Spectral Modeling of Periodic Scientific Time Series via Fourier–Galerkin Regression

Isabella Li
Lexington High School
Lexington, Massachusetts, USA
isal@bu.edu

Zigan Wang*
Shenzhen International Graduate School, Tsinghua University
Shenzhen, Guangdong, China
wangzigan@sz.tsinghua.edu.cn

This paper is accepted for the ACM 19th International Conference on Web Search and Data Mining (WSDM 2026)

Abstract

Periodic phenomena are ubiquitous across various scientific domains, such as climatic cycles. However, identifying and modeling these periodic behaviors in large-scale, heterogeneous scientific datasets pose significant challenges due to the inherent complexity and variability of phase shifts and signal misalignment. In this work, we propose a Fourier–Galerkin (FG) regression model that decomposes the time series into a linear trend and a periodic residual, using a FG basis for the residual. Our core model, FG-Linear, captures periodic behavior effectively by learning a linear combination of cosine and sine components, while addressing phase non-uniqueness with phase offsets for each frequency. We also introduce a deep spectral variant (FG-Deep), which incorporates a neural network to model non-linearities in the spectral features. Experimental results on the NOAA Global Summary of the Year climate dataset show that FG-Linear outperforms trend-only models and deep learning baselines, achieving the lowest test MSE. While FG-Deep provides flexibility, its performance is similar to FG-Linear on this small and regular dataset, demonstrating the efficiency and interpretability of the linear FG approach.

CCS Concepts

• **Information systems** → **Data mining**; • **Computing methodologies** → *Spectral methods*; • **Mathematics of computing** → Mathematical optimization.

Keywords

Fourier-Galerkin Method, Periodic Time Series, Spectral Modeling, Phase Invariance, Scientific Data

ACM Reference Format:
Isabella Li and Zigan Wang. 2026. Spectral Modeling of Periodic Scientific Time Series via Fourier–Galerkin Regression. In *Proceedings of the Nineteenth ACM International Conference on Web Search and Data Mining (WSDM '26), February 22–26, 2026, Boise, ID, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/XXXXXX.XXXXXX

*Corresponding author.

WSDM '26, Boise, ID, USA.

ACM ISBN 979-8-4007-2292-9/2026/02
https://doi.org/10.1145/XXXXXX.XXXXXX

1 Introduction

Periodic phenomena are fundamental in numerous scientific domains, ranging from stock market cycles to environmental systems like climate change. Identifying and modeling periodic behaviors within these systems is crucial for understanding their underlying dynamics. However, extracting these periodic signals from large, complex, and often noisy datasets is not straightforward.

Classical Fourier–Galerkin (FG) analysis provides a natural language for periodic structure, and have a long history in numerical analysis of periodic differential equations [1–3]. And recent Fourier transforms [4, 5] assume periodic signals are perfectly aligned or that phase shifts are negligible, which is rarely the case in real-world data. However, in data mining practice, spectral modeling is often implemented ad-hoc via FFT features without a clear connection back to approximation theory, and without explicit handling of trends and phase issues.

In this work, we aim to bridge FG theory and practical spectral modeling for scientific time series. Starting from the error bounds for FG approximations of periodic functions, as Figure 1, we design a family of spectral models that: (1) Decompose each series into a trend and a periodic residual; (2) Project the residual into a finite-dimensional FG subspace; (3) Learn either a linear FG regression model or a deep spectral model on these features.

We validate the approach on real-world climate data from NOAA's Global Summary of the Year dataset. Although this dataset is relatively small and regular, it serves as a realistic testbed where trend and periodic components coexist. These results support the view that FG style spectral modeling is both theoretically grounded and empirically strong, and that deep spectral models are best viewed as an extension for more complex, larger-scale Web data rather than a universal replacement. The code and examples of our method could be found at https://github.com/fg-research/phasefg.

2 Methodology

Problem setup. We consider a univariate scientific data ${y_t}_{t=1}^{T}$, where each y_t is a statistic (e.g., annual mean temperature). Our goal is to model y_t in a way that: (1) Separates long-term trend from periodic fluctuations; (2) Represents periodic structure in a finite FG subspace; (3) Admits efficient learning and clear interpretation.

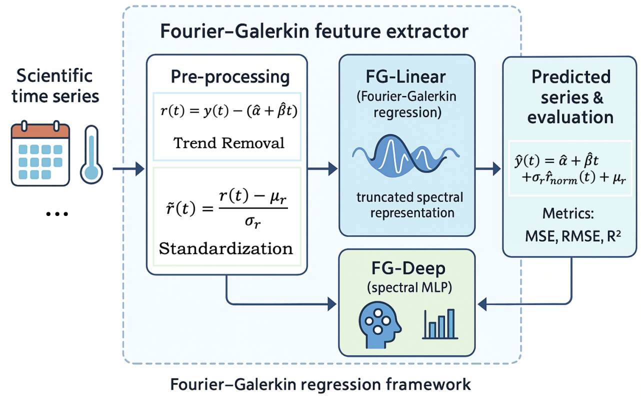

Figure 1: Fourier–Galerkin Regression framework.

Following the standard practice in time series decomposition, we assume $y(t) = \alpha + \beta t + r(t)$, where t is a possibly normalized time index, $\alpha + \beta t$ is a global linear trend, and $r(t)$ is a mean-zero residual that we aim to approximate by a truncated spectral expansion.

FG representation of residuals. We first fit α, β by least squares on the full series and compute $\hat{r}(t) = y(t) - (\hat{\alpha} + \hat{\beta}t)$. We then standardize the residual $\tilde{r}(t) = \frac{\hat{r}(t)-\mu_r}{\sigma_r}$, where μ_r, σ_r are the mean and standard deviation of $\hat{r}(t)$. Let τ denote a normalized time variable obtained by linearly re-scaling t to have zero mean and unit variance. Fix a base angular frequency ω based on the span of τ. For a chosen truncation order k, we define the standard FG basis:

$$\phi_0^{\cos}(\tau) = 1, \quad \phi_k^{\cos}(\tau) = \cos(k\omega\tau), \quad \phi_k^{\sin}(\tau) = \sin(k\omega\tau) \tag{1}$$

And, the truncated FG approximation of $\tilde{r}$ is:

$$\tilde{r}_K(\tau) = b + \sum_{k=0}^{K} w^{\cos} * k, \phi_k^{\cos}(\tau) + \sum *k = 1^K w_k^{\sin}, \phi_k^{\sin}(\tau) \tag{2}$$

Classical FG theory implies that, under mild regularity assumptions on $\tilde{r}$, the approximation error in L_2 decays at least at a polynomial rate in K. We leverage this as a theoretical justification for using finite-dimensional spectral regression as our core model class.

Linear FG regression. In practice, for each time point τ_t we construct a feature vector:

$$x_t = \left[\phi_0^{\cos}(\tau_t), \ldots, \phi_K^{\cos}(\tau_t), \phi_1^{\sin}(\tau_t), \ldots, \phi_K^{\sin}(\tau_t)\right]^\top \in \mathbb{R}^{2K+1} \tag{3}$$

We then fit a Ridge regression model on the normalized residuals $\tilde{r}(t) \approx f(x_t) = x_t^\top w + b$, where $w \in \mathbb{R}^{2K+1}$, $b \in \mathbb{R}$, tuned by minimizing:

$$\frac{1}{T_{\text{train}}} \sum_{t \in \mathcal{T}_{\text{train}}} (\tilde{r}(t) - f(x_t))^2 + \lambda ||w||_2^2 \tag{4}$$

At prediction time, we reconstruct $y(t)$ as $\hat{y}(t) = \hat{\alpha} + \hat{\beta}t + \sigma_r f(x_t) + \mu_r$. We refer to this model as FG regression. It is exactly a truncated FG projection with a regularized linear estimator on top, and is both computationally efficient and spectrally interpretable.

Spectral deep variant (FG-Deep). To explore the potential of non-linear spectral combinations, we also consider a deep variant, FG-Deep. Instead of directly regressing $\tilde{r}(t)$ on x_t with a linear model, we feed the spectral features into a small multi-layer perceptron $h_t = \sigma(W_1 x_t + b_1)$, $\hat{r}_{\text{norm}}(t) = W_2 h_t + b_2$, where σ is a ReLU activation and the hidden dimension is kept modest (e.g., 64–32 units) to avoid overfitting on small datasets.

FG-Deep still benefits from the FG feature space and thus inherits interpretability at the feature level, but its final mapping from spectral components to residuals is non-linear. As we show in experiments, on small and highly regular sequences FG-Deep does not always outperform the linear spectral model, but it provides a flexible template for more complex Web-scale settings.

3 Experiments

Datasets and Baselines. We use the NOAA Global Summary of the Year dataset [1], which aggregates annual statistics from meteorological stations worldwide. We (1) Select several long-running stations in the United States; (2) Aggregate their annual mean temperature values by averaging across stations for each year; (3) Obtain a single annual mean temperature time series spanning 80 years (1945–2024). This yields a realistic scientific time series that combines a clear global warming trend with year-to-year variability.

We compare against the following baselines: (1) LinearTrend: The global least-squares linear fit $y(t) = \alpha + \beta t$, evaluated directly on test years. (2) FCNN: A fully-connected feed-forward neural network that maps normalized time τ directly to normalized residual $\tilde{r}(t)$, with a 64–32 MLP and ReLU activations. Prediction is transformed back to the original scale using μ_r, σ_r and the trend.

Results and Performance Analysis. Table 1 summarizes the performance metrics for various models on the datasets. The LinearTrend, which only captures the linear trend, performs the worst with a test MSE of 1.98. The FCNN model, while showing some improvement, lacks the interpretability and spectral insights provided by the FG approach. Ours FG-Linear, based on FG regression, here K=8 for trade-off the complexity and performance, achieves the best performance with a test MSE of 0.79, demonstrating the effectiveness of spectral modeling in capturing periodic components. The FG-Deep model, incorporating a deep learning component, yields an MSE of 1.48, indicating that while non-linearities may improve performance in more complex datasets, they do not offer significant advantages on this small, regular dataset.

Table 1: Performance on annual mean temperature datasets.

Method	Test MSE	R^2	RMSE
LinearTrend	1.98	0.93	1.41
FCNN (time → resid)	1.08	0.97	1.04
FG-Deep (spectral MLP)	1.48	0.95	1.21
FG-Linear (K=8)	0.79	0.98	0.89

4 Conclusion

In this work, we propose a Fourier–Galerkin regression model that efficiently captures periodic structure by projecting the residual of a time series onto a finite-dimensional FG subspace. Our core model, FG regression, separates the time series into a trend and a periodic residual, with the residual represented using a truncated Fourier

[1] https://www.ncei.noaa.gov/

expansion. Our approach balances predictive accuracy, computational efficiency, and spectral interpretability, making it a strong tool for scientific discovery in time series data.

References

[1] Claudio Canuto, R Nochetto, and Marco Verani. 2014. Adaptive Fourier-Galerkin methods. *Math. Comp.* 83, 288 (2014), 1645–1687.

[2] Henning Lange, Steven L Brunton, and J Nathan Kutz. 2021. From fourier to koopman: Spectral methods for long-term time series prediction. *Journal of Machine Learning Research* 22, 41 (2021), 1–38.

[3] Juan Peña Miralles, Pedro José Jiménez Olivo, Damián Ginestar Peiró, Gumersindo Verdú Martín, and JoséLuis Muñoz-Cobo González. 1997. A fast Galerkin method to obtain the periodic solutions of a nonlinear oscillator. *Applied mathematics and computation* 86, 2-3 (1997), 261–282.

[4] Kun Yi, Qi Zhang, Shoujin Wang, Hui He, Guodong Long, and Zhendong Niu. 2023. Neural time series analysis with fourier transform: A survey. *arXiv preprint arXiv:2302.02173* 1046 (2023), 1047–1048.

[5] Qianru Zhang, Yuting Sun, Honggang Wen, Peng Yang, Xinzhu Li, Ming Li, Kwok-Yan Lam, Siu-Ming Yiu, and Hongzhi Yin. 2025. Time Series Analysis in Frequency Domain: A Survey of Open Challenges, Opportunities and Benchmarks. arXiv:2504.07099 [cs.CE] https://arxiv.org/abs/2504.07099