Zero-Shot Transfer Capabilities of the Sundial Foundation Model for Leaf Area Index Forecasting

Peining Zhang[1], **Hongchen Qin**[2], **Haochen Zhang**[1], **Ziqi Guo**[2], **Guiling Wang**[2], **Jinbo Bi**[1] *
[1]School of Computing
[2]School of Civil and Environmental Engineering
University of Connecticut, Storrs, CT 06269, USA
peining.zhang, hongchen.qin, haochen.zhang, ziqi.guo, guiling.wang, jinbo.bi@uconn.edu

Abstract

This work investigates the zero-shot forecasting capability of time series foundation models for Leaf Area Index (LAI) forecasting in agricultural monitoring. Using the HiQ dataset (U.S., 2000-2022), we systematically compare statistical baselines, a fully supervised LSTM, and the Sundial foundation model under multiple evaluation protocols. We find that Sundial, in the zero-shot setting, can outperform a fully trained LSTM provided that the input context window is sufficiently long—specifically, when covering more than one or two full seasonal cycles. We show that a general-purpose foundation model can surpass specialized supervised models on remote-sensing time series prediction without any task-specific tuning. These results highlight the strong potential of pretrained time series foundation models to serve as effective plug-and-play forecasters in agricultural and environmental applications.

Code — https://github.com/peiningzhang/sundial-lai
Datasets —
https://essd.copernicus.org/articles/16/1601/2024/

Introduction

Leaf Area Index (LAI) is a critical biophysical variable widely used in agricultural and ecological monitoring, quantifying canopy structure and influencing crop growth. Beyond its central role in agronomic decision-making and yield estimation, accurate LAI forecasting is vital for understanding land-surface feedbacks, as vegetation significantly influences regional climate through biogeophysical processes (e.g., albedo and roughness) and biogeochemical cycles (Alessandri et al. 2017). However, practical forecasting is challenged by diverse cropping systems, varying environmental conditions, and limited labeled data. This motivates models that can generalize across locations and years. Furthermore, supervised models trained on heterogeneous datasets tend to learn "global average" patterns and may miss site-specific phenological anomalies without extensive local fine-tuning.

Recent advances in time series foundation models, such as Sundial (Liu et al. 2025), offer a promising avenue for improving LAI forecasting. Unlike task-specific models, these large, pretrained networks leverage in-context learning, treating historical time series as "prompts" to deduce underlying dynamics. This paradigm has the potential to adapt to new tasks and domains more efficiently than traditional supervised approaches.

An important research question we address is: **Can a general-purpose, zero-shot time series foundation model match or even surpass a supervised model (e.g., LSTM) specifically trained for LAI prediction?**

This paper makes the following three contributions:

- We conduct a rigorous zero-shot benchmark of the Sundial foundation model for LAI remote sensing prediction using the comprehensive HiQ dataset.
- We demonstrate that, given a sufficiently long historical input context, Sundial (zero-shot) outperforms a fully supervised LSTM. This challenges the conventional assumption that task-specific fine-tuning is necessary for optimal performance and highlights the feasibility of "plug-and-play" forecasting.
- We analyze the context window threshold where zero-shot performance overtakes the supervised baseline, and discuss the distinct mechanisms at play: instance-level adaptation in foundation models versus dataset-level generalization in LSTMs.

Our findings offer new empirical insights into the potential and limitations of time series foundation models for remote sensing vegetation analysis, and motivate future research into scalable agricultural forecasting.

Related Work

LAI Time Series Forecasting

Leaf Area Index (LAI) is a key indicator for monitoring vegetation health, phenology, and long-term ecosystem dynamics. Most LAI forecasting approaches treat the signal as a univariate time series characterized by strong seasonality, low-frequency variation, and limited annotated samples (Jiang et al. 2010). Classical methods include harmonic models, autoregressive formulations, and

*Accepted at the First International Workshop on AI in Agriculture (Agri AI), co-located with AAAI 2026.
Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

exponential-smoothing variants, which capture seasonal cycles but struggle with nonlinear responses to climate and land-cover variability (Winters 1960). Recent studies explore data-driven forecasting of vegetation indices, yet robust LAI prediction remains challenging due to heterogeneous temporal patterns and spatial variability across regions (Liu et al. 2022b).

Baselines: Statistical and Supervised Models

Classical LAI forecasting has long relied on statistical time series models such as autoregressive (AR) and exponential-smoothing formulations, which are simple, interpretable, and effective for short-term dynamics. However, these models are limited in their ability to capture nonlinear seasonal patterns and interannual variability.

Supervised machine-learning baselines—including random forests (Breiman 2001) and gradient boosting (Friedman 2001; Ke et al. 2017)—can incorporate richer auxiliary features when available, but offer only marginal improvements when restricted to univariate LAI histories. Recurrent neural networks, particularly LSTMs, remain strong supervised baselines for vegetation-index forecasting and are widely used in agronomic and ecological applications (Hochreiter and Schmidhuber 1997; Liu et al. 2022a). Despite their capacity to model nonlinear temporal dependencies, these methods require substantial labeled data and typically exhibit limited gains when training data is scarce or spatially heterogeneous.

Importantly, none of these approaches leverage cross-domain temporal knowledge. As a result, their generalization under long sequences, sparse supervision, or unseen spatial regions remains constrained—highlighting the value of pretrained time series foundation models, which can transfer temporal priors learned from large and diverse corpora.

Time Series Foundation Models

Large pretrained time series models have recently demonstrated strong zero-shot and few-shot transfer across diverse domains. Models such as TimesFM and Sundial learn universal temporal representations from mixed-modality corpora spanning climate, sensor networks, and economic indicators (Das et al. 2024; Liu et al. 2025; Ansari et al. 2024). These foundation models can leverage broad temporal priors—seasonality, trends, and multi-scale structure—reducing dependence on in-domain training data and enabling robust downstream forecasting.

Data

HiQ Dataset

We conduct our experiments using the HiQ dataset (Yan et al. 2023), a large-scale benchmark for Leaf Area Index (LAI) prediction. HiQ reprocesses the satellite MODIS LAI product (Myneni, Knyazikhin, and Park 2015) using a spatiotemporal information compositing algorithm, resulting in improved spatiotemporal consistency and reduced noise. The dataset provides high-quality global LAI time series spanning 2000–2022, with broad geographic coverage across diverse crop types and ecological regions.

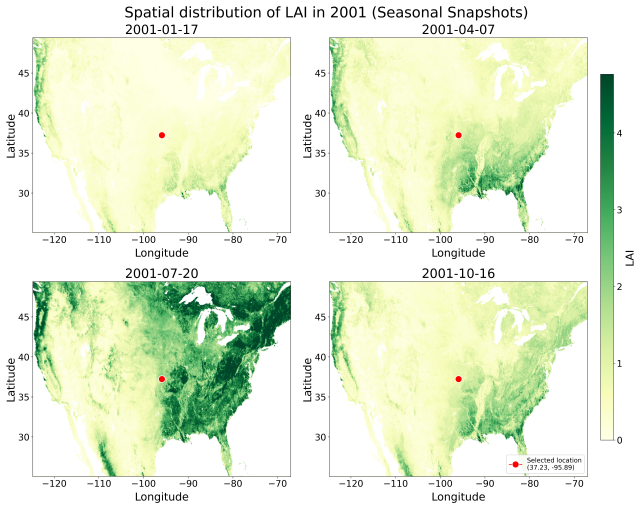

Figure 1: Spatial distribution of LAI across the study domain, shown as four seasonal snapshots from 2001. Higher LAI values (darker green) correspond to denser vegetation.

In this work, we restrict our analysis to the continental United States (CONUS) subset. Although HiQ offers fully available global coverage, this choice is primarily driven by computational constraints rather than data availability or methodological limitations. We argue that the CONUS domain already encompasses a wide range of vegetation regimes and climatic conditions, making it a representative testbed for evaluating zero-shot LAI forecasting performance.

We use the HiQ product at 5-km spatial resolution, where each pixel contains an LAI time series sampled at an 8-day cadence, resulting in approximately 1,050 time steps per location.

Seasonal snapshots

The HiQ dataset exhibits substantial spatial and temporal heterogeneity across the continental United States, as illustrated in Figure 1. The four seasonal snapshots from 2001 reveal pronounced geographic gradients in vegetation greenness, with distinct regional patterns that evolve throughout the year. This spatial variability, combined with the strong seasonal cycles evident in the temporal evolution, presents a challenging forecasting problem that requires models capable of capturing both long-term seasonal patterns and location-specific dynamics. The pronounced seasonal variations highlighted in these snapshots underscore the importance of sufficient context window length—covering multiple full seasonal cycles—for accurate LAI forecasting, which directly relates to our finding that foundation models like Sundial require extended input windows to achieve optimal zero-shot performance.

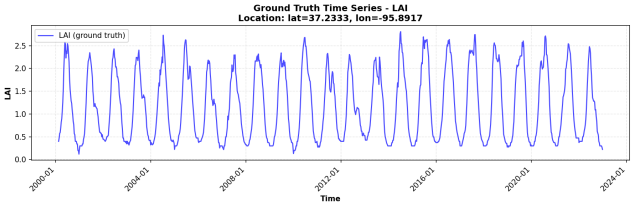

Figure 2: Temporal LAI evolution at the representative center pixel (red dot in Fig. 1). The series shows clear seasonal cycles and interannual variability.

Temporal Dynamics

To better understand the temporal behavior of the vegetation data, we examine the LAI time series of a representative pixel, referred to as the "center location." This specific location is geographically highlighted by a red dot in Figure 1. Figure 2 presents the complete temporal profile for this location, where the horizontal axis represents the date and the vertical axis indicates the LAI value. The series exhibits pronounced seasonal periodicity, reflecting the natural phenological cycle of vegetation growth and senescence typical for this region.

Preprocessing and Quality Control

To ensure consistency, all LAI time series undergo per-pixel normalization. For supervised models (e.g., LSTM), we normalize each pixel's time series using its mean and standard deviation computed from valid (non-missing) observations in the training period. Missing or masked values, common in satellite products due to clouds or sensor issues, are handled in two stages. For most models, missing values are either explicitly masked out during loss computation, or—where required—are imputed with linear interpolation over time. Cells with a very high fraction of missing data (i.e. > 0.1) are excluded from the training and evaluation splits.

Methods

Baseline Models

We consider a suite of established baselines for LAI forecasting:

- **Last Value**: Predicts the most recent observed LAI from the context window as the forecast for all future steps.
- **Mean(k)**: Uses the mean of the last k observed LAI values as the prediction for all horizons.
- **Trend**: Extrapolates a fitted linear trend from the previous k time points to predict future LAI values (Jensen et al. 2008).
- **ARIMA**: For each test window, we fit an $\text{ARIMA}(2, 1, 0)$ model (Box and Pierce 1970; Seabold and Perktold 2010) to the observed LAI values to forecast the next value. ARIMA is chosen as a simple yet reasonable baseline for univariate time series; model fitting uses state-space estimation.
- **LSTM**: A supervised Long Short-Term Memory model (Hochreiter and Schmidhuber 1997) trained on the entire training set to capture nonlinear dependencies in LAI time series. Our LSTM architecture consists of two stacked LSTM layers with 64 hidden units each, followed by a fully connected output layer. The model is trained using the Adam (Kingma 2014) optimizer with a learning rate of 0.001, batch size of 256, and early stopping based on validation loss.

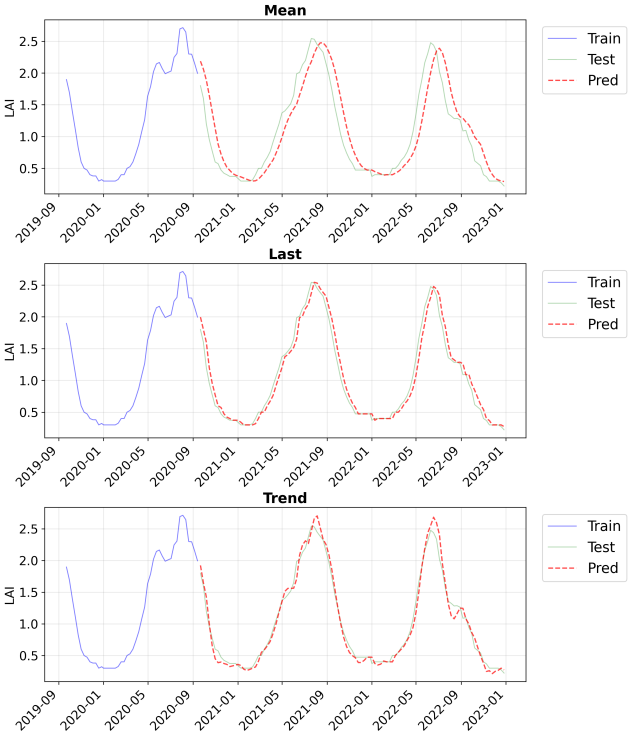

Figure 3: **Performance of simple statistical baselines on the final 150 time steps of a representative LAI pixel.** The first 46 time steps show ground-truth LAI, followed by 104 forecasted steps. The *Mean* baseline underestimates seasonal amplitude, the *Last* baseline lags the seasonal cycle, and the *Trend* baseline tracks short-term tendencies but fails at seasonal turning points.

Forecasting Protocol

We adopt a unified fixed-window forecasting protocol across all models. For each pixel, we generate samples using a sliding window with input length T_{in} and forecast horizon H. To evaluate scalability, we test multiple window sizes $T_{\text{in}} \in \{32, 64, 128, 256, 512, 1024\}$ and forecast horizons $H \in \{1, 4, 8, 12\}$ (corresponding to 8, 32, 64, and 96 days ahead, respectively, given the 8-day temporal resolution). Evaluation follows a strict forward-in-time split: models are trained (or conditioned) on historical segments and tested on the most recent unseen data. For the supervised baselines (e.g., LSTM), parameters are optimized by minimizing the Mean Squared Error (MSE) between predictions and ground truth. In contrast, the foundation model operates in strict inference-only mode without task-specific loss opti-

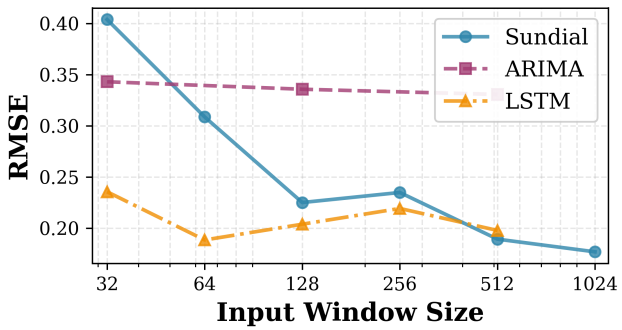

Figure 4: **Effect of input window size (T_{in}) on forecasting accuracy at horizon $H = 4$.** RMSE comparison across different window sizes. Larger input windows consistently improve Sundial performance, while ARIMA and LSTM saturate at moderate window lengths. Metrics are averaged per pixel.

mization.

Zero-Shot Inference with Sundial

We formulate the LAI forecasting task as a conditional generation problem. Instead of fine-tuning model weights, we leverage the pre-trained capabilities of Sundial directly (Liu et al. 2025). The historical input window $X \in \mathbb{R}^{T_{\text{in}}}$ serves as a continuous *context prompt*, analogous to a text prompt in Large Language Models (LLMs). The model processes this prompt to deduce the underlying phenological dynamics (e.g., seasonality and trend) and generates the future sequence. This approach relies entirely on the model's in-context learning ability acquired during pre-training, enabling a strictly zero-shot evaluation. We use the publicly available Sundial model checkpoint without any modifications or fine-tuning. For inference, we generate forecasts autoregressively for multi-step horizons.

Experiments

Evaluation Metrics

We evaluate model performance using standard error- and scale-aware regression metrics:

- **MAE**: Mean Absolute Error
- **RMSE**: Root Mean Square Error
- **R^2**: Coefficient of Determination
- **MAPE**: Mean Absolute Percentage Error
- **CVRMSE**: Coefficient of Variation of RMSE

Metrics are aggregated over 100 randomly sampled locations to ensure computational efficiency. With approximately 100 rolling test windows per pixel, each reported metric represents an average across $\sim 10{,}000 \times H$ prediction-ground truth pairs, where H is the forecast horizon.

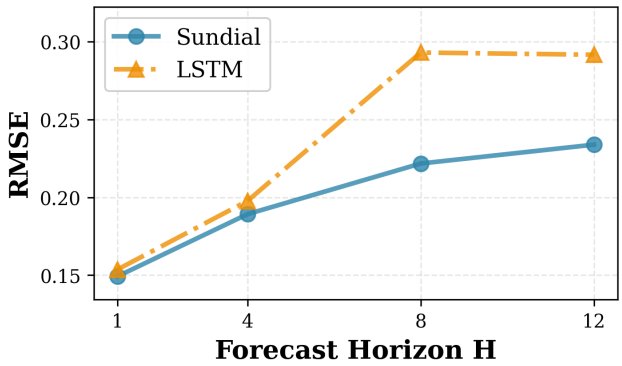

Figure 5: **Effect of forecasting horizon (H) on forecasting accuracy.** Multi-horizon forecasting comparison between Sundial and the supervised LSTM at a fixed context window (T_{in} = 512). The x-axis denotes the forecasting horizon H $\in \{1, 4, 8, 12\}$, and the y-axis shows the corresponding average RMSE. Although both models degrade as H increases, the zero-shot Sundial model exhibits a substantially slower performance deterioration, outperforming the supervised LSTM consistently for all $H \geq 4$. This demonstrates that Sundial retains long-range temporal structure more effectively, making it particularly suitable for extended-horizon agricultural forecasting.

Rule-based Baselines Results

To contextualize the difficulty of LAI forecasting, we first evaluate three simple rule-based baselines: Mean, Last Value, and Trend. As shown in Fig. 3, these heuristics struggle to follow the seasonal dynamics of LAI, with Mean lagging transitions, Last Value producing phase-shifted predictions, and Trend failing near turning points. These behaviors illustrate the limitations of simple historical extrapolation and motivate the need for more expressive forecasting models.

Model Results

One-Step-Ahead Forecasting Performance Table 1 reports the one-step-ahead forecasting results ($H = 1$). Classical statistical baselines and ARIMA establish a lower-bound benchmark, but their errors remain noticeably higher than those of learning-based methods.

Among supervised models, the LSTM achieves its best performance at relatively short windows, with the lowest RMSE (0.140) and highest R^2 (97.34%) obtained at $T_{\text{in}} = 64$. Performance degrades when the window becomes either too small or too large, reflecting the model's limited ability to exploit extended temporal context.

In contrast, as illustrated in Figure 4, **Sundial** benefits steadily from longer input windows. Its accuracy improves monotonically with window size and reaches its best performance at $T_{\text{in}} = 1024$, achieving the lowest MAE (0.096). Despite requiring no task-specific training, Sundial matches or surpasses the supervised baseline when provided with sufficient historical context, highlighting its strong plug-and-play capability for pixel-level LAI forecasting.

Model	Window	MAE (↓)	RMSE (↓)	R^2 (%) (↑)	MAPE (%) (↓)	CVRMSE (%) (↓)
Statistical Baselines						
Mean	4	0.274	0.375	72.6	36.37	32.90
Last	4	0.144	0.210	87.6	17.88	19.30
Trend	4	0.143	0.210	85.7	20.84	19.36
ARIMA						
ARIMA(2,1,0)	32	0.138	0.204	87.9	17.46	18.80
ARIMA(2,1,0)	128	0.134	0.197	88.6	17.31	18.18
ARIMA(2,1,0)	512	0.134	0.196	88.7	17.15	18.08
LSTM (Supervised Baseline)						
LSTM	32	0.107	0.155	96.82	20.20	17.73
LSTM	64	0.097	0.140	**97.34**	**18.83**	15.78
LSTM	128	0.097	0.141	97.33	19.53	15.97
LSTM	256	0.105	0.148	97.23	22.86	18.56
LSTM	512	0.099	0.143	97.26	20.44	16.45
Sundial (Plug-and-Play)						
Sundial	32	0.154	0.209	94.12	33.66	22.70
Sundial	64	0.132	0.185	95.18	26.32	19.97
Sundial	128	0.107	0.158	96.64	20.76	17.49
Sundial	256	0.112	0.162	96.41	21.80	17.71
Sundial	512	0.099	0.145	97.15	20.29	15.94
Sundial	1024	**0.096**	**0.140**	97.32	19.15	**15.44**

Table 1: **Performance benchmark for one-step-ahead LAI forecasting ($H = 1$).** All entries reflect the updated LSTM and Sundial results across input window sizes (T_{in}). Best values within each model category are **bolded**.

Missing ratio	RMSE increase over baseline (%)				
	1024	512	256	128	64
5%	8.2%	5.1%	1.5%	7.0%	0.8%
10%	10.2%	10.9%	2.1%	8.9%	4.2%
20%	25.6%	21.5%	12.8%	22.1%	7.5%
40%	57.1%	53.7%	42.3%	54.9%	36.7%

Table 2: **Sundial robustness under random missing observations at forecast horizon $H = 1$.** Entries are relative RMSE increases (%) compared to the no-missing-data baseline.

Robustness to Missing Data To assess the robustness of **Sundial** to missing observations, we randomly mask a fraction of values in the input context at rates of 5%, 10%, and 20% without interpolation. Notably, Sundial is equipped with a native pipeline for handling NaN inputs. We report performance at forecast horizon $H = 1$ as the relative RMSE increase (%) compared to the no-missing-data baseline (Table 2).

Multi-step Forecasting Across Longer Horizons We evaluate multi-horizon forecasting for $H \in \{1, 4, 8, 12\}$ at a representative context $T_{\text{in}} = 512$ (see Figure 5). Both models' errors increase with H, but Sundial's degradation is substantially slower: for $H \geq 4$ Sundial consistently outperforms the supervised LSTM despite remaining zero-shot. Although the per-pixel RMSE varies significantly across the dataset (high standard deviation due to spatial heterogeneity), a paired t-test confirms that Sundial's performance improvement over LSTM at this horizon ($H = 4$) is highly statistically significant ($p < 0.001$, $N \approx 10,000$). This pattern reinforces that, given sufficient context, the foundation model better preserves pixel-specific temporal structure for

Model	H\T_{in}	1024	512	256	128	64
LSTM (Supervised Baseline) — RMSE (↓)						
	1	–	0.143	0.157	0.150	0.149
	4	–	0.198	0.219	0.204	0.188
	8	–	0.251	0.293	0.238	0.247
	12	–	0.269	0.292	0.265	0.261
Sundial (Zero-Shot) — RMSE (↓)						
	1	0.140	0.145	0.162	0.158	0.185
	4	0.177	0.189	0.235	0.225	0.309
	8	0.202	0.222	0.290	0.283	0.433
	12	0.210	0.234	0.316	0.303	0.532
LSTM - Sundial (Δ RMSE)						
	1	–	-0.002	-0.019	-0.022	-0.052
	4	–	0.009	-0.016	-0.021	-0.120
	8	–	0.030	0.003	-0.045	-0.186
	12	–	0.035	-0.025	-0.038	-0.271

Table 3: **Unified comparison of multi-horizon (H) and multi-window (T_{in}) forecasting performance.** Cells contain RMSE (lower is better). Sundial exhibits monotonic improvement with longer context windows and consistently outperforms LSTM (Supervised Baseline) for $T_{\text{in}} \geq 512$, especially at longer horizons ($H \geq 4$). In the third section, positive values indicate Sundial achieves lower RMSE (better performance).

extended horizons.

Scalability in Multi-Step Forecasting The unified results in Table 3 reveal a clear divergence in how the two models respond to longer input windows and forecasting horizons. The supervised LSTM reaches its best accuracy at short windows (e.g., $T_{\text{in}} = 64$) but consistently degrades as the window grows, reflecting its limited ability to retain long-range temporal dependencies.

In contrast, Sundial benefits steadily from additional historical context. Its RMSE improves monotonically with increasing window size and eventually surpasses the LSTM baseline once $T_{\text{in}} \geq 512$, a trend that becomes more pronounced at longer horizons ($H \geq 4$). This behavior highlights Sundial's capacity to leverage extended seasonal information and maintain stability under multi-step forecasting, a key requirement for agricultural monitoring.

Discussion

Our experiments reveal a compelling crossover: while the supervised LSTM dominates in short-context regimes, the zero-shot Sundial foundation model significantly outperforms it given sufficient history. This finding highlights the trade-off between two distinct learning paradigms.

Global Priors vs. Local Adaptation. The performance disparity observed across window sizes stems from how each model utilizes information. The supervised LSTM relies on **dataset-level generalization**, excelling at $T_{\text{in}} \leq 64$ by leveraging learned "global average" priors to extrapolate incomplete seasonal cycles. In contrast, Sundial relies on **instance-level in-context learning**. While it struggles with short prompts, in long-context regimes ($T_{\text{in}} \geq 512$), the extended window enables it to capture **unique pixel-specific signatures**—such as specific drought responses or phenological shifts—that the generalized LSTM tends to smooth over. Effectively, Sundial builds a temporary, local model for each pixel in real-time, surpassing the static global template of the LSTM.

Towards "Plug-and-Play" Remote Sensing. These results suggest a potential paradigm shift. We demonstrate that a pre-trained foundation model can be deployed "off-the-shelf" to achieve state-of-the-art forecasting accuracy. This **"plug-and-play"** capability offers a scalable alternative to the resource-intensive "collect, annotate, and train" workflow, particularly valuable for agricultural monitoring in regions where training data is scarce or rapid deployment is required.

Limitations and Future Directions. While promising, several limitations warrant discussion. First, the requirement for long context windows ($T_{\text{in}} \geq 512$, approx. 11 years) may limit applicability in regions with sparse historical records. Second, our evaluation focuses on univariate forecasting; incorporating multi-modal inputs (e.g., climate variables, soil properties) could further enhance accuracy. Third, the inference computational cost of foundation models remains higher than lightweight baselines. Future work will focus on exploring few-shot fine-tuning to reduce context requirements and expanding geographic validation to assess global transferability.

Conclusion and Future Work

In this work, we conducted a systematic evaluation of the zero-shot capabilities of the **Sundial** time series foundation model for remote-sensing LAI forecasting. We found that, when provided with a sufficiently long historical context window, Sundial's zero-shot performance **can surpass** that of a supervised **LSTM** model specifically trained for the task, challenging the conventional view that task-specific training guarantees superior performance. Our findings suggest that these general-purpose models already encode powerful temporal patterns and vegetation phenology, enabling them to serve as highly efficient and powerful **plug-and-play** forecasting tools without the need for any task-specific training.

Looking ahead, we plan several extensions to further enhance model applicability and robustness:

- Incorporating additional climatological drivers (e.g., temperature, precipitation) as input features to improve forecasts under varying environmental conditions;
- Exploring multi-variate modeling to enable the joint modeling of multiple vegetation indices and ecosystem variables for more comprehensive ecological status prediction; and
- Conducting broader spatial transfer experiments at larger geographic scales to rigorously evaluate model generalization, thereby supporting global vegetation monitoring.

Acknowledgments

This work was partially supported by the U.S. National Science Foundation (NSF) under Grant AGS-2426655. We thank the anonymous reviewers for their constructive comments and suggestions. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

References

Alessandri, A.; Catalano, F.; De Felice, M.; Van Den Hurk, B.; Doblas Reyes, F.; Boussetta, S.; Balsamo, G.; and Miller, P. A. 2017. Multi-scale enhancement of climate prediction over land by increasing the model sensitivity to vegetation variability in EC-Earth. *Climate dynamics*, 49(4): 1215–1237.

Ansari, A. F.; Stella, L.; Turkmen, C.; Zhang, X.; Mercado, P.; Shen, H.; Shchur, O.; Rangapuram, S. S.; Arango, S. P.; Kapoor, S.; et al. 2024. Chronos: Learning the language of time series. *arXiv preprint arXiv:2403.07815*.

Box, G. E.; and Pierce, D. A. 1970. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American statistical Association*, 65(332): 1509–1526.

Breiman, L. 2001. Random forests. *Machine learning*, 45(1): 5–32.

Das, A.; Kong, W.; Sen, R.; and Zhou, Y. 2024. A decoder-only foundation model for time-series forecasting. In *Forty-first International Conference on Machine Learning*.

Friedman, J. H. 2001. Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189–1232.

Hochreiter, S.; and Schmidhuber, J. 1997. Long short-term memory. *Neural computation*, 9(8): 1735–1780.

Jensen, J. L.; Humes, K. S.; Vierling, L. A.; and Hudak, A. T. 2008. Discrete return lidar-based prediction of leaf area index in two conifer forests. *Remote Sensing of Environment*, 112(10): 3947–3957.

Jiang, B.; Liang, S.; Wang, J.; and Xiao, Z. 2010. Modeling MODIS LAI time series using three statistical methods. *Remote Sensing of Environment*, 114(7): 1432–1444.

Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; and Liu, T.-Y. 2017. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30.

Kingma, D. P. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Liu, Q.; Yang, M.; Mohammadi, K.; Song, D.; Bi, J.; and Wang, G. 2022a. Machine learning crop yield models based on meteorological features and comparison with a process-based model. *Artificial Intelligence for the Earth Systems*, 1(4): e220002.

Liu, T.; Jin, H.; Li, A.; Fang, H.; Wei, D.; Xie, X.; and Nan, X. 2022b. Estimation of vegetation leaf-area-index dynamics from multiple satellite products through deep-learning method. *Remote sensing*, 14(19): 4733.

Liu, Y.; Qin, G.; Shi, Z.; Chen, Z.; Yang, C.; Huang, X.; Wang, J.; and Long, M. 2025. Sundial: A Family of Highly Capable Time Series Foundation Models. *arXiv preprint arXiv:2502.00816*.

Myneni, R. B.; Knyazikhin, Y.; and Park, T. 2015. MYD15A2H MODIS/Aqua Leaf Area Index/FPAR 8-Day L4 Global 500m SIN Grid.

Seabold, S.; and Perktold, J. 2010. statsmodels: Econometric and statistical modeling with python. In *9th Python in Science Conference*.

Winters, P. R. 1960. Forecasting sales by exponentially weighted moving averages. *Management science*, 6(3): 324–342.

Yan, K.; Wang, J.; Peng, R.; Yang, K.; Chen, X.; Yin, G.; Dong, J.; Weiss, M.; Pu, J.; and Myneni, R. B. 2023. HiQ-LAI: A high-quality reprocessed MODIS LAI dataset with better spatio-temporal consistency from 2000 to 2022. *Earth System Science Data Discussions*, 2023: 1–32.