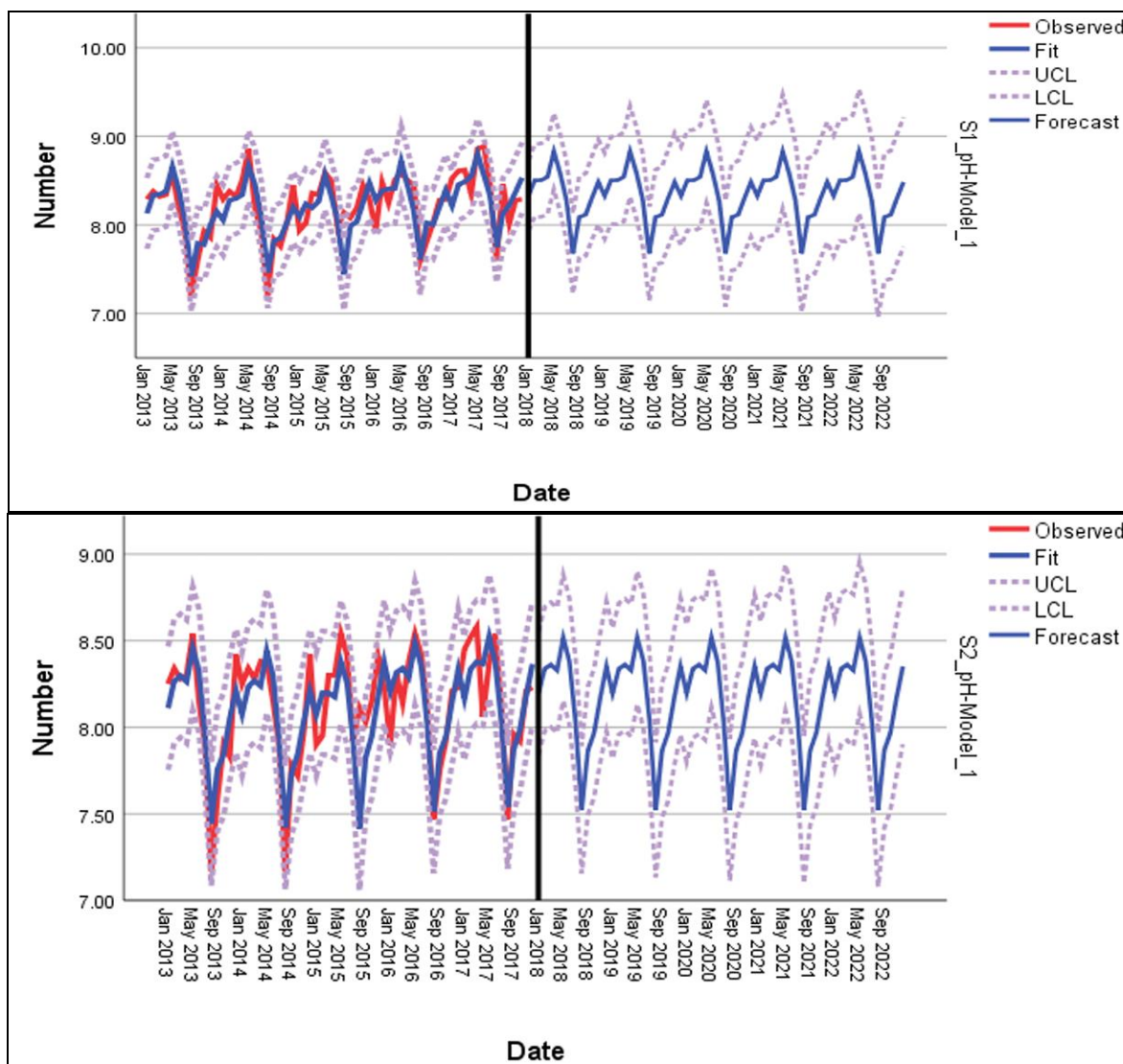
III. RESULTS AND DISCUSSION

Using statistical and time series analysis calculated by SPSS 26, the water quality at sampling stations S1, S2, S3, S4, S5, S6 and S7 of the river Gomti at Lucknow (India) has been discussed. Tab. 1 depicts the mean, median, mode, minimum, maximum, range, standard deviation (SD), skewness, kurtosis, and coefficient of variation for the parameters pH, TDS (mg/l), TH (CaCO₃, mg/l), Chlorides (mg/l), DO (mg/l), BOD (mg/l), and TC (MPN/100 ml) at all sampling stations. For all water quality parameters at all sampling sites, Trend and time series analysis are explained in Tab. 2 using the ARIMA model, stationary R squared, R squared, RMSE, MAPE, MaxAPE, MAE, Ljung-Box Q (18), ACF, and PACF.

➤ pH

All sampling stations (S1–S7) recorded pH values ranging from 6.89-8.59, 6.89-8.86, 7.11-8.59, 7.23-8.59, and 7.23-8.89 for 2013, 2014, 2015, 2016, and 2017 respectively. According to IS 2296:1992, Class A, the permissible pH range is 6.5-8.5. The pH of the Gomti river water ranged from 6.89 to 8.89, indicating that it is either mildly acidic or alkaline. Iqbal et al. (2019) [17] have also found the pH of the

Gomti river in the range of 6.22 to 8.6. The mean, median, and mode are close to each other and reflect normal behavior, a standard deviation which indicates that the curve is symmetrical and platykurtic because Kurtosis value is lower than 3 at all sampling stations. The curve is moderately skewed at S1, S4, S5, S6, and S7 while highly skewed at S2 and symmetrical at S3. The prediction model outperforms the baseline model because stationary R squared and R squared values behave identically. The dependent series is relatively close to the level that the model anticipated because the RMSE values are relatively low. Parmar and Bhardwaj (2015) [23] observed a similar value and pattern in the Yamuna River. Using Ljung-Box Q (18) model, the statistic value ranges from 22.52 to 36.93, the significance level varies from 0.00 to 0.13, and the degree of freedom is 16 for all sampling stations. For a period of 5 years, from January 2018 to December 2022, a simple seasonal ARIMA model was used for prediction, and it was discovered that the pH value ranges from 7.20 to 8.83, and the quality of the water is nearly constant at all sampling stations, as calculated at 95% confidence limits. Fig. 2 depicts the plot of time series, observed data, best fit, lower confidence limit (LCL), and upper confidence limit (UCL) for pH at all sampling stations.



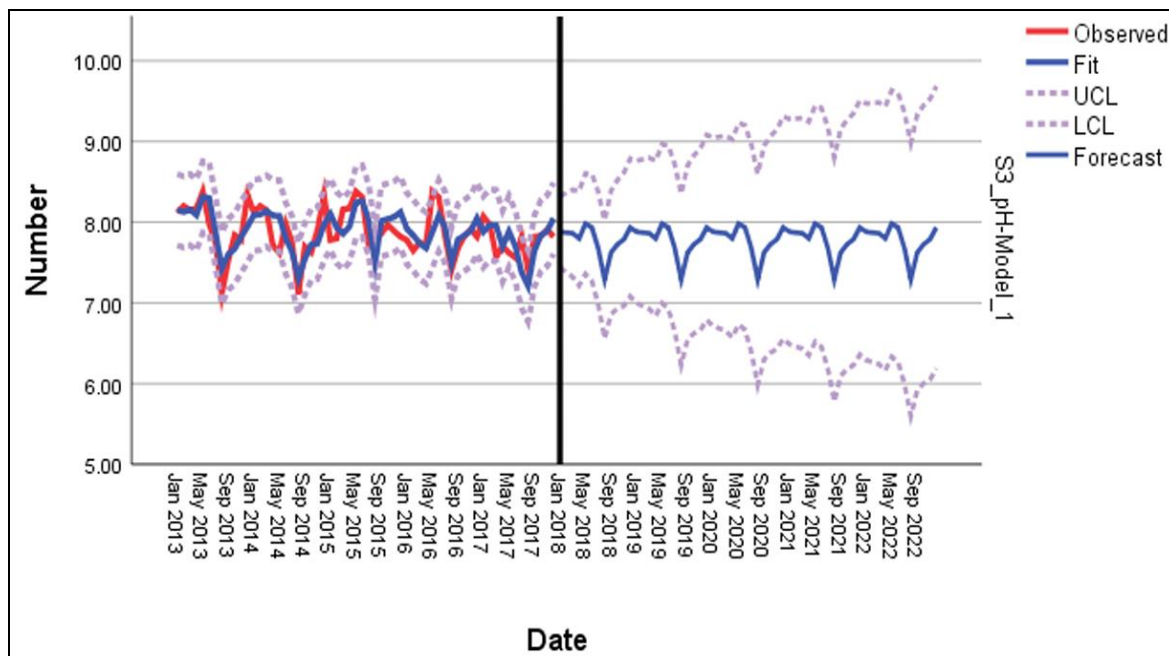


Fig 2 Statistical Analysis of the Gomti River Water Quality Parameters

Table 1 Statistical Analysis of the Gomti River Water Quality Parameters

Parameters	S1	S2	S3	S4	S5	S6	S7
pH							
Mean	8.22	8.12	7.87	7.80	7.70	7.64	7.48
Median	8.29	8.21	7.83	7.77	7.69	7.64	7.50
Mode	8.29	8.42	7.82	7.71	7.69	7.62	7.43
Minimum	7.20	7.17	7.11	7.04	7.00	6.95	6.89
Maximum	8.89	8.58	8.38	8.26	8.15	8.09	7.82
Range	1.69	1.41	1.27	1.22	1.15	1.14	0.93
SD	0.36	0.32	0.29	0.25	0.23	0.22	0.19
Kurtosis	0.86	1.08	0.35	1.43	1.78	2.15	1.60
Skewness	-0.77	-1.05	-0.23	-0.53	-0.80	-0.94	-0.92
Coeff. of variation	0.04	0.04	0.04	0.03	0.03	0.03	0.03
TDS							
Mean	231.37	240.83	274.23	286.53	304.77	316.47	343.73
Median	233.00	240.00	255.00	266.00	282.00	293.00	314.00
Mode	184.00	192.00	224.00	232.00	274.00	254.00	312.00
Minimum	112.00	118.00	128.00	136.00	152.00	160.00	172.00
Maximum	332.00	346.00	482.00	496.00	538.00	554.00	616.00
Range	220.00	228.00	354.00	360.00	386.00	394.00	444.00
SD	47.21	49.84	84.81	85.77	88.70	89.24	99.63
Kurtosis	0.11	0.03	0.56	0.51	0.61	0.58	0.89
Skewness	-0.10	-0.08	1.09	1.03	1.05	0.97	1.03
Coeff. of variation	0.20	0.21	0.31	0.30	0.29	0.28	0.29
TH							
Mean	196.77	202.70	211.97	218.80	228.47	235.87	249.42
Median	206.00	212.00	220.00	228.00	238.00	244.00	258.00
Mode	188.00	232.00	236.00	244.00	254.00	258.00	258.00
Minimum	120.00	128.00	134.00	140.00	148.00	156.00	170.00
Maximum	236.00	244.00	256.00	262.00	274.00	282.00	294.00
Range	116.00	116.00	122.00	122.00	126.00	126.00	124.00
SD	30.16	31.70	31.90	32.14	32.25	32.24	31.90
Kurtosis	-0.13	-0.21	-0.27	-0.23	-0.18	-0.18	-0.17
Skewness	-0.79	-0.80	-0.72	-0.78	-0.80	-0.80	-0.76
Coeff. of variation	0.15	0.16	0.15	0.15	0.14	0.14	0.13
Chloride							

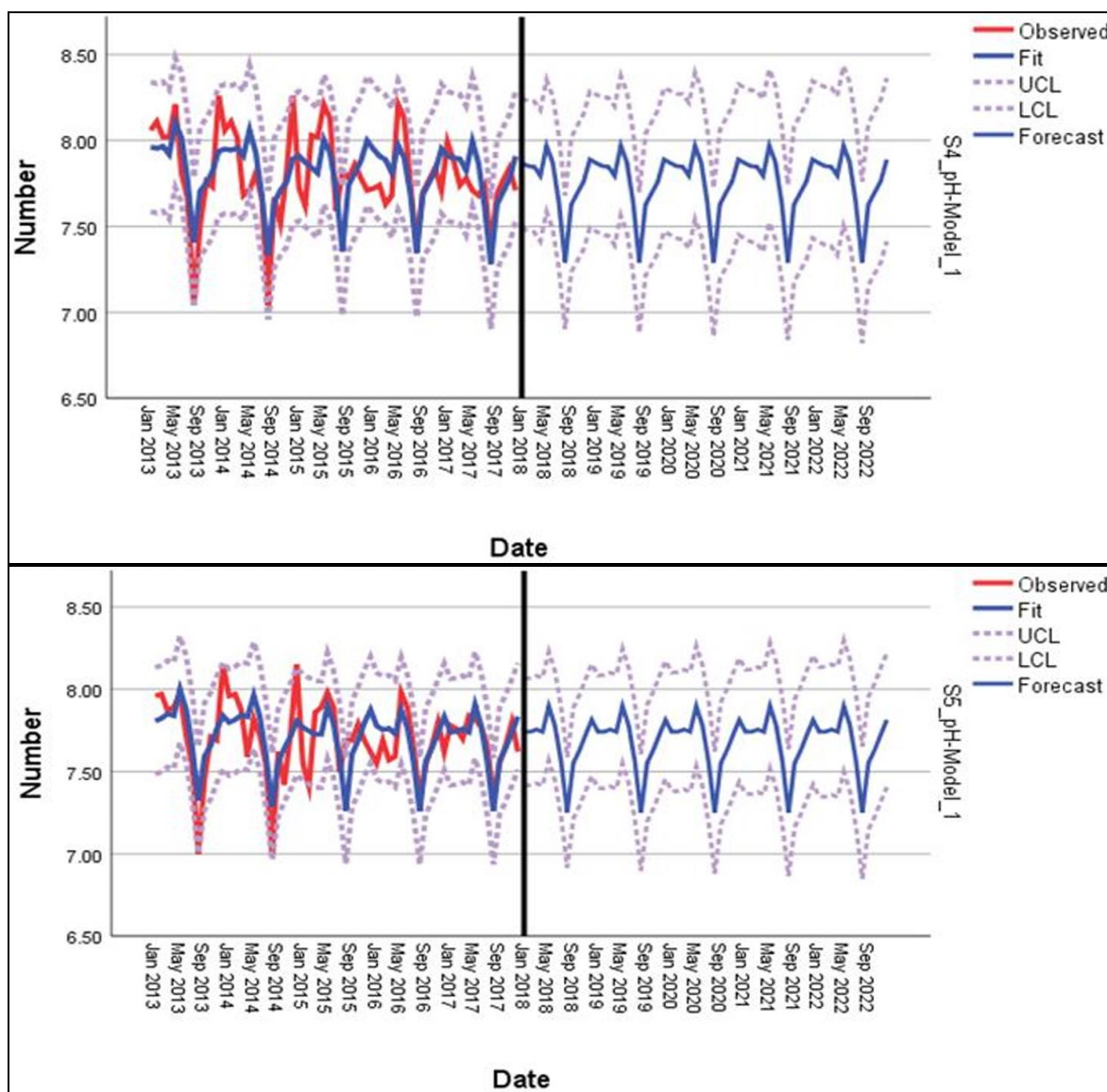
Mean	13.05	14.53	16.72	18.42	21.33	23.33	27.18
Median	13.00	14.00	16.00	18.00	21.00	23.00	28.00
Mode	13.00	14.00	15.00	20.00	24.00	25.00	29.00
Minimum	8.00	11.00	12.00	13.00	14.00	15.00	18.00
Maximum	17.00	19.00	22.00	24.00	27.00	29.00	33.00
Range	9.00	8.00	10.00	11.00	13.00	14.00	15.00
SD	1.73	1.86	2.39	2.59	2.93	3.23	3.44
Kurtosis	0.45	-0.54	-0.59	-0.74	-0.48	-0.41	-0.17
Skewness	-0.10	0.31	0.25	0.12	-0.16	-0.39	-0.62
Coeff. of variation	0.13	0.13	0.14	0.14	0.14	0.14	0.13
DO							
Mean	6.81	7.26	5.02	4.36	3.02	2.45	1.52
Median	6.90	6.70	5.05	4.20	2.80	2.10	1.20
Mode	2.80	6.50	3.50	4.20	3.20	3.00	1.00
Minimum	2.80	2.90	0.70	1.30	1.10	0.60	0.30
Maximum	11.40	10.90	9.50	8.40	6.00	5.10	3.80
Range	8.60	8.00	8.80	7.10	4.90	4.50	3.50
SD	2.57	2.13	2.41	1.74	1.19	1.12	0.86
Kurtosis	-0.81	-0.77	-0.50	-0.03	0.75	0.07	0.73
Skewness	-0.13	0.23	0.19	0.55	1.07	0.83	1.11
Coeff. of variation	0.38	0.29	0.48	0.40	0.40	0.46	0.57
BOD							
Mean	4.19	3.41	5.19	6.39	8.37	9.77	12.48
Median	3.00	3.30	4.60	6.00	8.50	10.00	13.00
Mode	2.90	3.30	4.00	6.00	8.50	10.00	13.00
Minimum	2.30	3.10	3.60	4.50	6.00	6.50	8.50
Maximum	11.00	4.20	9.00	10.00	11.50	12.50	18.60
Range	8.70	1.10	5.40	5.50	5.50	6.00	10.10
SD	2.57	0.24	1.49	1.48	1.49	1.53	1.70
Kurtosis	2.06	1.51	-0.32	-0.18	-0.86	-0.80	2.48
Skewness	1.89	1.23	0.96	0.92	0.25	-0.16	-0.07
Coeff. of variation	0.61	0.07	0.29	0.23	0.18	0.16	0.14
TC							
Mean	2618	4015	15953	30583	74933	99450	151033
Median	2600	3950	9400	32000	70000	94000	150000
Mode	2800	3500	7000	22000	70000	94000	170000
Minimum	1400	2400	4900	13000	32000	70000	94000
Maximum	4000	7000	40000	58000	110000	130000	240000
Range	2600	4600	35100	45000	78000	60000	146000
SD	587	1047	11643	10488	15050	16377	30930
Kurtosis	-0.12	1.42	-0.63	0.08	0.61	-0.42	0.47
Skewness	0.45	0.92	0.93	0.62	-0.09	0.02	0.39
Coeff. of variation	0.22	0.26	0.73	0.34	0.20	0.17	0.21

Table 2 Trend and Time Series Analysis of the Gomti River Water Quality

	S1	S2	S3	S4	S5	S6	S7
<i>Model Fit statistics for pH</i>							
Stationary R-squared	0.72	0.71	0.61	0.68	0.71	0.71	0.71
R-squared	0.69	0.70	0.42	0.45	0.49	0.46	0.39
RMSE	0.20	0.18	0.22	0.19	0.16	0.16	0.15
MAPE	1.96	1.71	2.18	1.96	1.66	1.61	1.56
MAE	0.16	0.14	0.17	0.15	0.13	0.12	0.12
MaxAPE	8.59	8.58	6.03	5.94	5.56	5.70	5.43
MaxAE	0.70	0.70	0.50	0.47	0.43	0.43	0.41
Normalized BIC	-3.09	-3.31	-2.89	-3.21	-3.51	-3.51	-3.65
<i>Ljung-Box Q (18) for pH</i>							
Statistics	24.88	24.19	27.98	35.77	36.93	22.52	22.98
DF	16	16	16	16	16	16	16

Sig.	0.07	0.09	0.03	0.00	0.00	0.13	0.11
<i>Model Fit statistics for TDS</i>							
Stationary R-squared	0.73	0.73	0.66	0.67	0.68	0.68	0.68
R-squared	0.22	0.22	0.32	0.32	0.35	0.37	0.38
RMSE	42.04	44.32	70.46	71.33	72.20	71.40	79.10
MAPE	14.47	14.81	19.96	19.74	18.65	17.54	18.19
MAE	30.63	32.83	50.85	53.14	53.54	51.79	58.99
MaxAPE	109.44	106.48	117.05	101.96	99.88	104.83	100.93
MaxAE	122.57	125.64	217.30	205.61	209.28	230.16	227.40
Normalized BIC	7.61	7.72	8.65	8.67	8.70	8.67	8.88
<i>Ljung-Box Q (18) for TDS</i>							
Statistics	19.94	18.94	23.56	24.92	22.28	21.16	25.68
DF	16	16	16	16	16	16	16
Sig.	0.22	0.27	0.10	0.07	0.13	0.17	0.06
<i>Model Fit statistics for TH</i>							
Stationary R-squared	0.71	0.67	0.73	0.70	0.74	0.75	0.73
R-squared	0.82	0.83	0.86	0.84	0.85	0.84	0.79
RMSE	13.06	13.21	12.28	12.84	12.74	13.30	14.85
MAPE	5.29	5.13	4.47	4.49	4.22	4.17	4.31
MAE	9.57	9.47	8.70	9.13	8.97	9.26	10.32
MaxAPE	23.29	20.69	18.59	21.53	18.57	18.90	20.60
MaxAE	38.91	38.15	34.41	43.91	40.47	43.08	48.61
Normalized BIC	5.28	5.30	5.22	5.24	5.30	5.38	5.53
<i>Ljung-Box Q (18) for TH</i>							
Statistics	16.53	20.35	22.30	21.35	23.80	18.70	22.05
DF	16	16	15	16	15	15	16
Sig.	0.42	0.21	0.10	0.17	0.07	0.23	0.14
<i>Model Fit statistics for Chloride</i>							
Stationary R-squared	0.60	0.62	0.68	0.59	0.67	0.67	0.71
R-squared	0.35	0.38	0.49	0.50	0.46	0.46	0.44
RMSE	1.41	1.48	1.74	1.84	2.17	2.41	2.60
MAPE	8.09	7.84	8.32	7.93	8.34	8.50	7.84
MAE	1.01	1.14	1.37	1.44	1.73	1.90	2.01
MaxAPE	42.58	27.44	28.06	24.70	33.76	33.22	40.28
MaxAE	3.56	3.41	4.20	5.35	5.46	5.77	7.25
Normalized BIC	0.82	0.92	1.32	1.35	1.68	1.89	2.05
<i>Ljung-Box Q (18) for Chloride</i>							
Statistics	22.01	25.84	10.96	11.12	9.36	10.62	18.24
DF	16	16	15	16	16	16	16
Sig.	0.14	0.06	0.76	0.80	0.90	0.83	0.31
<i>Model Fit statistics for DO</i>							
Stationary R-squared	0.51	0.67	0.73	0.68	0.53	0.56	0.60
R-squared	0.79	0.92	0.64	0.71	0.49	0.36	0.33
RMSE	1.19	0.63	1.46	0.95	0.86	0.90	0.71
MAPE	14.53	7.46	37.56	20.74	24.60	32.78	45.07
MAE	0.81	0.46	1.06	0.70	0.66	0.69	0.54
MaxAPE	98.93	41.80	397.91	118.82	120.74	182.35	186.01
MaxAE	4.26	1.80	4.78	2.51	2.58	2.92	2.46
Normalized BIC	0.55	-0.71	0.96	0.11	-0.17	-0.07	-0.55
<i>Ljung-Box Q (18) for DO</i>							
Statistics	16.50	30.65	15.7	26.13	37.97	16.86	17.28
DF	15	15	15	15	16	16	16
Sig.	0.35	0.01	0.40	0.04	0.00	0.40	0.37
<i>Model Fit statistics for BOD</i>							
Stationary R-squared	0.59	0.52	0.65	0.63	0.73	0.72	0.70
R-squared	0.90	0.50	0.64	0.67	0.81	0.83	0.60
RMSE	0.82	0.17	0.90	0.86	0.67	0.65	1.09
MAPE	18.83	3.65	11.72	10.23	6.20	5.48	6.23

MAE	0.64	0.13	0.63	0.66	0.52	0.53	0.76
MaxAPE	68.06	14.51	44.27	32.60	17.41	20.46	31.15
MaxAE	2.11	0.61	2.50	2.56	1.85	1.59	4.99
Normalized BIC	-0.20	-3.29	-0.07	-0.09	-0.60	-0.65	0.38
<i>Ljung-Box Q (18) for BOD</i>							
Statistics	41.92	22.93	22.05	25.45	19.76	30.31	12.42
DF	15	15	16	15	15	15	15
Sig.	0.000	0.09	0.14	0.04	0.18	0.01	0.65
<i>Model Fit statistics for TC</i>							
Stationary R-squared	0.59	0.55	0.60	0.60	0.58	0.56	0.50
R-squared	0.26	0.28	0.55	0.51	0.46	0.43	0.54
RMSE	510	899	7909	7400	11117	12604	21100
MAPE	16.99	16.38	42.68	19.07	11.46	10.13	11.41
MAE	421	632	5488	5290	8049	9480	16220
MaxAPE	81.73	68.52	278.25	182.19	119.15	33.76	41.03
MaxAE	1144	2365	23557	23685	38126	40173	55087
Normalized BIC	12.61	13.74	18.09	17.96	18.77	19.09	20.05
<i>Ljung-Box Q (18) for TC</i>							
Statistics	32.07	23.51	17.53	18.44	14.56	39.21	19.66
DF	16	16	16	16	16	15	16
Sig.	0.01	0.10	0.35	0.30	0.56	0.00	0.24



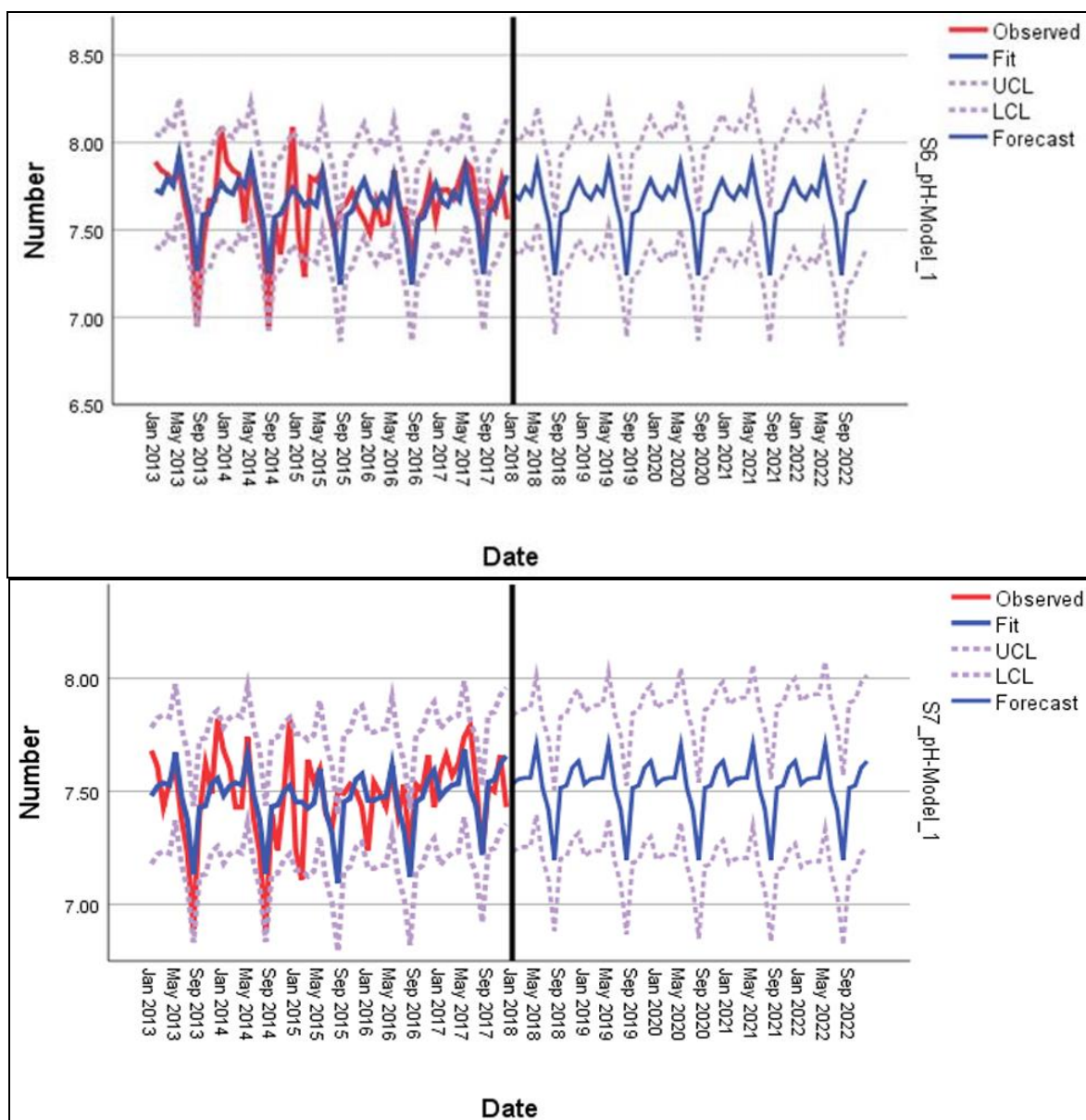
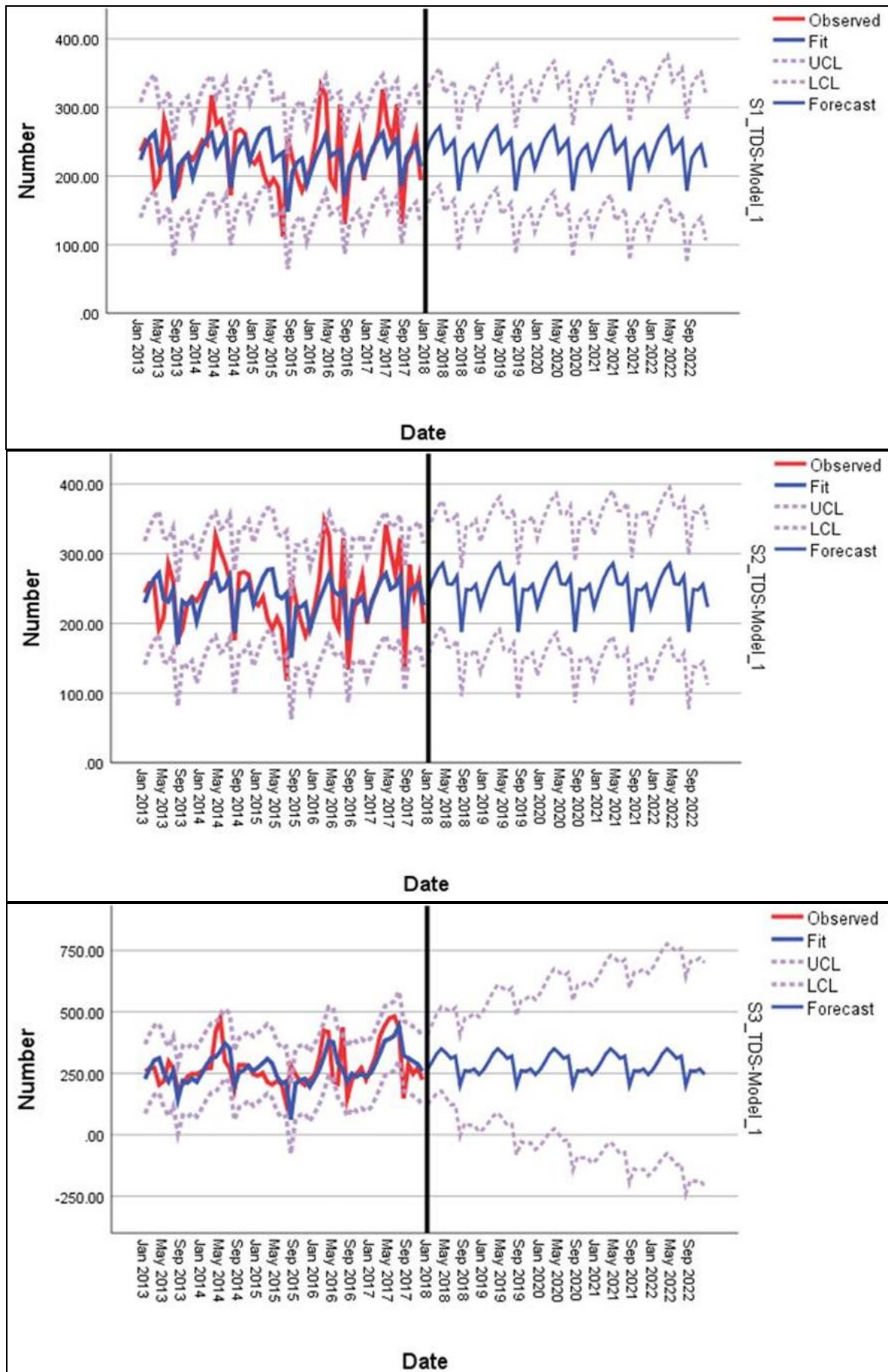


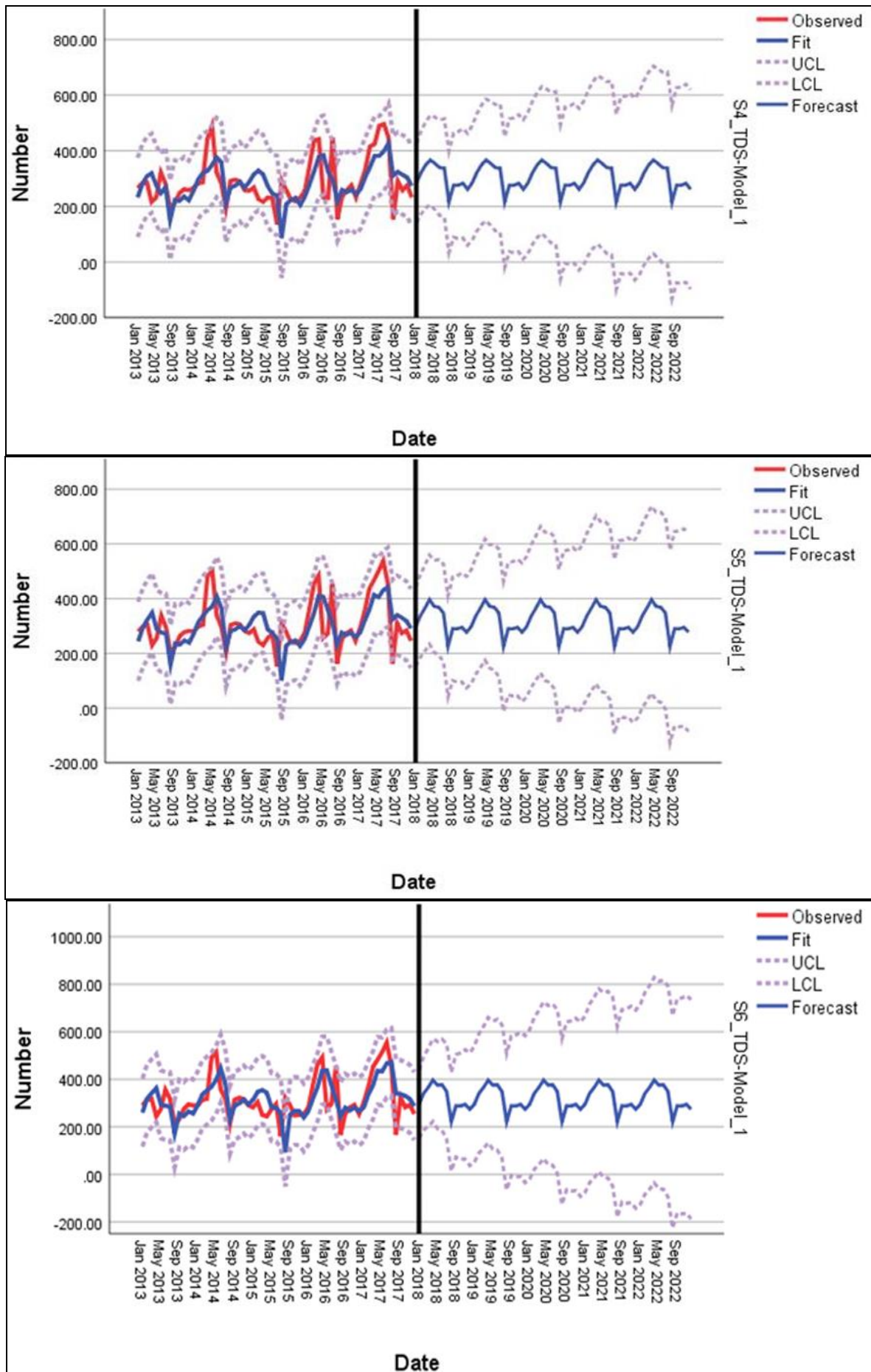
Fig 3 Representation of Trend and Time Series Analysis by Observed, Best Fit, LCL and UCL by Graphically of pH at all Sampling Stations for the River Gomti

➤ *Total Dissolved Solids (TDS)*

TDS concentrations were under the desired range of 500 mg/l at S1, S2, S3, and S4 but exceeded it at S5, S6, and S7, according to the IS 2296:1992, Class A. The lowest TDS was found at S1, while the highest was found at S7. Areas with high population density release significant amounts of solid waste, domestic wastewater discharge, encroachment, and sewage, resulting in a gradual increase from S1 to S7 [17]. Mean, median, and mode values are not equal and the Standard deviation is exceptionally high at all sampling stations, indicating the values of the TDS are not comparable to one another and the curve is not normal. The curve is highly skewed at all sampling stations except at S6, which is moderately skewed. RMSE number is extraordinarily high, indicating that the dependent series is not even close to the predicted level at all sampling stations. The time series model

outperforms the baseline model because stationary R squared and R squared values represent identical behavior. The RMSE value is extremely high, indicating that the dependent series is not near the model predicted level. Using the Ljung-Box Q (18) model, the value of statistics for all sites ranges from 18.94 to 25.68, with a significant level of 0.06 to 0.27, degree of freedom is 16 and a simple seasonal ARIMA model was used for prediction for all sampling stations. Using plots of ACF, PACF, lag, and time series, it is observed that the value of TDS lies between 178.64 to 272.24 (S1), 187.81 to 286.20 (S2), 200.94 to 350.14 (S3), 215.78 to 366.98 (S4), 225.41 to 396.21 (S5), 221.10 to 396.30 (S6) and 255.36 to 441.36 (S7). The water quality does not get affected at all sampling stations, as shown in Fig. 3 for the forecasted period of January 2018 to December 2022, which is calculated at a 95% confidence interval.





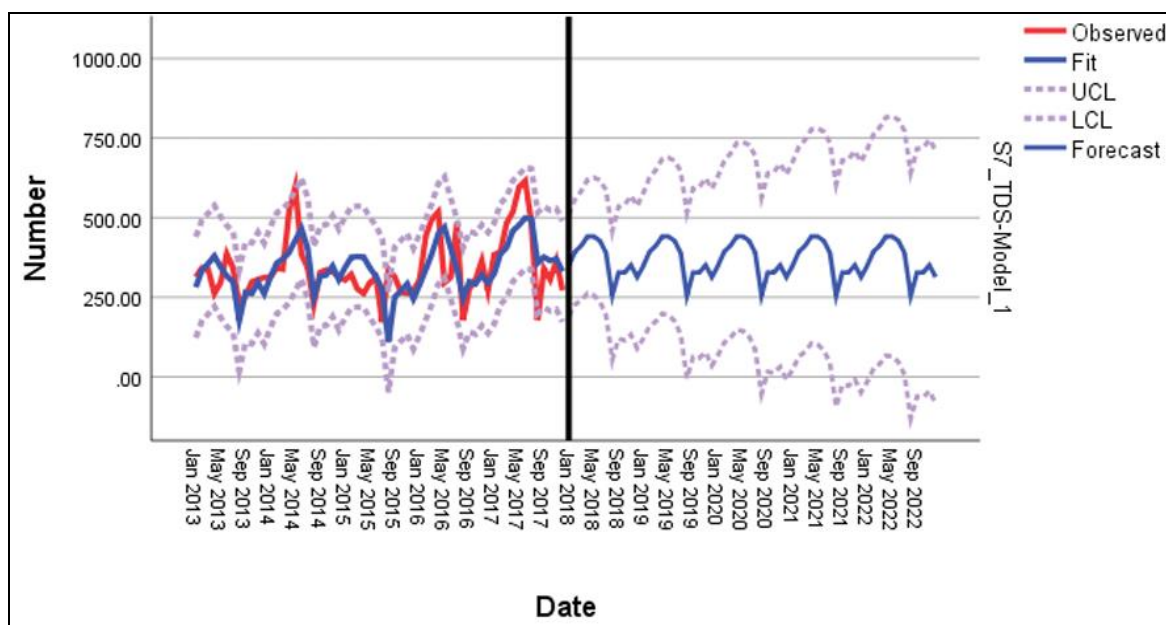


Fig 4 Representation of Trend and Time Series Analysis by Observed, Best Fit, LCL and UCL by Graphically of TDS at all Sampling Stations for the River Gomti

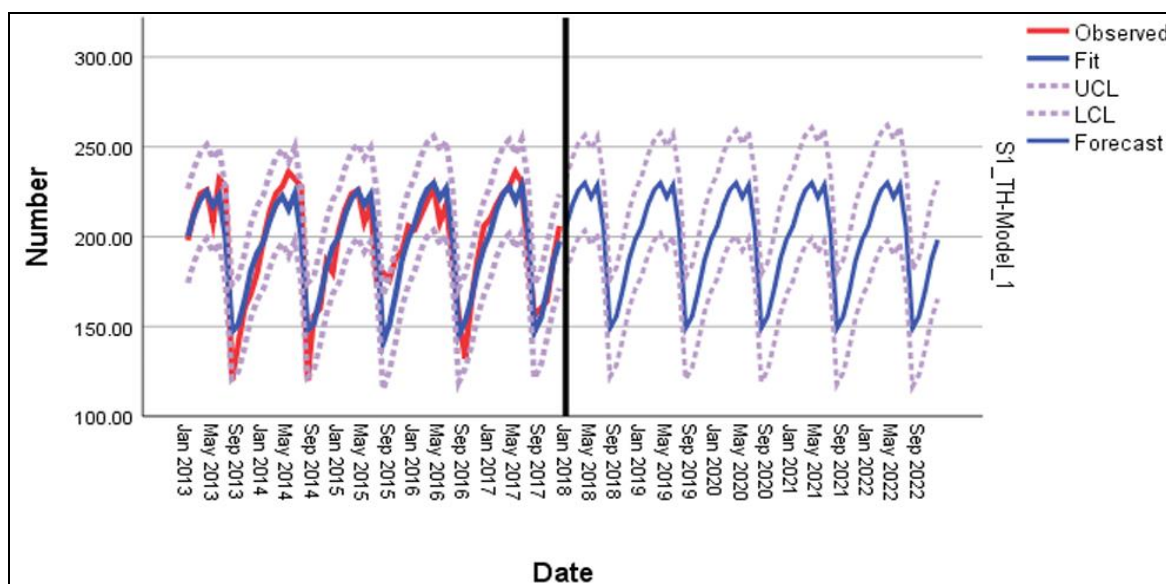
➤ *Total Hardness (TH)*

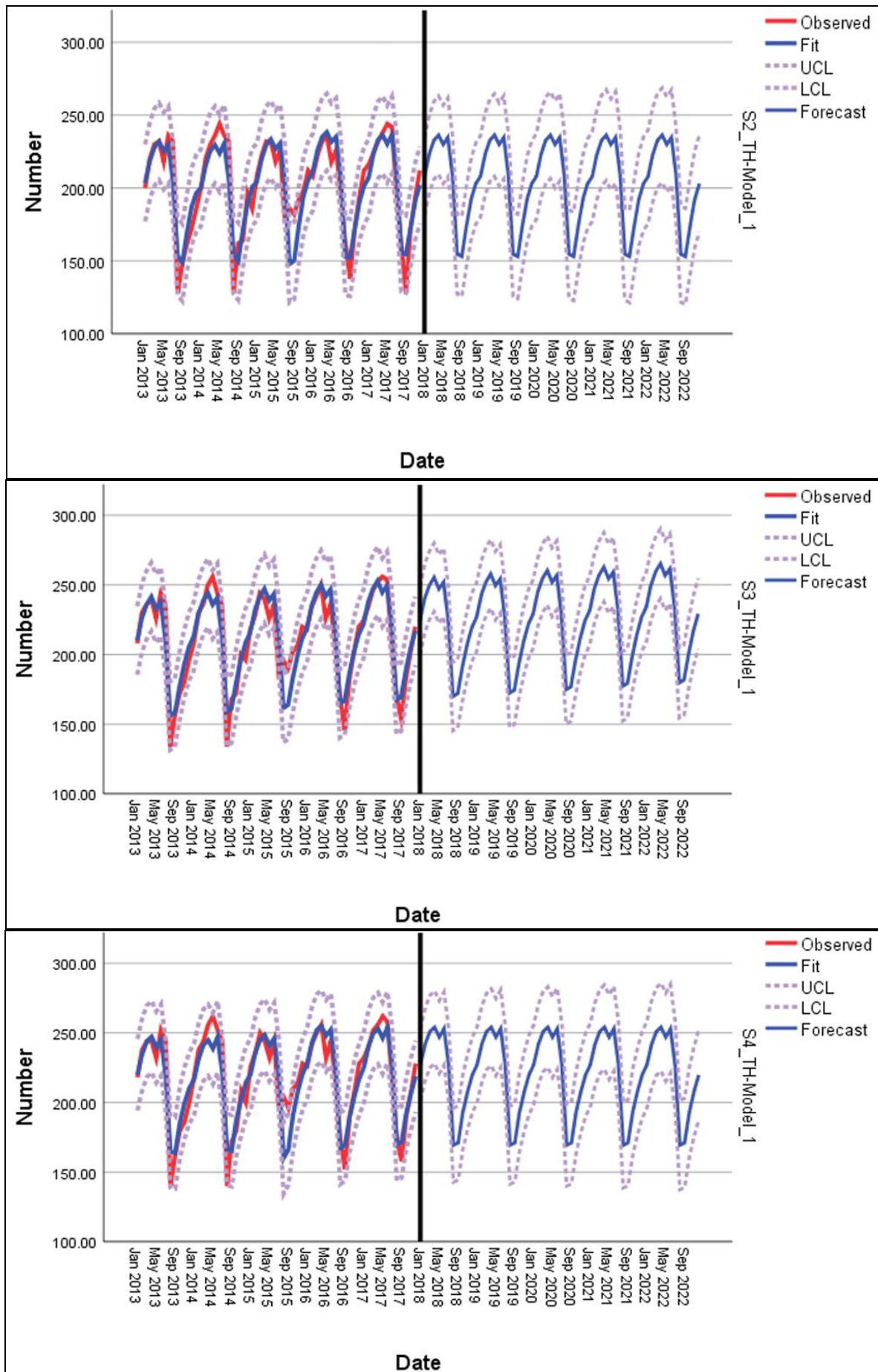
During the study period, the lowest TH values were 120, 120, 176, 132, and 128 mg/l at S1, while the highest values were 294, 294, 276, 282, and 292 mg/l at S7. When advancing from S1 to S7, TH concentrations increased significantly, and the recommended limit for TH is 200 mg/l, according to IS 2296:1992, Class A. The Gomti river water is classified as moderately hard to hard water-based on TH values [27]. The behavior is consistent across all sampling stations. The data spread is high for all sample stations, and the curve is symmetrical and platykurtic. The model performance was superior to the baseline model because stationary R squared and R squared values reflect the same behavior across all sites. Because the RMSE value is high, the dependent series is not close to the model level predicted. Statistics range from 16.53 to 23.80 using the Ljung-Box Q(18) model, with significance levels ranging from 0.07 to 0.42 and degrees of

freedom ranging from 15 to 16. The ARIMA (Simple Seasonal) model was fitted at S1, S2, S4, and S7, while the ARIMA (Winters' Additive) model was fitted at S3, S5, and S6, with 95% confidence limits of boundary lines. ACF, PACF, lag, and time series show that the value of TH ranges from 149.12 to 229.92 (S1), 153.07 to 236.27 (S2), 169.87 to 265.27 (S3), 169.37 to 254.17 (S4), 187.06 to 281.15 (S5), 193.15 to 285.29 (S6) and 196.51 to 281.31(S7) (Fig. 4). Water quality is affected at all sampling stations for the predicted five years, from January 2018 to December 2022.

➤ *Chloride*

During the years 2013, 2014, 2015, 2016, and 2017, the average chloride concentration in selected water samples was 10-32 mg/l, 11-32 mg/l, 10-31 mg/l, 11-33 mg/l, and 8-33 mg/l, respectively.





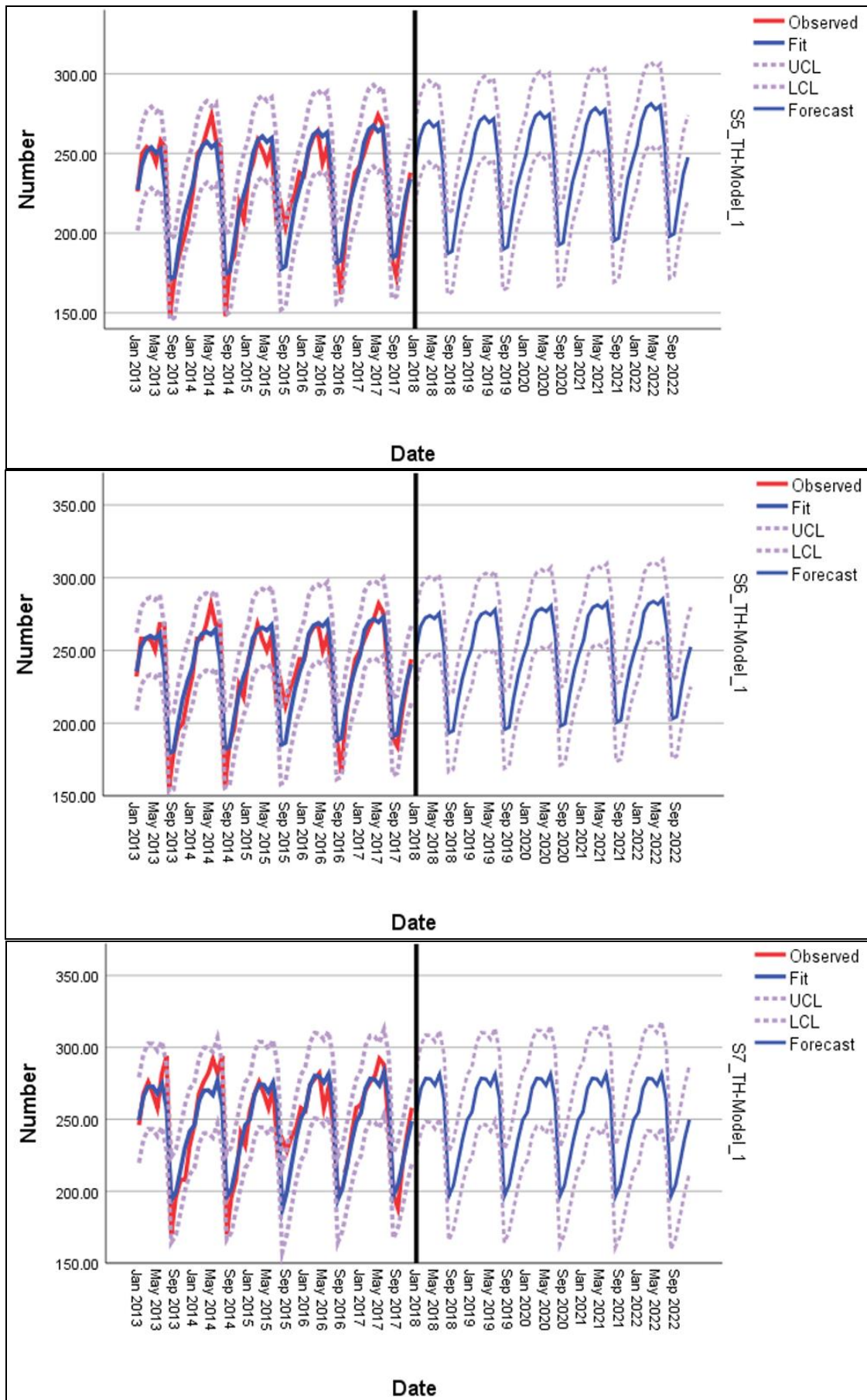
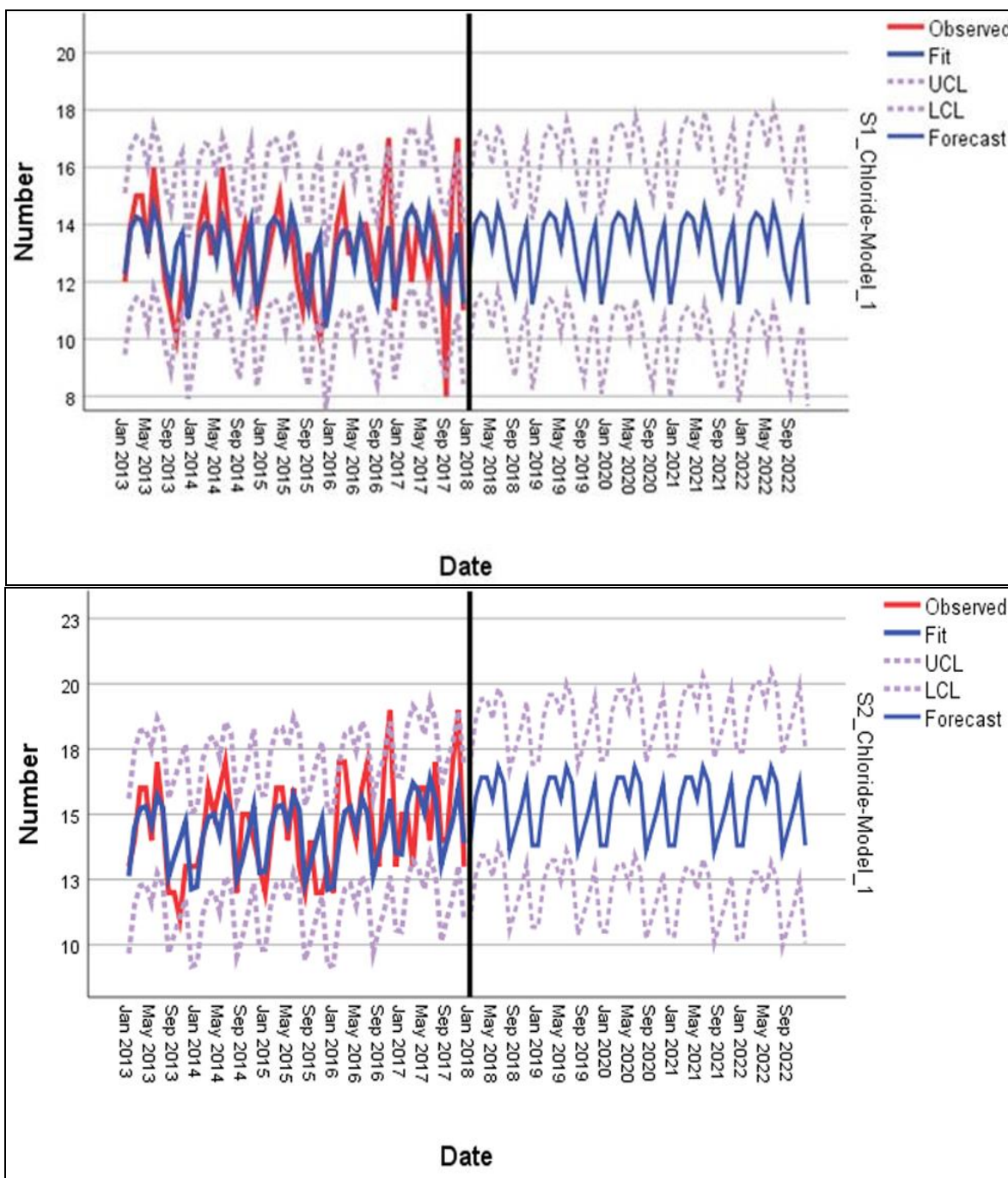
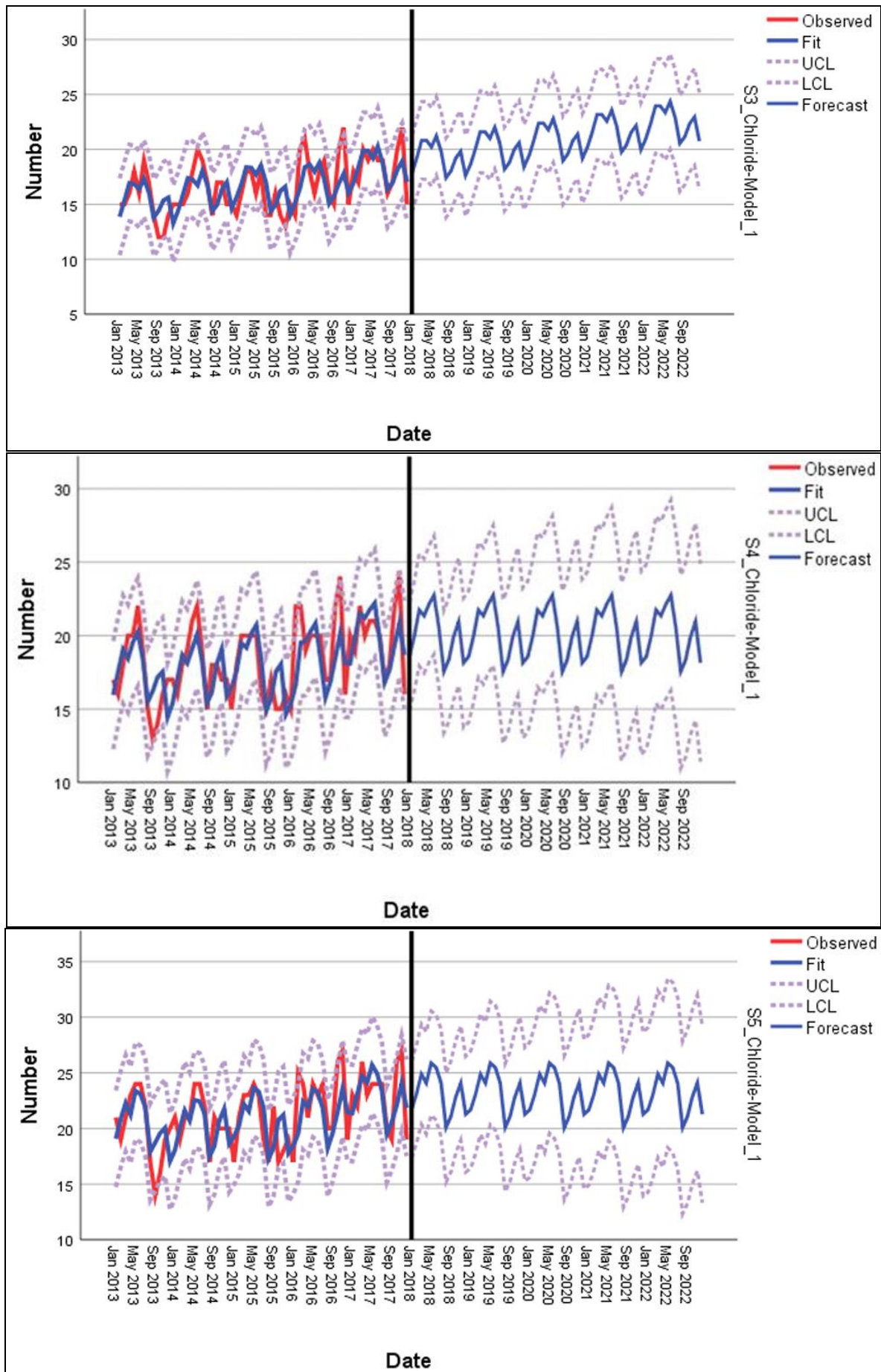


Fig 5 Representation of Trend and Time Series Analysis by Observed, Best Fit, LCL and UCL by Graphically of TH at all Sampling Stations for the River Gomti

The observed chloride levels were well below the recommended limit as IS 2296:1992, Class A i.e. 250 mg/l. The minimum and maximum chloride concentrations at S1 and S7 during the study period were 8 and 33 mg/l, respectively. It behaves normally; however, the data spread is large across all sampling stations. The curve is symmetrical and platykurtic for all sample stations, and stationary R squared and R squared values reveal similar behavior, indicating that the prediction model outperforms the baseline model. The RMSE value is high; therefore, the dependent series is not close to the model predicted level. The Ljung-Box Q(18) model yields statistics ranging from 9.36 to 25.84,

a significance level of 0.06 to 0.90, and a degree of freedom of 15 to 16. All sampling stations except S3 have ARIMA (Simple Seasonal) models fitted, whereas S3 has ARIMA (Winters' Additive) models fitted and the boundary lines are at 95% confidence limits. ACF, PACF, lag, and time series show that the value of Chloride ranges from 11 to 15 for S1, 14 to 17 for S2, 17 to 24 for S3, 18 to 23 for S4, 20 to 26 for S5, 21 to 28 for S6, and 25 to 33 for S7 (Fig. 5) and that the water quality is unaffected at all sampling stations for the predicted 5-year period from January 2018 to December 2022.





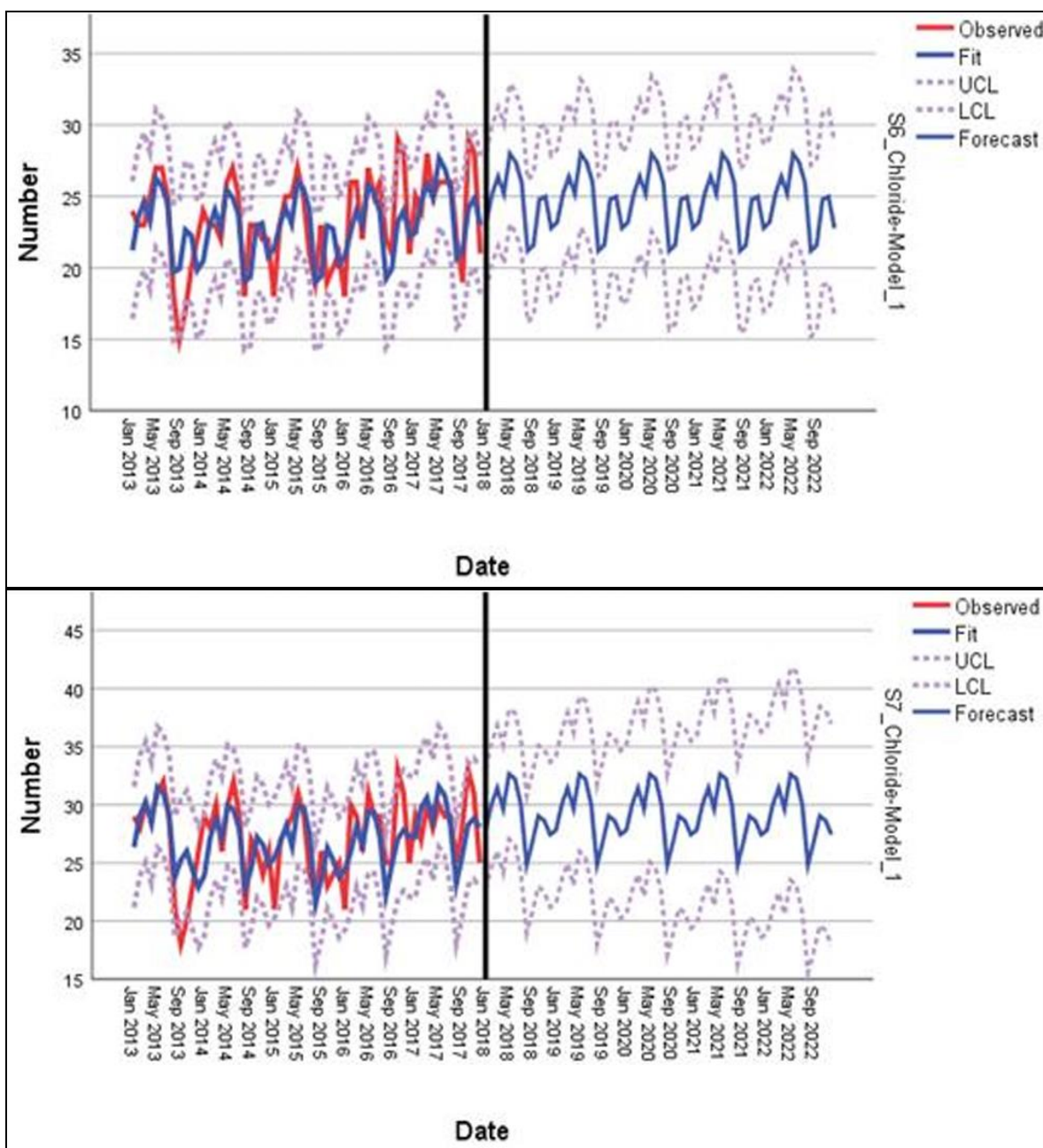
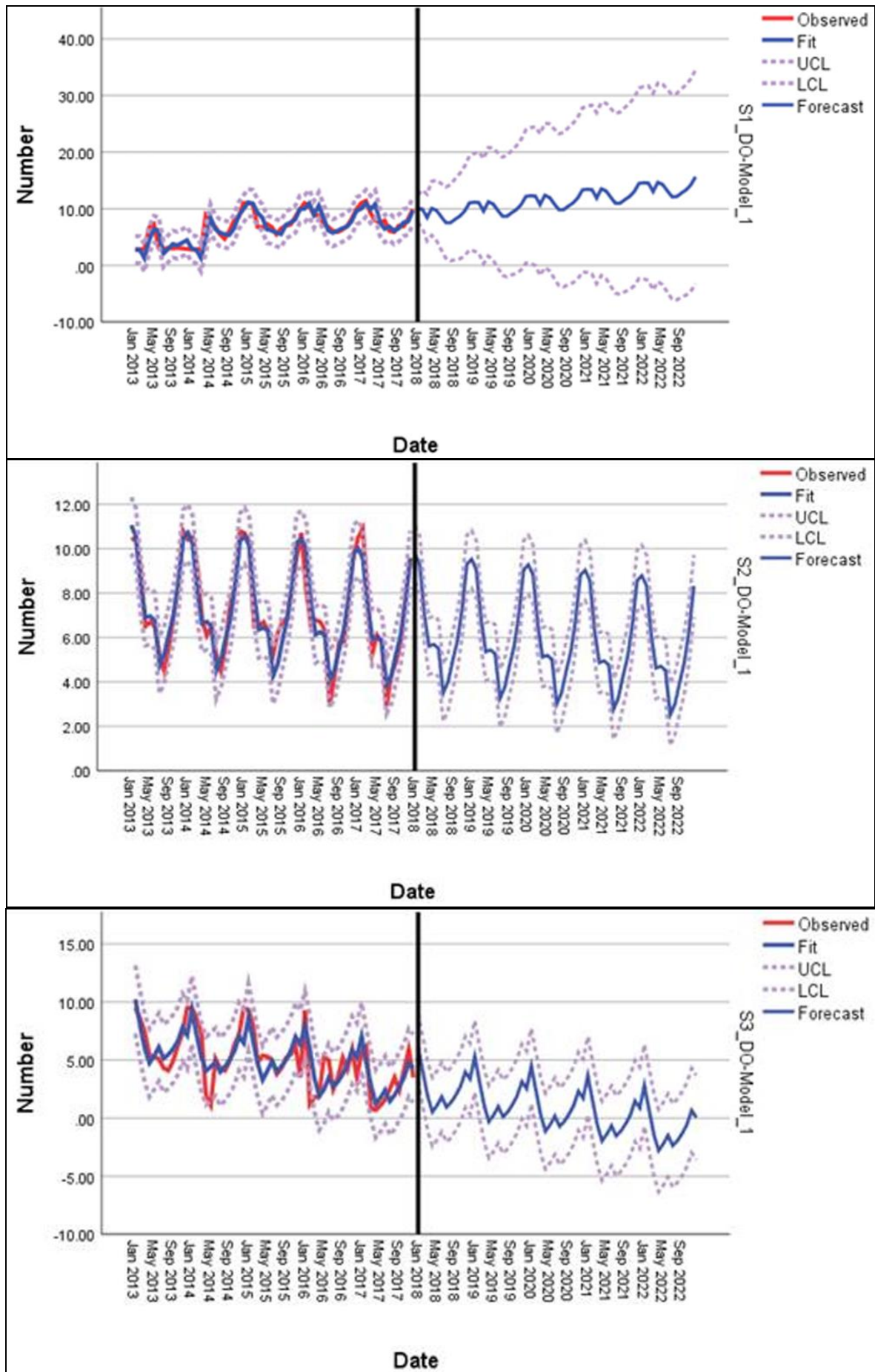


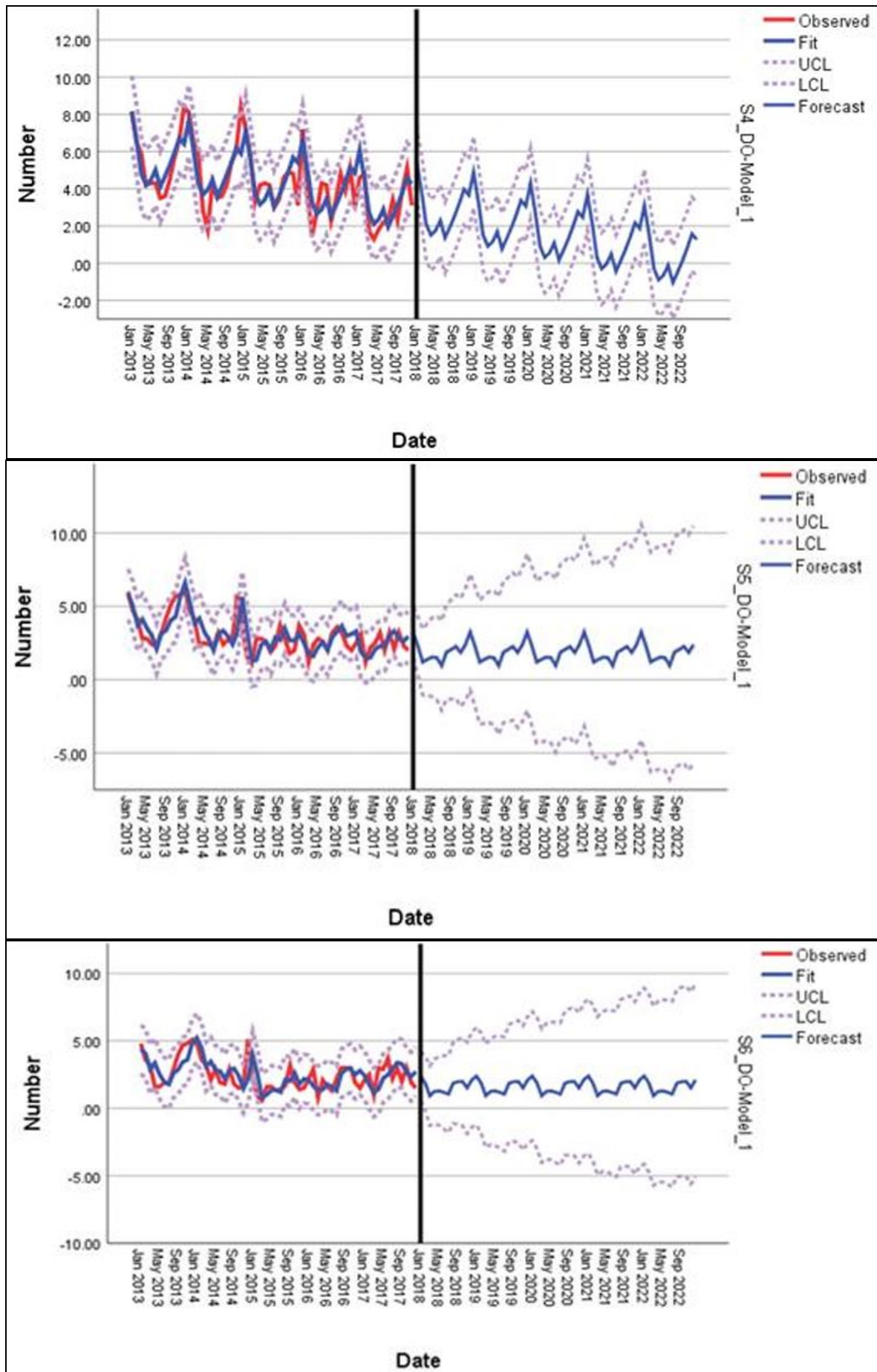
Fig 6 Representation of Trend and Time Series Analysis by Observed, Best Fit, LCL and UCL by Graphically of Chloride at all Sampling Stations for the River Gomti

➤ *Dissolved Oxygen (DO)*

DO levels at S1, S2, S3, S4, S5, S6, and S7 were 2.8-11.4 mg/l, 2.9-10.9 mg/l, 0.7-9.5 mg/l, 1.3-8.4 mg/l, 1.1-6.0 mg/l, 0.6-5.1 mg/l, and 0.3-3.8 mg/l, respectively, during the study period. The minimum value observed at all sample locations during the summer and lean rainfall periods was significantly below the permitted limit of 6 mg/l as per IS 2296:1992. Except at S4, the mean, median, and mode are not the same at all sample stations; hence, they only perform normally at S4. Data dispersion is high and asymmetrical at all sample stations, and the curve is platykurtic. The time series model outperforms the baseline model for all sample stations because stationary R squared and R squared values have comparable behavior. Except for S1 and S3, RMSE values appear to be low and dependent series near the model's

anticipated level. Parmar and Bhardwaj (2015) [23] discovered a similar value and pattern in the Yamuna River. For all sampling stations, the value of the statistic is between 15.7 and 37.97, the significance level is between 0.00 and 0.40, and the degree of freedom is between 15 and 16. S1, S2, S3, and S4 were fitted with the ARIMA (Winters' Additive) model, while S5, S6, and S7 were fitted with the ARIMA (Simple Seasonal) model, with the boundary lines at 95% confidence limits. Fig. 6 shows that the value of DO ranges from 7.53 to 15.63 (S1), 2.53 to 9.77 (S2), -2.79 to 6.13 (S3), -1.03 to 5.49 (S4), 0.96 to 3.24 (S5), 0.93 to 2.39 (S6), and 0.19 to 1.27 (S7), and the quality of water is affected at all sampling stations except S1 over the predicted period of 5 years from January 2018 to December 2022.





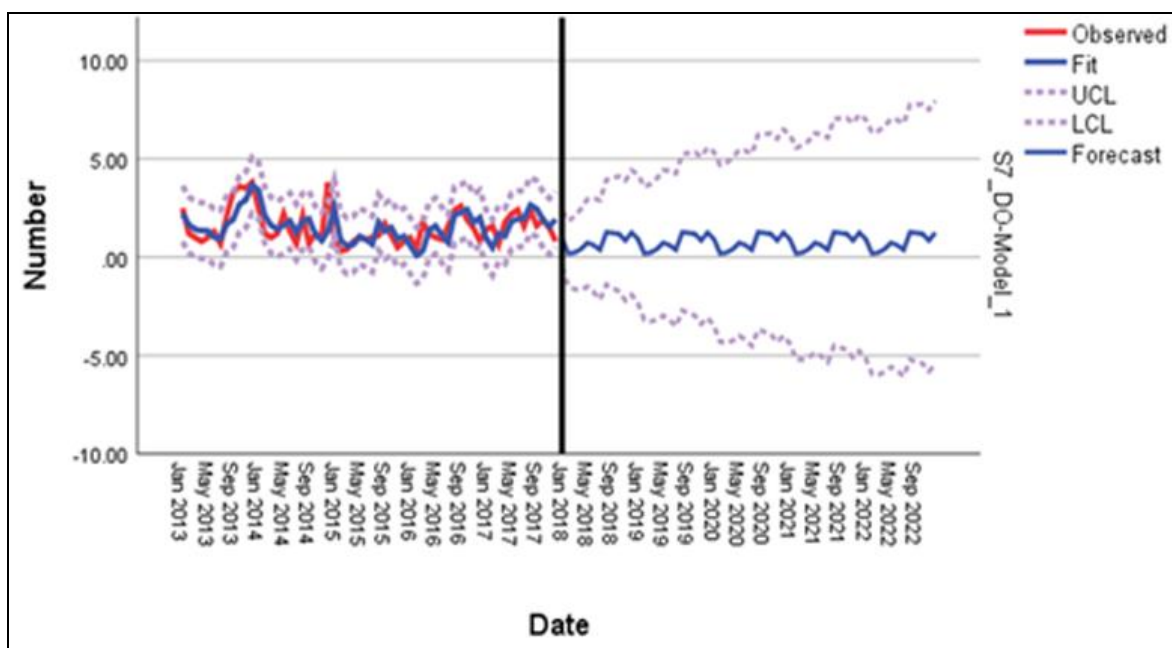
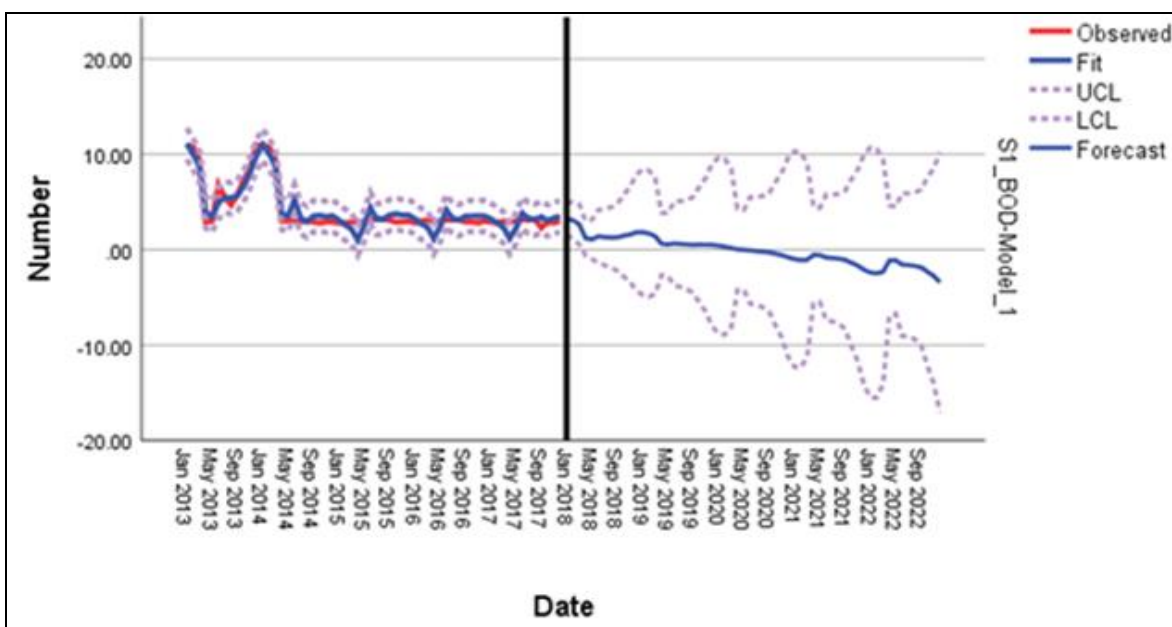


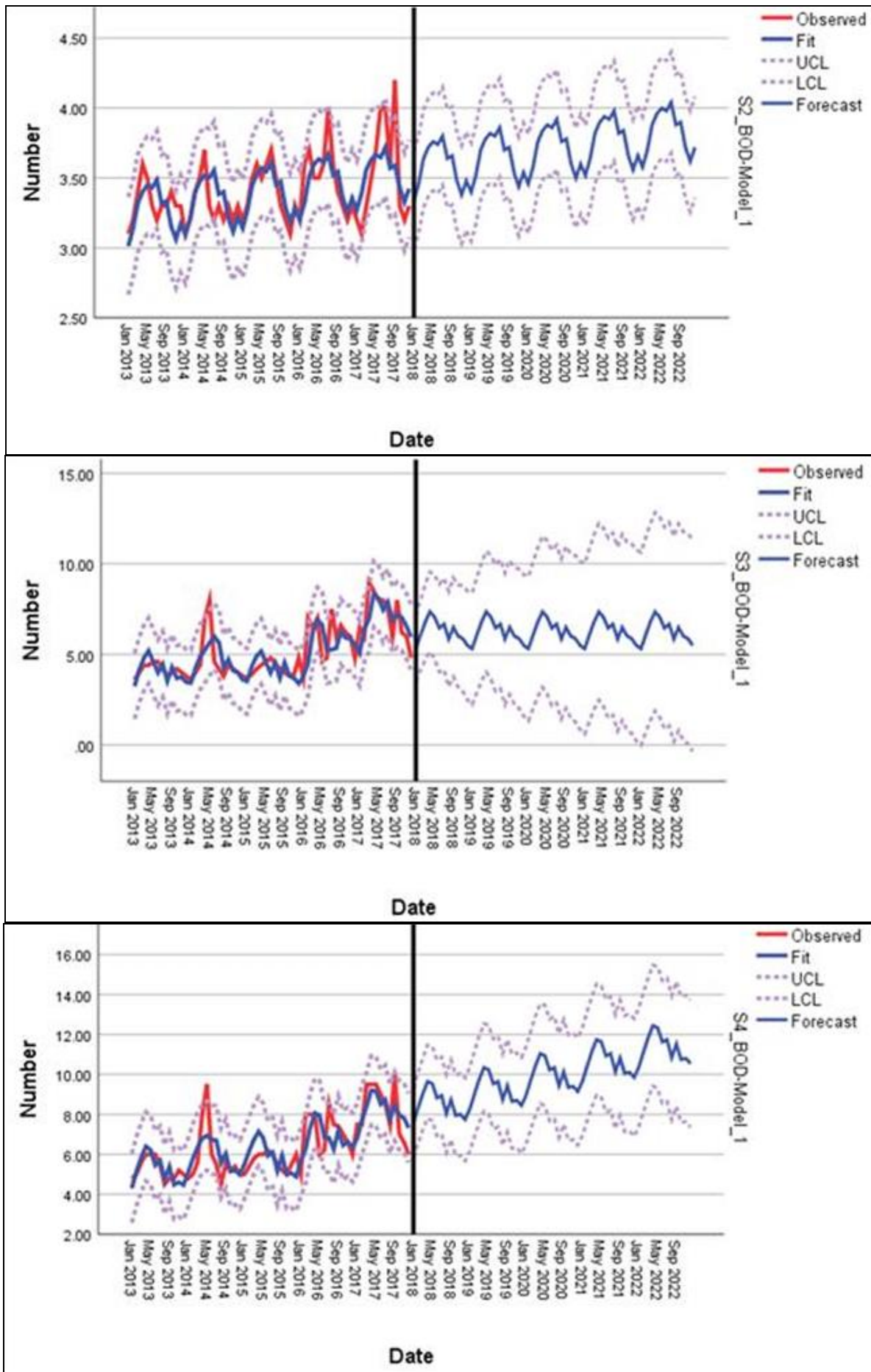
Fig 7 Representation of Trend and Time Series Analysis by Observed, Best Fit, LCL and UCL by Graphically of DO at all Sampling Stations for the River Gomti

➤ *Biochemical Oxygen Demand (BOD)*

At S1, S2, S3, S4, S5, S6, and S7, respectively, BOD values in river water varied from 2.3-11.0 mg/l, 3.1-4.2 mg/l, 3.6-9.0 mg/l, 4.5-10.0 mg/l, 6.0-11.5 mg/l, 6.5-12.5 mg/l, and 8.5-18.6 mg/l. The BOD value observed at all sampling stations during the study period was unacceptable, according to IS 2296:1992, Class A (2 mg/l), and it increased from S1 to S7. Except for S1 and S3, all sampling stations have a high dispersion of data points. BOD behaves symmetrically. For all sampling stations, the curve is platykurtic. The model outperforms the baseline model because stationary R squared and R squared values display identical behavior. Because the RMSE is minimal, the dependent series is close to the model predicted level for all sampling stations. Parmar and Bhardwaj (2015) [23] also discovered a similar value and pattern in the Yamuna River. According to the Ljung-Box

Q(18) model, statistics range from 12.42 to 41.92 for all sampling sites, significance ranges from 0.00 to 0.65, and degrees of freedom range from 15 to 16. The ARIMA (Winters' Multiplicative) model was fitted at S1, the ARIMA (Winters' Additive) model was fitted at S2, the ARIMA (Simple Seasonal) model was fitted at S4, the ARIMA (Winters' Additive) model was fitted at S5, the ARIMA (Simple Seasonal) model was fitted at S6, the ARIMA (Simple Seasonal) model was fitted at S7, and the boundary lines were at 95% confidence limits. ACF, PACF, lag, and time series explain that BOD values range from -3.42 to 3.23 for S1, 3.34 to 4.04 for S2, 5.31 to 7.35 for S3, 7.57 to 12.44 for S4, 9.01 to 14.86 for S5, 10.19 to 16.10 for S6, and 12.26 to 18.09 for S7 (Fig. 7) and that water quality is affected at all sampling stations for the predicted period of January 2018 to December 2022.





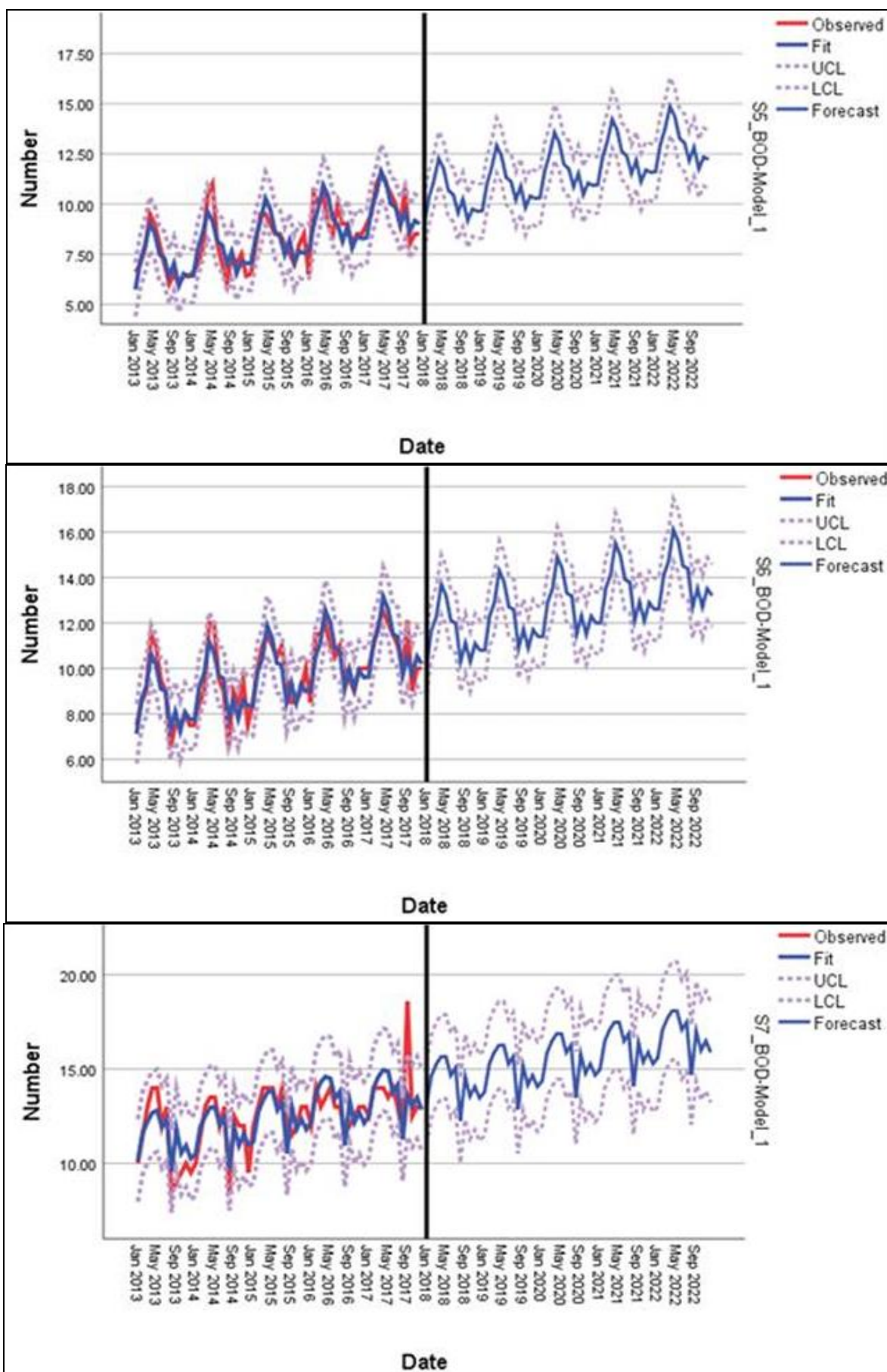
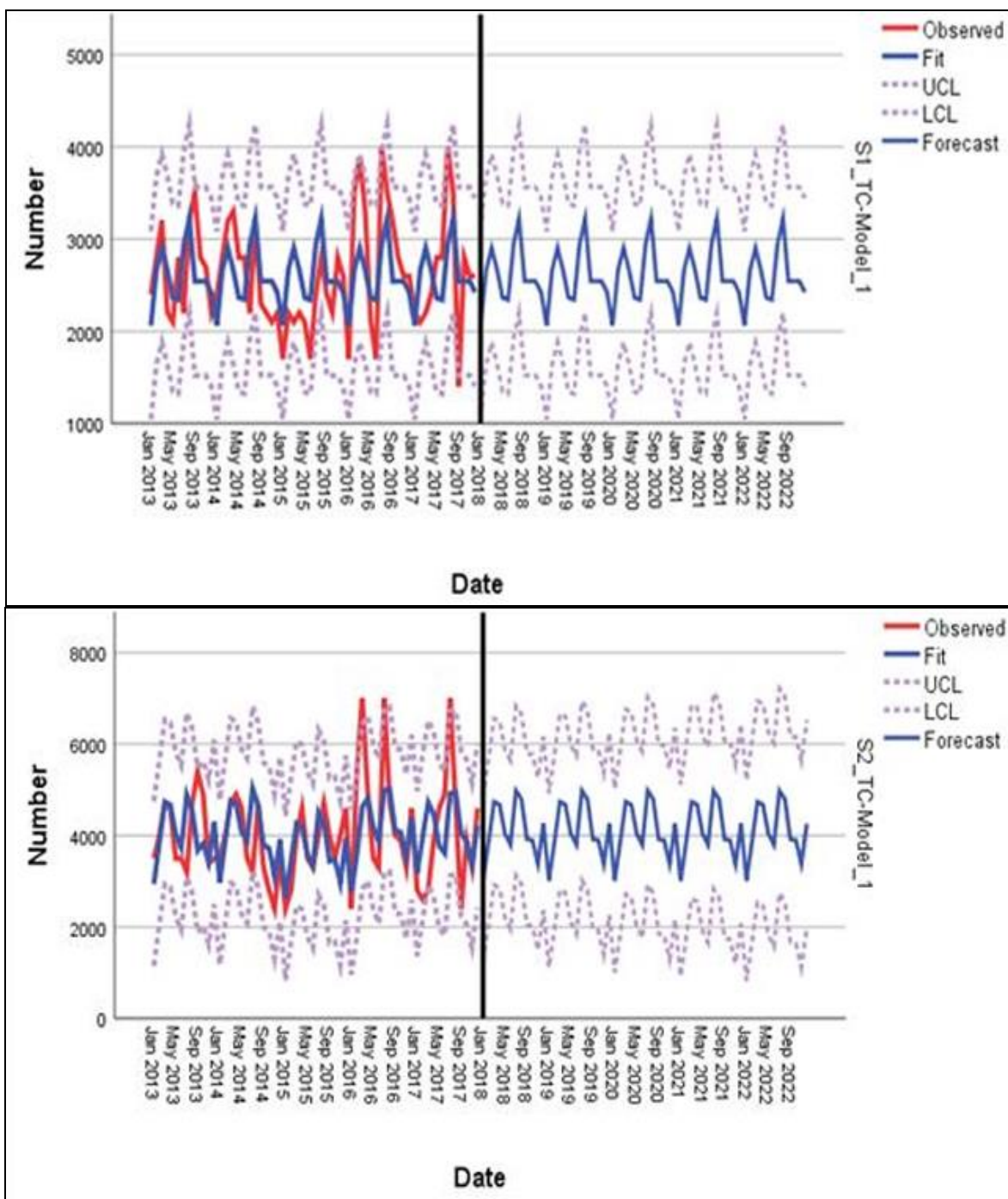


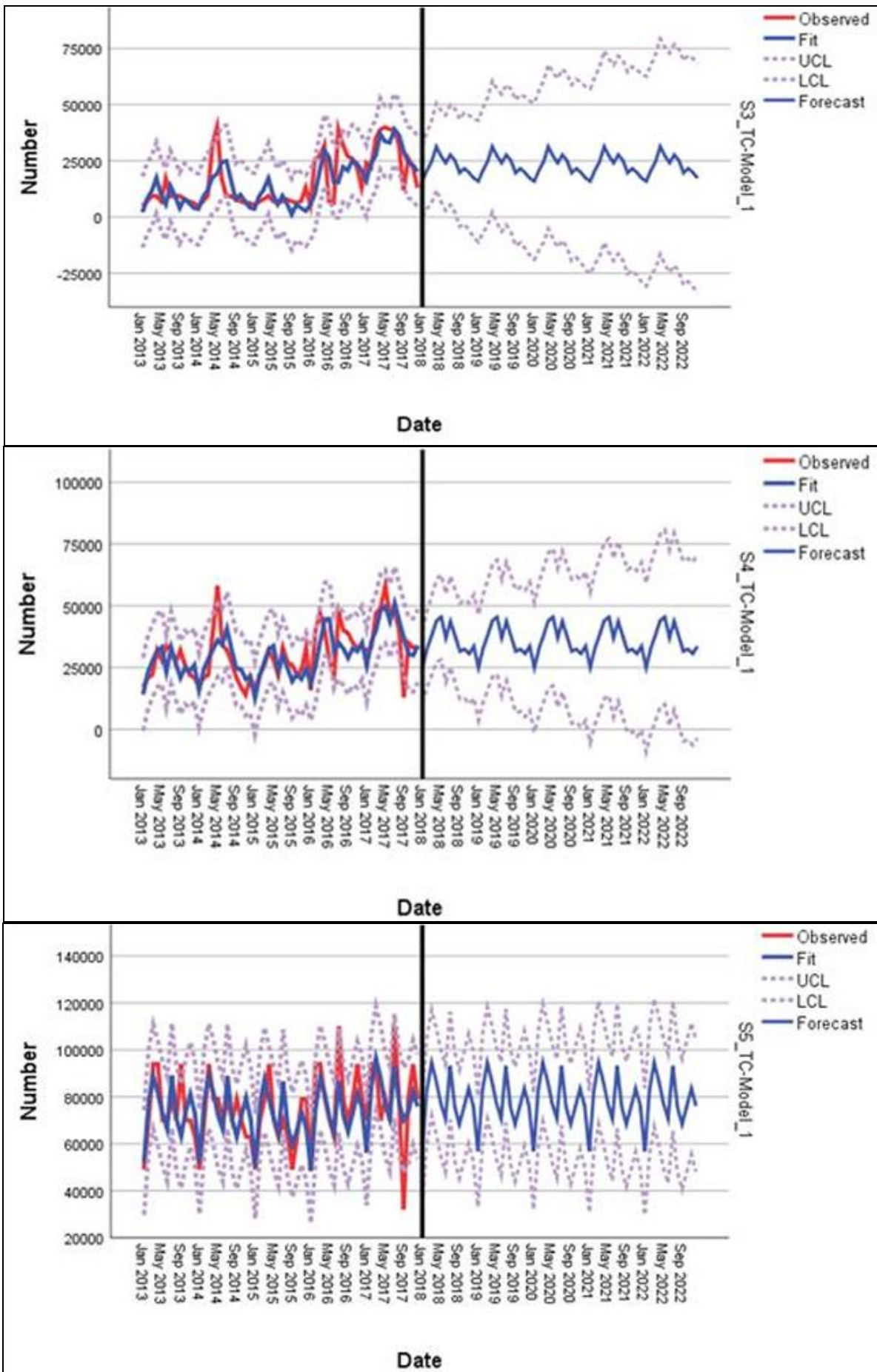
Fig 8 Representation of Trend and Time Series Analysis by Observed, Best Fit, LCL and UCL by Graphically of BOD at all Sampling Stations for the River Gomti

➤ *Total Coliform (TC)*

The mean, median, and mode for all sampling stations exceed the permitted limits of IS 2296:1992, Class A (50 MPN/100 ml), exhibits abnormal behavior, the standard deviation is large, and the spread is high and asymmetrical. At all sampling stations, the curve is platykurtic. TC values in river water ranged from 1400-4000 MPN/100 ml, 2400-7000 MPN/100 ml, 4900-40000 MPN/100 ml, 13000-58000 MPN/100 ml, 32000-110000 MPN/100 ml, 70000-130000 MPN/100 ml, and 94000-240000 MPN/100 ml at S1, S2, S3, S4, S5, S6, and S7, respectively, because stationary R squared and R squared value operate similarly, the model outperforms the baseline model. The RMSE values are incredibly high, which indicates that the dependent series is not even close to

reaching the level that the model anticipated. The value of statistics in the Ljung-Box Q(18) model is between 14.56 and 39.21, the significance level is between 0.00 and 0.56, and the degree of freedom is found between 15 and 16. All sampling stations except S6 have ARIMA (Simple Seasonal) models fitted, whereas S6 has ARIMA (Winters' Additive) models fitted and the boundary lines are at 95% confidence limits. Using plots of ACF, PACF, lag and time series, it is discovered that the value of TC ranges from 2064 to 3224 (S1), 3018 to 4978 (S2), 15908 to 31348 (S3), 25013 to 45413 (S4), 56755 to 94555 (S5), 89578 to 140387 (S6), and 150600 to 218600 (S7), as shown in Fig. 8, and the water quality is affected at all sampling stations for the predicted period of 5 years from January 2018 to December 2022.





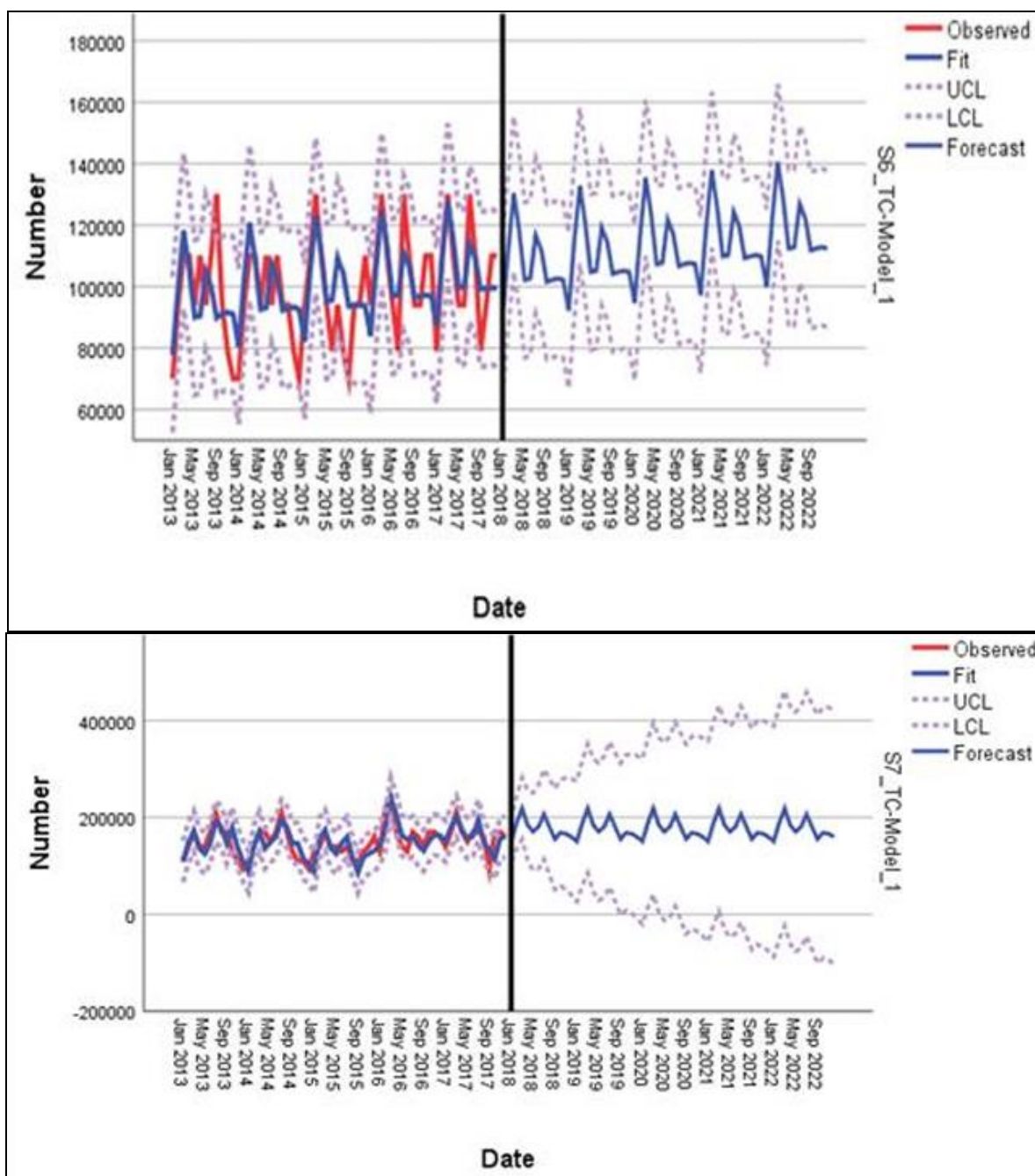


Fig 9 Representation of Trend and Time Series Analysis by Observed, Best Fit, LCL and UCL by Graphically of TC at all Sampling Stations for the River Gomti

IV. CONCLUSIONS

The different monitored water quality parameters (pH, TDS, TH, Chlorides, DO, BOD, and TC) for five years (2013-2017) at all seven selected sampling stations by UPPCB, Lucknow, India, have been used to conduct this study. Various statistical tests are used to analyse the model's performance. The ARIMA model was used to make projections regarding the monthly values of water quality parameters for the subsequent five years (2018–2022). The time series model outperforms the baseline model for all sampling stations because stationary R squared and R squared values showed similar behavior. The value of the RMSE for pH, DO, and BOD is comparably low, which indicates that

the dependent series is close to the model-predicted level, while high for TDS, TH, Chloride, and TC, indicating that the predictive model is useful at 95% confidence limits. As a result, the data does not need to be differentiated, which means that the series are independent of each other serially.

As the river approaches Lucknow's metropolitan area, the water quality metrics' observed values rise from upstream (S1) to downstream (S7), and they continue to do so throughout the anticipated period (2018- 2022). The water quality deteriorates across the board for all parameters at all test stations, except for pH, TDS, and chloride levels. Statistical modelling will be of enormous use to legislators, managers of water quality, and scientists as they develop the

strategies for managing the water quality of water body. The accuracy of the model is dependent on the user's ability to assess the water quality at various locations utilizing the data that is available to them.

ACKNOWLEDGMENT

The authors express their sincere gratitude to all individuals and organizations who contributed to the successful completion of this research study titled "Assessment and Forecasting of Water Quality Variables of the River Gomti, Lucknow, India."

We are especially thankful to the Uttar Pradesh Pollution Control Board (UPPCB), Lucknow, for providing the water quality monitoring data used in this study. Their continuous efforts in environmental monitoring and management have been invaluable for conducting this research.

Special appreciation is extended to all researchers, academicians, and environmental professionals whose published studies and scientific contributions provided useful references and guidance during the preparation of this manuscript.

➤ Conflict of Interest:

The authors declare that there is no conflict of interest regarding the publication of this manuscript.

➤ Funding:

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

- [1]. A. Krishan, R. K. Mishra, and A. Khursheed, "Assessment of water quality using water quality index: A case study of the River Gomti, Lucknow, Uttar Pradesh, India," *Urban Water Journal*, vol. 19, no. 5, pp. 520–530, 2022, doi: 10.1080/1573062X.2022.2032210.
- [2]. A. Krishan, A. Khursheed, and R. K. Mishra, "Evaluation of water quality using water quality index, synthetic pollution index, and GIS technique: A case study of the River Gomti, Lucknow, India," *Environmental Science and Pollution Research*, vol. 29, no. 54, pp. 81954–81969, 2022, doi: 10.1007/s11356-022-21493-3.
- [3]. D. P. Loucks and E. van Beek, "Water quality modeling and prediction," in *Water Resource Systems Planning and Management: An Introduction to Methods, Models, and Applications*. Cham, Switzerland: Springer International Publishing, 2017, pp. 417–467, doi: 10.1007/978-3-319-44234-1_10.
- [4]. A. Krishan, S. Yadav, and A. Srivastava, "Water pollution's global threat to public health: A mini-review," *International Journal of Scientific Research in Science, Engineering and Technology*, vol. 10, no. 6, pp. 321–334, 2023, doi: 10.32628/IJSRSET2310643.
- [5]. T. Abbasi and S. A. Abbasi, *Water Quality Indices*. Amsterdam, Netherlands: Elsevier, 2012, doi: 10.1016/C2010-0-69472-7.
- [6]. A. Krishan, R. K. Mishra, and D. Mohan, "Groundwater quality assessment: A case study of Manpur, India," *Journal of Environmental Nanotechnology*, vol. 7, no. 4, pp. 25–33, 2018, doi: 10.13074/jent.2018.12.184328.
- [7]. A. Krishan and R. K. Mishra, "Study on groundwater quality and status of Amas Block in Gaya District of Bihar with special reference to fluoride contamination," *Algerian Journal of Environmental Science and Technology*, vol. 6, no. 4, pp. 1631–1637, 2020.
- [8]. A. Krishan and A. Srivastava, "Recovery of nutrients from dairy wastewater by struvite crystallization," *International Journal of Engineering Research and General Science*, vol. 3, no. 5, pp. 591–597, 2015.
- [9]. A. Krishan and A. Srivastava, "Climate resilience: Safeguarding air and water quality," *International Research Journal of Environmental Sciences*, vol. 14, no. 3, pp. 77–85, 2025.
- [10]. S. Gupta and S. K. Gupta, "A critical review on water quality index tool: Genesis, evolution and future directions," *Ecological Informatics*, vol. 63, p. 101299, 2021, doi: 10.1016/j.ecoinf.2021.101299.
- [11]. J. Gao, G. Deng, H. Jiang, Y. Wen, S. Zhu, C. He, C. Shi, and Y. Cao, "Water quality pollution assessment and source apportionment of lake wetlands: A case study of Xianghai Lake in the Northeast China Plain," *Journal of Environmental Management*, vol. 344, p. 118398, 2023, doi: 10.1016/j.jenvman.2023.118398.
- [12]. H. Ding, X. Niu, D. Zhang, M. Lv, Y. Zhang, Z. Lin, and M. Fu, "Spatiotemporal analysis and prediction of water quality in Pearl River, China, using multivariate statistical techniques and data-driven model," *Environmental Science and Pollution Research*, vol. 30, pp. 63036–63051, 2023, doi: 10.1007/s11356-023-26209-9.
- [13]. A. Srivastava, A. Krishan, S. K. Yadav, and V. P. Gupta, "Sustainable business resilience: Navigating global environmental challenges for a greener future," in *Green Economy and Sustainable Development*, V. P. Gupta, A. K. Haghi, and S. Sharief, Eds. Cham, Switzerland: Springer, 2025, doi: 10.1007/978-3-031-92636-5_8.
- [14]. A. Srivastava, A. Krishan, and D. Srivastava, "Environmental regulations and compliance," in *Advances in Green Engineering for Sustainable Industrial Developments*, D. B. Tripathy, A. Gupta, and A. Ghosal, Eds. Elsevier, 2026, ch. 2, pp. 41–60, doi: 10.1016/B978-0-443-33637-9.00007-4.
- [15]. R. Khan, A. Saxena, S. Shukla, S. Sekar, V. Senapathi, and J. Wu, "Environmental contamination by heavy metals and associated human health risk assessment: A case study of surface water in Gomti River Basin, India," *Environmental Science and Pollution Research*, vol. 28, pp. 56105–56116, 2021, doi: 10.1007/s11356-021-14592-0.
- [16]. P. Goel, A. Saxena, D. S. Singh, and D. Verma, "Impact of rapid urbanization on water quality index

- in groundwater fed Gomati River, Lucknow, India,” *Current Science*, vol. 114, no. 3, pp. 650–654, 2018, doi: 10.18520/cs/v114/i03/650-654.
- [17]. K. Iqbal, S. Ahmad, and V. Dutta, “Pollution mapping in the urban segment of a tropical river: Is water quality index (WQI) enough for a nutrient-polluted river?” *Applied Water Science*, vol. 9, p. 197, 2019, doi: 10.1007/s13201-019-1083-9.
- [18]. R. Khan, A. Saxena, and S. Shukla, “Evaluation of heavy metal pollution for River Gomti, in parts of Ganga Alluvial Plain, India,” *SN Applied Sciences*, vol. 2, p. 1451, 2020, doi: 10.1007/s42452-020-03233-9.
- [19]. R. Khan, A. Saxena, S. Shukla, P. Goel, P. Bhattacharya, P. Li, E. F. Ali, and S. M. Shaheen, “Appraisal of water quality and ecological sensitivity with reference to riverfront development along the River Gomti, India,” *Applied Water Science*, vol. 12, p. 13, 2022, doi: 10.1007/s13201-021-01560-9.
- [20]. D. A. Attah and G. M. Bankole, “Time series analysis model for annual rainfall data in Lower Kaduna Catchment Kaduna, Nigeria,” *International Journal of Research in Chemistry and Environment*, vol. 2, no. 1, pp. 82–87, 2012.
- [21]. R. Seth, P. Singh, M. Mohan, R. Singh, and R. S. Aswal, “Monitoring of phenolic compounds and surfactants in water of Ganga Canal, Haridwar (India),” *Applied Water Science*, 2013, doi: 10.1007/s13201-013-0116-z.
- [22]. K. S. Parmar and R. Bhardwaj, “Water quality management using statistical analysis and time-series prediction model,” *Applied Water Science*, vol. 4, pp. 425–434, 2014, doi: 10.1007/s13201-014-0159-9.
- [23]. K. S. Parmar and R. Bhardwaj, “Statistical, time series, and fractal analysis of full stretch of River Yamuna (India) for water quality management,” *Environmental Science and Pollution Research*, vol. 22, pp. 397–414, 2015, doi: 10.1007/s11356-014-3346-1.
- [24]. P. Ravikumar, M. A. Mehmood, and R. K. Somashekar, “Water quality index to determine the surface water quality of Sankey Tank and Mallathahalli Lake, Bangalore Urban District, Karnataka, India,” *Applied Water Science*, vol. 3, pp. 247–261, 2013.
- [25]. G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*, 4th ed. West Sussex, U.K.: John Wiley & Sons, 2008.
- [26]. W. X. Lu, Y. Zhao, H. B. Chu, and L. L. Yang, “The analysis of groundwater levels influenced by dual factors in western Jilin Province by using time series analysis method,” *Applied Water Science*, 2013, doi: 10.1007/s13201-013-0111-4.
- [27]. M. Bora and D. C. Goswami, “Water quality assessment in terms of water quality index (WQI): Case study of the Kolong River, Assam, India,” *Applied Water Science*, vol. 7, pp. 3125–3135, 2017, doi: 10.1007/s13201-016-0451-y.