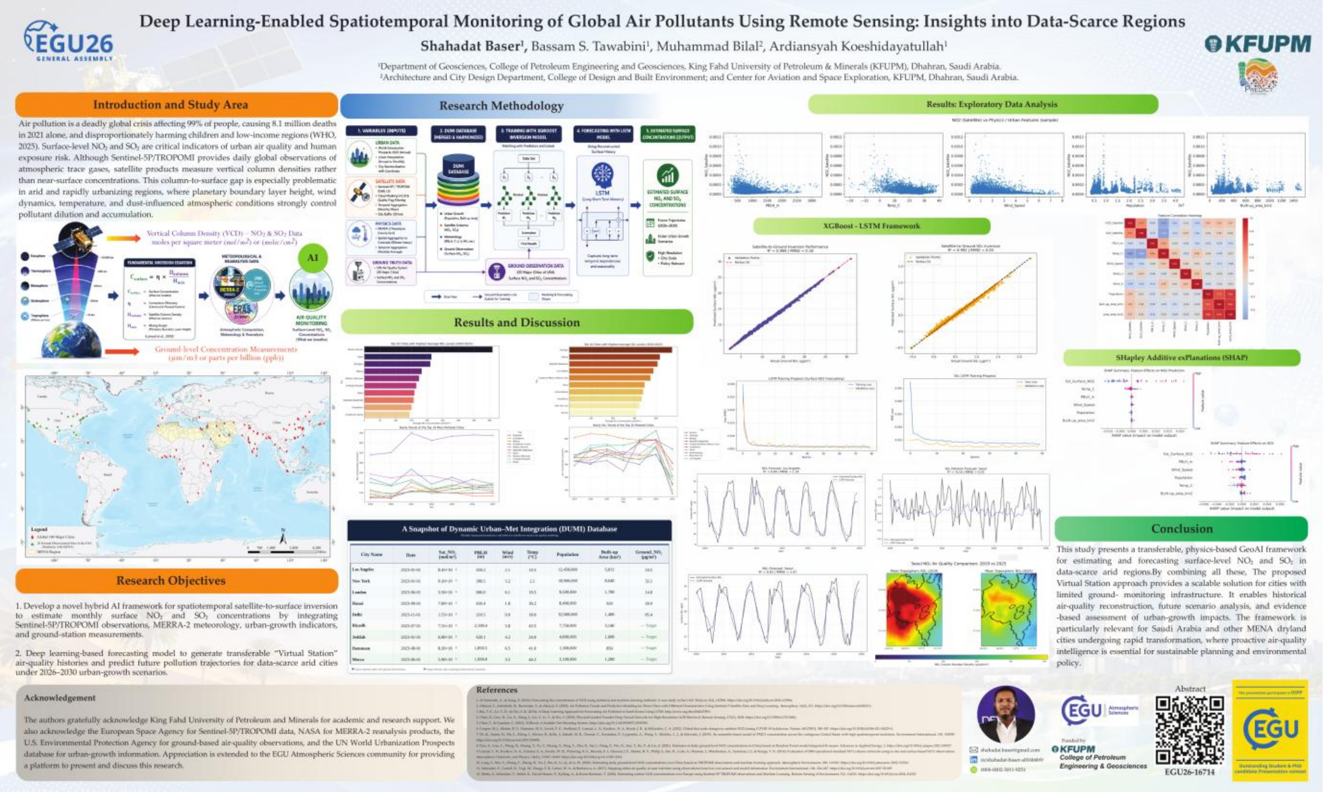
Supplementary Scientific Materials

EGU General Assembly 2026 | Abstract EGU26-16714 | Atmospheric Sciences Division

Deep Learning-Enabled Spatiotemporal Monitoring of Global Air Pollutants Using Remote Sensing: Insights into Data-Scarce Regions

Shahadat Baser, *Bassam S. Tawabini, Muhammad Bilal, Ardiansyah Koeshidayatullah*

King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia.



Deep Learning-Enabled Spatiotemporal Monitoring of Global Air Pollutants Using Remote Sensing: Insights into Data-Scarce Regions

Shahadat Baser[1], Bassam S. Tawabini[1], Muhammad Bilal[2], Ardiansyah Koeshidayatullah[1]

[1]Department of Geosciences, College of Petroleum Engineering and Geosciences, King Fahd University of Petroleum & Minerals (KFUPM), Dhahran, Saudi Arabia.
[2]Architecture and City Design Department, College of Design and Built Environment; and Center for Aviation and Space Exploration, KFUPM, Dhahran, Saudi Arabia.

Introduction and Study Area

Air pollution is a deadly global crisis affecting 99% of people, causing 8.1 million deaths in 2021 alone, and disproportionately harming children and low-income regions (WHO, 2025). Surface-level NO_2 and SO_2 are critical indicators of urban air quality and human exposure risk. Although Sentinel-5P/TROPOMI provides daily global observations of atmospheric trace gases, satellite products measure vertical column densities rather than near-surface concentrations. This column-to-surface gap is especially problematic in arid and rapidly urbanizing regions, where planetary boundary layer height, wind dynamics, temperature, and dust-influenced atmospheric conditions strongly control pollutant dilution and accumulation.

Research Objectives

1. Develop a novel hybrid AI framework for spatiotemporal satellite-to-surface inversion to estimate monthly surface NO_2 and SO_2 concentrations by integrating Sentinel-5P/TROPOMI observations, MERRA-2 meteorology, urban-growth indicators, and ground-station measurements.

2. Deep learning-based forecasting model to generate transferable "Virtual Station" air-quality histories and predict future pollution trajectories for data-scarce arid cities under 2026–2030 urban-growth scenarios.

Research Methodology

Results: Exploratory Data Analysis

XGBoost - LSTM Framework

Results and Discussion

A Snapshot of Dynamic Urban-Met Integration (DUMI) Database

SHapley Additive exPlanations (SHAP)

Conclusion

This study presents a transferable, physics-based GeoAI framework for estimating and forecasting surface-level NO_2 and SO_2 in data-scarce arid regions.By combining all these, The proposed Virtual Station approach provides a scalable solution for cities with limited ground- monitoring infrastructure. It enables historical air-quality reconstruction, future scenario analysis, and evidence -based assessment of urban-growth impacts. The framework is particularly relevant for Saudi Arabia and other MENA dryland cities undergoing rapid transformation, where proactive air-quality intelligence is essential for sustainable planning and environmental policy.

Acknowledgement

The authors gratefully acknowledge King Fahd University of Petroleum and Minerals for academic and research support. We also acknowledge the European Space Agency for Sentinel-5P/TROPOMI data, NASA for MERRA-2 reanalysis products, the U.S. Environmental Protection Agency for ground-based air-quality observations, and the UN World Urbanization Prospects database for urban-growth information. Appreciation is extended to the EGU Atmospheric Sciences community for providing a platform to present and discuss this research.

References

Funded by KFUPM College of Petroleum Engineering & Geosciences

Abstract

EGU26-16714

Poster preview used as the visual anchor for this file.

This supplementary material strengthens the poster by supplying methodological traceability, analytical validation, uncertainty interpretation, and a concise scientific conclusion. It is not a replacement for the poster; it is designed to support judges and technical viewers who want more detail.

1. Executive Scientific Summary

Problem: Surface-level NO2 and SO2 are critical for exposure assessment, but satellite missions retrieve vertical column densities. In arid and rapidly urbanizing environments, column-to-surface translation is strongly modulated by planetary boundary layer height, wind, temperature, and urban emissions.

Approach: The study builds the Dynamic Urban-Met Integration (DUMI) database by fusing Sentinel-5P/TROPOMI trace-gas columns, MERRA-2 meteorology, UN WUP 2025 urban-growth indicators, and U.S. EPA ground observations. XGBoost learns the nonlinear satellite-to-surface inversion, and LSTM forecasts future pollution trajectories.

Evidence: The inversion model captures monthly ground-station variability with R2 = 0.998 for NO2 and R2 = 0.992 for SO2. SHAP confirms physically interpretable relationships, especially the inverse PBLH-surface concentration relation. LSTM tests in Los Angeles and Seoul show R2 = 0.84 and 0.82, respectively.

Impact: The framework creates scalable virtual monitoring stations for data-scarce dryland cities. This enables historical baseline reconstruction and policy-relevant forecasting under 2026-2030 urban-growth scenarios.

Scientific contribution: The work is strongest where it demonstrates a physically constrained bridge between satellite columns and surface exposure, rather than treating satellite columns as direct ground-level concentrations.

2. Scientific Problem and Objectives

Scientific problem. Sentinel-5P/TROPOMI offers daily synoptic coverage of atmospheric trace gases, but the retrieved quantity is a column density integrated through the atmosphere. Public-health assessment requires near-surface concentration. This disconnect becomes severe in arid regions where boundary-layer depth can change rapidly and wind, temperature, and dust conditions control dilution and accumulation.

Data-scarcity problem. Many MENA cities, including cities in Saudi Arabia, lack dense, open, long-term surface monitoring networks. This prevents direct local training of supervised models and motivates domain adaptation from data-rich source cities to data-poor target cities.

Objectives. The work has two poster-level objectives: (1) construct a physics-based satellite-to-surface inversion framework for monthly NO2 and SO2, and (2) generate transferable virtual station histories and 2026-2030 forecasts for data-scarce arid cities using LSTM-based spatiotemporal modeling.

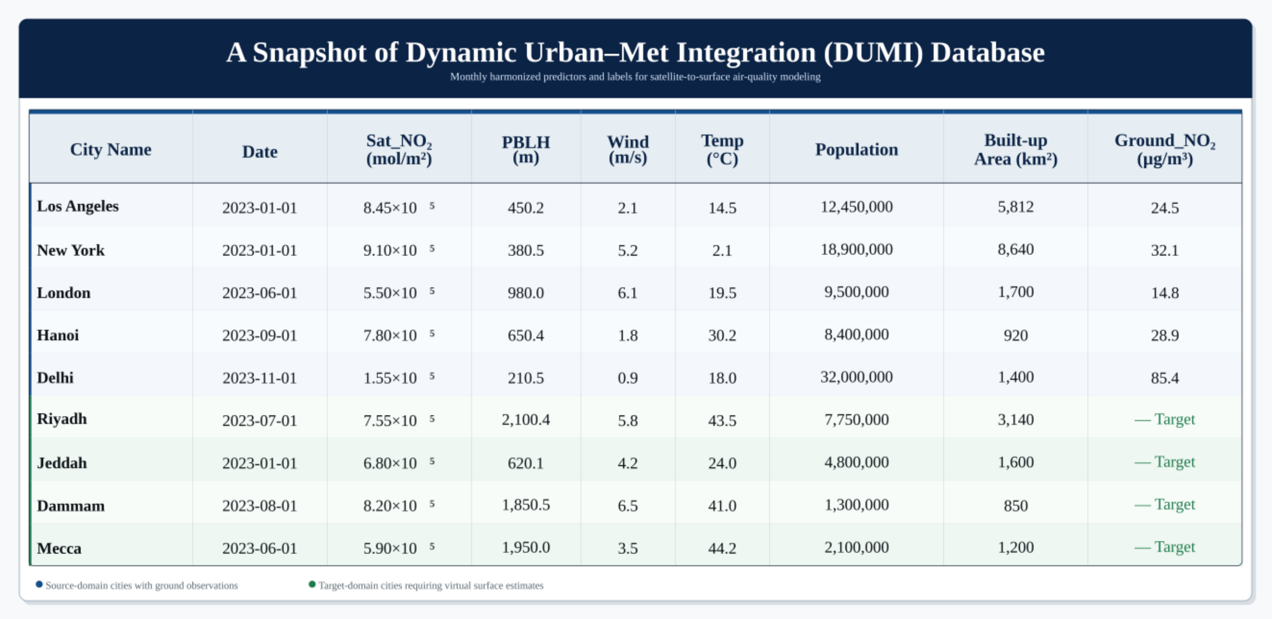
3. DUMI Data Architecture

DUMI is the modelling backbone of the poster. It is designed to reduce data heterogeneity by aligning satellite trace-gas columns, reanalysis meteorology, urban growth indicators and source-domain ground observations at a monthly city scale.

Table 1. Distinct data collection for column to surface measurement.

Category	**Variables**	**Source**	**Period**	**Resolution**	**Role in model**
Satellite	NO_2 and SO_2 vertical columns	Sentinel-5P/TROPOMI	2019-2025	Monthly city mean	Primary predictor
Meteorology	PBLH, wind U/V, temperature, humidity	*NASA MERRA-2*	2019-2025	Monthly city mean	Physical constraints
Urban growth	Population, built-up area	*UN WUP 2025*	2019-2025; projected 2026-2030	Annual to monthly	Anthropogenic drivers
Ground truth	Surface NO_2 and SO_2	*U.S. EPA AQS*	2019-2025	Monthly mean	Training labels

The source domain consists of 20 data-rich U.S. cities with available EPA surface observations. The target domain includes data-scarce arid and semi-arid regions, with emphasis on Saudi cities such as Riyadh, Jeddah, Dammam, Mecca, and Medina. The broader DUMI database covers 100 global cities to expose the model to diversity.



A Snapshot of Dynamic Urban–Met Integration (DUMI) Database

Monthly harmonized predictors and labels for satellite-to-surface air-quality modeling

City Name	Date	Sat_NO_2 (mol/m²)	PBLH (m)	Wind (m/s)	Temp (°C)	Population	Built-up Area (km²)	Ground_NO_2 (μg/m³)
Los Angeles	2023-01-01	8.45×10^{5}	450.2	2.1	14.5	12,450,000	5,812	24.5
New York	2023-01-01	9.10×10^{5}	380.5	5.2	2.1	18,900,000	8,640	32.1
London	2023-06-01	5.50×10^{5}	980.0	6.1	19.5	9,500,000	1,700	14.8
Hanoi	2023-09-01	7.80×10^{5}	650.4	1.8	30.2	8,400,000	920	28.9
Delhi	2023-11-01	1.55×10^{5}	210.5	0.9	18.0	32,000,000	1,400	85.4
Riyadh	2023-07-01	7.55×10^{5}	2,100.4	5.8	43.5	7,750,000	3,140	— Target
Jeddah	2023-01-01	6.80×10^{5}	620.1	4.2	24.0	4,800,000	1,600	— Target
Dammam	2023-08-01	8.20×10^{5}	1,850.5	6.5	41.0	1,300,000	850	— Target
Mecca	2023-06-01	5.90×10^{5}	1,950.0	3.5	44.2	2,100,000	1,200	— Target

● Source-domain cities with ground observations

● Target-domain cities requiring virtual surface estimates

Figure S1. DUMI snapshot showing source-domain records and target-domain examples requiring inferred surface concentration estimates.

4. Methodological Workflow

The workflow is deliberately cascaded. Ground observations do not directly fill the target cities; instead, they provide labels for the XGBoost inversion model. The trained inversion model reconstructs synthetic surface histories, and the LSTM uses those histories to forecast future trajectories.

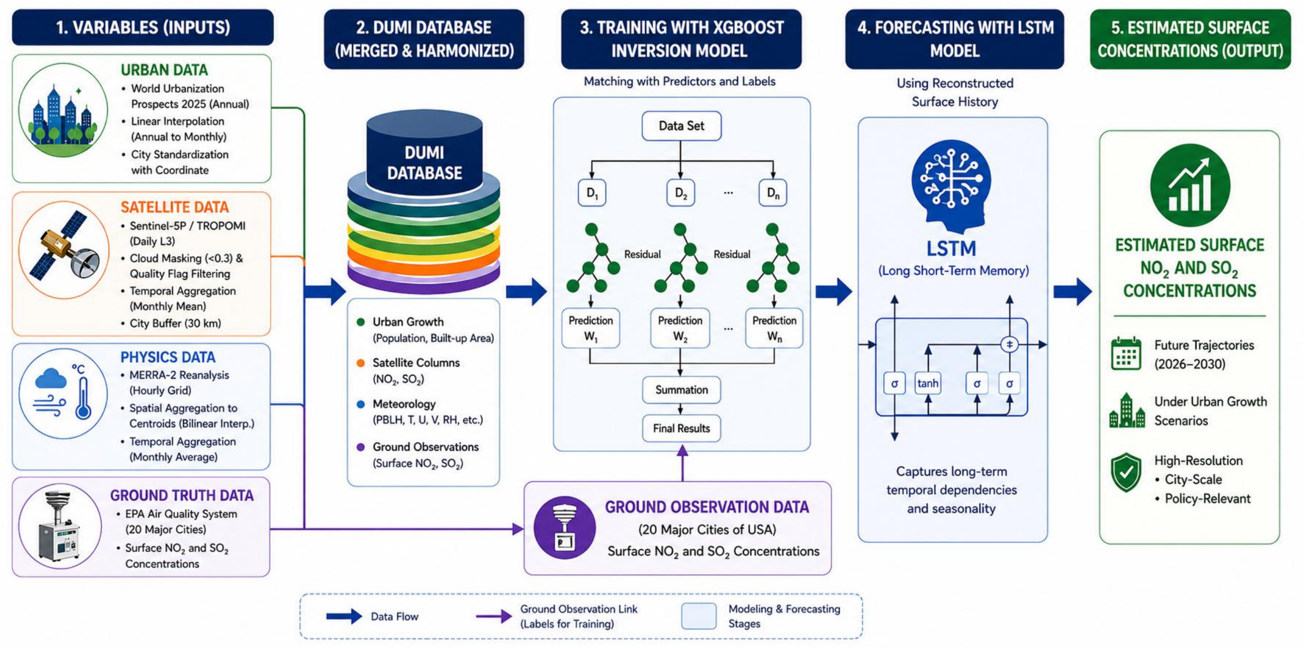


Figure S2. Simplified methodological workflow. Ground observation data are connected directly to the XGBoost inversion model as training labels; only LSTM is used for forecasting.

XGBoost inversion model acquired satellite column density into the monthly surface concentration. Inputs: satellite NO_2/SO_2 columns, PBLH, wind, temperature, and related meteorological constraints. Labels: EPA ground observations from 20 U.S. cities.

Appropriateness: tree ensembles capture threshold behavior such as stagnant low-wind periods and shallow boundary-layer accumulation.

LSTM forecasting model future surface pollution trajectories from reconstructed histories. Input sequence: 12-month sliding window to capture annual seasonality. Forecasting horizon: 2026-2030 under urban-growth scenarios.

Appropriateness: LSTM memory cells retain long-term temporal dependencies while representing seasonal volatility.

Model Implementation Notes

- XGBoost is calibrated on source-domain cities and applied to target-domain cities through homogeneous domain adaptation.
- The inversion target is surface NO2 or SO2 concentration; satellite column density alone is not sufficient without meteorological constraints.

- LSTM uses reconstructed surface histories as synthetic observations, combined with meteorological and urban-growth predictors.
- Scenario forecasting isolates the effect of urban growth by using projected population and built-up area while representing meteorology with baseline climatology.

5.1 Physics-based Inversion Model

The physical motivation is that satellite instruments observe a vertical column, whereas exposure is governed by surface concentration. A simplified mixed-layer relationship is $C_{surface}$ proportional to N_{column} / H_{mix}, but real atmospheric conditions require a non-linear model because mixing efficiency, wind dispersion and temperature-dependent chemistry vary across cities and seasons.

The inversion target is expressed as: $C_{surface} = f(N_{column}, PBLH, T, U, V) + epsilon$. XGBoost is appropriate because it handles tabular predictors, non-linear thresholds, missing values and interactions such as shallow boundary layer plus weak wind conditions.

XGBoost component	Configuration used	Analytical rationale
Predictors	NO2/SO2 satellite column, PBLH, temperature, wind *U*, and wind *V*.	Represents trace-gas load, mixing depth, thermodynamic state, and ventilation.
Training labels	U.S. EPA AQS surface NO_2/SO_2 monthly concentrations from source-domain cities.	Provides physically observed target values for satellite-to-surface mapping.
Core hyperparameters	n_estimators=1000, learning_rate=0.05, max_depth=6, tree_method=hist, objective=reg:squarederror.	Balances non-linear capacity with regularization against overfitting.
Transfer strategy	Homogeneous domain adaptation from data-rich source cities to data-scarce arid target cities.	Assumes the learned physical mixing/dispersion relationship remains transferable when the feature space overlaps.

5.2 LSTM Forecasting Model

Forecasting is treated as a temporal problem after the surface history is reconstructed. The LSTM uses a 12-month sliding window so that one full annual cycle is available at each prediction step. The feature vector includes reconstructed surface concentration, temperature, PBLH, wind speed, population and built-up area.

LSTM element	Implementation	Reason
Input sequence	12 months	Captures annual seasonality and meteorological recurrence.
Feature dimension	6 features: estimated surface pollutant, temperature, PBLH, wind speed, population, built-up area.	Combines pollution memory, atmospheric state and urban forcing.
Architecture	Two LSTM layers, 64 hidden units, dropout 0.2, dense output layer.	Learns temporal dependencies while limiting overfitting.
Optimizer/loss	Adam optimizer and mean squared error loss.	Standard for continuous pollutant forecasting.
Forecast horizon	2026-2030 recursive forecasting.	Supports urban-growth scenario analysis under Vision 2030 timeframe.

6. Validation Evidence and Analytical Interpretation

The strongest evidence is the inversion validation: XGBoost reproduced source-domain monthly ground observations with very high R^2 for both NO_2 and SO_2. The forecasting evidence is strongest for NO2 in Los Angeles and Seoul; temporal forecasting of SO_2 is more challenging and should be presented with greater uncertainty, as SO_2 can be influenced by episodic point sources and exhibits lower dynamic variance.

Physical credibility

- SHAP confirms inverse PBLH relationship.
- Wind and mixing variables act as dispersion controls.
- Satellite columns are modified by meteorological context.

Forecasting caveat

- NO2 forecasting is strongest in Los Angeles and Seoul tests.
- SO2 inversion is strong, but SO2 LSTM forecasting is less stable across cities.
- Scenario forecasts should be reported with uncertainty bands.

Figure S3. Analytical validation dashboard summarizing high-performing inversion results and representative NO_2 LSTM forecast tests.

Evaluation task	Model	Metrics	Interpretation
Surface NO_2 inversion	XGBoost	$R^2 = 0.9983; RMSE = 0.1843\ ug/m3$	Very strong source-domain fit; should be paired with spatial/temporal holdout to avoid optimistic interpretation.
Surface SO_2 inversion	XGBoost	$R^2 = 0.9921; RMSE = 0.0326\ ug/m3$	Strong inversion performance, but SO_2 remains sensitive to point-source and episodic behavior.
NO_2 forecast - Los Angeles	LSTM	$R^2 = 0.838; RMSE = 1.535$	Strong representative temporal generalization.
NO_2 forecast - Seoul	LSTM	$R^2 = 0.814; RMSE = 1.475$	Strong cross-city NO_2 seasonal tracking.
NO_2 forecast - Dubai/Baghdad/London	LSTM	$R^2 = 0.394 / 0.426 / 0.176$	Moderate or weak city-specific skill; suggests target-domain uncertainty must be reported.
SO_2 forecast tests	LSTM	City tests ranged from negative to modest positive R^2	SO_2 forecast component should be treated as exploratory and uncertainty-heavy.

The study emphasizes the strong inversion model and the NO_2 forecasting success, while transparently stating that SO_2 temporal forecasting requires further robustness testing because SO_2 signals are often sparse, episodic and point-source dominated.

7. Exploratory and Diagnostic Findings

The EDA supports the physical framing of the model. NO_2 satellite columns show a strong negative association with temperature and weaker negative associations with wind speed and PBLH, consistent with seasonal dilution and ventilation effects. Positive correlations with SO_2 and built-up area indicate shared combustion/urban source structure.

Diagnostic signal	Observed Correlation	Scientific meaning
Temp vs NO_2 column	~ -0.515	Seasonality and stronger photochemical/removal or mixing conditions in warmer periods.
Wind speed vs NO_2 column	~ -0.215	Ventilation reduces local accumulation.
PBLH vs NO_2 column	~ -0.096	Mixing-layer depth has a dilution signal, though non-linear effects are expected.
SO_2 vs NO_2 column	~ +0.467	Potential shared combustion/industrial source signatures.

Diagnostic signal	Observed Correlation	Scientific meaning
Built-up area vs NO_2 column	~ +0.210	Urban footprint acts as a proxy for anthropogenic intensity.
PBLH range	~ 134 to 3289 m in the cleaned table	Supports the need for atmospheric constraints in arid and urban settings.

8. SHAP Interpretability and Physical Consistency

SHAP analysis is used to test whether the AI model behaves like an atmospheric model rather than a black-box curve fitter. The expected physical signatures are: higher PBLH should reduce surface concentrations, higher wind should reduce local accumulation through dispersion, and satellite column density should remain important but should be modulated by meteorological context.

- *PBLH effect:* a strong inverse relationship supports the atmospheric lid/dilution mechanism.
- *Wind effect:* ventilation should reduce accumulated pollution, especially under weak mixing conditions.
- *Source-load effect:* NO_2 and SO_2 columns should provide the main satellite signal, but the column-to-surface relationship should not be treated as purely linear.
- Transferability argument: if SHAP confirms universal mixing and dispersion behavior, the model is more defensibly transferred to arid target cities.

9. Domain Adaptation and Virtual Station Logic

The Virtual Station concept is not a simple interpolation. It is a two-step modelling construct: first, a source-domain labelled model learns how column observations translate to surface concentrations under physical constraints; second, the trained function is applied to target cities where ground data are sparse. The resulting reconstructed surface histories become Virtual Stations.

Element	Source-domain implementation	Target-domain
Ground labels	U.S. EPA AQS observations for 20 data-rich cities.	Local Saudi/MENA ground labels are not required for first-order reconstruction.
Physical variables	PBLH, wind and temperature learned in the source-domain feature space.	Same atmospheric mechanisms govern dilution and dispersion in target cities.
Urban	Population and built-up area provide	Supports forecasting under rapid urban

Element	Source-domain implementation	Target-domain
variables	anthropogenic context.	expansion.
Risk control	Feature-space overlap and arid analog checks should be reported.	Prevents unsupported extrapolation outside trained meteorological regimes.

10. Uncertainty, Caveats, and Robustness Controls

A scientifically strong poster should explicitly acknowledge uncertainty rather than hiding it. The following controls enhance the credibility of the poster and help address potential technical questions from atmospheric science reviewers.

Uncertainty source	Implication	Recommended robustness control
Spatial representativeness	A 30 km airshed may smooth intra-urban hot spots.	Test sensitivity to buffer size and compare with known urban plumes.
Reanalysis resolution	MERRA-2 is coarser than Sentinel-5P and cannot resolve street-scale meteorology.	Use zonal aggregation rather than artificial downscaling; report scale mismatch.
Temporal aggregation	Monthly means reduce noise but smooth short pollution episodes.	Keep monthly framework for climate-scale signals; avoid claiming event-scale forecasting.
Domain shift	Source cities and Saudi target cities may differ in emissions and dust chemistry.	Report feature-space overlap and arid analog performance.
Autocorrelation leakage	Random splits can inflate performance when records from the same city/time are related.	Add city-blocked and time-blocked holdout tests for final publication.
SO_2 episodicity	SO_2 may be dominated by point sources and episodic industrial events.	Report SO_2 forecast uncertainty separately from NO_2.
Scenario assumptions	2026-2030 projections depend on urban-growth and climatology assumptions.	Present forecasts as scenario trajectories, not deterministic predictions.

11. Contribution and Strength

The poster is competitive because it integrates a timely atmospheric-science problem with an interpretable AI workflow. The contribution is not merely model accuracy; it is the construction of a physically constrained monitoring architecture for regions where conventional ground networks are sparse.

Contributed area	**Strength**	**Remarks**
Originality	Cascaded XGBoost-to-LSTM Virtual Station framework.	Combines inversion, forecasting and domain adaptation in one coherent workflow.
Atmospheric relevance	PBLH, wind and temperature constrain the column-to-surface translation.	Links machine learning to known atmospheric mixing and dispersion mechanisms.
Data engineering	DUMI harmonizes satellite, meteorology, urban and ground observations.	Addresses a major barrier in multi-source urban air-quality studies.
Analytical transparency	SHAP interpretation and explicit uncertainty controls.	Preempts black-box criticism and strengthens scientific trust.
Policy relevance	Saudi/MENA focus under Vision 2030 urban transformation.	Connects technical outputs to environmental planning and public-health relevance.

12. Conclusion

This supplementary file strengthens the EGU poster by showing that the work is not only visually attractive but also methodologically traceable and scientifically defensible. The DUMI database provides a harmonized multi-source foundation; the XGBoost inversion model offers a high-performing and physically constrained route from satellite columns to surface concentrations; and the LSTM model extends the reconstructed surface histories into future scenario space. The strongest evidence currently lies in the satellite-to-surface inversion, where NO_2 and SO_2 show excellent source-domain validation. The NO_2 LSTM forecasts also demonstrate encouraging skill in representative cities such as Los Angeles and Seoul. At the same time, the supplement transparently identifies weaker city-specific and SO_2 forecast performance, reinforcing the need for uncertainty reporting, spatial holdout validation and pollutant-specific interpretation. Overall, the poster presents a scalable Virtual Station framework for data-scarce arid regions. Its value lies in combining remote sensing, atmospheric physics, machine learning interpretability and urban-growth scenario modelling to support proactive air-quality assessment in Saudi Arabia and other rapidly urbanizing dryland environments.

Acknowledgement

The authors gratefully acknowledge King Fahd University of Petroleum and Minerals for academic and research support. The study acknowledges the European Space Agency for Sentinel-5P/TROPOMI data, NASA for MERRA-2 reanalysis products, the U.S. Environmental Protection Agency for ground-based air-quality observations, and the UN World Urbanization Prospects database for urban-growth information. Appreciation is extended to the EGU Atmospheric Sciences community for providing a platform to present and discuss this research.

14. Selected References

Al Yammahi, A., & Aung, Z. (2023). Forecasting the concentration of NO2 using statistical and machine learning methods: A case study in the UAE. *Heliyon*, *9*(2), e12584. https://doi.org/10.1016/j.heliyon.2022.e12584

Bui, T.-C., Le, V.-D., & Cha, S.-K. (2018). *A Deep Learning Approach for Forecasting Air Pollution in South Korea Using LSTM*. http://arxiv.org/abs/1804.07891

Chen, T., & Guestrin, C. (2016). *XGBoost: A Scalable Tree Boosting System*. https://doi.org/10.1145/2939672.2939785

Cooper, M. J., Martin, R. V., Hammer, M. S., Levelt, P. F., Veefkind, P., Lamsal, L. N., Krotkov, N. A., Brook, J. R., & McLinden, C. A. (2022). Global fine-scale changes in ambient NO2 during COVID-19 lockdowns. *Nature 2022 601:7893*, *601*(7893), 380–387. https://doi.org/10.1038/s41586-021-04229-0

Dou, X., Liao, C., Wang, H., Huang, Y., Tu, Y., Huang, X., Peng, Y., Zhu, B., Tan, J., Deng, Z., Wu, N., Sun, T., Ke, P., & Liu, Z. (2021). Estimates of daily ground-level NO2 concentrations in China based on Random Forest model integrated K-means. *Advances in Applied Energy*, *2*. https://doi.org/10.1016/j.adapen.2021.100017

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Lamsal, L. N., Krotkov, N. A., Celarier, E. A., Swartz, W. H., Pickering, K. E., Bucsela, E. J., Gleason, J. F., Martin, R. V., Philip, S., Irie, H., Cede, A., Herman, J., Weinheimer, A., Szykman, J. J., & Knepp, T. N. (2014). Evaluation of OMI operational standard NO 2 column retrievals using in situ and surface-based NO 2 observations. *Atmospheric Chemistry and Physics*, *14*(21), 11587–11609. https://doi.org/10.5194/acp-14-11587-2014

Long, S., Wei, X., Zhang, F., Zhang, R., Xu, J., Wu, K., Li, Q., & Li, W. (2022). Estimating daily ground-level NO2 concentrations over China based on TROPOMI observations and machine learning approach. *Atmospheric Environment*, *289*, 119310. https://doi.org/10.1016/j.atmosenv.2022.119310

Lundberg, S., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *31st Conference on Neural Information Processing Systems (NIPS)*. http://arxiv.org/abs/1705.07874

Shetty, S., Schneider, P., Stebel, K., David Hamer, P., Kylling, A., & Koren Berntsen, T. (2024).

Estimating surface NO2 concentrations over Europe using Sentinel-5P TROPOMI observations and Machine Learning. *Remote Sensing of Environment*, *312*, 114321. https://doi.org/10.1016/j.rse.2024.114321

WHO. (2021). *WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide*. https://www.who.int/publications/i/item/9789240034228