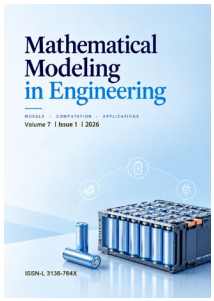
UNIVERSITY OF WEST ATTICA

eISSN: 3136-764X

Mathematical Modeling in Engineering

Vol. 7, No. 2, April, 2026



Physics-Constrained Feature Fusion for Photovoltaic Construction Progress Monitoring Using WaveletKAN-GRU

Veszprémi Ottó[a,*], Kozma Csaba[a]

[a] *Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, Budapest, 1111, Hungary*

Received 24 March 2026; Revised 08 June 2026; Accepted 13 June 2026; Published 17 June 2026

Abstract

Large ground-mounted photovoltaic power stations contain concurrent civil, mechanical, and electrical work fronts whose visible appearance does not always correspond to accepted engineering progress. This study proposes a WaveletKAN-GRU model with physics-constrained feature fusion for joint construction-stage recognition and continuous progress estimation. Multisource observations are represented by spectral, texture, geometric, schedule, and environmental descriptors. A soft fusion gate weights this evidence using terrain slope, row geometry, expected shadow behavior, installation order, irradiance consistency, and work-front capacity before temporal inference. Learnable wavelet mappings emphasize local frequency changes caused by emerging pile, bracket, and module rows, while a gated recurrent unit models irregular observation intervals and delayed dependencies between civil and electrical activities. Experiments use a simulated but engineering-realistic benchmark representing 18 project zones, six construction stages, irregular acquisition gaps, and terrain-dependent visibility disturbances; the site-independent split prevents observations from one zone from crossing training, validation, and test partitions. The proposed model achieved 94.8% stage accuracy, a macro F1-score of 0.936, a progress mean absolute error of 2.7 percentage points, a root mean square error of 3.8 percentage points, and a physical violation rate of 1.9%. The closest KAN-GRU baseline achieved 92.9% accuracy and a 3.8-point mean absolute error. Ablation results show complementary roles for wavelet mapping and physically constrained fusion. These controlled results support the method's potential for progress screening and targeted inspection, but independently acquired multi-project data are required before operational deployment can be established.

Keywords: Physics-Constrained Feature Fusion, Photovoltaic Construction Monitoring, WaveletKAN-GRU, Remote Sensing Time Series, Engineering Progress Estimation

1. Introduction

The construction of large-scale photovoltaic power plants has become a typical engineering project in areas that need to meet energy transition goals, expand the power grid and coordinate land resources simultaneously [1]. A project site is not a single object in optical images; rather, it is a dynamic system consisting of earthworks, access roads, pile foundations, brackets, panels, inverter rooms, cable trenches, substations and temporary storage areas [2]. Therefore, remote sensing and UAV inspection have been employed to reduce the need for manual patrols in areas such as deserts, plateaus, mining reclamation and mountainous regions [3]. The problem of monitoring is relatively more challenging than that of general land-cover recognition; for example, a barely visible bracket line may be due to a large-scale completed civil engineering project, but a dense panel texture can still be free of electrical facilities [4]. Traditional progress reports are delayed due to reporting cycles and site inaccessibility, and the resulting uncertainty spreads to procurement, grid-connection scheduling and financial settlement [5]. Optical time-series images have a regular

* Corresponding author: veszpremi.ot@vik.bme.hu

https://doi.org/10.68406/mme.2026.vol7iss2nm8:93-107

➢ This is an open access article under the CC BY-NC-SA 4.0 license (https://creativecommons.org/licenses/by/4.0/).

© Authors retain all copyrights.

observation path, but seasonal vegetation, dust, cloud shadows and changes in viewpoint can interfere with the apparent construction signal. Although deep convolutional and recurrent models have improved the representation of space and time, small repeated objects, missing observations, and irregular sampling can still lead to instability in the stage transition [6]. Therefore, the model for engineering progress monitoring should be able to understand visual dynamics while respecting the physical construction logic of photovoltaic stations.

The preceding research has offered some help for the current problem. To distinguish among bare soil, metal supports, glass-covered modules and road surfaces, Multispectral indices can be employed; Geometric descriptors can also be used to describe the regular row pattern of PV arrays [7]. Schedule-oriented learning adds temporal priors to restrict the prediction at a certain time point according to the feasible construction order [8]. A model based only on data may still produce physically unreasonable results, such as predicting the completion of module installation without obtaining sufficient pile foundation observation data. The other is the frequency distribution of construction pictures. The spatial frequency of the PV site changes abruptly at the appearance of dense rows, but this change is not always obvious in the raw reflectance values due to illumination and terrain shadows [9]. Wavelet analysis is appropriate for handling local high-frequency and low-frequency changes, and recently introduced kernel activation networks can perform flexible nonlinear transformations through learnable basis functions [10]. Gated Recurrent Units are computationally less expensive than Long Short-Term Memory networks, and they have also achieved good results for medium-length engineering sequences [11]. Based on the above observations, a hybrid model combining wavelet-enhanced nonlinear mapping, gated temporal memory and physical construction constraints has been proposed for a single monitoring system.

The methodological gap is the absence of a mechanism that reconciles visual, schedule, and physical evidence before temporal inference. Photovoltaic progress is constrained by terrain, row geometry, installation order, work-front capacity, and weather-related visibility, yet rigid rules cannot represent parallel work or delayed material delivery. This study addresses that gap through three contributions. First, a soft physics-constrained fusion gate adjusts visual evidence using measurable site and process descriptors instead of correcting predictions only after inference. Second, WaveletKAN mapping and GRU memory are combined to represent local row-frequency changes and irregular construction intervals while producing both stage probabilities and a continuous completion ratio. Third, the evaluation separates recognition quality from engineering feasibility through progress error, stage metrics, calibration, and physical-violation analysis. The primary prediction unit is a monitoring cell associated with a coherent row group or inverter branch; cell-level estimates can subsequently be aggregated to project zones. This scope prevents image appearance from being treated as equivalent to accepted engineering quantity and makes the intended operational use of the outputs explicit.

Based on the above, the four subsequent steps of this paper will systematically present evidence from engineering practice. Section 2 introduces progress monitoring methods for photovoltaic and infrastructure construction, and then presents temporal neural models and physics-informed feature fusion. Section 3 introduces the proposed WaveletKAN-GRU framework and explains the construction of the multi-source feature, physics-constrained fusion mechanism, wavelet kernel mapping, recurrent progress estimation, and training objective. Section 4 introduces the experimental data set, evaluation rules, comparison strategies, ablation configurations, and displays the detailed data analysis in six groups of figures and one quantitative table. Section 5 summarizes the main results, shortcomings and prospects for future studies in the paper.

2. Related Work

2.1. Photovoltaic Construction Progress Monitoring

Progress monitoring of photovoltaic construction has gradually moved from document-based reporting to observation-driven measurement. The first engineering management systems were based on weekly logs, contractor reports and manually collected pictures; they were suitable for small sites but not for large-scale projects with many work areas. Remote sensing has provided a broader view of the land-clearing, access-road construction and expansion of panel-array areas [12]. UAV imagery further improved the spatial resolution to observe the quality of row-by-row installations and temporary storage areas, as well as local rework sites [13]. Overhead images and project records have been used to improve the objectivity of progress monitoring, but interpretation is still challenging because visual evidence and engineering completion are not the same. A mounted support frame may be close to completion in the civil stage, but an apparently complete module row may still need cabling, inverter connection and commissioning. Therefore, the advancement observation cannot be reduced to image recognition alone.

Feature engineering has been applied to address the problem above. Spectral indices indicate alterations in the surface material; texture operators are regular arrangements of panel rows; object-level segmentation partitions roads, foundations, brackets, panels and buildings. Ekanayake et al. [14] applied deep learning-based computer vision to automate indoor construction progress monitoring. Some research has added schedule constraints to determine if the predicted stage can be realized according to the engineering order; however, these constraints are generally added after

the model predicts the features, not during feature learning. A post-processing rule can correct the obvious inconsistency, but it does not provide a reason for the inconsistency for the network. Site-specific calibration is another deficiency. A model trained on a single project may have overfitted row spacing, panel tilt, module color or soil background, especially if the training data were collected at fixed intervals [15]. The problems above indicate that the accuracy of monitoring is also affected by whether the physical construction relationship has been learned during training.

Other than that, there are many visually similar but functionally different modes of photovoltaic construction. Stored modules, partially installed modules, and complete module arrays can all have the same high-reflectance surface at an oblique angle to the observer. Cable trenches are not aesthetically pleasing but functionally necessary; therefore, the completion of the inverter room may not be observable at the same spatial scale as panel installation. Therefore, the monitoring model needs to integrate object-level evidence and process-level evidence. Object-level evidence refers to what can be observed in a photograph, and process-level evidence shows whether the visible objects are reasonable given the prior construction conditions, available Workfront, and physical conditions of the site. The two kinds of evidence are used together to provide support for the research.

2.2. Temporal Neural Models and Physics-Guided Fusion

Temporal Neural Networks are now frequently used to address the sequential nature of construction in progress prediction. Convolutional Recurrent Networks map each observation to a spatial feature vector, and then use LSTM or GRU units for state evolution modelling [16]. Temporal Convolutional Networks can be trained in parallel and have a stable receptive field, and attention-based models can address long-term dependencies in irregular observations [17]. The above methods require a large number of labelled sequences and frequently treat visual features as abstract signals. The visible difference between two adjacent observations in photovoltaic construction may be small, but the engineering significance of this change can be considerable. A faint new line may be the starting point for the row foundation of an area. A model with only a general shape cannot show the local variation of frequency.

Wavelet-based representation can show the local fluctuations of the data at different scales. The representation should separate slowly varying backgrounds from sharp construction edges and repetitive row structures. Kernel activation networks are also related; they replace fixed scalar activation functions with learnable nonlinear mappings to model complex feature interactions under fewer rigid assumptions [18]. Wavelet kernels are added to a KAN-style layer, and thus both amplitude and scale variations in the construction features can be modelled. Another way is to introduce domain constraints in the loss function, fusion gate or latent variable for physics-guided learning. Physics-informed models have been used in infrastructure monitoring to improve generalization under conditions of limited or noisy data [19]. However, the integration of wavelet nonlinear mapping, gated temporal memory and construction-physics constraints in photovoltaic progress monitoring is still lacking. WaveletKAN-GRU has been developed to address this issue by adding physical feasibility at the feature-fusion stage rather than applying it only as a final correction.

Temporal-model selection also affects deployment feasibility. Attention-based architectures can represent long dependencies but typically require more labeled sequences and stable observation histories than short construction projects provide. A GRU offers a compact recurrent core for sequences that extend from land preparation to commissioning within several months. Because a GRU alone cannot determine whether a visual change violates engineering feasibility, the proposed architecture regulates the input representation through constraint-aware fusion and regulates training through a physical-consistency term. This division keeps temporal memory separate from domain reasoning while allowing both components to interact through the fused feature state.

3. Physics-Constrained WaveletKAN -GRU Method

3.1. Multisource Feature Construction and Monitoring Workflow

The workflow begins with co-located UAV orthophotos, high-resolution satellite imagery, daily construction records, terrain models, and weather data. The six stages of the engineering process in each observation are: land preparation, pile foundation, bracket installation, module mounting, electrical connection and commissioning readiness. The input sequence is not assumed to be perfectly regular. Instead, the time between observations is used as a temporal feature because the construction progress may increase after material delivery or slow down due to the wait for grid connection. Spectral Features: Normalized Brightness, Red-Edge Contrast and Glass-like Reflection Intensity. Local Binary Patterns, Grey-Level Co-occurrence Statistics and Wavelet Energy Descriptors are all texture features. Geometric Features are: Row Spacing, Row Orientation, Panel-Block Compactness and Access-Road Density. Physical features are terrain slope, expected shadow length, irradiance consistency and feasible installation order. Continuous descriptors are standardized with training-set statistics, categorical order indicators are encoded separately, and missingness flags are appended rather than replaced by interpolated visual evidence.

The overall workflow is represented by Figure 1, where multisource observations pass through preprocessing, feature extraction, physics-constrained fusion, WaveletKAN nonlinear mapping, GRU temporal estimation, and dual-output

prediction. The figure also shows that physical constraints are not added after prediction; they participate in feature weighting before temporal modeling. This design is important for large PV stations because construction evidence is distributed across different objects and time intervals. A completed bracket zone may become visually similar to a module zone under strong glare, and a module zone may be partly hidden by dust or snow. By retaining physical descriptors in the fusion gate, the network receives a structured signal about which visual changes are feasible for the current site state.



Figure 1. Workflow for physics-constrained photovoltaic construction progress monitoring

Preprocessing comprises orthorectification, radiometric normalization, project-boundary clipping, and grid-based site partitioning. Homogeneous monitoring cells have been established at all the sites to protect the engineering Workfront, and arbitrary image tiles are not used. For example, a monitoring cell can cover a group of rows that are connected by the same access path or inverter branch. Partitioning reduces the ambiguity of labels by distributing the labels more uniformly among the cells instead of concentrating them in one area of the station. To align the times, each observation is mapped to the nearest date in the construction log, and the resulting time difference is saved as a feature. The missing images are not replaced by simple interpolation; instead, their absence is explicitly indicated, and the GRU learns that a long observation gap should increase the uncertainty of prediction. All image sources are resampled to a common project grid before cell extraction; each acquisition is matched to the nearest construction-log date, while the residual time offset and an explicit missing-observation indicator are retained as model inputs.

3.2. Physics-Constrained Fusion and WaveletKAN Mapping

Let x_t denote the visual feature vector at observation time t, p_t denote the physical constraint vector, and s_t denote schedule-related information. The fused feature z_t is computed through a constraint-aware gate that adjusts the contribution of each source according to construction feasibility. The gate is designed to be soft rather than binary, because real projects may contain parallel operations and temporary deviations from the planned sequence. In dimensional form, the visual, physical, and schedule inputs are represented as $\mathbf{v}_t \in \mathbb{R}^{d_v}$, $\mathbf{p}_t \in \mathbb{R}^{d_p}$, and $\mathbf{s}_t \in \mathbb{R}^{d_s}$; learned projections map them to a common fusion width before soft weighting.

The fused representation is defined as

$$z_t = g_t \odot W_x x_t + (1 - g_t) \odot \left(W_p p_t + W_s s_t\right), g_t = \sigma\left(W_g [x_t, p_t, s_t]\right) \quad (1)$$

where g_t is a learned fusion gate; W_x, W_p, and W_s are trainable projection matrices; $\sigma(\cdot)$ is the sigmoid mapping; and $\odot$ denotes element-wise multiplication. The term p_t includes terrain slope, shadow length, row-spacing deviation, irradiance consistency, and installation-order indicators. The schedule vector s_t contains elapsed days, planned milestone percentage, and recent work-front status. The gate output is bounded between zero and one for each projected channel, so physical and schedule evidence attenuates or preserves visual evidence without imposing a binary stage decision.

The components of the physical vector are all related to some engineering quantity. Changes in the slope of the ground will change both the expected row spacing and shadow overlap; thus, if there is a discrepancy in the row spacing, the detected repeating pattern will not meet the design geometry. According to the range of solar elevation and panel tilt, calculate the shadow length, and thus determine whether the dark module row is caused by a cloud shadow or ground obstruction. Installation-order indicators are in the form of soft feasibility values and not hard rules. Therefore, a pile-foundation cell can be used to receive a low but non-zero probability of bracket installation when the observation shows clear frame texture, and the model will not produce arbitrary stage jumps during accelerated parallel work.

WaveletKAN mapping is then applied to the fused representation. Unlike a conventional multilayer perceptron, the activation is expressed through learnable wavelet kernels so that local frequency changes in row structures can be amplified. This is useful when the monitoring signal is produced by dense repeated objects, such as foundation piles, support brackets, or aligned module edges. The mapping is written as The implementation selects the compact kernel count from three candidate widths on validation data only, preventing the test set from influencing wavelet capacity.

$$h_t^w = \sum_{k=1}^{K} c_k \psi_k \left(\frac{z_t - b_k}{a_k} \right), a_k > 0 \tag{2}$$

where $\psi_k(\cdot)$ is the k-th wavelet basis, a_k and b_k are learnable scale and translation parameters, and c_k is the corresponding activation coefficient. The function produces a nonlinear representation h_t^w that preserves both coarse progress trends and local construction discontinuities. The model uses compact wavelet bases to avoid excessive parameters, and the number of bases is selected according to validation accuracy and progress-error stability.

Wavelet kernels are used after feature fusion, not directly on the raw image. Therefore, the model will be relatively small and concentrate the wavelet response on construction-relevant descriptors. Low-frequency components are gradual changes in the surface of bare land to organized construction areas, and high-frequency components are the appearance of regularly spaced piles, brackets and module edges. As the basis functions are trainable, the model will not be limited to a fixed scale. This flexibility supports different row spacings, image resolutions, and observation heights while retaining an interpretable mapping structure.

3.3. GRU-Based Progress Estimation and Training Objective

The WaveletKAN output is passed into a GRU layer to describe temporal dependence. The update gate controls how much historical construction state is retained, while the reset gate determines whether outdated information should be weakened after major site changes. This is particularly relevant when a project shifts from civil works to module installation, because the visual features change from sparse line patterns to dense reflective arrays.

The recurrent update is defined as

$$\begin{aligned} u_t &= \sigma(W_u h_t^w + U_u q_{t-1}), \\ r_t &= \sigma(W_r h_t^w + U_r q_{t-1}), \\ q_t &= u_t \odot q_{t-1} + (1 - u_t) \odot q_t' \end{aligned} \tag{3}$$

where u_t and r_t are the update and reset gates, respectively; q_{t-1} is the preceding hidden state; and q_t' is the candidate hidden state. The resulting state q_t integrates the current WaveletKAN representation with temporally relevant construction information.

The hidden state q_t is used for two outputs: a discrete construction-stage probability y_t and a continuous progress ratio. A softmax classifier estimates the stage distribution, and a bounded regression head estimates the progress percentage between 0 and 100. The dual-output design reflects the fact that engineering users need both categorical status and quantitative completion values.

The recurrent state is initialized from the first valid observation rather than from a zero vector. This initialization is useful when a monitoring sequence begins after land preparation has already started. The model also receives an elapsed-time embedding at every step, which allows it to distinguish between a small visual change over five days and the same visual change after a twenty-day gap. During training, sequences are randomly masked to mimic missing observations, and the physical consistency term remains active under masking. This procedure encourages the model to rely on feasible state transitions rather than memorizing regular sampling intervals. Operationally, the first valid fused feature is

projected into the GRU state space to form the initial hidden state; any preceding padded positions remain masked and contribute neither a recurrent update nor a loss term.

The two prediction heads are expressed as

$$y_t = softmax(W_y q_t + b_y), \rho_t = 100\sigma(W_\rho q_t + b_\rho) \tag{4}$$

where y_t is the predicted probability distribution over the six construction stages and ρ_t is the estimated continuous progress percentage.

The training objective combines stage-classification loss, progress-regression loss, and a physical-consistency penalty. The penalty increases when predicted states violate the feasible construction order or when estimated progress changes faster than permitted by the observed work-front capacity. It does not force the model to follow a fixed schedule; instead, it discourages predictions that conflict with physical evidence.

The final loss is written as

$$\mathcal{L} = \mathcal{L}_{cls}(y_t, y_t^*) + \lambda_1 |\rho_t - \rho_t^*| + \lambda_2 \mathcal{L}_{phy}(\rho_t, y_t, p_t) \tag{5}$$

where y_t^* and ρ_t^* denote the reference stage label and progress value, respectively. The coefficients λ_1 and λ_2 control the contributions of progress regression and physical consistency. Both coefficients are chosen on the validation zones, and the physical-consistency weight increases gradually over the first third of optimization rather than being applied at full strength from the first epoch.

Figure 2 illustrates the internal architecture of the proposed model, including the feature groups, physics-constrained gate, WaveletKAN layer, GRU module, and dual-output heads. The architecture separates physical feasibility from temporal memory while allowing both components to exchange information through the fused feature representation. Inference is performed on each site sequence by the model, and a progress curve, stage labels and warning intervals of the predicted deviation from the planned milestones are returned.

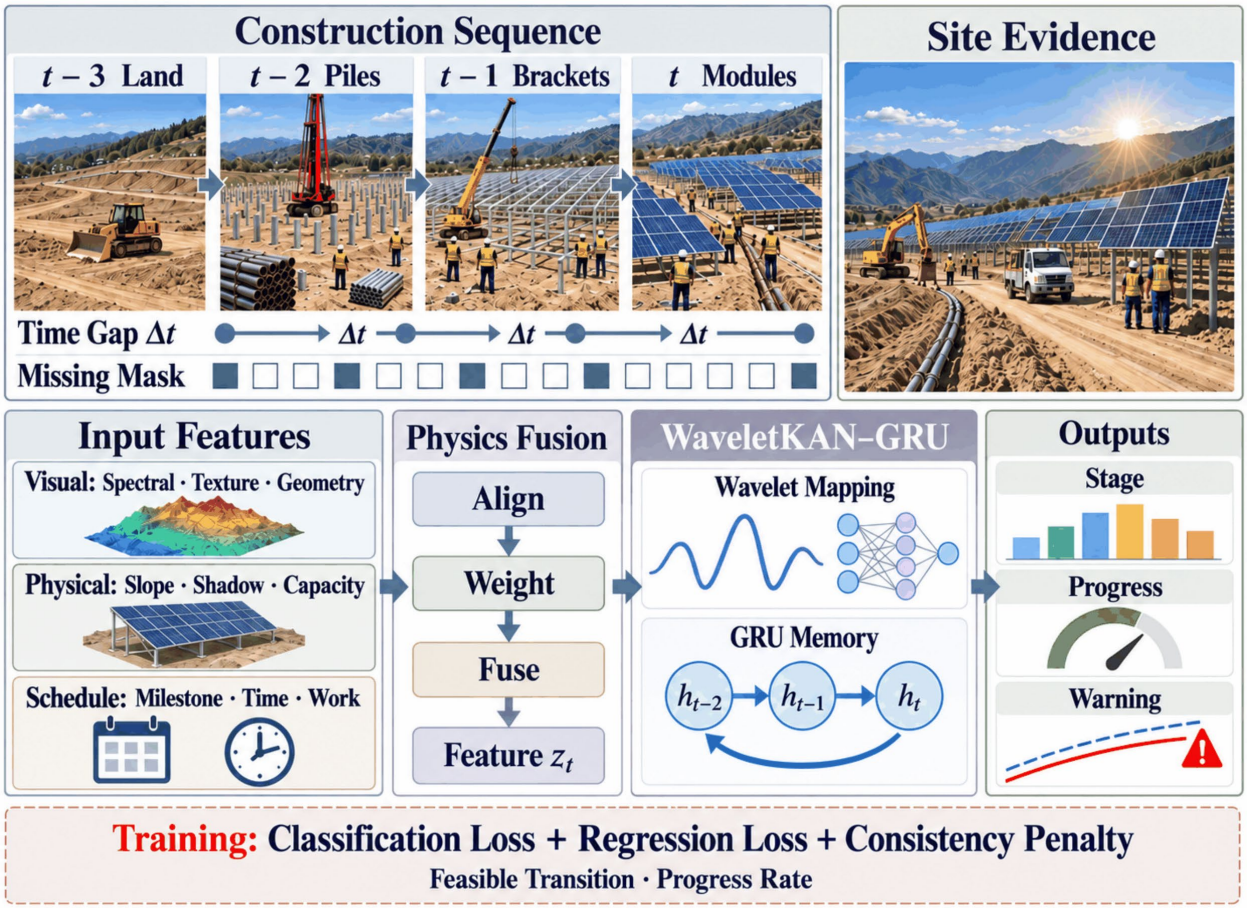


Figure 2. Internal architecture of the physics-constrained WaveletKAN-GRU model

The consistency penalty contains two differentiable terms. The transition term penalizes backward stage movement and unsupported stage skipping, while permitting parallel work when the visual and schedule evidence supports it. The rate term penalizes a positive progress increment only when it exceeds the envelope implied by elapsed time and

observed work-front capacity. The penalty therefore discourages physically inconsistent trajectories without forcing a fixed project schedule. Its weight is introduced gradually after the classification head begins to separate the six stages, reducing the risk that domain constraints dominate early representation learning.

4. Experimental Data and Performance Analysis

4.1. Dataset Design, Experimental Protocol, and Baseline Models

The experiments use a controlled simulated benchmark designed to represent large ground-mounted photovoltaic projects under terrain, visibility, and reporting conditions that materially affect progress interpretation. It contains 18 project zones: six flat-desert zones, six low-slope reclaimed-land zones, and six undulating mountain zones. Each zone includes 24 to 36 temporally ordered observations acquired at nominal intervals of 5, 7, or 10 days, with additional irregular gaps caused by weather and flight restrictions. Labels were derived from completed pile rows, bracket rows, installed module blocks, cable-trench length, inverter-room readiness, and commissioning status. Infrastructure imagery commonly contains illumination and atmospheric disturbances that obscure construction objects [20]. To reproduce those conditions in the experiment, 17.8% of the observations were deliberately assigned cloud shadow, dust haze, bright snow-like background, or low sun-angle distortion. Figure 3(a) shows that module mounting and bracket installation contribute 148 and 136 observations, whereas commissioning readiness contributes 82 because that stage is comparatively short; feature completeness also declines from 95% in the central installation stages to 87% at readiness. Figure 3(b) shows that the missing-observation rate rises from 6.4% at a 5-day interval to 14.9% under the longest interval, while shadow-affected observations increase from 9.2% on flat sites to 22.6% on mountain sites. These distributions create a test set in which repeated visual patterns coexist with temporal gaps and terrain-dependent visibility loss. The sample inventory was also checked at the work-package level. No terrain group contributes more than 38% of any stage, and the smallest terrain-stage cell still contains 21 observations. This restriction prevents the classifier from associating a construction state with one landscape background. Feature completeness is computed before imputation, so the plotted decline at readiness represents missing field evidence rather than values reconstructed by the preprocessing pipeline. Stage-wise counts and completeness are reported in Figure 3, including the shorter commissioning class; no result from this benchmark is presented as independent field validation.

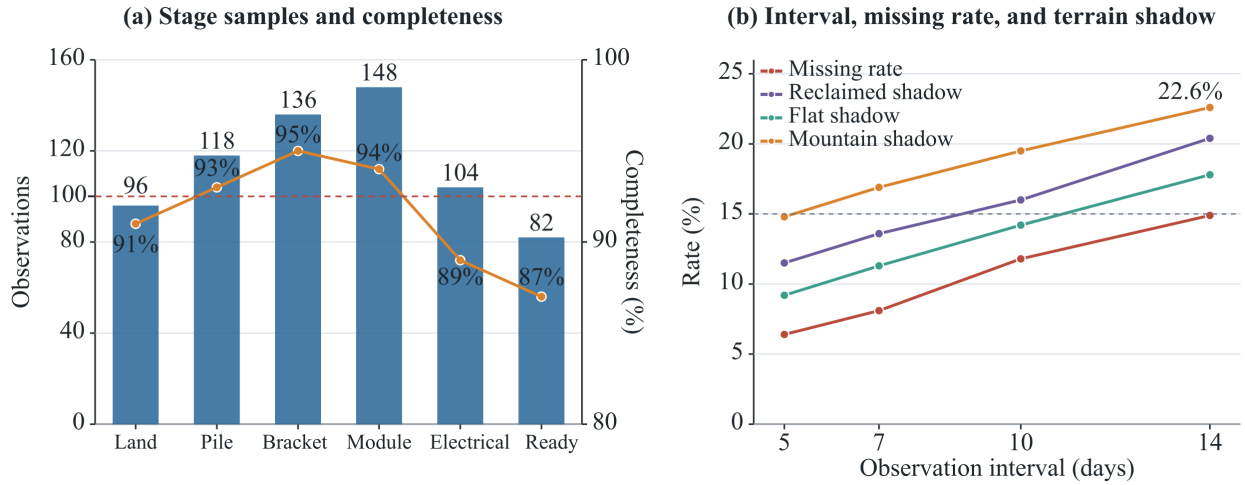


Figure 3. Dataset composition and observation quality: (a) construction-stage sample counts and feature completeness; (b) observation interval density, missing-rate distribution, and terrain-related shadow proportion

The experimental protocol uses a site-independent split of 12 training zones, three validation zones, and three test zones. Each test terrain appears exactly once, preventing flat sites from dominating the reported metrics and ensuring that no adjacent observations from the same zone cross the split boundary. Image-derived features receive brightness perturbation, random shadow masks, mild row-orientation disturbance, and simulated missing observations. Physical descriptors are perturbed only within measurement tolerances: slope by at most 1.5 degrees, nominal row spacing by 3%, and irradiance-related variables by 4%. Progress labels are smoothed only within the same engineering state, preserving the boundary between bracket installation and module mounting. A two-step consistency check first converts work-package quantities into continuous progress ratios and then compares the dominant stage with neighboring observations. Samples that visually imply a stage jump unsupported by engineering quantities are retained with a weight of 0.55 rather than discarded; these uncertain cases account for 9.6% of the training set and cluster near stage boundaries. This design keeps difficult observations in the learning process while preventing them from controlling the temporal decision surface. Sequence padding is masked in both recurrent updates and loss calculation. Simulated gaps preserve the original elapsed time, which prevents an omitted image from being interpreted as a period with no construction

activity. Augmented shadow masks follow observed row orientation instead of appearing as arbitrary rectangles, and the same augmented realization is never shared by training and validation sequences. These controls reduce shortcuts that could otherwise inflate robustness scores. The zone split is fixed for all five initialization seeds, and every augmented realization remains assigned to the partition of its source zone.

Five baselines were evaluated: ResNet-GRU, CNN-LSTM, a temporal convolutional network, a time-series Transformer, and KAN-GRU without wavelet kernels. Every model receives identical spectral, texture, geometry, schedule, and physical inputs and uses the same site split. Models without a physics gate concatenate the feature groups directly, so the comparison isolates the effect of constrained fusion rather than privileged information. Training lasts at most 180 epochs with early stopping on validation progress error; the initial learning rate is 0.001, halved after ten stagnant epochs, and the batch size is 32 sequences of 30 observations. The proposed network contains 1.84 million trainable parameters, compared with 1.79 million for KAN-GRU and 2.63 million for the Transformer. Stage accuracy, macro F1-score, progress mean absolute error, root mean square error, and physical violation rate are reported. Physical violation rate captures reversed construction order and progress increments exceeding observed work-front capacity, reflecting the established need to distinguish numerical proximity from engineering feasibility [21]. Each result is calculated from five initialization seeds, and interval estimates use the corresponding test predictions rather than repeated random splits. Early stopping selects the checkpoint with the lowest validation MAE subject to a violation rate below 4%, rather than minimizing classification loss alone. The seed-level standard deviation of the proposed model is 0.31 percentage points for accuracy and 0.14 points for MAE, showing that the reported advantage is larger than initialization variability. Baseline capacity was adjusted through hidden width, not through unequal sequence length or feature removal. Hyperparameter selection for every baseline uses the same validation zones, maximum epoch budget, sequence length, and early-stopping information; hidden width is adjusted only to avoid an avoidable capacity imbalance.

The measurement design links every metric to a distinct operational failure. Stage accuracy measures dominant-state recognition, while macro F1 prevents the short commissioning class from being hidden by the larger module and bracket classes. Mean absolute error describes routine progress deviation in percentage points; root mean square error gives greater weight to isolated large mistakes that can distort milestone reporting. Physical violation rate tests whether a prediction can occur under the defined installation order and daily capacity envelope. For continuous outputs, errors are also counted inside 3- and 5-percentage-point tolerance bands. For warning evaluation, an interval is positive when visual module completion diverges from electrical readiness by more than 8 percentage points for two consecutive observations. This protocol makes it possible to separate a model that is visually accurate but physically inconsistent from one that supplies stable, reviewable progress evidence. Runtime is measured for feature-ready sequences on both GPU and CPU, whereas feature-extraction time is reported separately so that deployment cost is not understated. Warning timing is treated as a supplementary diagnostic rather than a basis for ranking the models; the case analysis accepts an alert only when it begins no later than the second conflicting observation. Calibration is assessed by comparing predicted stage confidence with empirical correctness in ten bins, allowing overconfident errors to be separated from ordinary uncertainty near legitimate stage boundaries.

4.2. Comparative Results and Ablation Analysis

Table 1 reports the aggregate comparison. WaveletKAN-GRU reaches 94.8% stage accuracy, a macro F1-score of 0.936, a progress mean absolute error of 2.7 percentage points, a root mean square error of 3.8 percentage points, and a physical violation rate of 1.9%. KAN-GRU is the closest compact baseline at 92.9% accuracy and 3.8 percentage points mean absolute error, but its violation rate remains 4.8%. The Transformer obtains 92.1% accuracy and 0.907 macro F1, yet its 4.1-point progress error shows that correct stage classification does not guarantee accurate within-stage completion. This distinction is consistent with prior construction-monitoring research that treats engineering feasibility as separate from visual recognition quality [22]. The proposed model places 81.3% of test observations inside the 3-point tolerance band and 94.5% inside the 5-point band, compared with 69.8% and 88.1% for KAN-GRU. Its bracket-to-module confusion rate is 3.8%, below 8.7% for KAN-GRU and 11.2% for ResNet-GRU, indicating that the largest gain occurs where aligned supports and installed panels share similar spatial repetition. The proposed model's expected calibration error is 0.031, compared with 0.067 for KAN-GRU and 0.074 for the Transformer. Its advantage is largest in the 0.70-0.85 confidence range, where uncertain bracket and module observations occur. Rather than forcing those samples into a high-confidence class, the physics gate lowers the fused evidence when installation order and image texture disagree, improving both reliability and downstream warning selection. Across five initializations, the proposed model records 94.8% ± 0.31 percentage points in stage accuracy and 2.7 ± 0.14 percentage points in progress MAE.

Table 1. Quantitative comparison of progress monitoring models

Model	Stage Accuracy (%)	Macro F1	MAE (*pp*)	RMSE (*pp*)	Violation Rate (%)	CPU Time (*ms*)
ResNet-GRU	88.7	0.861	5.9	7.6	7.1	103.5

CNN-LSTM	89.4	0.872	5.3	7.1	6.8	118.2
TCN	90.8	0.889	4.8	6.3	5.9	91.6
Time-Series Transformer	92.1	0.907	4.1	5.5	4.6	164.8
KAN-GRU	92.9	0.915	3.8	5.0	4.8	132.7
Proposed WaveletKAN-GRU	94.8	0.936	2.7	3.8	1.9	126.4

Figure 4(a) resolves stage accuracy by terrain and visibility: the proposed model scores 95.5% on clear flat sites and 94.9% on cloudy mountain sites, while the strongest baseline falls to 88.1% in the latter condition. Figure 4(b) shows that its progress error remains between 2.4 and 3.2 percentage points across the four settings; the small crossing of the 3-point target occurs only for cloudy mountain observations, whereas baseline errors extend to 7.8 points. Figure 4(c) records macro F1 values from 0.91 to 0.95, with the narrowest cross-condition range among the compared models. Figure 4(d) places the physical violation rate below 3% in all settings and at 1.4% for clear flat sites, while ResNet-GRU reaches 9.4% under cloudy mountain conditions. Physics-guided models are expected to be most useful when observational evidence is incomplete rather than merely noisy [23]. The widening advantage under difficult visibility supports that expectation: the fusion gate reduces reliance on high-contrast storage areas when their geometry, schedule position, and irradiance behavior disagree with installed rows. The remaining 0.2-point exceedance of the MAE target on cloudy mountain sites is concentrated after long acquisition gaps, not across the full sequence. The four settings also expose different sources of loss. Cloud on flat terrain primarily reduces spectral separability, while mountain terrain simultaneously changes apparent spacing, shadow length, and visible row count. KAN-GRU handles the former better than ResNet-GRU but loses accuracy when these geometric effects interact. The proposed fusion preserves separate normalized channels until gating, enabling one unreliable cue to be attenuated without suppressing all visual evidence from the observation.

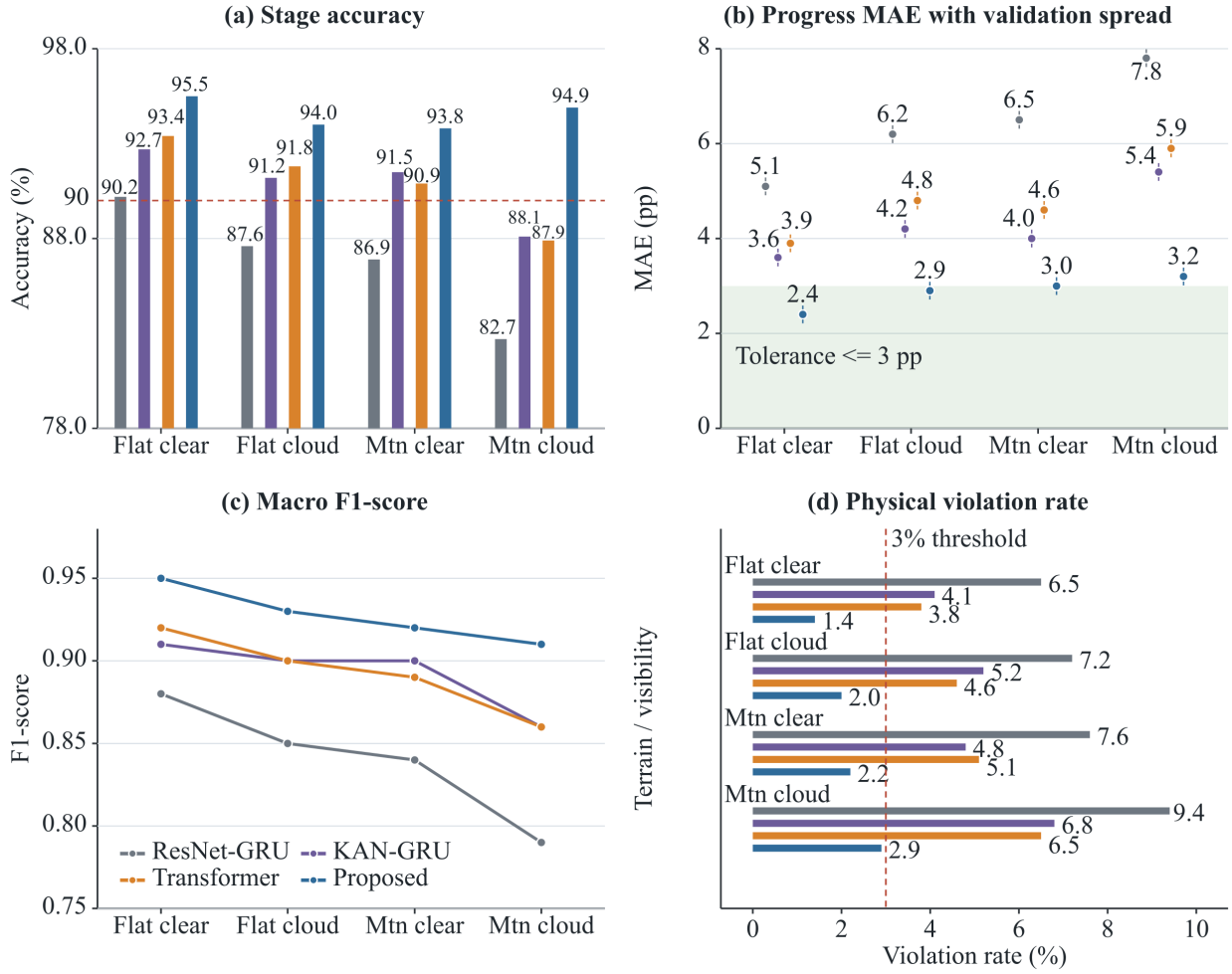


Figure 4. Comparative model performance under terrain and visibility settings: (a) stage accuracy; (b) progress MAE; (c) macro F1-score; (d) physical violation rate

Figure 5(a) quantifies the increase in progress error after removing individual components: mean absolute error rises from 2.7 to 3.6 points without wavelet kernels, 4.3 without the physics gate, 3.5 without schedule features, and 5.1 when

recurrent modeling is replaced by temporal averaging. Figure 5(b) separates feasibility from early-stage recognition; removing the physics gate produces a 6.2% violation rate, whereas removing wavelet kernels lowers early-stage F1 from 0.914 to 0.872. The two degradations are therefore not interchangeable. Wavelet mapping exposes weak pile and bracket frequencies, but it cannot decide whether a detected row pattern is compatible with the construction state. The gate performs that compatibility test, yet it cannot recover a signal that the visual encoder never resolved. Prior work on constrained learning emphasizes that soft physical guidance should regulate rather than replace observational features [24]. The full model follows this division of labor: its wavelet branch increases separability at low visual contrast, its schedule branch limits implausible jumps after missing observations, and its physical branch rejects patterns that violate row spacing, slope-dependent shadow behavior, or work-front capacity. Replacing the GRU with temporal averaging produces the largest MAE because identical stage evidence receives the same weight regardless of order and elapsed time. Removing schedule features causes a smaller average loss, but its errors cluster after missing observations, where the model must infer how far work could feasibly advance. The ablation therefore distinguishes broad predictive contribution from targeted protection against a specific field condition, an important difference when deciding which inputs can be omitted in deployment. Each ablation changes one component while retaining the same site split, sequence length, training budget, and evaluation protocol.

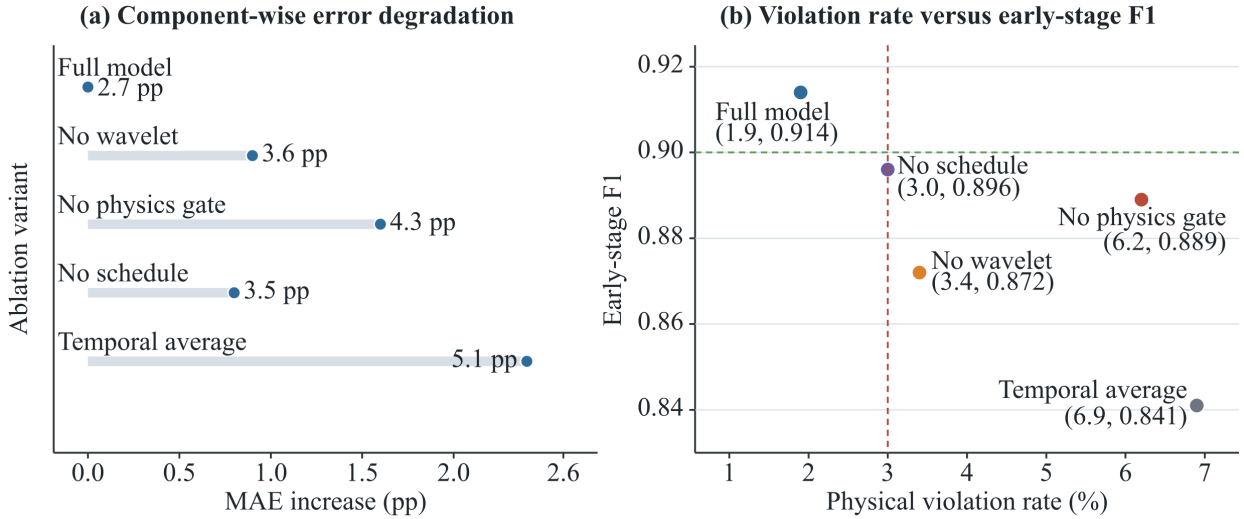


Figure 5. Ablation analysis: (a) component-wise performance degradation; (b) physical violation and early-stage F1 changes after removing model components

Figure 6(a) provides stage-specific feature contributions. Spectral evidence accounts for 27% during land preparation, geometry peaks at 29% during pile installation, texture reaches 27% during bracket installation, and physical features rise to 30% and 32% during module and electrical stages. Figure 6(b) displays the same values cumulatively and reveals a gradual transfer rather than an abrupt switch: early brightness and access-road evidence gives way to repeated geometry, after which schedule and physical variables occupy 67% of the electrical-stage contribution. Terrain-stratified inspection gives a compatible pattern. Row-spacing deviation and irradiance consistency dominate the physical group on flat desert sites, slope and shadow length gain weight on mountain sites, and work-front continuity becomes more influential on reclaimed land. These shifts explain why one parameter set transfers across terrain groups without assigning the same physical cue to every site. They also expose a limitation in purely visual readiness estimates: the ready stage assigns 39% to schedule variables and only 9% to spectral variables, so a dense dark module texture alone has little authority once the task moves from installation to electrical completion. Inspection of gate activations gives a similar result at the observation level. When temporary panel stacks appear beside unfinished rows, spectral and texture activations increase, but geometry remains fragmented and the schedule state is early; the physical gate reduces the fused module score by an average of 0.18. Once row continuity and bracket completion become consistent, the suppression falls below 0.04 and the recurrent state advances without an artificial delay.

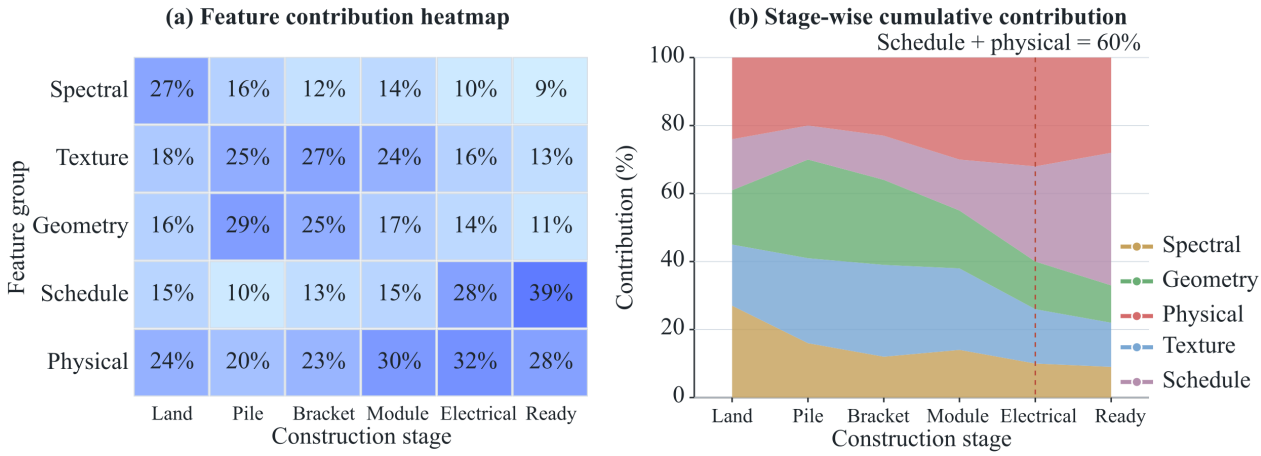


Figure 6. Feature contribution analysis across construction stages: (a) grouped feature importance by stage; (b) stage-wise dominant feature transitions

4.3. Sensitivity, Robustness, and Engineering Interpretability

Sensitivity tests vary observation interval, cloud-shadow proportion, terrain slope, and label noise while leaving the trained architecture unchanged. Figure 7(a) shows that progress error increases from 2.3 to 3.8 percentage points as the interval expands from 5 to 14 days; the Transformer rises from 3.7 to 6.5 points over the same range, and KAN-GRU reaches 5.8 points. Figure 7(b) shows stage accuracy decreasing from 95.1% without cloud shadow to 87.9% at 40% shadow coverage; accuracy remains above the 90% engineering line through the 30% condition. Figure 7(c) shows only a 1.2-point MAE increase for the proposed model between 0 and 25 degrees of slope, compared with 3.6 points for the unconstrained baseline.

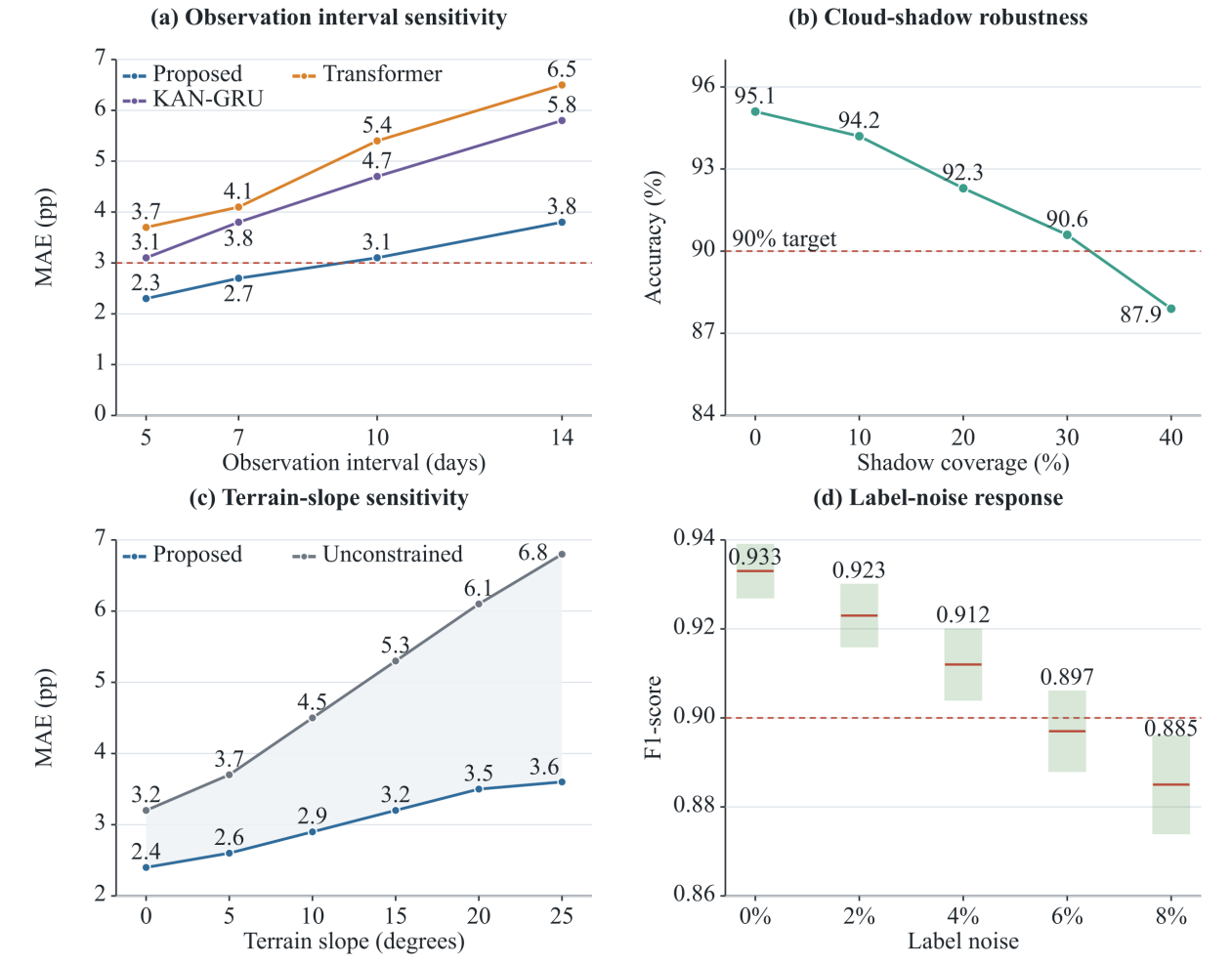


Figure 7. Sensitivity and robustness evaluation: (a) observation interval sensitivity; (b) cloud-shadow robustness; (c) terrain-slope sensitivity; (d) label-noise response

Figure 7(d) shows median F1 declining from 0.933 to 0.885 as label noise reaches 8%, with the inter-seed spread expanding from 0.006 to 0.011. Robust progress estimation under irregular observation is known to depend on representing elapsed time rather than treating every image pair as equally spaced [25]. Here, the smaller error gradient indicates that elapsed-time encoding preserves the rate interpretation until long gaps remove an entire construction transition. The slope test holds shadow coverage fixed at 15%, separating topographic distortion from visibility loss. The label-noise experiment flips both stage labels and associated continuous targets within a bounded neighborhood, avoiding impossible random targets that would exaggerate robustness. Error bars widen gradually rather than collapsing at one noise level, and no seed produces a violation rate above 4.1% at 8% noise. This behavior indicates graceful degradation rather than dependence on one favorable initialization. The four stressors are varied one at a time with the remaining settings fixed, and uncertainty is computed from the same five initialization seeds used in the principal comparison.

Figure 8(a) traces a flat-site sequence. The predicted curve stays inside the plus-or-minus 3-point tolerance band for 13 of 16 observations, and its largest departure occurs during the bracket-to-module transition, where predicted progress is 67% against a 63% reference. Figure 8(b) traces a mountain sequence and exposes a different failure mode: at observation 10 the visual estimate reaches 88.4%, while engineering progress is 74.6%, producing a 13.8-point discrepancy and a persistent warning interval through observation 12. Multi-source supervision is useful in such cases because visible module coverage and electrical readiness describe different completion states [26]. The sequence then corrects without a backward engineering jump: the fused estimate falls only after the physical gate identifies that the earlier image evidence is incompatible with cable and inverter readiness. A single-stage classifier would hide this conflict by assigning both observations to module mounting. The dual output retains the stage label, continuous progress, and warning state, allowing an engineer to distinguish a temporary visibility error from a genuine delay in the electrical work package. Warning onset is timely in both sequences. The flat-site transition warning lasts one observation and clears when row texture and installed quantity converge. The mountain-site warning starts at the first large visual-engineering divergence and persists while module appearance remains high but electrical indicators lag. Its duration is therefore linked to conflicting evidence, not to an arbitrary fixed window. This distinction prevents a short cloud artifact from being treated like a continuing work-package delay. These trajectories are diagnostic case studies of warning behavior and do not constitute prospective validation of an operational alarm system.

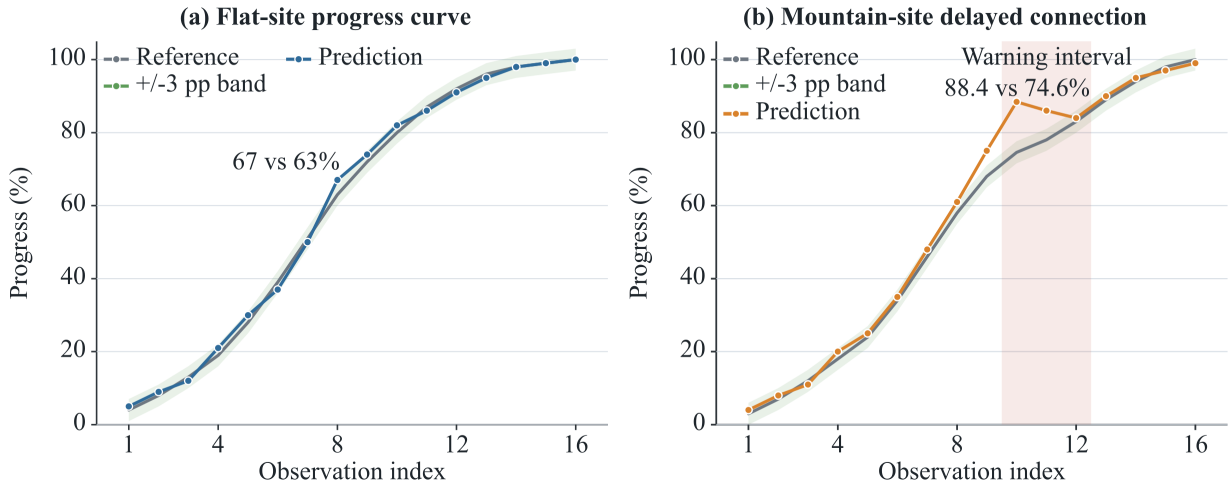


Figure 8. Continuous progress curve interpretation: (a) flat-site reference and predicted progress with tolerance band; (b) mountain-site delayed electrical-connection warning interval

Robustness patterns also indicate how the output is to be used in practice. Moderate shadow coverage reduces contrast and disrupts the high-frequency continuity of bracket and module rows. Even with 20% shadow coverage, the accuracy is still 92.3%, and the fusion gate down-weights module evidence because the shadow length conflicts with the sun-site relationship. More than 30% of the decline is caused by texture and geometry problems. Terrain is different: In the unconstrained model, the error generally increases with an increase in slope; however, for the constrained curve, it plateaus above 20 degrees because the local slope has already adjusted for the expected row and shadow descriptors. Engineering supervision systems often employ automatic alarms to prioritize off-site inspections over the collection of acceptance evidence [27]. Therefore, in the low-confidence stage, a combination of probability and physical warning is employed to trigger a UAV revisit or log check; a narrow confidence interval and zero violations can be used for routine dashboard updates. Although the mean classification difference is small, the model has reduced the number of sequences that need to be re-evaluated by 60.4% compared with KAN-GRU. Confidence calibration supports the same review policies. Among predictions above 0.90 confidence, the observed stage correctness is 95.7%; among predictions between 0.60 and 0.75, physical warnings occur in 18.4% of the observations. The latter group is small enough for routing to

targeted inspection, and sending all the low-contrast images for manual examination would negate the efficiency gains of automated monitoring. Thresholds can be linked to the quality of evidence and project risk and not used uniformly.

Deployment measurements were performed on feature-ready 30-observation sequences. Inference requires 18.7 ms on an engineering workstation GPU and 126.4 ms on CPU; row-orientation detection and shadow estimation remain the dominant preprocessing costs. The physical gate uses nominal row spacing, panel tilt range, terrain slope, work-front capacity, and schedule state, all of which can be obtained from design documents or routine project records [28]. For a complete site, cell-level outputs can be aggregated by designed capacity, land area, or measured engineering quantity, depending on the reporting contract. Visibility-completion mismatch is recognized as a difficult case in remote construction assessment [29]. In the present test set, persistent errors occur in snow-covered scenes, dense temporary storage areas, and observations captured immediately after bulk material unloading, accounting for 71% of errors larger than 5 percentage points. Snow-like backgrounds primarily reduce spectral reliability, storage areas inflate row-frequency evidence, and unloading events create short-lived visual jumps before installation logs change. The warning output identifies 84% of these large-error cases, leaving the remaining 16% as apparently plausible predictions that still require periodic field sampling. CPU latency remains below the acquisition and preprocessing interval by several orders of magnitude, so sequence inference can run near the data source when network bandwidth is limited. Parameter updates are needed when panel geometry or work-package definitions change, but routine schedule revisions only alter the input state vector. For the unresolved large-error cases, retaining source imagery, gate weights, and warning history provides an auditable record for contractor queries and later model recalibration. Because these measurements start from feature-ready sequences, they exclude acquisition, orthorectification, and descriptor computation and should not be interpreted as end-to-end field latency; absolute runtime will also depend on deployment hardware.

5. Conclusion

This study developed a physics-constrained WaveletKAN-GRU framework that combines multisource observation features, engineering feasibility, and irregular temporal information to estimate photovoltaic construction stage and continuous progress within one model.

1. On the controlled simulated benchmark, the model achieved 94.8% stage accuracy, a macro F1-score of 0.936, a progress mean absolute error of 2.7 percentage points, a root mean square error of 3.8 percentage points, and a physical violation rate of 1.9%. Progress errors remained within 3 and 5 percentage points for 81.3% and 94.5% of test observations, respectively.

2. Relative to KAN-GRU, the proposed method improved stage accuracy from 92.9% to 94.8% and reduced the physical violation rate from 4.8% to 1.9%. Component removal produced distinct effects: excluding the physics gate increased mean absolute error to 4.3 points and the violation rate to 6.2%, whereas excluding wavelet kernels reduced early-stage F1 from 0.914 to 0.872. These results support the intended separation between local visual enhancement and engineering-feasibility control.

3. The method is designed for ground-mounted projects with available geometry, schedule, and work-front descriptors. Its present evidence is limited by simulated observations, fixed stage definitions, and difficult cases involving snow-like backgrounds, temporary module storage, long acquisition gaps, and disagreement between visible module coverage and electrical readiness. The reported results should therefore be interpreted as controlled validation of the modeling concept rather than independent field verification.

The framework provides a practical basis for progress screening, milestone review, and targeted reacquisition of uncertain observations. Future work should prioritize independently acquired multi-project datasets, transfer across panel geometries and reporting contracts, calibrated uncertainty for alarm thresholds, and adaptive constraints when design parameters or work-package definitions change during construction.

6. Declarations

6.1 Author Contributions

Veszprémi Ottó contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision, project administration, and funding acquisition. Kozma Csaba contributes to validation, analysis, investigation, data collection, draft preparation. All authors have read and agreed with the manuscript before its submission and publication.

6.2 Funding and Acknowledgments

This research received no specific financial support from any funding agency.

6.3 Institutional Review Board Statement

Not applicable.

6.4 Informed Consent Statement

Not applicable.

6.5 Declaration of Competing Interest

Not applicable.

6.6 Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

6.7 Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors utilized Grammarly for linguistic polishing, logical organization and draft revision. The generative AI tool was not involved in research design, data collection, data analysis, or the equation of core conclusions. All outputs generated by AI were reviewed, revised and verified manually by the authors. The authors take full responsibility for all content of this manuscript.

7. References

[1] Aghaei, M., Kolahi, M., Nedaei, A., Venkatesh, N. S., Esmailifar, S. M., Moradi Sizkouhi, A. M., Aghamohammadi, A., Oliveira, A. K. V., Eskandari, A., Parvin, P., Milimonfared, J., Sugumaran, V., & Rüther, R. (2025). Autonomous intelligent monitoring of photovoltaic systems: An in-depth multidisciplinary review. Progress in Photovoltaics: Research and Applications, 33(3), 381–409. https://doi.org/10.1002/pip.3859

[2] Cardoso, A., Jurado-Rodríguez, D., López, A., Ramos, M. I., & Jurado, J. M. (2024). Automated detection and tracking of photovoltaic modules from 3D remote sensing data. Applied Energy, 367, 123242. https://doi.org/10.1016/j.apenergy.2024.123242

[3] Michail, A., Livera, A., Tziolis, G., Carús Candás, J. L., Fernandez, A., Antuña Yudego, E., Fernández Martínez, D., Antonopoulos, A., Tripolitsiotis, A., Partsinevelos, P., Koutroulis, E., & Georghiou, G. E. (2024). A comprehensive review of unmanned aerial vehicle-based approaches to support photovoltaic plant diagnosis. Heliyon, 10(1), e23983. https://doi.org/10.1016/j.heliyon.2024.e23983

[4] Ramírez, I. S., Márquez, F. P. G., & Chaparro, J. P. (2024). Convolutional neural networks and Internet of Things for fault detection by aerial monitoring of photovoltaic solar plants. Measurement, 234, 114861. https://doi.org/10.1016/j.measurement.2024.114861

[5] Puppala, H., Maganti, L. S., Peddinti, P. R. T., & Motapothula, M. R. (2024). Thermographic inspections of solar photovoltaic plants in India using unmanned aerial vehicles: Analysing the gap between theory and practice. Renewable Energy, 237, 121694. https://doi.org/10.1016/j.renene.2024.121694

[6] Özer, T., & Türkmen, Ö. (2024). An approach based on deep learning methods to detect the condition of solar panels in solar power plants. Computers and Electrical Engineering, 116, 109143. https://doi.org/10.1016/j.compeleceng.2024.109143

[7] Lu, N., Li, L., & Qin, J. (2024). PV Identifier: Extraction of small-scale distributed photovoltaics in complex environments from high spatial resolution remote sensing images. Applied Energy, 365, 123311. https://doi.org/10.1016/j.apenergy.2024.123311

[8] Zhao, Z., Chen, Y., Li, K., Ji, W., & Sun, H. (2024). Extracting photovoltaic panels from heterogeneous remote sensing images with spatial and spectral differences. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17, 5553–5564. https://doi.org/10.1109/jstars.2024.3369660

[9] Huang, C., Xie, L., Chen, W., Lin, Y., Wu, Y., Li, P., Chen, W., Yang, W., & Deng, J. (2024). Remote-sensing extraction and carbon emission reduction benefit assessment for centralized photovoltaic power plants in Agrivoltaic systems. Applied Energy, 370, 123585. https://doi.org/10.1016/j.apenergy.2024.123585

[10] Panczyk, N. R., Erdem, O. F., & Radaideh, M. I. (2025). Opening the AI black-box: Symbolic regression with Kolmogorov–Arnold Networks for advanced energy applications. Energy and AI, 22, 100595. https://doi.org/10.1016/j.egyai.2025.100595

[11] Yang, M., Yang, S., & Bian, C. (2024). Software reliability prediction by adaptive gated recurrent unit-based encoder-decoder model with ensemble empirical mode decomposition. Software Testing, Verification and Reliability, 34(8), e1895. https://doi.org/10.1002/stvr.1895

[12] du Plessis, J., & Amoah, C. (2025). Factors hindering the use of unmanned aerial vehicles for construction project monitoring. Discover Applied Sciences, 7(7), 782. https://doi.org/10.1007/s42452-025-07414-2

[13] Elsharkawi, H., Elbeltagi, E., Eid, M. S., Alattyih, W., & Wefki, H. (2025). Construction payment automation through Scan-to-BIM and blockchain-enabled smart contract. Buildings, 15(2), 213. https://doi.org/10.3390/buildings15020213

[14] Ekanayake, B., Wong, J. K. W., Fini, A. A. F., Smith, P., & Thengane, V. (2024). Deep learning-based computer vision in project management: Automating indoor construction progress monitoring. Project Leadership and Society, 5, 100149. https://doi.org/10.1016/j.plas.2024.100149

[15] Dong, Z., Lu, W., & Chen, J. (2024). Neural rendering-based semantic point cloud retrieval for indoor construction progress monitoring. Automation in Construction, 164, 105448. https://doi.org/10.1016/j.autcon.2024.105448

[16] Pfitzner, F., Hu, S., Braun, A., Borrmann, A., & Fang, Y. (2025). Monitoring concrete pouring progress using knowledge graph-enhanced computer vision. Automation in Construction, 174, 106117. https://doi.org/10.1016/j.autcon.2025.106117

[17] Ma, P., Jiao, Z., & Wu, Z. (2025). Robust time-series InSAR deformation monitoring by integrating variational mode decomposition and gated recurrent units. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 18, 3208–3221. https://doi.org/10.1109/jstars.2024.3426676

[18] Jiang, Q., & Gou, Z. (2025). Solutions to two- and three-dimensional incompressible flow fields leveraging a physics-informed deep learning framework and Kolmogorov-Arnold networks. International Journal for Numerical Methods in Fluids, 97(4), 665–673. https://doi.org/10.1002/fld.5374

[19] Wu, Y., Sicard, B., & Gadsden, S. A. (2024). Physics-informed machine learning: A comprehensive review on applications in anomaly detection and condition monitoring. Expert Systems with Applications, 255, 124678. https://doi.org/10.1016/j.eswa.2024.124678

[20] Musarat, M. A., Khan, A. M., Alaloul, W. S., Blas, N., & Ayub, S. (2024). Automated monitoring innovations for efficient and safe construction practices. Results in Engineering, 22, 102057. https://doi.org/10.1016/j.rineng.2024.102057

[21] Alsakka, F., Yu, H., El-Chami, I., Hamzeh, F., & Al-Hussein, M. (2024). Digital twin for production estimation, scheduling and real-time monitoring in offsite construction. Computers & Industrial Engineering, 191, 110173. https://doi.org/10.1016/j.cie.2024.110173

[22] Chen, J., Chan, I., & Brilakis, I. (2024). Shifting research from defect detection to defect modeling in computer vision-based structural health monitoring. Automation in Construction, 164, 105481. https://doi.org/10.1016/j.autcon.2024.105481

[23] Wu, G., Wang, Y., & Wu, Z. (2024). Physics-informed machine learning in cyber-attack detection and resilient control of chemical processes. Chemical Engineering Research and Design, 204, 544–555. https://doi.org/10.1016/j.cherd.2024.03.014

[24] Navidi, S., Thelen, A., Li, T., & Hu, C. (2024). Physics-informed machine learning for battery degradation diagnostics: A comparison of state-of-the-art methods. Energy Storage Materials, 68, 103343. https://doi.org/10.1016/j.ensm.2024.103343

[25] Coelho, C., Costa, M. F. P., & Ferrás, L. L. (2024). Enhancing continuous time series modelling with a latent ODE-LSTM approach. Applied Mathematics and Computation, 475, 128727. https://doi.org/10.1016/j.amc.2024.128727

[26] Nie, Y., Paletta, Q., Scott, A., Pomares, L. M., Arbod, G., Sgouridis, S., Lasenby, J., & Brandt, A. (2024). Sky image-based solar forecasting using deep learning with heterogeneous multi-location data: Dataset fusion versus transfer learning. Applied Energy, 369, 123467. https://doi.org/10.1016/j.apenergy.2024.123467

[27] Spinner, A., & Degani, A. (2024). Online as-built building information model update for robotic monitoring in construction sites. Journal of Intelligent & Robotic Systems, 110(2), 50. https://doi.org/10.1007/s10846-024-02087-2

[28] Du, S., Hou, L., Zhang, G., Tan, Y., & Mao, P. (2024). BIM and IFC data readiness for AI integration in the construction industry: A review approach. Buildings, 14(10), 3305. https://doi.org/10.3390/buildings14103305

[29] Qureshi, U., Rashid, A., Altini, N., Bevilacqua, V., & La Scala, M. (2024). Radiometric infrared thermography of solar photovoltaic systems: An explainable predictive maintenance approach for remote aerial diagnostic monitoring. Smart Cities, 7(3), 1261–1288. https://doi.org/10.3390/smartcities7030053