Boletim da Sociedade Astronômica Brasileira, **37**, no. 1, 341-342

Sunspot Number Forescasting via a Hybrid Hilbert-Huang Transform and LSTM Neural Network Model

Y. Lopes, Y. Barros & J. D. do Nascimento

[1] Universidade Federal do Rio Grande do Norte. e-mail: ylanaakarolina@gmail.com, jdonascimento@fisica.ufrn.br

Abstract. Sunspot numbers exhibit strong nonlinear and nonstationary variability, challenging classical spectral tools such as the Fourier Transform, which assumes global linearity and stationarity. As a result, Fourier-based models often miss amplitude modulation and cycle-to-cycle changes in solar activity.
We compare Fourier analysis with a hybrid forecasting framework that couples the Hilbert–Huang Transform (HHT) to a Long Short-Term Memory (LSTM) neural network. Monthly sunspot data (1749–2010) were decomposed into intrinsic mode functions isolating high-frequency variations, the 11-year cycle, and longer-term modulations. These components were then used as structured inputs to train the LSTM model for short- and long-term prediction.
Forecasts were evaluated by comparing reconstructed series against observed sunspot numbers, with accuracy assessed through correlation metrics and long-horizon predictive stability. Fourier-based models captured the dominant periodicity but failed to reproduce nonlinear modulations, while the HHT–LSTM approach maintained high fidelity over multi-decadal timescales.
The hybrid method consistently outperformed Fourier analysis, highlighting the importance of adaptive decomposition and sequence learning for solar activity prediction. Future applications may extend this framework to solar irradiance proxies or incorporate physics-informed constraints to further refine cycle forecasting.

Resumo. Os números de manchas solares exibem variabilidade fortemente não linear e não estacionária, o que desafia ferramentas espectrais clássicas como a Transformada de Fourier, que assume linearidade e estacionariedade globais. Como resultado, modelos baseados em Fourier frequentemente não capturam modulações de amplitude e variações de ciclo a ciclo na atividade solar.
Comparamos a análise de Fourier com uma abordagem híbrida de previsão que combina a Transformada de Hilbert–Huang (HHT) a uma rede neural Long Short-Term Memory (LSTM). Dados mensais de manchas solares (1749–2010) foram decompostos em funções modais intrínsecas, isolando variações de alta frequência, o ciclo de 11 anos e modulações de longo prazo. Esses componentes foram então usados como entradas estruturadas para treinar o modelo LSTM em previsões de curto e longo alcance.
As previsões foram avaliadas comparando-se as séries reconstruídas com os números observados de manchas solares, utilizando correlações e estabilidade preditiva de longo horizonte como métricas. Os modelos baseados em Fourier capturaram a periodicidade dominante, mas falharam em reproduzir modulações não lineares, enquanto a abordagem HHT–LSTM manteve alta fidelidade ao longo de escalas multidecadais.
O método híbrido superou consistentemente a análise de Fourier, destacando a importância de decomposição adaptativa e aprendizado sequencial para previsões da atividade solar. Aplicações futuras podem estender esse framework para proxies de irradiância solar ou incorporar restrições físico-informadas para refinar ainda mais a previsão dos ciclos.

Keywords. (Sun:) sunspots – Sun: activity – Methods: data analysis

1. Introduction

Sunspot numbers are among the most important tracers of solar magnetic activity, encoding the rhythmic rise and fall of the ~ 11-year solar cycle as well as longer-term modulations driven by the Sun's complex dynamo (Pietrafesa et al. (2024); Barnhart & Eichinger (2010)). Yet, despite centuries of observations, predicting their temporal evolution remains challenging (Wang (2021)). The difficulty arises from the inherently nonlinear and nonstationary nature of solar variability, which violates the assumptions of classical spectral techniques such as the Fourier Transform (Huang et al. (1998)). Because these traditional methods require global stationarity and fixed-frequency components, they often fail to capture amplitude modulation, intermittency, and the multiscale structure characteristic of real solar cycles (Kolotkov et al. (2015)).

To overcome these limitations, adaptive time–frequency methods have gained attention (Wu & Huang (2008)). In particular, the Hilbert–Huang Transform (HHT) provides a fully empirical framework capable of decomposing a signal into intrinsic mode functions (IMFs), each representing a locally defined oscillatory component (Huang et al. (1998); Barnhart & Eichinger (2010)). This decomposition naturally isolates high-frequency activity, the fundamental 11-year cycle, and even multi-decadal or centennial variations, offering a physically interpretable view of solar magnetic behavior (Kolotkov et al. (2015); Pietrafesa et al. (2024)). Meanwhile, Long Short-Term Memory (LSTM) neural networks have emerged as powerful tools for learning temporal dependencies in complex datasets (Yang et al. (2023); Zhu et al. (2023)). Their gated architecture allows for long-term memory retention (as illustrated in Fig. 2), but they rely on clean, structured inputs and can struggle when confronted with raw, mixed-frequency data typical of direct sunspot records (Wang (2021); Leung & Zhao (2021)).

Motivated by these complementary strengths, we develop a hybrid HHT–LSTM approach that integrates adaptive decomposition with deep sequence learning (Yang et al. (2023)). Monthly sunspot numbers from the SIDC (1749–2010) were first decomposed into IMFs that capture high-frequency fluctuations, the 11-year cycle, multi-decadal variability, centennial-scale trends, and a long-term baseline (Barnhart & Eichinger (2010)). These modes were then used as structured inputs for an LSTM model trained to learn the dynamical behavior encoded in each component (Zhu et al. (2023)). This strategy allows the network to focus on the temporal relationships within individual scales,

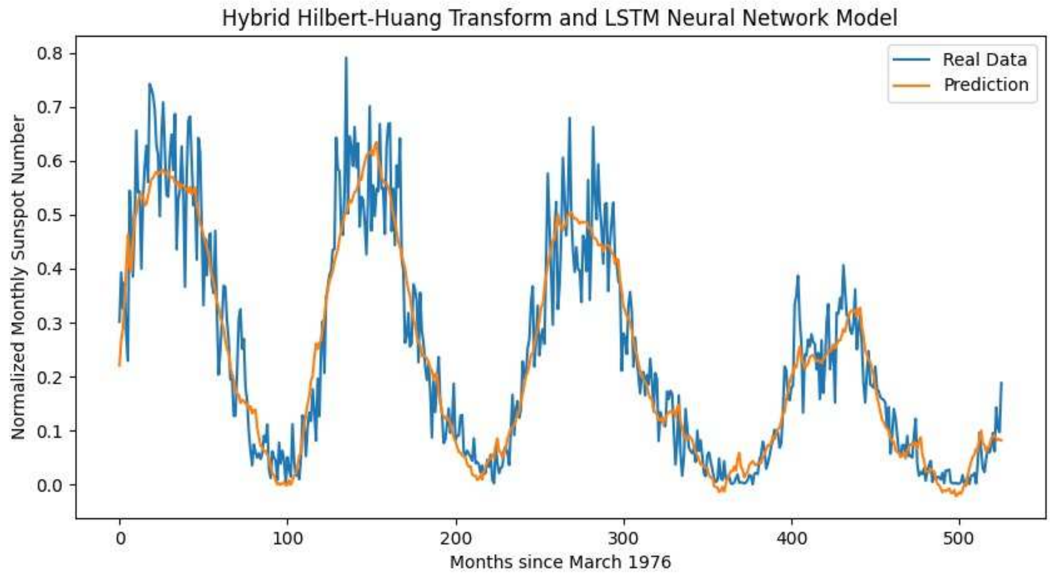

Figure 1. Long-term prediction of normalized monthly sunspot numbers using a hybrid Hilbert-Huang Transform (HHT) and LSTM model. The model's forecast (orange line) closely aligns with the observed real data (blue line) over a 500-month (~ 40-year) period from March 1976, demonstrating its efficacy in capturing the quasi-periodic nature of solar activity.

rather than attempting to untangle all variability at once (Wu & Huang (2008)).

The physical relevance of the underlying data is illustrated by the evolution of sunspots across activity phases: sparse and simple structures during Solar Minimum give way to numerous, complex magnetic regions at Solar Maximum, reflecting the modulation of the global dynamo (Pietrafesa et al. (2024)). Figure 1 presents both the empirical mode decomposition and the long-term forecasting results. When predicting normalized monthly sunspot numbers over a 500-month (~ 40-year) interval starting in March 1976, the hybrid model's forecast (orange) closely tracks the observed data (blue), successfully reproducing the quasi-periodic behavior and amplitude variations of multiple solar cycles (Zhu et al. (2023); Wang (2021)).

Overall, the HHT–LSTM framework captures the nonlinear and nonstationary features of solar activity far more effectively than Fourier-based or other traditional approaches (Huang et al. (1998); Kolotkov et al. (2015)). By pairing adaptive signal decomposition with deep learning, this method provides a powerful new direction for long-term solar cycle prediction (Yang et al. (2023); Zhu et al. (2023)). It opens avenues for applying data-driven modeling to irradiance proxies, space weather forecasting, and future solar dynamo studies where nonstationarity plays a central role (Leung & Zhao (2021)).

References

Barnhart, B. L. & Eichinger, W. E., 2010, Solar Physics, 269(2), 439–449
Huang, N. E. et al., 1998, Proc. R. Soc. Lond. A, 454(1971), 903–995
Kolotkov, D. Y., Broomhall, A.-M., & Nakariakov, V. M., 2015, MNRAS, 453(3)
Leung, T. & Zhao, T., 2021, Applied Stochastic Models in Business and Industry
Pietrafesa, L. J., Bao, S. & Gayes, P. T., 2024, European Journal of Science, Innovation and Technology, 4, 339-345
Wang, Y., 2021, Research in Astronomy and Astrophysics, 21, 23
Wu, Z. & Huang, N. E., 2008, Advances in Data Science and Adaptive Analysis
Yang, J. et al., 2023, IEEJ Transactions on Electrical and Electronic Engineering
Zhu, H. et al., 2023, Advances in Space Research, 71(8), 3521–3531