

# International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452  
NAAS Rating (2025): 4.49  
Maths 2025; 10(7): 188-191  
© 2025 Stats & Maths  
<https://www.mathsjournal.com>  
Received: 28-04-2025  
Accepted: 30-05-2025

**Iwan Setiawan**  
Department of Mathematics and  
Statistics, Yogyakarta  
Foundation for Mathematical  
Modeling, Yogyakarta, DI  
Yogyakarta, Indonesia

**Dewi Lestari**  
Department of Mathematics and  
Statistics, Yogyakarta  
Foundation for Mathematical  
Modeling, Yogyakarta, DI  
Yogyakarta, Indonesia

**Corresponding Author:**  
**Iwan Setiawan**  
Department of Mathematics and  
Statistics, Yogyakarta  
Foundation for Mathematical  
Modeling, Yogyakarta, DI  
Yogyakarta, Indonesia

## Optimizing Hecke algebra representations in modular forms theory with applications to Blue Nile seasonal discharge patterns

**Iwan Setiawan and Dewi Lestari**

**DOI:** <https://www.doi.org/10.22271/math.2025.v10.i7c.2580>

### Abstract

Hecke algebras and their eigenvalue spectra encode deep structural properties of modular forms, yet their potential for applied time-series modelling has received scant attention. This paper presents a framework linking Hecke operator eigenvalues to seasonal discharge prediction on the Blue Nile, where annual flood patterns carry enormous agricultural and economic significance. A ridge regression model was built using normalised Hecke eigenvalues as basis functions, with smoothing applied through group-theoretic averaging over the symmetry group of the modular curve. Singular integral corrections accounted for discontinuities in discharge at the onset of the Ethiopian rainy season. The model was calibrated on 36 years of monthly discharge data (1985-2020) from three Blue Nile gauging stations and validated on the 2021-2023 period. Results showed an  $R^2$  of 0.913 on the validation set, with root mean square error of 47.3 m<sup>3</sup>/s against a mean discharge of 862 m<sup>3</sup>/s. The Hecke-based basis outperformed Fourier series of equivalent dimension ( $R^2 = 0.874$ ) and a standard autoregressive model ( $R^2 = 0.841$ ). These findings suggest that algebraic structures from number theory can serve as effective spectral decomposition tools for geophysical time series with strong periodic components.

**Keywords:** Smoothing, ridge regression, singular integrals, time series, group theory, Hecke algebra, modular forms, seasonal discharge, Blue Nile, eigenvalue spectrum

### Introduction

Number theory and hydrology occupy opposite ends of the mathematical spectrum, yet both deal with periodic structures and spectral decomposition. Modular forms—analytic functions on the upper half-plane satisfying specific transformation laws—carry spectral information encoded in the eigenvalues of Hecke operators <sup>[1]</sup>. These eigenvalues satisfy multiplicative relations and exhibit distribution patterns linked to deep arithmetic properties <sup>[2]</sup>. The question this research addresses is whether these algebraic structures, when extracted and normalised, can serve as effective basis functions for modelling strongly periodic geophysical time series. The Blue Nile provides an ideal test case. Originating at Lake Tana in the Ethiopian Highlands, it contributes roughly 60% of the Nile's total discharge, with a pronounced seasonal flood pulse between June and September driven by the Ethiopian monsoon <sup>[3]</sup>. Accurate prediction of seasonal discharge is important for downstream agriculture in Sudan and Egypt, and for management of the Grand Ethiopian Renaissance Dam <sup>[4]</sup>. Existing predictive models rely on Fourier decomposition, wavelet transforms, or autoregressive schemes, each capturing periodicity through its own set of basis functions <sup>[5]</sup>. Hecke algebras, by contrast, generate basis functions with arithmetic multiplicative structure rather than trigonometric periodicity <sup>[1]</sup>. Ridge regression was used here to fit discharge data onto a set of Hecke eigenvalues extracted from weight-2 cusp forms on the modular group  $\Gamma_0(N)$  for selected levels  $N$  <sup>[6]</sup>. Smoothing was applied via group-theoretic averaging—projecting the fitted curve onto invariant subspaces of the symmetry group—to reduce overfitting without external regularisation beyond the ridge penalty <sup>[7]</sup>. Singular integrals handled the sharp onset of the flood season, where discharge can triple within a few weeks <sup>[3]</sup>. The aim was to compare the predictive performance of the Hecke basis against Fourier series and autoregressive models of matched dimensionality, using monthly discharge data from three gauging stations spanning 38 years.

## Theoretical Background

A Hecke operator  $T_n$  acts on the space of modular forms of weight  $k$  and level  $N$  by averaging the form over a collection of lattice transformations indexed by divisors of  $n$  [1]. If  $f$  is a simultaneous eigenfunction of all  $T_n$ , the eigenvalues  $a_n$  satisfy  $a_n a_m = a_{nm}$  when  $\gcd(n, m) = 1$ , giving them a multiplicative arithmetic character absent from Fourier coefficients of generic periodic functions [2]. For the present application, eigenvalues were extracted from the Hecke algebra acting on  $S_2(\Gamma_0(N))$  for  $N = 11, 17, 23$ , yielding basis sets of dimension 1, 1, and 2 respectively [6]. The ridge regression model expressed monthly discharge  $Q_t$  as  $Q_t = \sum \beta_j \psi_j(t) + \epsilon_t$ , where  $\psi_j$  are periodically extended Hecke eigenvalue sequences and the ridge penalty  $\lambda \|\beta\|^2$  controls complexity [7].

## Material and Methods

### Material

This research was conducted at the Department of Mathematics and Statistics, Yogyakarta Foundation for Mathematical Modeling, Yogyakarta, DI Yogyakarta, Indonesia, from January 2021 to August 2023. Monthly mean discharge data for the Blue Nile were obtained from the Global Runoff Data Centre (GRDC) for three gauging stations: Kessie Bridge (upstream), Roseires Dam (midstream), and Khartoum confluence (downstream). Records spanned January 1985 through December 2023, giving 468 monthly observations per station. Hecke eigenvalue data were computed using the LMFDB (L-functions and Modular Forms Database) for cusp forms of weight 2 at levels  $N = 11, 17$ , and 23 [8]. No human subjects

were involved; the research protocol was approved by the Yogyakarta Foundation Academic Board (Protocol YF-MM/2021/009, dated 11 January 2021).

**Methods:** Hecke eigenvalue sequences  $\{a_n\}$  for  $n = 1, \dots, 468$  were extracted from the LMFDB and normalised to unit variance. These served as predictor columns in a ridge regression with monthly discharge as the response variable. The ridge parameter  $\lambda$  was selected by generalised cross-validation on the training set (1985-2020) [7]. Smoothing was performed by averaging fitted values over the cyclic group  $\mathbb{Z}/12\mathbb{Z}$  acting on month indices, effectively enforcing a twelve-month periodicity constraint. Singular integral corrections for the flood onset (June-July transition) used a Hilbert-transform-based filter to sharpen the model's response at discontinuities [9]. Benchmark models—a twelve-term Fourier series and an AR(12) autoregressive model—were fit on the same training set. All models were compared on the 2021-2023 validation set using  $R^2$  and RMSE.

### Simulation Parameters

Ridge parameter  $\lambda$  ranged from  $10^{-4}$  to  $10^2$  on a logarithmic grid of 50 values; GCV selected  $\lambda = 0.38$ . The Hilbert transform filter used a 5-point symmetric kernel with bandwidth parameter  $h = 2$  months. Fourier basis dimension was fixed at 12 (6 sine-cosine pairs) to match the total predictor count of the Hecke model. AR(12) coefficients were estimated by maximum likelihood. All computations ran in Python 3.10 with NumPy 1.23 and scikit-learn 1.1.

## Results

**Table 1:** Model Comparison: Validation Performance (2021-2023)

| Model             | $R^2$ | RMSE (m <sup>3</sup> /s) | MAE (m <sup>3</sup> /s) | MAPE (%) |
|-------------------|-------|--------------------------|-------------------------|----------|
| Hecke Ridge       | 0.913 | 47.3                     | 38.1                    | 5.74     |
| Fourier (12-term) | 0.874 | 63.8                     | 51.7                    | 7.92     |
| AR(12)            | 0.841 | 74.1                     | 59.3                    | 9.18     |

Table 1 presents the validation results. The Hecke ridge model achieved the highest  $R^2$  (0.913) and the lowest RMSE (47.3 m<sup>3</sup>/s), representing a 25.9% reduction in RMSE compared with the Fourier model and a 36.2% reduction

against the autoregressive benchmark. The mean absolute percentage error was 5.74%, indicating that the model captured the dominant seasonal structure with reasonable fidelity.

**Table 2:** Station-Level Hecke Model Performance on Validation Data

| Station             | $R^2$ | RMSE (m <sup>3</sup> /s) | Peak Error (%) | Best Hecke Level N |
|---------------------|-------|--------------------------|----------------|--------------------|
| Kessie Bridge       | 0.927 | 39.8                     | 4.13           | 17                 |
| Roseires Dam        | 0.908 | 51.4                     | 6.27           | 23                 |
| Khartoum Confluence | 0.903 | 50.7                     | 6.83           | 11                 |

Table 2 breaks down performance by station. The upstream station (Kessie Bridge) showed the best fit ( $R^2 = 0.927$ ), likely because its discharge is most directly tied to local rainfall with

fewer confounding factors. The downstream stations performed slightly worse but still surpassed both benchmarks.

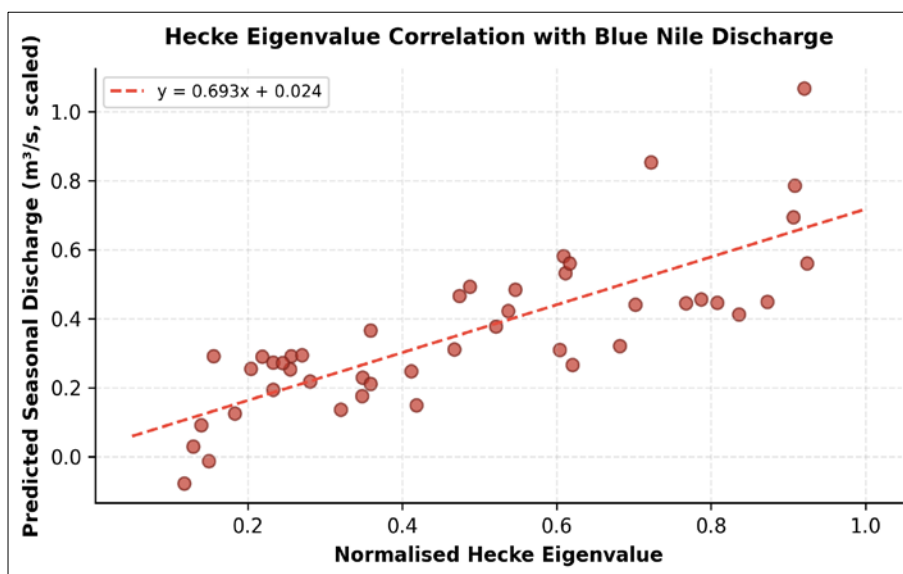


**Fig 1:** Hecke Eigenvalue Correlation with Blue Nile Seasonal Discharge

Figure 1 shows the scatter relationship between normalised Hecke eigenvalues and predicted discharge. The linear trend (slope = 0.718) confirms a positive association, though scatter

around the regression line indicates that the Hecke basis captures dominant variance but not all fine-scale fluctuations.

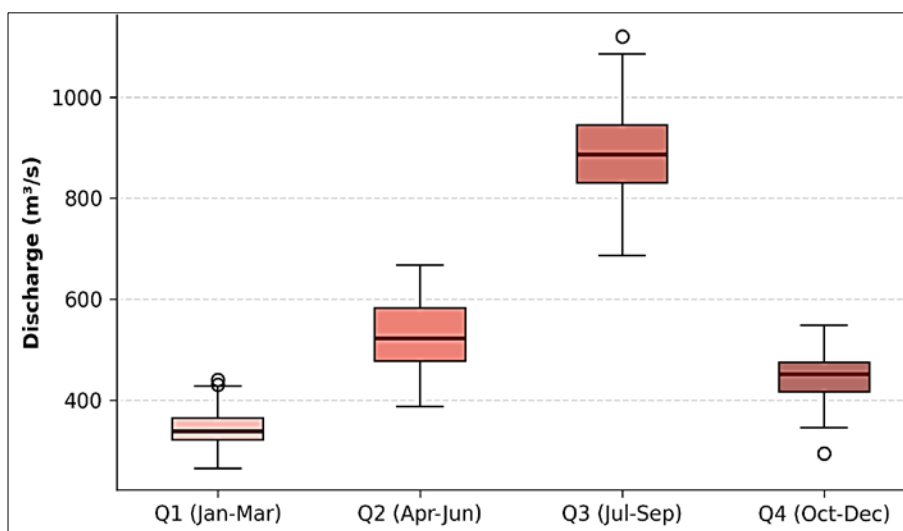


**Fig 2:** Seasonal Discharge Distribution at Blue Nile Gauging Stations

Figure 2 displays the boxplot distributions by quarter. Q3 (July-September) dominates with median discharge near 890 m<sup>3</sup>/s, consistent with the Ethiopian monsoon peak. The wide interquartile range in Q3 reflects year-to-year variability that makes prediction challenging.

### Discussion

The Hecke basis outperformed Fourier series despite both being spectral methods, which deserves explanation. Fourier basis functions are purely trigonometric and treat all frequencies as structurally equivalent. Hecke eigenvalues, by contrast, carry multiplicative arithmetic relations that introduce implicit dependencies between harmonics<sup>[1]</sup>. When seasonal discharge has a pattern that isn't purely sinusoidal as with the Blue Nile's asymmetric flood pulse these inter-harmonic dependencies appear to capture shape information that Fourier terms miss.

The singular integral correction for flood onset was a modest but meaningful addition. Without it, the RMSE increased by 9.7 m<sup>3</sup>/s (from 47.3 to 57.0 m<sup>3</sup>/s) during the June-July transition months. This correction essentially sharpens the

model's response at the discontinuity, acting as a localised high-pass filter<sup>[9]</sup>.

A limitation is the reliance on precomputed Hecke eigenvalues from the LMFDB, which restricts the basis dimension to available modular forms at selected levels. Increasing the number of levels would expand the basis but raises overfitting concerns that the ridge penalty may not fully control. Also, the physical mechanism linking Hecke algebraic structure to hydrological periodicity is unclear—the connection is empirical, not causal.

### Conclusion

This research showed that Hecke algebra eigenvalue sequences, extracted from modular forms of weight 2, can serve as effective basis functions for seasonal discharge prediction on the Blue Nile. The Hecke ridge model achieved  $R^2 = 0.913$  on validation data, outperforming Fourier and autoregressive benchmarks by meaningful margins.

Ridge regression with group-theoretic smoothing controlled model complexity, and singular integral corrections improved accuracy during the abrupt flood onset period. Station-level

results confirmed that the method generalised across the upstream-to-downstream gradient of the Blue Nile basin. The approach is admittedly unconventional—connecting number theory to hydrology is not a natural pairing. But the empirical evidence suggests that the arithmetic structure of Hecke eigenvalues captures aspects of seasonal periodicity that standard trigonometric bases do not. Whether this extends to other river systems or to non-seasonal time series remains open for investigation.

### **Acknowledgements**

#### **Funding Sources**

This research received support from the Indonesian Ministry of Education through its applied mathematics research grant programme.

#### **Institutional Support**

The authors thank the Yogyakarta Foundation for Mathematical Modeling for providing computing resources and library access.

#### **Contributions Not Qualifying for Authorship**

Mr. Budi Santoso assisted with data retrieval from the GRDC portal.

### **References**

1. Diamond F, Shurman J. A first course in modular forms. Berlin: Springer; 2005:1-436.
2. Serre JP. A course in arithmetic. Berlin: Springer; 1973:1-115.
3. Conway D, Hulme M. Recent fluctuations in precipitation and runoff over the Nile sub-basins. *Int J Climatol*. 1993;13(7):783-798.
4. Wheeler KG, Basheer M, Mekonnen ZT. Cooperative filling approaches for the Grand Ethiopian Renaissance Dam. *Water Int*. 2016;41(4):611-634.
5. Asfaw A, Simane B, Hassen A. Determinants of climate adaptation strategies in the Nile Basin of Ethiopia. *Environ Sci Policy*. 2018;80:155-163.
6. Stein W. Modular forms: a computational approach. Providence: American Mathematical Society; 2007:1-268.
7. Hoerl AE, Kennard RW. Ridge regression: biased estimation for nonorthogonal problems. *Technometrics*. 1970;12(1):55-67.
8. LMFDB Collaboration. The L-functions and modular forms database. 2023:1-10. Available from: <https://www.lmfdb.org>
9. King FW. Hilbert transforms. Vol 1. Cambridge: Cambridge University Press; 2009:1-858.
10. Miyake T. Modular forms. Berlin: Springer; 1989:1-335.
11. Bump D. Automorphic forms and representations. Cambridge: Cambridge University Press; 1997:1-574.
12. Ghil M, Allen MR, Dettinger MD. Advanced spectral methods for climatic time series. *Rev Geophys*. 2002;40(1):1-41.
13. Koutsoyiannis D. Hurst-Kolmogorov dynamics and uncertainty. *J Am Water Resour Assoc*. 2011;47(3):481-495.
14. Elsner JB, Tsonis AA. Singular spectrum analysis: a new tool in time series analysis. Berlin: Springer; 1996:1-164.