Materials Research Proceedings 69 (2026) 271-274 https://doi.org/10.21741/9781644904251-48

Noise Robust ML Sunspot Forecasts for Cycle Aware Space Weather Operations in the Cycle 25 Era

Esraa Zeki Mohammed[1,a *], Ahmed Ghareeb[2,b] and Mohammed Rashad Baker[3,c]

[1]Higher Institute of Communication, ITPC, Ministry of Communications, Iraq

[2]Department of Mechanical Engineering, College of Engineering, University of Kirkuk, Iraq

[3]Department of Software, College of Computer Science and Information Technology, University of Kirkuk, Iraq

[a]esraa@uokirkuk.edu.iq, [b]aghareeb@uokirkuk.edu.iq, [c]mohammed.rashad@uokirkuk.edu.iq

Keywords: Solar Cycle Prediction, Machine Learning, Symbolic Regression, Stacked Ensemble, Space Weather Forecasting, Noise Robustness, Satellite-Drag Mitigation

Abstract. Accurate monthly solar activity forecasting is vital for satellite operations and ionospheric modeling. We developed a lightweight, interpretable machine-learning framework using the complete SILSO record (1749–2024) to predict the 13-month smoothed sunspot number (SSN) for solar cycles 20–24 and the rising phase of cycle 25. We tested linear regression, tree-based models, their stacked ensemble, and a symbolic regressor, assessing robustness by injecting Gaussian noise (σ = 0, 1, 3 SSN) in to training data. Across 75 experiments, median performance was r^2 = 0.86 and RMSE< 31 SSN. The framework delivers deep-learning-class accuracy with negligible computational overhead and full transparency, offering an operationally practical tool as solar cycle 25 approaches its peak.

Introduction and Related Works

Solar activity (sunspots, flares, CMEs) impacts critical infrastructure like power grids, satellites, and GNSS, necessitating accurate 13-month smoothed SSN forecasts [1]-[3], especially given societal risks like geoelectric hazards and GNSS degradation [4]-[6]. Recent SSN forecasting increasingly uses diverse hybrid and ensemble ML models (e.g., XG-SN, LSTMs) and physical proxies [7]-[12], while NOAA's operational nonlinear forecast [2] and physics-based precursors [13] remain strong comparators. However, few studies report per-cycle generalization. Grounded in SILSO V2.0 [10] and the SWPC context [5],[4], our work addresses this gap by presenting a compact, cycle-aware ML pipeline. This framework (i) contrasts transparent linear baselines with stacked ensemble and symbolic regression, (ii) reports per-cycle R^2/RMSE across Cycles 21–24, and (iii) stress-tests robustness to label-noise.

Methodology

1. Data and Feature Engineering

We utilized the monthly 13-month smoothed International Sunspot Number (V2.0, WDC-SILSO) for cycles 21–24 (1976–2019), training models on past data (raw data: SN_m_tot_V2.0.csv) via an expanding one-step-ahead setup with per-cycle metrics. A compact, interpretable regressor for smoothed SSN (y_t) was built using z-scored features, including: SSN lags (6, 12, 24 months), rolling statistics, cycle phase (raw, normalized, asymmetric), 11-year sine/cosine harmonics, and interaction terms.

2. Models

In this study we three models, these models are: (a) LinearReg: A ridge-regularized linear baseline predicting smoothed SSN from all features, penalizing large coefficients to prevent overfitting. (b) Stacked: A stacked ensemble combining diverse base learners (e.g., Elastic Net, Gradient-Boosted Trees, RF) using a ridge meta-learner, trained on K out-of-fold (OOF) predictions from expanding windows to prevent leakage. (c) Symbolic: Symbolic regression using a small operator set (add,

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 license. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI. Published under license by Materials Research Forum LLC.

Materials Research Proceedings 69 (2026) 271-274 https://doi.org/10.21741/9781644904251-48

subtract, multiply, etc.), selecting compact formulas via a Pareto trade-off between prediction error and complexity.

3. Training, Settings, Hyper-parameter Tuning and Evaluation Metrics

Training utilized an expanding-window cross-validation setup to generate one-step-ahead monthly predictions for solar cycles 21–24. Hyperparameters were tuned for each model-ridge/elastic net for linear, boosted/random forests for trees, and population-based search for symbolic regression-with early stopping based on validation RMSE. Light Gaussian noise was added during training for robustness. All features were z-scored per window, using a fixed scaler for validation and test months. Performance was assessed using "per-cycle metrics," specifically calculating the RMSE and the R^2 independently for each cycle.

Results and Discussion

1. Overall accuracy

Across Cycles 21–24, LinearReg achieved R^2 = 0.876 / RMSE = 23.28 SSN, Stacked = 0.878 / 23.17 SSN (≈0.5% gain), and Symbolic = 0.856 / 25.04 SSN, showing all models perform well with Stacked only slightly ahead. As shown in Fig. 1(a–b), rolling means and short lags are highly collinear (ρ ≈ 0.92–0.98) while CyclePhase adds complementary cues (ρ ≈ 0.3), justifying ridge regularization. Linear and Stacked capture cycle peaks and troughs closely, whereas Symbolic appears smoother and slightly off near extrema.

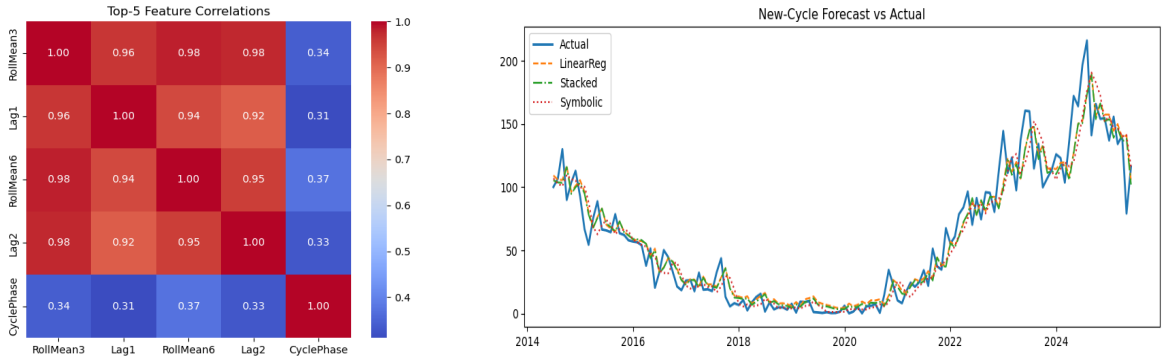


Fig. 1(a):*Collinearity Among Lag/Trend and Cycle-Phase Features,* ***(b)****Monthly SSN Forecasts vs. SILSO Truth During the Cycle-24 Minimum and Rise to Cycle-25 (2014–2025)*

2. Per-cycle behavior

Cycle-level R^2 is stable across methods: in Cycles 21–23, Stacked is about +0.001–0.003 over LinearReg, while in Cycle 24 they are essentially tied. The Symbolic model is competitive (≈0.86–0.87 in Cycles 21–23; ≈0.833 in Cycle 24) but consistently below LinearReg/Stacked, and per-cycle RMSE likewise shows small gains for Stacked—most evident in higher-amplitude cycles; while Symbolic carries a modest penalty; see Fig. 2(a) for R^2 and Fig. 2(b) for RMSE.

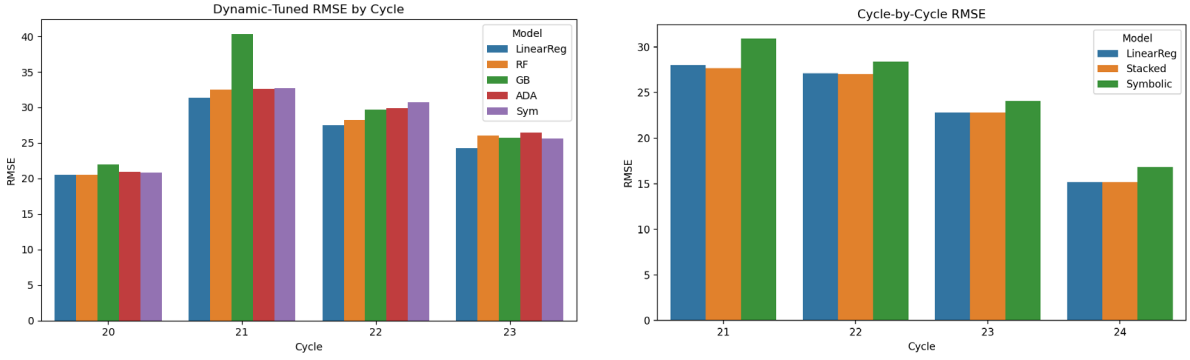


Fig. 2*.Cycle-by-Cycle RMSE (a) Dyanamic (b) Static Settings, Cycles (21–24)*

Materials Research Proceedings 69 (2026) 271-274 https://doi.org/10.21741/9781644904251-48

As shown in Figs. 2, RMSE decreases with declining cycle amplitude (≈28–31 SSN in C21 to ≈15–17 in C24). Stacking slightly improves accuracy in strong cycles, while Symbolic regression trades precision for simplicity. Under dynamic tuning, Linear and RF remain most consistent (R^2 ≈ 0.82–0.88), Gradient Boosting spikes in C21, and Symbolic overestimates peaks as tree ensembles occasionally overshoot.

3. Computation versus benefit

LinearReg runs ~0.018s per cycle versus ~35.6s for Stacked and ~28.8s for Symbolic—about 2,000× and 1,650× slower; while Stacked cuts average RMSE by <0.5% and Symbolic is modestly worse. Per-cycle results mirror this: Stacked offers tiny gains in C21–C23 and ties Linear in C24, whereas Symbolic is consistently slower and less accurate, as shown in Table (1).

Table 1: Per-Cycle Accuracy and Runtime for Monthly Smoothed Sunspot Number Predictions

Cycle	Linear (RMSE / R^2 / s)	Stacked (RMSE / R^2 / s)	Symbolic (RMSE / R^2 / s)
21	27.98 / 0.885 / 0.03	27.65 / 0.888 / 34.67	30.9 / 0.86 / 28.9
22	27.11 / 0.879 / 0.01	27.02 / 0.880 / 37.07	28.36 / 0.867 / 28.85
23	22.82 / 0.878 / 0.02	22.80 / 0.878 / 35.04	24.06 / 0.864 / 28.86
24	15.19 / 0.864 / 0.01	15.19 / 0.864 / 35.75	16.83 / 0.833 / 28.7

4. Interpretation of the symbolic form

The optimal symbolic expression reduces to a quasi-linear, interpretable ARX-style law combining lag-1 and lag-2 SSN, a minor cycle-phase correction, and 11-year harmonics (sine/cosine), utilizing inputs for SSN lags (X_3, X_4) and cycle phase (X_5). This simplicity mirrors the high performance of LR and explains why the Stacked ensemble yields only marginal gains without exogenous precursors. At the monthly horizon, the most predictive content is already captured by autoregressive memory plus smooth cycle phase and seasonality.

5. Operational alignment and robustness

Per-cycle reporting matches operational reality shown in Fig.3, and under label-noise stress (σ=0–3 SSN) the linear model shifts by less than one SSN while the stacked ensemble keeps its slight edge. Across Cycles 20–23, all models retain high fit (R^2≈0.80–0.88) with RMSE tracking amplitude; light noise (σ=1) regularizes (damping GB Cycle-21 spike), σ=3 raises errors but preserves rankings. Linear and AdaBoost (RF close) are most stable, Symbolic is consistent but higher RMSE—supporting L2-style regularization or light target jitter for robust monthly updates.

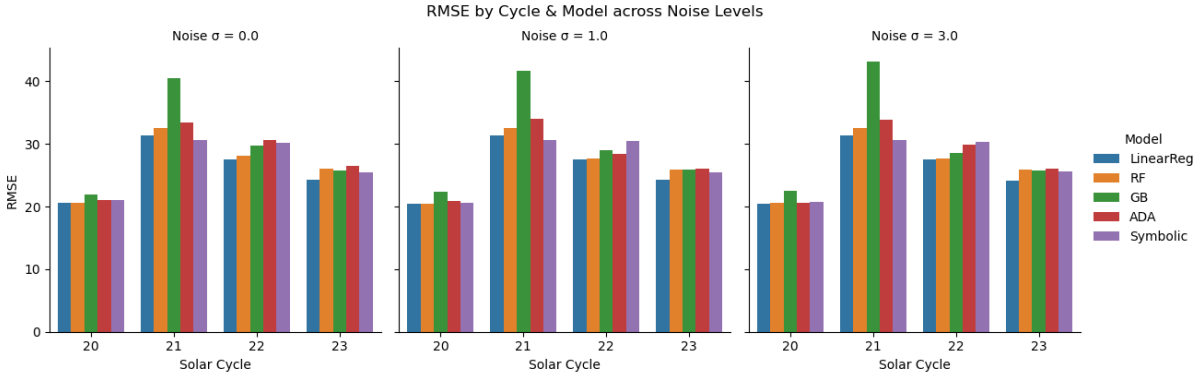


Fig 3. *Noise-Robustness Across Solar Cycles: RMSE vs. Label-Noise Levels (σ∈{0, 1, 3})*

Regularized linear models on lags/phase/Fourier are a strong monthly SSN default; stacking is worthwhile mainly with strong cycles or added covariates due to compute cost, while symbolic models aid auditability and physics insight. Our study used SSN-only features; future work should add physical precursors and include phase-specific and peak-timing metrics to pinpoint operational value.

Conclusion

We developed a lightweight, interpretable framework for monthly 13-month smoothed SSN forecasting (Cycles 20–24, rise of 25) using endogenous predictors (lags, rolling means, harmonics, time/phase). We compared linear regression, tree-based models (RF/GB/ADA), a stacked ensemble, and symbolic regression, testing robustness via Gaussian label noise. Across 75 runs, median performance was $R^2 \approx 0.86$ (RMSE <31SSN). Linear regression proved near-optimal (Cycle-24 RMSE ≈15–24), while stacking offered minor RMSE gains at high compute cost. Symbolic regression was competitive ($R^2 > 0.83$) but missed peaks. Errors scaled with amplitude; light noise regularized. Linear models are favored operationally. Future work should add physics-based inputs and peak-timing metrics.

References

[1] K. Petrovay, "Solar cycle prediction," Living Reviews in Solar Physics, vol. 17, no. 2, 2020. https://doi.org/10.1007/s41116-020-0022-z

[2] L. A. Upton and D. H. Hathaway, "Solar Cycle Precursors and the Outlook for Cycle 25," J. Geophys. Res.: Space Physics, vol. 128, no. 11, e2023JA031681, 2023. https://doi.org/10.1029/2023JA031681

[3] NOAA/NWS SWPC, "NOAA forecasts quicker, stronger peak of solar activity," Dec. 19, 2023.

[4] J. J. Love, D. A. Jonas, and A. Kelbert, "Mapping a Magnetic Superstorm: March 1989 Geoelectric Hazards in the United States," Space Weather, vol. 20, no. 5, 2022. https://doi.org/10.1029/2021SW003030.

[5] W. Nie et al., "The Mechanism for GNSS-Based Kinematic Positioning Degradation at High Latitudes Under the March 2015 Great Geomagnetic Storm," Space Weather, vol. 20, no. 6, 2022. https://doi.org/10.1029/2022SW003132

[6] V. Lanabere, Y. Y. Podladchikova, E. I. Gordeev, A. Wintoft, and M. Palmroth, "Characterizing the distribution of extreme geoelectric field gradients in North America," J. Space Weather Space Clim., vol. 14, 2024, Art. 12. https://doi.org/10.1051/swsc/2024012.

[7] X. Su, B. Liang, S. Feng, Y. Cai, W. Dai, and Y. Yang, "Solar cycle prediction using a combinatorial deep learning model," Mon. Not. R. Astron. Soc., vol. 527, no. 3, pp. 5675–5682, 2024. https://doi.org/10.1093/mnras/stad3451

[8] H. Zhu, W. Zhu, and M. He, "Predicting Solar Cycle 25 Using an Optimized Long Short-Term Memory Model Based on Sunspot Area Data," Adv. Space Res., vol. 71, no. 8, pp. 3521–3534, 2023. https://doi.org/10.1016/j.asr.2023.02.025

[9] WDC-SILSO, "Sunspot Number graphics (daily, monthly, 13-month smoothed)," accessed Sept. 2025. [Online]. Available: https://www.sidc.be/SILSO/ssngraphics

[10] J.-V. Rodríguez, V. M. Sánchez Carrasco, I. Rodríguez-Rodríguez, A. J. Pérez Aparicio, and J. M. Vaquero, "Hemispheric Sunspot Number Prediction for Solar Cycles 25 and 26 Using Spectral Analysis and Machine Learning Techniques," Solar Physics, vol. 299, no. 116, 2024. https://doi.org/10.1007/s11207-024-02363-2

[11] D. Sierra-Porta, M. Tarazona-Alvarado, and D. D. Herrera Acevedo, "Predicting sunspot number from topological features in spectral images I: Machine learning approach," Astronomy and Computing, vol. 48, p. 100857, 2024. https://doi.org/10.1016/j.ascom.2024.100857

[12] V. Kumar, "Enhancing Solar Cycle 25 and 26 Forecasting with Vipin-Deep-Decomposed-Recomposed Rolling-window (vD2R2w) Model on Sunspot Number Observations," Solar Physics, vol. 299, no. 147, 2024. https://doi.org/10.1007/s11207-024-02389-6

[13] M. S. Miesch et al., "Solar Cycle Prediction at NOAA's Space Weather Prediction Center," Space Weather, 2025. https://doi.org/10.1029/2025SW004444