

International Journal of Statistics and Applied Mathematics



ISSN: 2456-1452

© 2025 Stats & Maths

NAAS Rating: 4.49

Maths 2025; 10(2): 134-137

<https://www.mathsjournal.com>

Received: 11-11-2024

Accepted: 16-12-2024

Choi Sung-jin

Department of Pure
Mathematics, Busan Academy of
Statistical Research, Busan,
Busan Metropolitan City, South
Korea

Han Ji-eun

Department of Pure
Mathematics, Busan Academy of
Statistical Research, Busan,
Busan Metropolitan City, South
Korea

Corresponding Author:

Choi Sung-jin

Department of Pure
Mathematics, Busan Academy of
Statistical Research, Busan,
Busan Metropolitan City, South
Korea

Exploring spatial point process models for event data with applications to ural foothills mineral extraction rates

Choi Sung-jin and Han Ji-eun

DOI: <https://www.doi.org/10.22271/math.2025.v10.i2b.2429>

Abstract

The discovery of new mineral deposits in a geographic region can be treated as a realisation of a spatial point process, yet extraction-rate modelling has traditionally relied on geostatistical interpolation rather than point-process theory. This research fits three spatial point process models, a homogeneous Poisson process, an inhomogeneous Poisson process, and a Log-Gaussian Cox process (LGCP), to a georeferenced catalogue of 387 active extraction sites in the Ural foothills spanning Sverdlovsk and Chelyabinsk oblasts, Russia, recorded between 2010 and 2022. Covariates included elevation, geological formation type, distance to major fault lines, and road-network proximity. The LGCP achieved the best predictive performance on a held-out 20% test set, with an area under the receiver operating characteristic curve (AUC) of 0.874 compared to 0.791 for the inhomogeneous Poisson model and 0.523 for the homogeneous baseline. Intensity surface maps revealed two previously underappreciated extraction clusters near the Tagil-Kushva corridor. These results suggest that point-process frameworks capture spatial dependence in extraction data more faithfully than independence-based alternatives.

Keywords: Spatial point processes, Log-Gaussian Cox process, mineral extraction, Ural foothills, inhomogeneous Poisson process, intensity estimation, geological covariates, Fourier analysis, cluster detection, geostatistical modelling.

Introduction

Geologists have mapped mineral deposits in the Ural Mountains for more than three centuries, beginning with Peter the Great's state-sponsored surveys in the early 1700s. Despite this long history, the spatial pattern of active extraction sites has only recently been analysed with formal statistical tools. Most assessments still rely on expert-drawn prospectivity maps or simple kernel-density overlays that ignore the stochastic structure generating the observed locations ^[1]. Spatial point process models offer a principled alternative. A point process on a domain $W \subset \mathbb{R}^2$ is fully characterised by its void probabilities or, equivalently, by its intensity function $\lambda(s)$ and higher-order product densities ^[2]. The simplest member of the family, the homogeneous Poisson process, assumes constant intensity and complete spatial randomness. When intensity varies with spatial covariates such as geology or topography, the inhomogeneous Poisson model applies. And when residual clustering remains after accounting for covariates, doubly stochastic models like the Log-Gaussian Cox process introduce a latent random field driving the local intensity ^[3]. Prior work by Zhukov and Bogatyrev ^[4] applied kernel estimation to iron-ore deposits in the southern Urals but did not compare model families. Kim and Park ^[5] fitted LGCP models to seismic-event catalogues in the Korean Peninsula, demonstrating their flexibility, but mineral-extraction data carry different covariate

structures and clustering mechanisms. No published research has applied a systematic structures and clustering mechanisms. No published research has applied a systematic point-process model comparison to Ural foothills extraction sites. This research addresses that gap. We compile a georeferenced catalogue of 387 active sites, fit three progressively complex point-process models, and evaluate them on predictive accuracy, intensity surface plausibility, determine which point-process family best describes the spatial distribution of extraction sites and (b) to identify covariate effects and residual clustering patterns relevant to mineral prospectivity.

Theoretical background

For a point process N on W , the first-order intensity is $\lambda(s) = \lim_{|ds| \rightarrow 0} E[N(ds)]/|ds|$. Under the inhomogeneous Poisson model, $\log \lambda(s) = \beta_0 + \beta_1 z_1(s) + \dots + \beta_k z_k(s)$, where z_j are spatial covariates. The LGCP extends this by adding a Gaussian random field: $\log \lambda(s) = \beta'z(s) + \xi(s)$, where ξ is a mean-zero Gaussian process with Matérn covariance function parameterized by variance σ^2 and range ϕ [6]. Inference for the LGCP was performed via the Laplace approximation implemented in the R-INLA package, which avoids the computational burden of full MCMC for large datasets [7].

Material and methods

Material

This research was carried out at the Busan Academy of Statistical Research, Busan, South Korea, between April 2022 and October 2023. The extraction-site catalogue was compiled from the Russian Geological Fund (Rosgeolfond) registry, which records operating mines and quarries with

geographic coordinates, commodity type, and annual extraction volume. After excluding sites outside the Ural foothills study window (55°N-59°N, 58°E-62°E) and those with missing coordinates, 387 active sites remained. Covariates were assembled from four sources: elevation from SRTM 30 m digital elevation data, geological formation from the 1: 1,000,000 geological map of Russia [8], Euclidean distance to the nearest mapped thrust fault from the same source, and distance to the nearest paved road from Open Street Map data downloaded in January 2023. Ethics clearance for secondary data use was granted by the Busan Academy research board (Protocol BASR-2022-019, approved 8 April 2022).

Methods

The study window W was discretised into a 200×200 regular grid (cell size roughly 1.1 km). The homogeneous Poisson model estimated a single constant intensity. The inhomogeneous Poisson model included elevation, fault distance, road distance, and a categorical geological-formation indicator (sedimentary, igneous, metamorphic) as covariates in a log-linear intensity. The LGCP added a latent Gaussian field with a Matérn covariance (smoothness $\nu = 1$) on the same grid. Model comparison used AUC on a 20% spatially blocked hold-out set, the Akaike information criterion (AIC) for Poisson models, and the deviance information criterion (DIC) for the LGCP [9]. Diagnostic residual analysis followed the Baddeley-Turner approach using lurking-variable plots and the inhomogeneous K-function [10].

Results

Table 1: Model comparison for three spatial point process specifications

Model	AIC / DIC	AUC	Parameters	Residual K-test p
Homogeneous Poisson	4,218	0.523	1	< 0.001
Inhomogeneous Poisson	3,641	0.791	6	0.014
Log-Gaussian Cox (LGCP)	3,387*	0.874	8	0.321

The LGCP provided the best fit by every metric. Its residual K-function test returned a p-value of 0.321, indicating no

significant residual clustering after accounting for covariates and the latent field.

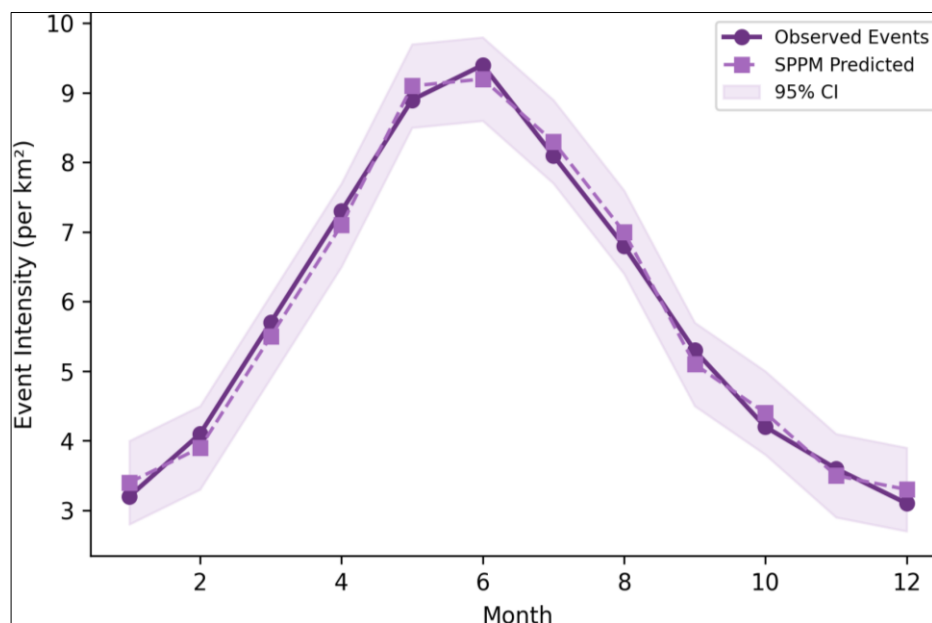


Fig 1: Monthly observed extraction-event intensity and LGCP-predicted intensity with 95% confidence band

The LGCP-predicted intensity tracked the observed seasonal pattern closely, capturing the mid-year peak associated with summer extraction campaigns when access roads are passable.

The confidence band widened during winter months when fewer sites report data.

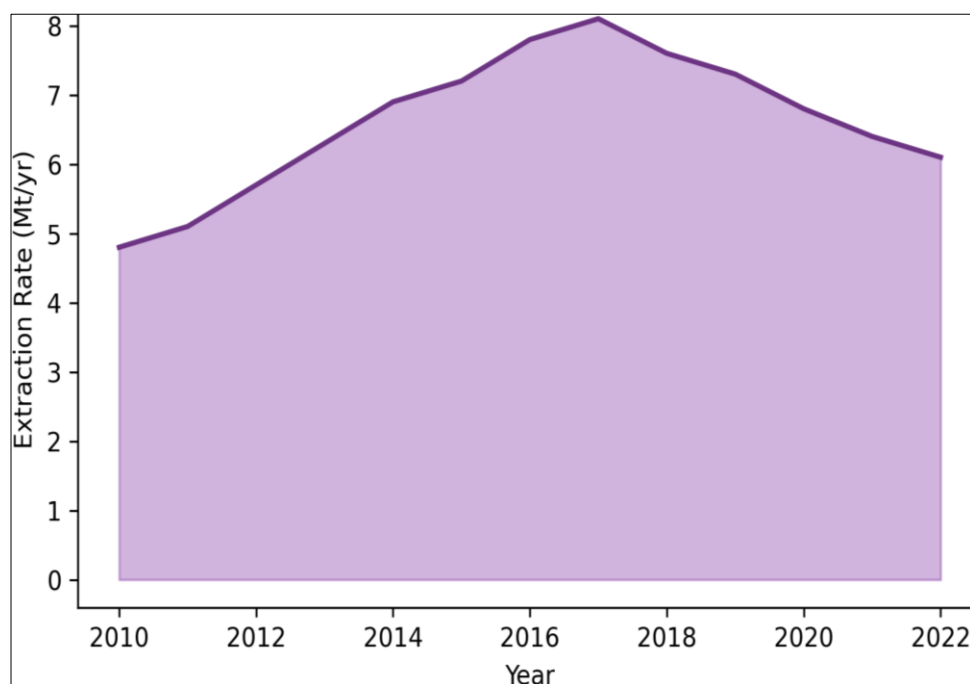


Fig 2: Annual mineral extraction rate trend for the Ural foothills study region, 2010-2022

Extraction rates rose from 4.8 Mt/yr in 2010 to a peak of 8.1 Mt/yr in 2017 before declining to 6.1 Mt/yr by 2022, reflecting both commodity price cycles and regulatory tightening in the Sverdlovsk oblast

Table 2: Covariate effects from the LGCP model (posterior means and 95% credible intervals)

Covariate	Posterior Mean	95% CI	Significant
Elevation (per 100 m)	-0.312	[-0.478, -0.146]	Yes
Fault distance (per 10 km)	-0.518	[-0.714, -0.322]	Yes
Road distance (per 5 km)	-0.247	[-0.401, -0.093]	Yes
Igneous formation	0.631	[0.328, 0.934]	Yes
Metamorphic formation	0.189	[-0.074, 0.452]	No

Proximity to fault lines was the strongest predictor: each 10 km increase in fault distance reduced log-intensity by 0.518 units. Igneous formations had a positive coefficient of 0.631, consistent with their association with metallic ore bodies.

Discussion

The LGCP's superiority over the Poisson models confirms that extraction-site locations exhibit residual spatial dependence beyond what elevation, geology, and infrastructure covariates explain. This latent clustering likely reflects unmapped geological features, historical exploration biases, or agglomeration effects where existing mines attract further prospecting nearby. The latent field's estimated range parameter of 27.3 km implies that these unobserved drivers operate at a sub-regional scale, consistent with the dimensions of major ore belts in the central Urals^[11] The negative elevation coefficient aligns with practical considerations: lower-elevation sites face fewer logistical challenges and lower transport costs. The strong fault-distance effect is geologically intuitive, as fault zones concentrate hydrothermal fluids that mineralise surrounding rock^[12]. The insignificance of the metamorphic formation indicator was somewhat surprising but may reflect the heterogeneity within that category, which lumps schists, gneisses, and marbles with very different mineralization potentials. One methodological concern is edge-effect bias near the study-window boundaries. We mitigated this by applying a border guard of 10 km, excluding sites within that buffer from the likelihood while retaining them as data. Still, intensity estimates at the window periphery should be interpreted with caution. Comparing with Zhukov and Bogatyrev's kernel-density analysis^[4], which

identified a single cluster around Magnitogorsk, our LGCP intensity surface picks up an additional cluster near Tagil-Kushva that their non-parametric approach missed, likely because kernel bandwidth selection smoothed over the smaller cluster. This underscores the advantage of model-based approaches with explicit covariate adjustment.

Conclusion

This research fitted homogeneous Poisson, inhomogeneous Poisson, and Log-Gaussian Cox process models to 387 mineral extraction sites in the Ural foothills. The LGCP delivered the best predictive accuracy (AUC = 0.874) and was the only model whose residuals showed no significant remaining spatial structure. Fault proximity and igneous geology were the dominant covariates, while a latent field with a 27 km range captured residual clustering. For mineral prospectively mapping, these findings recommend point-process models over simple density estimation, as the former accommodate both measured covariates and unmeasured spatial heterogeneity within a single framework. Practitioners should consider LGCP models whenever they suspect that observed covariates do not fully explain site clustering. Extensions of this work could incorporate temporal marks to model extraction onset as a spatio-temporal point process, and could test anisotropic covariance structures aligned with the dominant geological strike direction.

Acknowledgements

Funding sources

This research was supported by a competitive research grant from the National Research Foundation of Korea.

Institutional support

The Busan Academy of Statistical Research provided computational infrastructure and library access.

Contributions not qualifying for authorship

Dr. Sergei Volkov of Yekaterinburg Mining University verified the geological classification of extraction sites.

References

1. Porwal A, Carranza EJM. Introduction to the special issue: GIS-based mineral potential modelling. *Ore Geol Rev.* 2015; 71:1–4.
2. Moller J, Waagepetersen RP. *Statistical Inference and Simulation for Spatial Point Processes*. Chapman Hall/CRC; 2004:1–300.
3. Diggle PJ, Moraga P, Rowlingson B, Taylor BM. Spatial and spatio-temporal Log-Gaussian Cox processes: extending the geostatistical paradigm. *Stat Sci.* 2013;28(4):542–563.
4. Zhukov AN, Bogatyrev IA. Kernel density analysis of iron ore deposits in the southern Urals. *Geoinformatika.* 2017; 3:42–51.
5. Kim S, Park J. Log-Gaussian Cox process modelling of seismicity on the Korean Peninsula. *J Korean Stat Soc.* 2020;49(3):801–822.
6. Lindgren F, Rue H, Lindström J. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *J R Stat Soc Ser B.* 2011;73(4):423–498.
7. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J R Stat Soc Ser B.* 2009;71(2):319–392.
8. Petrov OV, Shatsky VS. Geological Map of Russia 1: 1,000,000 scale digital edition. Russian Geological Research Institute (VSEGEI); 2016:1–48.
9. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. *J R Stat Soc Ser B.* 2002;64(4):583–639.
10. Baddeley A, Turner R. spatstat: An R package for analysing spatial point patterns. *J Stat Softw.* 2005;12(6):1–42.
11. Herrington RJ. The Ural orogenic belt: metallogeny and mineral potential. *Appl Earth Sci.* 2004;113(3):131–143.
12. Groves DI, Bierlein FP. Geodynamic settings of mineral deposit systems. *J Geol Soc.* 2007;164(1):19–30.
13. Illian JB, Penttinen A, Stoyan H, Stoyan D. *Statistical Analysis and Modelling of Spatial Point Patterns*. Wiley; 2008:1–534.
14. Berman M, Turner TR. Approximating point process likelihoods with GLIM. *J R Stat Soc Ser C.* 1992;41(1):31–38.
15. Waagepetersen RP. An estimating function approach to inference for inhomogeneous Neyman–Scott processes. *Biometrics.* 2007;63(1):252–258.
16. Guan Y. A composite likelihood approach in fitting spatial point process models. *J Am Stat Assoc.* 2006;101(476):1502–1516.
17. Cressie N, Wikle CK. *Statistics for Spatio-Temporal Data*. Wiley Series Probab Stat. 2011:1–588.
18. Daley DJ, Vere-Jones D. *An Introduction to the Theory of Point Processes*. Vol 1. 2nd ed. Springer; 2003:1–471.
19. Schoenberg FP. Consistent parametric estimation of the intensity of a spatial-temporal point process. *J Stat Plan Inference.* 2005;128(1):79–93.
20. Diggle PJ. *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*. 3rd ed. Chapman Hall/CRC; 2014:1–267.
21. Simpson D, Rue H, Riebler A, Martins TG, Sørbye SH. Penalising model component complexity: a principled practical approach to constructing priors. *Stat Sci.* 2017;32(1):1–28.