

Multi-Horizon Forecasting of Thermal and Ventilation Risks in Naturally Ventilated Hospitals using Multi-Output Transformers

Zeynep Duygu Tekler^{*1,2}, Ben Roberts¹, Raymond Low², Malcolm Cook¹

*1 Loughborough University
Epinal Way, Loughborough
LE11 3TU*

*2 University of Oxford
Wellington Square, Oxford
OX1 2JD*

**Corresponding author: z.d.tekler@lboro.ac.uk*

ABSTRACT

Hospital wards in tropical regions face sustained challenges in maintaining acceptable thermal comfort and indoor air quality, particularly where mechanical cooling and ventilation are not available. This study presents a multi-horizon forecasting framework for indoor environmental risks in naturally ventilated healthcare settings. The framework draws on more than two years of 15-minute interval sensor measurements collected from 14 wards across two regional hospitals in Tamale, Ghana, paired with weather data, to forecast indoor air temperature, relative humidity (RH), and carbon dioxide (CO₂) concentration over a 24-hour horizon. We benchmark a Temporal Fusion Transformer (TFT) against a multi-output XGBoost baseline, while current environmental risk is summarised against international standards at ward-level. The risk summary indicates that overheating and elevated humidity are widespread across the monitored wards, whereas CO₂ exceedance is concentrated in a smaller number of high-risk spaces, particularly neonatal and paediatric wards. Model results show that the TFT meaningfully outperforms XGBoost for temperature and RH forecasting, reducing error by approximately 32% and 35–42%, respectively, across MAE, RMSE, and MAPE, while CO₂ forecasting remains more challenging and is better served by the XGBoost. The proposed framework provides a basis for targeted ventilation management and demonstrates the feasibility of low-cost, health-oriented indoor climate forecasting in low-resource tropical hospitals.

KEYWORDS

Overheating; Indoor environmental quality; Natural ventilation; Forecasting; Machine learning; Hospitals

1 INTRODUCTION

Maintaining comfortable indoor environments in hospitals plays a central role in patient recovery, infection control, and the wellbeing of clinical staff. In tropical regions, achieving these conditions is particularly difficult because mechanical cooling and ventilation infrastructure is frequently absent, limited or outside routine staff control. Sustained overheating has been linked to thermal discomfort, heat-related illness, and elevated mortality among vulnerable patient groups (Kenny et al., 2024), while inadequate ventilation, often inferred from elevated indoor CO₂ concentrations, is associated with increased airborne infection risk (Roberts et al., 2025).

A growing body of work has applied machine learning to indoor environmental quality (IEQ) and thermal comfort prediction. Most existing applications, however, have focused on relatively controlled environments such as offices, classrooms, or experimental testbeds (Godasiaei et al., 2025; Tekler et al., 2024; Boutahri & Tilioua, 2024; Tekler et al., 2023), where occupancy patterns and ventilation systems are stable and well-instrumented. By contrast, naturally ventilated hospital wards, particularly in low-resource regions, present a substantially harder problem. They typically lack continuous environmental monitoring or automated control, and their indoor conditions are shaped by continuous patient occupancy, mixed clinical

functions, variable airflow patterns, and high vulnerability to environmental stressors (Roberts et al., 2022). World Health Organization (WHO) guidance on natural ventilation in healthcare facilities highlights that such settings are particularly susceptible to fluctuating indoor conditions and airborne transmission risk (WHO, 2009). Fixed ventilation schedules and reactive control strategies are therefore rarely sufficient to protect occupant safety and comfort, and there is a clear need for predictive tools that are both accurate and interpretable, so that healthcare staff can anticipate emerging risks and apply timely, low-cost mitigation measures (e.g., adjusting window openings, deploying fans, or relocating patients).

This study addresses this gap by developing a multi-horizon forecasting framework for naturally ventilated healthcare environments. Using over two years of 15-minute interval indoor measurements from 14 hospital wards in Tamale, Ghana, together with weather data, we forecast indoor air temperature, relative humidity (RH), and CO₂ over a 24-hour horizon. We adopt the Temporal Fusion Transformer (TFT) architecture, a recent transformer-based model that combines sequence modelling with variable selection and uncertainty estimation, and benchmark it against a multi-output XGBoost baseline. We also summarise current risk at ward level against international thresholds. The resulting probabilistic forecasts support interpretable day-ahead operational planning in resource-constrained hospital settings.

2 METHODOLOGY

2.1 Study Context and Data

The study was carried out at two regional healthcare facilities in Tamale, Northern Ghana, namely a Central Hospital and a Teaching Hospital. Tamale lies in a tropical savannah climate with high daytime temperatures, distinct wet and dry seasons, and very limited access to mechanical cooling. Both hospitals serve a broad patient population and rely predominantly on natural ventilation. Fourteen wards were instrumented to capture a wide range of clinical functions, including maternity, paediatric, neonatal, emergency, and general inpatient areas. All monitored spaces were primarily naturally ventilated, with airflow facilitated by operable architectural features such as louvred windows, sliding panes, and bottom-hung casements. Mechanical conditioning was limited to two neonatal wards with ducted air conditioning without staff control, and one paediatric ward with wall-mounted unit. Ceiling and pedestal fans were used intermittently across wards.

Indoor environmental data were recorded using Eltek MS47B wall-mounted sensors deployed in each ward. The sensors recorded indoor air temperature (°C), relative humidity (%), and CO₂ concentration (ppm) at 15-minute intervals (Table 1). CO₂ was used as a proxy for occupancy intensity and ventilation effectiveness. Sensors were installed between 1.6 and 2.2 m above floor level and positioned away from direct sunlight, doors, and local heat sources. An additional outdoor sensor was deployed at each site to monitor ambient conditions.

Table 1: Measured indoor variables and sensor specifications

Sensor	Variable	Frequency (min)	Unit	Range	Resolution / Accuracy
Eltek MS47B	Temperature	15	°C	−30 to +65	0.1 / ±0.4 (+5 to +40)
Eltek MS47B	Relative humidity	15	%	0 to 100	0.1 / ±2 (10–90); ±4 (0–100)
Eltek MS47B	CO ₂ concentration	15	ppm	0 to 5000	3 / ±50

Measurements were collected continuously from 8 June 2022 to 18 January 2025, spanning both wet and dry seasons. Implausible outliers, including CO₂ spikes caused by tampering or

sensor displacement were removed, and short missing intervals of less than one hour were linearly interpolated. Ward-level metadata, including patient type, floor area, ventilation type, and sensor mounting height are summarised in Table 2.

Table 2: Ward-level metadata for the 14 monitored hospital wards in Tamale, Ghana.

Ward ID	Hospital	Ward type	Patient type	Floor area (m ²)	Natural ventilation	Sensor height (m)
6584	Central	Outdoor	All	–	–	2.4
6597	Central	Labour	Pregnant women	20.5	Adj. louveres	1.6
6598	Central	Waiting Area	Pregnant women & neonates	133	Fire-exit door	2.2
6600	Central	Emergency	All	36	Adj. louveres	1.6
6608	Central	Maternity	Pregnant & postpartum	106	Side-hung casement	1.6
6609	Central	Male	Adult males	114	Adj. & fixed louveres	1.6
6578	Teaching	Paediatrics (<6m)	Infants under 6 months	34	Adj. & fixed louveres	1.7
6579	Teaching	Paediatrics (>6m)	Children over 6 months	18	Adj. louveres	1.6
6588	Teaching	Neonatal Int. Care	Unstable neonates	33	Sliding panes	2.2
6586	Teaching	Kangaroo	Stable neonates	39	Adj. louveres	1.6
6581	Teaching	Post-natal	Postpartum & neonates	46	Adj. & fixed louveres	1.7
6589	Teaching	Pre-natal	Pregnant women	63	Adj. & fixed louveres	1.7
6585	Teaching	Emergency (Y)	All	105	Sliding panes	1.8
6587	Teaching	Outdoor	All	–	–	2.4

Outdoor weather data were obtained from the ERA5 reanalysis dataset produced by the European Centre for Medium-Range Weather Forecasts (ECMWF, 2024). ERA5 provides hourly estimates of key atmospheric variables, including ambient temperature, dew point, surface pressure, and solar radiation, at a horizontal resolution of approximately 31 km. The data were spatially interpolated to the hospital locations and temporally aligned with the indoor sensors by forward-filling the hourly ERA5 series to a 15-minute resolution.

2.2 Indoor Air Quality Standards

To benchmark environmental risk in each ward, we compared observed values against internationally recognised IEQ thresholds. CIBSE TM40 (2020) and WHO (2009) recommend that indoor temperatures in patient-care spaces should not exceed 27 °C, particularly in tropical regions. ASHRAE Standard 170 (2021) and WHO (2021) recommend keeping CO₂ concentrations below 1000 ppm to ensure adequate ventilation. For RH, a band of 30%–60% is recommended to balance comfort and infection control by limiting pathogen viability while preserving mucosal health. The thresholds adopted in this study are summarised in Table 3.

Table 3: Reference indoor air quality thresholds adopted in this study.

Variable	Recommended range	Reference
Temperature (°C)	≤ 27 °C	WHO (2009); CIBSE TM40 (2020)
CO ₂ concentration (ppm)	≤ 1000 ppm	ASHRAE 170 (2021); WHO (2021)
Relative humidity (%)	30%–60%	ASHRAE 170 (2021); WHO (2009)

2.3 Forecasting Model Development

We employed two complementary forecasting approaches: TFT as a deep-learning sequence model, and XGBoost as a gradient-boosted tree ensemble baseline. Both models predict indoor temperature T_t , relative humidity RH_t , and CO₂ concentration CO_{2t} , over future time steps based on historical measurements and contextual features.

The forecasting task is formulated as a multivariate, multi-output, multi-step time-series regression problem. Given a sliding input window of v historical steps and static metadata Z , the models learn a mapping from past observations to the next k steps of all three target variables:

$$f: T_{t-v:t}, RH_{t-v:t}, CO_{2t-v:t}, W_{t-v:t}, Z \rightarrow [T_{t+1:t+k}, RH_{t+1:t+k}, CO_{2t+1:t+k}] \quad (1)$$

where $T_{t-v:t}$, $RH_{t-v:t}$, and $CO_{2t-v:t}$ represent the lagged indoor environmental variables, $W_{t-v:t}$ represents the lagged weather variables, and Z encodes static attributes (city, hospital, ward type). We set $v = 288$ steps (3 days) and $k = 96$, corresponding to a 24-hour forecast horizon at 15-minute resolution. A sliding window with stride 4 (1 hour) was used to reduce redundancy while retaining good temporal coverage. Both models were trained using a per-ward chronological split, with the first 80% of each ward’s series used for training and the remaining 20% reserved for testing, ensuring evaluation reflects generalisation to unseen time periods. Forecast accuracy was assessed using mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE) across the 24-hour horizon.

2.3.1 Temporal Fusion Transformer

TFT (Lim et al., 2021) is a sequence-to-sequence architecture for multi-horizon forecasting with built-in interpretability. It combines long short-term memory (LSTM) encoders for short-term sequential dependencies with multi-head self-attention for longer-range temporal structure, and performs dynamic variable selection via gated residual networks. Static and temporal covariates are fused through learned embeddings. We configured the model with a hidden state size of 64, four attention heads, a dropout rate of 0.1, and a hidden continuous size of 32. Static categorical features (city, hospital, ward type) were encoded as learned embeddings; time-varying known covariates (hour of day, day of week, month, week of year, outdoor weather) provided future temporal context; and the three indoor targets served as time-varying unknown variables, with per-ward group normalisation.

Training used the Adam optimiser (learning rate 10^{-3}), a batch size of 256, gradient clipping at 0.1, and a maximum of 50 epochs. A 10% chronological hold-out from the training partition was used as a validation set for early stopping. The model was optimised using a quantile (pinball) loss to produce probabilistic forecasts. For predicted quantile $\hat{q}$ and target value y , the pinball loss is:

$$L_q(y, \hat{q}) = \max[q(y - \hat{q}), (q - 1)(y - \hat{q})] \quad (2)$$

with $q \in \{0.02, 0.1, 0.25, 0.5, 0.75, 0.9, 0.98\}$. This yields a full predictive distribution per target across the forecast horizon, with the 50th percentile used as the point forecast.

2.3.2 XGBoost baseline

As a baseline, we trained a single multi-output XGBoost model (Chen & Guestrin, 2016) using the native multi-output tree strategy, jointly predicting all three targets across the 96-step horizon. The input feature vector was constructed by flattening the 288-step lookback of dynamic features—the three indoor variables, cyclical time encodings (sine/cosine of hour, day of week, month, and week of year), and outdoor weather variables—concatenated with one-hot encodings of static ward metadata. The model was configured with 100 estimators, a maximum tree depth of 3, a learning rate of 0.15, row subsampling of 0.8, column subsampling of 0.7 per tree, the histogram-based tree method, and squared-error loss. Early stopping was applied using a 90/10 chronological train/validation split within the training partition.

3 RESULTS

3.1 Ward-level Risk Summary

This section summarises the prevalence and timing of threshold exceedance across the 14 wards using a compact ward-level summary (Table 4). Temperature non-compliance was widespread, with most wards exceeding 27 °C for much of the day and peak exposure occurring in the early-to-mid afternoon. Relative humidity was also frequently outside the recommended range, although exceedance was more variable and tended to be stronger overnight and in the early morning. In contrast, CO₂ exceedance was concentrated in neonatal and paediatric wards, indicating that ventilation risk was spatially localised and more strongly linked to ward activity.

Table 4: Summary of current environmental risk by indicator and ward group.

Indicator	Spatial pattern	Temporal pattern	Operational implication
Temperature (°C)	Widespread across most wards	Peak exposure early-to-mid afternoon	Facility-wide shading and cross-ventilation
Relative humidity (%)	Widespread, but more variable by ward	Higher overnight and early morning	Timed ventilation and moisture-source management
CO ₂ (ppm)	Concentrated in neonatal and paediatric spaces	Episodic peaks linked to occupancy	Targeted airflow management and crowding control

Overall, the tabulated summary indicates that heat and humidity mitigation must be implemented at the facility level, whereas CO₂ control can be targeted to a smaller number of high-risk spaces and periods.

3.2 Forecasting Model Performance

Table 5: Overall forecasting performance for XGBoost and TFT, aggregated across all 14 wards and the full 24-hour forecast horizon. Best results per metric are shown in bold.

Model	Temperature (°C)			Relative humidity (%)			CO ₂ (ppm)		
	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE
XGBoost	1.19	1.69	4.09	6.21	7.78	14.3	103	211	14.8
TFT	0.812	1.15	2.75	3.67	5.05	8.34	117	269	14.6

Table 5 summarises overall forecasting performance aggregated across all 14 wards and the full 24-hour horizon. The TFT delivers consistently stronger results than XGBoost for indoor temperature and RH. In particular, it reduces error for indoor temperature by approximately 31.8%–32.8% across MAE, RMSE, and MAPE, and improves RH forecasting by approximately 35.1%–41.7% on the same metrics. Together, these results indicate that TFT

achieves a substantial gain in predictive performance for variables that exhibit relatively smooth temporal evolution and pronounced diurnal structure. By contrast, CO₂ dynamics are harder to forecast, Table 5 shows that CO₂ prediction is better served by the XGBoost baseline. This divergence suggests that forecasting performance is variable-dependent and tightly linked to the underlying drivers of each indoor variable.

A plausible explanation for TFT’s advantage on temperature and RH is that it is explicitly designed for multi-horizon sequence forecasting and includes architectural components that match the ward setting well. These include learned embeddings for static categorical context (e.g., ward and hospital identity), future-known time covariates (hour, day, week, month), gated variable selection, and attention mechanisms for long-range dependencies. Together, these allow TFT to (i) represent multi-scale temporal structure, combining short-term autocorrelation with diurnal and seasonal patterns; (ii) condition predictions on static ward context; and (iii) dynamically select and weight time and meteorological covariates. CO₂, by contrast, often shows event-like dynamics with abrupt spikes driven by occupancy and ventilation disturbances. This pattern is comparatively well suited to tree-based ensembles such as XGBoost, which can model sharp non-linear thresholds without smoothing them.

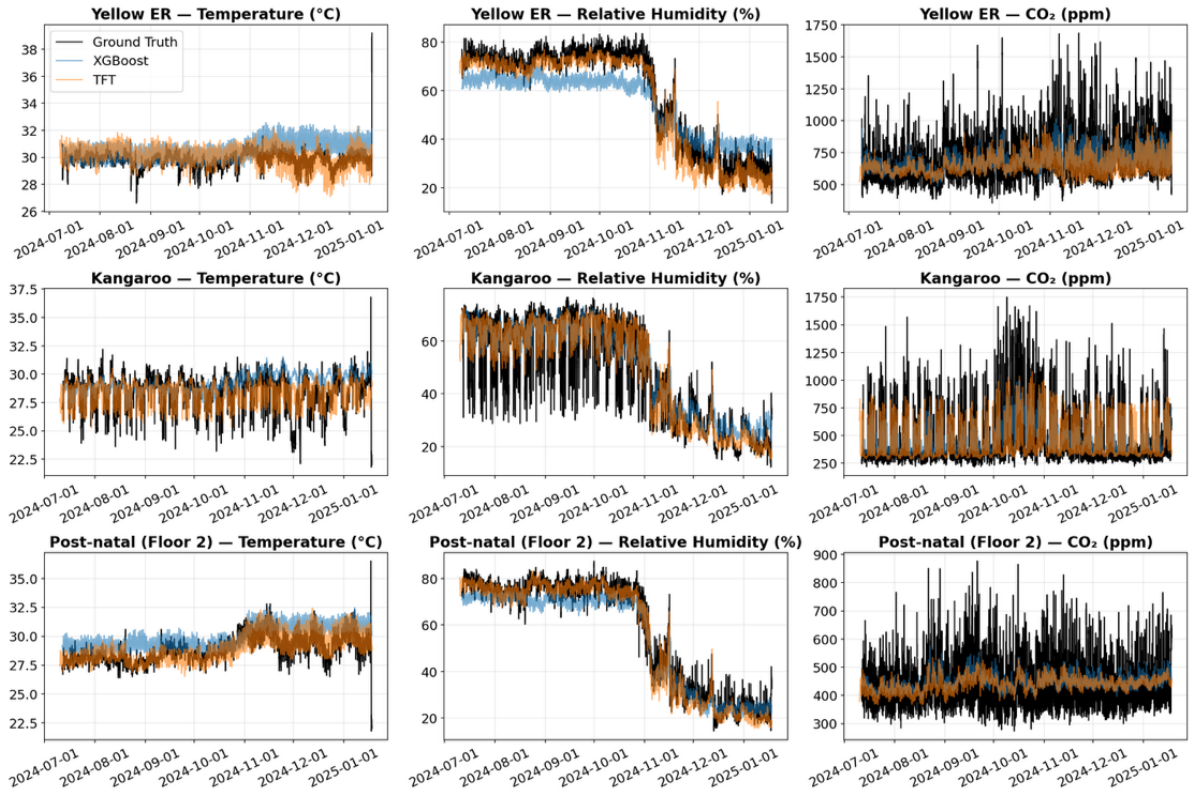


Figure 1: Multi-step forecasts by XGBoost and TFT against the ground truth for three representative wards (Yellow ER, Kangaroo, and Post-natal) for temperature (°C), relative humidity (%), and CO₂ (ppm).

Beyond the aggregate metrics in Table 5, Figure 1 shows that the TFT forecasts for temperature and RH track the ground-truth baseline level and longer-term drift more closely than XGBoost, particularly during periods of gradual change, while XGBoost more frequently exhibits bias when the trend shifts. For CO₂, both models capture the general time-series trend but under-represent the magnitude and frequency of sharp spikes, indicating that peak-driven dynamics remain insufficiently captured by the current feature set.

4 DISCUSSION

The risk summary in Table 4 indicates a distinct dual profile across the naturally ventilated tropical wards studied. Overheating is widespread and strongly diurnal, relative humidity is similarly prevalent but more temporally variable, and CO₂ exceedance is concentrated in a limited number of neonatal and paediatric spaces. This pattern is consistent with the forecasting results in Table 5, where temperature and relative humidity benefit from the smoother temporal structure captured by TFT, whereas CO₂ remains more difficult to predict, with XGBoost achieving a lower absolute error while the two models perform comparably in relative terms. Together, these findings point to two operational regimes. Heat and moisture call for anticipatory action across the whole facility, where night flushing and cross-ventilation are initiated the evening before a forecast hot day rather than in response to reported discomfort. CO₂ calls for something narrower, namely threshold-triggered responses in a small number of wards, such as staged louver opening or moving new admissions away from spaces forecast to exceed. The persistent underestimation of CO₂ spikes further implies that short-term occupancy and airflow disturbances are not adequately represented in the current feature set, and that future gains are more likely to arise from richer occupancy or ventilation proxies than from additional model complexity alone.

5 CONCLUSION

Using two years of 15-minute interval sensor measurements across 14 wards in two Ghanaian hospitals, this study demonstrates a multi-step forecasting framework for early warning of indoor environmental risks in naturally ventilated tropical hospitals. The risk summary indicates chronic overheating and frequent relative humidity non-compliance across most wards, whereas CO₂ exceedance is concentrated in a limited number of neonatal and paediatric spaces. TFT outperforms XGBoost for temperature and relative humidity forecasting, while XGBoost is more effective for CO₂ forecasting. The results suggest that future improvements are more likely to arise from richer occupancy and ventilation proxies and from peak- or exceedance-aware training objectives than from further increases in model complexity alone, thereby supporting anticipatory rather than reactive ventilation control in low-resource hospital settings, where operable façade elements and portable fans are the only available means of managing indoor environmental risk.

REFERENCES

- ASHRAE (2021). ANSI/ASHRAE/ASHE Standard 170-2021: Ventilation of Health Care Facilities. American Society of Heating, Refrigerating and Air-Conditioning Engineers.
- Boutahri, Y. and Tilioua, A. (2024). Machine learning-based predictive model for thermal comfort and energy optimization in smart buildings. *Results in Engineering*, 22, 102148.
- Chen, T. and Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.
- CIBSE (2020). CIBSE TM40: Health and Wellbeing in Building Services. Chartered Institution of Building Services Engineers.
- ECMWF (2024). ERA5 hourly data on single levels from 1940 to present. European Centre for Medium-Range Weather Forecasts, Copernicus Climate Change Service.
- Godasiaei, S.H., Ejohwomu, O.A., Zhong, H. and Booker, D. (2025). Integrating experimental analysis and machine learning for enhancing energy efficiency and indoor air quality in educational buildings. *Building and Environment*, 276, 112874.

- Kenny, G.P., Tetzlaff, E.J., Journeay, W.S., Henderson, S.B. and O'Connor, F.K. (2024). Indoor overheating: a review of vulnerabilities, causes, and strategies to prevent adverse human health outcomes during extreme heat events. *Temperature*, 11(3), 203–246.
- Lim, B., Arik, S.Ö., Loeff, N. and Pfister, T. (2021). Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4), 1748–1764.
- Roberts, B.M., Kasei, R., Codjoe, S.N., Amankwaa, E.F., Gough, K.V., Abdullah, K., Mensah, P. and Lomas, K.J. (2022). Comparing indoor air quality in naturally ventilated and air-conditioned hospitals in the tropics. *Building and Environment*, 218, 109138.
- Roberts, B.M., Adzic, F., Hamilton-Smith, A., Iddon, C., Wild, O., Cook, M., Gwynne, S., Hunt, A. and Malki-Epshtein, L. (2025). Air quality at mass gatherings: assessing ventilation and occupancy in marquees to evaluate airborne infection risk during the COVID-19 pandemic. *Building and Environment*, 113847.
- Tekler, Z.D., Lei, Y., Peng, Y., Miller, C. and Chong, A. (2023). A hybrid active learning framework for personal thermal comfort models. *Building and Environment*, 234, p.110148.
- Tekler, Z.D., Lei, Y. and Chong, A. (2024). Data-efficient comfort modeling: Active transfer learning for predicting personal thermal comfort using limited data. *Energy and Buildings*, 319, 114507.
- WHO (2009). *Natural Ventilation for Infection Control in Health-Care Settings*. Geneva: World Health Organization.
- WHO (2021). *Roadmap to Improve and Ensure Good Indoor Ventilation in the Context of COVID-19*. Geneva: World Health Organization.