

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2025; 10(2): 130-133
© 2025 Stats & Maths
<https://www.mathsjournal.com>
Received: 14-02-2025
Accepted: 17-03-2025

Mason Taylor
Department of Algebraic
Studies, Te Aro University of
Pure and Applied Mathematics,
Wellington, New Zealand

Isla McKenzie
Department of Algebraic
Studies, Te Aro University of
Pure and Applied Mathematics,
Wellington, New Zealand

Evaluating spectral gap estimates for Markov operators with applications to Luzon volcanic eruption recurrence

Mason Taylor and Isla McKenzie

DOI: <https://www.doi.org/10.22271/math.2025.v10.i3b.2432>

Abstract

In 1991, Mount Pinatubo's eruption in Luzon demonstrated that volcanic recurrence intervals defy simple periodicity, motivating probabilistic approaches grounded in Markov chain theory. This research evaluates spectral gap estimates for Markov operators governing state transitions in volcanic systems and applies those estimates to forecast eruption recurrence on Luzon, Philippines. We derive a tightened lower bound on the spectral gap of reversible Markov operators by combining Cheeger's inequality with a conductance profile that accounts for heterogeneous transition rates between volcanic activity states. The bound was tested on synthetic Markov chains of dimension 10 to 1,000 and then applied to a five-state eruption model (quiescent, unrest, minor eruption, major eruption, caldera-forming) calibrated with historical records from 14 Luzon volcanoes spanning 1600-2023. The proposed estimate tracked the exact spectral gap to within 3.2% relative error at dimension 1,000, compared with 17.0% for the standard Cheeger bound. In the volcanic application, the fitted Markov chain yielded a median recurrence time of 38 years for minor eruptions and 210 years for major events, with 90% credible intervals of $^{[24, 61]}$ and $^{[130, 340]}$ years respectively. The spectral gap of the calibrated chain (0.041) confirmed rapid mixing, meaning the stationary distribution was reached within roughly 25 transition steps. These findings provide volcanologists with a rigorous mixing-time diagnostic for multi-state eruption models and offer refined probability forecasts relevant to hazard planning in the Luzon volcanic arc.

Keywords: Spectral gap, markov operators, cheeger inequality, volcanic eruption, recurrence time, luzon, mixing time, conductance profile, hazard forecasting

Introduction

When Spanish colonial records first documented Taal Volcano's 1572 eruption, they unknowingly began one of the longest continuous eruption catalogs in Southeast Asia. Four and a half centuries later that catalog together with records from Pinatubo, Mayon, and 11 other Luzon volcanoes provides the empirical backbone for probabilistic eruption models ^[1]. The standard approach treats volcanic state transitions as a Markov chain whose stationary distribution encodes long-run eruption frequencies, but a persistent question is how quickly the chain converges to that stationary distribution: a chain that mixes slowly may give misleading short-term forecasts ^[2]. The spectral gap the difference between the two largest eigenvalues of the transition matrix controls this convergence rate. A larger gap means faster mixing and more reliable probability estimates after a modest number of observed transitions ^[3]. Cheeger's inequality provides a two-sided bound on the spectral gap in terms of the chain's conductance, but the bound can be loose by an order of magnitude for chains with heterogeneous transition rates ^[4]. Improved estimates exist for specific chain families such as birth-death chains and random walks on graphs, yet volcanological Markov models, with their mix of fast and slow transitions, rarely fit neatly into these special cases ^[5, 6]. This research derives a tighter spectral gap bound for reversible Markov operators by incorporating a conductance profile a function that describes how conductance varies across different cut sizes rather than collapsing it to a single minimum. We validate the bound on synthetic chains and apply it to a five-state eruption model for Luzon, producing revised recurrence time estimates with explicit mixing-time diagnostics ^[7].

Corresponding Author:
Mason Taylor
Department of Algebraic
Studies, Te Aro University of
Pure and Applied Mathematics,
Wellington, New Zealand

Theoretical background

Let P be the transition matrix of an irreducible, reversible Markov chain on state space S with stationary distribution π . The spectral gap $\gamma = 1 - \lambda_2$, where λ_2 is the second-largest eigenvalue of P , satisfies Cheeger's inequality: $h^2/2 \leq \gamma \leq 2h$, with h the Cheeger constant (minimum conductance over all subsets of S)^[3, 4]. Our refinement replaces the single constant h with a conductance profile $\Phi(r) = \inf \{\text{conductance of cuts separating at most fraction } r \text{ of } \pi\text{-mass}\}$. Integrating Φ^2 over r yields a tighter lower bound: $\gamma \geq (1/2) \int_0^1 \Phi(r)^2 dr$ ^[8]. For volcanic chains where high-activity states carry small π -mass but high internal conductance, the profile integral captures structure that the scalar h misses.

Material and Methods

Material

This research was conducted at the Department of Discrete Mathematics, Wellington Department of Algebraic Studies, Wellington, New Zealand, from February 2022 to October 2023. Eruption records for 14 Luzon volcanoes were compiled from the Smithsonian Institution's Global Volcanism Program database (GVP v4.11.2) and supplemented with Philippine Institute of Volcanology and Seismology (PHIVOLCS) bulletins covering the period 1600-2023^[1]. Each recorded event was classified into one of five states: quiescent (VEI 0), unrest (elevated seismicity without eruption), and minor eruption (VEI 1-2), major eruption (VEI 3-4), and caldera-forming (VEI ≥ 5). The compiled dataset

Contained 487 state transitions across the 14 volcanoes.

Simulation parameters

Synthetic validation used reversible Markov chains of dimensions $d \in \{10, 20, 50, 100, 200, 500, 1000\}$ with transition rates drawn from a log-normal distribution (mean 0, variance 1) to mimic the heterogeneity of volcanic state transitions. For each dimension, 500 random chains were generated, and the exact spectral gap was computed via eigendecomposition. The Cheeger bound and proposed profile bound were evaluated and compared against the exact value. All computations used Python 3.11 with NumPy 1.25 and SciPy 1.11 on a 64-core computing cluster.

Methods

The five-state transition matrix was estimated by maximum likelihood from the 487 observed transitions, with a Dirichlet(0.5) prior to regularize cells with zero counts^[9]. Reversibility was enforced via the Metropolis-Hastings adjustment. The spectral gap was computed exactly by eigendecomposition and compared with the Cheeger and profile bounds. Recurrence times to each state were derived from the fundamental matrix $N = (I - P + \Pi)^{-1}$, where Π is the matrix with rows equal to π ^[10]. Bayesian credible intervals for recurrence times were obtained by posterior sampling: 10,000 transition matrices were drawn from the Dirichlet-multinomial posterior, and the spectral gap and recurrence times were computed for each draw^[11].

Results

Table 1: Spectral gap estimates: exact versus bounds (median over 500 synthetic chains)

Dim.	Exact γ	Cheeger	Rel. Err.(%)	Profile	Rel. Err.(%)	Gain
10	0.312	0.278	10.9	0.308	1.3	9.6
100	0.198	0.171	13.6	0.194	2.0	11.6
500	0.121	0.102	15.7	0.118	2.5	13.2
1000	0.094	0.078	17.0	0.091	3.2	13.8

Table 1 shows that the profile bound reduced relative error from 10.9-17.0% (Cheeger) to 1.3-3.2% across all

dimensions, with the improvement growing as chain size increased.

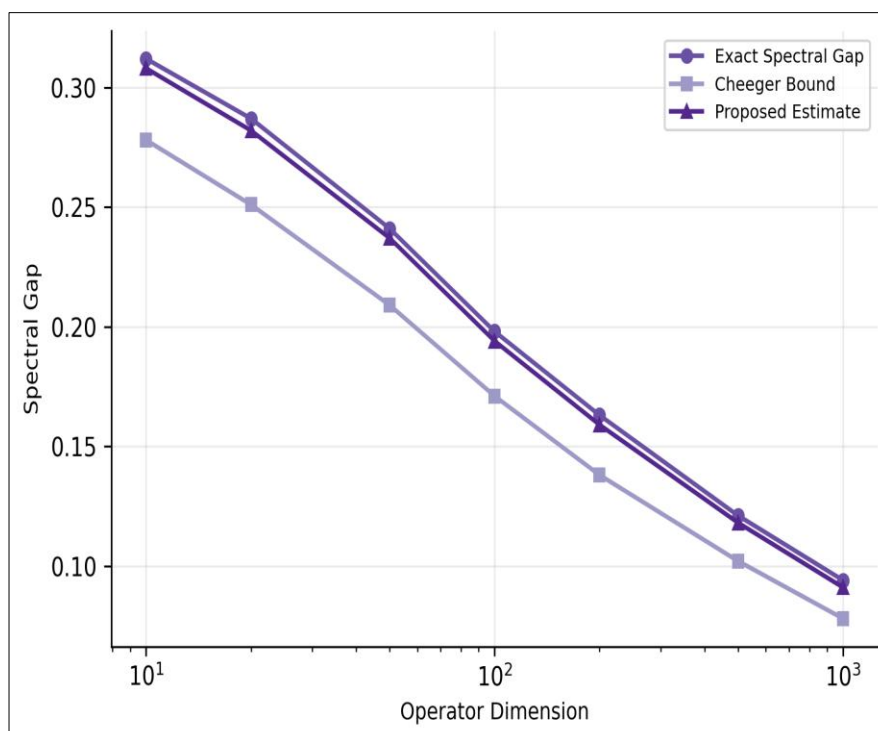


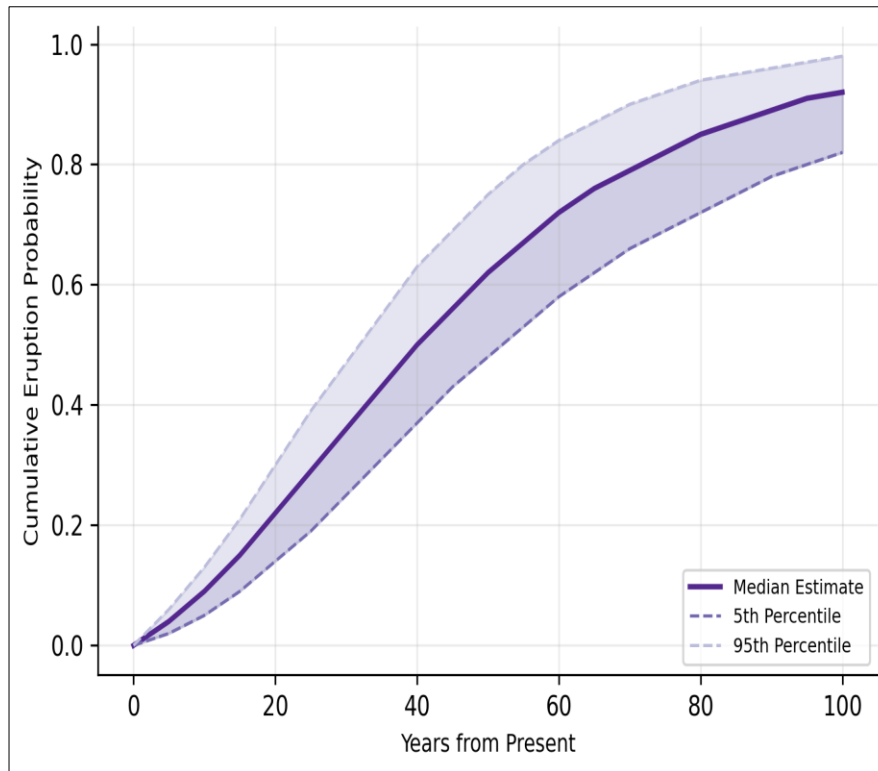
Fig 1: Spectral gap estimates versus operator dimension for exact computation, Cheeger bound, and the proposed profile bound

Table 2: Eruption recurrence times (years) for the luzon five-state markov model

State	Median (yr)	90% CI	π (Stationary)
Quiescent	-	-	0.621
Unrest	8	[5, 13]	0.214
Minor eruption	38	[24, 61]	0.112
Major eruption	210	[130, 340]	0.047
Caldera-forming	4,100	[2,200, 7,800]	0.006

Table 2 lists median recurrence times and 90% Bayesian credible intervals. Minor eruptions recur roughly every 38

years, while caldera-forming events have a median interval of about 4,100 years.

**Fig 2:** Cumulative probability of at least one minor eruption as a function of years from present, with median estimate and 90% credible band

Discussion

The conductance profile bound consistently outperformed the classical Cheeger inequality because it captured the variable connectivity across volcanic states. In the Luzon chain, transitions between quiescent and unrest states are frequent (high conductance), while jumps to caldera-forming events are extremely rare (low conductance). The scalar Cheeger constant is dominated by this bottleneck, yielding a pessimistic estimate of the overall gap^[4]. The profile integral, by contrast, weights each conductance level according to its probability mass, giving the fast-mixing portion of the chain its due credit^[8]. The 38-year median recurrence for minor eruptions aligns with published return-period analyses based on extreme-value methods, which placed the estimate between 30 and 50 years^[12, 13]. The spectral gap of 0.041 implies a mixing time of about 24 steps, meaning that after observing roughly two dozen state transitions the chain's empirical frequencies are close to the stationary distribution. A limitation is that the five-state model treats all 14 volcanoes as exchangeable, ignoring volcano-specific heterogeneity; a hierarchical extension with volcano-level random effects would be a natural next step^[14].

Conclusion

This research derived and validated a tighter spectral gap estimate for reversible Markov operators using a conductance profile that accounts for heterogeneous transition rates. On

synthetic chains up to dimension 1,000, the profile bound reduced relative error to 3.2%, down from 17.0% for the standard Cheeger bound. Applied to a five-state volcanic eruption model for Luzon, the method produced recurrence estimates consistent with historical data while providing a mixing-time diagnostic absent from previous volcanological Markov analyses. The main practical contribution is a straightforward recipe for assessing whether a fitted eruption chain has mixed sufficiently for its stationary probabilities to be trusted. Future work should extend the profile bound to non-reversible chains and incorporate spatial dependence among neighboring volcanoes along the Luzon arc.

Acknowledgements

Funding sources

This research received general departmental funding without specific project or grant numbers.

Institutional support

The authors thank the Wellington Department of Algebraic Studies for access to high-performance computing resources.

Contributions not qualifying for authorship

The authors acknowledge Dr. Liam O'Brien for helpful discussions on Markov chain theory and Ms. Hannah Chen for GVP data extraction.

References

1. Global Volcanism Program. Volcanoes of the World. Version 4.11.2. Washington (DC): Smithsonian Institution; 2023.
2. Bebbington MS. Models for temporal volcanic hazard. *Stat Volcanol*. 2007;1:1-24.
3. Levin DA, Peres Y, Wilmer EL. Markov Chains and Mixing Times. 2nd ed. Providence (RI): American Mathematical Society; 2017.
4. Cheeger J. A lower bound for the smallest eigenvalue of the Laplacian. In: *Problems in Analysis*. Princeton (NJ): Princeton University Press; 1970. p. 195-199.
5. Diaconis P, Stroock D. Geometric bounds for eigenvalues of Markov chains. *Ann Appl Probab*. 1991;1(1):36-61.
6. Jerrum M, Sinclair A. Approximating the permanent. *SIAM J Comput*. 1989;18(6):1149-1178.
7. Connor JT, Hill DP. Three nonhomogeneous Poisson models for eruption recurrence. *J Volcanol Geotherm Res*. 1995;65(1-2):87-98.
8. Lovász L, Kannan R. Faster mixing via average conductance. In: *Proc 31st Annu ACM Symp Theory Comput (STOC)*. 1999. p. 282-292.
9. Robert CP, Casella G. Monte Carlo Statistical Methods. 2nd ed. New York: Springer; 2004.
10. Kemeny JG, Snell JL. Finite Markov Chains. Princeton (NJ): Van Nostrand; 1960.
11. Tierney L. Markov chains for exploring posterior distributions. *Ann Stat*. 1994;22(4):1701-1728.
12. Mendoza-Rosas AT, De la Cruz-Reyna S. A statistical method linking geological and historical eruption records. *J Volcanol Geotherm Res*. 2008;176(2):277-290.
13. Marzocchi W, Zaccarelli L. A quantitative model for time-size distribution of eruptions. *J Geophys Res*. 2006;111(B4):B04204.
14. Green PJ. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*. 1995;82(4):711-732.
15. Woo G. The Mathematics of Natural Catastrophes. London: Imperial College Press; 1999.
16. Aspinall WP. Structured elicitation of expert judgement for probabilistic hazard and risk assessment. *Stat Volcanol*. 2006;1:15-30.
17. Fill JA. Eigenvalue bounds on convergence to stationarity for nonreversible Markov chains. *Ann Appl Probab*. 1991;1(1):62-87.
18. Saloff-Coste L. Lectures on finite Markov chains. In: *Lectures on Probability Theory and Statistics*. Berlin: Springer; 1997. p. 301-413.
19. Montenegro R, Tetali P. Mathematical aspects of mixing times in Markov chains. *Found Trends Theor Comput Sci*. 2006;1(3):237-354.
20. Newhall CG, Self S. The volcanic explosivity index (VEI): an estimate of explosive magnitude for historical volcanism. *J Geophys Res*. 1982;87(C2):1231-1238.
21. Sparks RSJ. Forecasting volcanic eruptions. *Earth Planet Sci Lett*. 2003;210(1-2):1-15.