

Satellite-Based Forecasting of NO₂ and SO₂ Concentrations in Sri Lanka using SARIMA Time-Series Models

P. V. U. Kalpana¹

uden1722@gmail.com

University of Peradeniya

P. G. R. N. I. Pussella²

Sabaragamuwa University of Sri Lanka

J. Gunatilake³

University of Peradeniya

Research Article

Keywords: SARIMA, Sentinel-5P, TROPOMI, NO₂, SO₂, Sri Lanka, time-series forecasting, Google Earth Engine

Posted Date: October 7th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10874523/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Satellite-Based Forecasting of NO₂ and SO₂ Concentrations in Sri Lanka using SARIMA Time-Series Models

P. V. U. Kalpana¹, P. G. R. N. I. Pussella², J. Gunatilake³

¹ Postgraduate Institute of Science, University of Peradeniya, Peradeniya, Sri Lanka

² Department of Remote Sensing & GIS, Faculty of Geomatics, Sabaragamuwa University of Sri Lanka,
Belihuloya, Sri Lanka

³ Department of Geology, Faculty of Science, University of Peradeniya, Peradeniya, Sri Lanka

Corresponding author: udeni722@gmail.com

Abstract

Air pollution poses significant environmental and public health challenges in countries where continuous ground-based monitoring is limited. This study evaluates the use of Sentinel-5P TROPOMI observations for forecasting nitrogen dioxide (NO₂) and sulfur dioxide (SO₂) using Seasonal Autoregressive Integrated Moving Average (SARIMA) time-series models in five representative districts of Sri Lanka: Colombo, Kandy, Matara, Trincomalee, and Jaffna. Monthly observations from the Near Real-Time (NRTI) and Offline (OFFL) Sentinel-5P products were processed using Google Earth Engine for January 2019–December 2024. Sentinel-5P provides atmospheric column information rather than direct ground-level measurements; therefore, the study used an air-column-density formulation and unit conversion to express the satellite-derived quantities as concentration estimates in parts per billion (ppb). Negative retrieved values were masked before temporal aggregation. A two-phase forecasting framework was applied. In the first phase, observations from 2019–2022 were used for training and January 2023–December 2024 was retained for out-of-sample validation. In the second phase, the models were retrained using the complete 2019–2024 dataset and used to forecast 2025–2027. SARIMA models with a 12-month seasonal period were applied separately to each district, pollutant, and satellite product. Forecast performance was evaluated using Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). NRTI NO₂ MAPE ranged from 8.91% in Kandy to 19.69% in Colombo, while NRTI SO₂ produced substantially larger errors, reaching 124.92% in Kandy. OFFL NO₂ showed similarly low errors, whereas OFFL SO₂ remained more variable. The findings demonstrate the potential of satellite-derived time series combined with

seasonal statistical modelling for air-pollution forecasting, while highlighting the limitations of column-to-concentration conversion and the lower predictability of irregular SO₂ behaviour.

Keywords: SARIMA; Sentinel-5P; TROPOMI; NO₂; SO₂; Sri Lanka; time-series forecasting; Google Earth Engine

1 Introduction

Air pollution is a growing environmental and public-health concern, particularly in developing countries experiencing urban growth and industrial expansion (Longhurst et al. 2000). Nitrogen dioxide (NO₂) and sulfur dioxide (SO₂) are important atmospheric pollutants associated with combustion and industrial activities. NO₂ is mainly generated by combustion engines and industrial processes, whereas SO₂ is primarily associated with fossil-fuel combustion, power generation, and industrial activities (Delaram and Khedmati 2021).

In Sri Lanka, increasing traffic, urban development, and local industrial activity have raised concerns regarding air quality (Manawadu et al. 2009). The limited spatial coverage of ground-based monitoring makes continuous national-scale assessment difficult. Satellite remote sensing provides an opportunity to obtain spatially extensive atmospheric observations and to examine pollutant behaviour over areas not covered by monitoring stations.

The Sentinel-5 Precursor satellite carries the TROPOspheric Monitoring Instrument (TROPOMI), which provides observations of atmospheric trace gases including NO₂ and SO₂. TROPOMI has demonstrated the capability to monitor atmospheric trace-gas columns at high spatial and temporal resolution (Theys et al. 2016; Theys et al. 2019; Van Geffen et al. 2022). Google Earth Engine (GEE) facilitates processing of large satellite datasets and their temporal aggregation for environmental applications (Gorelick et al. 2017).

An important methodological distinction is that TROPOMI retrieves atmospheric column quantities rather than direct surface-level concentrations. Column number density represents the integrated amount of a gas through the atmospheric column and therefore differs from a fixed-point surface observation (Judd et al. 2020). This distinction is particularly important when satellite-derived values are expressed in ppb for time-series modelling.

Time-series models are widely used for air-pollution forecasting because they can represent temporal dependence and recurring seasonal behaviour. SARIMA extends ARIMA by incorporating seasonal

autoregressive, differencing, and moving-average components and has been applied to environmental pollutant forecasting (Bhatti et al. 2021; Doreswamy et al. 2020; Delaram and Khedmati 2021).

This study focuses specifically on satellite-based forecasting of NO₂ and SO₂ in five Sri Lankan districts. The objectives were to: (i) prepare monthly Sentinel-5P NO₂ and SO₂ time series for 2019–2024; (ii) develop and validate SARIMA models using 2019–2022 training data and 2023–2024 observations; (iii) compare NRTI and OFFL products; and (iv) retrain the models using 2019–2024 data to generate 2025–2027 forecasts.

2 Materials and Methods

This research used a step-by-step approach that integrates satellite monitoring of air pollution with statistical forecasting over time (Jia 2019). The main goal was to create models that could predict the monthly levels of nitrogen dioxide (NO₂) and sulfur dioxide (SO₂) in five selected districts of Sri Lanka. Data from the Sentinel-5P satellite was obtained through Google Earth Engine (GEE), and SARIMA models were created and assessed in Google Colab using Python. The process took place in two steps: first, a validation phase using data from 2023 to 2024, followed by a forward forecast phase estimating values for 2025 to 2027. The overall methodological framework, including data acquisition, processing, modelling, and forecasting phases, is illustrated in Fig. 1.

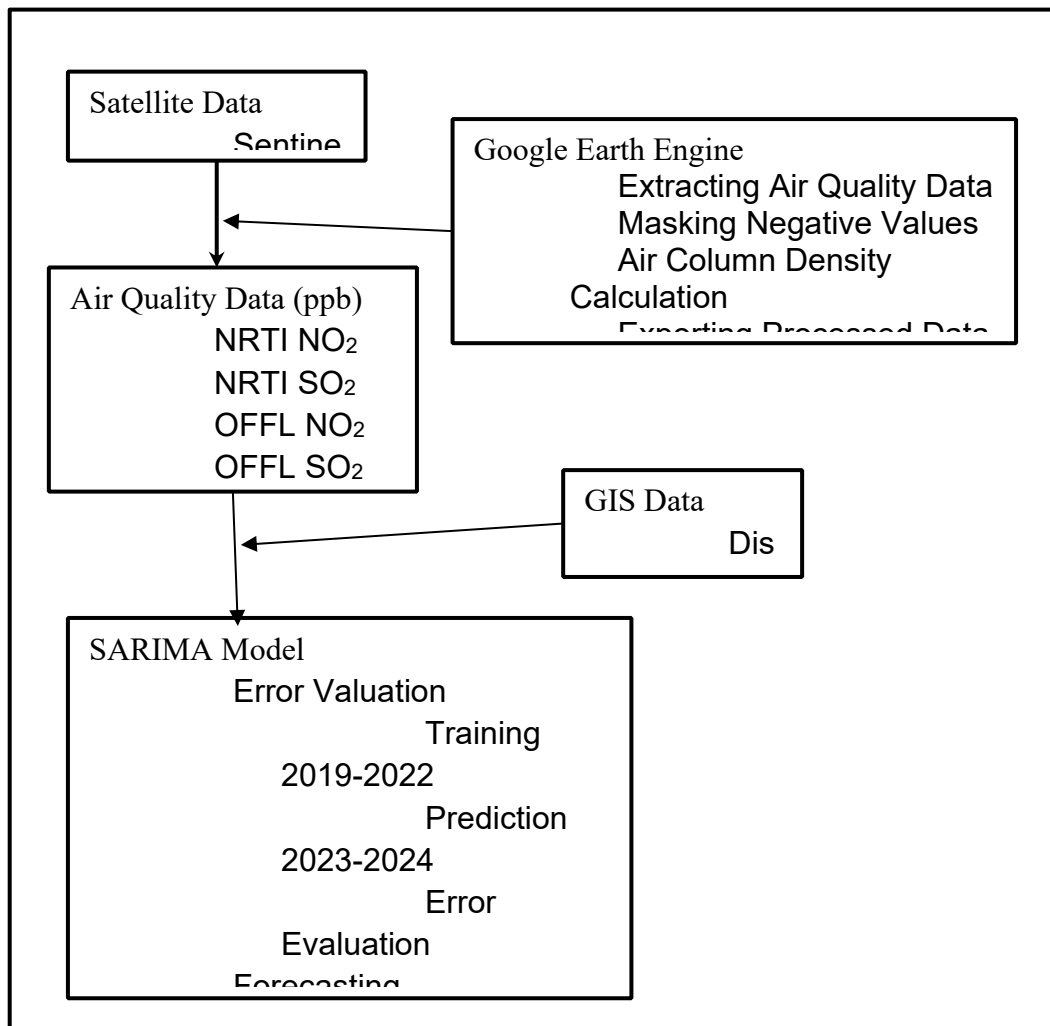


Fig. 1 Methodological Framework

2.1 Study area and district selection

Five districts—Colombo, Kandy, Matara, Trincomalee, and Jaffna—were selected to represent contrasting geographical and environmental settings. Colombo represents a densely urbanized environment; Kandy represents the Central Highlands; Matara represents a southern coastal setting; Trincomalee represents an eastern area with industrial and power-generation activities; and Jaffna represents the northern dry-zone environment. This selection provides a range of settings for examining district-level temporal behaviour (Sokhi et al. 2022).

2.2 Sentinel-5P data

Monthly NO₂ and SO₂ observations were obtained from Sentinel-5P TROPOMI through Google Earth Engine. Four datasets were prepared: NRTI NO₂, NRTI SO₂, OFFL NO₂, and OFFL SO₂. The monthly series cover January 2019 to December 2024. The NRTI and OFFL products were retained

separately because they represent complementary processing streams with different processing timeliness and levels of post-processing (Van Geffen et al. 2022; Oxoli et al. 2020).

The satellite retrievals represent column number density. TROPOMI retrieval algorithms use spectroscopic information and atmospheric radiative-transfer assumptions to estimate vertical columns; they do not directly measure pollutant concentration at the ground surface (Judd et al. 2020; Theys et al. 2016).

2.3 Satellite data preprocessing

Satellite data were processed in Google Earth Engine. The study region was spatially filtered and monthly observations were generated from the available satellite observations. Negative retrieved values were masked before subsequent processing because negative pollutant column values are not physically interpretable as concentrations. The resulting observations were temporally aggregated to monthly values for each district.

The use of GEE allowed the spatial filtering and temporal aggregation of Sentinel-5P observations before exporting the resulting datasets for statistical modelling (Gorelick et al. 2017).

2.4 Conversion of column number density to concentration estimates

Because Sentinel-5P provides atmospheric column number density rather than direct surface concentration, the study applied an air-column-density formulation to express the retrieved pollutant quantity as a ppb concentration estimate. This conversion is an approximation and does not constitute a direct retrieval of near-surface concentration.

Atmospheric pressure at elevation h was calculated using the barometric relationship:

$$P(h) = P_0 \exp(-h/H_s)$$

where $P_0 = 101325$ Pa is the reference sea-level pressure and $H_s = 8500$ m is the atmospheric scale height used in the thesis methodology.

The altitude-adjusted air-column density was calculated from the ideal-gas relationship:

$$N_{air}(h) = [P(h)/(R \cdot T)] \times [N_{air}(0)/(P_0/(R \cdot T))]$$

where $R = 8.314$ J mol⁻¹ K⁻¹, $T = 298$ K, and $N_{air}(0) = 42230$ mol m⁻² is the reference sea-level air-column density documented in the thesis. The resulting pollutant concentration estimate was calculated as:

$$C_{ppb} = [N_{pollutant} / N_{air}(h)] \times 10^9$$

The formulation and parameter values are based on the satellite-processing methodology documented in the MSc thesis (Judd et al. 2020; Oxoli et al. 2020; Theys et al. 2019). The resulting ppb values are therefore treated as satellite-derived column-based concentration estimates. They should not be interpreted as direct ground-level measurements because vertical pollutant distribution, boundary-layer structure, and meteorological conditions are not explicitly resolved in this conversion.

2.5 Missing observations and preprocessing

The monthly series were inspected for missing observations before modelling. The original analysis used interpolation/statistical imputation to maintain continuous monthly series and also applied Z-score-based outlier detection with month-wise median replacement. The Z-score was calculated as:

$$Z = (x - \mu) / \sigma$$

where x is an observation, μ is the series mean, and σ is the standard deviation. Because the original analysis did not retain a complete numerical table of missing-observation counts, missing percentages, or maximum consecutive gaps, these quantities should be calculated directly from the original monthly datasets and reported alongside the revised manuscript. No missing-data percentage is fabricated here.

2.6 SARIMA formulation

SARIMA was selected because the datasets contain monthly observations and the thesis identified recurring seasonal behaviour in the pollutant series. The general form is SARIMA(p,d,q)(P,D,Q)_s, where p , d , and q denote the non-seasonal autoregressive, differencing, and moving-average orders; P , D , and Q denote the seasonal orders; and s is the seasonal period.

Because the observations are monthly, the seasonal period was set to $s = 12$. The implemented model configuration was SARIMA(1,1,1)(1,1,1)₁₂. The thesis states that first-order differencing was used to obtain stationarity and that the selected parameters were based on empirical testing and recommendations from previous environmental time-series studies (Bunnag 2024; Bhatti et al. 2021).

2.7 Stationarity, ACF/PACF and parameter identification

To address reproducibility, the revised workflow explicitly distinguishes the required model-identification steps. Stationarity should be evaluated before final fitting using the Augmented Dickey–Fuller (ADF) test, with test statistics and p-values reported for the original and differenced series. ACF and PACF plots should be examined to identify candidate autoregressive and moving-average structures. Candidate SARIMA specifications should then be compared using information criteria such as AIC and BIC, followed by residual diagnostics.

The available thesis documents the final SARIMA(1,1,1)(1,1,1)₁₂ configuration and first-order differencing, but it does not contain numerical ADF results or a complete candidate-model comparison table. Therefore, these numerical diagnostics must be generated from the original modelling datasets before journal submission. This statement is included to maintain methodological transparency and avoid introducing unsupported statistical results.

2.8 Model training and validation

The forecasting procedure comprised two phases. In Phase 1, monthly observations from January 2019 to December 2022 were used to train the SARIMA models. The fitted models were then used to generate forecasts for January 2023 to December 2024. These forecasts were compared against the observed values to provide an out-of-sample assessment of model performance.

In Phase 2, the models were retrained using the complete January 2019–December 2024 series. The retrained models were then used to generate monthly forecasts for January 2025–December 2027. This two-stage framework follows the forecasting procedure documented in the thesis and manuscript.

2.9 Forecast accuracy metrics

Four performance measures were used: MAE, MSE, RMSE, and MAPE.

$$MAE = (1/n) \sum |A_i - F_i|$$

$$MSE = (1/n) \sum (A_i - F_i)^2$$

$$RMSE = \sqrt{MSE}$$

$$MAPE = (100/n) \sum |(A_i - F_i)/A_i|$$

where A_i is the observed value and F_i is the forecast value. MAPE was interpreted cautiously because very small observed concentrations can produce disproportionately large percentage errors.

3 Results

The comparison between observed and predicted NO₂ and SO₂ concentrations for each district is illustrated in Figures 2–6. These figures demonstrate that SARIMA models effectively capture the seasonal variability and general trends in NO₂ concentrations across all study areas, with closer alignment observed in districts such as Kandy and Matara compared to Colombo and Jaffna.

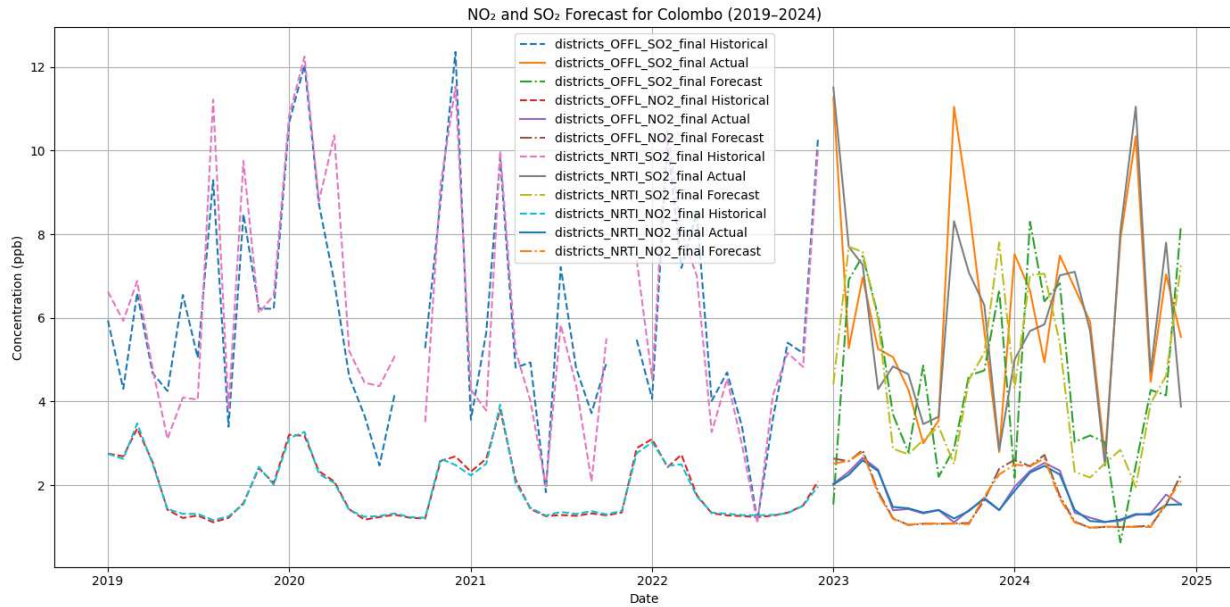


Fig. 2 Comparison of historical observations and SARIMA forecasts of monthly NO₂ concentrations in Colombo (2019–2024)

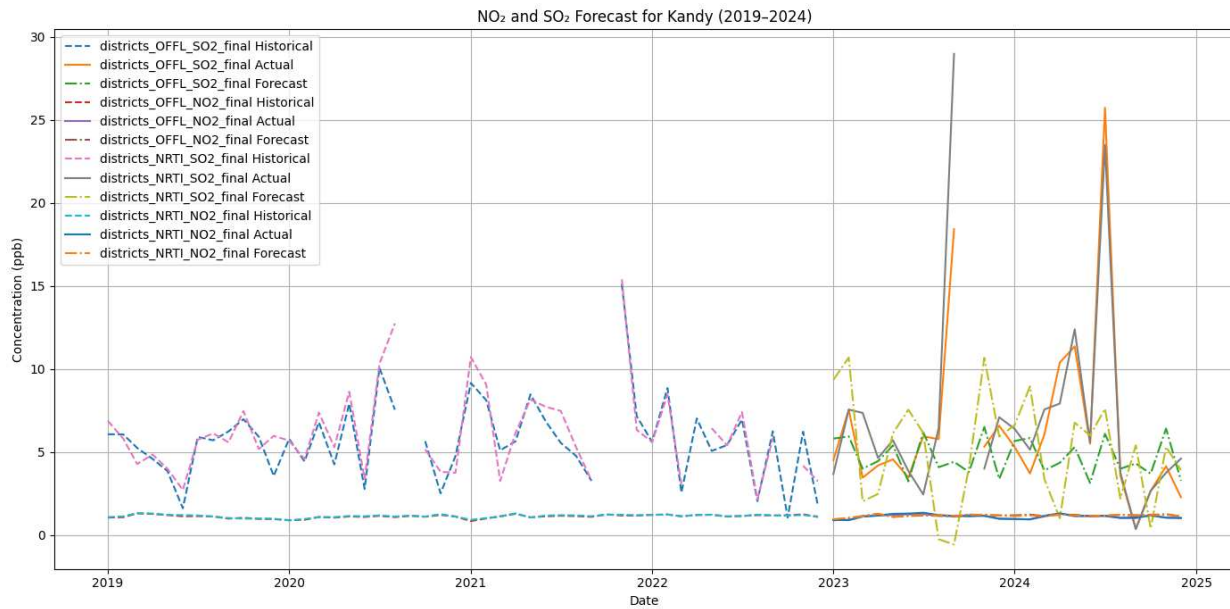


Fig. 3 Comparison of historical observations and SARIMA forecasts of monthly NO_2 concentrations in Kandy (2019–2024)

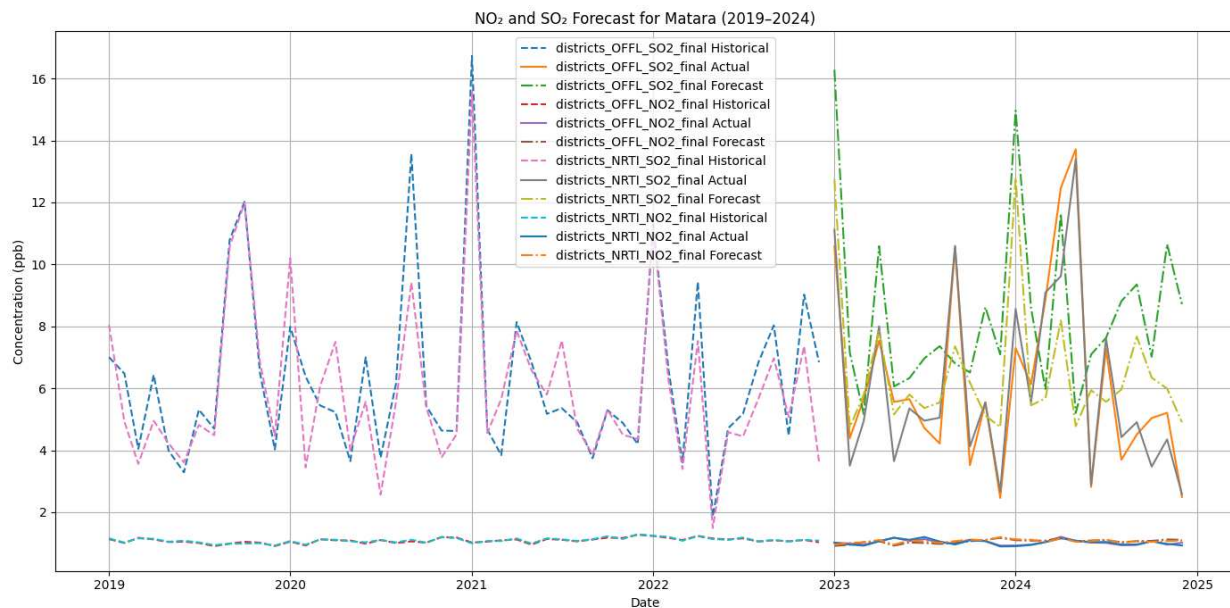


Fig. 4 Comparison of historical observations and SARIMA forecasts of monthly NO_2 concentrations in Matara (2019–2024)

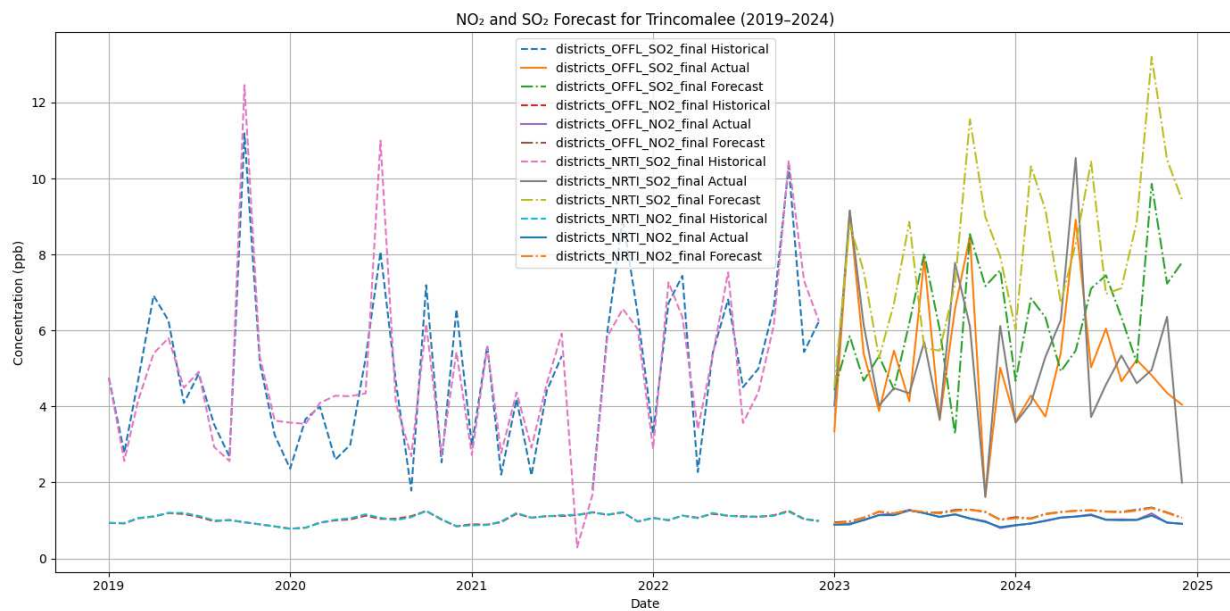


Fig. 5 Comparison of historical observations and SARIMA forecasts of monthly NO_2 concentrations in Trincomalee (2019–2024)

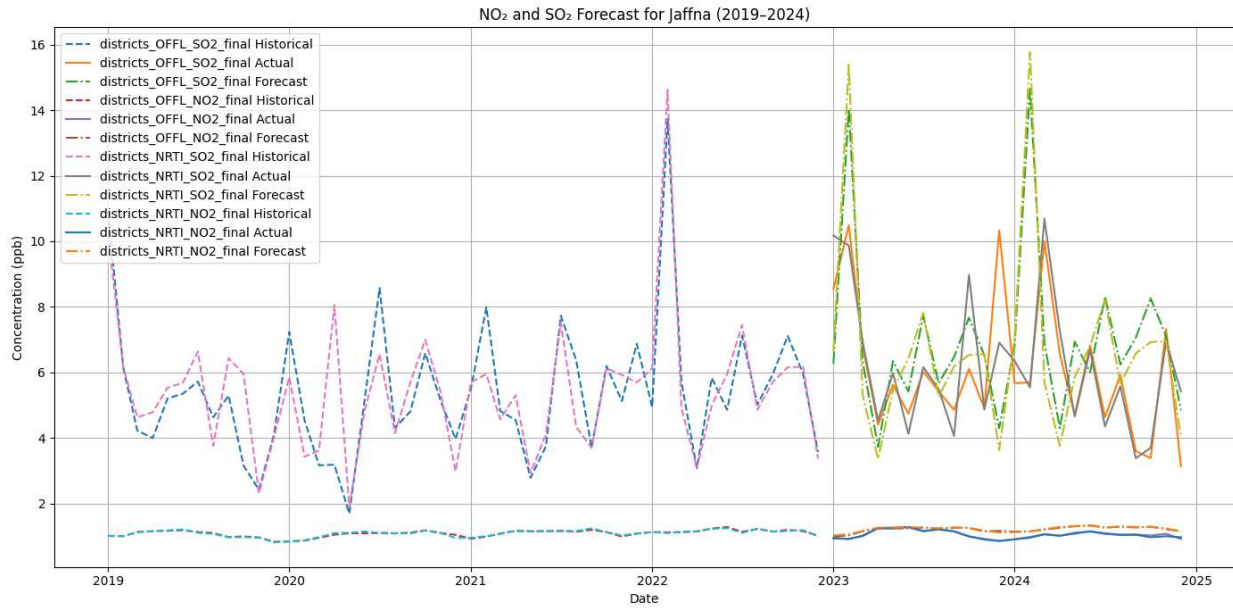


Fig. 6 Comparison of historical observations and SARIMA forecasts of monthly NO₂ concentrations in Jaffna (2019–2024)

3.1 NRTI NO₂ forecasting performance

Table 1 SARIMA model performance metrics for NO₂ forecasting in selected districts using NRTI Sentinel-5P data (2019–2024)

District	MAE	MSE	RMSE	MAPE (%)
Colombo	0.3181	0.1403	0.3746	19.6937
Kandy	0.0956	0.0143	0.1196	8.9072
Matara	0.0891	0.0137	0.1169	9.0131
Trincomalee	0.1396	0.0250	0.1583	13.9793
Jaffna	0.1721	0.0375	0.1936	17.3003

As shown in Table 1, the NRTI NO₂ models showed comparatively low forecasting errors across all five districts. MAPE ranged from 8.91% in Kandy to 19.69% in Colombo. Kandy and Matara produced the lowest errors, while Colombo produced the highest MAPE.

3.2 NRTI SO₂ forecasting performance

Table 2 SARIMA model performance metrics for SO₂ forecasting in selected districts using NRTI Sentinel-5P data (2019–2024)

District	MAE	MSE	RMSE	MAPE (%)
Colombo	2.6201	12.5382	3.5409	41.9278
Kandy	5.0599	63.7617	7.9851	124.9223
Matara	2.0218	7.1976	2.6828	37.6831
Trincomalee	3.2347	16.5647	4.0700	89.0368
Jaffna	2.4427	10.8764	3.2979	42.2485

As shown in Table 2, the NRTI SO₂ models showed substantially greater forecasting errors than the NO₂ models. Kandy had the highest MAPE at 124.92%, whereas Matara had the lowest MAPE at 37.68%. The larger percentage errors indicate weaker predictability of the SO₂ series within the univariate seasonal framework.

3.3 OFFL NO₂ forecasting performance

Table 3 SARIMA model performance metrics for NO₂ forecasting in selected districts using OFFL Sentinel-5P data (2019–2024)

District	MAE	MSE	RMSE	MAPE (%)
Colombo	0.3284	0.1617	0.4022	20.1479
Kandy	0.0934	0.0149	0.1221	8.8895
Matara	0.0877	0.0124	0.1112	8.7099
Trincomalee	0.1439	0.0269	0.1640	14.4685
Jaffna	0.1606	0.0333	0.1826	16.0194

As shown in Table 3, OFFL NO₂ produced similarly low errors. MAPE ranged from 8.71% in Matara to 20.15% in Colombo. The close agreement between NRTI and OFFL NO₂ performance suggests that the recurring temporal structure of NO₂ was reasonably well represented in both products.

3.4 OFFL SO₂ forecasting performance

Table 4 SARIMA model performance metrics for SO₂ forecasting in selected districts using OFFL Sentinel-5P data (2019–2024)

District	MAE	MSE	RMSE	MAPE (%)
Colombo	3.0206	16.3126	4.0389	45.3069
Kandy	3.1525	31.0001	5.5678	78.1335
Matara	3.4834	16.9105	4.1122	71.4716
Trincomalee	2.1121	6.5603	2.5613	53.0861
Jaffna	2.2493	9.2056	3.0341	40.8022

As shown in Table 4, OFFL SO₂ also showed considerably larger errors than NO₂. MAPE ranged from 40.80% in Jaffna to 78.13% in Kandy. The pattern is consistent with the greater temporal irregularity of SO₂ and the sensitivity of MAPE to low observed values.

3.5 2025–2027 forecasts

After validation, the models were retrained using the complete 2019–2024 dataset and used to generate monthly forecasts for 2025–2027. The forecasts preserve the seasonal structure learned from the historical data. NO₂ forecasts show comparatively consistent recurring behaviour, whereas SO₂ forecasts show greater variability. The forecasts should be interpreted as statistical projections conditional on the continuation of historical temporal patterns rather than as deterministic predictions of future emissions.

The projected trends for NO₂ and SO₂ concentrations from 2025 to 2027 for each district are presented in Figures 7–11. These figures illustrate the continuation of seasonal patterns for NO₂ and the higher uncertainty associated with SO₂ forecasts, particularly in districts influenced by irregular emission sources.

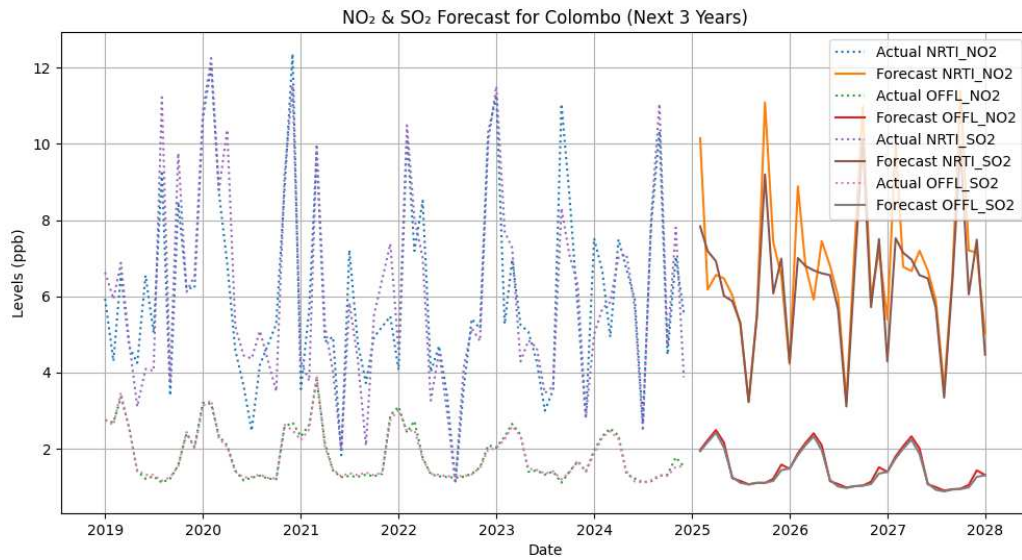


Fig. 7 SARIMA forecasts of NO₂ and SO₂ concentrations in Colombo from 2025 to 2027, based on models trained with NRTI and OFFL Sentinel-5P data (2019–2024)

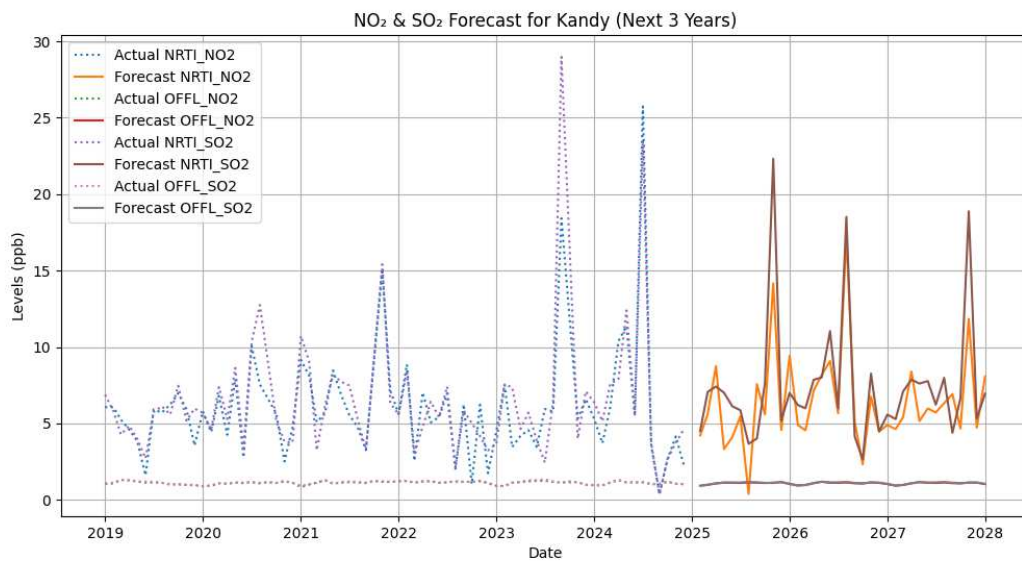


Fig. 8 SARIMA forecasts of NO₂ and SO₂ concentrations in Kandy from 2025 to 2027, based on models trained with NRTI and OFFL Sentinel-5P data (2019–2024)

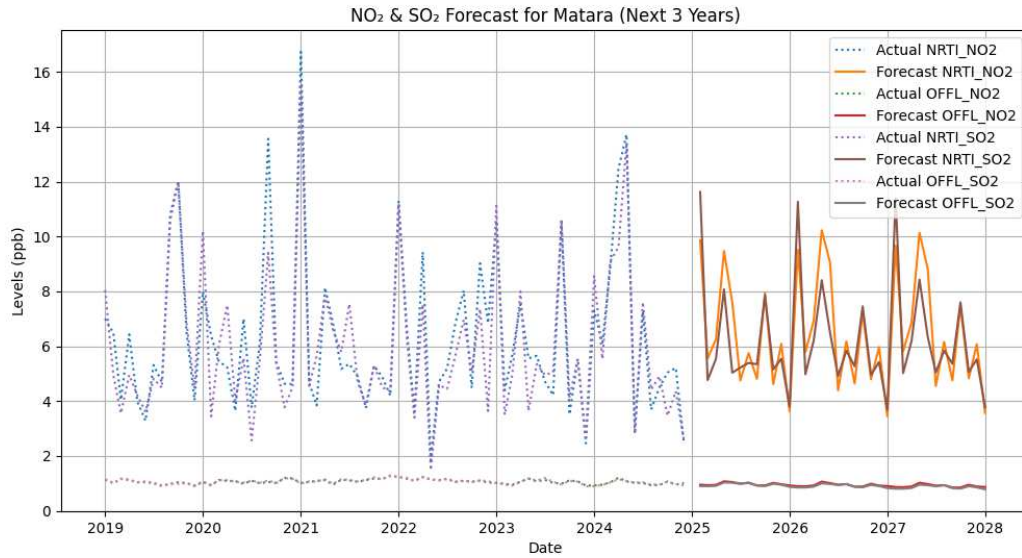


Fig. 9 SARIMA forecasts of NO₂ and SO₂ concentrations in Matara from 2025 to 2027, based on models trained with NRTI and OFFL Sentinel-5P data (2019–2024)

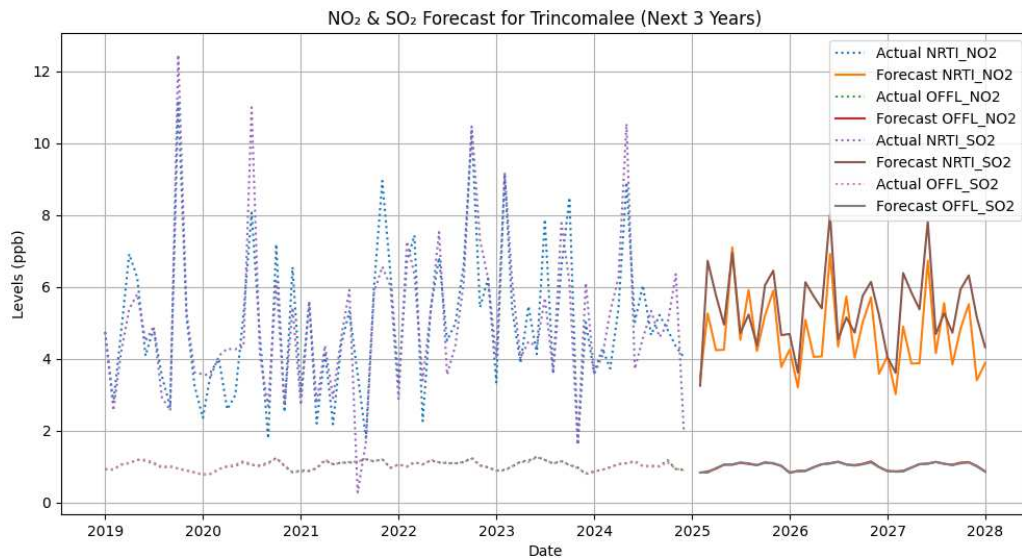


Fig. 10 SARIMA forecasts of NO₂ and SO₂ concentrations in Trincomalee from 2025 to 2027, based on models trained with NRTI and OFFL Sentinel-5P data (2019–2024)

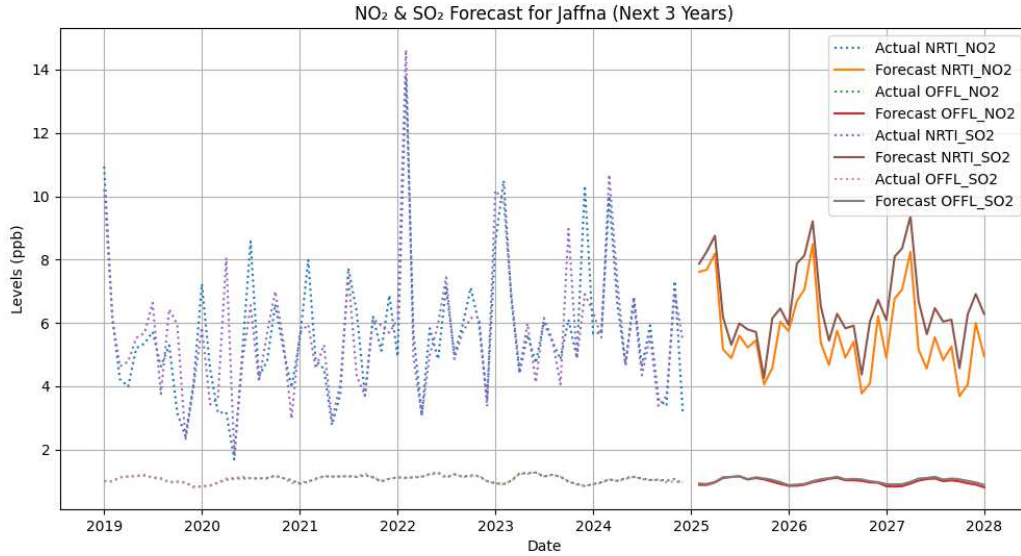


Fig. 11 SARIMA forecasts of NO₂ and SO₂ concentrations in Jaffna from 2025 to 2027, based on models trained with NRTI and OFFL Sentinel-5P data (2019–2024)

4 Discussion

4.1 SARIMA performance for NO₂

The comparatively low errors for NO₂ indicate that the monthly NO₂ series contained temporal dependence and recurring patterns that could be represented by the selected seasonal model. The results are consistent with previous applications of SARIMA to air-pollution time series, where seasonal autoregressive and moving-average components can capture recurring temporal structure (Bhatti et al. 2021; Doreswamy et al. 2020).

The best NO₂ performance occurred in Kandy and Matara, while Colombo showed the largest NO₂ MAPE. This district-level variation indicates that model performance depends not only on the statistical model but also on the temporal characteristics of the underlying satellite-derived series.

4.2 SO₂ forecasting uncertainty

SO₂ forecasts were substantially less accurate than NO₂ forecasts. The thesis identifies the greater variability of SO₂ as a consequence of irregular industrial activity, unexpected environmental changes, and meteorological influences. Such episodic behaviour is difficult to capture using a univariate model that relies primarily on the historical pollutant series.

MAPE should be interpreted with particular caution for SO₂. When observed concentrations approach zero, a small absolute error can correspond to a very large percentage error. Consequently, MAE and RMSE provide important complementary information when assessing SO₂ forecasting performance.

4.3 NRTI versus OFFL products

The NRTI and OFFL NO₂ results were broadly comparable. Kandy, for example, produced MAPE values of 8.91% and 8.89% for NRTI and OFFL, respectively. The similarity indicates that the dominant temporal signal in NO₂ was retained in both processing streams. For SO₂, both products produced substantially larger errors, indicating that product selection alone does not resolve the underlying temporal unpredictability.

4.4 Interpretation of satellite-derived concentration estimates

A central methodological consideration is the distinction between satellite column measurements and surface observations. Sentinel-5P TROPOMI retrieves vertical atmospheric columns, while ground stations measure near-surface conditions. Differences can arise from vertical pollutant distributions, atmospheric mixing, spatial averaging, viewing geometry, and the timing of observations (Judd et al. 2020; Van Geffen et al. 2022; Oxoli et al. 2020).

The present study converts column number density to ppb using an altitude-adjusted total air-column formulation. This approach provides a consistent concentration scale for constructing the time series, but it does not explicitly account for boundary-layer height or the detailed vertical distribution of the pollutant. The resulting values are therefore best described as satellite-derived concentration estimates rather than direct ground-level concentrations.

The thesis also notes differences in measurement time and spatial representation between Sentinel-5P and ground monitoring. Sentinel-5P provides observations at the satellite overpass time and represents a spatially averaged atmospheric column, whereas ground stations provide continuous or frequent point-based measurements. These differences are important when interpreting satellite-derived values and should be acknowledged in forecasting applications.

4.5 Missing-data and interpolation considerations

Interpolation can help maintain a continuous monthly series required by SARIMA, but it can also alter temporal structure when gaps are long or systematically distributed. The revised manuscript therefore explicitly identifies missing-data reporting as a required quality-control step. The original thesis documents interpolation as part of the preprocessing workflow, but it does not provide a complete numerical summary of missing percentages and gap lengths. Such statistics should be generated from the original datasets before submission, together with a sensitivity check where feasible.

4.6 Implications and future improvements

The results indicate that satellite-derived time series can support pollutant forecasting where ground-monitoring coverage is limited. However, the performance difference between NO₂ and SO₂ demonstrates that a univariate seasonal model has limitations for pollutants strongly affected by irregular events.

Future studies should investigate SARIMAX or hybrid models that incorporate meteorological variables such as temperature, rainfall, wind speed, and boundary-layer height, as well as relevant emission indicators. Hybrid SARIMA–machine-learning approaches may also improve the modelling of nonlinear residual behaviour (Liu et al. 2022; Doreswamy et al. 2020).

5 Limitations

- Sentinel-5P observations represent atmospheric column quantities rather than direct surface-level measurements.
- The column-to-ppb conversion relies on specified atmospheric assumptions and does not explicitly resolve the vertical pollutant profile or boundary-layer height.
- The original analysis used interpolation for missing observations, but a complete numerical summary of missing-data proportions and gap lengths was not retained in the thesis.
- The final SARIMA configuration is documented as SARIMA(1,1,1)(1,1,1)₁₂, but the thesis does not contain numerical ADF results or a complete candidate-model comparison.
- The univariate SARIMA framework does not explicitly include meteorological or emission-related explanatory variables.
- Long-range forecasts assume that historical temporal relationships remain sufficiently representative of future conditions.

6 Conclusions

This study evaluated the use of Sentinel-5P-derived NO₂ and SO₂ time series for seasonal forecasting in five districts of Sri Lanka using SARIMA models. Monthly NRTI and OFFL observations from January 2019 to December 2024 were processed using Google Earth Engine and converted to satellite-derived concentration estimates using an altitude-adjusted air-column formulation.

The out-of-sample validation results showed that NO₂ was substantially more predictable than SO₂. NRTI NO₂ MAPE ranged from 8.91% to 19.69%, while OFFL NO₂ MAPE ranged from 8.71% to 20.15%. SO₂ produced considerably larger errors, with NRTI MAPE reaching 124.92% in Kandy and OFFL MAPE reaching 78.13% in Kandy.

The findings demonstrate that SARIMA can represent recurring seasonal behaviour in monthly satellite-derived NO₂ series. However, the weaker SO₂ performance shows the limitations of relying only on historical pollutant values when emissions are irregular or strongly affected by external environmental conditions.

The study also emphasizes that satellite-derived ppb estimates should not be interpreted as direct ground-level measurements. The conversion used here provides a consistent basis for time-series analysis, but future work should incorporate boundary-layer information, meteorological variables, and ground observations to improve the relationship between satellite columns and surface concentrations.

Overall, Sentinel-5P observations combined with seasonal time-series modelling provide a potentially useful forecasting framework for data-limited air-quality monitoring. More rigorous model-identification diagnostics, transparent missing-data reporting, improved satellite-to-surface conversion, and multivariate or hybrid forecasting models are recommended for future research.

Author Contributions

Conceptualization: P. V. U. K., P. G. R. N. I. P., J. G. Methodology: P. V. U. K., P. G. R. N. I. P., J. G. Formal analysis and investigation: P. V. U. K. Data curation: P. V. U. K. Writing – original draft preparation: P. V. U. K. Writing – review and editing: P. V. U. K., P. G. R. N. I. P., J. G. Supervision: P. G. R. N. I. P., J. G.

Funding

No funding was received to assist with the preparation of this manuscript.

Competing Interests

The authors declare that they have no competing interests.

Ethics Declaration

Not applicable.

Consent to Participate

Not applicable.

Consent to Publish

Not applicable.

Data Availability Statement

The Sentinel-5P TROPOMI datasets analysed during the current study are publicly available from the Copernicus Open Access Hub via Google Earth Engine (<https://developers.google.com/earth-engine/datasets/catalog/sentinel-5p>). The remaining datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

Acknowledgements

We acknowledge the European Space Agency for Sentinel-5P TROPOMI data access through Google Earth Engine. We thank Prof. P. G. R. N. I. Pussella and Dr. J. Gunatilake for their supervision and guidance. We also acknowledge support from the Postgraduate Institute of Science, University of Peradeniya.

References

- Bhatti UA, Yan Y, Zhou M, Ali S, Hussain A, Qingsong H, Yu Z, Yuan L (2021) Time series analysis and forecasting of air pollution particulate matter (PM_{2.5}): an SARIMA and factor analysis approach. *IEEE Access* 9:41019–41031. <https://doi.org/10.1109/ACCESS.2021.3060744>
- Bunnag T (2024) Forecasting PM₁₀ caused by Bangkok's leading greenhouse gas emission using the SARIMA and SARIMA-GARCH model. *Int J Energy Econ Policy* 14(1):418–426. <https://doi.org/10.32479/ijeep.15275>
- Delaram J, Khedmati M (2021) Forecasting ambient air pollutants by box-Jenkins stochastic models in Tehran. *Sci Iran* 28(6E):3551–3568. <https://doi.org/10.24200/sci.2020.52893.2937>

- Doreswamy, Harish Kumar KS, Gad I (2020) Time series analysis for prediction of PM_{2.5} using seasonal autoregressive integrated moving average (SARIMA) model on Taiwan air quality monitoring network data. *J Comput Theor Nanosci* 17(9):3964–3969. <https://doi.org/10.1166/jctn.2020.8997>
- Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, Moore R (2017) Google Earth Engine: planetary-scale geospatial analysis for everyone. *Remote Sens Environ* 202:18–27. <https://doi.org/10.1016/j.rse.2017.06.031>
- Jia Q (2019) Urban air quality assessment method based on GIS technology. *Appl Ecol Environ Res* 17(4):9367–9375. https://doi.org/10.15666/aecer/1704_93679375
- Judd LM, Al-Saadi JA, Szykman JJ, Valin LC, Janz SJ, Kowalewski MG, Eskes HJ, Veefkind JP, Cede A, Mueller M, Gebetsberger M, Swap R, Pierce RB, Nowlan CR, González Abad G, Nehrir A, Williams D (2020) Evaluating Sentinel-5P TROPOMI tropospheric NO₂ column densities with airborne and Pandora spectrometers near New York City and Long Island Sound. *Atmos Meas Tech* 13(11):6113–6140. <https://doi.org/10.5194/amt-13-6113-2020>
- Longhurst JWS, Brebbia CA, Power H (2000) Air pollution VIII. WIT Press, Southampton.
- Manawadu L, Wijsekera S, Manawadu L (2009) Understanding air pollution dynamics and exploring reasons for aggravation of air pollution in the city of Colombo using geostatistical techniques. ResearchGate. <https://www.researchgate.net/publication/272686901>
- Oxoli D, Ceden Jimenez JR, Brovelli MA (2020) Assessment of Sentinel-5P performance for ground-level air quality monitoring: preparatory experiments over the Covid-19 lockdown period. *Int Arch Photogramm Remote Sens Spatial Inf Sci XLIV-3-W1-2020*:111–116. <https://doi.org/10.5194/isprs-archives-XLIV-3-W1-2020-111-2020>
- Sokhi RS, Moussiopoulos N, Baklanov A, Bartzis J, Coll I, Finardi S, Friedrich R, Geels C, Grönholm T, Halenka T, Ketzel M, Maragkidou A, Matthias V, Moldanova J, Ntziachristos L, Schäfer K, Suppan P, Tsegas G, Carmichael G, Franco V, Hanna S, Jalkanen JP, Velders GJM, Kukkonen J (2022) Advances in air quality research – current and emerging challenges. *Atmos Chem Phys* 22:4615–4703. <https://doi.org/10.5194/acp-22-4615-2022>
- Theys N, De Smedt I, Yu H, Danckaert T, van Gent J, Hörmann C, Wagner T, Hedelt P, Bauer H, Romahn F, Pedernana M, Loyola D, Van Roozendaal M (2016) Sulfur dioxide retrievals from TROPOMI onboard Sentinel-5 Precursor: algorithm theoretical basis. <https://doi.org/10.5194/amt-2016-309>

- Theys N, Hedelt P, De Smedt I, Lerot C, Yu H, Vlietinck J, Pedergrana M, Arellano S, Galle B, Fernandez D, Carlito CJM, Barrington C, Taisne B, Delgado-Granados H, Loyola D, Van Roozendaal M (2019) Global monitoring of volcanic SO₂ degassing with unprecedented resolution from TROPOMI onboard Sentinel-5 Precursor. *Sci Rep* 9(1). <https://doi.org/10.1038/s41598-019-39279-y>
- Van Geffen J, Eskes H, Compernelle S, Pinardi G, Verhoelst T, Lambert JC, Snee M, Ter Linden M, Ludewig A, Boersma KF, Veefkind JP (2022) Sentinel-5P TROPOMI NO₂ retrieval: impact of version v2.2 improvements and comparisons with OMI and ground-based data. *Atmos Meas Tech* 15(7):2037–2060. <https://doi.org/10.5194/amt-15-2037-2022>
- Virghileanu M, et al. (2020) Nitrogen dioxide (NO₂) pollution monitoring with Sentinel-5P satellite imagery over Europe during the coronavirus pandemic outbreak. *Remote Sensing* 12(21):1–29. <https://doi.org/10.3390/rs12213575>
- Wang Y, et al. (2022) Monitoring of urban ecological environment including air quality using satellite imagery. *PLoS ONE* 17(8). <https://doi.org/10.1371/journal.pone.0266759>
- Liu F, et al. (2022) A new method for inferring city emissions and lifetimes of nitrogen oxides from high-resolution nitrogen dioxide observations: a model study. *Atmos Chem Phys* 22(2):1333–1349. <https://doi.org/10.5194/acp-22-1333-2022>