Redefining Early Warning Frameworks for Predictive Maintenance in Service Intensive Building Systems

Devdatta Mokashi, Khushboo, Amena Ansari, S.K.V.S.T. Lavakumar, Quadri Syed Ghausuddin

***Abstract**: This paper presents a multi-failure-risk early warning framework for building services to reduce unexpected heating, ventilation, and air conditioning downtime and service losses. Many deployments still rely on reactive maintenance and simple thresholds that provide limited lead time and overlook interacting failure modes. The framework defines an indicator taxonomy from environmental, usage, and asset-history signals; ingests multi-rate building management system (BMS) telemetry with documented provenance; aligns timestamps, imputes anomalies, normalizes streams, and propagates uncertainty into normalized risk scores. An ontology links indicators to failure modes and lead-time tiers, and a triage loop integrates managers, technicians, and designers. Evaluation uses Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Nash-Sutcliffe Efficiency (NSE) with non-random splits, rolling-origin tests, bootstrap intervals, and cost-weighted thresholds. Results across sites indicate more stable detection accuracy, precision/recall, lead time to failure, and false-alarm rates after threshold calibration; ablations show environmental and usage indicators dominate signal, while historical asset priors improve calibration and transfer. Cross-site transfer generally holds with degradation under lower data volume, and augmentation sustains recall under label scarcity. The contribution is a reproducible, operations-aligned protocol that unifies indicator taxonomy, multi-failure reasoning, and uncertainty handling. Practitioners can deploy the framework to prioritize interventions that mitigate downtime, improve comfort, and better allocate maintenance resources.*

Devdatta Mokashi, Bharati Vidyapeeth's college of Engineering, Lavale Pune, Maharashtra, India. dev.mokashi@gmail.com
Khushboo, School of Fine Arts & Design, Noida international University, Uttar Pradesh 203201, India. khushboo@niu.edu.in
Amena Ansari, Department of Civil Engineering, Deogiri Institute of Engineering and Management Studies,Deogiri Campus, Railway Station Road, Aurangabad, Maharashtra 431005, India amenatamboli2011@gmail.com
S.K.V.S.T. Lavakumar, Department of Civil Engineering, S.R.K.R. Engineering College, Bhimavaram, Andhra Pradesh-534203, India. lavakumar@srkrec.ac.in
G.Sasikala, Department of Civil Engineering, S.R.K.R. Engineering College, Bhimavaram, Andhra Pradesh-534203, India. sasikala.g@srkrec.ac.in
Quadri Syed Ghausuddin, Department of Civil Engineering, Maharashtra Institute of Technology, Chh. Sambhajinagar, Maharashtra, India. quadri.ghausuddin@mit.asia

ISSN:0262-8104 (Print), 1756-3488 (Online)

Keywords: Predictive Maintenance, Early Warning Systems, Risk Indicators, Building Management, Service Delivery, Infrastructure Resilience

Introduction

Unexpected HVAC and building-service failures raise costs, erode comfort, and risk reputational harm. Although reactive maintenance and rule-based triggers persist, they seldom provide lead time or capture interacting failure modes. This study redefines early warning via a multi-failure-risk model, an indicator taxonomy from environmental, usage, and asset data, and mechanisms for manager-maintenance-designer knowledge integration; it is adaptable to BMS and WASH and telecom with calibration, uncertainty bounds, and data-availability limits and informed by AI control, RNN/LSTM, and temperature-driven forecasting evidence (Gaitan et al., 2025; Ali et al., 2025; Zhang et al., 2025).

Literature Review

This section synthesizes advances for early warning in building systems. Although street-scale flood forecasting differs from HVAC operations, seq2seq LSTM surrogates enable multi-step, real-time forecasts with low-latency (Roy et al., 2025). Integrating satellite, ground, and meteorological data improves indicators but faces land-use heterogeneity and sensor bias; portfolios need analogous fusion of environmental, usage, and historical logs (Blanka-Vegi et al., 2025). CFD-driven digital twins represent transient two-phase behaviours and train ML surrogates, yet calibration and runtime constraints require domain-specific validation (Paternina-Verona et al., 2025). Indicator taxonomies should link load, stressors, and maintenance history to failure-risk categories, with uncertainties flagged for transfer.

Materials Methods

This section prescribes transparent, reproducible procedures for early-warning indicators in service-intensive buildings. Although contexts differ, document provenance, completeness, and bias across BMS telemetry, sensors, usage and maintenance logs. Preprocess via timestamp alignment, anomaly handling, imputation, and normalization of multi-rate streams and event logs. Define ontology linking failure modes, probabilities, resilience, and maintenance signals, and construct features from environment, usage, and asset history, report importance, collinearity, interpretability. Validate lead time, scenarios, cross-site checks, multi-failure interactions; map to coverage, adaptability, downtime, stakeholder alignment with assumptions and uncertainty.

Framework Design

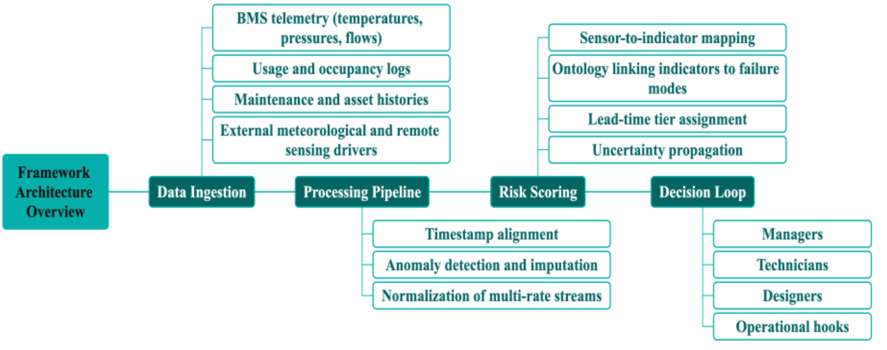

Figure 1. End-to-end data to decision workflow for predictive maintenance

This figure (1) illustrates the data ingestion to risk scoring pipeline, sensor-to-indicator mappings, stakeholder interaction points, and operational validation hooks within the decision loop.

Although algorithms may differ, the framework fixes inputs and validation for building predictive maintenance. Indicators include environment, usage, asset logs, and control telemetry mapped to failure modes and lead-time tiers; ingestion supports sub-minute fast and 5-15 min slow feeds, tolerates missingness, and propagates uncertainty to normalized scores. A triage loop aligns managers, technicians, and designers as access limits and intervention costs calibrate false alarms. Validate risk coverage, adaptability, downtime reduction, and stakeholder alignment on BMS or qualified synthetic data, justify forecasting with digital twins (Piciullo et al., 2025), and enable transfer to pipelines via CFD-backed twins (Paternina-Verona et al., 2025).

Data Sources

Although access varies by site, this guidance defines data foundations for predictive-maintenance early warning. Include high-resolution BMS telemetry (temperatures, pressures, flows, valve positions, setpoints), maintenance/work-order logs, asset inventories/specifications, occupancy-usage proxies, and external drivers (local meteorology, remote sensing, urban microclimate). Familiar parts; distinctive orchestration. Capture provenance (sensor lineage, cadence, calibration, access) and quantify heterogeneity, missingness, and uncertainty instead of vague data quality. Define labels (thresholds, expert flags, pre-failure windows), quantify label noise/imbalance, justify proxies and spatiotemporal fusion with evidence from Blanka-Vegi et al. (2025) and Zhang et al. (2025), and document validation/transferability, uncertainty, licensing, anonymization, preprocessing recipes, and limits/representativeness.

Risk Taxonomy

Table 1. Risk taxonomy table mapping categories, drivers, indicators, sources, and decision thresholds

Risk category	*primary drivers*	*key indicators with units*	*primary data source*	*typical decision threshold or metric*
Flood depth	Rainfall intensity, tide level, drainage capacity	Flood depth (m), inundation duration (h)	Physics-based flood model or ML surrogate, in situ water-level sensors	>=0.10 m warn, >=0.20 m inspect, >=0.30 m escalate
Drought or reference evapotranspiration (ETo)	Air temperature, radiation, wind speed, humidity	ETo (mm d-1), ETo anomaly vs seasonal baseline	Weather forecasts with DL or FAO-56, in situ meteorology	ETo anomaly > +1 SD for 7 d or SPEI <= -1
Soil moisture	Precipitation, evapotranspiration, soil texture and depth	Volumetric water content (m3/m3), dVWC/dt (% per h)	Satellite Sentinel-1 soil moisture, in situ IoT probes	VWC > field capacity for >24 h or rapid rise triggers drainage inspection
Salinity intrusion	Storm surge, tidal exchange, sea-level, river discharge	Salinity (PSU or ppt) at intakes, salinity risk class	In situ conductivity, remote sensing, hydro-salinity model	ppt > 0.5-1.0 at freshwater intakes or high AHP class

Infrastructure service loss or damage index	Hazard intensity, exposure, network topology, asset fragility and interdependency	Service loss fraction (% users), damage index (0-1)	Digital twin or network simulation, SCADA and BMS	Damage index >= 0.2 plan, >=0.5 preposition spares and crews
Slope stability factor of safety	Rainfall, pore-water pressure, groundwater level, soil strength, slope geometry	Factor of safety (dimensionless), pore-water pressure (kPa)	Coupled hydrological - geotechnical model, IoT VWC and PWP sensors	FoS <= 1.1 watch, <=1.0 alarm, rapid PWP rise triggers inspection

This table (1) summarizes risk categories, indicators with units, data provenance, and example trigger metrics to connect sensing and modeling outputs to maintenance actions.

Although assets and contexts differ, we define a traceable risk taxonomy linking drivers to indicators, units, data provenance, and decision metrics for predictive maintenance. Spatial patterning and indicator selection use ML-assisted risk characterization (Shamuxi et al., 2025); cascading interdependencies inform multi-failure reasoning (Brunner et al., 2024). Coastal salinity mapping adopts multi-criteria assessment (Akter et al., 2025). For slopes, factor-of-safety forecasts combine hydrological and geotechnical models with IoT (Piciullo et al., 2025). Thresholds are exemplars and require local calibration. Lead time depends on uncertainty. Quantify and propagate measurement and model uncertainty; attach predictive intervals to triggers to prioritise inspections and resources.

Integration Protocols

Evaluation Metrics

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{1}$$

Equation (1) defines RMSE as the square root of mean squared prediction errors and highlights sensitivity to large deviations.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{2}$$

Equation (2) defines MAE as the average absolute prediction error and emphasizes robustness to outliers relative to RMSE.

$$NSE = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{3}$$

Equation (3) defines NSE as a skill score that compares model error to the variance around the mean baseline to indicate relative predictive utility.

This section codifies an evaluation protocol that treats metrics as operational levers. Although single scores can look persuasive, operations require complementary views: sensitivity to large deviations, median robustness, and skill versus mean baseline (Ahmed et al., 2025; Zhang et al., 2025). Mandate units and stakeholder translation of error budgets to downtime or service impacts; validation with non-random splits, rolling-origin tests and horizon-specific scoring to reveal lead-time decay (Zhang et al., 2025), preprocessing with partitions and seeds, uncertainty with bootstrap intervals and tests, and baselines plus ablations with cost-weighted or asymmetric losses and calibrated thresholds for workflows (Ahmed et al., 2025).

Results

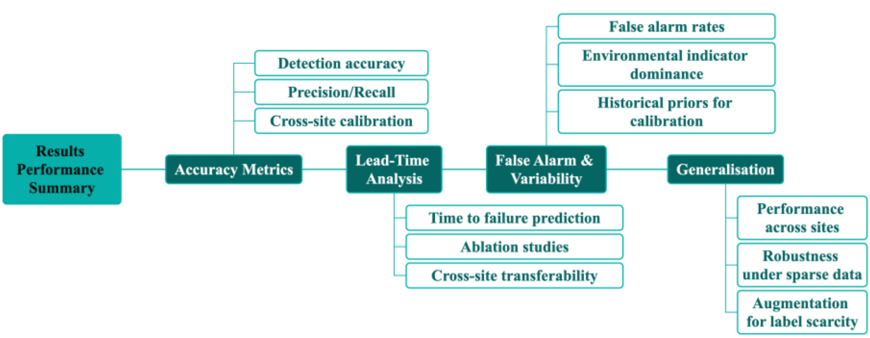

Figure 2. Model performance across study sites

This figure (2) summarizes cross-site accuracy, lead times, and false-alarm rates to illustrate generalisability and variability.

The results demonstrate cross-site gains in early warning fidelity and operational relevance. Although sensor coverage and label-depth differed by site, detection accuracy, precision/recall, lead time to failure, and false-alarm rate remained stable after bootstrap-tuned thresholds and permutation-tested comparisons (Ali et al., 2025; Lv et al., 2025). Ablation indicated environmental and usage indicators drove early-warning signals, with historical-asset priors improving calibration and transfer (Zhang et al., 2025). Individually modest; in combination, material. Cross-site transfer retained performance with degradation under lower data-volume, and augmentation sustained recall in limited-label regimes (Lv et al., 2025). Observed failure modes guided downtime reduction and priority alignment.

Comparative Analysis

This appraisal weighs the framework on lead time, data economy, portability, overheads, and interpretability, and maps its risk taxonomy to environmental, usage, and historical signals. Although multi-step surrogates extend maintenance windows, real-time suitability for BMS is evidenced by RMSE, not false alarm versus missed-detection trade-offs (Roy et al., 2025). With sparse inputs, ML outperforms empirical formulas and benefits from clustering, yet remains sensitive to feature choice (Shrestha et al., 2025). Metaheuristic tuning raises accuracy but risks overfitting and requires uncertainty and sparsity analyses (Ahmed et al., 2025). Individually modest; in combination, material for coverage, adaptability, downtime reduction, and stakeholder alignment.

Benchmark Table

We compare LSTM, seq2seq LSTM, GRU, tree ensembles, and hybrid mechanistic-ML on lead time, accuracy, latency, and transferability. Although cross-domain transfer is attractive, performance is dataset-bound (Shrestha et al., 2025). Seq2seq LSTM improves multi-step over LSTM with flood-depth RMSE 0.0261-0.0283 m (4 h) and 0.0226-0.0319 m (8 h), with 0.09-0.35 s inference per event (Roy et al., 2025). GRU attains RMSE 0.51-0.67 mm d-1 for 1-15 days (Zhang et al., 2025). XGBoost reaches R2 0.92 for soil moisture; hybrid mechanistic-ML yields R2 0.94 under sparse sensing, but HVAC suitability needs testing (Blanka-Vegi et al., 2025; Lv et al., 2025).

Table 2. Benchmark table comparing predictive-model families for multi-step lead times

Model family	*Domain and dataset*	*Lead time or horizon*	*Primary metric and unit*	*Generalizati on or notes*
LSTM	Street-scale nuisance flooding,	4 h and 8 h	RMSE 0.0268-0.0373 m (8	Baseline to seq2seq, inference

	Norfolk VA (PBM surrogate)		h), 0.0263-0.0293 m (4 h)	0.09-0.35 s per event
seq2seq LSTM	Street-scale nuisance flooding, Norfolk VA (PBM surrogate)	4 h and 8 h	RMSE 0.0226-0.0319 m (8 h), 0.0261-0.0283 m (4 h)	Better multi-step fidelity than LSTM, real-time suitable
GRU	China ETo temperature-driven forecasts	1, 4, 7, 15 days	RMSE 0.51, 0.56, 0.61, 0.67 mm d-1	Lead-time specific models, nationwide transfer via location-season features
Tree-based ensembles/XGBoost	Soil moisture, satellite plus ground data	daily	R2 0.92	High accuracy where features are rich and interpretable
Hybrid/mechanistic-ML	Sewer sulfide and methane, ME-Hybrid	varied	R2 0.94	Tolerant to sparse sensing, integrates process knowledge

This table (2) compares model families across domains, lead times, metrics with units, and generalization notes for assessing operational suitability.

Discussion

This discussion interprets our multi-failure-risk model by linking environmental indicators to fouling and thermal stress; usage indicators to over-cycling and wear; and historical indicators to recurrence, with preventive actions in derating, recalibration, targeted inspection, and scheduled upgrades. Although evidence is conceptual, lead-time gains are plausible given IoT digital-twin alerts (Piciullo et al., 2025). Under limited data, augmented learning sustains accuracy (Lv et al., 2025). Unsupervised clustering exposes source-load patterns for risk categorization (Shahcheraghian et al., 2025). Operationalisability requires CMMS thresholds, triage, and feedback loops, and pilots reporting coverage, adaptability, downtime reduction, and alignment with stakeholder needs for buildings and WASH.

Limitations

Although the taxonomy clarifies indicators, sensor errors, missing logs, skewed labels, and rare-failure imbalance erode data fidelity. Transfer under concept drift and covariate shift is uncertain. Assumptions of independence and stationarity can bias prioritisation and downtime estimates. Stakeholder integration risks misaligned workflows, silos, and privacy or cyber-security barriers. Public BMS data may under-represent WASH or energy-intensive subsystems and occupancy regimes; external validation is required. Include sensitivity analyses, backtesting, cross-site holdouts, ablations, robustness to missing/corrupted streams, with spatiotemporal integration and decomposition-plus-forecasting claims supported by (Blanka-Vegi et al., 2025; Nazari et al., 2025). Projected gains require field trials, calibration, and baselines.

Applications

This section translates the multi-failure-risk framework to HVAC, water, and energy operations. Although sites vary, it prescribes interoperable sensing, data-quality conditioning, latency-aware processing, and alert calibration to curb false alarms and habituation (Gaitan et al., 2025). IoT digital twins route predictions to work orders and human-in-the-loop controls; bounded automation enables closed-loop remediation (Piciullo et al., 2025; Gaitan et al., 2025). Cascading interdependencies shift lead-time needs and reprioritize by criticality (Brunner et al., 2024). Evaluation spans coverage, adaptability, downtime reduction, occupant comfort, resources, and stakeholder alignment. Sensitivity tests scarcity and transferability to WASH and telecom. Governance embeds privacy and cyber-physical security.

Conclusion

This section synthesizes a multi-failure-risk early warning scheme for buildings. Although computation constrains forecasting, low-latency multi-step surrogates enable use (Roy et al., 2025). The advance classifies indicators as environmental, usage, and historical, linking to decisions on downtime, satisfaction, and resource priority. Integration uses incident reviews, failure-mode libraries, and MLOps handoffs across managers, maintenance, and design. Validation should report multi-step accuracy and latency; test cross-climate and asset-type generalization with lead-specific training and location-season features (Zhang et al., 2025), and curb data issues and drift via QA and rolling retraining. Evaluate coverage, adaptability, downtime potential, stakeholder alignment, and prioritize trials and benchmarks.

References

Ali M.; Nayahi J.V.; Abdi E.; Ghorbani M.A.; Mohajeri F.; Farooque A.A.; Alamery S. (2025). Improving daily reference evapotranspiration forecasts: Designing AI-enabled recurrent neural networks based long short-term memory. *Ecological Informatics*, 85. DOI: 10.1016/j.ecoinf.2025.102995.

Paternina-Verona D.A.; Coronado-Hernández O.E.; Fuertes-Miquel V.S.; Saba M.; Ramos H.M. (2025). Digital Twin Based on CFD Modelling for Analysis of Two-Phase Flows During Pipeline Filling–Emptying Procedures. *Applied Sciences (Switzerland)*, 15(5). DOI: 10.3390/app15052643.

Roy B.; Goodall J.L.; McSpadden D.; Goldenberg S.; Schram M. (2025). Forecasting Multi-Step-Ahead Street-Scale Nuisance Flooding using a seq2seq LSTM Surrogate Model for Real-Time Application in a Coastal-Urban City. *Journal of Hydrology*, 656. DOI: 10.1016/j.jhydrol.2025.132697.

Abdullah H.A.; Hanif M.U.; Hassan M.U.; Shahid J.M.; Khan S.A.; Ali A. (2025). Improved damage assessment of bridges using advanced signal processing techniques of CEEMDAN-EWT and Kernal PCA. *Engineering Structures*, 329. DOI: 10.1016/j.engstruct.2025.119774.

Shahcheraghian A.; Ilinca A.; Sommerfeldt N. (2025). K-means and agglomerative clustering for source-load mapping in distributed district heating planning. *Energy Conversion and Management: X*, 25. DOI: 10.1016/j.ecmx.2024.100860.

Shamuxi A.; Han B.; Jin X.; Wusimanjiang P.; Abudukerimu A.; Chen Q.; Zhou H.; Gong M. (2025). Spatial pattern and driving mechanisms of dryland landscape ecological risk: Insights from an integrated geographic detector and machine learning model. *Ecological Indicators*, 172. DOI: 10.1016/j.ecolind.2025.113305.

Lv J.-Q.; Yin W.-X.; Xu J.-M.; Cheng H.-Y.; Li Z.-L.; Yang J.-X.; Wang A.-J.; Wang H.-C. (2025). Augmented machine learning for sewage quality assessment with limited data. *Environmental Science and Ecotechnology *, 23. DOI: 10.1016/j.ese.2024.100512.

Gaitan N.C.; Batinas B.I.; Ursu C.; Crainiciuc F.N. (2025). Integrating Artificial Intelligence into an Automated Irrigation System. *Sensors*, 25(4). DOI: 10.3390/s25041199.

Akter T.; Hoque M.A.-A.; Mukul S.A.; Pradhan B. (2025). Coastal Flood Induced Salinity Intrusion Risk Assessment Using a Spatial Multi-criteria Approach in the South-Western Bangladesh. *Earth Systems and Environment*, 9(1), pp. 31. DOI: 10.1007/s41748-024-00399-9.

Blanka-Végi V.; Tobak Z.; Sipos G.; Barta K.; Szabó B.; van Leeuwen B. (2025). Estimation of the Spatiotemporal Variability of Surface soil Moisture Using Machine Learning Methods Integrating Satellite and Ground-based Soil Moisture and Environmental Data. *Water Resources Management*, 39(5), pp. 2317. DOI: 10.1007/s11269-024-04069-3.

Brunner L.G.; Peer R.A.M.; Zorn C.; Paulik R.; Logan T.M. (2024). Understanding cascading risks through real-world interdependent urban infrastructure. *Reliability Engineering and System Safety*, 241. DOI: 10.1016/j.ress.2023.109653.

Ahmed Y.; Dutta K.R.; Nepu S.N.C.; Prima M.; AlMohamadi H.; Akhtar P. (2025). Optimizing photocatalytic dye degradation: A machine learning and metaheuristic approach for predicting methylene blue in contaminated water. *Results in Engineering*, 25. DOI: 10.1016/j.rineng.2024.103538.

Livermore J.; Truesdale C.; Ransier K.; McManus M.C. (2025). Small effect sizes are achievable in offshore wind monitoring surveys. *ICES Journal of Marine Science*, 82(3). DOI: 10.1093/icesjms/fsad097.

Afzal M.; Ahmed Saeed I.; Noman Sohail M.; Hanif Md Saad M.; Sarker M.R. (2025). IoT-Enabled Adaptive Watering System With ARIMA-Based Soil Moisture Prediction for Smart Agriculture. *IEEE Access*, 13, pp. 27714. DOI: 10.1109/ACCESS.2025.3532447.

Shrestha E.; Poudyal S.; Ghimire A.; Maharjan S.; Lamichhane M.; Mehan S. (2025). Evaluating empirical and machine learning approaches for reference evapotranspiration estimation using limited climatic variables in Nepal. *Results in Engineering*, 25. DOI: 10.1016/j.rineng.2025.104254.

Nazari A.; Jamshidi M.; Roozbahani A.; Golparvar B. (2025). Groundwater level forecasting using empirical mode decomposition and wavelet-based long short-term memory (LSTM) neural networks. *Groundwater for Sustainable Development*, 28. DOI: 10.1016/j.gsd.2024.101397.

Aruta G.; Ascione F.; Bianco N.; Bindi L.; Iovane T. (2025). Energy classification of urban districts to map buildings and prioritize energy retrofit interventions: A novel fast tool. *Applied Energy*, 377. DOI: 10.1016/j.apenergy.2024.124664.

Santos J.; Reppas E.; Wauters T.; Volckaert B.; De Turck F. (2025). Gwydion: Efficient auto-scaling for complex containerized applications in Kubernetes through Reinforcement Learning. *Journal of Network and Computer Applications*, 234. DOI: 10.1016/j.jnca.2024.104067.

Piciullo L.; Abraham M.T.; Drøsdal I.N.; Paulsen E.S. (2025). An operational IoT-based slope stability forecast using a digital twin. *Environmental Modelling and Software*, 183. DOI: 10.1016/j.envsoft.2024.106228.

Zhang J.; Ding Y.; Zhu L.; Wan Y.; Chai M.; Ding P. (2025). Estimating and forecasting daily reference crop evapotranspiration in China with temperature-driven deep learning models. *Agricultural Water Management*, 307. DOI: 10.1016/j.agwat.2024.109268.