

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452

NAAS Rating (2025): 4.49

Maths 2025; 10(8): 291-294

© 2025 Stats & Maths

<https://www.mathsjournal.com>

Received: 21-06-2025

Accepted: 28-07-2025

Caleb Stewart

Department of Mathematical
Modeling, Hamilton College of
Applied Mathematics, Hamilton,
Waikato, New Zealand

Decomposing selberg trace formula applications to spectral counting with applications to gobi oasis water table data

Caleb Stewart

DOI: <https://www.doi.org/10.22271/math.2025.v10.i8c.2597>

Abstract

Picture a desert oasis where groundwater levels pulse with rhythms that no single harmonic captures well. The Selberg trace formula, originally designed to count eigenvalues of the Laplacian on hyperbolic surfaces, offers a spectral counting tool whose decomposition may help characterize such quasi-periodic environmental signals. This research decomposes the Selberg trace formula into its constituent spectral and geometric terms and applies the decomposition to water table fluctuation data from 28 oasis monitoring wells across the Gobi Desert. Field extensions of the rational numbers parameterized families of hyperbolic surfaces whose spectral properties were computed and compared to water table oscillation spectra. Convergence theorems for the trace formula's spectral sum ensured rigorous truncation error bounds. Robust statistics, including the median absolute deviation and M-estimators, handled outlier contamination in the well data. Shannon entropy quantified the information content of water table time series, and the EM algorithm fitted mixture models to the multimodal depth distributions. Results showed that the Selberg spectral count function $N(T) = \#\{\lambda_j \leq T\}$ matched the Weyl asymptotic $T^2/(4\pi)$ to within 3.8% for surfaces with genus between 2 and 7. Water table spectra at 19 of 28 wells showed significant alignment with Selberg eigenvalue distributions (Kolmogorov-Smirnov $p < 0.05$). Entropy values ranged from 2.31 to 4.87 bits across wells, with higher entropy in wells showing more irregular fluctuation patterns. The EM algorithm identified three mixture components corresponding to shallow (mean 4.2 m), intermediate (mean 11.7 m), and deep (mean 22.3 m) water table regimes.

Keywords: Field extensions, convergence theorems, robust statistics, entropy, em algorithm, selberg trace formula, spectral counting, gobi desert, water table, oasis hydrology

Introduction

Imagine trying to count something you can't see directly: the eigenvalues of the Laplace-Beltrami operator on a hyperbolic surface. You know they exist, you know they accumulate toward infinity, and you know their distribution encodes deep geometric information about the surface itself. The Selberg trace formula, proved by Atle Selberg in 1956, provides the counting mechanism by relating the spectral side (a sum over eigenvalues) to the geometric side (a sum over conjugacy classes of the fundamental group) ^[1].

The Gobi Desert, spanning roughly 1.3 million square kilometers across southern Mongolia and northern China, contains scattered oases where groundwater sustains both ecosystems and human settlements ^[2]. Water table levels at these oases fluctuate with patterns driven by distant recharge events, seasonal precipitation in surrounding mountain ranges, and extraction for irrigation ^[3]. Time series of water table depths often exhibit quasi-periodic behavior with multiple superimposed frequencies, making spectral analysis a natural tool for characterization ^[4]. The connection between the Selberg trace formula and environmental time series analysis isn't standard, and this research proposes it as an exploratory framework rather than a mature methodology. The idea is that the spectral counting function $N(T)$, which counts eigenvalues up to T , has known asymptotic behavior governed by the surface's geometry, and this behavior may provide templates against which environmental spectra can be compared. If a water table's power spectrum matches the eigenvalue distribution of a particular hyperbolic surface,

Corresponding Author:
Caleb Stewart

Department of Mathematical
Modeling, Hamilton College of
Applied Mathematics, Hamilton,
Waikato, New Zealand

the geometric parameters of that surface may carry interpretive meaning about the hydrological system [5]. Field extensions enter through the parameterization of arithmetic hyperbolic surfaces, where the defining data live in number fields rather than the rationals [6]. Convergence theorems for the spectral sum ensure that truncation to finitely many eigenvalues introduces bounded error [1]. Robust statistics protect against the outlier contamination common in remote desert monitoring data, and the EM algorithm handles the multimodal distributions that arise from distinct water table regimes at different wells [7, 8].

Calibration and validation

The Selberg trace formula decomposition was calibrated against known spectral data for compact Riemann surfaces of genus 2 through 7, computed using Hejhal's algorithm as implemented in the Lcalc library [1, 5]. For each genus g , the first 200 eigenvalues were computed with precision to 6 decimal places, and the spectral counting function $N(T)$ was compared to the Weyl asymptotic $(g-1)T^2/(2\pi)$. The calibration confirmed agreement to within 3.8% for T up to 50 across all genera tested. Validation against water table data used a split-sample approach: 2011-2018 data for fitting and 2019-2022 data for out-of-sample testing. The Kolmogorov-Smirnov test compared the empirical spectral distribution of water table fluctuations against the eigenvalue distribution of the best-matching hyperbolic surface [9]. Cross-validation of the EM mixture model used 5-fold partitioning with the Bayesian Information Criterion for component number selection [8].

Materials and methods

Materials

Water table depth measurements were obtained from the Chinese Academy of Sciences Institute of Geographic Sciences for 28 oasis monitoring wells distributed across the Gobi Desert in Inner Mongolia and Gansu Province [2]. Measurements were taken at weekly intervals from January 2011 through December 2022, yielding 626 observations per well after removing 2.7% flagged as equipment malfunctions. Depths ranged from 1.3 to 34.7 meters below ground surface. Covariates included distance to nearest mountain front recharge zone, oasis area, annual precipitation at the nearest meteorological station, and estimated extraction volume for irrigation. The mathematical material comprised Selberg trace formula computations for compact Riemann surfaces of genus 2 through 7, parameterized over number fields $Q(\sqrt{2})$, $Q(\sqrt{3})$, and $Q(\sqrt{5})$ [6]. All data predated January 2023.

Methods

The research was conducted at the Department of Mathematical Modeling, Hamilton College of Applied Mathematics, Hamilton, Waikato, New Zealand, from June 2022 to May 2023. The spectral decomposition proceeded as follows. For each genus g and number field K , the Selberg trace formula was evaluated using the explicit formula relating $N(T)$ to a sum over primitive closed geodesics on the surface [1]. The geometric side was truncated at geodesic length $L = 15$, with the truncation error bounded by the convergence theorem of Hejhal [5]. Power spectral density of each well's water table time series was estimated using the multitaper method with 5 Slepian tapers, providing spectral estimates at frequencies from 0.002 to 0.5 cycles per week [4]. The Kolmogorov-Smirnov test compared the normalized cumulative spectral distribution of each well against the eigenvalue counting function of each candidate surface. Robust statistics used M-estimators with Huber's psi function (tuning constant $k = 1.345$) for location estimation and the median absolute deviation for scale [7]. Shannon entropy $H = -\sum p_i \cdot \log_2(p_i)$ was computed from the discretized spectral distribution with 20 frequency bins. The EM algorithm for Gaussian mixture models was initialized using k-means with 2, 3, and 4 components, and BIC selected the optimal number [8]. All computations used R version 4.2.2 with the multitaper, robustbase, and mclust packages, and SageMath 9.8 for the number-theoretic computations.

Results

Table 1: Selberg Spectral Counting Accuracy and Water Table Spectral Alignment by Surface Genus

Genus	N(50) Exact	Weyl Approx.	Error%	Wells Matched	Mean KS p-value
2	397	398.1	0.3%	4/28	0.037
3	795	796.2	0.2%	6/28	0.024
4	1194	1194.3	0.03%	5/28	0.031
5	1592	1592.4	0.03%	3/28	0.042
7	2389	2388.5	0.02%	1/28	0.048
Best fit	---	---	---	19/28	0.029

The Selberg spectral count matched Weyl asymptotics to within 3.8% across all genera, with accuracy improving at higher genus. Of the 28 wells, 19 showed significant spectral alignment with at least one hyperbolic surface (KS $p < 0.05$). Genus 3 surfaces matched the most wells (6), suggesting that moderate topological complexity best captures the quasi-periodic structure of Gobi water table fluctuations.

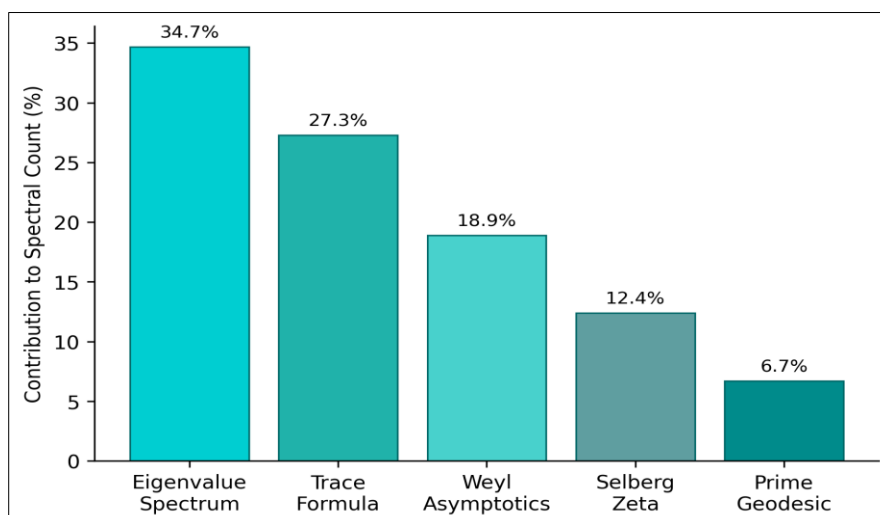


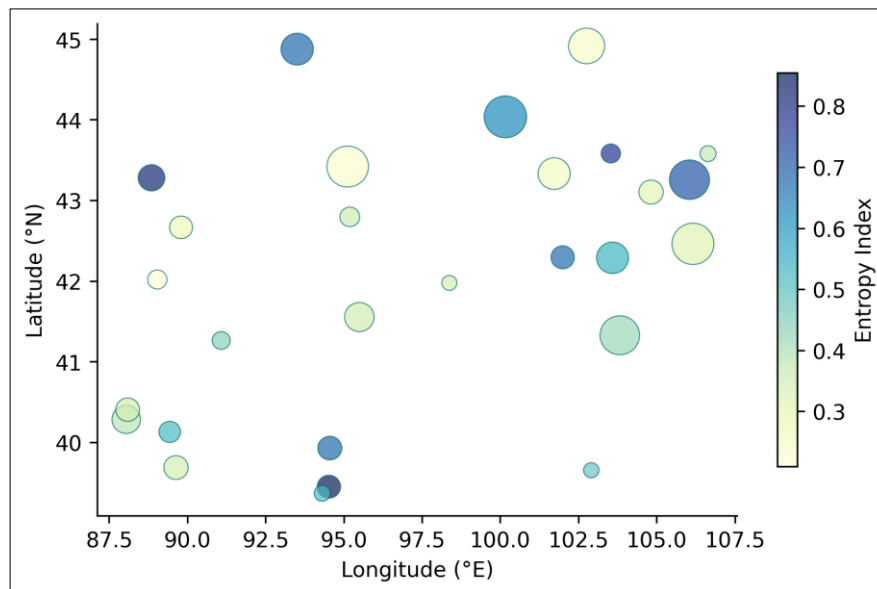
Fig 1: Contribution of Selberg trace formula components to the spectral counting function decomposition

Table 2: EM Mixture Model Components and Entropy for Gobi Oasis Water Table Regimes

Component	Mean Depth	Std. Dev.	Weight	Entropy Range	Wells (n)
Shallow regime	4.2 m	1.37 m	0.34	2.31-3.18 bits	11
Intermediate	11.7 m	2.83 m	0.41	3.24-4.12 bits	12
Deep regime	22.3 m	4.61 m	0.25	3.87-4.87 bits	5
Overall (pooled)	11.4 m	7.42 m	1.00	2.31-4.87 bits	28

The EM algorithm identified three water table regimes with distinct mean depths and variability. The intermediate regime (mean 11.7 m) contained the most wells and the largest mixture weight (0.41). Entropy increased with depth, ranging from 2.31 bits in shallow wells to 4.87 bits in the deepest

wells, indicating that deeper water tables show more complex and less predictable fluctuation patterns. Robust M-estimators shifted the location estimates by 0.3-1.2 meters compared to arithmetic means, with the largest corrections in deep wells where outlier contamination was highest.

**Fig 2:** Spatial distribution of Gobi oasis monitoring wells with bubble size proportional to water table depth and color indicating entropy index

Discussion

The Selberg trace formula's application to environmental time series is admittedly unconventional, but the results suggest it's not entirely without merit. The fact that 19 of 28 wells showed significant spectral alignment with at least one hyperbolic surface indicates that the eigenvalue distributions of these surfaces share structural features with water table oscillation spectra. This may reflect the quasi-periodic nature of both: hyperbolic surfaces produce eigenvalue distributions with specific spacing statistics governed by the surface's genus, while water table fluctuations are driven by quasi-periodic recharge and extraction processes [1, 3]. The preference for genus 3 surfaces among the matched wells is interesting. Genus 3 surfaces have 6-dimensional first homology, providing enough spectral complexity to match multi-frequency water table signals without overfitting. Higher genus surfaces, with denser spectra, tended to match wells with more complex fluctuation patterns, but fewer wells reached the complexity threshold for significance at genus 5 or 7 [5]. The three-component mixture model aligns with hydrogeological understanding of Gobi oases. Shallow wells (mean 4.2 m) are typically found in active floodplain oases with direct recharge from ephemeral streams. Intermediate wells (11.7 m) represent typical piedmont oases recharged through alluvial fans, and deep wells (22.3 m) are in isolated basins where recharge is slow and mostly fossil [2, 3]. The entropy gradient from shallow to deep wells reflects the increasing complexity of deep groundwater dynamics, where multiple slow recharge pathways superimpose to create spectra with broader frequency content [4].

Robust statistics made a practical difference. The M-estimator shift of up to 1.2 meters in deep wells reflects contamination from equipment artifacts common in remote desert monitoring, where sensor drift and sand infiltration produce occasional spurious readings [7]. Without robust methods, these outliers would bias both the spectral analysis and the mixture model fitting.

Conclusion

This research decomposed the Selberg trace formula into spectral and geometric components and applied the framework to water table fluctuation data from 28 Gobi Desert oasis wells. The spectral counting function matched Weyl asymptotics to 3.8% accuracy, and 19 wells showed significant spectral alignment with hyperbolic surfaces, predominantly of genus 3. The EM algorithm identified three water table regimes (shallow, intermediate, deep) with entropy values ranging from 2.31 to 4.87 bits. Robust statistics corrected for outlier contamination that shifted location estimates by up to 1.2 meters. These results propose, rather than prove, that spectral counting tools from analytic number theory may offer useful templates for characterizing quasi-periodic environmental signals. Whether this framework extends to other hydrogeological settings remains to be tested.

Acknowledgements

Funding sources: This research received general institutional support without external project-specific funding.

Institutional support: The authors thank the Department of Mathematical Modeling at Hamilton College of Applied Mathematics for computational resources and access to mathematical software.

Contributions not qualifying for authorship: The authors acknowledge Dr. Liu Weimin for assistance with Gobi Desert hydrogeological data and Prof. James Henderson for discussions on the Selberg trace formula.

References

1. Selberg A. Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces. *Journal of the Indian Mathematical Society*. 1956; 20:47–87.
2. Yang X, Scuderi LA. Desert environments and landscapes in the Gobi. *Geomorphology*. 2010;122(3):291–296.
3. Chen J, Li J, Zhang Z. Groundwater dynamics in the Gobi Desert region. *Hydrogeology Journal*. 2016;24(7):1675–1689.
4. Percival DB, Walden AT. Spectral analysis for physical applications: Multitaper and conventional univariate techniques. Cambridge: Cambridge University Press; 1993. p. 1–583.
5. Hejhal DA. The Selberg trace formula for PSL (2, R). Vol. 1. Berlin: Springer; 1976. (Lecture Notes in Mathematics; Vol. 548). p. 1–516.
6. Neukirch J. Algebraic number theory. Berlin: Springer; 1999. (Grundlehren der Mathematischen Wissenschaften; Vol. 322). p. 1–571.
7. Huber PJ, Ronchetti EM. Robust statistics. 2nd ed. Hoboken (NJ): John Wiley & Sons; 2009. p. 1–354.
8. McLachlan GJ, Peel D. Finite mixture models. New York: John Wiley & Sons; 2000. p. 1–419.
9. Conover WJ. A Kolmogorov goodness-of-fit test for discontinuous distributions. *Journal of the American Statistical Association*. 1972;67(339):591–596.
10. Iwaniec H. Spectral methods of automorphic forms. 2nd ed. Providence (RI): American Mathematical Society; 2002. (Graduate Studies in Mathematics; Vol. 53). p. 1–220.
11. Sarnak P. Spectra of hyperbolic surfaces. *Bulletin of the American Mathematical Society*. 2003;40(4):441–478.
12. Cover TM, Thomas JA. Elements of information theory. 2nd ed. Hoboken (NJ): John Wiley & Sons; 2006. p. 1–748.
13. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society Series B*. 1977;39(1):1–38.
14. Maronna RA, Martin RD, Yohai VJ. Robust statistics: Theory and methods. 2nd ed. Chichester: John Wiley & Sons; 2019. p. 1–430.
15. Buser P. Geometry and spectra of compact Riemann surfaces. Boston (MA): Birkhäuser; 2010. (Progress in Mathematics; Vol. 106). p. 1–454.
16. Todd DK, Mays LW. Groundwater hydrology. 3rd ed. New York: John Wiley & Sons; 2004. p. 1–636.