agriculture

MDPI

*Article*

# DRAMM-Pest: Direct Multi-Horizon Forecasting of Rice Leaf Roller Activity from Light-Trap and Field-Survey Records

**Qianxin Xie [1], Jinfeng Xu [1,*], Yuxuan Wang [2], Yuchen Lu [3] and Yuxuan Zhang [2,4,*]**

[1] School of Information Engineering, Guilin University, Guilin 541006, China; xieqianxin@glc.edu.cn
[2] College of Intelligent Science and Engineering, Beijing University of Agriculture, Beijing 102206, China; 202520742202@bua.edu.cn
[3] Yantai Research Institute, Harbin Engineering University, Yantai 264006, China; luyuchen@hrbeu.edu.cn
[4] Department of Computer and Electrical Engineering, Mid Sweden University, 85170 Sundsvall, Sweden
* Correspondence: xujinfeng@glc.edu.cn (J.X.); yuxuan.zhang@miun.se (Y.Z.)

**Abstract**

Multi-horizon forecasts of rice leaf roller activity support advance surveillance planning. Routine light-trap and field-survey records differ in measurement meaning and availability, complicating their complementary use across future monitoring periods. We developed DRAMM-Pest, a mask-aware dual-measurement model in which available field histories condition temporal states derived from light-trap history before horizon-specific aggregation. The model jointly predicts positive-catch probability and light-trap rate over six successive monitoring periods, with observed field-survey targets providing auxiliary supervision during training. On a public 24-year rice leaf roller dataset covering 11 sites, chronological testing on 2020–2023 yielded a light-trap-rate RMSLE of 0.837 ± 0.029 and a Brier score of 0.084 ± 0.004. The corresponding absolute reductions relative to XGBoost were 0.25023 and 0.00859, respectively. Removing all field information or retaining only its binary availability indicator increased mean light-trap-rate MAE by 3.12 and 2.74, respectively. The findings support using complementary monitoring records as historical context and training targets within a horizon-specific forecasting framework for pest-surveillance planning.

**Keywords:** pest early warning; *Cnaphalocrocis medinalis*; multi-horizon time-series forecasting; dual-measurement forecasting; light-trap monitoring

## 1. Introduction

Rice leaf roller (RLR; *Cnaphalocrocis medinalis*) is a migratory pest of rice whose larval feeding damages photosynthetically active leaves and increases yield risk when populations exceed management thresholds. Its population dynamics are shaped by seasonal development and regional migration, making continuous monitoring and timely warning important components of sustainable pest management [1–3]. Regional forecasting of RLR occurrence and migration has used monitoring and meteorological records to construct prediction models for different rice-growing regions and crop-development stages [4]. Routine monitoring provides observations of current activity, whereas agricultural decisions concerning field scouting, trap maintenance, and intervention preparation must often be made several monitoring periods in advance. Direct multi-horizon forecasting can convert historical monitoring records into anticipated positive-catch probabilities and predicted activity levels, thereby extending pest surveillance from current-state assessment to advance planning.

RLR monitoring produces two complementary observation streams. Light-trap catches provide relatively dense records of adult flight activity, including signals associated with

Check for updates

Academic Editor: Xianming Yang

Received: 7 September 2026
Revised: 26 October 2026
Accepted: 28 October 2026
Published: 29 October 2026

**Copyright:** © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

immigration, whereas field surveys quantify pest abundance within the crop environment at scheduled calendar periods [2,5]. Searchlight-trap observations of RLR have identified different nocturnal catch rhythms during immigration and emigration periods [6]. Trajectory analyses have further linked regional migration to wind fields, flight duration, and the distribution of suitable rice habitats [7]. The two streams therefore differ in measurement meaning, unit, sampling rhythm, site coverage, and observation availability. Each observation stream follows a source-specific availability pattern represented by its own mask. Structured missingness can itself carry information about the observation process, while agricultural datasets commonly exhibit uneven temporal and spatial coverage [8,9]. Pest activity also varies among monitoring locations because temporal population dynamics are embedded in regional environmental and management contexts [10]. Forecasting from these records consequently requires the joint treatment of complementary measurement semantics, incomplete histories, temporal variation, and regional context.

Deep learning for plant-disease identification has combined residual learning with adaptive feature recalibration to represent disease-related visual patterns [11]. Data-driven pest forecasting has used expert systems, machine-learning models, and recurrent or spatiotemporal neural networks to learn relationships among monitoring records, environmental variables, crop-development information, and future pest activity [12–14]. ESD-TripleStream applies exponential smoothing decomposition to pest records and combines trend, seasonal, and temporal-feature streams for short- and medium-term forecasting [15]. MeteoSCOPE represents historical pest and multidimensional meteorological descriptors as structured tokens and uses horizon-specific queries for retrospective multi-horizon forecasting [16]. Related agricultural models integrate heterogeneous sensor, field, image, and remote-sensing inputs or represent geographic dependence among monitoring locations [10,17–21]. Multi-sensor condition-monitoring studies have also combined individual measurements into integrated state indicators for prediction [22].

Existing pest-forecasting models provide approaches for temporal decomposition, integration of environmental and heterogeneous predictors, and prediction at multiple lead times. For the paired monitoring records considered here, the specific modeling question is how to use field-survey abundance to support light-trap forecasting when the two measurements differ in meaning and availability. Available field histories can provide predictive context, while observed field-survey targets can provide auxiliary supervision during training. The challenge is to coordinate these two roles with the temporal information relevant to each forecast horizon. We investigate this question through direct forecasts over six successive monitoring periods.

We address this question with DRAMM-Pest, which separately encodes masked light-trap and field-survey histories. Available field context modulates temporal component states derived from light-trap history before horizon-specific aggregation. The resulting representations support joint forecasts of positive-catch probability and light-trap rate. During training, observed field-survey targets provide auxiliary supervision for the shared representation. Calendar variables and site coordinates supply temporal and spatial context. The main contributions are as follows:

- We formulate a direct six-period RLR forecasting task in which light-trap records define the main prediction target and field-survey records supply historical context and auxiliary training targets, with a separate observation mask for each stream.
- We develop mask-aware temporal modeling in which available field history modulates light-trap component states before horizon-specific aggregation, producing forecasts of positive-catch probability and light-trap rate.
- We evaluate chronological and held-out-site forecasting and sensitivity to missing historical inputs. Field-information controls, component ablations, and within-site

comparisons assess the predictive contributions of temporal modeling, adaptive field guidance, and the binary field-availability indicator.

## 2. Materials and Methods

### *2.1. Dataset and Preprocessing*

Table 1 summarizes the observation modalities, derived model variables, and chronological data partitions.

**Table 1.** Summary of the RLR observation modalities, aligned model variables, and chronological forecasting samples.

| **Panel A. Observation Modalities and Model Variables** | | | | |
|---|---|---|---|---|
| **Data Component** | **Aligned Representation** | **Resolution and Unit** | **Sites** | **Model Role** |
| Light-trap observations | Catch rate, positive-catch indicator, valid trap nights, and observation mask | Daily to calendar period; individuals trap-night$^{-1}$, nights, and binary indicators | 10 | Historical primary stream and main forecast target |
| Field-survey observations | Calendar-period abundance and an independent observation mask | Calendar period; individuals ha$^{-1}$ and binary indicator | 10 | Historical field context and auxiliary forecast target |
| Calendar variables | Historical and future cyclic features derived from month and ordered calendar-period indices | Calendar period; four-dimensional continuous encoding | All | Temporal conditioning |
| Spatial variables | Continuous longitude and latitude | Static; degrees | 11 | Spatial conditioning |

| **Panel B. Chronological Partitions and Accepted Forecasting Samples** | | | |
|---|---|---|---|
| **Partition** | **Years** | **Number of Years** | **Accepted 12-to-6 Samples** |
| Training | 2000–2015 | 16 | 3151 |
| Validation | 2016–2019 | 4 | 855 |
| Test | 2020–2023 | 4 | 836 |
| Total | 2000–2023 | 24 | 4842 |

This study uses publicly archived rice leaf roller monitoring records collected during 2000–2023 in the Lower Yangtze River Valley. The records were obtained from Dryad and originated from local plant-protection stations coordinated by the National Agro-Tech Extension and Service Centre (NATESC). The original study documents nightly adult-moth monitoring using standardized 20-W ultraviolet light traps and daily transect surveys during the local moth-occurrence season, in which field staff counted adults disturbed from rice plants. The archived field-survey variable records adult abundance per hectare at calendar-period resolution [2,23]. The light-trap and field-survey modalities each covered 10 georeferenced sites, of which nine were shared. One site was exclusive to each modality, resulting in 11 sites in their union.

Table 1 distinguishes source-site coverage from accepted forecasting windows. Table 2 reports site coordinates, valid observation periods, and accepted windows for each data partition. Ten sites contributed forecasting windows. The field-only site, S010, contributed none because window acceptance required valid light-trap targets at all six forecast horizons.

**Table 2.** Site coordinates, observation availability, and contributions to the chronological forecasting dataset. Availability percentages use the 1152 calendar-period positions per site over 2000–2023 as the denominator. Valid light-trap periods contain at least four valid trap nights; valid field-survey periods contain an observed abundance value. Observed zeros are retained. These percentages describe coverage of the common analysis grid, rather than completion of scheduled surveys. The final three columns report accepted 12-to-6 forecasting windows.

| Site | Longitude | Latitude | Valid Light-Trap Periods, $n$ (%) | Valid Field-Survey Periods, $n$ (%) | Train | Validation | Test |
|---|---|---|---|---|---|---|---|
| S001 | 116.3489 | 27.9223 | 940 (81.6) | 586 (50.9) | 436 | 110 | 98 |
| S002 | 116.9401 | 30.9699 | 759 (65.9) | 294 (25.5) | 314 | 84 | 84 |
| S003 | 117.7753 | 29.3243 | 858 (74.5) | 381 (33.1) | 368 | 107 | 94 |
| S004 | 118.3166 | 29.8148 | 784 (68.1) | 384 (33.3) | 404 | 71 | 63 |
| S005 | 119.4908 | 32.8314 | 699 (60.7) | 407 (35.3) | 278 | 84 | 82 |
| S006 | 119.6336 | 31.9581 | 687 (59.6) | 539 (46.8) | 255 | 84 | 84 |
| S007 | 120.2633 | 32.0347 | 589 (51.1) | 355 (30.8) | 165 | 79 | 78 |
| S008 | 120.5020 | 30.0893 | 841 (73.0) | 637 (55.3) | 343 | 104 | 96 |
| S009 | 120.6294 | 31.9002 | 788 (68.4) | 440 (38.2) | 356 | 84 | 84 |
| S010 | 121.2313 | 30.8334 | 0 (0.0) | 386 (33.5) | 0 | 0 | 0 |
| S011 | 121.5232 | 31.6507 | 657 (57.0) | 0 (0.0) | 232 | 48 | 73 |
| Total | — | — | 7602 | 4409 | 3151 | 855 | 836 |

Values coded as −999 were converted to missing values, whereas observed zeros were retained. Light-trap records were indexed by date, longitude, and latitude, and field-survey records were indexed by year, month, calendar period, longitude, and latitude. Exact duplicate records were reduced to a single entry. When the same key contained both a valid value and a missing entry, the valid value was retained.

Each site-year sequence was represented on a March–October calendar grid containing 48 periods, with six periods per month covering days 1–5, 6–10, 11–15, 16–20, 21–25, and 26–end of month. Actual observation coverage varied among sites and years. This resolution follows the archived field-survey records and provides a common time scale for aligning field abundance with nightly catches aggregated over the same periods. The periods describe within-season adult activity and are not defined as developmental stages or generations.

Let $n_{s,d}$ denote the catch count recorded at site $s$ on trap night $d$, and let $E_{s,t}$ denote the number of valid trap nights in calendar period $t$. The calendar-period light-trap catch rate and positive-catch indicator were calculated as

$$r_{s,t} = \frac{\sum_{d \in \mathcal{D}_{s,t}} n_{s,d}}{E_{s,t}}, \qquad o_{s,t} = \mathbb{I}\left( \sum_{d \in \mathcal{D}_{s,t}} n_{s,d} > 0 \right), \tag{1}$$

where $\mathcal{D}_{s,t}$ is the set of valid trap nights in the corresponding calendar period. A calendar period was treated as a valid light-trap observation when $E_{s,t} \geq 4$. Periods below this threshold remained at their original positions in the aligned calendar grid and were represented by $m^L_{s,t} = 0$. Field-survey abundance was mapped to the same site-by-calendar-period grid and retained an independent mask $m^F_{s,t}$.

The two observation modalities were aligned by site, year, month, and calendar period. Historical and future calendar covariates were generated for every aligned period. Count and abundance inputs were transformed using $\log(1 + x)$. Missing inputs were filled using training-derived medians for the corresponding variables, while observed zeros and modality-specific availability masks were retained. Continuous inputs to the neural networks and Direct Ridge were standardized using training-set means and standard deviations, except that DRAMM-Pest coordinates used the fixed encoding in Equation (11)

without training-set standardization. Binary masks and cyclic sine–cosine features were not standardized. Valid field observations were retained independently of light-trap availability. Missing targets were not imputed and were excluded from loss and metric calculations. All preprocessing statistics were estimated exclusively from the training partition.

For the main chronological evaluation, the data were partitioned before sample construction into training (2000–2015), validation (2016–2019), and test (2020–2023) sets. In all three partitions, sliding windows were generated separately within each site-year monitoring season using a stride of one calendar period. Adjacent windows could overlap, and every window remained within a single year and monitoring season. Each sample contained 12 historical calendar periods followed by six consecutive target periods. On the March–October calendar grid, the 12-period history places the earliest eligible forecast target at period 13, corresponding to 1–5 May. With six periods per month, this corresponds to approximately two months of historical context and one month of forecasts within a monitoring season. The history provides context for recent activity levels and temporal changes, while the six-step output describes anticipated activity over successive periods for near-term surveillance planning. All model comparisons used this fixed 12-to-6 forecasting task.

Missing historical observations remained at their corresponding temporal positions and were represented by modality-specific masks [8]. Accepted windows contained valid light-trap targets at all six future horizons. Field-survey targets retained independent horizon-specific masks and contributed to the auxiliary objective at their observed positions. The resulting training, validation, and test sets contained 3151, 855, and 836 samples, respectively.

Figure 1 shows how the two observation streams were aligned on a common calendar grid while retaining independent masks, then partitioned chronologically before 12-to-6 window construction.

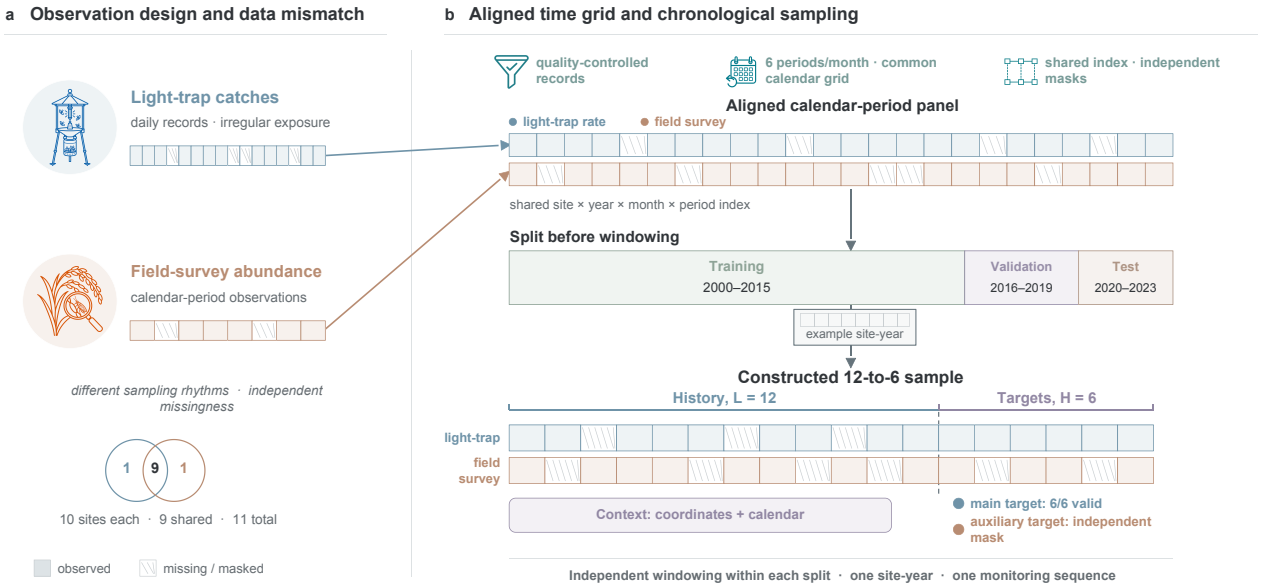



**Figure 1.** Construction of chronologically partitioned RLR forecasting samples. (**a**) Light-trap and field-survey records were obtained from partially overlapping monitoring sites. (**b**) The two observation streams were mapped to a common six-period-per-month calendar grid with independent masks and divided into training, validation, and test years before window construction. Each sample contained a 12-period history and six future targets. All six light-trap targets were valid, whereas field-survey targets retained independent masks; calendar variables and coordinates provided contextual inputs. Arrows indicate the progression from source observations through calendar alignment and chronological partitioning to forecasting-window construction.

### 2.2. Forecasting Task

For site $s$ and forecast origin $t$, the historical light-trap input $\mathcal{X}^{L}_{s,t}$ comprised $\{r_{s,\tau}, o_{s,\tau}, E_{s,\tau}, m^{L}_{s,\tau}\}$ over $\tau = t - L + 1, \ldots, t$. The corresponding field-survey input $\mathcal{X}^{F}_{s,t}$ comprised $\{f_{s,\tau}, m^{F}_{s,\tau}\}$, where $f_{s,\tau}$ denotes field-survey abundance. The model additionally received historical calendar features $\mathbf{C}^{\text{hist}}_{s,t}$, future calendar features $\mathbf{C}^{\text{fut}}_{s,t}$, and the site coordinate vector $\mathbf{q}_s = [\text{longitude}_s, \text{latitude}_s]$.

For calendar-grid position $\kappa \in \{1, \ldots, 48\}$ and month $m$, the four-dimensional calendar representation was defined below. The position index increased from 1 to 48 within each site-year monitoring sequence and was reset at the start of the following monitoring season.

$$\mathbf{c}_{\kappa,m} = \left[\sin\left(\frac{2\pi(\kappa-1)}{48}\right), \cos\left(\frac{2\pi(\kappa-1)}{48}\right), \sin\left(\frac{2\pi(m-1)}{12}\right), \cos\left(\frac{2\pi(m-1)}{12}\right)\right]. \quad (2)$$

DRAMM-Pest generated all six future horizons jointly in a direct multi-horizon formulation. The forecasting function was defined as

$$\left(\hat{\mathbf{p}}_{s,t+1:t+H}, \hat{\mathbf{r}}_{s,t+1:t+H}, \hat{\mathbf{f}}_{s,t+1:t+H}\right) = \mathcal{F}_{\theta}\left(\mathcal{X}^{L}_{s,t}, \mathcal{X}^{F}_{s,t}, \mathbf{C}^{\text{hist}}_{s,t}, \mathbf{C}^{\text{fut}}_{s,t}, \mathbf{q}_s\right), \quad (3)$$

where $\hat{\mathbf{p}}_{s,t+1:t+H}$ contains the positive-catch probabilities, $\hat{\mathbf{r}}_{s,t+1:t+H}$ is the main light-trap-rate forecast, and $\hat{\mathbf{f}}_{s,t+1:t+H}$ is the auxiliary field-survey forecast. The historical input length and forecasting horizon were fixed at $L = 12$ and $H = 6$, respectively.

### 2.3. DRAMM-Pest Architecture

Figure 2 summarizes the DRAMM-Pest workflow from measurement-specific encoding to horizon-specific outputs.

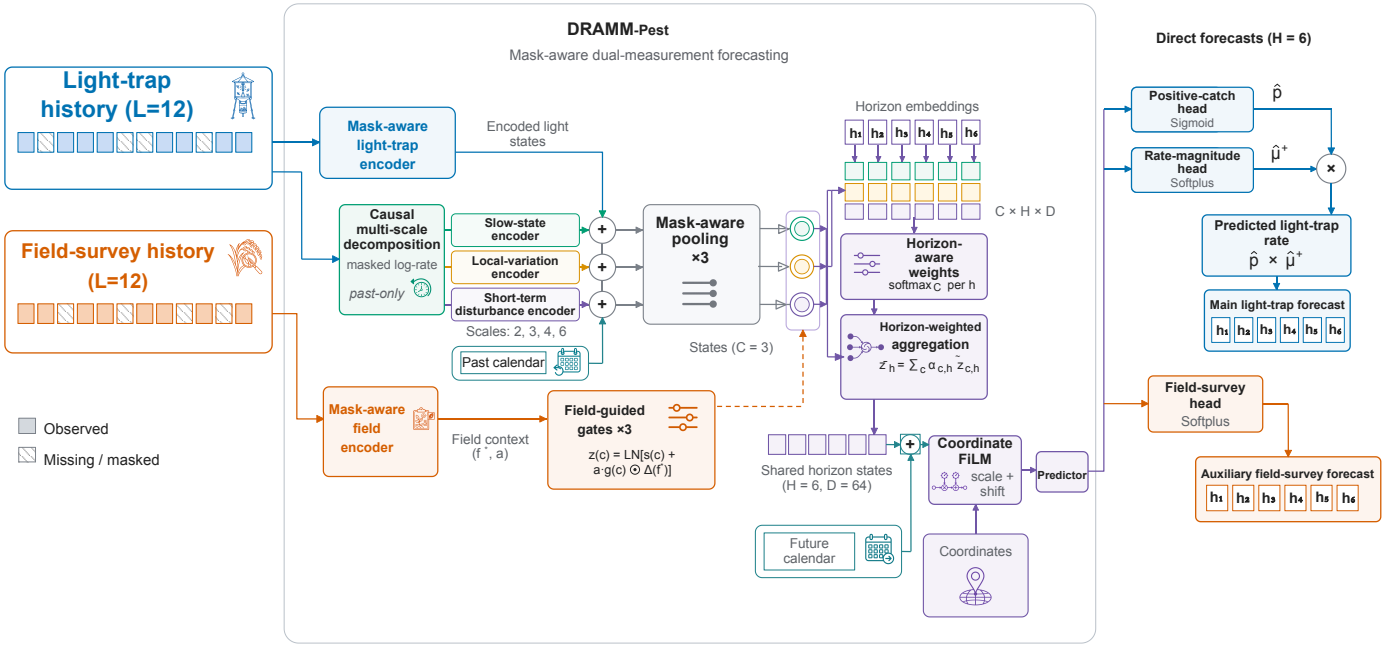



**Figure 2.** Overall workflow of DRAMM-Pest for dual-measurement RLR activity forecasting. Separate mask-aware encoders represent light-trap and field-survey histories. Causal component modeling separates the transformed light-trap rate history, $x = \log(1 + r)$, into slow-state, local-variation, and short-term-disturbance states, which are modulated by available field context and aggregated using horizon-specific weights. Future calendar features and Coordinate FiLM condition the shared horizon representations. The positive-catch head produces $\hat{p}$, and the rate-magnitude head produces a learned positive scaling factor $\hat{\mu}^{+}$. Their product gives the light-trap-rate point forecast, $\hat{r} = \hat{p}\hat{\mu}^{+}$,

trained under the joint objective; the field-survey head provides auxiliary supervision. Arrows indicate information flow, and the dashed orange arrow denotes field-guided modulation. Blue and orange identify the light-trap and field-survey pathways; green, gold, and purple distinguish the three temporal-component branches. The asterisk in $\mathbf{f}^*$ denotes the pooled field representation. The symbols + and ⊙ denote addition and element-wise multiplication, respectively; × denotes multiplication, and ×3 indicates application to all three components.

#### 2.3.1. Mask-Aware Measurement Encoding

The light-trap encoder used $\mathbf{x}^L_{s,\tau} = [\log(1 + r_{s,\tau}), o_{s,\tau}, \log(1 + E_{s,\tau})]$, whereas the field-survey encoder used $\mathbf{x}^F_{s,\tau} = [\log(1+f_{s,\tau})]$. The continuous components of these input vectors were standardized using training-set statistics before projection. For modality $v \in \{L, F\}$, the validity indicator was concatenated with these features before projection. The projected sequence was masked and processed by causal convolutions with kernel widths 2 and 3:

$$\begin{aligned} \mathbf{p}^v_{s,\tau} &= W_v[\mathbf{x}^v_{s,\tau}; m^v_{s,\tau}] + \mathbf{b}_v, \\ \tilde{\mathbf{p}}^v_{s,\tau} &= m^v_{s,\tau}\mathbf{p}^v_{s,\tau}, \\ \mathbf{e}^v_{s,\tau} &= m^v_{s,\tau}\,\mathrm{LN}\Big[\tilde{\mathbf{p}}^v_{s,\tau} + \mathrm{CausalConv}_{\{2,3\}}\big(\tilde{\mathbf{p}}^v_{s,1:\tau}\big)\Big]. \end{aligned} \tag{4}$$

The causal bank used separate left-padded convolution branches with widths 2 and 3. Each branch applied GELU activation, and the concatenated outputs were linearly projected to $D$ dimensions and normalized.

For a latent sequence $\mathbf{E}$ and observation mask $\mathbf{m}$, mask-aware temporal pooling combined the masked mean and last valid state. Applying this operator to the field-survey sequence produced its historical summary:

$$\begin{aligned} \Pi(\mathbf{E}, \mathbf{m}) &= \frac{1}{2}\left[\frac{\sum_{\tau=1}^{L} m_\tau \mathbf{E}_\tau}{\max\left(1, \sum_{\tau=1}^{L} m_\tau\right)} + \mathrm{LastValid}(\mathbf{E}, \mathbf{m})\right], \\ \mathbf{f}^*_s &= \Pi\left(\mathbf{E}^F_s, \mathbf{m}^F_s\right). \end{aligned} \tag{5}$$

The operator $\mathrm{LastValid}(\mathbf{E}, \mathbf{m})$ returned the zero vector when the sequence contained no valid observation.

#### 2.3.2. Causal Multi-Scale Component Modeling

The decomposition module operated on the transformed light-trap sequence $x_{s,\tau} = \log(1 + r_{s,\tau})$. Exponential-smoothing decomposition has been used in pest forecasting [15]. We defined a learnable one-sided exponential moving average for masked calendar-period observations to generate the slow state:

$$s_{s,\tau} = \begin{cases} s_{s,\tau-1}, & m^L_{s,\tau} = 0, \\ x_{s,\tau}, & m^L_{s,\tau} = 1 \text{ and } \tau \text{ is the first valid period}, \\ \alpha x_{s,\tau} + (1-\alpha)s_{s,\tau-1}, & m^L_{s,\tau} = 1 \text{ and a previous valid state exists}. \end{cases} \tag{6}$$

The smoothing coefficient $\alpha$ was learnable, initialized at 0.25, and constrained to $[0.01, 0.99]$. For the causal scale set $\mathcal{K} = \{2, 3, 4, 6\}$, the detrended residual, left-sided masked averages, local variation, and short-term disturbance were computed as

$$\begin{aligned} d_{s,\tau} &= m^L_{s,\tau}(x_{s,\tau} - s_{s,\tau}), \\ A_k(d)_{s,\tau} &= \begin{cases} \dfrac{\sum_{j=0}^{k-1} m^L_{s,\tau-j} d_{s,\tau-j}}{\sum_{j=0}^{k-1} m^L_{s,\tau-j}}, & \sum_{j=0}^{k-1} m^L_{s,\tau-j} > 0, \\ 0, & \text{otherwise}, \end{cases} \\ \ell_{s,\tau} &= \frac{1}{|\mathcal{K}|}\sum_{k\in\mathcal{K}} A_k(d)_{s,\tau}, \qquad u_{s,\tau} = m^L_{s,\tau}(d_{s,\tau} - \ell_{s,\tau}). \end{aligned} \tag{7}$$

The recursion used $s_{s,0} = 0$, $\Delta s_{s,1} = 0$, and $\Delta s_{s,\tau} = s_{s,\tau} - s_{s,\tau-1}$ for $\tau = 2, \ldots, L$. For indices $\tau - j < 1$, both $m^L_{s,\tau-j}$ and $d_{s,\tau-j}$ were set to zero.

Let $\mathcal{C} = \{S, V, U\}$ denote the slow-state, local-variation, and short-term-disturbance components. Their encoders received $\mathbf{v}^S_{s,\tau} = [s_{s,\tau}, \Delta s_{s,\tau}]$, $\mathbf{v}^V_{s,\tau} = [\ell_{s,\tau}, m^L_{s,\tau}]$, and $\mathbf{v}^U_{s,\tau} = [u_{s,\tau}, \ell_{s,\tau}, m^L_{s,\tau}]$, respectively. The slow-state and short-term encoders used Linear–GELU–LayerNorm blocks. The local-variation encoder used parallel left-padded convolutions at scales $\mathcal{K}$, followed by GELU activation, concatenation, projection, and LayerNorm. Each encoded component was combined with the light-trap representation and historical calendar embedding, then summarized by mask-aware pooling:

$$\begin{aligned} \mathbf{y}_{c,s,\tau} &= m^L_{s,\tau}\,\mathrm{LN}\big[E_c(\mathbf{v}^c_{s,\tau}) + \mathbf{e}^L_{s,\tau} + W^{\mathrm{hist}}_{\mathrm{cal}}\mathbf{c}^{\mathrm{hist}}_{s,\tau}\big], \quad c \in \mathcal{C}, \\ \mathbf{s}_{c,s} &= \Pi\big(\mathbf{Y}_{c,s}, \mathbf{m}^L_s\big) \in \mathbb{R}^D, \qquad D = 64. \end{aligned} \tag{8}$$

Here, $\mathbf{Y}_{c,s} = \{\mathbf{y}_{c,s,\tau}\}_{\tau=1}^{L}$ denotes the encoded sequence for component $c$.

### 2.3.3. Field-Guided Latent Modulation

Field-history availability was summarized by $a_s$, which activated field guidance when at least one historical field observation was available. For each temporal component, a gate combined the light-trap state, field-survey summary, and availability indicator. The resulting field-guided state was

$$\begin{aligned} a_s &= \mathbb{I}\Big(\sum_{\tau=1}^{L} m^F_{s,\tau} > 0\Big), \\ \mathbf{g}_{c,s} &= \sigma\big(W_c[\mathbf{s}_{c,s}; \mathbf{f}^*_s; a_s] + \mathbf{b}_c\big), \\ \Delta(\mathbf{f}^*_s) &= W_\Delta \mathbf{f}^*_s + \mathbf{b}_\Delta, \\ \mathbf{z}_{c,s} &= \mathrm{LN}\big[\mathbf{s}_{c,s} + a_s \mathbf{g}_{c,s} \odot \Delta(\mathbf{f}^*_s)\big]. \end{aligned} \tag{9}$$

This operation produced three $D$-dimensional component states for each sample. When the historical window contains no valid field observation, the field-modulation term is zero in both the adaptive-gate and fixed-gate configurations.

### 2.3.4. Horizon-Aware Component Aggregation and Coordinate Conditioning

Horizon-specific decoding has been used to distinguish future prediction steps in multi-period forecasting [16]. Here, a learnable horizon embedding was added to each of the three component states, and a shared network computed normalized weights for combining them before the future calendar embedding was added.

$$\begin{aligned} \tilde{\mathbf{z}}_{c,s,h} &= \mathbf{z}_{c,s} + \mathbf{e}_h, \\ \xi_{c,s,h} &= W_2 \tanh\big(W_1 \tilde{\mathbf{z}}_{c,s,h} + \mathbf{b}_1\big) + b_2, \\ \alpha_{c,s,h} &= \frac{\exp(\xi_{c,s,h})}{\sum_{c' \in \mathcal{C}} \exp(\xi_{c',s,h})}, \\ \bar{\mathbf{z}}_{s,h} &= \sum_{c \in \mathcal{C}} \alpha_{c,s,h} \tilde{\mathbf{z}}_{c,s,h}, \\ \mathbf{u}_{s,h} &= \bar{\mathbf{z}}_{s,h} + W^{\mathrm{fut}}_{\mathrm{cal}} \mathbf{c}^{\mathrm{fut}}_{s,h}, \qquad h = 1, \ldots, H. \end{aligned} \tag{10}$$

The coordinate encoder used six bounded features and a hidden width of 32. Coordinate-dependent scale and shift vectors were applied through feature-wise linear modulation (FiLM) [24] and shared across the six forecast horizons:

$$\begin{aligned}
\boldsymbol{\phi}(\mathbf{q}_s) &= [\mathrm{lon}_s/180,\ \mathrm{lat}_s/90,\ \sin(\mathrm{lon}_{s,\mathrm{rad}}),\ \cos(\mathrm{lon}_{s,\mathrm{rad}}),\\
&\qquad \sin(\mathrm{lat}_{s,\mathrm{rad}}),\ \cos(\mathrm{lat}_{s,\mathrm{rad}})],\\
\mathbf{v}_s^{\mathrm{coord}} &= \mathrm{LN}\big[\mathrm{GELU}\big(W_{\mathrm{coord}}\boldsymbol{\phi}(\mathbf{q}_s) + \mathbf{b}_{\mathrm{coord}}\big)\big],\\
[\tilde{\boldsymbol{\gamma}}_s; \tilde{\boldsymbol{\beta}}_s] &= W_F\mathbf{v}_s^{\mathrm{coord}} + \mathbf{b}_F,\\
\boldsymbol{\gamma}_s &= 0.25\tanh(\tilde{\boldsymbol{\gamma}}_s), \qquad \boldsymbol{\beta}_s = 0.25\tanh(\tilde{\boldsymbol{\beta}}_s),\\
\mathbf{u}'_{s,h} &= \mathbf{u}_{s,h} \odot (1 + \boldsymbol{\gamma}_s) + \boldsymbol{\beta}_s.
\end{aligned} \tag{11}$$

#### 2.3.5. Observation-Specific Prediction Heads

Dual-task frameworks have combined categorical recognition with continuous estimation, including gait-phase recognition and hip-angle prediction [25]. The coordinate-conditioned representation was processed by a shared LayerNorm–Linear–GELU–Dropout forecast block to obtain $\mathbf{h}_{s,h}$. Each prediction head then applied a 64-to-32 linear layer, a GELU activation, and a 32-to-1 output layer. The prediction heads produced the positive-catch probability, a positive magnitude factor, and auxiliary field-survey abundance. The light-trap-rate forecast was obtained by multiplying the probability by the magnitude factor:

$$\begin{aligned}
\hat{p}_{s,h} &= \sigma\big[F_{\mathrm{occ}}(\mathbf{h}_{s,h})\big],\\
\hat{\mu}^{+}_{s,h} &= \mathrm{softplus}\big[F_{\mathrm{rate}}(\mathbf{h}_{s,h})\big] + 10^{-6},\\
\hat{r}_{s,h} &= \hat{p}_{s,h}\hat{\mu}^{+}_{s,h},\\
\hat{f}_{s,h} &= \mathrm{softplus}\big[F_{\mathrm{field}}(\mathbf{h}_{s,h})\big].
\end{aligned} \tag{12}$$

Here, $\hat{\mu}^{+}_{s,h}$ is a learned positive scaling factor optimized through the loss on the final rate prediction, rather than a separately fitted conditional mean for positive catches. The product $\hat{r}_{s,h}$ is a point forecast under the joint objective defined below; this parameterization does not specify a hurdle likelihood or identify the conditional mean of the light-trap rate.

### *2.4. Learning Objective and Experimental Protocol*

Let $i$ index forecasting samples and $h$ index forecast horizons. For compact notation, $r_{i,h}$, $f_{i,h}$, $m^{L}_{i,h}$, and $m^{F}_{i,h}$ denote their target quantities. The positive-catch target was $o_{i,h} = \mathbb{I}(r_{i,h} > 0)$. The positive-catch loss was calculated over the valid light-trap targets:

$$\mathcal{L}_{\mathrm{occ}} = -\frac{1}{M_L}\sum_{i,h} m^{L}_{i,h}\big[o_{i,h}\log\hat{p}_{i,h} + (1 - o_{i,h})\log(1 - \hat{p}_{i,h})\big], \qquad M_L = \max\Big(1, \sum_{i,h} m^{L}_{i,h}\Big). \tag{13}$$

The rate regression loss optimized the product $\hat{r} = \hat{p}\hat{\mu}^{+}$ over all valid light-trap targets, including zeros, using Huber loss on the $\log(1 + r)$ scale. Both factors received gradients from this loss, while $\hat{p}$ additionally received binary cross-entropy supervision from the positive-catch labels.

$$\mathcal{L}_{\mathrm{rate}} = \frac{1}{M_L}\sum_{i,h} m^{L}_{i,h}\rho_\delta\big[\log(1 + \hat{r}_{i,h}) - \log(1 + r_{i,h})\big]. \tag{14}$$

The auxiliary field-survey loss was

$$\mathcal{L}_{\mathrm{field}} = \frac{1}{M_F}\sum_{i,h} m^{F}_{i,h}\rho_\delta\big[\log(1 + \hat{f}_{i,h}) - \log(1 + f_{i,h})\big], \qquad M_F = \max\Big(1, \sum_{i,h} m^{F}_{i,h}\Big). \tag{15}$$

Here, $\rho_\delta$ denotes the Huber function with $\delta = 1.0$: it equals $e^2/2$ when $|e| \le \delta$ and $\delta(|e| - \delta/2)$ otherwise. The joint learning objective was

$$\mathcal{L} = \mathcal{L}_{\mathrm{occ}} + \mathcal{L}_{\mathrm{rate}} + \lambda_{\mathrm{field}}\mathcal{L}_{\mathrm{field}}, \qquad \lambda_{\mathrm{field}} = 0.1. \tag{16}$$

DRAMM-Pest was implemented in PyTorch 2.4.1 with $D = 64$, $L = 12$, $H = 6$, causal local scales $\{2, 3, 4, 6\}$, and a shared forecast-block dropout rate of 0.05. The model was trained using AdamW with a learning rate of 0.001, weight decay of $10^{-4}$, and a batch size of 32, for a maximum of 30 epochs. Early stopping used a patience of five epochs and monitored validation light-trap MAE pooled across all six forecast horizons. With epochs numbered from one, early stopping occurred at epochs 23, 26, 21, 28, and 25 for seeds 2024–2028, respectively. The corresponding checkpoints with the lowest validation MAE, from epochs 18, 21, 16, 23, and 20, were used for testing. Training was performed on an NVIDIA GeForce RTX 4090 GPU (NVIDIA Corporation, Santa Clara, CA, USA).

Two additional controls examined the role of field information. The no-field control used the light-trap backbone, calendar variables, and coordinates, with the field encoder, field-modulation branch, availability indicator, and auxiliary prediction head and loss removed. The availability-only control retained the binary indicator of whether any valid field observation occurred in the historical window, while excluding field-abundance values, the full field-mask sequence, and field-target supervision. Both controls used the same chronological partitions, forecasting windows, training and early-stopping procedures, and five random seeds as the complete model.

Within each site, test windows were grouped by the number of valid field observations in the 12-period history: zero, 1–11, or 12. For each seed, MAE was calculated over identical valid target positions across the six horizons within each site and stratum. Paired differences were calculated as fixed-gate minus full, no-field minus availability-only, and availability-only minus full, and summarized as the mean and sample standard deviation across five seeds.

Six comparison models were evaluated under the same chronological partitions and forecasting windows. Persistence repeated the last valid light-trap rate over all future horizons and converted the repeated rates to deterministic positive-catch indicators. For positive-catch evaluation, Persistence therefore served as a hard-classification reference with outputs of zero or one; its Brier score equals its misclassification rate. For field-survey forecasts, it carried the most recent valid abundance in the 12-period history forward over all six horizons. Histories without valid field observations used the training-set site median, followed by the training-set global median when the site had no valid training records. Seasonal Mean calculated positive-catch prevalences and original-scale mean light-trap rates and field-survey abundances separately from valid training labels, retaining true zeros. Each horizon queried the target site and calendar period, defined by month and within-month monitoring period. Unavailable site–period estimates were replaced successively by the corresponding cross-site calendar-period estimate and the global training estimate.

Direct Ridge, GRU [26], XGBoost [27], and DLinear [28] used the same information categories, comprising light-trap and field-survey histories, separate availability masks, historical and future calendar variables, and site coordinates. Direct Ridge and XGBoost used flattened features and fitted horizon-specific models, whereas GRU encoded the historical sequence. Direct Ridge used regularized regression with log-transformed rate targets and a logistic model for positive-catch prediction. DLinear decomposed only the light-trap rate, with separate 12-to-6 linear mappings for the trend and remainder and an exogenous projection for the other variables. Table 3 reports the core configurations and hyperparameter-search ranges.

The architectures of DRAMM-Pest and GRU were fixed, with validation data used for early stopping. Hyperparameters for DLinear, Direct Ridge, and XGBoost were selected according to validation performance within the search ranges in Table 3. XGBoost row and column sampling ratios were fixed at 0.8 throughout the search.

**Table 3.** Core configurations and hyperparameter-search ranges of the trainable forecasting models.

| Model | Reported Configuration | Validation Procedure |
|---|---|---|
| DRAMM-Pest | Hidden dimension 64; local scales $\{2, 3, 4, 6\}$; dropout 0.05 | Fixed architecture; validation-based early stopping and checkpoint selection |
| GRU | One layer with 64 hidden units | Fixed architecture; validation-based early stopping |
| DLinear | Decomposition window 3; separate 12-to-6 trend and remainder mappings; exogenous projection | Decomposition window searched over $\{3, 5\}$ |
| Direct Ridge | Horizon-specific regression; $\alpha = 10$; $\log(1 + \text{rate})$ target | $\alpha$ searched over $\{0.001, 0.01, 0.1, 1, 10, 100\}$ |
| XGBoost | 300 trees; maximum depth 3; learning rate 0.05; row and column sampling ratios 0.8 | Maximum depth $\{3, 5, 7\}$, learning rate $\{0.03, 0.05, 0.1\}$, and number of trees $\{200, 300\}$; 18 combinations |

Positive-catch forecasts were evaluated at valid light-trap target positions using the area under the precision–recall curve (AUPRC), the area under the receiver operating characteristic curve (AUROC), the Brier score, and the Brier skill score (BSS). For each forecast horizon, the climatological reference probability was estimated from the positive-catch frequency among valid training-window targets. BSS was calculated as one minus the ratio of the model Brier score to the climatological Brier score, with both scores evaluated at the same test positions. The observed test prevalence at each horizon defined its no-skill AUPRC reference [29]. Reliability diagrams for DRAMM-Pest and XGBoost were constructed separately for each seed by pooling valid predictions across all six horizons and comparing mean predicted probabilities with observed positive-catch frequencies in ten equal-width probability bins between zero and one. MAE, RMSE, and MAE on positive light-trap targets were calculated on the original light-trap-rate scale, whereas RMSLE was calculated on the $\log(1 + x)$ scale. Observed field-survey target positions defined the evaluation set for auxiliary MAE and RMSLE. For the overall results, all valid prediction positions from $h_1$ to $h_6$ were pooled within each run. DRAMM-Pest, GRU, DLinear, and XGBoost were each trained and evaluated using five random seeds, 2024, 2025, 2026, 2027, and 2028. Results for GRU, XGBoost, DLinear, and DRAMM-Pest were summarized as the mean and sample standard deviation over five random seeds. Persistence, Seasonal Mean, and Direct Ridge were reported as single-run results.

Target distributions were summarized over valid test prediction positions using the mean, median, zero proportion, and upper percentiles. Site-specific MAE skill was calculated as one minus the ratio of model MAE to Seasonal Mean MAE at identical target positions. Within each seed, skill scores were averaged equally across the ten sites. Results were summarized as the mean and sample standard deviation across five seeds.

High-abundance targets were defined as calendar-period light-trap rates strictly greater than 11.20 individuals trap-night$^{-1}$, the 90th percentile of 5001 unique valid site–calendar-period records in the 2000–2015 training set. MAE and mean signed error, defined as prediction minus observation, were calculated at these test positions. Over all valid test positions, hit rate was the fraction of observed exceedances also predicted to exceed this threshold, and false-alarm rate was the fraction of observed non-exceedances predicted to exceed it. Peak-timing MAE was evaluated over windows containing at least one high-abundance target as the mean absolute difference between predicted and observed maximum-rate positions across the six horizons, measured in calendar periods; ties were resolved using the earliest position. Each metric was calculated separately for each seed and summarized as the five-seed mean and sample standard deviation. High-abundance MAE and mean-signed-error differences were evaluated using the paired site-year bootstrap procedure described below, with the training threshold held fixed across resamples.

Seed-level standard deviations characterize training variability. For the comparison between DRAMM-Pest and XGBoost, paired performance differences were evaluated using 5000 bootstrap resamples of the 39 site-year blocks in the test set. Each block retained its accepted windows and all six forecast horizons. Identical block draws were used for both models and all five seeds. Within each resample, pooled metrics were recalculated separately for each model and seed, and paired differences were averaged across seeds without ensembling predictions. The 2.5th and 97.5th percentiles defined the 95% confidence intervals, conditional on the five fitted seeds. For the Coordinate FiLM and auxiliary-supervision ablations, paired $t$ intervals were calculated from the five seed-level MAE differences with four degrees of freedom. These intervals describe seed-level variability rather than site-year sampling variability. All intervals were reported individually without multiplicity adjustment.

Sensitivity analyses evaluated performance separately for 2020–2023 and on a subset with non-overlapping target intervals. Within each site-year, the earliest accepted window was retained, followed successively by the next window whose target interval began after the preceding retained target interval ended. This procedure retained 161 windows and 966 prediction positions. Historical inputs could still overlap.

Spatial transfer was evaluated using five two-site holdout folds across the ten sites contributing forecasting windows, as listed in Table 2. The held-out pairs were S001 and S004 in F1, S008 and S011 in F2, S003 and S007 in F3, S002 and S009 in F4, and S005 and S006 in F5. Each fold used the remaining eight sites for training during 2000–2015 and validation during 2016–2019. Held-out sites were excluded from training, validation, and preprocessing-statistic estimation in all years; evaluation used their 2020–2023 windows. Models and preprocessing statistics were fitted independently within each fold. DRAMM-Pest, its Coordinate FiLM ablation, XGBoost, and GRU used seeds 2024–2028. Seasonal Mean used calendar-period means calculated from valid 2000–2015 records of the eight training sites, with the same training subset's global mean used when a calendar-period estimate was unavailable. MAE was pooled across held-out windows and six horizons within each seed and summarized as the five-seed mean and sample standard deviation; Seasonal Mean was reported as a single deterministic result.

Missing-input sensitivity was evaluated by randomly masking an additional 10%, 30%, or 50% of historical light-trap observations, field-survey observations, or both during inference. The trained model and forecast targets were held fixed, and masked inputs were filled using training-derived stratified medians. Under dual-history masking, the mask-aware condition updated the observation masks to reflect the added missingness. The original-mask reference used the same checkpoints and filled inputs while retaining the masks recorded before additional masking. Both conditions therefore produced identical predictions at 0% additional masking. The compared conditions shared ten random masking realizations at each masking level. Within each seed, pooled MAE and RMSLE were calculated for each realization and averaged across the ten realizations, then summarized as the mean and sample standard deviation across five seeds.

## 3. Results

### *3.1. Overall Forecasting Performance*

Table 4 summarizes the overall forecasting performance of DRAMM-Pest and six baseline models on the held-out 2020–2023 RLR test set. Under the pooled six-horizon evaluation, DRAMM-Pest yielded the highest reported AUPRC, AUROC, and BSS point estimates and the lowest reported error point estimates.

**Table 4.** Overall forecasting performance on the held-out 2020–2023 test set. Metrics were computed using pooled aggregation over all valid prediction positions across the six forecast horizons. Values for GRU, XGBoost, DLinear, and DRAMM-Pest are reported as the mean ± sample standard deviation over five random seeds. Persistence, Seasonal Mean, and Direct Ridge are reported as single-run results. Field-survey metrics were calculated only at observed target positions. BSS is relative to horizon-specific training climatology, whose pooled test Brier score is 0.195106. The pooled no-skill AUPRC reference is 0.806818. Bold values indicate the most favorable mean or single-run estimate, rather than statistical significance. Upward and downward arrows indicate that higher and lower metric values are preferable, respectively.

**Panel A. Positive-Catch Forecasting**

| Model | AUPRC ↑ | AUROC ↑ | Brier Score ↓ | BSS ↑ |
|---|---|---|---|---|
| Persistence | 0.918 | 0.8035 | 0.200 | −0.0249 |
| Seasonal Mean | 0.946 | 0.8314 | 0.170 | 0.1271 |
| Direct Ridge | 0.979 | 0.9128 | 0.105 | 0.4606 |
| GRU | 0.980 ± 0.003 | 0.9181 ± 0.0095 | 0.094 ± 0.005 | 0.5180 ± 0.0258 |
| XGBoost | 0.984 ± 0.001 | 0.9332 ± 0.0040 | 0.093 ± 0.003 | 0.5231 ± 0.0154 |
| DLinear | 0.982 ± 0.002 | 0.9249 ± 0.0081 | 0.094 ± 0.005 | 0.5159 ± 0.0259 |
| **DRAMM-Pest** | **0.987 ± 0.001** | **0.9464 ± 0.0059** | **0.084 ± 0.004** | **0.5671 ± 0.0228** |

**Panel B. Light-Trap Rate Forecasting**

| Model | MAE ↓ | RMSE ↓ | RMSLE ↓ | Positive-Target MAE ↓ |
|---|---|---|---|---|
| Persistence | 72.19 | 258.04 | 1.593 | 89.15 |
| Seasonal Mean | 62.32 | 252.26 | 1.610 | 76.93 |
| Direct Ridge | 56.95 | 242.17 | 1.306 | 70.46 |
| GRU | 61.03 ± 0.41 | 252.50 ± 0.69 | 1.325 ± 0.047 | 75.53 ± 0.50 |
| XGBoost | 55.06 ± 0.38 | 238.64 ± 0.83 | 1.087 ± 0.028 | 68.13 ± 0.47 |
| DLinear | 58.60 ± 0.49 | 248.59 ± 0.95 | 1.185 ± 0.042 | 72.52 ± 0.61 |
| **DRAMM-Pest** | **51.00 ± 0.72** | **230.47 ± 1.89** | **0.837 ± 0.029** | **63.11 ± 0.89** |

**Panel C. Auxiliary Field-Survey Forecasting (Observed Targets Only)**

| Model | Field-Survey MAE ↓ | Field-Survey RMSLE ↓ |
|---|---|---|
| Persistence | 157.92 | 2.105 |
| Seasonal Mean | 171.18 | 1.787 |
| Direct Ridge | 154.78 | 1.627 |
| GRU | 157.39 ± 2.59 | 1.638 ± 0.024 |
| XGBoost | 145.85 ± 2.21 | 1.502 ± 0.021 |
| DLinear | 152.11 ± 2.54 | 1.628 ± 0.024 |
| **DRAMM-Pest** | **142.75 ± 3.45** | **1.471 ± 0.032** |

For positive-catch forecasting, DRAMM-Pest achieved AUROC and BSS values of 0.9464 ± 0.0059 and 0.5671 ± 0.0228, compared with 0.9332 ± 0.0040 and 0.5231 ± 0.0154 for XGBoost. Its Brier score of 0.084 ± 0.004 represented a 56.71% reduction relative to the training-derived climatological reference. Persistence and Seasonal Mean yielded BSS values of −0.0249 and 0.1271, respectively. The paired comparison with XGBoost showed an AUPRC increase of 0.00327 (95% CI, 0.00197–0.00508) and a Brier-score reduction of 0.00859 (95% CI, 0.00743–0.00987).

For light-trap-rate forecasting, the DRAMM-Pest point estimates were an MAE of 51.00±0.72, an RMSE of 230.47±1.89, an RMSLE of 0.837±0.029, and an MAE of 63.11 ± 0.89 on positive light-trap targets. The corresponding XGBoost values were 55.06 ± 0.38, 238.64 ± 0.83, 1.087 ± 0.028, and 68.13 ± 0.47, respectively. The absolute RMSLE reduction was 0.25023 (95% CI, 0.22893–0.27095).

For the auxiliary field-survey task, DRAMM-Pest yielded an MAE of 142.75 ± 3.45 and an RMSLE of 1.471 ± 0.032 at observed target positions. XGBoost produced corresponding

values of 145.85 ± 2.21 and 1.502 ± 0.021, respectively. Both mean errors were lower for DRAMM-Pest.

The paired site-year bootstrap intervals favored DRAMM-Pest over XGBoost for all four evaluated metrics. The mean MAE reduction was 4.0661 (95% CI, 3.0607–5.1844). MAE differences favored DRAMM-Pest in each test year, ranging from 2.3579 to 5.0389, and remained positive at 3.8485 in the target-non-overlapping subset. These analyses show that the direction of the MAE difference was retained across the evaluated years and after removing overlap between target intervals.

Across the 5016 valid test prediction positions, the mean and median light-trap rates were 65.71 and 3.90 individuals per valid trap night, respectively, with 19.32% zero values. The 90th, 95th, and 99th percentiles were 127.00, 316.20, and 1124.60, indicating a strongly right-skewed target distribution. Site-specific mean rates ranged from 0.63 to 178.05.

Table 5 reports site-specific forecasting performance. DRAMM-Pest achieved lower mean MAE than XGBoost and Seasonal Mean at all ten sites. Its site-specific MAE skill scores ranged from 0.1284 to 0.5955. The equal-site mean skill was 0.2651 ± 0.0131 for DRAMM-Pest and 0.1817 ± 0.0044 for XGBoost, corresponding to average site-specific MAE reductions of 26.51% and 18.17% relative to Seasonal Mean.

**Table 5.** Site-specific light-trap-rate forecasting performance on the 2020–2023 test set. Counts represent valid prediction positions across six horizons. Target means and MAEs are expressed in individuals per valid trap night. MAE skill is relative to Seasonal Mean. DRAMM-Pest and XGBoost results are five-seed means ± sample standard deviations; Seasonal Mean results are single-run values.

| Site | Valid Positions | Target Mean | DRAMM-Pest MAE | XGBoost MAE | Seasonal Mean MAE | DRAMM-Pest MAE Skill | XGBoost MAE Skill |
|---|---|---|---|---|---|---|---|
| S001 | 588 | 21.97 | 13.77 ± 0.34 | 15.87 ± 0.07 | 19.82 | 0.3052 ± 0.0170 | 0.1991 ± 0.0035 |
| S002 | 504 | 43.88 | 31.21 ± 0.55 | 34.77 ± 0.26 | 40.29 | 0.2253 ± 0.0137 | 0.1369 ± 0.0065 |
| S003 | 564 | 13.95 | 8.16 ± 0.20 | 9.51 ± 0.05 | 12.69 | 0.3568 ± 0.0160 | 0.2507 ± 0.0040 |
| S004 | 378 | 0.63 | 0.71 ± 0.02 | 0.86 ± 0.02 | 1.76 | 0.5955 ± 0.0098 | 0.5142 ± 0.0130 |
| S005 | 492 | 30.99 | 22.40 ± 0.36 | 24.86 ± 0.21 | 27.06 | 0.1722 ± 0.0132 | 0.0811 ± 0.0077 |
| S006 | 504 | 145.72 | 125.23 ± 1.37 | 131.49 ± 1.11 | 143.68 | 0.1284 ± 0.0095 | 0.0848 ± 0.0077 |
| S007 | 468 | 178.05 | 147.36 ± 1.39 | 155.67 ± 1.29 | 173.79 | 0.1521 ± 0.0080 | 0.1043 ± 0.0074 |
| S008 | 576 | 14.06 | 8.48 ± 0.22 | 9.97 ± 0.03 | 12.50 | 0.3213 ± 0.0178 | 0.2020 ± 0.0020 |
| S009 | 504 | 103.77 | 76.33 ± 1.23 | 83.41 ± 0.44 | 95.97 | 0.2047 ± 0.0129 | 0.1309 ± 0.0046 |
| S011 | 438 | 123.39 | 92.80 ± 1.74 | 101.62 ± 0.51 | 114.54 | 0.1898 ± 0.0152 | 0.1128 ± 0.0044 |

### *3.2. Horizon-Wise Performance and Illustrative Forecasts*

Figure 3 presents horizon-wise forecasting performance and positive-catch probability reliability on the held-out 2020–2023 test set.

In Figure 3a, the no-skill AUPRC values were 0.7201, 0.7608, 0.7967, 0.8301, 0.8565, and 0.8768 at $h_1$–$h_6$, respectively, reflecting the increase in observed positive-catch prevalence. DRAMM-Pest achieved the highest reported AUPRC and lowest reported Brier-score point estimates at every horizon. In Figure 3e, observed positive-catch frequencies exceeded mean predicted probabilities in most intermediate-probability bins for both DRAMM-Pest and XGBoost, indicating underprediction of positive-catch probabilities.

For light-trap-rate forecasting, MAE and RMSLE generally increased from $h_1$ to $h_6$ for all methods. DRAMM-Pest yielded the lowest reported error point estimates throughout the six-horizon range, with the largest separation from XGBoost appearing in late-horizon RMSLE.

Figure 4 further illustrates the horizon-wise behavior through three contrasting test cases: near-zero activity, a rapid rise to an extreme peak, and partially missing light-trap history.

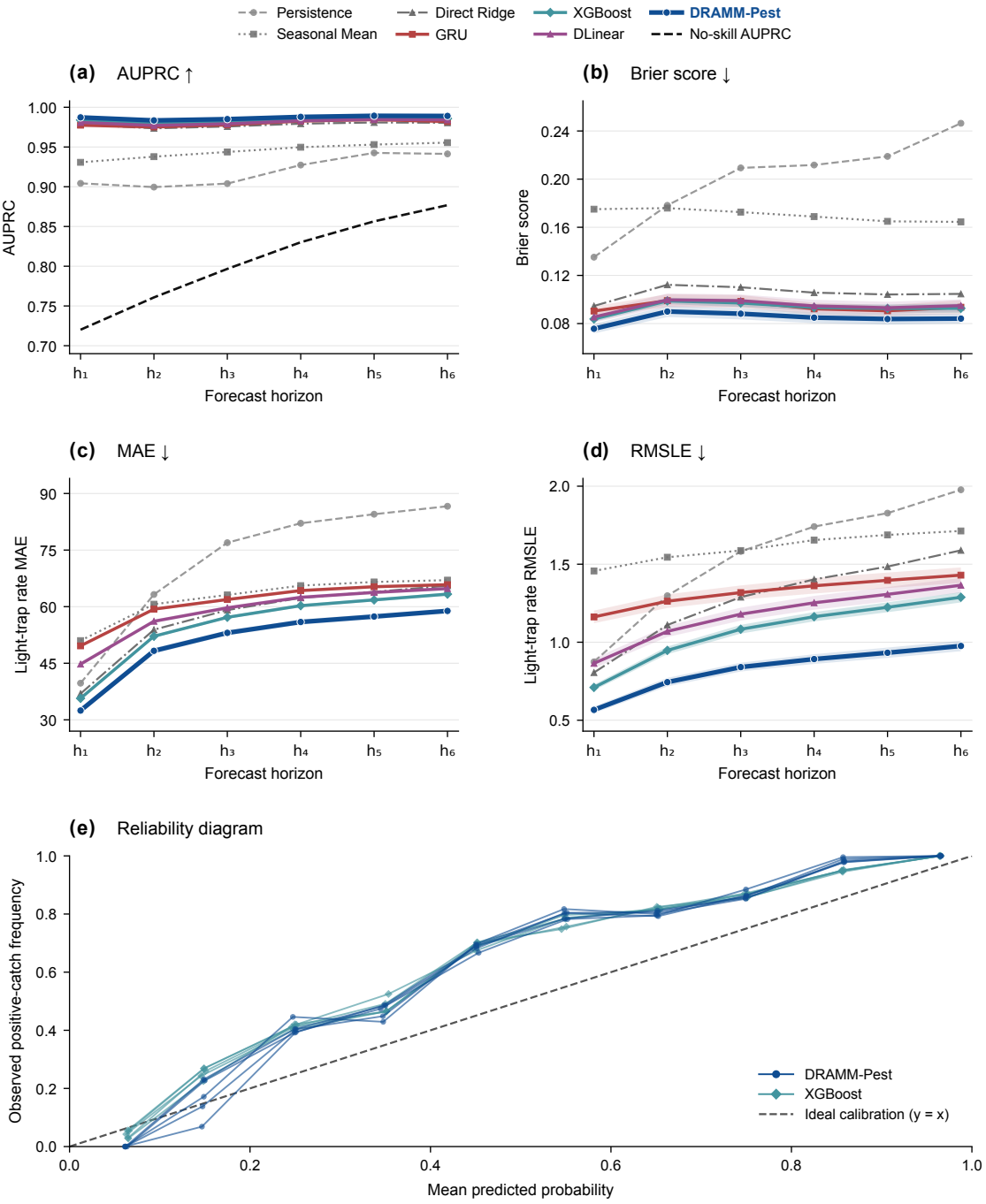



**Figure 3.** Horizon-wise forecasting performance and probability reliability on the held-out 2020–2023 test set. (**a**) AUPRC and (**b**) Brier score for positive-catch forecasting; (**c**) MAE and (**d**) RMSLE for light-trap-rate forecasting; (**e**) reliability diagrams for DRAMM-Pest and XGBoost. In panels (**a**–**d**), each horizon was evaluated on 836 accepted test windows. Curves and shaded bands for GRU, XGBoost, DLinear, and DRAMM-Pest show the five-seed mean and ±1 sample standard deviation; Persistence, Seasonal Mean, and Direct Ridge are single-run results. The black dashed line in (**a**) shows the no-skill AUPRC reference defined by the test prevalence at each horizon. In (**e**), each curve represents one seed evaluated over 5016 valid prediction positions using ten equal-width probability bins. The gray dashed line denotes ideal calibration. Upward and downward arrows indicate that higher and lower metric values are preferable, respectively.

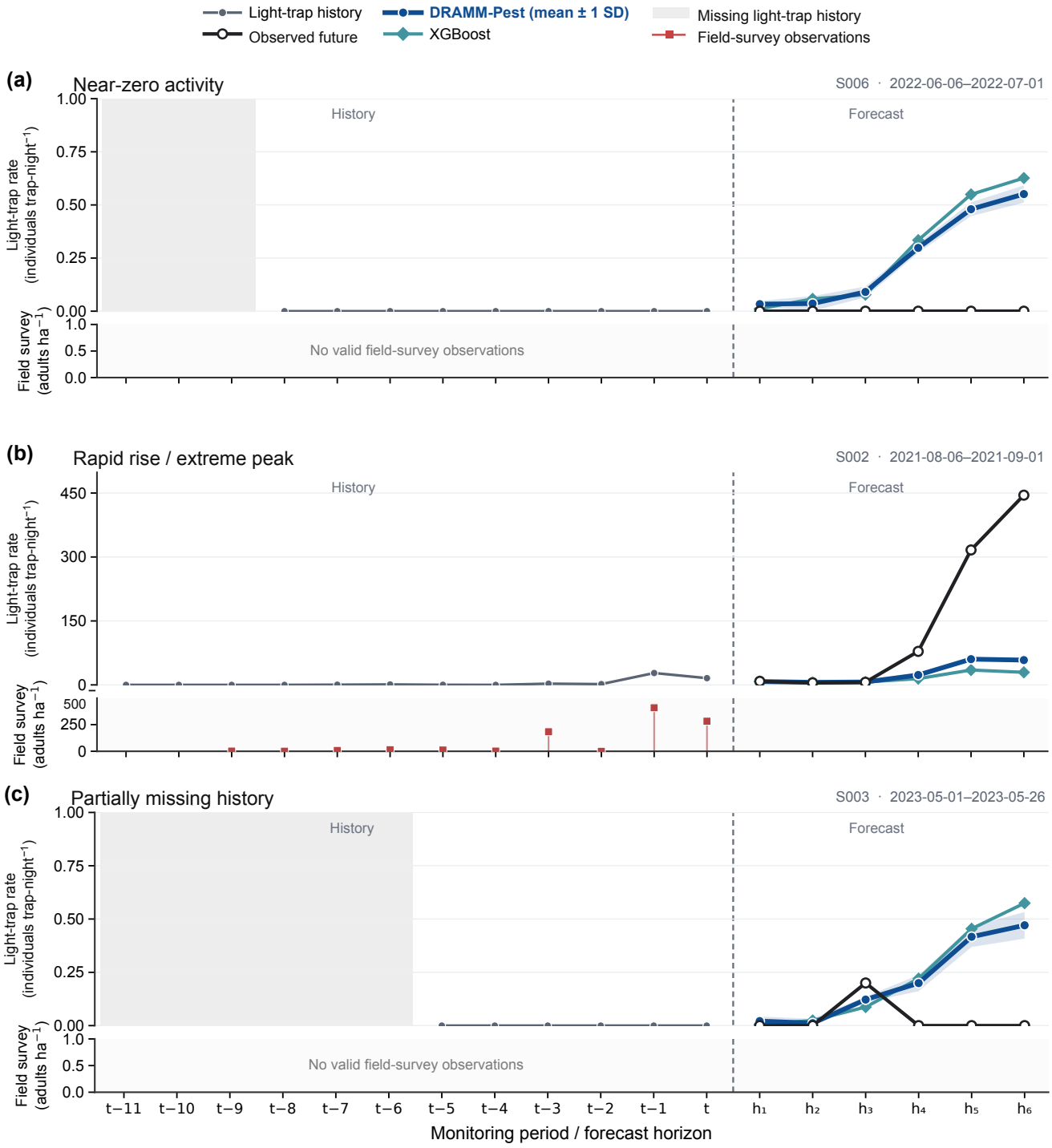



**Figure 4.** Illustrative multi-horizon forecasts under contrasting pest-activity conditions on the held-out 2020–2023 test set. (**a**) Near-zero activity at site S006; (**b**) a rapid rise to an extreme peak at site S002; and (**c**) partially missing light-trap history at site S003. Gray filled markers denote observed light-trap history, black open markers denote observed future values, the blue line and shaded band denote the DRAMM-Pest mean and ±1 sample standard deviation over five random seeds, and teal diamonds denote the XGBoost mean prediction over the same five seeds. Gray regions indicate missing light-trap history. Red stems and squares denote observed field-survey abundance; the absence of a red marker indicates an unobserved field-survey position rather than a zero value. The vertical dashed line separates the historical input from the six forecast horizons. Each panel uses an independent vertical-axis range.

In the near-zero-activity case, the observed future light-trap rate remained zero across all six horizons, whereas both models generated increasingly positive predictions toward the later horizons. DRAMM-Pest generally produced lower late-horizon predictions than XGBoost but still overestimated the observed activity. In the partially missing-history case, the first six historical periods were unavailable and the observed future contained a small pulse at $h_3$, followed by a return to zero. DRAMM-Pest responded more closely to the $h_3$ pulse and remained below XGBoost at the later horizons, but it projected a continuing increase from $h_4$ to $h_6$. Across both low-activity examples, the late-horizon forecasts remained above the observed near-zero values.

In the rapid-rise case, the observed light-trap rate increased sharply after $h_3$ and reached an extreme peak at $h_5$ and $h_6$. DRAMM-Pest produced a stronger response than XGBoost and was closer to the observed values at the peak horizons, but both models substantially underestimated the magnitude of the extreme increase. These illustrative cases show underestimation of an abrupt extreme peak and late-horizon overprediction in near-zero sequences.

The training-defined high-abundance subset contained 1988 of the 5016 valid test prediction positions across 489 of the 836 test windows. DRAMM-Pest achieved an MAE of $126.31 \pm 1.73$ individuals trap-night$^{-1}$, compared with $136.01 \pm 1.01$ for XGBoost. The paired MAE reduction was 9.70 (95% CI, 7.77–11.65). Mean signed errors were $-121.17 \pm 1.63$ and $-129.59 \pm 1.56$, respectively. The DRAMM-Pest-minus-XGBoost mean-signed-error difference was 8.42 (95% CI, 6.75–10.34), indicating a smaller average underprediction for DRAMM-Pest.

DRAMM-Pest also achieved a higher mean threshold-exceedance hit rate than XGBoost, $0.7721 \pm 0.0099$ versus $0.6426 \pm 0.0098$, and a lower false-alarm rate, $0.0300 \pm 0.0013$ versus $0.0375 \pm 0.0010$. Peak-timing MAE was $0.447 \pm 0.020$ calendar periods for DRAMM-Pest and $0.758 \pm 0.053$ for XGBoost.

### *3.3. Ablation and Within-Site Analysis*

Table 6 compares component ablations and field-information controls. The complete DRAMM-Pest model yielded the most favorable pooled point estimates across the four reported metrics. All variants used the same chronological data split, preprocessing procedure, training configuration, and five random seeds. Metrics were calculated by pooling valid predictions across $h_1$–$h_6$ on the held-out 2020–2023 test set and are reported as the mean ± sample standard deviation.

**Table 6.** Component ablations and field-information controls on the held-out 2020–2023 test set. Values were computed using pooled aggregation across $h_1$–$h_6$ and are reported as the mean ± sample standard deviation over five random seeds. Bold and underlined values indicate the first- and second-ranked point estimates, respectively. Upward and downward arrows indicate whether higher or lower values are preferable.

| Variant | Positive-Catch Forecasting | | Light-Trap Rate Forecasting | |
|---|---|---|---|---|
| | AUPRC ↑ | Brier Score ↓ | MAE ↓ | RMSLE ↓ |
| Without Three-Component Causal Decomposition | $0.980 \pm 0.003$ | $0.105 \pm 0.007$ | $58.45 \pm 1.25$ | $1.071 \pm 0.055$ |
| Fixed Field-Guided Gates ($\mathbf{g} = 1$) | $0.984 \pm 0.002$ | $0.094 \pm 0.005$ | $54.38 \pm 0.94$ | $0.927 \pm 0.039$ |
| Without Horizon-Aware Weighting | $0.982 \pm 0.002$ | $0.097 \pm 0.006$ | $55.72 \pm 1.05$ | $0.966 \pm 0.044$ |
| Without Coordinate FiLM | $\underline{0.986 \pm 0.002}$ | $\underline{0.088 \pm 0.005}$ | $\underline{52.18 \pm 0.78}$ | $\underline{0.865 \pm 0.032}$ |
| Without Auxiliary Supervision ($\lambda_{\text{field}} = 0$) | $0.985 \pm 0.002$ | $0.089 \pm 0.005$ | $52.82 \pm 0.84$ | $0.886 \pm 0.034$ |
| No Field Information | $0.985 \pm 0.001$ | $0.090 \pm 0.004$ | $54.12 \pm 0.83$ | $0.924 \pm 0.033$ |
| Availability Indicator Only | $0.985 \pm 0.001$ | $0.089 \pm 0.004$ | $53.74 \pm 0.79$ | $0.908 \pm 0.031$ |
| **Full DRAMM-Pest** | $\mathbf{0.987 \pm 0.001}$ | $\mathbf{0.084 \pm 0.004}$ | $\mathbf{51.00 \pm 0.72}$ | $\mathbf{0.837 \pm 0.029}$ |

The no-field control yielded a light-trap-rate MAE of 54.12 ± 0.83, while the availability-only control yielded 53.74 ± 0.79, compared with 51.00 ± 0.72 for the complete model. Based on the reported means, the corresponding MAE increases were 3.12 and 2.74. Both controls also had lower mean AUPRC and higher mean Brier scores and RMSLE than the complete model.

Replacing the three-component causal decomposition pathway with its undecomposed variant produced the largest deterioration across all four metrics. Relative to the complete model, AUPRC decreased by 0.007, while the Brier score, MAE, and RMSLE increased by 0.021, 7.45, and 0.234, respectively. Replacing horizon-aware component weighting with equal weights produced the second-largest deterioration, reducing AUPRC by 0.005 and increasing the Brier score, MAE, and RMSLE by 0.013, 4.72, and 0.129, respectively.

The remaining ablations produced smaller increases in error point estimates. Fixing the field-guided gates increased MAE from 51.00 to 54.38. Removing Coordinate FiLM and auxiliary supervision increased MAE by 1.1824 and 1.8224, respectively. Their paired-seed 95% $t$ intervals reported in Table 7 were 1.1062–1.2587 and 1.6717–1.9732. Both intervals lay above zero, supporting positive mean MAE differences across the evaluated seeds; they do not quantify variation associated with resampling site-year blocks.

**Table 7.** Paired performance differences and sensitivity analyses on the chronological test set. Positive differences favor DRAMM-Pest: AUPRC is DRAMM-Pest minus XGBoost; errors are comparator minus DRAMM-Pest. Panel A: 95% CIs from 5000 paired bootstrap resamples of 39 site-year blocks. Panel B: paired $t$ intervals across five seeds ($df = 4$), reflecting training variability. Panel C: descriptive five-seed mean MAE differences.

| **Panel A. DRAMM-Pest Versus XGBoost** | | |
|---|---|---|
| **Metric** | **Difference** | **95% Site-Year Bootstrap CI** |
| MAE | 4.0661 | [3.0607, 5.1844] |
| AUPRC | 0.00327 | [0.00197, 0.00508] |
| Brier score | 0.00859 | [0.00743, 0.00987] |
| RMSLE | 0.25023 | [0.22893, 0.27095] |
| **Panel B. Ablation Comparisons** | | |
| **Variant** | **MAE Increase** | **95% Paired-Seed *t* CI** |
| Without Coordinate FiLM | 1.1824 | [1.1062, 1.2587] |
| Without auxiliary supervision | 1.8224 | [1.6717, 1.9732] |
| **Panel C. Sensitivity Analyses** | | |
| **Evaluation Subset** | **Windows** | **MAE Difference** |
| 2020 | 204 | 5.0389 |
| 2021 | 216 | 4.9558 |
| 2022 | 200 | 2.3579 |
| 2023 | 216 | 3.8392 |
| Non-overlapping target intervals | 161 | 3.8485 |

Figure 5 presents paired MAE differences for the field-information configurations within each monitoring site, stratified by historical field-observation availability.

Figure 5a shows a pooled MAE increase of 3.38 ± 0.22 for fixed gates relative to the complete model. Among the nine sites containing both partially observed and complete field histories, this increase was larger for complete histories at eight sites. At S007, the increases were 4.58 and 4.47 for partial and complete histories, respectively.

Adding the binary availability indicator to the no-field model reduced pooled MAE by 0.38 ± 0.04 in Figure 5b. This reduction was larger for complete than for partial histories at two of the nine sites. In Figure 5c, the complete model achieved lower mean MAE than

the availability-only control in all 18 populated site–stratum combinations with at least one valid field observation. The pooled MAE reduction relative to the availability-only control was 2.74 ± 1.01.

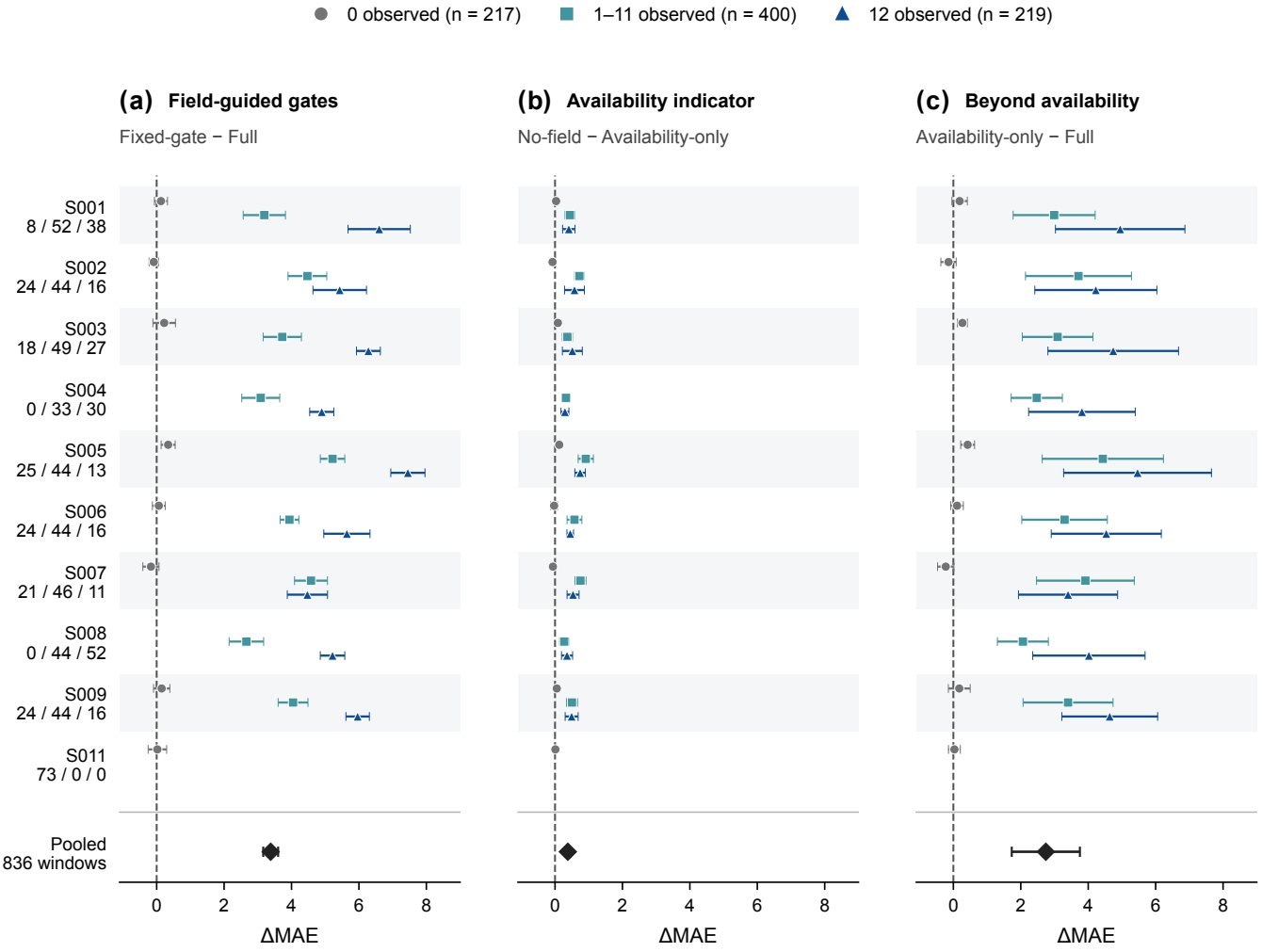



**Figure 5.** Within-site comparisons of light-trap-rate MAE across field-observation availability strata on the 2020–2023 test set. (**a**) Fixed-gate minus full DRAMM-Pest; (**b**) no-field minus availability-only; (**c**) availability-only minus full DRAMM-Pest. Circles, squares, and triangles denote windows with zero, 1–11, and 12 valid field observations in the 12-period history, respectively. Points and horizontal error bars show mean paired MAE differences and ±1 sample standard deviation across five matched seeds. Counts beneath each site give window numbers in the same stratum order. Black diamonds show pooled differences over all 836 windows and 5016 valid target positions, calculated separately for each seed before summarization. Positive values indicate lower MAE for the second model named in each comparison. Vertical dashed lines mark zero MAE difference.

### *3.4. Spatial Holdout Performance and Sensitivity to Missing Historical Observations*

Figure 6 compares forecasting performance across five site-held-out folds and sensitivity to random historical-input masking. Panel (a) presents the spatial benchmark, and panels (b,c) show MAE and RMSLE under additional masking on the chronological test set.

DRAMM-Pest achieved lower mean MAE than XGBoost, GRU, and the Coordinate FiLM ablation in all five site-held-out folds. Removing Coordinate FiLM increased mean MAE by 1.90–3.80. Seasonal Mean achieved the lowest MAE in F1, which held out S001 and S004, with 17.37 compared with 61.20 for DRAMM-Pest. It also achieved the lowest MAE in

F2, which held out S008 and S011, with 56.64 compared with 59.70. DRAMM-Pest achieved the lowest mean MAE among the evaluated models in F3–F5.

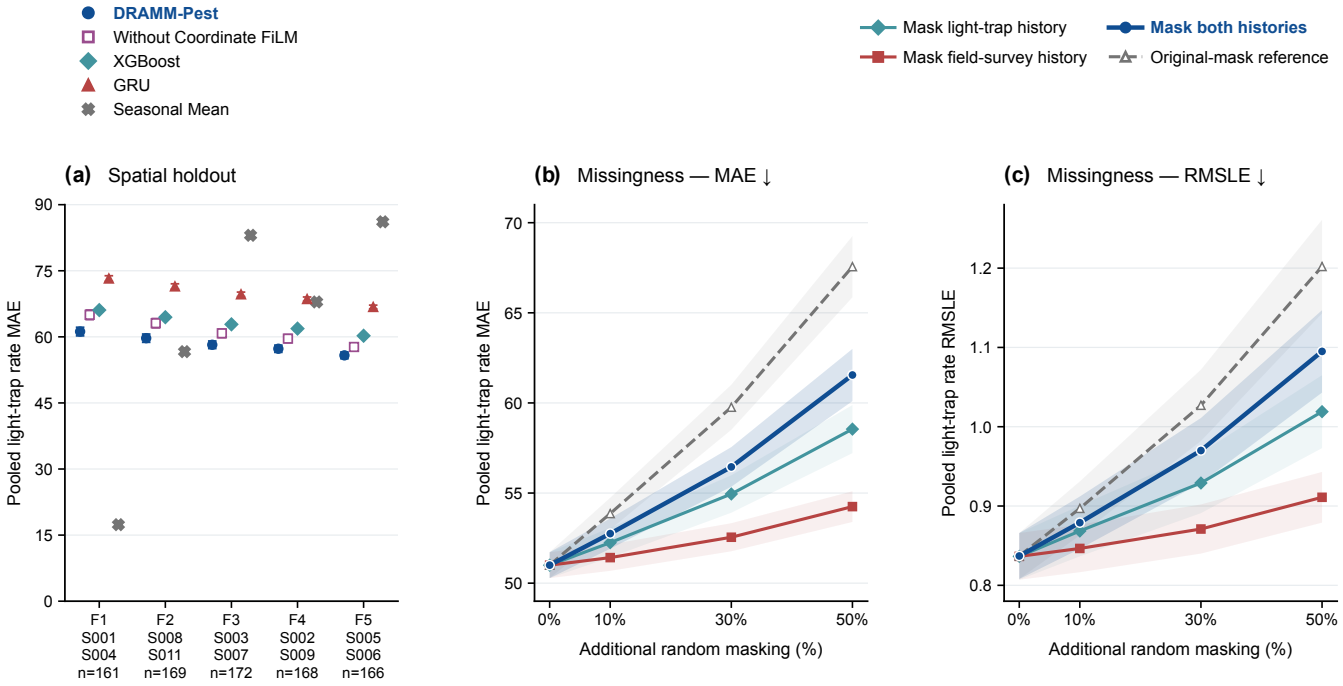



**Figure 6.** Spatial holdout performance and sensitivity to random historical-input masking. (**a**) Pooled light-trap-rate MAE for DRAMM-Pest, its Coordinate FiLM ablation, XGBoost, GRU, and Seasonal Mean across five site-held-out folds. Labels identify the held-out sites and forecast-window counts ($n$). Points and error bars show the mean and ±1 sample standard deviation across five seeds; Seasonal Mean is deterministic. (**b**,**c**) Pooled MAE and RMSLE under 0%, 10%, 30%, and 50% additional random masking of light-trap history, field-survey history, or both. Gray dashed curves represent the dual-history original-mask reference, which uses the same checkpoints and filled inputs while retaining masks from before additional masking. Metrics were averaged across ten shared random masking realizations within each seed; lines and shaded bands show the five-seed mean and ±1 sample standard deviation. Downward arrows indicate that lower metric values are preferable.

MAE and RMSLE increased with additional random masking. Among the three mask-aware conditions, errors were lowest when only field-survey history was masked, higher when only light-trap history was masked, and highest when both histories were masked. Under matched dual-history masking, updating the masks yielded lower mean MAE and RMSLE than the original-mask reference at all three nonzero masking levels. At 50% additional masking, the mask-aware condition achieved an MAE of 61.55 ± 1.45 and an RMSLE of 1.095 ± 0.052, compared with 67.56 ± 1.70 and 1.202 ± 0.059 for the original-mask reference.

## 4. Discussion

### *4.1. Principal Findings and Comparison with Existing Studies*

On the held-out 2020–2023 test set, DRAMM-Pest yielded the most favorable reported point estimates across the positive-catch, light-trap-rate, and auxiliary field-survey metrics. Relative to XGBoost, the absolute reductions in light-trap-rate RMSLE and Brier score were 0.25023 (95% CI, 0.22893–0.27095) and 0.00859 (95% CI, 0.00743–0.00987), respectively. DRAMM-Pest retained the most favorable reported point estimates across $h_1$–$h_6$, while light-trap-rate errors increased toward the later horizons.

Previous pest-forecasting studies have combined monitoring records with weather, crop-development, spatial, and heterogeneous observational information [10,12,13,19,30,31]. Choquet-integral ensembles have also combined convolutional-network predictions for tomato leaf disease identification [32]. DRAMM-Pest integrates complementary evidence from light-trap and field-survey records through their distinct forecasting roles. It links light-trap temporal modeling to field-survey information through historical conditioning and an auxiliary training objective, with temporal states subsequently aggregated for each forecast horizon.

The coupled outputs describe complementary aspects of light-trap monitoring: whether a monitoring period has a positive catch and the predicted catch rate. The learned positive factor links these outputs within the forecasting model, without representing a separate biological process. Their practical relevance lies in jointly characterizing positive-catch occurrence and activity magnitude for surveillance planning [1,3,33,34].

### *4.2. Temporal Component Modeling and Horizon-Aware Aggregation*

Temporal component modeling and horizon-aware aggregation formed a connected pathway from a shared history to six horizon representations. Their ablations produced the largest metric changes within the tested design, supporting the predictive value of this temporal modeling pathway in the complete model.

Temporal decomposition has been used in pest forecasting [15] and general time-series forecasting [28,35]. In DRAMM-Pest, horizon-dependent aggregation combines the learned temporal component representations for each forecast step. The horizon-aware weighting ablation in Table 6 quantifies the contribution of this aggregation to forecasting performance.

Positive-catch prevalence increased from 72.0% at $h_1$ to 87.7% at $h_6$, raising the prevalence-dependent reference level of AUPRC. Brier scores at $h_2$–$h_6$ remained above the $h_1$ value, while MAE and RMSLE generally increased toward later horizons. The high-abundance evaluation shows that DRAMM-Pest improved threshold-exceedance identification and within-window peak timing relative to XGBoost, alongside lower absolute error. Negative mean signed errors in both models identify persistent underestimation of high-abundance magnitude.

Related pest-forecasting studies identify broad target ranges and sparse high-abundance observations as challenges for multi-step prediction [36,37].

### *4.3. Dual-Measurement Guidance, Coordinate Conditioning, and Missing-Input Sensitivity*

The pooled controls in Table 6 and the within-site comparisons in Figure 5 support the predictive value of combining field histories with auxiliary supervision. The complete model achieved lower mean MAE than the availability-only configuration in both partially observed and complete-history strata across the nine sites containing field observations. The fixed-gate and auxiliary-supervision ablations further characterize the contributions of adaptive field modulation and field-target supervision.

The spatial experiment evaluated prediction at sites excluded from model fitting within the same regional monitoring network. Coordinate conditioning was associated with lower mean errors in the matched ablation comparisons. F1 held out S001 and S004, whose test target means were 21.97 and 0.63 individuals per valid trap night, respectively, as reported in Table 5. For this relatively low-abundance site pair, calendar-period averages estimated from the eight training sites yielded substantially lower pooled MAE than DRAMM-Pest. Spatial transfer performance varied across held-out site pairs, and DRAMM-Pest did not consistently outperform Seasonal Mean.

Random masking revealed greater forecast sensitivity to the loss of light-trap history than to the loss of field-survey history. Under matched dual-history masking, explicitly

updating observation masks improved both MAE and RMSLE using the same fitted model. Continuous trap outages and prolonged survey gaps provide further scenarios for evaluating temporally structured input loss.

### *4.4. Scope and Limitations*

The evaluation used retrospective monitoring records from 2000–2023 in the Lower Yangtze River Valley, with forecasting samples contributed by ten sites. The model incorporated monitoring histories, observation masks, calendar information, and coordinates, but did not include weather, crop phenology, or management practices. Its results therefore establish the forecasting value of complementary monitoring records within the studied regional network; the additional value of environmental and agronomic covariates remains a subject for further evaluation. Remote-sensing estimates of crop productivity provide spatial and temporal descriptors that could be evaluated as additional crop-context inputs [38]. Although DRAMM-Pest improved high-abundance forecasts relative to XGBoost, persistent underprediction indicates a continuing challenge in estimating elevated activity levels.

Temporal testing covered a single held-out period from 2020 to 2023, and spatial hold-out assessed transfer within the same regional network. The 12-period history requirement restricts forecast evaluation to targets from May onward, excluding March–April activity and immigration events occurring during those months. Year-specific and target-non-overlapping analyses assess sensitivity within this test period but do not replace evaluation with retraining at multiple forecast origins. Site-year bootstrap intervals preserve within-block dependence and are conditional on the fitted seeds; residual dependence between blocks and the limited number of test years constrain broader generalization. The results support the potential of six-period forecasts to help prioritize surveillance under the evaluated conditions. Independent-network and prospective-season evaluations, together with assessment of their effects on monitoring decisions and intervention outcomes, are needed to establish operational value.

## 5. Conclusions

This study developed DRAMM-Pest for direct six-horizon forecasting of rice leaf roller activity from aligned light-trap and field-survey records. Available field history conditions light-trap temporal states before horizon-specific aggregation, while observed field-survey targets provide auxiliary supervision during training. The model jointly predicts positive-catch probability and light-trap rate. On the chronologically held-out 2020–2023 test set, its Brier score and light-trap-rate RMSLE were lower than the corresponding XGBoost values by 0.00859 and 0.25023, respectively. Component ablations identified temporal component modeling and horizon-aware aggregation as the modules associated with the largest numerical changes. The complete model achieved lower mean forecasting errors than controls excluding field information or retaining only its binary availability indicator, supporting the predictive value of the dual-measurement framework under the evaluated regional conditions. Further evaluation incorporating weather, crop phenology, and management information, alongside validation in independent networks and prospective monitoring seasons, will help establish the framework's broader applicability and operational value.

**Author Contributions:** Conceptualization, Q.X., J.X., Y.W. and Y.L.; methodology, Q.X.; software, Q.X.; validation, Q.X., J.X., Y.W. and Y.L.; formal analysis, Q.X.; investigation, Q.X.; resources, J.X., Y.W., Y.L. and Y.Z.; data curation, Q.X.; writing—original draft preparation, Q.X.; writing—review and editing, J.X., Y.W., Y.L. and Y.Z.; visualization, Q.X.; supervision, J.X. and Y.Z.; project administration, J.X. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The original monitoring data are publicly available in Dryad [23]. The processed datasets used in this study are available from the corresponding authors on request.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Yang, Y.; Lu, Y.; Tian, J.; Zheng, X.; Guo, J.; Liu, X.; Lü, Z.; Xu, H. Sustainable Management Strategies for Rice Leaffolder, *Cnaphalocrocis medinalis* (Guenée): Progress and Prospects. *Rice Sci*. **2025**, *32*, 322–338. [CrossRef]
2. Zhang, S.Y.; Zhang, Y.Y.; Yang, F.; Zhou, C.; Shen, H.M.; Wang, B.B.; Zeng, J.; Reynolds, D.R.; Chapman, J.W.; Hu, G. Climate Change Is Leading to an Ecological Trap in a Migratory Insect. *Proc. Natl. Acad. Sci. USA* **2025**, *122*, e2422595122. [CrossRef] [PubMed]
3. Gao, L.; Yan, R.; He, W.; Wu, K. Use of Food Attractant to Monitor and Forecast Population Dynamics of *Cnaphalocrocis medinalis* (Lepidoptera: Pyralidae), a Long-Distance Migratory Pest. *Agronomy* **2023**, *13*, 2141. [CrossRef]
4. Liu, W.; Lu, M.; Wang, C.; Bao, Y. Prediction Model of Regional Meteorological Level of Rice Leaf Roller in China. *Chin. Agric. Sci. Bull*. **2016**, *32*, 172–178. [CrossRef]
5. Zhang, J.H.; Zhang, X.Y.; Zhang, Y.Y.; Tian, J.; Ouyang, X.Y.; Yin, L.; Wu, Y.; Zeng, J.; Zhang, S.Y.; Hu, G. Migration Patterns and Meteorological Drivers of the Rice Leaf Roller in Western Hunan Province, China. *Insects* **2026**, *17*, 466. [CrossRef] [PubMed]
6. Qi, H.H.; Zhang, Y.H.; Wang, J.; Peng, H.; Zhang, Z.; Cheng, D.F. Rhythm of Rice Leaf Folder (*Cnaphalocrocis medinalis*) to the Searchlight Trap. *Sci. Agric. Sin*. **2014**, *47*, 4436–4444. [CrossRef]
7. Wang, F.Y.; Yang, F.; Lu, M.H.; Luo, S.Y.; Zhai, B.P.; Lim, K.S.; McInerney, C.E.; Hu, G. Determining the Migration Duration of Rice Leaf Folder (*Cnaphalocrocis medinalis* (Guenée)) Moths Using a Trajectory Analytical Approach. *Sci. Rep*. **2017**, *7*, 39853. [CrossRef] [PubMed]
8. Mitra, R.; McGough, S.F.; Chakraborti, T.; Holmes, C.; Copping, R.; Hagenbuch, N.; Biedermann, S.; Noonan, J.; Lehmann, B.; Shenvi, A.; et al. Learning from Data with Structured Missingness. *Nat. Mach. Intell*. **2023**, *5*, 13–23. [CrossRef]
9. Kebede, E.A.; Abou Ali, H.; Clavelle, T.; Froehlich, H.E.; Gephart, J.A.; Hartman, S.; Herrero, M.; Kerner, H.; Mehta, P.; Nakalembe, C.; et al. Assessing and Addressing the Global State of Food Production Data Scarcity. *Nat. Rev. Earth Environ*. **2024**, *5*, 295–311. [CrossRef]
10. Zhang, H.; He, B.; Xing, J.; Lu, M. Deep Spatial and Temporal Graph Convolutional Network for Rice Planthopper Population Dynamic Forecasting. *Comput. Electron. Agric*. **2023**, *210*, 107868. [CrossRef]
11. Thanammal Indu, V.; Suja Priyadharsini, S. Deep Residual Squeeze Excited Region Convolutional Network: A Robust Deep Learning Framework for Precision Agriculture and Plant Disease Diagnosis. *Cybern. Syst*. **2026**, *57*, 1356–1385. [CrossRef]
12. Ibrahim, E.A.; Salifu, D.; Mwalili, S.; Dubois, T.; Collins, R.; Tonnang, H.E.Z. An Expert System for Insect Pest Population Dynamics Prediction. *Comput. Electron. Agric*. **2022**, *198*, 107124. [CrossRef]
13. Wang, J.; Zhang, D. Intelligent Pest Forecasting with Meteorological Data: An Explainable Deep Learning Approach. *Expert Syst. Appl*. **2024**, *252*, 124137. [CrossRef]
14. Palma, G.R.; Mello, R.F.; Godoy, W.A.C.; Engel, E.; Lau, D.; Markham, C.; Moral, R.A. Forecasting Insect Abundance Using Time Series Embedding and Machine Learning. *Ecol. Inform*. **2025**, *85*, 102934. [CrossRef]
15. He, C.; Peng, Z.; Peng, L.; Liu, Y.; Zhang, C.; Zhu, L.; Tan, S.; Zou, L. Chilo suppressalis Population Dynamics Forecasting by Exponential Smoothing Decomposition and Multi-Stream Network. *Agriculture* **2025**, *15*, 2474. [CrossRef]
16. Zhang, D.; Wang, J. Interpretable Multidimensional Meteorological Memory Modeling for Diamondback Moth Forecasting. *Agronomy* **2026**, *16*, 1114. [CrossRef]
17. Deforce, B.; Baesens, B.; Diels, J.; Serral Asensio, E. Harnessing the Power of Transformers and Data Fusion in Smart Irrigation. *Appl. Soft Comput*. **2024**, *152*, 111246. [CrossRef]
18. Li, J.; Wang, L.; Liu, J.; Tang, J. ViST: A Ubiquitous Model with Multimodal Fusion for Crop Growth Prediction. *ACM Trans. Sens. Netw*. **2023**, *20*, 1–23. [CrossRef]
19. Chacón-Maldonado, A.M.; Asencio-Cortés, G.; Troncoso, A. A Multimodal Hybrid Deep Learning Approach for Pest Forecasting Using Time Series and Satellite Images. *Inf. Fusion* **2025**, *124*, 103350. [CrossRef]
20. Saad, F.; Burnim, J.; Carroll, C.; Patton, B.; Köster, U.; Saurous, R.A.; Hoffman, M. Scalable Spatiotemporal Prediction with Bayesian Neural Fields. *Nat. Commun*. **2024**, *15*, 7942. [CrossRef] [PubMed]
21. Kopton, J.; de Bruin, S.; Schulz, D.; Luedeling, E. Combining Spatio-Temporal Pest Risk Prediction and Decision Theory to Improve Pest Management in Smallholder Agriculture. *Comput. Electron. Agric*. **2025**, *236*, 110426. [CrossRef]
22. Liu, M.; Dong, Z.; Shi, H. Multi-Sensor Information Fusion and Multi-Model Fusion-Based Remaining Useful Life Prediction of Fan Slewing Bearings with the Nonlinear Wiener Process. *Sustainability* **2023**, *15*, 12010. [CrossRef]
23. Zhang, S.Y.; Zhang, Y.Y.; Hu, G. Climate Change Is Leading to an Ecological Trap in a Migratory Insect [Dataset]. *Dryad* **2025**. [CrossRef]

24. Perez, E.; Strub, F.; de Vries, H.; Dumoulin, V.; Courville, A. FiLM: Visual Reasoning with a General Conditioning Layer. *AAAI Conf. Artif. Intell.* **2018**, *32*, 3942–3951. [CrossRef]
25. Luo, M.; Yin, M.; Li, J.; Li, Y.; Kobsiriphat, W.; Yu, H.; Xu, T.; Wu, X.; Cao, W. Lateral Walking Gait Recognition and Hip Angle Prediction Using a Dual-Task Learning Framework. *Cyborg Bionic Syst.* **2025**, *6*, 0250. [CrossRef] [PubMed]
26. Cho, K.; van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning Phrase Representations Using RNN Encoder–Decoder for Statistical Machine Translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*; Association for Computational Linguistics: Stroudsburg, PA, USA, 2014; pp. 1724–1734. [CrossRef]
27. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; Association for Computing Machinery: New York, NY, USA, 2016; pp. 785–794. [CrossRef]
28. Zeng, A.; Chen, M.; Zhang, L.; Xu, Q. Are Transformers Effective for Time Series Forecasting? *Proc. AAAI Conf. Artif. Intell.* **2023**, *37*, 11121–11128. [CrossRef]
29. Saito, T.; Rehmsmeier, M. The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. *PLoS ONE* **2015**, *10*, e0118432. [CrossRef] [PubMed]
30. Mittal, M.; Gupta, V.; Aamash, M.; Upadhyay, T. Machine Learning for Pest Detection and Infestation Prediction: A Comprehensive Review. *WIREs Data Min. Knowl. Discov.* **2024**, *14*, e1551. [CrossRef]
31. Palma, G.R.; Godoy, W.A.C.; Engel, E.; Lau, D.; Galvan, E.; Mason, O.; Markham, C.; Moral, R.A. Pattern-Based Prediction of Population Outbreaks. *Ecol. Inform.* **2023**, *77*, 102220. [CrossRef]
32. Hang, Y.; Meng, X.; Wang, R.; Tian, X.; Cui, J. Ensemble Technique of Deep Learning Model for Identifying Tomato Leaf Diseases Based on Choquet Fuzzy Integral. *Cybern. Syst.* **2026**, *57*, 1036–1059. [CrossRef]
33. Rincon, D.F.; Esch, E.D.; Gutierrez-Illan, J.; Tesche, M.; Crowder, D.W. Predicting Insect Population Dynamics by Linking Phenology Models and Monitoring Data. *Ecol. Model.* **2024**, *493*, 110763. [CrossRef]
34. Dai, M.; Shen, Y.; Li, X.; Liu, J.; Zhang, S.; Miao, H. Digital Twin System of Pest Management Driven by Data and Model Fusion. *Agriculture* **2024**, *14*, 1099. [CrossRef]
35. Zhou, T.; Ma, Z.; Wen, Q.; Wang, X.; Sun, L.; Jin, R. FEDformer: Frequency Enhanced Decomposed Transformer for Long-Term Series Forecasting. In *Proceedings of the 39th International Conference on Machine Learning*; Proceedings of Machine Learning Research; PMLR: Cambridge, MA, USA, 2022; Volume 162, pp. 27268–27286.
36. Chen, X.; Hassan, M.M.; Yu, J.; Zhu, A.; Han, Z.; He, P.; Chen, Q.; Li, H.; Ouyang, Q. Time Series Prediction of Insect Pests in Tea Gardens. *J. Sci. Food Agric.* **2024**, *104*, 5614–5624. [CrossRef] [PubMed]
37. Chacón-Maldonado, A.M.; Melgar-García, L.; Asencio-Cortés, G.; Troncoso, A. A Novel Method Based on Hybrid Deep Learning with Explainability for Olive Fruit Pest Forecasting. *Neural Comput. Appl.* **2025**, *37*, 3245–3264. [CrossRef]
38. Zhang, Z.; Huo, L.; Su, Y.; Shen, H.; Yang, G. Estimation of Corn Net Primary Productivity and Carbon Sequestration Based on the CASA Model: A Case Study of the Fen River Basin. *Sustainability* **2024**, *16*, 2942. [CrossRef]


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.