Indicator Variograms as an Aid for Geological Interpretation and Modeling of Ore Deposits

Mohammad Maleki 1,2, Xavier Emery 1,2,* and Nadia Mery 1

1 Department of Mining Engineering, University of Chile, Santiago 8370448, Chile; mohammad.maleki@ing.uchile.cl (M.M.); nmery@ing.uchile.cl (N.M.)
2 Advanced Mining Technology Center, University of Chile, Santiago 8370448, Chile
*
Correspondence: xemery@ing.uchile.cl; Tel.: +56-2-2978-4477

Received: 1 November 2017; Accepted: 30 November 2017; Published: 5 December 2017

Abstract: Geostatistics offers a set of methods for modeling, predicting, or simulating geological domains in space. In addition of being an input of some of these methods, indicator direct and cross-variograms convey valuable information on the geometry of the domain layouts and on their contact relationships, in particular, on the surface area of a domain boundary, on the surface area of the contact between two domains, on the propensity for a domain to be in contact with, or separated from, another domain, and on the minimum and maximum distances between points from two domains. Accordingly, the indicator variograms inferred from sparse sampling data can be used to determine whether or not an interpreted model of the subsurface is consistent with the sampling information. The previous concepts are illustrated through a case study corresponding to a porphyry copper deposit.

Keywords: geological domaining; indicator direct and cross-variograms; edge effects; contact relationships

1. Introduction

Identifying and modeling rock type, mineralization and alteration domains in ore deposits has become a crucial step in the assessment of geological and geometallurgical regionalized variables, such as metal grades or metal recoveries, due to the controls that these domains often exert on the distribution of these variables [1–5].

One issue for mining geologists and geometallurgists is to define the layout of each domain with the greatest possible accuracy, in order to minimize misclassifications (material predicted as belonging to a domain, but actually belonging to another one). Another issue relates to the consistency between the interpreted domain layouts and the information obtained from sampling data, so as to determine whether or not the geological interpretation is a faithful representation of reality. Of particular interest are the spatial regularity of the domain layouts and the contact relationships between domains, as some domains may have a propensity to be or not to be in contact depending on the geological setting.

The appraisal of the spatial and geometrical properties of geological domains often relies on geostatistical tools like vertical or regionalized proportion curves [6,7], contact matrices [8], transition probabilities under Markov chain models [9,10], or multiple-point metrics [11,12]. This paper focuses on alternative tools (namely, the indicator direct and cross-variograms) that are often disregarded in the geological modeling stage, although they can easily be inferred from a set of sampling data. The aim is to show that these tools convey geometrical information about the domain layouts and can be used to validate the consistency between sampling data (geological logs) and an interpreted geological model.

The outline is the following one: Section 2 provides theoretical background about the direct and cross-variograms of the indicators of non-overlapping domains and links the variogram properties...
with geometrical features of the domains. Section 3 presents a case study of a porphyry copper deposit. A general discussion and conclusions follow in Sections 4 and 5, respectively.

2. Materials and Methods

2.1. Indicator Modeling

A geological model can be viewed as a partition of a mineral deposit or, more generally, of a region in the subsurface, into non-overlapping domains, such as rock type, mineralization or alteration domains. Each domain \( D_k \), with \( k = 1, \ldots, K \), can be codified by a binary variable called indicator:

\[
I_k(x) = \begin{cases} 
1, & \text{if } x \text{ belongs to } D_k \\
0, & \text{otherwise} 
\end{cases}
\]  

(1)

where \( x \) stands for a generic point in the three-dimensional Euclidean space \( \mathbb{R}^3 \). Viewing this indicator as a spatial random field in \( \mathbb{R}^3 \) opens the way for the use of the geostatistical formalism, which allows for prediction or simulation of the indicator at any point based on the values observed at surrounding sampling points.

The expected value of \( I_k(x) \) represents the probability of finding domain \( D_k \) at point \( x \):

\[
p_k(x) = E\{I_k(x)\} = P\{x \in D_k\}
\]

(2)

where \( E \) and \( P \) denote the expected value and the probability measure associated with the random field definition, respectively. Provided that the indicator random field is first-order stationary, \( p_k \) does not depend on \( x \) and stands for the proportion of space covered by domain \( D_k \). First-order stationarity means that the probability of finding domain \( D_k \) is constant over space.

2.2. Theoretical Indicator Direct Variogram

A tool commonly used in geostatistical modeling for quantifying the spatial continuity of domain \( D_k \) is the direct variogram of the indicator random field \( I_k \), defined for each pair of points \((x, x')\) as:

\[
\gamma_k(x, x') = \frac{1}{2} \text{var}\{I_k(x) - I_k(x')\}
\]

(3)

where \( \text{var} \) denotes the variance. Under the assumption of first-order stationarity, one has [13–15]:

\[
\gamma_k(x, x') = \frac{1}{2} E\{|I_k(x) - I_k(x')|^2\} = \frac{1}{2} [P\{x \in D_k, x' \notin D_k\} + P\{x \notin D_k, x' \in D_k\}]
\]

(4)

Furthermore, \( P\{x \in D_k, x' \notin D_k\} = p_k - P\{x \in D_k, x' \in D_k\} = P\{x \notin D_k, x' \in D_k\} \), so that one obtains:

\[
\gamma_k(x, x') = P\{x \in D_k, x' \notin D_k\} = P\{x \notin D_k, x' \in D_k\}
\]

(5)

Provided that the indicator random field \( I_k \) is second-order stationary, \( \gamma_k \) only depends on the separation vector between \( x \) and \( x' \) and is commonly denoted as \( \gamma_k(h) \) with \( h = x - x' \). The indicator direct variogram \( \gamma_k(h) \) represents the probability of transitioning from inside \( D_k \) at point \( x \) to outside \( D_k \) at point \( x' = x - h \) (equivalently, from outside \( D_k \) at point \( x \) to inside \( D_k \) at point \( x - h \)), the transition happening somewhere over the extent of the separation vector \( h \).

When the norm of \( x - x' \) becomes infinitely large, the indicator variables \( I_k(x) \) and \( I_k(x') \) are independent, so that \( P\{x \in D_k, x' \notin D_k\} = P\{x \in D_k\} P\{x' \notin D_k\} = p_k(1 - p_k) \). For separation
vectors with a large norm, the indicator direct variogram presents a sill, whose value relates to the proportion of space ($p_k$) covered by domain $D_k$:

$$\gamma_k(h) \xrightarrow{\|h\| \to +\infty} p_k(1 - p_k)$$

(6)

The distance $\|h\|$ at which the indicator direct variogram reaches the sill (correlation range), corresponds to the distance beyond which the two indicator variables $I_k(x)$ and $I_k(x + h)$ are independent. This distance may depend on the direction of $h$ and does not relate to the size of domain $D_k$ [9] (under the first-order stationarity assumption, $D_k$ has a constant probability of occurrence over $R^3$, so it is unbounded) nor to the size of its convex parts [16] (based on Equation (5), $D_k$ has the same indicator direct variogram as its complement, whose convex parts differ from that of $D_k$).

More information on the geometry of domain $D_k$ can be obtained from the indicator direct variogram, in particular, from its behavior at the origin. The first-order derivative conveys information about the regularity of the boundary of $D_k$ [13,17]:

$$\frac{1}{\pi} \int_{S^3} \gamma_k^{(u)}(0) \, du = E\{|\partial D_k|\}$$

(7)

where $E\{|\partial D_k|\}$ is the expected surface area of the boundary of $D_k$ per unit volume of $R^3$, $S_3$ is the unit sphere of $R^3$, $u$ a unit-length vector and $\gamma_k^{(u)}(0)$ is the first-order derivative of the variogram at the origin in the direction of vector $u$:

$$\gamma_k^{(u)}(0) = \lim_{\varepsilon \to 0} \frac{P\{x \in D_k, x + \varepsilon u \notin D_k\}}{\varepsilon} = \lim_{\varepsilon \to 0} \frac{P\{x \notin D_k, x + \varepsilon u \in D_k\}}{\varepsilon}$$

(8)

Accordingly, unless $D_k$ is void ($p_k = 0$) or fills the whole space $R^3$ ($p_k = 1$), a null derivative (indicator direct variogram with a zero slope at the origin) in every direction cannot be observed [18]. A finite non-zero derivative means that the boundary of $D_k$ is almost always regular (Figure 1A,B), while an infinite derivative means that $D_k$ has an irregular or fractal boundary (Figure 1C,D). In the former case, the second-order derivative of the indicator direct variogram at the origin in the direction of vector $u$ relates to the existence of singular points along the boundary of $D_k$ [19]:

$$\gamma_k^{(u)}(0) = -\lim_{\varepsilon \to 0} \frac{P\{x - \varepsilon u \in D_k, x + \varepsilon u \notin D_k\}}{\varepsilon^2} = -\lim_{\varepsilon \to 0} \frac{P\{x - \varepsilon u \notin D_k, x + \varepsilon u \in D_k\}}{\varepsilon^2}$$

(9)

If the boundary has singular points (such as angular points or “vertices”), then the second-order derivative is negative and the indicator direct variogram is strictly concave near the origin (Figure 2A,B). On the contrary, if any point of the boundary has a single tangent plane (i.e., if $D_k$ has a smooth boundary everywhere), then the second-order derivative of the indicator direct variogram at the origin is zero, which means that the variogram is straight at the origin (Figure 2C,D).

Finally note that the properties of the indicator direct variogram (correlation range, slope at the origin, concavity at the origin and overall shape) can depend on the direction of space under consideration, which corresponds to an anisotropic behavior of the indicator random field $I_k$. In such a case, the geometric properties of domain $D_k$ and of its boundary are not the same in all the directions of space.
2.3. Experimental Indicator Direct Variogram

Given a set of points \( \{x_\alpha : \alpha = 1, \ldots, n\} \) where the indicator is known, the experimental indicator direct variogram is defined as:

\[
\hat{\gamma}_k(h) = \frac{1}{2 \text{card}\{N^+(h)\}} \sum_{(\alpha, \beta) \in N^+(h)} \left[ I_k(x_\alpha) - I_k(x_\beta) \right]^2
\]  \hspace{1cm} (10)

where \( N^+(h) = \{(\alpha, \beta) \in \{1, \ldots, n\}^2 : x_\alpha - x_\beta = h \text{ up to distance and angle tolerances}\} \) and \( \text{card}\{N^+(h)\} \) is the cardinality of \( N^+(h) \). Under a second-order stationarity assumption, this is an unbiased estimator of the indicator direct variogram \( \gamma_k(h) \), so that its properties (behavior at large distances, behavior near the origin, anisotropy) should be the same, up to statistical fluctuations.

An experimental variogram that does not converge to a positive sill may incite one to revise the second-order stationarity assumption. This situation happens when the probability of occurrence of \( D_k \) is not constant in space (presence of a systematic trend or zonation, which can be explained by geological considerations). In such a situation, one may assume that the second-order stationarity assumption is valid only at small scales (local stationarity), and restrict the interpretation of the experimental variogram to small distances. In particular, the regularity and smoothness of the domain boundary can still be inferred from the behavior of the variogram near the origin.
2.4. Theoretical Indicator Cross-Variogram

Consider now two domains \( D_k \) and \( D_{k'} \) that do not overlap. The joint behavior of these two domains can be modeled by the cross-variogram of the associated indicator fields \( I_k \) and \( I_{k'} \):

\[
\gamma_{k,k'}(x, x') = \frac{1}{2} \text{cov}\{I_k(x) - I_k(x'), I_{k'}(x) - I_{k'}(x')\}
\]

where \( \text{cov} \) denotes the covariance. Under the assumption of first-order stationarity, one has [14–16]:

\[
\gamma_{k,k'}(x, x') = \frac{1}{2} E\{|I_k(x) - I_k(x')||I_{k'}(x) - I_{k'}(x')|\}
= -\frac{1}{2} \left[ P\{x \in D_k, x' \in D_{k'}\} + P\{x \in D_{k'}, x' \in D_k\} \right]
= -\frac{1}{2} P\{x \in D_k, x' \in D_{k'}\} \text{ or } (x \in D_{k'}, x' \in D_k) \]

Figure 2. Plan views of geological domains (A,C) and indicator variograms of domain \( D_1 \) (B,D).

The smoothness of the domain boundary is reflected in the second-order derivative of the indicator direct variogram at the origin: domain boundary with vertices highlighted in red (A) associated with a strictly concave variogram near the origin (B); smooth domain boundary (C) associated with a straight variogram near the origin (D).

The indicator cross-variogram is negative and represents, up to the sign, the probability of transitioning from \( D_k \) to \( D_{k'} \) (or vice-versa, from \( D_{k'} \) to \( D_k \)) between two points separated by vector \( h \). If the two indicator random fields \( I_k \) and \( I_{k'} \) are jointly second-order stationary, \( \gamma_{k,k'} \) only depends on \( h = x - x' \) and is commonly denoted as \( \gamma_{k,k'}(h) \).
When the norm of \( x - x' \) becomes infinitely large, the indicator variables \( I_k(x) \) and \( I_{k'}(x') \) are independent of \( I_k(x') \) and \( I_{k'}(x') \), so that [14]:

\[
\gamma_{k,k'}(h) \xrightarrow{\|h\| \to +\infty} \frac{P\{x \in D_k\}P\{x + h \in D_{k'}\} + P\{x \in D_{k'}\}P\{x + h \in D_k\}}{2} = -p_k p_{k'} \tag{13}
\]

The indicator cross-variogram presents a sill whose value relates to the proportions of space covered by domains \( D_k \) and \( D_{k'} \). The distance \( \|h\| \) at which the sill is reached (correlation range) corresponds to the distance beyond which the indicator variables \( I_k(x) \) and \( I_{k'}(x) \) are independent of \( I_k(x + h) \) and \( I_{k'}(x + h) \). This distance may differ from the correlation ranges of the indicator direct variograms and may also depend on the direction of \( h \). Note that the two domains \( D_k \) and \( D_{k'} \) cannot be independent, insofar as they are assumed not to overlap; as a consequence, the cross-variogram cannot be identically zero, which is corroborated by the existence of a negative sill (Equation (13)), unless one of the domains is void (\( p_k = 0 \) or \( p_{k'} = 0 \)).

The behavior at the origin of the indicator cross-variogram conveys information about the boundary common to both domains \( D_k \) and \( D_{k'} \). In particular, one has:

\[
\frac{1}{\pi} \int_{S_3} \gamma_{k,k'}^{(u)}(0) \, du = -E\{ |\partial D_k \cap \partial D_{k'}| \} \tag{14}
\]

where \( E\{ |\partial D_k \cap \partial D_{k'}| \} \) is the expected surface area of the boundary common to \( D_k \) and \( D_{k'} \) per unit volume of \( \mathbb{R}^3 \) and \( \gamma_{k,k'}^{(u)}(0) \) is the first-order derivative of the indicator cross-variogram at the origin in the direction of vector \( u \):

\[
\gamma_{k,k'}^{(u)}(0) = -\lim_{\varepsilon \to 0} \frac{P\{x \in D_k, x + \varepsilon u \cap D_{k'}\} + P\{x \in D_{k'}, x + \varepsilon u \cap D_k\}}{2\varepsilon} \tag{15}
\]

A null derivative (indicator cross-variogram with a zero slope at the origin) in all directions indicates that \( D_k \) and \( D_{k'} \) have no contact (Figure 3A,B). Even more, the distance up to which the cross-variogram is equal to zero in a given direction represents the minimum separation distance between the two domains \( D_k \) and \( D_{k'} \) along this direction. In contrast, an indicator cross-variogram with a negative derivative (negative slope) at the origin indicates the existence of contacts between \( D_k \) and \( D_{k'} \); the greater the slope, the more contact area exists (Figure 3C–F).

### 2.5. Experimental Indicator Cross-Variogram

Given a set of points \( \{x_a : a = 1, \ldots, n\} \) where the indicators are known, the experimental indicator cross-variogram is defined as:

\[
\hat{\gamma}_{k,k'}(h) = \frac{1}{2\text{card}\{N^+(h)\}} \sum_{(a,b) \in N^+(h)} [I_k(x_a) - I_k(x_b)] [I_{k'}(x_a) - I_{k'}(x_b)] \tag{16}
\]

with the same notation as in Section 2.3. Under a second-order stationarity assumption, this is an unbiased estimator of the indicator cross-variogram \( \gamma_{k,k'}(h) \) and should therefore exhibit the same properties, up to statistical fluctuations.

An experimental indicator cross-variogram that does not converge to a negative sill may incite one to revise the second-order stationarity assumption. For instance, if \( D_k \) and \( D_{k'} \) are bounded, the experimental indicator cross-variogram will be zero for distances larger than the maximum separation distance between points from the two domains (all the terms in the summation of Equation (16) are equal to zero). As stated in Section 2.3, one may then assume that the second-order stationarity assumption is valid only locally and restrict the interpretation of the experimental indicator cross-variogram to small distances. In particular, the slope at the origin reflects the amount of contact area between \( D_k \) and \( D_{k'} \).
Figure 3. Plan views of geological domains (A,C,E) and indicator cross-variograms of domains $D_1$ and $D_2$ (B,D,F). The surface area of the contact between domains $D_1$ and $D_2$ is reflected in the first-order derivative of the indicator cross-variogram at the origin: no contact between domains (A) associated with a zero cross-variogram slope at the origin (B); existence of a contact with a small surface area (C) associated with a small negative cross-variogram slope at the origin (D); existence of a contact with a large surface area (E) associated with a large negative cross-variogram slope at the origin (F).
2.6. Link between Indicator Direct and Cross-Variograms

Due to the compositional nature of the indicator random fields, the sum of the indicator cross-variograms \( \gamma_{k,k'} \) for \( k' \neq k \) is equal to the opposite of the indicator direct variogram \( \gamma_k \), which can be demonstrated using Equations (5) and (12):

\[
\sum_{k' \neq k} \gamma_{k,k'}(x,x') = -\frac{1}{2} \left[ \sum_{k' \neq k} P\{x \in D_k, x' \in D_{k'}\} + \sum_{k' \neq k} P\{x \in D_{k'}, x' \in D_k\} \right] \\
= -\frac{1}{2} \left[ P\{x \in D_k, x' \notin D_k\} + P\{x \notin D_k, x' \in D_k\} \right] \\
= -\gamma_k(x,x')
\]  

Equation (17)

2.7. Ratio of Indicator Cross-to-Direct Variograms and Edge Effects

Under the assumption of first-order stationarity, based on Equations (5) and (12), the cross-to-direct variogram ratio corresponds, up to the sign, to the probability that a point \( x' \) when leaving \( D_k \) at point \( x \) or, vice-versa, of reaching domain \( D_{k'} \) at point \( x \) when leaving \( D_k \) at point \( x' \):

\[
\frac{\gamma_{k,k'}(x,x')}{\gamma_k(x,x')} = -\frac{1}{2} \left[ \frac{P\{x \in D_k, x' \in D_{k'}\}}{P\{x \in D_k \}} + \frac{P\{x \in D_{k'}, x' \in D_k\}}{P\{x \notin D_k \}} \right] \\
= -\frac{1}{2} \frac{P\{x' \in D_{k'}|x \in D_k\}}{P\{x \notin D_k\}} + P\{x \notin D_k|D_k \}
\]  

Equation (18)

If the indicator random fields are second-order stationary, this ratio can be rewritten as:

\[
\frac{\gamma_{k,k'}(h)}{\gamma_k(h)} = -P\{x + \delta h \in D_{k'}|x \in D_k\} \\
\]  

Equation (19)

with \( \delta = 1 \) or \(-1\) with equal probability 0.5. The sill of this ratio is (Equations (6) and (13)) \[15\]:

\[
\frac{\gamma_{k,k'}(h)}{\gamma_k(h)} \xrightarrow{|h| \to \infty} -P\{x' \in D_{k'}|x' \notin D_k\} = -\frac{pk'}{1-p_k}
\]  

Equation (20)

and corresponds, up to the sign, to the probability that a point \( x' \) belongs to \( D_{k'} \) knowing that it does not belong to \( D_k \), which is equal to the proportion of domain \( D_{k'} \) relative to the proportion of what is not \( D_k \). Such a sill value can be used as a reference to determine whether or not the transition from domain \( D_k \) to domain \( D_{k'} \) between two points separated by vector \( h \) or \(-h\) is preferential, a situation known as an edge effect or a border effect \[14,20\]. In particular, for small separation vectors (\( h = \varepsilon u \) with \( u \) a unit vector and \( \varepsilon \) a scalar tending to zero), the following cases may be found (Figure 4):

- \( \frac{\gamma_{k,k'}(\varepsilon u)}{\gamma_k(\varepsilon u)} < \frac{p_{k'}}{1-p_k} \): due to Equations (19) and (20), \( P\{x' \in D_{k'}|x \in D_k, x' \notin D_k\} < \frac{p_{k'}}{1-p_k} \) where \( x' = x + \delta \varepsilon u \) is a point located close to \( x \). Knowing that \( x \) belongs to \( D_k \) decreases the probability that \( x' \) belongs to \( D_{k'} \), given that it does not belong to \( D_k \). This means that the contact area between \( D_k \) and \( D_{k'} \) is smaller than it would be in the absence of edge effect, i.e., there is a propensity for both domains not to be in contact (Figure 4A,B).

- \( \frac{\gamma_{k,k'}(\varepsilon u)}{\gamma_k(\varepsilon u)} = \frac{p_{k'}}{1-p_k} \): the contact area between domains \( D_k \) and \( D_{k'} \) coincides with the expected contact area in the absence of edge effect (no preferential contact) (Figure 4C,D).

- \( \frac{\gamma_{k,k'}(\varepsilon u)}{\gamma_k(\varepsilon u)} > \frac{p_{k'}}{1-p_k} \): the contact area between \( D_k \) and \( D_{k'} \) is larger than it would be in the absence of edge effect. There is a propensity for both domains to be in contact together (Figure 4E,F).

Because the indicator cross-to-direct variogram ratios \( \gamma_{k,k'}/\gamma_k \) for \( k' \neq k \) add to \(-1\) (Equation (17)), the existence of an edge effect between domains \( D_k \) and \( D_{k'} \) necessarily implies the existence of an opposite edge effect between \( D_k \) and \( D_{k''} \) for some \( k'' \in [1, \ldots, n] \).
Because the indicator cross-to-direct variogram ratios $\gamma_k / \gamma'_k$, for $k \neq k'$ add to $-1$ (Equation (17)), the existence of an edge effect between domains $D_k$ and $D_{k'}$ necessarily implies the existence of an opposite edge effect between $D_k$ and $D_{k''}$ for some $k'' \in \{1, \ldots, n\}$.

Figure 4. Plan views of geological domains ($A, C, E$) and absolute values of the ratios of indicator cross-to-direct variograms $\gamma_{1,2} / \gamma_1$ (red curves) ($B, D, F$); the expected sill value (Equation (20)) is represented by a black solid line.

3. Case Study and Results

3.1. Presentation of the Deposit

The previous concepts are now illustrated through a case study corresponding to a porphyry copper deposit. For confidentiality reasons, the name and location of the deposit are not disclosed.
The host rock is composed of a sequence of acidic and andesite volcanic rocks with intercalations of sedimentary rocks. Six main rock types have been recognized in the region under study (Figure 5):

- Andesite (code 1): this volcanic rock is distributed as an interbedded sequence spread over the whole deposit and consists of lavas, auto-breccias and ocoites, with porphyritic to aphanitic texture and, locally, with amygdaloidal to ocoitic texture.
- Other volcanic rocks (code 2): this is a sequence of differentiated volcanic rocks with textures that vary from porphyritic to tuffaceous with an aphanitic matrix.
- Sedimentary rocks (code 3): this rock type includes undifferentiated, volcaniclastic and calcareous sediments, which are mostly interbedded with dacite and andesite, forming the host rock of the intrusive porphyry.
- Granodioritic porphyry (code 4): this is a medium grain granodiorite of Triassic age.
- Monzonitic porphyry (code 5): this porphyry rock is associated with the hypogene mineralization.
- Hydrothermal breccia (code 6) with tourmaline, quartz, chalcopyrite and pyrite cements, which can be found in the contact between the monzonitic porphyry with the host rock as a breccia with intrusive fragments.

![Figure 5](image_url). Interpreted rock type model in a cross-section under study. Sample drill hole data have been superimposed. Local coordinates are used to preserve the confidentiality of the data.

3.2. Data Presentation and Pre-Processing

The deposit has been recognized by a set of exploration diamond drill holes, along which the rock types have been logged. In the following, we will focus on a vertical cross-section in which 9060 data points are available, with information on the spatial coordinates and prevailing rock types of core samples. In addition to these data, a rock type model interpreted by the expert geologists is available (Figure 5).

In the forthcoming sections, the indicator direct and cross-variograms will be calculated, and geometrical information on the rock type layout will be extracted. To investigate the consistency of interpreted rock type model, the information obtained from the variograms calculated with the drill hole data will be compared with the information obtained from variograms calculated from the interpreted model.

To avoid misinterpretations, the experimental variograms calculated from the drill hole data are declustered to account for the irregular sampling pattern of these data [21–23]. To this end, each data is assigned a declustering weight equal to its area of influence within the corresponding rock type layout in the interpreted geological model. The weight assigned to a data pair in experimental variogram calculations corresponds to the product of the two data weights [21]. This way, the declustered mean (domain proportion) and variance (variogram sill) of each rock type indicator inferred from the drill hole data coincide with the mean and variance inferred from the interpreted model, facilitating
the comparison between the indicator variograms calculated from the drill hole data and from the interpreted model.

3.3. Indicator Direct Variograms

For each rock type, the indicator direct variogram is calculated from the drill hole data and from the interpreted rock type model, along the directions experimentally recognized as the main directions of anisotropy (these are found to be the directions dipping 30° and −60° for rock types 1, 2 and 3, and the directions dipping −15° and 75° for rock types 4, 5 and 6, with the dip being measured positively upwards from the north direction). As an illustration, the direct variograms for rock types 1 (andesite) and 3 (sedimentary rocks) are displayed in Figures 6 and 7, respectively. The following comments and interpretations can be made.

**Figure 6.** Indicator direct variogram of rock type 1 (andesite) along directions dipping 30° (blue) and −60° (red) with respect to the north direction, calculated from (A) drill hole data, (B) interpreted model. Calculations consider an angle tolerance of 20° around the target directions. The same lag values have been used for both variograms.

**Figure 7.** Indicator direct variogram of rock type 3 (sedimentary rocks) along directions dipping 30° (blue) and −60° (red) with respect to the north direction, calculated from (A) drill hole data, (B) interpreted model. Calculations consider an angle tolerance of 20° around the target directions. The same lag values have been used for both variograms.
3.3.1. Existence of a Sill

For both the drill hole data (Figure 6A) and the interpreted model (Figure 6B), the experimental indicator variogram of rock type 1 (andesite) does not converge to a sill, which makes the stationarity assumption questionable for this rock type. This is due to the fact that andesite mainly appears in the center of the region under study, so that its proportion is not constant in space. In contrast, the experimental indicator variogram of rock type 3 (sedimentary rocks) fluctuates around a sill (Figure 7), which agrees with the stationarity assumption at the working scale (2000–2500 m along the direction dipping 30° and 700–900 m along the direction dipping −60°). This assumption is reasonable, as the occurrences of sedimentary rocks are intermingled with that of the other rock types throughout the region under study. The variogram sill (0.16) is related to the proportion of the region covered by sedimentary rocks (0.20) (Equation (6)).

3.3.2. Shape and Anisotropy

Figures 6 and 7 show that there is more spatial continuity along the direction dipping 30° than along the direction dipping −60°. The shapes of the experimental variograms calculated from the drill hole data (Figures 6A and 7A) are comparable to those of the experimental variograms calculated from the interpreted geological model (Figures 6B and 7B). In terms of spatial continuity, the interpreted model is therefore consistent with the drill hole data. As an exception, at short distances (0–300 m) along the direction dipping −60°, the experimental variograms calculated from the drill hole data increase faster than the experimental variograms calculated from the interpreted model, but this discrepancy can be provoked by statistical fluctuations and may not be significant.

3.3.3. Slope at the Origin

The experimental indicator variograms calculated from the drill hole data (Figures 6A and 7A) have a large (almost infinite) slope at the origin; the same feature can be observed along other directions (not displayed in Figures 6 and 7). This indicates that the boundaries of both andesite (rock type 1) and sedimentary rocks (rock type 3) are irregular (Equation (7)). The same indicator variograms calculated from the interpreted model (Figures 6B and 7B) exhibit a more moderate slope at the origin and lower values at distances less than 100 m, suggesting slightly more regular domain boundaries and short-scale behavior. However, as mentioned above, the differences are deemed not significant in comparison with expected statistical fluctuations caused by the limited number of drill hole data, their irregular sampling design, and the possible presence of errors in the data (mislogged rock types in some drill hole samples). Several analytical and numerical approaches have been proposed to quantify experimental variogram fluctuations, based on resampling techniques or on the fitting of a theoretical random field model [24–27], which is out of the scope of this work.

3.4. Indicator Cross-Variograms

Let us now examine a few indicator cross-variograms calculated from the drill hole data (Figures 8A, 9A and 10A) and from the interpreted rock type model (Figures 8B, 9B and 10B) along the directions recognized as the main directions of anisotropy (directions dipping 30° and −60° for rock types 2 and 3; directions dipping −15° and 75° for rock types 4, 5 and 6).

3.4.1. Behavior at Large Distances

In case of joint second-order stationarity, the indicator cross-variogram should converge to a sill equal to minus the product of the rock type proportions (Equation (13)), but this behavior is not present in any of the three examples shown in Figures 8–10. Instead, the presented cross-variograms tend to zero at large distances. As mentioned in Section 2.5, the distance at which an experimental indicator cross-variogram reaches a zero sill defines the maximum separation distance between points from the two associated rock types. It is deduced that points located in the other volcanic rocks (rock type 2)
and sedimentary rocks (rock type 3) are distant no more than 1000 m along the direction dipping $-60^\circ$, but can be distant more than 2000 m along the direction dipping $30^\circ$ (Figure 8). Also, points located in granodioritic porphyry (rock type 4) are separated from points located in monzonitic porphyry (rock type 5) less than 800 m along the direction dipping $75^\circ$ and 1200 m along the direction dipping $-15^\circ$ (Figure 9). Likewise, points located in granodioritic porphyry (rock type 4) are separated from points located in hydrothermal breccia (rock type 6) less than 1800 m along the direction dipping $-15^\circ$ (the two rock types are never found one above the other along the direction dipping $75^\circ$, as the corresponding cross-variogram is identically zero) (Figure 10).

Figure 8. Indicator cross-variogram between rock types 2 (other volcanic rocks) and 3 (sedimentary rocks) along directions dipping $30^\circ$ (blue) and $-60^\circ$ (red) with respect to the north direction, calculated from (A) drill hole data, (B) interpreted model. Calculations consider an angle tolerance of $20^\circ$ around the target directions. The same lag values have been used for both cross-variograms.

Figure 9. Indicator cross-variogram between rock types 4 (granodioritic porphyry) and 5 (monzonitic porphyry) along directions dipping $75^\circ$ (blue) and $-15^\circ$ (red) with respect to the north direction, calculated from (A) drill hole data, (B) interpreted model. Calculations consider an angle tolerance of $20^\circ$ around the target directions. The same lag values have been used for both cross-variograms.
3.4.2. Behavior at Short Distances

Since the associated indicator cross-variograms have a large negative slope at the origin, the other volcanic (rock type 2) and sedimentary (rock type 3) rocks on the one hand (Figure 8), and the granodioritic (rock type 4) and monzonitic (rock type 5) porphyries on the other hand (Figure 9) are in contact along the two anisotropy directions. In both cases, up to expected statistical fluctuations, the cross-variograms calculated from the interpreted model are comparable to the cross-variograms calculated from the drill hole data at short distances (0 to 300 m). Therefore, there is no evidence that the amounts of contacts between the rock type domains in the interpreted model are over- or underestimated with respect to reality: the interpreted model agrees with the sampling information in this aspect.

In contrast, it is observed that the indicator cross-variogram between granodioritic porphyry (rock type 4) and hydrothermal breccia (rock type 6) is zero until a distance of 1000 m (according to drill hole data, Figure 10A) or 500 m (according to the interpreted model, Figure 10B) along the direction dipping $-15^\circ$. These figures indicate the minimum distance between both rock types along the mentioned direction, being twice larger in the drill hole data. A revision of the interpreted geological model is therefore advisable not to underestimate the true separation distance between granodioritic porphyry and hydrothermal breccia.

3.5. Ratios of Indicator Cross-to-Direct Variograms

We now focus on the ratios of indicator cross-to-direct variograms (Equation (19)). As an illustration, Figures 11 and 12 show these ratios for rock types 2 (other volcanic rocks) and 3 (sedimentary rocks). From these two examples, one concludes that, at short distances, there is an edge effect between other volcanic and sedimentary rocks, with both domains being prone to be in contact, which can be corroborated by a look at the cross-section displayed in Figure 5. The edge effect is stronger when leaving the other volcanic rocks to reach the sedimentary rocks than in the opposite way, as the ratio plotted in Figure 11 is only a little higher than the expected sill, while the ratio in Figure 12 is significantly higher than the expected sill (one mostly finds rock type 3 when leaving rock type 2, while one mostly finds both rock types 1 and 2 when leaving rock type 3). Such edge effects can be explained by lithostratigraphic considerations, as the region is characterized by a volcano-sedimentary sequence consisting of volcanic rocks (mainly rhyolitic and dacitic ash-flow tuffs, pyroclastic rocks,
as well as basaltic and andesitic lavas) followed by sedimentary intercalations, which imply spatial alternations between rock type 3 (sedimentary rocks) and rock types 1 and 2 (andesite and other volcanic rocks, respectively).

**Figure 11.** Absolute value of the ratio of indicator cross-to-direct variograms between rock types 2 (other volcanic rocks) and 3 (sedimentary rocks) along directions dipping 30° (blue) and −60° (red) with respect to the north direction, calculated from (A) drill hole data, (B) interpreted model. Calculations consider an angle tolerance of 20° around the target directions. The same lag values have been used for both variogram ratios. The expected sill value (Equation (20)) is represented by a black solid line.

**Figure 12.** Absolute value of the ratio of indicator cross-to-direct variograms between rock types 3 (sedimentary rocks) and 2 (other volcanic rocks) along directions dipping 30° (blue) and −60° (red) with respect to the north direction, calculated from (A) drill hole data, (B) interpreted model. Calculations consider an angle tolerance of 20° around the target directions. The same lag values have been used for both variogram ratios. The expected sill value (Equation (20)) is represented by a black solid line.

4. **Discussion**

4.1. **Geological Interpretation and Modeling**

A first axis for discussion concerns the geological domaining of ore deposits and the modeling of the domain layouts for further use in the evaluation of mineral resources and ore reserves or in the quantification of geological uncertainty. As shown in the previous sections, the indicator direct and cross-v variograms give a profound insight into the spatial and geometrical properties of geological
domains and into their contact relationships, since they convey information about the regularity and smoothness of the domain boundaries, the contact area between two domains, and their propensity to be in contact with, or separated from, each other at short distances.

Besides the aforementioned properties, one aspect not captured by indicator variograms is the presence of a directional edge effect, when the transition from a domain to another one preferentially occurs in one way rather than in the opposite way (for instance, an oxide mineralogical domain located above a sulfide domain, while the reverse does not happen). To investigate such an edge effect, it is necessary to resort to the indicator cross-covariance (Appendix A), which is not an even function, unlike the indicator cross-variogram.

Indicator variograms or covariances therefore appear as a helpful tool to determine whether or not an interpreted geological model is consistent with sampling data. For instance, a hand-drawn interpreted model that is intended to be a locally accurate representation of reality is likely to exhibit over-smoothed domain boundaries or to exaggerate the domain spatial continuity, a situation that should be avoided if the interpreted model is used as a reference (training image) for constructing multiple realizations of the geological domain layouts.

4.2. Geostatistical Modeling and Simulation

A second axis for discussion concerns the spatial prediction or simulation of geological domains for modeling the subsurface and quantifying geological uncertainty. Two types of geostatistical approaches have been developed for such a purpose. The first one consists of model-driven approaches, which rely on the definition of a categorical random field and the specification of the parameters characterizing its spatial distribution; the reader is referred to [13] for a review of the most widespread model-driven approaches. The second type is the family of data-driven approaches. These approaches do not completely specify the random field, but rely on the fitting of some of its parameters directly from the available data.

The latter approaches are likely to suffer from consistency problems in the definition of the model parameters. As a first example, indicator-based approaches such as indicator kriging, indicator cokriging or sequential indicator simulation [28] rely on the fitting of indicator variogram models. However, as seen in Sections 2 and 3, the behaviors of the direct and cross-variograms near the origin is likely to be very different, with a possible zero slope for the cross-variograms, a situation that is prohibited for the direct variograms. The anisotropy directions and the behavior at large distances may also not be the same for the indicator direct and cross-variograms. In this context, the widely used linear model of coregionalization, based on the fitting of the direct and cross-variograms with the same set of basic nested structures [28,29], is inadequate. So is the separate modeling of the indicator direct variograms, as used in indicator kriging and in sequential indicator simulation, because it neglects the indicator cross-variograms and therefore omits important information about the contact relationships between the geological domains. Instead, a non-linear model of coregionalization should be used, a task that is made complex due to the mathematical restrictions that affect the indicator variograms [13,18,27]. As a second example, the simulation approaches based on multiple-point statistics [30] rely on the definition of a training image that is representative of the spatial distribution of the geological domains under study. Following the statements of the previous subsection, the choice of a training image consistent with the indicator direct and cross-variograms that can be inferred from available sampling data (in particular, reproducing the domain proportions, boundary regularity, contact relationships and edge effects) turns out to be crucial and non-trivial.

5. Conclusions

In geostatistical applications, geological domains are often codified through indicators and partially characterized by their direct and cross-variograms. The sills, correlation ranges, shapes and possible anisotropies of such indicator variograms convey information about the domain proportions and on their spatial correlation structure. Furthermore, their behavior at short distances brings a wealth
of knowledge on the domain boundaries. In particular, the average slope at the origin is proportional to the surface area of the domain boundary (in case of a direct variogram) or to the surface area of the contact between two domains (in case of a cross-variogram): the more slope, the more surface area. Also, the ratio of cross-to-direct variograms allows identification for whether two domains are preferentially in contact (border or edge effect) or whether they tend to be separated. Finally, the range of distances for which an indicator cross-variogram differs from zero provides information about the minimum and maximum distances between points from two domains.

In practice, all of these properties can be investigated on the basis of sampling information, through the analysis of experimental indicator variograms. Such an analysis is helpful for validating the consistency of an interpreted geological model of the subsurface or for constructing a representative training image to be used for simulating the layout of the domains with data-driven geostatistical approaches based on multiple-point statistics. A case study of a copper ore deposit has been presented, in which the interpreted layout of rock type domains is globally consistent with the actual layout, as revealed from the analysis of exploration drill hole data; as an exception, the distance separating two rock type domains (granodioritic porphyry and hydrothermal breccia) may be underestimated in the interpreted model and needs to be revised.

**Acknowledgments:** The first two authors acknowledge the funding of the Chilean Commission for Scientific and Technological Research (CONICYT), through Projects CONICYT PIA Anillo ACT1407 and CONICYT/FONDECYT/REGULAR/N°1170101, respectively. Costs to publish in open access were covered by CONICYT/FONDECYT/REGULAR/N°1170101 funds.

**Author Contributions:** Mohammad Maleki, Xavier Emery and Nadia Mery conceived and designed the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

**Conflicts of Interest:** The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

**Appendix A**

Consider two domains \(D_k\) and \(D_{k'}\) that do not overlap. The non-centered cross-covariance of the domains indicators \(I_k\) and \(I_{k'}\) is defined as:

\[
C_{k,k'}(x, x') = E\{I_k(x) I_{k'}(x')\} = P\{x \in D_k, x' \in D_{k'}\}
\]

and is related to the indicator cross-variogram in the following fashion (Equation (12)):

\[
\gamma_{k,k'}(h) = \frac{-C_{k,k'}(x, x') + C_{k,k'}(x', x)}{2}
\]

If the two indicator random fields are jointly second-order stationary, \(C_{k,k'}\) only depends on \(h = x' - x\) and tends to \(p_k p_{k'}\) when the norm of \(x' - x\) becomes infinitely large. Furthermore, the boundary common to both domains \(D_k\) and \(D_{k'}\) is related with the average derivative of \(C_{k,k'}\) at the origin (Equation (14)):

\[
\frac{1}{\pi} \int_{S_3} C_{k,k'}^{(u)}(0) \, du = E\{|\partial D_k \cap \partial D_{k'}|\}
\]

Unlike the cross-variogram, the cross-covariance provides information on the directional contact between the two domains \(D_k\) and \(D_{k'}\) and on asymmetrical relationships between these domains (for instance, \(D_k\) being preferentially above \(D_{k'}\) rather than below), a pattern that can sometimes be explained by chemical or sedimentological processes \[9,31,32\]. In particular, the ratio between indicator cross-covariance and direct variogram is:

\[
\frac{C_{k,k'}(h)}{\gamma_k(h)} = \frac{P\{x \in D_k, x + h \in D_{k'}\}}{P\{x \in D_k, x + h \not\in D_k\}} = P\{x + h \in D_{k'}|x \in D_k, x + h \not\in D_k\}
\]
and corresponds, up to the sign, to the case \( \delta = 1 \) of Equation (19). The sill of this ratio is \( \frac{p_x}{1 - p_x} \) and represents the reference value against which one can determine the existence of directional edge effects at very small distances:

- \( \frac{C_{x'}(x)}{\gamma_1(x)} < \frac{p_x}{1 - p_x} \): the contact area between \( D_k \) and \( D_{k'} \) when moving along vector \( u \) (i.e., leaving \( D_k \) at \( x \) and reaching \( D_{k'} \) at \( x + cu \)) is smaller than it would be in the absence of a directional edge effect.

- \( \frac{C_{x'}(x)}{\gamma_1(x)} = \frac{p_x}{1 - p_x} \): the contact area between \( D_k \) and \( D_{k'} \) when moving along vector \( u \) is equal to the expected area in the absence of a directional edge effect.

- \( \frac{C_{x'}(x)}{\gamma_1(x)} > \frac{p_x}{1 - p_x} \): the contact area between \( D_k \) and \( D_{k'} \) when moving along vector \( u \) is larger than it would be in the absence of an edge effect.

References


22. Emery, X.; Ortiz, J.M. Weighted sample variograms as a tool to better assess the spatial variability of soil properties. *Geoderma* 2007, 140, 81–89. [CrossRef]


© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).