1ST IEEE INTERNATIONAL CONFERENCE ON SUSTAINABLE SCIENCE AND INFRASTRUCTURE DEVELOPMENT (ICSSID 2026)

Reliability-Aware PM2.5 Mapping in Africa via Satellite-Reanalysis Fusion and Sparse Monitors

BY

YAW OSEI ADJEI, DAVIS OPOKU, EPHRAIM ABOTSI, BENJAMIN TEI PARTEY, PhD.

Kwame Nkrumah University of Science and Technology (KNUST), Kumasi, Ghana

Presentation Outline

1. Introduction
2. Problem Statement
3. Literature Review
4. Methodology
5. Experimental Evaluation
6. Results and Discussion
7. Conclusion
9. Reference

Introduction

Mapping PM2.5 across a continent with almost no ground truth

- Fine particulate matter (PM2.5, 2.5 micrometers or less) drives severe cardiovascular and respiratory harm worldwide.
- The WHO 2021 guidelines set a strict annual mean target of 5 micrograms per cubic meter (WHO, 2021).
- Vast parts of Africa have no continuous ground monitoring to check compliance.
- Exposure studies and air quality policy face a basic observational deficit.
- This study fuses satellite, reanalysis, and meteorology into a reliability-aware PM2.5 map. It shows not just what the model predicts, but where the prediction can be trusted.

2.5 µm

Aerodynamic diameter or less, the size fraction that lodges deep in the lungs

5 µg/m³

WHO 2021 annual mean guideline, unmeasurable across most of Africa

847

Station-months from 404 monitors across 2016 to 2025, the entire training base

Problem Statement

2

Four problems break naive PM2.5 models in Africa

Health burden without data. PM2.5 causes severe cardiovascular and respiratory harm, yet most of Africa cannot measure it (Burnett et al., 2018).

Sparse, clustered monitors. 404 stations exist, packed into South and West Africa. East Africa has 8, Central and North Africa almost none.

Spatial leakage inflates scores. Random cross-validation breaks Tobler's First Law and inflates R^2 by 20 to 30 percent (Ploton et al., 2020).

Covariate shift invalidates uncertainty. Training and test regions differ so much that standard uncertainty bounds fail (Barber et al., 2023).

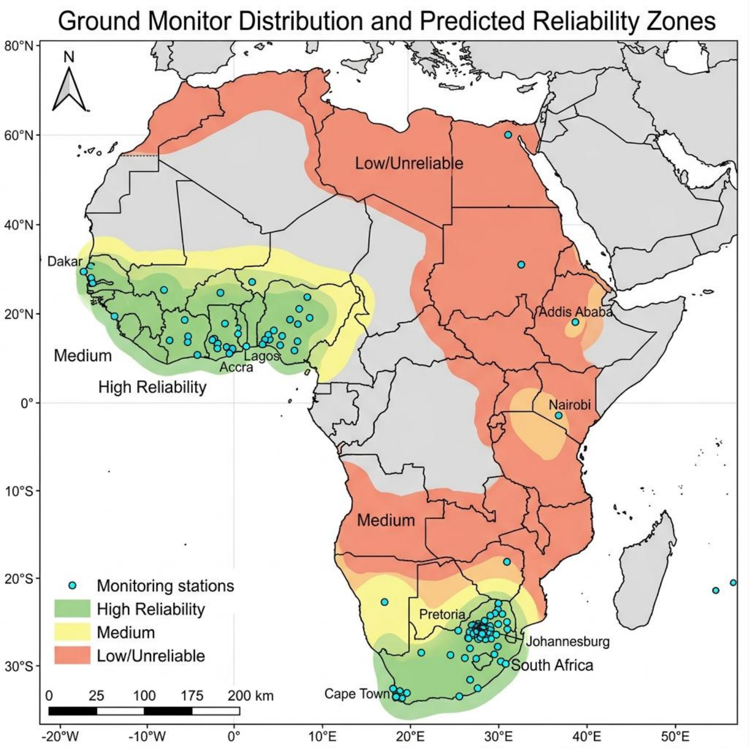


Fig. 1 - Ground monitor distribution and predicted reliability zones.

Literature Review

1 Satellite–Reanalysis Fusion

- Hammer et al. (2020) and van Donkelaar et al. (2021) joined satellite AOD with GEOS-Chem to widen global coverage.
- In data-sparse regions, satellite-driven methods can beat transport models (Diao et al., 2019).
- Tang et al. (2024) flagged severe data gaps in East African cities.
- Porcheddu et al. (2024) corrected model output with machine learning, cutting error sharply.

2 Spatial Cross-Validation

- Roberts et al. (2017) required validation that respects spatial structure.
- Ploton et al. (2020) showed random splits hide poor predictive skill.
- Schratz et al. (2019) nested hyperparameter tuning inside spatial folds.
- Wadoux et al. (2021) gave a counterpoint on when spatial CV is too strict.

3 Conformal Prediction

- Split-conformal prediction gives distribution-free intervals (Vovk et al., 2005).
- Barber et al. (2023) showed exchangeability breaks under distribution shift.
- Kang et al. (2024) proposed GeoConformal, weighting scores by spatial proximity.
- Mao et al. (2024) developed smoothed local spatial conformal prediction.

Methodology

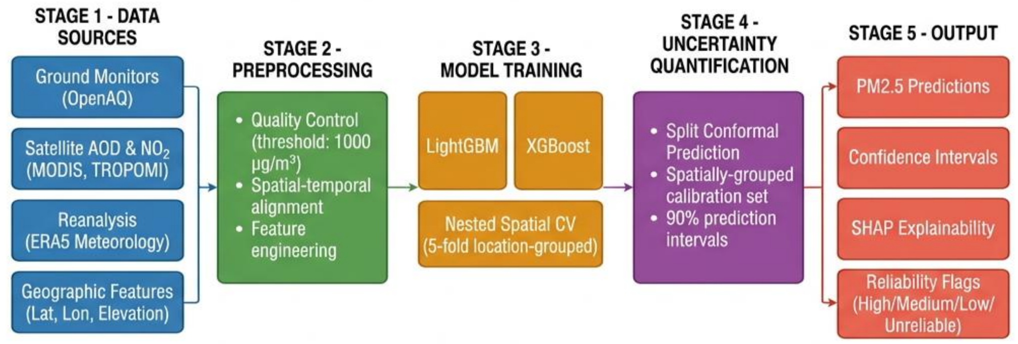


Fig. 2 - Five stages: data sources, quality-controlled preprocessing, model training, uncertainty quantification, and reliability-flagged output.

Multi-source data fusion
15 predictors from satellite, ERA5 reanalysis, and geography.

Nested spatial cross-validation
5-fold outer evaluation, 3-fold inner tuning, no geographic leakage.

Conformal prediction
90 percent intervals from spatially disjoint calibration sets.

Four-tier reliability flags
Tiers set by AQI bin widths and fold-level R^2.

Experimental Evaluation

Data and honest evaluation protocol

- 847 station-months from 404 OpenAQ monitors, 2016 to 2025.
- Upper cleaning threshold of 1000 micrograms per cubic meter, justified by Harmattan dust and biomass burning.
- 15 predictors from satellite, ERA5, and geography.
- 5-fold outer evaluation, 3-fold inner tuning (Optuna, 100 trials per fold).

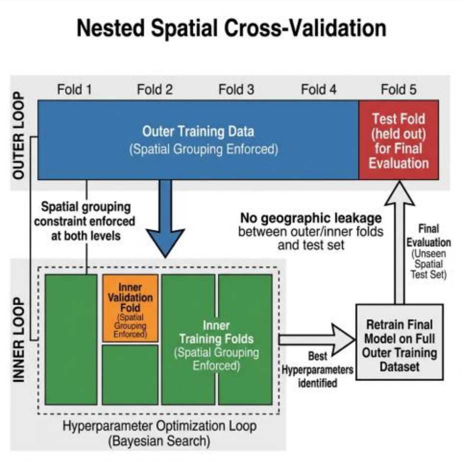
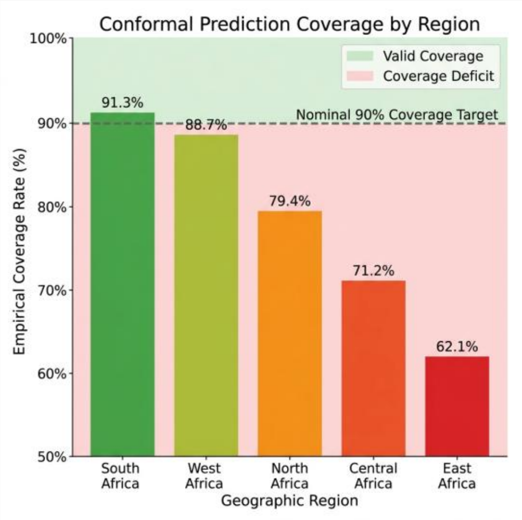


Fig. 3 - Nested spatial CV. Inner loop tunes ; outer loop tests on unseen regions .

Parameter	LightGBM	XGBoost
Learning rate	0.05	0.03
Max depth	6	7
Num. estimators	800	1,000
Min. data in leaf	20	-
Min. child weight	-	5
Feature fraction	0.7	0.8
Bagging fraction	0.8	0.85
Reg. alpha (L1)	0.1	0.05
Reg. lambda (L2)	1.0	1.5
Num. leaves	31	-
Subsample	-	0.8

Results and Discussion

Model / Region	R^2	RMSE	Cvg.
LightGBM – Overall	**0.134**	**28.7**	**83.4%**
South Africa	0.312	19.2	91.3%
West Africa	0.187	31.5	88.7%
East Africa	**−0.749**	52.3	**62.1%**
XGBoost – Overall	**0.155**	**27.4**	**84.1%**
South Africa	0.341	18.6	91.8%
West Africa	0.203	30.1	89.2%
East Africa	**−0.682**	49.8	**64.3%**



What the numbers say

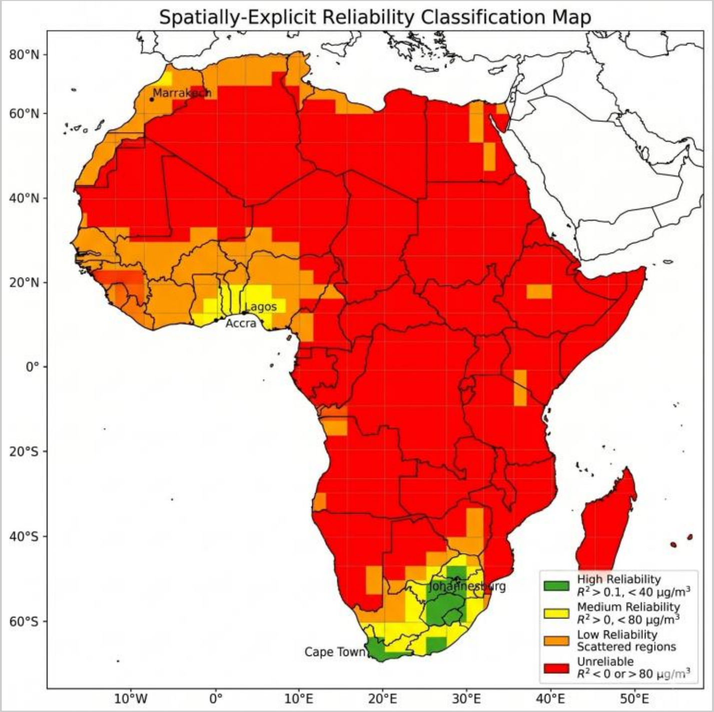
- Spatial CV cuts R^2 from above 0.90 to 0.13–0.15, the true interpolation versus prediction gap.
- South Africa is reliable: R^2 up to 0.34, coverage near 91 percent.
- SHAP ranks latitude, longitude, satellite AOD, and NO_2 as the top drivers (Lundberg & Lee, 2017).

East Africa fails by design, not by chance

$R^2 = -0.75$ and coverage 62.1 percent against a 90 percent target. A severe latitude shift (KS = 0.682) forces the model to extrapolate, and the Harmattan flag misleads where the ITCZ governs the weather.

Conclusion

Where the map can be trusted



Reliability map. Green zones are actionable for AQI. Red marks unreliable predictions across Central and East Africa.

Honest spatial scores

Gradient-boosted models reach R^2 of 0.13 to 0.15 under spatial cross-validation. This shows the real gap between interpolation and true prediction (Ploton et al., 2020).

East Africa breaks down

The East Africa fold drops to $R^2 \approx -0.75$. Harmattan and ITCZ divergence plus a severe latitude shift drive the failure (KS = 0.682).

Coverage collapses

Conformal coverage falls to 62.1% in East Africa. Broken exchangeability makes the prediction intervals invalid there (Barber et al., 2023).

Maps that show limits

Four reliability flags mark where predictions are safe to use. They guide monitor deployment and climate finance checks.

Conclusion

8

Implications

- **Health risk.** AQI advisories hold only in High zones where width stays under 40 μg/m³. Unreliable zones cannot replace physical measurements.
- **Climate finance.** An Unreliable flag gates carbon crediting and industrial permitting under SDGs 3.9, 9, and 11.6 (United Nations, 2015).
- **Monitor deployment.** Interval width paired with WorldPop density (Tatem, 2017) ranks high-burden, high-uncertainty areas. East Africa comes first.

Future Work

- **CTM covariates.** Add MERRA-2 (Randles et al., 2017) and CAMS (Inness et al., 2019), with post-process correction (Porcheddu et al., 2024).
- **Spatial conformal.** Move to GeoConformal (Kang et al., 2024) or SLSCP (Mao et al., 2024) to restore valid coverage.
- **Region models.** Build sub-continental models with ITCZ position indicators for East Africa, replacing one pan-African model.
- **More sensors.** Deploy targeted low-cost sensor arrays in the highest-priority Unreliable zones, led by East African cities.

References

[1] R. Burnett, H. Chen, M. Szyszkowicz et al., "Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter," *Proceedings of the National Academy of Sciences*, vol. 115, no. 38, pp. 9592–9597, 2018.
[2] World Health Organization, "WHO global air quality guidelines: Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide," Geneva, 2021.
[3] E. P. Petkova, D. W. Jack, N. H. Volavka-Close, and P. L. Kinney, "Particulate matter pollution in African cities," *Air Quality, Atmosphere & Health*, vol. 6, no. 3, pp. 603–614, 2013.
[4] P. D. M. C. Katoto, L. Byamungu, A. S. Brand et al., "Ambient air pollution and health in Sub-Saharan Africa: Current evidence, perspectives and a call to action," *Environmental Research*, vol. 173, pp. 174–188, 2019.
[5] M. S. Hammer, A. van Donkelaar, C. Li et al., "Global estimates and long-term trends of fine particulate matter concentrations (1998–2018)," *Environmental Science & Technology*, vol. 54, no. 13, pp. 7879–7890, 2020.
[6] A. van Donkelaar, M. S. Hammer, L. Bindle et al., "Monthly global estimates of fine particulate matter and their uncertainty," *Environmental Science & Technology*, vol. 55, no. 22, pp. 15 287–15 300, 2021.
[8] D. R. Roberts, V. Bahn, S. Ciuti et al., "Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure," *Ecography*, vol. 40, no. 8, pp. 913–929, 2017.
[9] P. Ploton, F. Mortier, M. Réjou-Méchain et al., "Spatial validation reveals poor predictive performance of large-scale ecological mapping models," *Nature Communications*, vol. 11, no. 1, p. 4540, 2020.
[11] R. F. Barber, E. J. Candès, A. Ramdas, and R. J. Tibshirani, "Conformal prediction beyond exchangeability," *The Annals of Statistics*, vol. 51, no. 2, pp. 816–845, 2023.
[12] W. Tang, R. Kumar, A. Del Moral Méndez et al., "East African megacity air quality: Rationale and framework for a measurement and modeling program," *Bulletin of the American Meteorological Society*, vol. 105, no. 6, pp. E1079–E1098, 2024.
[13] A. Porcheddu, C. Rundel, S. Safieddine et al., "Post-process correction improves the accuracy of satellite PM2.5 retrievals," *Atmospheric Measurement Techniques*, vol. 17, no. 18, pp. 5747–5764, 2024.
[14] P. Schratz, J. Muenchow, E. Iturritxa, J. Richter, and A. Brenning, "Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data," *Ecological Modelling*, vol. 406, pp. 109–120, 2019.
[15] A. M. J.-C. Wadoux, G. B. M. Heuvelink, S. de Bruin, and D. J. Brus, "Spatial cross-validation is not the right way to evaluate map accuracy," *Ecological Modelling*, vol. 457, p. 109692, 2021.
[16] V. Vovk, A. Gammerman, and G. Shafer, "Algorithmic learning in a random world," Springer, 2005.
[17] Y. Kang, R. Cao, and S. Wang, "GeoConformal prediction: A model-agnostic framework for measuring the uncertainty of spatial prediction," *Annals of the American Association of Geographers*, vol. 114, no. 6, pp. 1291–1312, 2024.
[18] H. Mao, R. Martin, and B. J. Reich, "Valid model-free spatial prediction," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 86, no. 3, pp. 749–768, 2024.
[27] C. A. Randles, A. M. da Silva, V. Buchard et al., "The MERRA-2 aerosol reanalysis, 1980 onward. Part I: System description and data assimilation evaluation," *Journal of Climate*, vol. 30, no. 17, pp. 6823–6850, 2017.
[29] A. Inness, M. Ades, A. Agustí-Panareda et al., "The CAMS reanalysis of atmospheric composition," *Atmospheric Chemistry and Physics*, vol. 19, no. 6, pp. 3515–3556, 2019.
[32] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 30, 2017.
[33] United Nations, "Transforming our world: The 2030 agenda for sustainable development," 2015.
[34] A. J. Tatem, "WorldPop, open data for spatial demography," *Scientific Data*, vol. 4, p. 170004, 2017.

THANK YOU!