Versatile Parametric Classes of Covariance Functions that Interlace Anisotropies and Hole Effects

Alfredo Alegría[1] and Xavier Emery[2,3]

[1] *Departamento de Matemática, Universidad Técnica Federico Santa María, Valparaíso, Chile*

[2] *Department of Mining Engineering, Universidad de Chile, Santiago, Chile*

[3] *Advanced Mining Technology Center, Universidad de Chile, Santiago, Chile*

Supplementary Material

S1 Literature Review on Spatial Hole Effects

Spatial covariances with hole effects have been reported in different scientific and engineering disciplines, including

- atmospheric science (Bleck, 1975; Thiébaux, 1976, 1985);
- soil science (Webster, 1977; Pierson and Wight, 1991; Ciollaro and Romano, 1995; Guillobez and Arnaud, 1998; Sharifi et al., 2020);
- agronomy (San Martín et al., 2018; Bosaz et al., 2019);
- ecology (Curran, 1988; Cohen et al., 1990; Pastor et al., 1998);

- physics (Price and Kozlowski, 2021);
- manufacturing and materials science (Bonetto et al., 2002; Yang and Shao, 2018; Everett et al., 2020);
- biology (Mary-Huard et al., 2004; Dong et al., 2015);
- image analysis (Balaguer et al., 2010; Balaguer-Beser et al., 2013);
- geotechnical engineering (Chang et al., 2021; Ching et al., 2023);
- geomorphology (Jordan, 2003);
- geodesy (Varbla and Ellmann, 2023);
- glaciology (Irvine-Fynn et al., 2022);
- hydrology (Fiori et al., 2003; Chen, 2005; Salamon et al., 2007);
- outcrop-based geology (Jennings et al., 2000; Budd et al., 2006; Matonti et al., 2015);
- subsurface geology (Journel and Froidevaux, 1982; Jones and Ma, 2001; Lefranc et al., 2008; Parra and Emery, 2013; Emery and Parra, 2013; Le Blévec et al., 2018).

Many of the above references have pointed out the importance of accounting for hole effects for improved spatial predictions, uncertainty modeling and decision-making.

S2 Visual Illustrations

S2.1 Basic Models

Figure S1 shows the following basic constructions in dimension $d = 2$:

I. $2\exp(-0.8\,\boldsymbol{h}^\top \mathbf{A}\boldsymbol{h}) - \exp(-0.4\,\boldsymbol{h}^\top \mathbf{A}\boldsymbol{h})$, with $\mathbf{A} = \begin{pmatrix} 1 & -0.5 \\ -0.5 & 1 \end{pmatrix}$.

II. $\exp(-0.2\,\|\boldsymbol{h}\|^2)\,[3.41\exp(-0.8\,h_2^2) - 2.41\exp(-0.4\,h_2^2)]$.

III. $\exp(-0.5\,\|\boldsymbol{h}\|)\mathcal{W}(5|h_2|)$, with $\mathcal{W}$ being the wave model.

The positive semidefiniteness of the first two models, where differences of covariance functions are involved, is a consequence of Theorem 1(i) in Ma (2005).

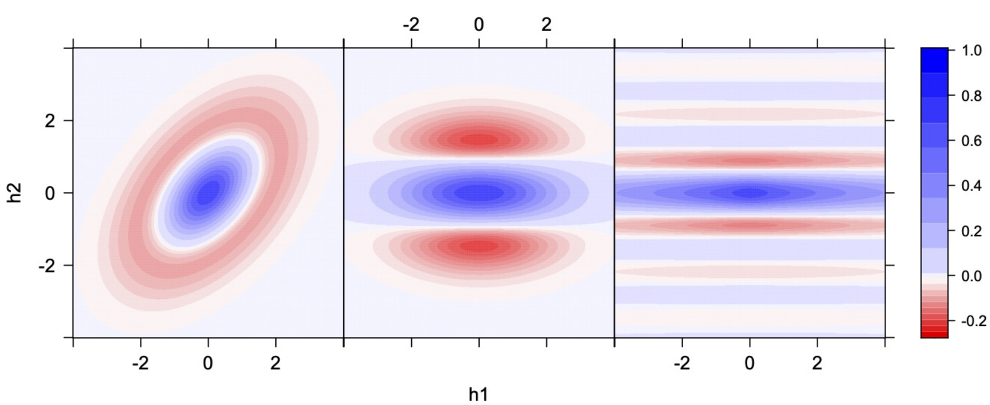

Figure S1: Basic models **I-III** (from left to right).

S2.2 Proposed Models

We now illustrate the various shapes that can be achieved with the proposed models in dimension $d = 2$. We consider the following scenarios:

I. The models in Corollary 3 with $\mathbf{A}_1 = \mathbf{I}_2$ and $\mathbf{A}_2 = \mathbf{P}\,\mathrm{diag}(\mu_1, \mu_2)\,\mathbf{P}^\top$, with $\mu_1, \mu_2 > 0$ and

$$\mathbf{P} = \begin{bmatrix} \cos(\pi/4) & -\sin(\pi/4) \\ \sin(\pi/4) & \cos(\pi/4) \end{bmatrix}$$

being a rotation matrix. The conditions of Corollary 3 are satisfied if, and only if, $\max(\mu_1, \mu_2) \leq 1$ and $b_1\sqrt{\mu_1\mu_2} \geq b_2$. Thus, we fix $b_1 = 2.5$, $b_2 = 1$, $\mu_1 = 0.2$ and $\mu_2 = 0.8$.

II. The models in Corollary 4, with $b_1 = 2$, $b_2 = 1$, $a_1 = 0.8$ and $a_2 = 0.4$, with a shift vector given by $\boldsymbol{\eta} = [1, 1]^\top$.

III. The models in Corollary 5, with $b_1 = 1$, $b_2 = 2$, $a_1 = 1$ and $a_2 = 0.5$, and the unit vector $\boldsymbol{u} = [1/\sqrt{2}, 1/\sqrt{2}]^\top$.

Figure S2 shows the contour plots of the Matérn model with $\nu = 1.5$, the Cauchy model with $\delta = 1$ and the Gauss hypergeometric model with $\alpha = 3, \beta = 7/2$ and $\gamma = 6$, after the application of the transformations described in Corollaries 3-5 under scenarios **I-III**, respectively, together with a normalization in order to obtain correlation functions. To improve the visualization of each individual model, we have chosen specific ranges for plotting. We consider $\boldsymbol{h} = [h_1, h_2]^\top \in [-10, 10]^2$ for the first two models, and $\boldsymbol{h} = [h_1, h_2]^\top \in [-2, 2]^2$ for the last model. All the covariance functions have been designed to present a hole effect around the northeast direction.

Figure S3 shows $\mathcal{T}^{(3)}_{a_1,a_2,b_1,b_2,\boldsymbol{u}}[\mathcal{W}]$ and $\mathcal{T}^{(3)}_{a_1,a_2,b_1,b_2,\boldsymbol{u}}[\mathcal{M}_{1/2}, \mathcal{W}]$ in dimension $d = 2$, with parameters $a_1 = a_2 = b_1 = 1$, $b_2 = 2$ and $\boldsymbol{u} = [1/\sqrt{2}, 1/\sqrt{2}]^\top$. While certain structural oscillations from the cardinal sine model persist, the proposed models exhibit a notably amplified hole effect in the $\boldsymbol{u}$ direction. Observe that $\mathcal{T}^{(3)}_{a_1,a_2,b_1,b_2,\boldsymbol{u}}[\mathcal{W}]$ exceeds the lower bound required for isotropic models in $\mathbb{R}^2$.

S3 Simulation Study

We use simulated data to compare covariance models with or without hole effect and anisotropy. Specifically, we consider $d = 2$ and the following covariance structures, where $\mathcal{H}_{\alpha,\beta,\gamma}$ stands for the radial part of the Gauss hypergeometric model defined in (4.21):

- **Model 1.** $\sigma^2 \mathcal{H}_{\alpha,\beta,\gamma}(\sqrt{a}\|\boldsymbol{h}\|)$

- **Model 2.** $b_1 \mathcal{H}_{\alpha,\beta,\gamma}(\sqrt{a}_1\|\boldsymbol{h}\|) - b_2 \mathcal{H}_{\alpha,\beta,\gamma}(\sqrt{a}_2\|\boldsymbol{h}\|)$

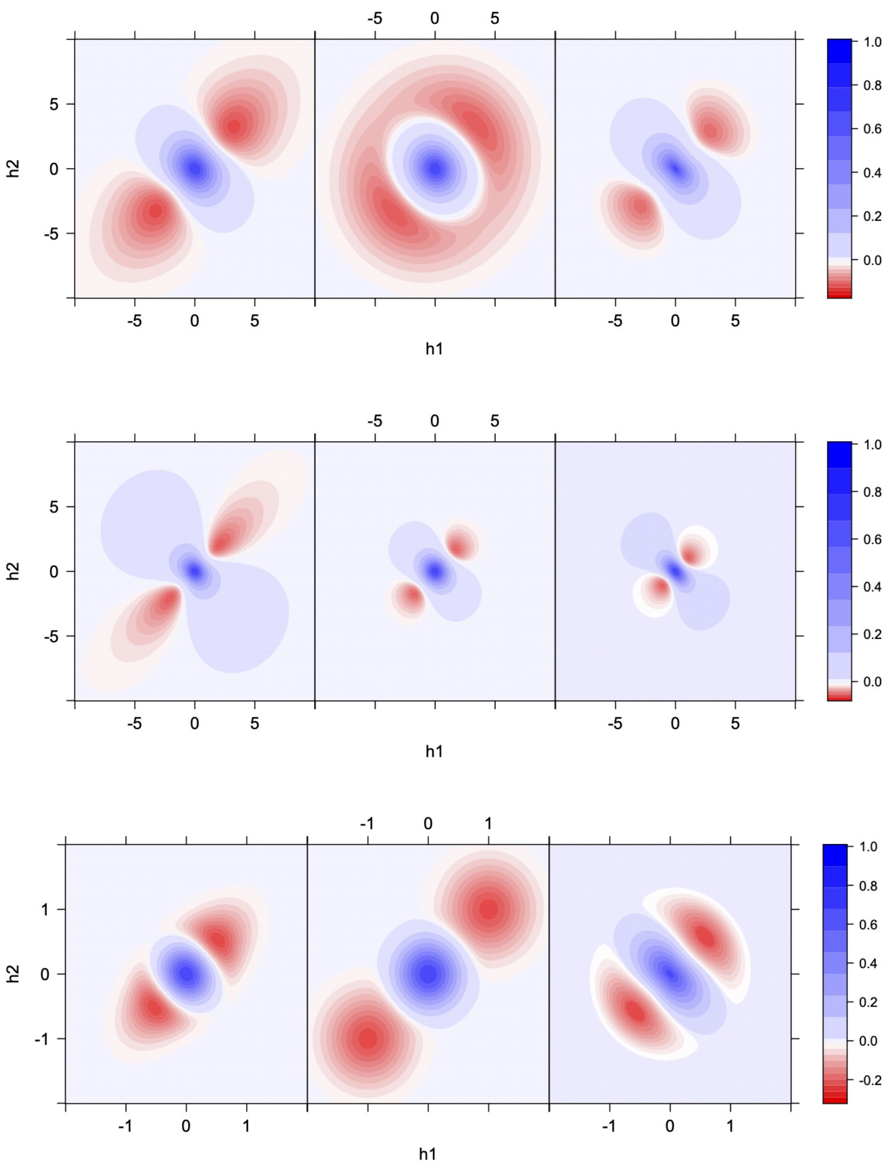

Figure S2: Different combinations of anisotropies and hole-effects for the transformed Matérn (top), the transformed Cauchy (middle) and the transformed Gauss hypergeometric (bottom) models. From left to right we consider the transformations introduced in Corollaries 3-5, respectively. The values of the parameters have been described in scenarios **I-III**.

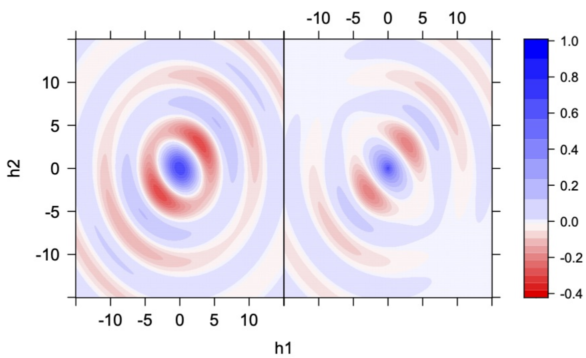

Figure S3: Models $\mathcal{T}^{(3)}_{a_1,a_2,b_1,b_2,\boldsymbol{u}}[\mathcal{W}]$ (Left) and $\mathcal{T}^{(3)}_{a_1,a_2,b_1,b_2,\boldsymbol{u}}[\mathcal{M}_{1/2},\mathcal{W}]$ (Right).

- **Model 3.** $b_1\mathcal{H}_{\alpha,\beta,\gamma}(\sqrt{a_1}\|\boldsymbol{h}\|)-\frac{b_2}{2}\big(\mathcal{H}_{\alpha,\beta,\gamma}(\sqrt{a_2}\|\boldsymbol{h}-\boldsymbol{\eta}\|)+\mathcal{H}_{\alpha,\beta,\gamma}(\sqrt{a_2}\|\boldsymbol{h}+\boldsymbol{\eta}\|)\big)$.

As in Section S2.2, we fix $\alpha = 3$, $\beta = 7/2$ and $\gamma = 6$, which corresponds to the well-known cubic model (see Emery and Alegría, 2022). Unlike Model 1, Models 2 and 3 can exhibit a hole effect. The hole effect in Model 2 appears in all directions, whereas Model 3 is the only anisotropic model and exhibits the hole effect in specific directions (it is a special case of construction (3.15)). These models offer a range of increasing complexity to make statistical comparisons, and further contribute to the ongoing discussion in this manuscript regarding the model versatility.

Figure S4 displays realizations of Gaussian random fields with these covariance models over a regular grid with 80×80 nodes in the square $[0, 4]^2$.

The simulation was performed by means of the Cholesky factorization of the covariance matrix; the same seed for random number generation was considered for each realization. The values of the parameters were $\sigma^2 = 1$ and $a = 1/2$ for Model 1, and $b_1 = 2$, $b_2 = 1$, $a_1 = 1$, $a_2 = 1/2$ for Models 2 and 3. In addition, for Model 3, we used $\boldsymbol{\eta} = \left[1/\sqrt{2}, 1/\sqrt{2}\right]^\top$. Note that these values satisfy the admissibility condition (3.14). Although some distinction emerges among these realizations, the presence or absence of hole effect or anisotropy may be difficult to perceive with a naked eye. To explore this aspect, we simulate 100 independent samples from each model, over a grid with 20×20 nodes in $[0, 4]^2$, with the parametric setting explained above, and calculate sample variograms across the northeast direction (parallel to $\boldsymbol{\eta}$) and the southeast direction (orthogonal to $\boldsymbol{\eta}$). The results are reported in Figure S5, together with the average sample variograms and the theoretical underlying models. For the first model, there is no hole effect in either direction. In the second model, a hole effect is observed along both directions. In the third model, a hole effect is observed in one direction but not in the other. These observations align closely with the expected properties of the underlying theoretical models, and serve as a reinforcement of their main features.

Now, we conduct a study to quantify the difference in terms of good-

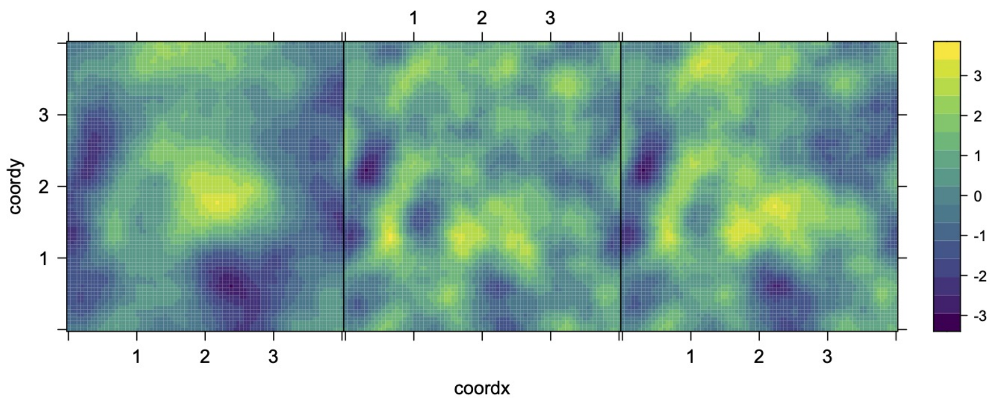

Figure S4: Realizations of Gaussian random fields from Models 1, 2 and 3 (from left to right), over a regular grid with 80×80 nodes in the square $[0, 4]^2$.

ness of fit of Model 3 with respect to the other models, when the inherent correlation structure presents a hole effect in a specific direction.

Let us first establish an adequate parameterization. Note that the intensity of the hole effect in Models 2 and 3 depends on the ratios b_1/b_2 and a_1/a_2 (see condition (3.14)). For instance, when b_1/b_2 decreases, not exceeding its minimum admissible value a_1/a_2, the significance of the negative term in the covariance structure amplifies and the hole effect emerges. Since only the ratios matter, allowing the parameters to vary freely could give rise to identifiability issues. To circumvent this problem, we parameterize Model 2 in the following way:

$$\sigma^2 \left(b\, \mathcal{H}_{\alpha,\beta,\gamma}(\|\boldsymbol{h}\|) - \mathcal{H}_{\alpha,\beta,\gamma}(\sqrt{a}\|\boldsymbol{h}\|) \right),$$

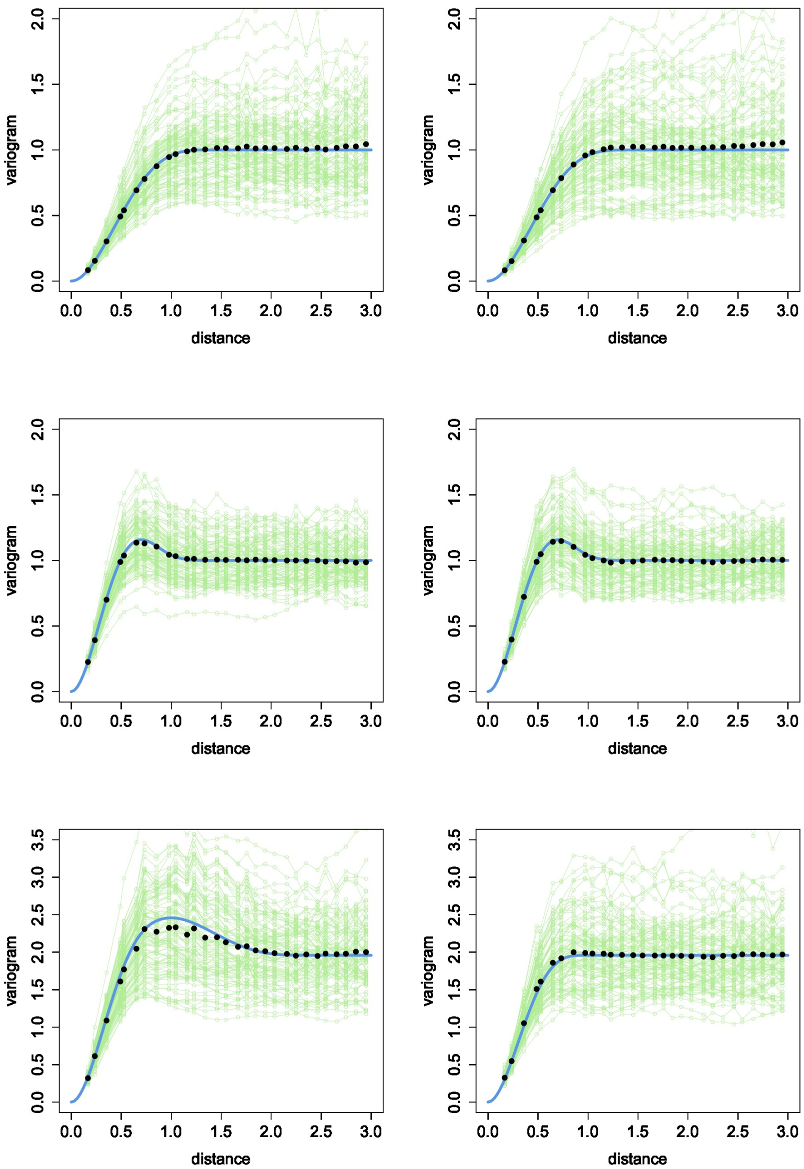

Figure S5: Sample directional variograms for 100 realizations (green lines), average of sample directional variograms (black dots) and theoretical directional variograms (blue solid lines). From the top, each row refers to Models 1, 2 and 3, respectively. Left panels are associated to northeast direction, whereas right panels correspond to southeast direction.

and the validity condition (3.14) simplifies into $b \geq 1/a$. A similar parameterization is used for Model 3. In addition, $\boldsymbol{\eta}$ is taken as a unit vector, parameterized through an angle $\theta \in [0, 2\pi)$, i.e., $\boldsymbol{\eta} = (\cos\theta, \sin\theta)^\top$. To sum up, Model 1 is parameterized by (a, σ^2), Model 2 by (b, a, σ^2) and Model 3 by (b, a, θ, σ^2). The parameters in Model 3 allow us to control the hole effect, the correlation range, the predominant direction of the hole effect, and the variance of the random field.

We simulate independent samples from Model 3, considering the following scenarios: **1.** $(a, \theta) = (1/2, \pi/4)$, **2.** $(a, \theta) = (3/4, \pi/4)$, **3.** $(a, \theta) = (1/2, 3\pi/4)$ and **4.** $(a, \theta) = (3/4, 3\pi/4)$. For each scenario, $b = 1/a$, i.e., we use its minimum admissible value, and $\sigma^2 = 1$. To understand the scenarios covered, let us analyze Model 2, as Model 3 inherits similar characteristics from it. Normalizing Model 2 to obtain a correlation structure, we find numerically that, for scenarios 2 and 4, its minimum value is -0.170 at a distance of 0.64. In contrast, for scenarios 1 and 3, the minimum value is -0.158 at a distance of 0.70. These scenarios also encompass two distinct directions regarding the presence of the hole effect.

For each scenario, we simulate 100 independent realizations from Model 3, fit each model through maximum likelihood and assess their performance through the Akaike Information Criterion (AIC), which inherently penalizes the number of parameters. The results are summarized in Table S1. Model 3 consistently shows smaller AIC averages in each of the scenarios, as expected. This improvement is more pronounced in scenarios 2 and 4, where the hole effect is more marked and the correlation range is smaller. Model 2 exhibits slightly better performance than Model 1 in all scenarios, yet both models demonstrate suboptimal results due to the misspecification. Figure S6 displays a more comprehensive panorama of the AIC values for Model 1 in comparison to those of Model 3, depicted through their ratios across 100 realizations. As AIC values are negative, a ratio less than one indicates superior performance of Model 3. While most values fall below one; it is worth noting that many of them are considerably smaller than one (it is common to observe improvements of 10% to 20%), providing further support for the superiority of Model 3. For completeness, Figure S7 displays centered boxplots for the maximum likelihood estimates of Model 3 involved in this experiment. These plots reveal unbiased estimates, emphasizing that the selected parameterization yields models with statistically meaningful and identifiable parameters.

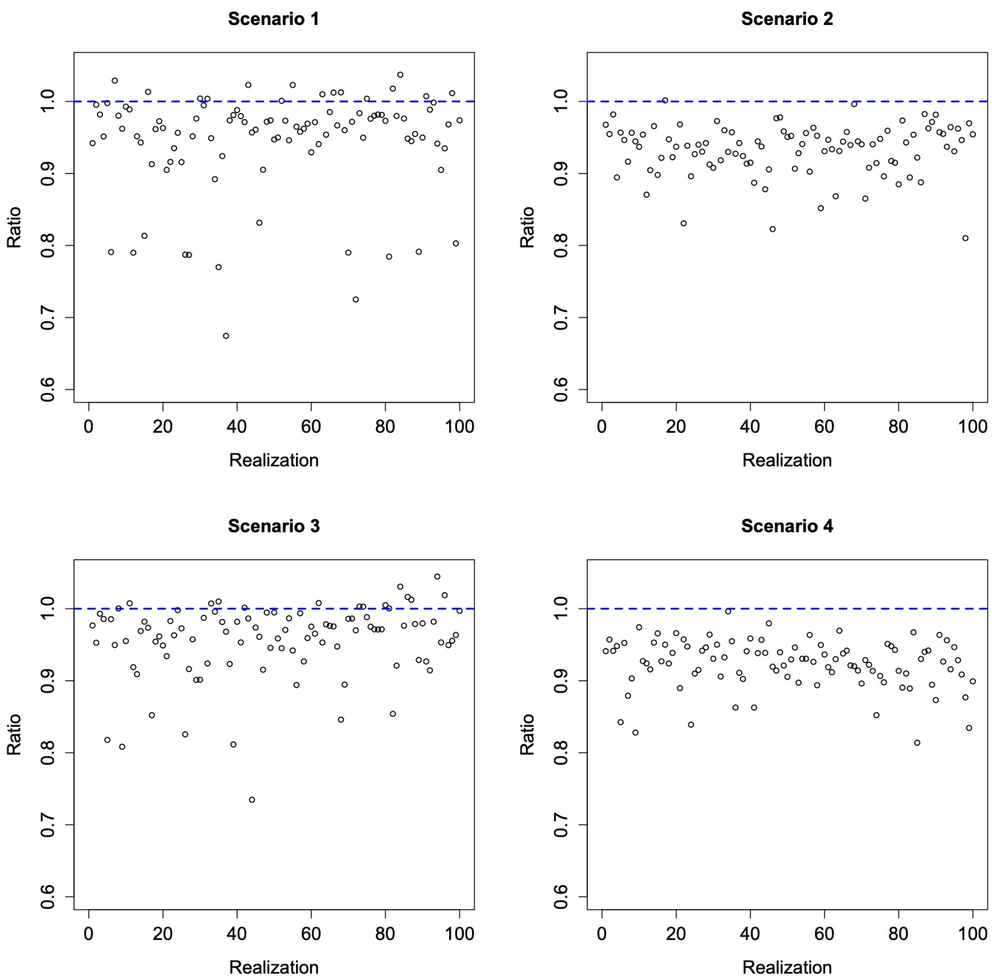

Figure S6: Ratios between AIC values for Models 1 and 3 across 100 realizations, for each scenario.

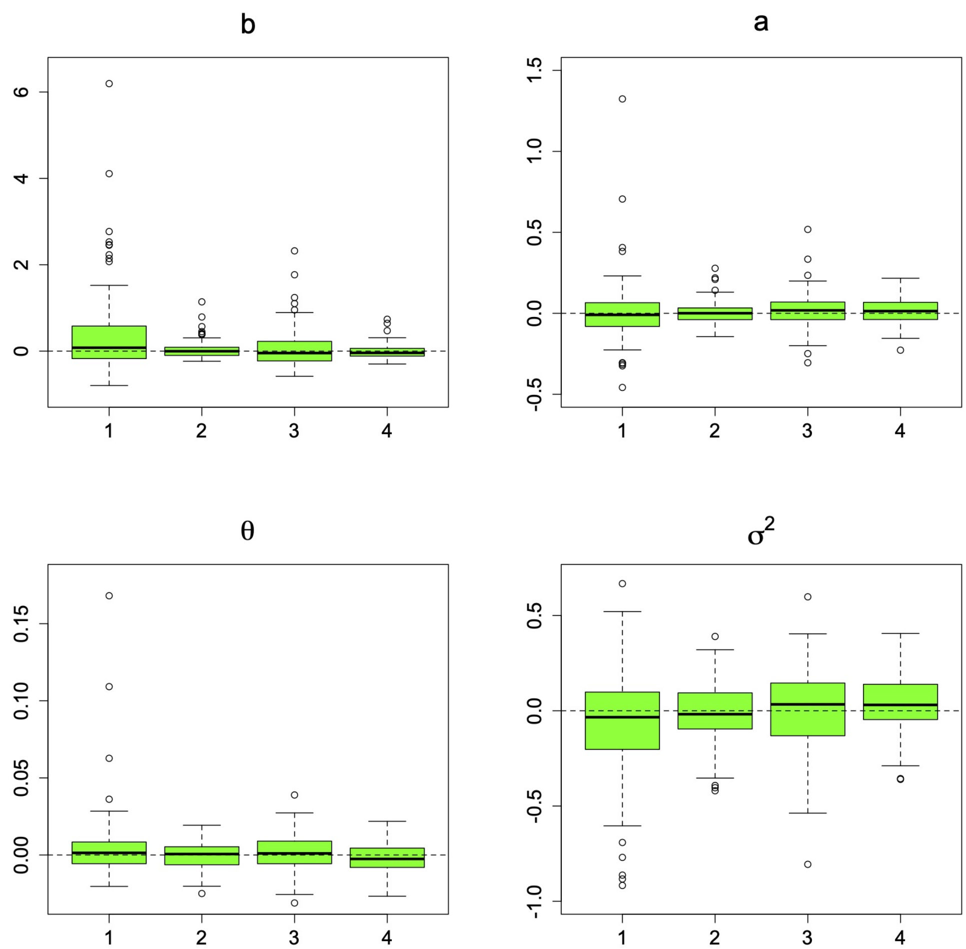

Figure S7: Centered boxplots of the maximum likelihood estimates of Model 3 for scenarios 1 to 4.

Table S1: Summary statistics of AIC values from fitting each model on 100 independent realizations.

		$a = 0.50$				$a = 0.75$			
		Q_1	Median	Q_3	Mean	Q_1	Median	Q_3	Mean
$\theta = \pi/4$	Model 1	−165.4	−147.8	−126.0	−146.5	−330.0	−313.1	−296.2	−313.4
	Model 2	−167.0	−151.5	−129.3	−148.9	−334.2	−313.6	−296.1	−314.8
	Model 3	−173.6	−157.0	−136.9	−155.1	−349.4	−336.1	−316.9	−336.0
$\theta = 3\pi/4$	Model 1	−166.8	−146.1	−125.8	−147.4	−331.7	−310.4	−294.3	−311.2
	Model 2	−165.5	−146.8	−126.3	−147.7	−332.9	−313.1	−293.6	−312.1
	Model 3	−172.2	−150.2	−132.4	−154.5	−356.6	−336.8	−314.2	−337.3

Bibliography

Balaguer, A., Ruiz, L., Hermosilla, T., and Recio, J. (2010). Definition of a comprehensive set of texture semivariogram features and their evaluation for object-oriented image classification. *Computers & Geosciences*, 36(2):231–240.

Balaguer-Beser, A., Ruiz, L., Hermosilla, T., and Recio, J. (2013). Using semivariogram indices to analyse heterogeneity in spatial patterns in remotely sensed images. *Computers & Geosciences*, 115–127:231–240.

Bleck, R. (1975). An economical approach to the use of wind data in the optimum interpolation of geo- and Montgomery potential fields. *Monthly Weather Review*, 103(9):807—816.

Bonetto, R., Forlerer, E., and Ladaga, J. (2002). Texture characterization of digital images which have a periodicity or a quasi-periodicity. *Measurement Science and Technology*, 13(9):1458–1466.

Bosaz, L. B., Gerde, J. A., Borrás, L., Cipriotti, P. A., Ascheri, L., Campos, M., Gallo, S., and Rotundo, J. L. (2019). Management and environmental factors explaining soybean seed protein variability in central Argentina. *Field Crops Research*, 240:34–43.

Budd, D., Pranter, M., and Reza, Z. (2006). Lateral periodic variations in the petrophysical and geochemical properties of dolomite. *Geology*, 34(5):373—376.

Chang, Y.-C., Ching, J., Phoon, K.-K., and Yue, Q. (2021). On the hole effect in soil spatial variability. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 7(4):04021039.

Chen, X. (2005). Statistical and geostatistical features of streambed hydraulic conductivities in the Platte River, Nebraska. *Environmental Geology*, 48(6):693—701.

Ching, J., Uzielli, M., Phoon, K.-K., and Xu, X. (2023). Characterization of autocovariance parameters of detrended cone tip resistance from a global CPT database. *Journal of Geotechnical and Geoenvironmental Engineering*, 149(10):04023090.

Ciollaro, G. and Romano, N. (1995). Spatial variability of the hydraulic-properties of a volcanic soil. *Geoderma*, 65(3-4):263–282.

Cohen, W. B., Spies, T. A., and Bradshaw, G. A. (1990). Semivariograms of digital imagery for analysis of conifer canopy structure. *Remote Sensing of Environment*, 34(3):167–178.

Curran, P. J. (1988). The semivariogram in remote sensing: An introduction. *Remote Sensing of Environment*, 24(3):493–507.

Dong, X., Pinninti, R., Tvinnereim, A., Lowe, T., Di Paolo, D., and Shirvaikar, M. (2015). Stochastic predictors from the DXA scans of human lumbar vertebrae are correlated with the microarchitecture parameters of trabecular bone. *Journal of Biomechanics*, 48(12):2968–2975.

Emery, X. and Alegría, A. (2022). The Gauss hypergeometric covariance kernel for modeling second-order stationary random fields in Euclidean spaces: Its compact support, properties and spectral representation. *Stochastic Environmental Research and Risk Assessment*, 36:2819––2834.

Emery, X. and Parra, J. (2013). Integration of crosswell seismic data for simulating porosity in a heterogeneous carbonate aquifer. *Journal of Applied Geophysics*, 98:254–264.

Everett, R., Duffy, M., Storck, S., and Zupan, M. (2020). A variogram analysis of build height effects in an additively manufactured AlSi10Mg part. *Additive Manufacturing*, 35:101306.

Fiori, A., Janković, I., and Dagan, G. (2003). Flow and transport in highly heterogeneous formations: 2. Semianalytical results for isotropic media. *Water Resources Research*, 39(9):1269.

Guillobez, S. and Arnaud, M. (1998). Regionalized soil roughness indices. *Soil and Tillage Research*, 45(3-4):419–432.

Irvine-Fynn, T. D. L., Holt, T. O., James, T. D., Smith, M. W., Rutter, N., Porter, P. R., and Hodson, A. J. (2022). Time-lapse photogrammetry reveals hydrological controls of fine-scale High-Arctic glacier surface roughness evolution. *Earth Surface Processes and Landforms*, 47(6):1635–1652.

Jennings, J., Ruppel, S., and Ward, W. (2000). Geostatistical analysis of permeability data and modeling of fluid-flow effects in carbonate outcrops. *SPE Reservoir Evaluation & Engineering*, 3(4):292—303.

Jones, T. and Ma, Y. (2001). Geologic characteristics of hole-effect variograms calculated from lithology-indicator variables. *Mathematical Geology*, 33(5):615–629.

Jordan, G. (2003). Morphometric analysis and tectonic interpretation of digital terrain data: A case study. *Earth Surface Processes and Landforms*, 28(8):807–822.

Journel, A. and Froidevaux, R. (1982). Anisotropic hole-effect modeling. *Mathematical Geology*, 14(3):217–239.

Le Blévec, T., Dubrule, O., John, C. M., and Hampson, G. J. (2018). Geostatistical modelling of cyclic and rhythmic facies architectures. *Mathematical Geosciences*, 50:609–637.

Lefranc, M., Beaudoin, B., Chilès, J., Guillemot, D., Ravenne, C., and Trouiller, A. (2008). Geostatistical characterization of Callovo–Oxfordian clay variability from high-resolution log data. *Physics and Chemistry of the Earth*, 33(Suppl. 1):S2–S13.

Ma, C. (2005). Linear combinations of space-time covariance functions and variograms. *IEEE transactions on signal processing*, 53(3):857–864.

Mary-Huard, T., Daudin, J.-J., Robin, S., Bitton, F., Cabannes, E., and Hilson, P. (2004). Spotting effect in microarray experiments. *BMC Bioinformatics*, 5:63.

Matonti, C., Guglielmi, Y., Viseur, S., Bruna, P., Borgomano, J., Dahl, C., and Marié, L. (2015). Heterogeneities and diagenetic control on the spatial distribution of carbonate rocks acoustic properties at the outcrop scale. *Tectonophysics*, 638:94–111.

Parra, J. and Emery, X. (2013). Geostatistics applied to cross-well reflection seismic for imaging carbonate aquifers. *Journal of Applied Geophysics*, 92:68–75.

Pastor, J., Dewey, B., Moen, R., Mladenoff, D. J., White, M., and Cohen, Y. (1998). Spatial patterns in the moose-forest-soil ecosystem on Isle Royale, Michigan, USA. *Ecological Applications*, 8(2):411–424.

Pierson, F. B. and Wight, J. R. (1991). Variability of near-surface soil temperature on sagebrush rangeland. *Journal of Range Management*, 44(5):491–497.

Price, D. and Kozlowski, T. (2021). A multiscale approach to analyze the effect of radial void fraction distributions on BWR neutronic lattice parameters. *Annals of Nuclear Energy*, 152:107988.

Salamon, P., Fernández-Garcia, D., and Gómez-Hernández, J. (2007). Modeling tracer transport at the MADE site: The importance of heterogeneity. *Water Resources Research*, 43:W08404.

San Martín, C., Milne, A., Webster, R., Storkey, J., Andújar, D., Fernández-Quintanilla, C., and Dorado, J. (2018). Spatial analysis of digital imagery of weeds in a maize crop. *Quarterly Journal of the Royal Meteorological Society*, 7(2):61.

Sharifi, M., Messiga, A., Asefpour Vakilian, K., Stopford, E., and Hutchinson, T. (2020). Spatial distribution of soil phosphorous fractions following 1-year farrowing sows in an outdoor hog-rearing farm in Eastern Canada. *Environmental Monitoring and Assessment*, 192(5):322.

Thiébaux, H. (1976). Anisotropic correlation functions for objective analysis. *Monthly Weather Review*, 104:994—1002.

Thiébaux, H. (1985). On approximation to geopotential and windfield correlation structures. *Tellus A: Dynamic Meteorology and Oceanography*, 37(2):126—131.

Varbla, S. and Ellmann, A. (2023). Iterative data assimilation approach for the refinement of marine geoid models using sea surface height and dynamic topography datasets. *Journal of Geodesy*, 97(3):24.

Webster, R. (1977). Spectral analysis of gilgai soil. *Australian Journal of Soil Research*, 15(3):191–204.

Yang, Y. and Shao, C. (2018). Spatial interpolation for periodic surfaces in manufacturing using a Bessel additive variogram model. *Journal of Manufacturing Science and Engineering*, 140(6):061001.