Proceedings of Machine Learning Research 329:1–3, 2026 Conformal and Probabilistic Prediction with Applications

Reliable Environmental Planning at HS2 Construction Sites

Jasmin Ahluwalia jasmin.ahluwalia.2024@live.rhul.ac.uk
Xu Feng xu.feng@rhul.ac.uk
Khuong An Nguyen khuong.nguyen@rhul.ac.uk
Royal Holloway University of London, Surrey TW20 0EX, United Kingdom

Matthew Tye matthew.tye@hs2.org.uk
High Speed 2, Snow Hill Queensway, Birmingham B4 6GA, United Kingdom

Editors: Ernst Ahlberg, Ulf Johansson, Henrik Boström, Alberto Carlevaro, Johan Hallberg Szabadváry and Lars Carlsson

Abstract

High Speed 2 (HS2) is a major UK railway infrastructure project with substantial economic and transport benefits. However, its scale and operational complexity have potentially created environmental impacts. Thus, forecasting air quality and noise levels may support more informed site management. Nevertheless, existing Machine Learning models do not reveal the uncertainty of their predictions, limiting their reliability in such high-stakes environments. Therefore, we propose a state-conditioned conformal prediction (CP) framework that uses recent operational-state information to produce reliable and adaptive prediction intervals. Experiments on real-world HS2 air-quality and noise monitors show that state-conditioned calibration reduced mean interval width by up to 40.38%.

1. Methodology

Conventional rolling CP pools recent residuals without accounting for operational state changes. This can yield inefficient intervals or coverage under temporal dependence and distribution shifts (Gibbs and Candès (2021); Zaffran et al. (2022)). Thus, we propose a state-conditioned calibration framework inspired by change-point detection that estimates intervals from residuals of the current operational state (Sun and Yu (2025)).

Instead of assuming that all recent observations follow the same calibration distribution, the proposed approach divides the time series into two states. For state-detection method (m), each observation is assigned a binary label $S_t^{(m)} \in \{0,1\}$ where $S_t^{(m)} = 1$ denotes an active or high-intensity construction-work state, whereas $S_t^{(m)} = 0$ denotes an inactive or low-intensity construction-work state. Let y_t denote the observed hourly environmental target at time t, and let $\widehat{y}_t = f(X_t)$ denote its point forecast, generated using only information available before time t. The absolute prediction error is defined as $e_t = |y_t - \widehat{y}_t|$, where e_t is used as the nonconformity score for conformal calibration. For residual-based state detection, the same error sequence is also used to identify model-specific error regimes.

For state-detection method m and a rolling calibration window of length W, the state-matched calibration index set is defined as $\mathcal{I}_t^{(m)} = \{i : t - W \leq i < t, S_{i-1}^{(m)} = S_{t-1}^{(m)}\}$ where $S_{t-1}^{(m)}$ denotes the state assigned by method m at time $t-1$. The corresponding set of state-matched nonconformity scores is $\mathcal{E}_t^{(m)} = \{e_i : i \in \mathcal{I}_t^{(m)}\}$. Let $n = |\mathcal{E}_t^{(m)}|$ denote the number of effective calibration scores. For a confidence level of $1-\alpha$, the finite-sample conformal rank is $k = \lceil (n+1)(1-\alpha) \rceil$. The state-conditioned conformal quantile is then defined as the k-th ordered value of the calibration scores with $q_{t,1-\alpha}^{(m)} = \text{Quantile}_k\left(\mathcal{E}_t^{(m)}\right)$. The resulting prediction interval is $C_{t,1-\alpha}^{(m)} = \left[\widehat{y}_t - q_{t,1-\alpha}^{(m)}, \widehat{y}_t + q_{t,1-\alpha}^{(m)}\right]$.

We investigated three state-detection methods: K-means, measurement-level, and residual-level, which identify operational states from working schedules, environmental measurement patterns, temporal segments, recent measurement levels or prediction-error regimes.

2. Empirical results

The framework was evaluated on two HS2 monitors in Birmingham: `BCC_WHD001` for air quality and `BCC_TOS_N1` for noise. Four hourly targets were considered: PM_{10}, $L_{\text{pAeq},T}$, $L_{\text{pA90},T}$, and $L_{\text{pAF,max}}$, representing particulate concentration, average noise exposure, background noise, and maximum short-duration noise, respectively. ARIMAX was used as the main forecasting model, trained on data from 8/25 to 11/25, with 12/25 used for validation and initial conformal calibration and 1/26 for evaluation. A rolling 30-day conformal interval served as the baseline, empirical coverage and mean interval width assessed reliability.

Table 1: State-conditioned CP results (ARIMAX) at 90% and 95% confidence. MAE and RMSE are in $\mu\text{g}\,\text{m}^{-3}$ for PM_{10} and dB(A) for noise targets. Width reduction is measured relative to the rolling 30-day calibrated conformal baseline.

Target	Monitor	State method	MAE	RMSE	Conf. (%)	Coverage (%)	Width Red. (%)
PM_{10}	BCC_WHD001	Residual-level	3.54	5.95	90	83.74	**40.38**
PM_{10}	BCC_WHD001	Residual-level	3.54	5.95	95	93.15	10.90
$L_{\text{pAeq},T}$	BCC_TOS_N1	Residual-level	3.14	3.84	90	90.99	**14.31**
$L_{\text{pAeq},T}$	BCC_TOS_N1	Residual-level	3.14	3.84	95	95.70	**19.49**
$L_{\text{pA90},T}$	BCC_TOS_N1	Measurement-level	4.47	5.31	90	90.99	-0.40
$L_{\text{pA90},T}$	BCC_TOS_N1	Measurement-level	4.47	5.31	95	95.43	-2.59
$L_{\text{pAF,max}}$	BCC_TOS_N1	K-means	4.27	5.45	90	94.76	**16.53**
$L_{\text{pAF,max}}$	BCC_TOS_N1	K-means	4.27	5.45	95	97.45	**14.58**

Table 1 shows that residual-level gave the strongest reliable improvement for $L_{\text{pAeq},T}$ and K-means maintained coverage for $L_{\text{pAF,max}}$. Figure 1 illustrates that at both the 85% and 95% confidence levels, residual-level intervals remain tighter than the rolling 30-day baseline whilst maintaining strong coverage.

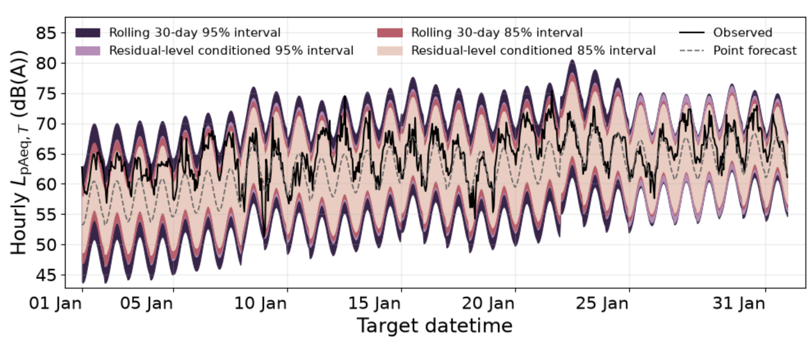

Figure 1: Interval comparison for $L_{\text{pAeq},T}$ at 85% and 95% confidence using ARIMAX.

Acknowledgments

Khuong An Nguyen is supported by UKRI DAFNI Fellowship grant, and Innovate UK grant for the 'TunnelGuard: A Ground-Air Robotic Safety Twin for HS2 Construction' project.

References

I. Gibbs and E. J. Candès. Adaptive conformal inference under distribution shift. In *Advances in Neural Information Processing Systems*, volume 34, pages 1660–1672, 2021.

S. Sun and R. Yu. Conformal prediction for time-series forecasting with change points. In *Advances in Neural Information Processing Systems*, volume 38, 2025.

M. Zaffran, O. Féron, Y. Goude, J. Josse, and A. Dieuleveut. Adaptive conformal predictions for time series. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162, pages 25834–25866, 2022.