Warping Away Nonstationarity: Benefits in Mineral Resource Estimation

F.S. Leal-Villaseca*
Department of Mathematics and Statistics, School of Physics, Mathematics and Computing
The University of Western Australia
35 Stirling Highway, Perth WA 6009, Australia
fabianlealv@gmail.com

E. Cripps
Department of Mathematics and Statistics, School of Physics, Mathematics and Computing
The University of Western Australia
35 Stirling Highway, Perth WA 6009, Australia

M. Jessell
Centre for Exploration Targeting, School of Earth Sciences
The University of Western Australia
35 Stirling Highway, Perth WA 6009, Australia

M. Lindsay
CSIRO Mineral Resources
Kensington WA 6151, Australia

Abstract

Mineral resource estimation predicts grade distributions between limited hole measurements. Spatial structure is directional (anisotropy) and changes with location (nonstationarity), so traditional stationary methods can oversmooth and leak continuity across boundaries. We introduce a quantitative framework that measures and separates the value of modelling anisotropy and depth-changing correlation by coupling a warped Gaussian process (GeoWarp) with depth-wise moving-window variogram diagnostics. GeoWarp decomposes a depth trend and a 3D residual process and learns a monotone coordinate warp; its vertical warp derivative, together with the variogram diagnostic computed in shallow sliding depth bins, provides an interpretable readout of vertical continuity. On Kevitsa drill-hole data we compare two variants: a linear axis-wise warp (anisotropy only) and the same model with a nonlinear depth warp. Using a geographically blocked split that mimics step-out drilling, the anisotropy-only variant delivers the best average MAE/RMSE across three of the four metals we made predictions on; the nonlinear depth warp adds value only where diagnostics reveal strong, regular short-scale vertical structure. The framework tells practitioners when depth warping helps, quantifies its marginal benefit over anisotropy, and offers a practical, scalable machine learning alternative for hole-based interpolation.

1 Introduction

Mineral resource estimation (MRE) asks a simple question with high stakes: given measurements from a limited number of holes, what is the concentration of a mineral in the rock volume between

*Correspondence: fabianlealv@gmail.com. All authors are affiliated with the ARC Centre for Data Analytics for Resources and Environments (DARE), The University of Sydney, 1 Central Avenue, South Eveleigh NSW 2015, Australia.

them? These holes provide many measurements along depth in a single location but are spaced far apart horizontally, so models must infer structure in the large unknown regions between holes. Decisions such as mine feasibility, planning and risk analysis depend on these estimates and their uncertainties [Rossi and Deutsch, 2014].

Two empirical facts drive the challenge. First, spatial patterns are often *directional*: continuity down the hole can be very different from continuity across holes. Second, patterns *change with depth*: the strength or scale of correlation near the surface can differ from deeper layers. Correlations are anisotropic (direction dependent) and nonstationary (location dependent) [Chilès and Delfiner, 2012]. Methods that assume a single, global correlation, common in stationary kriging workflows [Journel and Huijbregts, 1978], tend to oversmooth and can "leak" continuity across geological boundaries [Goovaerts, 1997, Maleki and Emery, 2019].

Spatial warping offers a way to relax these assumptions by learning a smooth, monotone transformation of coordinates so that, in the transformed space, correlation is closer to stationary before applying a standard kernel [Sampson and Guttorp, 1992]. We adopt GeoWarp [Bertolacci et al., 2025], a Bayesian warped Gaussian process that separates a depth-wise mean from a 3D deviation process and scales to large datasets via Vecchia approximations [Vecchia, 1988]. We use two variants: a linear, axis-wise warp to capture anisotropy and a nonlinear vertical warp to allow depth-dependent nonstationarity.

The primary contribution is a quantitative assessment framework for anisotropy and depth-dependent nonstationarity in *hole-based* interpolation problems. We pair a moving-window variogram [Li and Lake, 1994] over depth; empirical variograms computed in shallow sliding depth bins to track how spatial continuity changes; with GeoWarp's interpretable vertical warp derivative, an inverse proxy for vertical correlation strength [Bertolacci et al., 2025]. This combination detects where vertical continuity changes, and contrasting the accuracy gains measures how much predictive benefit is gained by modelling it explicitly. Although demonstrated on Kevitsa drill-hole data, the framework applies broadly to boreholes and penetrometer soundings where measurements are dense along depth and sparse laterally.

As a secondary contribution, this work presents a case study that brings GeoWarp; originally developed for cone penetration tests in subsea sediments; into MRE, which differs in geometry, anisotropy, and variability of the target variable. Despite these differences, with our configuration, GeoWarp consistently outperforms widely used baselines under a geographically blocked split. Our work positions warped Gaussian processes as a practical machine learning alternative that integrates with standard drill-hole workflows, preserves calibrated uncertainty, and is applicable to a broad set of hole-based settings beyond MRE.

2 Problem setup and objectives

We observe composite grades z_i at borehole coordinates (x_i, y_i, h_i), where h is depth and $\mathbf{s} = (x, y)$ is location. The goal is to predict the latent grade field $Y(\mathbf{s}, h)$ between holes and quantify uncertainty, so that decisions (e.g., cut-offs and planning) can reflect risk. In general terms, this is supervised regression with uncertainty under structured spatial dependence: measurements are dense along depth but sparse laterally, correlations are directional (anisotropy), and they can change with depth (nonstationarity). A simple data-generating view is $z_i = Y(\mathbf{s}_i, h_i) + \varepsilon_i$, with ε_i capturing assay/compositing noise; the modelling question is how to encode the spatial structure in Y so that predictions are accurate and well-calibrated.

Our primary objective is to study the marginal value of modelling anisotropy and depth-changing correlation, which we achieve using a warped Gaussian process (GeoWarp) paired with a depth-wise moving-window variogram. Practically, we compare two GeoWarp variants: a linear axis-wise warp (anisotropy only), and the same model with a nonlinear depth warp, against stationary baselines.

3 Methodology

Model. For location $\mathbf{s} = (x, y)$ and depth h, GeoWarp models

$$Y(\mathbf{s}, h) = \mu(h) + \delta\big(\mathbf{T}(\mathbf{s}, h)\big), \tag{1}$$

where $\mu(h)$ is a B–spline mean over depth and $\delta(\cdot)$ is a zero–mean Gaussian process with Matérn correlation. The covariance is

$$Cov\big[\delta(u_1), \delta(u_2)\big] = \sqrt{\sigma_\delta^2(h_1)\, \sigma_\delta^2(h_2)}\, \mathcal{M}_\nu(u_1 - u_2)\,,\ u_i = \mathbf{T}(\mathbf{s}_i, h_i). \tag{2}$$

The Matérn kernel with smoothness $\nu = 3/2$ is $\mathcal{M}_\nu(r) = \big(1 + \sqrt{3}\, r/\ell\big) \exp\big(-\sqrt{3}\, r/\ell\big)$ with range parameter ℓ. Nonstationarity enters through the monotone warp $\mathbf{T}$ and, optionally, a depth-varying variance $\sigma_\delta^2(h)$. Inference uses Vecchia approximations [Vecchia, 1988] for scalability: with m neighbours per location the cost scales approximately as $\mathcal{O}(nm^2)$ for n observations, enabling large datasets (up to $22,275$ data points in this study) to be executed in a singular machine.

Warping choices. *LinearWarp* (GW-Lin) applies axis-wise linear rescalings, yielding a purely anisotropic but stationary Matérn covariance. *DepthDeform* (GW-Depth) retains linear horizontal scalings but learns a flexible, monotone vertical warping via a Bernstein polynomial basis, allowing vertical correlation ranges to vary smoothly with depth. The derivative $\mathrm{d}T_3/\mathrm{d}h$ is inversely proportional to the vertical correlation.

To ensure a clean separation between large-scale depth trends and spatial warping, the hierarchical mean $\mu(h)$ captures all smooth, monotone depth effects through a low-order B-spline basis with independent priors on the coefficients. The warp $\mathbf{T}(\mathbf{s}, h)$ is applied only to the residual process $\delta(\cdot)$, so that it can modify spatial correlation without absorbing compositional or depth-driven shifts in the mean structure. The monotonicity of the vertical warp $f_3(h)$ is enforced by parameterising it as a cumulative sum of positive increments ($\gamma_{3,k} > 0$), guaranteeing $\mathrm{d}f_3/\mathrm{d}h > 0$ everywhere. Smoothness is controlled via a low-order Bernstein polynomial basis and weakly informative $\mathrm{Gamma}(1.01, 0.01)$ priors on $\gamma_{3,k}$, which penalise large local distortions and encourage gradual variation with depth.

Calibration and metrics. We report mean absolute error (MAE) and root mean squared error (RMSE) on log-transformed grades, plus empirical coverage at ± 2 standard deviations (SD) is the proportion of held-out points with $|y_i - \hat{\mu}_i| \leq 2\hat{\sigma}_i$, where $\hat{\mu}_i$ and $\hat{\sigma}_i$ are the predictive mean and SD. A well-calibrated Gaussian gives 95.45%. We report the across metal average absolute deviation from this value.

Depth-wise moving-window variogram diagnostic. To detect and quantify vertical nonstationarity, we compute variograms within sliding depth windows and track their fitted parameters as functions of depth. Composites are assigned to consecutive depth bins of fixed thickness. Within each bin, grades are locally standardised to remove slow trends and scale differences, and an empirical semi-variogram

$$\widehat{\gamma}(r) = \frac{1}{2|N(r)|} \sum_{(i,j) \in N(r)} \big[z(\mathbf{s}_i, h_i) - z(\mathbf{s}_j, h_j)\big]^2$$

is fit with a simple model to extract a sill and range for that bin. The resulting sill and range profiles reveal how spatial continuity evolves with depth. For a compact scalar summary, we detrend the sill profile with a smooth LOESS (Locally Estimated Scatterplot Smoothing) curve [Cleveland, 1979, Cleveland and Devlin, 1988] and compute the lag-1 autocorrelation of the residuals, which is a standard summary of short-lag dependence [Box et al., 2015, Shumway and Stoffer, 2017], together with a robust peak-to-peak spacing; strong negative lag-1 values and short spacing indicate regular alternating layering at short vertical scales. These diagnostics are complementary to GeoWarp's vertical warp derivative $\mathrm{d}T_3/\mathrm{d}h$, which is inversely related to vertical correlation strength and can be read directly from the fitted model.

Kevitsa dataset, experimental design and depth diagnostics Kevitsa (northern Finland) is a disseminated Ni–Cu deposit with Au and Ag by–products. We form four element–specific subsets (Ag, Au, Cu, Ni), log–transform grades, and standardise coordinates to the unit cube for numerical stability.

To emulate prediction into new push–backs, we withhold a spatially distinct test block identified via k–means clustering (k=2) on the horizontal (x, y) coordinates of drill-hole collars. This procedure partitions the dataset into two spatially coherent clusters, from which the smaller one, containing roughly 30% of composites, is designated as the test set. All samples from holes belonging to that cluster are withheld entirely, ensuring a clear spatial separation between training and test data. The resulting boundary lies well beyond the dominant horizontal variogram range (~ 150–200 m), effectively preventing information leakage that would arise if neighbouring samples from the same geological volume were split across subsets.

Metal	Lag-1 ACF (30 m)	Median peak spacing (m)
Ag	**-0.81**	**60**
Au	-0.19	60
Cu	-0.27	60
Ni	-0.22	90

Table 1: Moving-window variogram diagnostic using 30 m depth bins on detrended sill residuals. More negative lag-1 implies stronger short-scale alternation; smaller spacing implies finer layering.

	Ag MAE/RMSE	Au MAE/RMSE	Cu MAE/RMSE	Ni MAE/RMSE	Avg MAE	Avg RMSE	Avg $\lvert\Delta\rvert$ @2SD
GW-Lin	0.82/0.96	**1.12/1.46**	**0.51/0.60**	0.41/0.49	**0.715**	**0.878**	**1.78**
GW-Depth	**0.81/0.96**	1.14/1.47	0.51/0.61	0.42/0.51	0.720	0.888	2.00
OK	0.83/1.05	1.22/1.59	0.51/0.61	0.42/0.53	0.745	0.945	2.03
SK	0.86/0.99	1.16/1.50	0.53/0.61	0.41/0.48	0.740	0.895	2.65
IDW (p=2)	0.86/0.98	1.19/1.53	0.53/0.62	**0.40/0.47**	0.745	0.900	n/a

Table 2: Out-of-sample log-grade performance under a blocked split. Columns show MAE/RMSE (lower is better). The last column reports the mean absolute deviation of coverage within ± 2 SD from the Gaussian nominal 95.45% across Ag, Au, Cu, Ni. GW-Lin = GeoWarp linear warp; GW-Depth = nonlinear depth warp; OK = ordinary kriging; SK = simple kriging.

This geographically blocked design reflects the operational reality of step-out drilling, where new push–backs are developed sequentially and predictions must extrapolate into unmined areas. It forces the models to interpolate across geological discontinuities such as the Satovaara fault, rather than within densely sampled zones. Compared with a random split, which would allow models to "peek" at near-identical neighbouring samples and thus overstate accuracy, this approach provides a more rigorous and realistic assessment of predictive performance under conditions that mimic how resource models are applied in practice.

For further characterisations of the deposit, refer to Horrocks et al. [2019], Junno et al. [2020], Koivisto et al. [2012]. Comprehensive information on the geological setting of the intrusion can be found in Mutanen [1997].

Code availability. All code to reproduce data processing, model fitting (GW-Lin and GW-Depth), diagnostics (moving-window variograms and warp-derivative profiles), and figures is openly available at `https://anonymous.4open.science/r/GeoWarp-Application-Kevitsa-6F73`. The repository includes documented scripts and configuration files for end-to-end replication on the Kevitsa drill-hole data.

4 Results and Discussion

Predictive accuracy. Table 2 summarises performance on the log scale. GW-Lin attains the best average MAE and RMSE across metals, improving mean MAE by about three percent over the strongest stationary baseline and about seven percent over ordinary kriging. GW-Depth improves Ag marginally but not Au, Cu or Ni. Regarding coverage calibration at the ± 2 SD level, GW-Lin has the smallest average deviation at 1.78%, but all models show modest deviations, indicating broadly well-calibrated uncertainty. GW-Lin combines the best calibration with the top MAE/RMSE, while GW-Depth is a close second in both.

These improvements may seem small, but even modest percentage gains in MAE or RMSE can be operationally significant in mineral resource estimation, where millions of tonnes are extrapolated from sparse drill data; such improvements translate to more reliable grade forecasts and reduced financial uncertainty at deposit scale, justifying the added model complexity.

Depth-dependent nonstationarity and when warping helps. The depth-wise moving-window variogram and GeoWarp's vertical warp derivative $\mathrm{d}T_3/\mathrm{d}h$ give consistent readouts of vertical continuity. Table 1 shows that Ag has the strongest and most regular short-scale alternation (lag-1 $= -0.81$ with ~ 60 m spacing), Au and Cu show weaker effects at a similar scale, and Ni exhibits coarser layering (~ 90 m). The fitted models mirror this: Figure 1 shows that $\mathrm{d}T_3/\mathrm{d}h$ is persistently smaller for Ag, indicating stronger vertical correlation, and comparatively flat for Au, Cu, and Ni. Table 2 further corroborates this, showing that only for Ag, the nonlinear GW-Depth improves accuracy beyond GW-Lin.

Means and standard deviations. Figure 1 shows monotone increases in log grade with depth for all four metals, with GW-Lin and GW-Depth closely aligned and small near-surface offsets that fade at depth. Regarding standard deviations, only GW-Depth exhibits nonlinearity, even though both models can represent it, while overall levels are comparable: Ag and Cu are essentially flat (≈ 0.89, ≈ 0.62), whereas Au and Ni taper with depth ($\approx 1.70 \rightarrow 1.45$, $\approx 0.56 \rightarrow 0.48$).

5 Conclusion

We present a compact, interpretable framework that measures anisotropy and depth-dependent nonstationarity by pairing GeoWarp with a depth-wise moving-window variogram to measure marginal predictive value. These diagnostics are corroborated with the accuracy metrics: linear anisotropy accounts for most gains over OK/SK/IDW; a nonlinear depth warp helps only when diagnostics show strong, regular short-scale vertical structure (Ag), and is otherwise unnecessary.

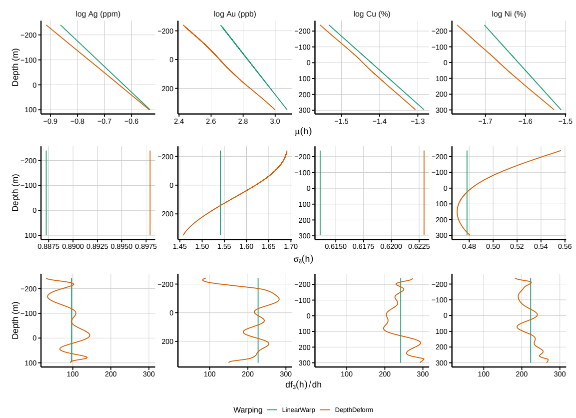

Figure 1: Depth-wise profiles of log Ag, Au, Cu, and Ni. Top: mean; middle: standard deviation; bottom: vertical warp derivative. Waps: linear (green) and nonlinear (orange).

Although GeoWarp was built for cone penetration test profiles, it outperforms stationary baselines in MRE despite different geometry and variability, in both accuracy and coverage calibration, suggesting broader use across hole-based settings with dense depth sampling and sparse lateral control. Evaluation uses a geographically blocked split beyond the horizontal range with harmonised neighbourhood sizes. The approach integrates with standard workflows, retains Matérn familiarity, and scales comparably to neighbourhood kriging via Vecchia. The depth-wise variance provide simple model-side diagnostics that align with data-side variograms.

While results across four metals at Kevitsa demonstrate clear benefits, the study is limited to a single deposit. Because geological structures and continuity patterns differ across settings, further testing on deposits with varied geometries and mineralisation styles is needed to assess robustness and generalisability.

References

M. Bertolacci, A. Zammit-Mangion, J. Valderrama Giraldo, M. O'Neill, F. Bransby, and P. Watson. Geowarp: Warped spatial processes for inferring subsea sediment properties. *Journal of the American Statistical Association*, 120(550):710–722, 2025.

G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung. *Time Series Analysis: Forecasting and Control*. Wiley, Hoboken, NJ, 5th edition, 2015. ISBN 978-1-118-67502-1.

J.-P. Chilès and P. Delfiner. *Geostatistics: Modeling Spatial Uncertainty*. Wiley Series in Probability and Statistics. Wiley, 2nd edition, 2012. ISBN 978-0-470-18315-1.

W. S. Cleveland. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74(368):829–836, 1979. doi: 10.1080/01621459.1979.10481038.

W. S. Cleveland and S. J. Devlin. Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association*, 83(403):596–610, 1988. doi: 10.1080/01621459.1988.10478639.

P. Goovaerts. *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York, 1997. ISBN 0-19-511538-4.

T. Horrocks, E.-J. Holden, D. Wedge, C. Wijns, and M. Fiorentini. Geochemical characterisation of rock hydration processes using t-sne. *Computers & geosciences*, 124:46–57, 2019.

A. G. Journel and C. J. Huijbregts. *Mining Geostatistics*. Academic Press, London, 1978.

N. Junno, E. Koivisto, I. Kukkonen, A. Malehmir, C. Wijns, and M. Montonen. Data mining of petrophysical and lithogeochemical borehole data to elucidate the origin of seismic reflectivity within the kevitsa ni–cu–pge bearing intrusion, northern finland. *Geophysical Prospecting*, 68 (1-Cost-Effective and Innovative Mineral Exploration Solutions):82–102, 2020.

E. Koivisto, A. Malehmir, P. Heikkinen, S. Heinonen, and I. Kukkonen. 2d reflection seismic investigations at the kevitsa ni-cu-pge deposit, northern finland. *Geophysics*, 77(5): WC149–WC162, 2012.

D. Li and L. W. Lake. A moving window semivariance estimator. *Water Resources Research*, 30(5): 1479–1489, 1994. doi: 10.1029/94WR00319.

M. Maleki and X. Emery. Geostatistics in the presence of geological boundaries: Application to mineral resources modeling. *Ore Geology Reviews*, 114:103124, 2019. doi: 10.1016/j.oregeorev.2019.103124.

T. Mutanen. *Geology and ore petrology of the Akanvaara and Koitelainen mafic layered intrusions and the Keivitsa-Satovaara layered complex, northern Finland*, volume 395. Geological Survey of Finland, 1997.

M. E. Rossi and C. V. Deutsch. *Mineral Resource Estimation*. Springer, Dordrecht, 2014. ISBN 978-1-4020-5716-8. doi: 10.1007/978-1-4020-5717-5.

P. D. Sampson and P. Guttorp. Nonparametric estimation of nonstationary spatial covariance structure. *Journal of the American Statistical Association*, 1992.

R. H. Shumway and D. S. Stoffer. *Time Series Analysis and Its Applications: With R Examples*. Springer Texts in Statistics. Springer, 4th edition, 2017. ISBN 978-3-319-52451-1.

A. V. Vecchia. Estimation and model identification for continuous spatial processes. *Journal of the Royal Statistical Society: Series B*, 1988.