A new approach to integrate geological information in geostatistical modelling

Yupeng Li and Clayton V. Deutsch

Subsurface geological formations are obtained from sparsely sampled well sites. Geostatistical techniques are commonly used to predict properties at unsampled locations. An important step in all geostatistical techniques is the calculation of the distance between data and/or grid cell locations. The conventional approach considers geometric anisotropy ratios applied in a stratigraphic coordinate system. This approach is extended to consider the depositional dip and strike as well as vertical changes for facies modelling. The facies variation pattern is different in each direction. The distance and the orientation must be accounted for. The distance upward is not the same as downward and the distance in a distal orientation is not the same as a proximal orientation. The variations along depositional strike are different again. A consistent 3-D approach to anisotropic distance calculation is proposed in this paper.

Introduction

A spatial data set consists of a collection of measurements on one or more attributes as well as the locations of those measurements. The locations are used to reference the relative positions between the sampled data and potential locations to predict. A coordinate system is needed to determine the “addresses” of different locations in a spatial space.

In classic geostatistic, the Cartesian coordinate system is used. Thus, a location $u$ is identified by means of three spatial coordinate axes as $u : (x, y, z)$. The Euclidean distance vector between two locations $u_1 : (x_1, y_1, z_1)$ and $u_2 : (x_2, y_2, z_2)$ is written as:

$$d = ||u_1 - u_2|| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \quad (1)$$

The information in this vector would include the separation intervals between two locations and the direction between these two locations in Euclidean space.

Further more, it is believed in geostatistics that the locations which are closer together will exhibit similar measured properties compare to locations farther apart. This similarity among the sampled spatial data set is dependent on their relative spatial locations. The direction information in the physical lag is important when the variability of spatial process behaves differently in different directions as quantified by spatial anisotropy. Defining this Euclidean distance vector $d$ as calculated from Equation (1) is the first step in measuring the spatial variability.

Next step in a traditional geostatistical approach would be using some statistic to characterize the spatial variability for distance vector $d$. It is done by introducing certain probability model to quantify the spatial variability with distance vector $d$. Variogram is the most commonly used measurement of spatial variability in geostatistics currently.
The experimental variogram denoted as $\hat{\gamma}(d)$, is a function of the spatial variability and its relative physical lag vector between locations [1].

After the experimental variogram calculation, these experimental variogram are modeled with certain mathematical correct functions in order to use them in the covariance based linear prediction model. As the mathematical functions are isotropic, a transformation to the anisotropic distances is required. Thus, the physical distance is transformed to statistical distance in probability space denoted as $h = \phi(d)$. Fitted with the mathematics function, the function will be written as $\gamma(h)$.

After setting the relationship of spatial variability with the distance vector, such as the variogram function, these location-to-location statistical distance $h$ in correlation space will be used to calculate the weights for the spatial prediction using some optimal approach such as all kinds of kriging in geostatistics [2]. The above procedure in traditional geostatistical practice is illustrated in Figure 1.

Although the above procedure is well practised in traditional geostatistical research, some assumptions behind the procedure need to be further checked. First is the choice of the coordinate system and the related distance calculation for two spatial locations. Because the spatial variability statistics are measured as a function of the spatial distance intervals, the choice of distance measure particularly in a spatial context becomes extremely important. Also, the way we calculate distance affects correlations as well as the selection of an conditioning neighbourhood for the potential predict location. Although Euclidean coordinate and some other coordinate system such as cylindrical coordinate system has been used recently [3], they are all in Euclidean space.

The chosen coordinate system should depend on the pertinent information about the modelling domain such as natural laws, topographical features etc. One such example is using the stream or water network as a coordinate to calculate the “water distances” in a water network [4]. In this paper, a new coordinate system is proposed to integrate the sedimentary stacking trends into the modelling. The geological understanding about the heterogeneity trends is enforced in this new proposed coordinate system. It will based on two fundamental geological principles: Walther’s law and the sequence stratigraphy theory. The details will be given in flowing section.

Based on the principle that the closer in distance will exhibit more similar properties, the distance vector in a more homogeneous distance will be shorter than the one in a more heterogeneous direction after both of them are transformed to the variability space. In earth science, soil or geological layers in environment research, it is very common to find different anisotropy characters along different directions. Such as the facies transition pattern along two direction $h_1$ and $h_2$ in Figure 2 are different. The distance vector $h_1$ is transforming along the direction where the facies (green) will have a higher transition probability to facies (blue). While distance vector $h_2$ is in a direction that facies (green) will extend to itself. Thus, the facies transition along distance vector $h_2$ will have a higher probability to stay in itself as green. So, those two distances will have different effective distances in probability space if the spatial variability between them are characterized by a same spatial statistics. The above situation is very common in geological situation such
as a delta front sedimentary environment where the sedimentary dip direction will have a totally different anisotropy character along sedimentary strike direction [5].

The traditional ellipsoid transforming as used in geometric anisotropic transforming assumes that the same transforming scale for all the directions would not appropriate for such kind of heterogeneous environment.

As discussed above, the geostatistical prediction of reservoir properties at an unsampled location almost always requires the calculation of the distance between nearby data to the unsampled location and between multiple nearby data. Also, it will affect the conditioning data neighbourhood in the predictor calculation or the correct configuration of data event in scanning a training image for multiple-point geostatistics. In this paper, a new distance transformation approach is proposed where the geological anisotropy in different directions are counted into final distance separately. Any spatial distance $d_{ij}$ between two locations $u_i, u_j$ will be scaled to an effective distance along the vertical direction where the spatial variability measurements are already obtained. This transformation is done with a decomposition in physical space based on the Cartesian coordinates and a combination in the new geological based coordinate system. The details is explained in the flowing section and with some examples illustration. Also, some possible applications are given in the discussion section.

This proposed methodology can be especially used in the horizontal spatial statistical inferring from vertical direction when the sampled location along horizontal is not dense enough for a reliable statistics inference.

**Geological Based Coordinate Definition**

There are two aspects of a successful reservoir modelling effort. One is the focus on the modelling algorithm itself while the other is the geological analysis work. Both are important. The final simulated model from the algorithm should be geologically realistic. The design of stochastic models should be tailored to a particular reservoir management problem. Also, the uncertainty in the underlying conceptual geologic model should be captured and reflected in the geostatistics model.

As the spatial distance is calculated from some spatial distributed locations, the address of each location will totally depended on the coordinate. Before quantity modelling the research domain, usually geologists already have a concepts that how the model should be looked like. In the proposed coordinate system definition, this conceptual heterogeneity distribution information will be exploited and integrated into the address calculation. It is based on two sound sedimentary theories: Walther’s law and sequence stratigraphy.

**Walther’s law**

The earth’s surface can be classified into different sedimentary realms or environments that are physically, chemically and biologically distinct from adjacent areas such as mountain ranges, sand deserts and deltas.
Each environment shows both abrupt and gradational lateral and vertical transitions. Sediments from those transitions will repeat vertically through a sedimentary sequence but may vary in character as a result of environmental and/or evolutionary changes through time. The relationship between depositional environments in space, and the resulting stratigraphic sequences developed through time was summarized by Walther’s Law which states that sedimentary environments that started out side-by-side will end up above one another over time due to transgressions and regressions [6, 7]. In other words, a vertical sequence of facies is similar to the original lateral distribution of sedimentary environments.

This is a fundamental principle of stratigraphy which allows geologists to predict the facies lateral changes from the vertical changes observed in outcrops, core or well logs. For example, in a fluvial dominated delta building sedimentary process, the sediments will have an upward-coarsening pattern along the vertical direction which is from prodelta, delta front to delta plain. This is the same as the lateral chronostratigraphical boundary, from the proximal to distal direction, the same contacting pattern can be found both from the lateral and vertical direction as shown in Figure 3.

Sedimentary facies model

In general, between 2 to 7 facies types are defined and their spatial relations are described according to a conceptual geologic model. This is usually done by using local criteria or by reference to an existing universal facies model that characterizes the facies spatial assemblage [8].

Facies models will show a pattern of environments that prograde side ways to deposit a series of facies arranged in a predictable vertical sequence. Facies models summarize the essential aspects of facies sequences and relate them to the inferred depositional environments.

Facies models have been proposed for most major depositional environments, and there is a large measure of consensus about their general aspects, particularly for fluvial, aeolian, deltaic, and other shallow-water settings [5]. They are usually most effectively conveyed graphically as three-dimensional block diagrams that relate the environment and its behaviour to the facies pattern as shown in Figure 4. The paleogeographic sketches, and vertical profile logs are also typical components of a published facies model as shown in Figure 5.

The universal facies model presented in classical sedimentary textbooks provides valuable facies stacking pattern information for geostatistical facies modelling.

Sequence stratigraphy

The theory of sequence stratigraphy is a relatively recent paradigm in the field of sedimentary geology. It is the study of genetically related facies within a framework of chronostratigraphically significant surfaces. The sequence stratigraphic approach has led to improved understanding of how stratigraphic units, facies tracts, and depositional elements relate to each other in time and space within sedimentary basins [9, 10]. It was first utilized
by the petroleum industry to interpret depositional surfaces on seismic sections [11, 12]. Now sequence stratigraphy is used extensively by geologists to explain vertical and lateral changes in sediment rock distribution [13, 14, 15, 16].

In sequence stratigraphy theory, the change in accommodation in conjunction with the rates of sedimentation represent a key control on depositional trends which is reflected by specific shoreline shifts. Two major depositional trends are usually recognized from the deposits: transgression and regression. A transgression is defined as the landward migration of the shoreline. This migration triggers a corresponding landward shift of facies. Transgression results in retragradational stacking patterns. A regression is defined as the seaward migration of the shoreline. This migration triggers a corresponding seaward shift of facies. Regressions result in progradational stacking patterns, for example, nonmarine facies shifting toward and overlying marine facies as shown in Figure 6.

Spatial heterogeneity prototype

Each reservoir is unique. The facies will be identified from wells, cores and perhaps from the analogue outcrops. The geologist will select a specific facies model and provide the sequence stratigraphic analysis for the reservoir from correlating a group of wells. The facies model and sequence stratigraphy analysis results show the facies associations and how each of them will be interpreted in context with others.

Based on the previous discussion, one conceptual heterogeneity prototype is proposed as plotted in Figure 7 to integrate the geological facies model and the sequence stratigraphy information in the geostatistical algorithms. In this prototype, three major heterogeneity axes are defined as: sedimentary dip, sedimentary strike and the vertical direction.

The vertical direction is the direction that the main sediments are observed and documented. The vertical profiles are the main form of data obtained from drilling. They are usually documented from logs. Their analysis is the starting point for Walther’s law and sequence stratigraphy to interpret depositional processes and sedimentary environments.

The dip axis is associated with the direction of the major facies transitions horizontally. Usually, in channelized sedimentary environments, it will be parallel to the direction from sedimentary source to sedimentary deposition. For example, it will be from proximal to distal in a coastal sedimentary environment. Along the dip and vertical direction, the facies stacking pattern is expected to be similar. For a fluvial dominated delta building sedimentary process, the sedimentary deposits will have an upward-coarsening pattern along the vertical direction as well as a similar pattern from the flood plain to the delta front direction as shown in Figure 3.

The strike axis will extend perpendicular to the the major sedimentary source direction. In most sedimentary environments, the shifting of sedimentary source is common. This sediments source switching phenomenon can be observed in the current Mississippi delta shown in Figure 8a. The Mississippi delta area is the modern area of land (the river delta) built up by alluvium deposited by the Mississippi River as it slows down and enters the Gulf of Mexico. The switching of the Mississippi River delta during the last 4,000 years is well documented [17, 18]. This kind of switching phenomenon is characterized by the
strike direction in the conceptual model also shown in Figure 8b.

Properties of prototype

The dip and strike axis can be predicted from the data if the pathway of deposition is directional such as in a delta environment, or in gravity sedimentary environments [19, 20]. The dip direction will be toward the source direction and the strike will be perpendicular to the dip direction. The vertical axis will be the main sedimentary stacking direction.

The sedimentary trends or stacking patterns are asymmetric along different directions. Along the vertical axis, the fining upward facies transition pattern would become a coarsening upward pattern in the reverse direction. The dip axis will have the same asymmetric property. Thus, it is necessary to define a positive direction along those two axes for computational implementation.

The direction from the earlier deposits to the later deposits will define a positive direction and usually it will be from bottom to top. For the dip axis, the positive direction will be from the depositional source to the margin, that is, from proximal to distal. The positive direction definition along dip and in the vertical axis should have the same stacking trend, that is, their positive direction should be consistent with each other. For example, in a delta building depositional process, the positive axes will be different from delta sediments in a transgressing process as shown in Figure 9.

The strike direction will have no positive or negative direction. The sediments along this axis will be characterized by a stochastic function defined below.

One more characteristic of these three axes is the anisotropy ratio along the strike and dip axes. With high-resolution chronostratigraphic correlation analysis, the sedimentologists could estimate how much of the heterogeneity stacking pattern from the vertical well profile will be valid along the lateral direction, especially along the sedimentary dip direction. In this research, this geological understanding is quantified by heterogeneity ratio along dip axis. This ratio is denoted as $a_x$, which is equal to:

$$a_x = \frac{H_x}{H_z}$$  \hspace{1cm} (2)

where the $H_x$ and $H_z$ would be the full pattern thickness along the dip and vertical direction.

While along the strike direction, the relative shifting area will also estimated from the sequence stratigraphy and facies analysis. Along the strike direction, the anisotropy ratio $a_y$ will control the prototype model. It will calculated as:

$$a_y = \frac{H_y}{H_z}$$  \hspace{1cm} (3)

where $H_y$ is the distance that the geologists believe the switching will happen for the particular reservoir.
3D Spatial Distance Transformation

After setting up the anisotropy axes in the reservoir, during the modelling, any arbitrary spatial distance vector \( \mathbf{h} \) will be decomposed along those three axes as shown in Figure 10. Three new vectors \( \mathbf{h}_{\text{vert}}, \mathbf{h}_{\text{dip}} \) and \( \mathbf{h}_{\text{strike}} \) will describe the spatial distance between two locations \((\mathbf{u}, \mathbf{u} + \mathbf{h})\).

The orientation of the anisotropy is controlled by the geological features of the reservoir and the orientation of the coordinate system. The anisotropy axes may not coincide with the axes of the model coordinate system. In this case, the components \( h_x, h_y, h_z \) of the distance vector \( \mathbf{h} \) in the data coordinate system will have different values when referenced in the coordinate system of the anisotropy axes. Thus, it is necessary to transform the vector from the data coordinate system to the coordinate system of the anisotropy axes. In the following discussions it is assuming that the \( x, y \) and \( z \) direction of the model coincide with the dip, strike and vertical axes of the prototype model. The decomposed distance vector along the three axes \( \mathbf{h}_{\text{vert}}, \mathbf{h}_{\text{dip}}, \mathbf{h}_{\text{strike}} \) will correspond to the \( h_z, h_x, h_y \) distance vectors.

Any spatial distance vector \( \mathbf{h} \) will be transformed to an effective distance vector \( \mathbf{h}_{\text{eff}} \) along the vertical direction. It is done in two steps shown in Figure 11. In the first step, the dip distance vector component \( h_x \) is modified to a vector \( D_x \) reflecting the sedimentary shifting along the strike direction controlled by strike component vector \( h_y \). Then, the modified distance vector \( D_x \) from the first step is combined with the vertical component vector \( h_z \) to obtain a final effective vertical vector \( \mathbf{h}_{\text{eff}} \) based on the sedimentary environment.

Integration of dip and strike direction

The anisotropy along the sedimentary dip and strike direction will be combined together first. The basic idea is that the shifting process along the strike direction follows some random process. The lag distance \( h_x \) along the dip direction will be modified by a factor from the strike directions which will mimic sediment source shifting. It will be the main effect on the final effective distance calculation \( D_x \) from the strike direction vector \( h_y \).

As the sediments source shifting along the strike direction, the distance increase along strike direction will not always increase the probability of transition to another facies as it does along dip direction. This geological variability along the strike axis direction is quantified by \( F_y \). For example, it could be modelled as sine wave function \( F_y = f(h_y) = A \cdot \sin(B \cdot h_y + C) \). For the sine function, the value from valley to peak will be the maximum deviation to the dip distance caused by strike. The period of this sine wave would depends on the research domain

\[
\frac{2\pi}{B} = H_y = nper \cdot ny \cdot ysz
\]

where \( H_y \) is the full switching distance; and the parameter \( nper \) will control how may full sine periods will be found along the strike direction. The larger \( nper \), the higher frequency of randomness modifications to the distance along the dip direction. In the simulation, each
realization would use a different random parameter $\frac{C}{F} = \text{rand}(\cdot)$ for the sine wave as shown in Figure 12.

The variety caused by the distance vector along strike direction $h_y$ is denoted as $\Delta F_y$ and will cause a modification to the distance vector $h_x$ along the dip distance. The effective distance along dip direction $D_x$ is combined from the factor $\Delta F_y$ and the actual distance along dip direction $h_x$ as:

$$D_x = h_x - \Delta F_y \tag{4}$$

There are different cases for the combination of strike and sedimentary dip together. Some possible cases are shown in Figure 13.

Case A : $h_x = \Delta F_y$. Although there is a dip distance difference ($h_x \neq 0$), the effect from the strike will cancel that dip difference. Thus, for those two locations, there will be no increment along the dip direction.

Case B : $\Delta F_y < 0$ the increment along the dip direction will be increased from the sedimentary effect according Equation (4). Thus, the combined distance will be larger than the original increment $h_x$.

Case C : $\Delta F_y = 0$ the increment along the dip direction will be kept as the same. The variability will only depend on the dip vector $h_x$.

Case D : $\Delta F_y > 0$, the increment along the dip direction will be decreased considering the effect from the strike direction;

Case E : $h_x = 0$ and $\Delta F_y < 0$, although there is no dip distance difference. But from the strike effect, there would be an increment along the dip direction as $D_x > 0$ from Equation (4).

By this approach, the distance along strike and dip will be combined together. As shown in the combination example listed above, the geological meaning is integrated into the “effective” sign-dependent distance calculation.

A small example that can be calculated manually will be used to show the combination of dip from strike direction. There are four locations in Cartesian coordinates shown in Figure 14a. The distance vectors between all the locations will follow the Euclidean geometry property. In Figure 14b, the coordinate system is the geological based as along strike direction and there is a sine wave of $(2 \cdot \sin(\pi/10 \cdot h_y))$ attached to. The X component of each distance vector is modified by the factor which is calculated from the random function along the strike direction using the Y component of the distance vector as a variable. As can be seen from this example, after accounting the fluctuation caused by Y components, the X component along the dip direction is embedded into a new space. Although this transformation is not Euclidean, the transformed distances along the dip direction satisfy the triangle qualities.

After taking the positive direction into consideration, the distance matrix between the four locations is shown in Table (1). In this matrix, the negative sign will affect the facies
Table 1: Distance matrix from coming dip and strike direction

<table>
<thead>
<tr>
<th></th>
<th>o</th>
<th>a</th>
<th>b</th>
<th>c</th>
</tr>
</thead>
<tbody>
<tr>
<td>o</td>
<td>0</td>
<td>-5.4797</td>
<td>14.4046</td>
<td>-0.5441</td>
</tr>
<tr>
<td>a</td>
<td>-5.4797</td>
<td>0</td>
<td>8.9249</td>
<td>-6.0238</td>
</tr>
<tr>
<td>b</td>
<td>-14.4046</td>
<td>-8.9249</td>
<td>0</td>
<td>-14.9487</td>
</tr>
<tr>
<td>c</td>
<td>0.5441</td>
<td>6.0238</td>
<td>14.9487</td>
<td>0</td>
</tr>
</tbody>
</table>

stacking pattern. The row index is the head of a distance vector. The column index will be the tail of a distance vector.

Integration of dip and vertical direction

Combining the $h_z$ and $D_x$ into an effective distance $h_{eff}$ based on the anisotropy ratio $a_x$ along the dip direction is the next step for a final effective distance calculation. The final combined distance vector is calculated as:

$$h_{eff} = h_z + \frac{D_x}{a_x}$$  (5)

The positive direction will be different for a regression and transgression stacking pattern. There will be four combination results for an arbitrary distance vector as shown in Figure 15.

Case (a) in Figure 15 is the situation where $h_z > 0$ and $D_x > 0$. The vector of $h_z > 0$ means that along the vertical direction, the facies will have a transition tendency from four to one. While the vector of $D_x > 0$ means that along the dip direction, the facies will have a same trend as the vertical trend. Thus, the effective vector $h_{eff}$ combined from Equation (5) will also keep the transition trend that the facies will have a high tendency from four to one. The distance interval $|h_{eff}|$ will be larger than any single one $|h_z|$ or $|\frac{D_x}{a_x}|$.

In Case (b), the vector of $D_x$ is still positive, while the vector $h_z$ is negative which means that along the vertical direction and the lateral direction the facies transition trends are different. The direction of the final combination result vector $h_{eff}$ from those two vectors will depend on the stronger one. If $|\frac{D_x}{a_x}| > |h_z|$, then the direction of vector $h_{eff}$ will be positive. Otherwise, it will be negative. In both situations, the distance interval $|h_{eff}|$ will be shorter than $|h_z|$ or $|\frac{D_x}{a_x}|$.

In case (c), both the transition trend from vertical and dip direction show that the facies has a higher probability from four to one. They are consistent. From the combination, the value $|\frac{D_x}{a_x}|$ is added to the vertical distance interval $|h_z|$. The direction vector will have a negative value to indicate the geological meaning.

In case (d), the trend from vertical and lateral contradict each other again. In this situation, if the $|\frac{D_x}{a_x}| > |h_z|$, the direction of vector $h_{eff}$ will follow the direction of vector $D_x$. Otherwise, if $|\frac{D_x}{a_x}| < |h_z|$, the direction of vector $D_x$ will follow the vector $h_z$, which is positive. The distance from the combination will be shorter than each one.
A small manually calculated example is given in Figure 16. This time, the data configuration is the same as the example given in Figure 14, but the coordinate axes changed to dip and vertical axes. To the left, all the distance vectors are decomposed along the two axes. To the right, all the distance vectors are combined from vertical component $h_z$ and effective dip vector component $D_x$ using Equation (5). Note that the geometrical relations are still valid in this new coordinate system.

**Conclusion and future work**

One spatial distance approach is proposed in this paper. In this distance calculation approach, the geological anisotropy constraints are integrated through define a new coordinate system.

It could be used in traditional geostatistics to construct a more meaning data searching distance. It can also be used in the covariance based facies modelling approach such as Indicator kriging estimation/simulation. The challenges would come from the positive definite in the covariance matrix. New approach is needed to integrate the statistics from these new transformed distances without involving the covariance matrix.

**References**


Figure 1: The transforming flow of a spatial physical distance to statistical distance

Figure 2: Two distance vector along different direction in a heterogeneity environment

Figure 3: The facies stacking pattern in fluvial dominated delta sedimentary system
Figure 4: One example of braided river facies model shown in three-dimensional block diagrams

Figure 5: Example of facies model diagram in paleogeographic sketches and vertical profile logs (http://fosi.iagi.or.id/mahakam/mah-facies-des.htm)
Figure 6: Transgression and regressions model in sequence stratigraphy theory (Modified from Catuneanu, 2006)

Figure 7: Sketch showing the heterogeneity prototype
Figure 8: The lateral switching phenomenon in Mississippi delta sedimentary environment and its prototype representation
Figure 9: The different positive definition for dip direction in case of transgression (top) and regression of the sea (bottom).

Figure 10: Decomposition along three major anisotropy axes.
Figure 11: Two steps in three dimensional distance combination

Figure 12: Sketch showing the random factor from the switching function along the strike direction

Figure 13: Illustration of random modification of the dip vector from the strike vector based on the random switching function
Figure 14: One example of the combination between dip and strike direction. In (a), all the distances are in their original space and will be decomposed along dip, and strike direction; in (b), the green locations are the relative dip locations for those four points; After modification by the factor from strike vector, the relative dip locations are shown in (c).
Figure 15: Distance combination from vertical direction and dip direction
Figure 16: One example of vertical and dip combination results. The green relative locations for those four points in (a) indicate their relationship before combination; in (b) is the relative location along the vertical direction. The distance interval between them will be the final effective distance.