Proceedings of the 4th International Conference on Functional Materials and Civil Engineering
DOI: 10.54254/2755-2721/2026.38639

Dynamic Environmental Risk Identification for Historic Buildings Based on Multi-Source Sensing Digital Twins

Yinran Du

Department of Landscape Architecture, University of Sheffield, Sheffield, UK
duyinran02580@outlook.com

Abstract. Buildings with heritage suffer from various disturbances like hygrothermal disturbances, moisture, vibration, and crack problems which cannot be easily identified from sensor readings independently. In this research, a multi-source sensing-based digital twin is developed for identifying environmental risks dynamically in a monitored two-story building made of bricks and timber structure. Variables such as temperature, relative humidity, wall moisture, vibration, crack displacement, and weather data outside are synchronized every five minutes and mapped into IFC-based Heritage Building Information Model in Unity. With the Temporal Fusion Transformer, a multivariate history with 24 time steps is taken into account, and normal, hygrothermal instability, moisture infiltration, vibration abnormalities, and crack displacement risk predictions are done. The architecture includes class weights, adaptive temporal consistency, and spatial regularization of components. With temporal testing, macro-F1 of 0.891±0.012, AUROC of 0.962±0.006, and mean warning delay of 6.7±1.1 min were obtained which is better than LSTM and XGBoost models with $p<0.001$. Removal of moisture sensors yielded maximum drop of macro-F1 to 0.847±0.014. Digital twin localization error was 0.43±0.17 m.

Keywords: historic buildings, digital twin, environmental risk, multi-source sensing, Temporal Fusion Transformer

1. Introduction

The degradation of historic buildings takes place due to environmental and structural degradation processes, combined and complex, rather than extreme measurement. Cycling of temperature alters the expansion of materials and moisture balance; prolonged relative humidity encourages condensation and bioactivity; water ingress increases masonry moisture; vibrations result from traffic or construction operations; and accumulated movements might eventually lead to measurable displacement of cracks. Inspection can document the visible degradation, but it would be impossible to trace temporary excursions or delayed interactions between inspections. New approaches based on research suggest using a digital twin as an essential instrument for prevention and conservation of heritage assets as it unites the physical object with continuously updated observations and a geospatially aware digital representation [1]. Ontologies for sensor-aware heritage highlight the fact that a digital twin needs to represent not only the geometry of the building but also sensors, the

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.38639

context of their observations, and relationships between the measurements and physical components [2].

HBIM is the necessary tool for associating monitoring data with walls, floors, timber elements, openings, and degradation information. HBIM-based workflows for heritage conservation proved that it was useful to include intervention and condition information scattered throughout the data in the same model [3]. At the sensing level, non-destructive moisture evaluation confirmed that geospatially-aware measurements allow revealing information about internal moisture conditions not captured by surface observations [4].

The next problem is dynamic risk detection. A static threshold allows detecting an extreme value, but does not provide a possibility to separate seasonality and other changes in the dynamics of measurements associated with compound degradation, such as prolonged humidity followed by increased wall moisture or vibrations combined with crack movement. Therefore, the purpose of this study is to create a multi-source sensing digital twin in which heterogeneous measurements are synchronized, interpreted by Temporal Fusion Transformer (TFT), and delivered to specific HBIM components as continuously updated risk states.

2. Literature review

2.1. Environmental monitoring of historic buildings

Environmental monitoring nowadays frequently includes measurements of micro-climate, material moisture levels, and structural responses since deterioration does not usually occur based on a single parameter. The fuzzy environmental monitoring of protected archaeological sites has illustrated the importance of interpreting the seasonal and interactional aspects of the environment, rather than focusing only on the threshold values [5]. Digital twin technology-based monitoring of historic buildings has similarly been able to show that changing structure states can be pinpointed using physical sensors and their consecutive digital twins [6]. Thus, continuous monitoring offers a way to identify both sudden changes and deterioration processes, but this requires synchronization of measurement channels.

2.2. Digital twins for heritage conservation

A heritage digital twin builds on HBIM technology by combining geometric-semantic models and real-time measurements, simulation results, and conservation actions. Building level studies have already combined HBIM, environmental sensing data, and numerical modeling to assess thermal and deformation effects [7], whereas digital twin-based structural inspections have enabled defect localization at the component level [8]. The outstanding challenge, however, is to sustain the operations: many heritage digital twins remain mere documentation or inspection tools and not real-time risk assessment systems.

2.3. Temporal risk identification

Predictive maintenance studies focus more and more on the integration of IoT sensors and digital twins to achieve real-time condition monitoring of cultural heritage assets [9]. The attention mechanism used in time series models is useful since the significance of humidity, moisture, vibration, weather conditions, and displacement varies over time. Applications of the TFT model

DOI: 10.54254/2755-2721/2026.38639

using interpretability of the attention mechanism on building time series show that this kind of approach is able to model nonlinear relationships while still providing explanation [10].

3. Methodology

3.1. Multi-source sensing network

The data from a historical building made of brick and timber with two floors was collected during 12 weeks. The SHT35 nodes sensed the temperature and humidity indoors once a minute, while the capacitive sensors measured moisture in the walls with an interval of five minutes. Moreover, the triaxial accelerometers ADXL355 collected vibration data once every twenty seconds, whereas the LVDTs registered crack movement once a tenth of a second. An outdoor station provided data on temperature, humidity, rain, wind speed, and barometric pressure. All the sensors monitored either the exterior walls, timber floor zones, moist zones, or cracks without invasive connection to building materials.

3.2. Digital twin construction

The measured drawings and terrestrial laser scans were then aligned in Autodesk Revit 2025. All walls, timber components, floor slabs, windows, roofing areas, and deterioration points got a unique component ID number. The IFC4 file was then imported into Unity 2022 LTS where the sensor IDs were assigned to semantically valid components and not simply to the coordinates. The raw and processed data was kept in InfluxDB 2.7, while MQTT updated data every five minutes. All components have information on materials, exposure direction, sensors attached, latest risk vector, and warnings issued.

3.3. Multi-source data processing

All channels were synced to five-minute time stamps using Python 3.11. The environmental and displacement channels were summed directly, where RMS acceleration, peak acceleration, crest factor, spectral centroid, and dominant frequency were computed on each vibration channel using FFT. In order to capture short transient events despite the fusion of data over five minutes, vibration features were first extracted from sub-windows of 10 seconds and then aggregated using their median and 95th percentile. Windows of two hours were comprised of 24 synchronized samples. Features consisted of average, extreme values, standard deviation, rate of change, trend slope, duration of exceedance, indoor/outdoor temperature gap, humidity/moisture delay, crack velocity, and vibration displacement synchronicity. Shorter gaps were filled using interpolation while longer gaps were masked out.

3.4. Dynamic risk identification model

The TFT used an embedding dimension size of 64, four attention heads, dropout rate of 0.20, and a 24-step encoder. Descriptors for static components went into the static covariate encoder, and the sensor data was filtered using gated variable selection networks. The LSTM layer modeled the sequence ordering locally, while the multi-head attention layer modeled the sequence ordering globally.

To address rare classes while avoiding unstable frame-to-frame warnings, the objective combined weighted cross-entropy with adaptive temporal and component-neighborhood regularization as

DOI: 10.54254/2755-2721/2026.38639

shown in Equation (1) - (2):

$$\mathscr{L} = -\frac{1}{N}\sum_{n=1}^{N}\sum_{c=0}^{4} w_c y_{n,c} log\widehat{p}_{n,c} + \lambda_t \frac{1}{N-1}\sum_{n=2}^{N}\omega_n \|\widehat{p}_n - \widehat{p}_{n-1}\|_2^2 \quad (1)$$

$$+\lambda_s \frac{1}{|\mathscr{E}|}\sum_{(i,j)\in\mathscr{E}} \kappa_{ij}\|\widehat{p}_i - \widehat{p}_j\|_2^2$$

$$\omega_n = exp[-\eta\|\boldsymbol{x}_n - \boldsymbol{x}_{n-1}\|_2] \quad (2)$$

Here, w_c is the inverse-frequency class weight; $y_{n,c}$ and $\widehat{p}_{n,c}$ are the target and predicted probability; $\boldsymbol{x}_n$ is the synchronized sensor state; $\mathscr{E}$ contains physically adjacent components; and κ_{ij} represents material continuity. The adaptive factor ω_n relaxes smoothing during genuine rapid changes, while the spatial term suppresses isolated warnings in continuous masonry zones. The coefficients λ_t and λ_s were selected by validation macro-F1. Training used AdamW with an initial learning rate of 0.001, batch size of 64, gradient clipping at 1.0, and early stopping after 20 epochs without validation improvement (see Figure 1).

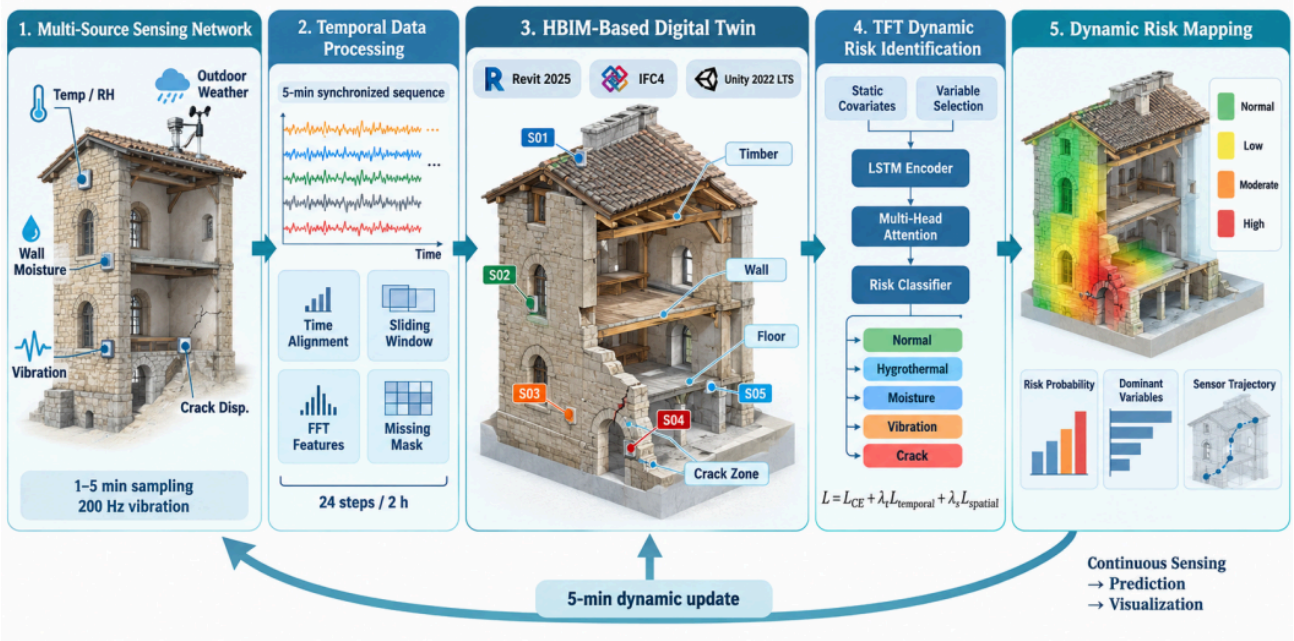

Figure 1. Multi-source sensing digital twin framework for dynamic environmental risk identification

3.5. Risk mapping and dynamic updating

Every five minutes, the probability vector was saved back into the Unity component associated with it. The maximum non-normal probability values established four categories of values, where anything less than 0.30 is considered normal, while 0.30 to 0.50 is at low risk; 0.50 to 0.70 is at medium risk and 0.70 and above at high risk. The Unity showed the state of the component along with its two-hour trajectory and top three selection weights in TFT variables.

DOI: 10.54254/2755-2721/2026.38639

4. Experimental design and validation

4.1. Field monitoring and risk annotation

The first 8 weeks were used for training data, weeks 9 to 10 for validation data, and weeks 11 to 12 for the test data set. A total of 24,192 synchronized 5-minute timestamps were available in the 12-week window prior to any gap filtration. The risk labeling process involved exceedance time duration, conservation check, presence of wetness, crack detection, vibration-event data, and analysis by conservation engineers. Two individuals performed the labeling for candidate risks independently, and discrepancies were sorted out via inspection of sensor and field data together.

4.2. Controlled perturbation experiment

Three non-destructive perturbations were used to measure response latency. The humidity rise was accomplished using a programmable humidifier in a localized manner without wetting any of the historic surfaces; the vibration was induced using an instrumented exciter at a low amplitude level within the allowable conservation range; and the LVDT displacement was generated using a micrometer device without changing the crack geometry. All three perturbations were performed five times each in a manner that involved a 30-min baseline stage, input stage, and recovery stage.

4.3. Comparative and robustness validation

TFT was compared to a two-layer LSTM, XGBoost, and rule-based threshold monitoring on the same temporal partitions. The macro-F1 was the primary metric, and precision, recall, AUROC, false alarm rate, and mean detection delay were the secondary metrics. Neural network models were trained on ten random seeds. One thousand bootstrap samples yielded 95% confidence intervals, and Holm-corrected paired tests were applied to determine the statistical significance of model differences. For sensor ablation, one family of sensors at a time was eliminated.

5. Results

5.1. Dynamic environmental risk identification performance

TFT produced the strongest temporal-test performance, with macro-F1 of 0.891±0.012, precision of 0.903±0.011, recall of 0.884±0.015, and AUROC of 0.962±0.006. Relative to LSTM, the macro-F1 gain was 0.049±0.009 with a bootstrap 95% interval of 0.033–0.066 and adjusted $p<0.001$. Mean warning delay decreased from 9.8±1.5 min to 6.7±1.1 min (see Table 1).

Table 1. Temporal test performance of environmental risk identification models

Model	Macro-F1	Precision	Recall	AUROC	False-Alarm Rate	Detection Delay (min)
Threshold Monitoring	0.671±0.026	0.704±0.024	0.648±0.031	0.806±0.018	0.142±0.016	15.6±2.7
XGBoost	0.806±0.018	0.821±0.017	0.798±0.021	0.914±0.011	0.091±0.012	10.9±1.8
LSTM	0.842±0.015	0.857±0.014	0.832±0.019	0.936±0.009	0.073±0.010	9.8±1.5
TFT	0.891±0.012*	0.903±0.011*	0.884±0.015*	0.962±0.006*	0.051±0.008*	6.7±1.1*

** $p<0.001$ versus LSTM after Holm correction.*

DOI: 10.54254/2755-2721/2026.38639

5.2. Multi-source sensor contribution and robustness

Moisture removal caused the largest macro-F1 reduction to 0.847±0.014, followed by vibration removal at 0.854±0.015 and crack-displacement removal at 0.862±0.013. Removing indoor temperature-relative-humidity sensing yielded 0.868±0.012, whereas outdoor-weather removal produced 0.879±0.011. Under 10%, 20%, and 30% simulated dropout, macro-F1 remained 0.879±0.013, 0.861±0.016, and 0.833±0.019, respectively.

5.3. Digital twin risk localization and warning response

Across verified events, component localization error was 0.43±0.17 m and database-to-Unity synchronization lag was 4.8±1.2 s. Controlled humidity perturbation generated a warning after 8.6±1.4 min, vibration after 3.9±0.7 min. The slower hygrothermal recovery corresponded to persistent moisture-related features rather than interface latency.

6. Conclusion

The digital twin for a brick and timber structure based on multi-sensor data for dynamic risk detection was created. Binding of heterogenous sensor signals to elements of HBIM along with the modeling of temporal dependencies using the TFT framework allowed the identification of hygrothermal, moisture, vibration, and crack-displacement risks while maintaining interpretability at the level of elements. Temporal experiments have demonstrated superior performance in terms of macro-F1 and AUROC as well as warning delay compared to LSTM, XGBoost, and thresholding methods. The ablation study has shown that the moisture and vibration channels were the most influential.

References

[1] Vuoto, A., Funari, M. F., & Lourenço, P. B. (2024). Shaping digital twin concept for built cultural heritage conservation: A systematic literature review. *International Journal of Architectural Heritage, 18*(11), 1762–1795.
[2] Niccolucci, F., & Felicetti, A. (2024). Digital twin sensors in cultural heritage ontology applications. *Sensors, 24*(12), 3978.
[3] Malovrh Rebec, K., Deanovič, B., & Oostwegel, L. (2022). Old buildings need new ideas: Holistic integration of conservation-restoration process data using Heritage Building Information Modelling. *Journal of Cultural Heritage, 55*, 30–42.
[4] Muradov, M., Kot, P., Markiewicz, J., Łapiński, S., Tobiasz, A., Onisk, K., Shaw, A., Hashim, K., Zawieska, D., & Mohi-Ud-Din, G. (2022). Non-destructive system for in-wall moisture assessment of cultural heritage buildings. *Measurement, 203*, 111930.
[5] Cabello-Briones, C., Verichev, K., & Prieto, A. J. (2022). Digital fuzzy control system for monitoring environmental conditions and technical suitability of shelters at archaeological sites. *Journal of Cultural Heritage, 57*, 226–234.
[6] Kong, X., & Hucks, R. G. (2023). Preserving our heritage: A photogrammetry-based digital twin framework for monitoring deteriorations of historic structures. *Automation in Construction, 152*, 104928.
[7] Cheng, J. C. P., Zhang, J., Kwok, H. H. L., & Tong, J. C. K. (2024). Thermal performance improvement for residential heritage building preservation based on digital twins. *Journal of Building Engineering, 82*, 108283.
[8] Yiğit, A. Y., & Uysal, M. (2024). Automatic crack detection and structural inspection of cultural heritage buildings using UAV photogrammetry and digital twin technology. *Journal of Building Engineering, 94*, 109952.
[9] Cecere, L., Colace, F., Lorusso, A., Messina, B., Tucker, A., & Santaniello, D. (2024). IoT and digital twin: A new perspective for cultural heritage predictive maintenance. *Procedia Structural Integrity, 64*, 2181–2188.
[10] Zheng, P., Zhou, H., Liu, J., & Nakanishi, Y. (2023). Interpretable building energy consumption forecasting using spectral clustering algorithm and temporal fusion transformers architecture. *Applied Energy, 349*, 121607.