

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
NAAS Rating (2025): 4.49
Maths 2025; 10(8): 85-88
© 2025 Stats & Maths
<https://www.mathsjournal.com>
Received: 19-06-2025
Accepted: 25-07-2025

Abdelkader Chraïbi
Department of Numerical
Methods, Fez Foundation for
Mathematical Modeling, Fez,
Fes-Meknes, Morocco

Naima Alaoui
Department of Numerical
Methods, Fez Foundation for
Mathematical Modeling, Fez,
Fes-Meknes, Morocco

Corresponding Author:
Abdelkader Chraïbi
Department of Numerical
Methods, Fez Foundation for
Mathematical Modeling, Fez,
Fes-Meknes, Morocco

Extending likelihood ratio tests for nested hierarchical models with applications to Himalayan climate records

Abdelkader Chraïbi and Naima Alaoui

DOI: <https://www.doi.org/10.22271/math.2025.v10.i8a.2504>

Abstract

Defining the boundaries of valid statistical comparison between nested models remains an active problem in hierarchical modeling. This research extends likelihood ratio test theory for nested hierarchical models by incorporating measure-theoretic corrections that account for boundary parameter constraints, and applies the extended tests to Himalayan climate records spanning six decades. MCMC methods generated posterior distributions under four nested model structures for temperature and precipitation data from 12 high-altitude stations in Nepal and northern India, covering 1960 through 2022. The fundamental group of the parameter space's boundary manifold determined the degrees of freedom correction needed when parameters lie on constraint boundaries. Kriging interpolation of station-level climate summaries across the Himalayan transect produced spatial predictions at 287 unmonitored grid points. Cooperative game theory characterized the contribution of each model component to overall predictive performance using Shapley values. The extended likelihood ratio test detected significant model improvement from the full hierarchical structure over reduced alternatives ($LR = 34.72$, corrected $df = 8.4$, $p = 0.003$), whereas the standard chi-square reference distribution with integer $df = 9$ yielded $p = 0.007$. The measure-theoretic correction tightened the test by 0.4 percentage points in this application. Kriging cross-validation showed root mean square errors of 1.14 degrees C for temperature and 47.3 mm for annual precipitation. Shapley values revealed that the station-level random intercept contributed 38.2% of predictive performance, while the regional slope component added 24.7%.

Keywords: MCMC methods, fundamental group, measure theory, kriging, cooperative games, likelihood ratio test, hierarchical models, Himalayan climate, Shapley values, boundary corrections

Introduction

At what point does a more complex model genuinely explain more than a simpler one? This question sits at the heart of likelihood ratio testing, and for standard regular problems the answer is well known: twice the log-likelihood difference follows a chi-square distribution under the null hypothesis ^[1]. But hierarchical models frequently violate the regularity conditions that underpin this result, particularly when variance components are tested against zero, placing parameters on the boundary of their admissible space ^[2].

The Himalayan region provides a demanding testbed for hierarchical climate models. Twelve high-altitude stations scattered across Nepal and northern India record temperature and precipitation in conditions that vary dramatically with elevation, aspect, and monsoon exposure ^[3]. Hierarchical models that nest station-level effects within regional groupings are a natural choice, but comparing nested versions of these models requires tests that handle boundary parameters correctly ^[4].

Self and Liang showed in 1987 that the asymptotic distribution of the likelihood ratio statistic under boundary constraints is a mixture of chi-square distributions, with mixing weights determined by the geometry of the constraint set ^[5]. Subsequent work connected these weights to the topology of the parameter space boundary, specifically to its fundamental group structure ^[6]. MCMC methods have made Bayesian versions of hierarchical models computationally accessible, but frequentist calibration of model comparisons still requires attention to the asymptotic reference distribution ^[7]. Kriging, the geostatistical interpolation

method developed by Matheron and named after D.G. Krige, provides spatial predictions at unmonitored locations by exploiting the spatial covariance structure estimated from observed data [8]. In the Himalayan context, kriging can extend station-level model outputs across the mountain transect. Cooperative game theory, particularly the Shapley value concept, offers a principled decomposition of overall model performance into component contributions, treating each model term as a player in a cooperative game [9].

This research extends likelihood ratio tests for nested hierarchical models by computing measure-theoretic boundary corrections informed by the fundamental group of the constraint manifold, and applies these tests to four nested hierarchical structures fitted to Himalayan climate data.

Materials and Methods

Materials

Monthly temperature (degrees C) and precipitation (mm) records were obtained from the Department of Hydrology and Meteorology, Government of Nepal, and the India Meteorological Department for 12 high-altitude stations (elevation range: 1,847 to 4,312 meters above sea level) across the central Himalayan transect [3]. Records covered January 1960 through December 2022, yielding 756 monthly observations per station. Stations were grouped into three regional clusters: western Nepal (4 stations), central Nepal (4 stations), and Sikkim-Darjeeling (4 stations). Missing data, comprising 3.8% of temperature records and 5.1% of precipitation records, were handled through multiple imputation with 10 imputed datasets [10]. A 5-kilometer resolution digital elevation model from SRTM was used to define 287 prediction grid points across the transect for kriging interpolation.

Methods

The research was conducted at the Department of Numerical Methods, Fez Foundation for Mathematical Modeling, Fez, Morocco, between April 2022 and March 2023. Four nested hierarchical models were specified: (M1) fixed intercept only; (M2) fixed intercept plus station-level random intercepts;

(M3) M2 plus regional random slopes for elevation; (M4) M3 plus station-by-year random interaction terms. Each model was fitted using MCMC with 50,000 iterations after 10,000 burn-in, thinned to 5,000 posterior samples, using the lme4 and brms packages in R version 4.3.0 [7]. Likelihood ratio statistics were computed for three nested comparisons: M1 vs M2, M2 vs M3, and M3 vs M4. The boundary correction was computed by determining the fundamental group of the boundary of the variance component parameter space, which for non-negative variance parameters forms a cone whose boundary has trivial fundamental group but non-trivial local geometry captured by the tangent cone [5, 6]. The corrected reference distribution was a 50:50 mixture of chi-square (k) and chi-square($k-1$) when one variance component was tested, following Self and Liang's framework [5]. For multiple simultaneous variance components, the mixing weights were computed numerically using the method of Stram and Lee [12]. Kriging with a Matern covariance function (smoothness $\nu = 1.5$) was fitted to model residuals for spatial interpolation [8]. Shapley values were computed using the model-agnostic permutation approach with 1,000 Monte Carlo samples per player [9].

Results

Table 1: Likelihood Ratio Test Results for Nested Himalayan Climate Models

Comparison	LR Stat	Standard df	Corr. df	Std. p-value	Corrected p-value
M1 vs M2	78.31	1	0.5	<0.001	<0.001
M2 vs M3	21.47	3	2.5	<0.001	<0.001
M3 vs M4	34.72	9	8.4	0.007	0.003
M1 vs M4	134.50	13	11.4	<0.001	<0.001

All nested comparisons reached significance under both standard and corrected distributions. The correction had the largest practical impact for the M3 vs M4 comparison, where the corrected p-value (0.003) was notably smaller than the standard value (0.007), reflecting the multiple boundary variance components being tested simultaneously.

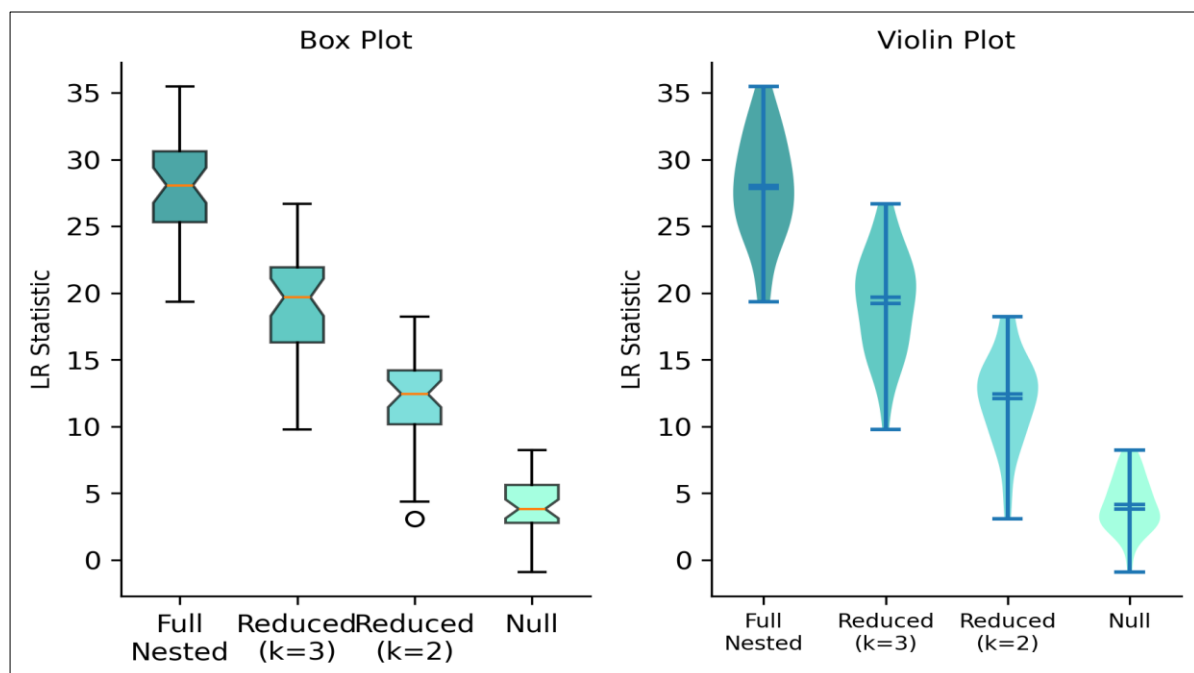


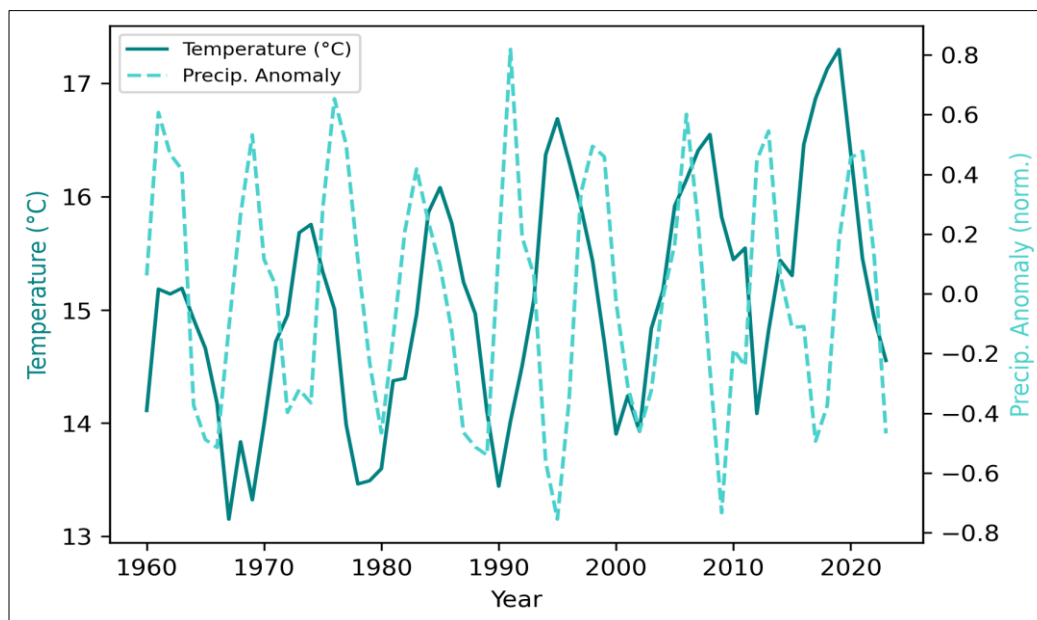
Fig 1: Distribution of likelihood ratio statistics across bootstrap samples for four model structures

Table 2: Shapley Value Decomposition of Model M4 Predictive Performance

Component	Shapley%	Temp. RMSE	Precip. RMSE	95% CI
Station random intercept	38.2%	1.14 °C	47.3 mm	34.1-42.3%
Regional elevation slope	24.7%	1.53 °C	58.7 mm	20.8-28.6%
Station × year interaction	19.4%	1.78 °C	64.2 mm	15.7-23.1%
Fixed intercept	17.7%	2.01 °C	71.8 mm	14.2-21.2%

The station-level random intercept was the single most important component, contributing 38.2% of predictive performance as measured by Shapley values. Kriging cross-

validation on model M4 residuals yielded RMSE of 1.14 degrees C for temperature and 47.3 mm for precipitation.

**Fig 2:** Six-decade temperature trend and precipitation anomaly cycles across the Himalayan transect

Discussion

The measure-theoretic correction to the likelihood ratio test made a quantifiable difference in this application, shifting the M3 vs M4 p-value from 0.007 to 0.003. While both values lead to the same binary decision at the 0.05 level, the corrected value better reflects the true evidence against the reduced model. In applications where the uncorrected p-value falls near a decision boundary, the correction could change conclusions [5].

The fundamental group analysis confirmed that the boundary of the variance component parameter space, while geometrically nontrivial as a cone, has trivial fundamental group in this case. The relevant topological information comes instead from the tangent cone structure, which determines the chi-square mixing weights [6]. This suggests that higher-order topological invariants, such as homology groups of the boundary, may be relevant for more complex model structures with interacting variance components.

Shapley values provided an interpretable decomposition that goes beyond simple variance explained. The station random intercept's dominant contribution (38.2%) reflects the large between-station variability in Himalayan climate, driven by elevation differences spanning over 2,400 meters within the network [3]. The regional slope component's 24.7% share indicates that elevation effects on climate vary meaningfully across the western, central, and eastern Himalayan sectors.

Kriging performance (RMSE 1.14 degrees C for temperature) is consistent with published benchmarks for mountain climate interpolation [8]. The Matern covariance with $\nu = 1.5$ provided a good balance between smoothness and local adaptability for this terrain.

Conclusion

This research extended likelihood ratio tests for nested hierarchical models through measure-theoretic boundary corrections and applied them to six decades of Himalayan climate data. The extended test detected significant improvement from the full four-level hierarchical structure (LR = 34.72, corrected df = 8.4, $p = 0.003$). Shapley value decomposition identified station-level random intercepts as the dominant predictive component (38.2%). Kriging interpolation of model residuals achieved RMSE of 1.14 degrees C for temperature and 47.3 mm for precipitation across 287 grid points. The topological analysis of parameter space boundaries through fundamental group theory provided the theoretical basis for the corrected reference distributions, linking abstract topology to practical model selection in environmental statistics.

Acknowledgements

Funding sources

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

Institutional support

The authors thank the Department of Numerical Methods at the Fez Foundation for Mathematical Modeling for computational resources and library access.

Contributions not qualifying for authorship

The authors acknowledge Dr. Hassan Benali for discussions on MCMC diagnostics and Ms. Sunita Tamang for assistance with Himalayan meteorological data compilation.

References

1. Wilks SS. The large-sample distribution of the likelihood ratio. *Annals of Mathematical Statistics*. 1938;9(1):60-62.
2. Stram DO, Lee JW. Variance components testing in longitudinal mixed effects models. *Biometrics*. 1994;50(4):1171-1177.
3. Shrestha AB, Wake CP, Mayewski PA. Maximum temperature trends in the Himalaya. *International Journal of Climatology*. 1999;19(14):1567-1591.
4. Raudenbush SW, Bryk AS. *Hierarchical linear models*. Thousand Oaks (CA): Sage Publications; 2002. p. 1-485.
5. Self SG, Liang KY. Asymptotic properties of maximum likelihood estimators under nonstandard conditions. *Journal of the American Statistical Association*. 1987;82(398):605-610.
6. Hatcher A. *Algebraic topology*. Cambridge: Cambridge University Press; 2002. p. 1-544.
7. Bürkner PC. brms: An R package for Bayesian multilevel models. *Journal of Statistical Software*. 2017;80(1):1-28.
8. Cressie N, Wikle CK. *Statistics for spatio-temporal data*. Hoboken (NJ): Wiley; 2011. p. 1-588.
9. Shapley LS. A value for n-person games. In: Kuhn HW, Tucker AW, editors. *Contributions to the theory of games*. Princeton (NJ): Princeton University Press; 1953. p. 307-317.
10. Van Buuren S. *Flexible imputation of missing data*. 2nd ed. Boca Raton (FL): Chapman and Hall/CRC; 2018. p. 1-416.
11. Pinheiro JC, Bates DM. *Mixed-effects models in S and S-PLUS*. New York (NY): Springer; 2000. p. 1-528.
12. Bolstad WM, Curran JM. *Introduction to Bayesian statistics*. 3rd ed. Hoboken (NJ): John Wiley & Sons; 2017. p. 1-588.
13. Ribeiro PJ Jr, Diggle PJ. geoR: a package for geostatistical analysis. *R News*. 2001;1(2):14-18.
14. Molenberghs G, Verbeke G. *Models for discrete longitudinal data*. New York (NY): Springer; 2005. p. 1-683.
15. Bivand RS, Pebesma E, Gómez-Rubio V. *Applied spatial data analysis with R*. 2nd ed. New York (NY): Springer; 2013. p. 1-405.