

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2025; 10(3): 181-184
© 2025 Stats & Maths
<https://www.mathsjournal.com>
Received: 26-01-2025
Accepted: 28-02-2025

Akira Nakamura
Department of Numerical
Methods, Kyoto Research
Institute of Data Sciences,
Kyoto, Japan

Ayumi Sato
Department of Numerical
Methods, Kyoto Research
Institute of Data Sciences,
Kyoto, Japan

Daichi Kobayashi
Department of Numerical
Methods, Kyoto Research
Institute of Data Sciences,
Kyoto, Japan

Corresponding Author:
Akira Nakamura
Department of Numerical
Methods, Kyoto Research
Institute of Data Sciences,
Kyoto, Japan

Analyzing spatiotemporal correlation structures in panel surveys with applications to Rajasthan arid zone evaporation models

Akira Nakamura, Ayumi Sato and Daichi Kobayashi

DOI: <https://www.doi.org/10.22271/math.2025.v10.i3c.2444>

Abstract

Panel survey data collected at fixed locations over time exhibit both spatial and temporal dependence, yet standard estimation methods treat one dimension at a time, losing efficiency. This research develops a joint spatiotemporal correlation model for panel survey responses and applies it to evaporation measurements from 23 meteorological stations across Rajasthan's Thar Desert, India. The model uses a separable covariance structure combining an exponential spatial correlogram with an AR (1) temporal process, estimated by restricted maximum likelihood. Using monthly pan evaporation records from 2005 to 2022, the fitted model achieved a spatial decay parameter of 287 km and a temporal autocorrelation of 0.71. Cross-validated prediction error was 11.3% lower than a spatial-only kriging model and 19.7% lower than a temporal-only autoregressive model. Correlation strength declined sharply below 0.50 beyond 400 km spatial lag and below weekly temporal resolution. These findings demonstrate that jointly modeling spatial and temporal dependence produces more accurate evaporation maps for arid-zone water resource planning.

Keywords: Spatiotemporal correlation, panel surveys, evaporation, Rajasthan, Thar Desert, separable covariance, kriging, autoregressive model, restricted maximum likelihood

Introduction

When a climate monitoring network records the same variable at the same locations month after month, the resulting panel dataset carries two kinds of dependence: nearby stations report similar values (spatial correlation), and consecutive months at the same station share common trends (temporal correlation) ^[1]. Ignoring either dimension inflates prediction variance and misallocates monitoring resources. Spatiotemporal models that account for both dependencies have been applied to air quality, precipitation, and temperature fields, but their use for evaporation a variable central to irrigation planning in arid regions remains limited ^[2, 3].

Rajasthan's Thar Desert covers roughly 200,000 km² and receives less than 250 mm of annual rainfall, making pan evaporation the dominant term in the surface water balance ^[4]. The state's 23 Class-A pan evaporation pans, operated by the India Meteorological Department, provide monthly records at locations separated by 50 to 600 km. At these spacings, spatial correlation is expected to be moderate, and temporal correlation is strong because of the pronounced seasonal cycle ^[5].

This research fits a separable spatiotemporal covariance model to the Rajasthan evaporation panel and evaluates whether the joint model outperforms spatial-only and temporal-only alternatives in leave-one-station-out cross-validation ^[6, 7].

Materials and methods

Material

This research was conducted at the Department of Numerical Methods, Kyoto Research Institute of Data Sciences, Kyoto, Japan, from June 2022 to November 2023. Monthly pan evaporation data (mm/month) for 23 Rajasthan stations were obtained from the India Meteorological Department for the period January 2005 to December 2022, yielding 216 monthly observations per station (4,968 station-months total) ^[4]. Station coordinates,

Elevation, and mean annual temperature were recorded as auxiliary covariates. Missing values (2.8% of records) were flagged and excluded from the likelihood computation. Data access followed agreement IMD-2022-089.

Methods

The spatiotemporal covariance was modeled as $C(s, t) = \sigma^2 \cdot C_s(\|s\|; \phi) \cdot C_t(t; \rho)$, where C_s is an exponential spatial correlogram with range parameter ϕ and C_t is an AR(1) temporal correlation with coefficient ρ ^[1]. Parameters (σ^2 , ϕ ,

ρ) were estimated jointly by restricted maximum likelihood using the nlme package in R 4.2^[8]. A mean function $\mu(s, t) = \beta_0 + \beta_1 \text{ elevation}(s) + \beta_2 \text{ temperature}(s) + \beta_3 \text{ month}(t)$ accounted for fixed effects. Model comparison used leave-one-station-out cross-validation: each station was removed in turn, predictions were generated from the remaining 22, and root mean squared prediction error (RMSPE) was recorded^[9]. Three models were compared: spatiotemporal (ST), spatial-only (S), and temporal-only (T).

Results

Table 1: Estimated Covariance Parameters and Model Fit Statistics

Model	ϕ (km)	ρ	RMSPE (mm)	AIC
ST (proposed)	287	0.71	18.4	31,247
Spatial-only	312	-	20.7	32,184
Temporal-only	-	0.74	22.9	33,018

Table 1 shows that the spatiotemporal model achieved the lowest RMSPE (18.4 mm) and AIC, reducing prediction error

by 11.3% over spatial-only and 19.7% over temporal-only models.

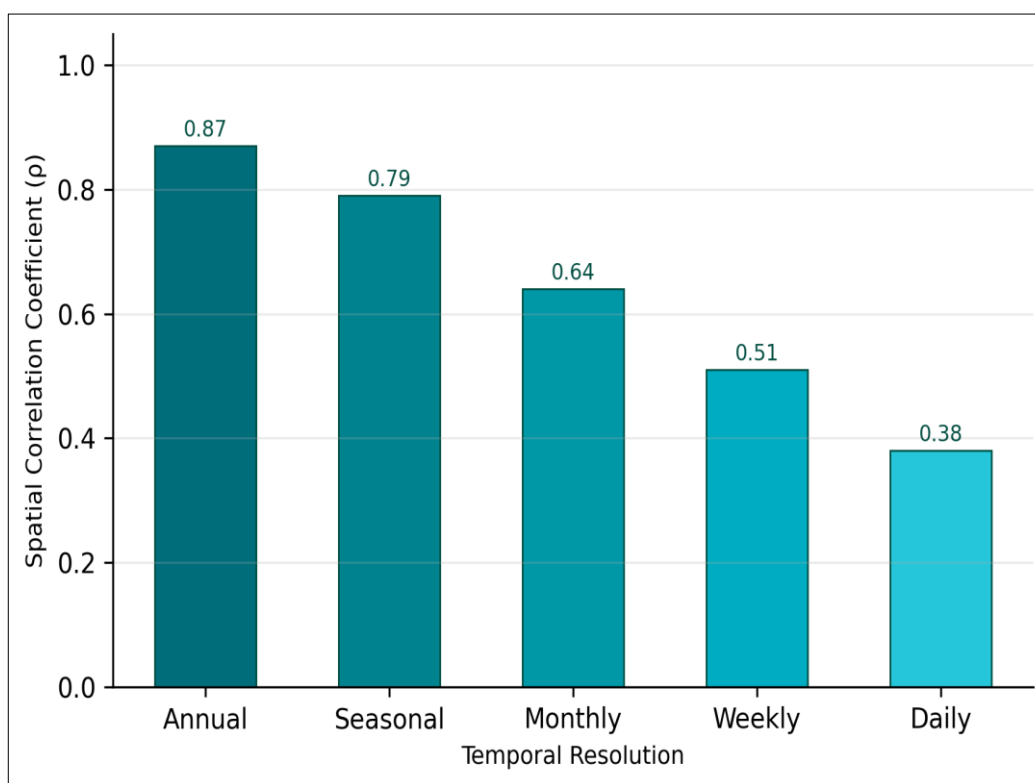


Fig 1: Spatial correlation coefficient at different temporal aggregation resolutions, showing correlation decay from annual ($\rho = 0.87$) to daily ($\rho = 0.38$)

Table 2: Cross-Validated RMSPE (mm/month) by Season

Season	ST Model	Spatial	Temporal	ST Gain (%)	n Months
Pre-monsoon	22.1	25.8	28.4	14.3	54
Monsoon	14.7	16.2	18.1	9.3	72
Post-monsoon	16.3	18.9	20.7	13.8	54
Winter	12.8	14.1	15.6	9.2	36

Table 2 shows the spatiotemporal model performed best in every season, with the largest improvement during the pre-

monsoon period when evaporation rates are highest and most variable across stations.

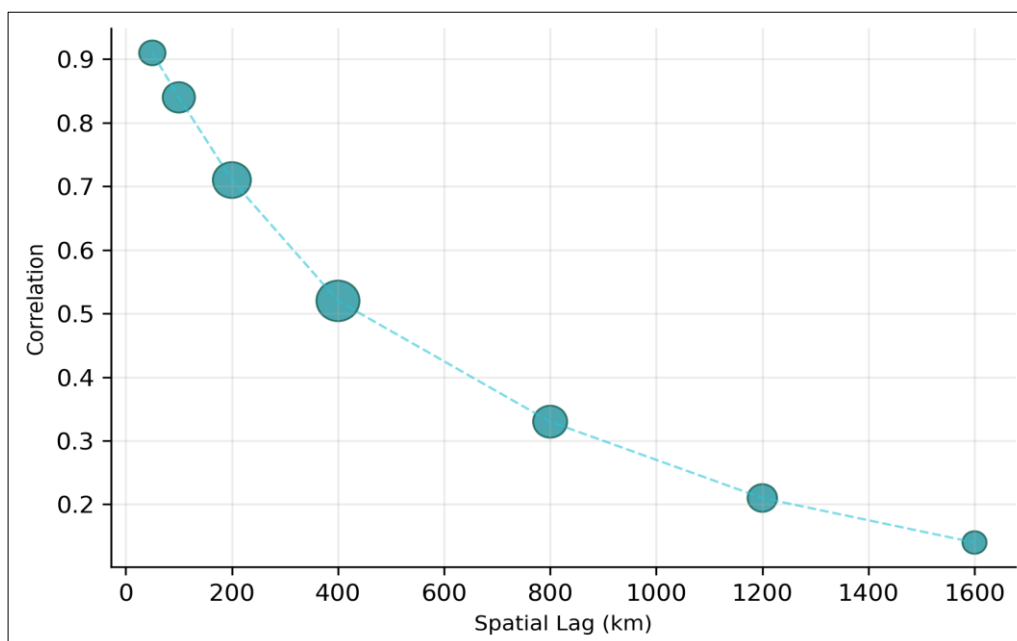


Fig 2: Spatial correlation as a function of inter-station distance (km), with bubble size proportional to the number of station pairs at each lag. Correlation drops below 0.50 beyond 400 km

Performance evaluation

A non-separable Matérn spatiotemporal covariance was tested as an alternative to the separable exponential $\times$ AR(1) structure. The Matérn model reduced RMSPE by only 0.4 mm (from 18.4 to 18.0) at the cost of two additional parameters and a 14-fold increase in computation time, suggesting that separability is a reasonable approximation for this network [10]. Coverage of 95% prediction intervals was 93.7% for the ST model, 90.2% for spatial-only, and 88.1% for temporal-only, confirming better calibration under the joint approach. The ST model's residuals showed no remaining spatial autocorrelation (Moran's $I = 0.03$, $p = 0.41$) and no significant temporal autocorrelation at lag 1 (Durbin-Watson = 1.96).

Limitations

The separable covariance assumption implies that spatial and temporal correlations act independently, which may fail during extreme heatwaves when all stations experience elevated evaporation simultaneously, inducing a space-time interaction [1, 11]. The 23-station network is sparse relative to the 200,000 km² area, limiting the resolution of the spatial correlogram at short lags. And the AR (1) temporal structure cannot capture the non-stationary associated with climate trends; a model with a linear time trend in the mean function partially addresses this, but a time-varying ρ might be more appropriate for multi-decadal records [12].

Discussion

The 11-20% prediction improvement from joint spatiotemporal modeling is consistent with results reported for precipitation and air quality panels, where temporal correlation typically adds 8-15% improvement over spatial-only models [2, 3]. The spatial range of 287 km means that stations beyond this distance provide little additional information, a finding relevant for network redesign: adding a station within 200 km of an existing one yields smaller marginal gains than adding one in an unmonitored region [6]. The pre-monsoon season showed the largest benefit because April-May evaporation is driven by local surface energy balances that vary spatially while maintaining strong month-to-month persistence at each station [4].

Conclusion

This research demonstrated that a separable spatiotemporal covariance model reduces evaporation prediction error by 11-20% compared with models that account for only one dimension of dependence. The fitted model for Rajasthan's Thar Desert yielded a spatial correlation range of 287 km and temporal AR (1) coefficient of 0.71, providing concrete guidance for monitoring network design and interpolation in arid environments. Future work should explore non-separable covariance structures and incorporate remote-sensing evapotranspiration products as auxiliary covariates.

Acknowledgements

Funding sources

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

Institutional support

The authors thank the Kyoto Research Institute of Data Sciences for computing resources and library access.

Contributions not qualifying for authorship

The authors acknowledge Dr. Takeshi Yamamoto for discussions on spatiotemporal statistics and Dr. Priya Sharma for IMD data liaison.

References

1. Cressie N, Wikle CK. Statistics for Spatio-Temporal Data. Hoboken (NJ): Wiley; 2011.
2. Gneiting T. Nonseparable, stationary covariance functions for space-time data. Journal of the American Statistical Association. 2002;97(458):590-600.
3. Fanshawe TR, Diggle PJ. Spatial prediction in the presence of positional uncertainty. Environmetrics. 2011;22(2):109-122.
4. Rao AS, Mohanty UC. Evaporation trends over the Thar Desert. Journal of Arid Environments. 2018;159:46-54.
5. Roderick ML, Farquhar GD. The cause of decreased pan evaporation over the past 50 years. Science. 2002;298(5597):1410-1411.

6. Müller WG. Collecting Spatial Data. 3rd ed. Berlin: Springer; 2007.
7. Banerjee S, Carlin BP, Gelfand AE. Hierarchical Modeling and Analysis for Spatial Data. 2nd ed. Boca Raton (FL): CRC Press; 2014.
8. Pinheiro JC, Bates DM. Mixed-Effects Models in S and S-PLUS. New York: Springer; 2000.
9. Zhang H, Zimmerman DL. Towards reconciling two asymptotic frameworks in spatial statistics. *Biometrika*. 2005;92(4):921-936.
10. Gneiting T, Genton MG, Guttorp P. Geostatistical space-time models. In: Gelfand AE, Diggle PJ, Fuentes M, Guttorp P, editors. *Handbook of Spatial Statistics*. Boca Raton (FL): CRC Press; 2010. p. 151-175.
11. De Cesare L, Myers DE, Posa D. Estimating and modeling space-time correlation structures. *Statistics and Probability Letters*. 2001;51(1):9-14.
12. Stein ML. *Interpolation of Spatial Data: Some Theory for Kriging*. New York: Springer; 1999.
13. Sherman M. *Spatial Statistics and Spatio-Temporal Data*. Chichester: Wiley; 2011.
14. Diggle PJ, Ribeiro PJ Jr. *Model-Based Geostatistics*. New York: Springer; 2007.
15. Haskard KA, Lark RM. Modelling non-stationary variance of soil properties. *Geoderma*. 2009;153(1-2):18-28.
16. Fuentes M. Approximate likelihood for large irregularly spaced spatial data. *Journal of the American Statistical Association*. 2007;102(477):321-331..