JSR data archive material for

HOW POROSITY AND PERMEABILITY VARY SPATIALLY WITH GRAIN SIZE, SORTING, CEMENT VOLUME AND MINERAL DISSOLUTION IN FLUVIAL TRIASSIC SANDSTONES: THE VALUE OF GEOSTATISTICS AND LOCAL REGRESSION.

J.M. MCKINLEY, P.M. ATKINSON, C. D. LLOYD, A.H. RUFFELL and R.H. WORDEN

Geostatistical Techniques

In geostatistics, spatial variation is modelled as comprising two distinct parts, a deterministic component $\mu(\mathbf{x})$ and a stochastic (or 'random') component $R(\mathbf{x})$ where $\mathbf{x}$ represents location. This is termed a random function (RF) model. The upper case Z refers to the RF whereas lower case z refers to a data realisation of the RF model, called a regionalised variable (ReV). In geostatistics, a spatially-referenced variable $z(\mathbf{x})$ is treated as an outcome of a RF $Z(\mathbf{x})$. The Theory of Regionalised Variables (Matheron, 1971) is the fundamental framework on which geostatistics is based. In a classical framework, geostatistical analysis usually involves two stages: (i) estimation of the variogram and fitting a model to it and (ii) use of the variogram model coefficients for spatial prediction (kriging) or simulation.

Variography

The variogram is a core tool in geostatistical analysis and is required for geostatistical spatial prediction. The variogram characterises spatial dependence in the property of interest such as permeability or porosity. In simple terms, the variogram is estimated by calculating half the average squared difference between all the available paired measurements separated by a given lag tolerance, where lag $\mathbf{h}$ is a separation vector (distance and direction). The experimental variogram, $\hat{\gamma}(\mathbf{h})$, can be estimated from $p(\mathbf{h})$ paired observations, $z(\mathbf{x}_\alpha)$, $z(\mathbf{x}_\alpha + \mathbf{h})$, $\alpha = 1, 2, \ldots p(\mathbf{h})$ using:

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2p(\mathbf{h})} \sum_{\alpha=1}^{p(\mathbf{h})} \{z(\mathbf{x}_\alpha) - z(\mathbf{x}_\alpha + \mathbf{h})\}^2 \qquad (1)$$

A mathematical model may be fitted to the experimental variogram and the coefficients of this model can be used for a range of geostatistical operations such as spatial prediction (kriging) and

conditional simulation. A model is usually selected from one of a set of authorised models. McBratney and Webster (1986) provide a review of some of the most widely used authorised models. The various parameters of the models fitted to experimental variograms can be related to structural features in a rock face or succession (McKinley et al. 2004). For example, in bounded or transitive types of model, the range parameter provides information about the maximum scale of spatial variation.

Kriging

Kriging is a smoothing interpolator since kriging predictions are weighted moving averages of the available data. The most widely used variant of kriging, termed ordinary kriging (OK) is used in this study. OK allows the mean to vary spatially: the mean is estimated for each prediction neighbourhood from the available data.

The OK weights define the best linear unbiased predictor (BLUP). The OK prediction, $\hat{z}_{\mathrm{OK}}(\mathbf{x}_0)$, is defined as:

$$\hat{z}_{\mathrm{OK}}(\mathbf{x}_0) = \sum_{\alpha=1}^{n} \lambda_{\alpha}^{\mathrm{OK}} z(\mathbf{x}_{\alpha}) \qquad (2)$$

given $\alpha = 1,...n$ available data $\hat{z}(\mathbf{x}_{\alpha})$, with the constraint that the weights, $\lambda_{\alpha}^{\mathrm{OK}}$, sum to 1 to ensure unbiased prediction:

$$\sum_{\alpha=1}^{n} \lambda_{\alpha}^{\mathrm{OK}} = 1 \qquad (3)$$

The kriging prediction error must have an expected value of 0:

$$E\{\hat{Z}_{\mathrm{OK}}(\mathbf{x}_0) - Z(\mathbf{x}_0)\} = 0 \qquad (4)$$

The kriging (or prediction) variance, σ_{OK}^2, is expressed as:

$$\hat{\sigma}^2_{\mathrm{OK}}(\mathbf{x}_0) = E[\{\hat{Z}_{\mathrm{OK}}(\mathbf{x}_0) - Z(\mathbf{x}_0)\}^2] \qquad (5)$$

$$= 2\sum_{\alpha=1}^{n} \lambda_\alpha^{\mathrm{OK}} \gamma(\mathbf{x}_\alpha - \mathbf{x}_0) \sum_{\alpha=1}^{n} \sum_{\beta=1}^{n} \lambda_\alpha^{\mathrm{OK}} \lambda_\beta^{\mathrm{OK}} \gamma(\mathbf{x}_\alpha - \mathbf{x}_\beta)$$

That is, we seek the values of $\lambda_1, \ldots, \lambda_n$ (the weights) that minimise this expression with the constraint that the weights sum to one (equation 3). This minimisation is achieved through Lagrange Multipliers. The conditions for the minimisation are given by the OK system comprising $n + 1$ equations and $n + 1$ unknowns:

$$\begin{cases} \sum_{\beta=1}^{n} \lambda_\beta^{\mathrm{OK}} \gamma(\mathbf{x}_\alpha - \mathbf{x}_\beta) + \psi_{\mathrm{OK}} = \gamma(\mathbf{x}_\alpha - \mathbf{x}_0) \qquad \alpha = 1, \ldots, n \\ \sum_{\beta=1}^{n} \lambda_\beta^{\mathrm{OK}} = 1 \end{cases} \qquad (6)$$

where $\gamma(\mathbf{x}_\alpha - \mathbf{x}_\beta)$ represents the semivariance between the available data and themselves, $\gamma(\mathbf{x}_\alpha - \mathbf{x}_0)$ represents the semivariance between the available data and the prediction location $\mathbf{x}_0$, and ψ_{OK} is a Lagrange multiplier. Knowing ψ_{OK}, the prediction variance of OK can be given as:

$$\hat{\sigma}^2_{\mathrm{OK}} = \psi_{\mathrm{OK}} + \sum_{\alpha=1}^{n} \lambda_\alpha^{\mathrm{OK}} \gamma(\mathbf{x}_\alpha - \mathbf{x}_0) \qquad (7)$$

The kriging variance is a measure of confidence in predictions and is a function of the form of the variogram, the sample configuration and the sample support (the area over which an observation is made, which may be approximated as a point or may be an area; Journel and Huijbregts, 1978). The kriging variance is not conditional on the data values locally and this has led some researchers to use alternative approaches such as conditional simulation (discussed in the next section) to build models of spatial uncertainty (Goovaerts, 1997).

Conditional simulation

As stated previously, kriging predictions are weighted moving averages of the available sample data. Kriging is, therefore, a smoothing interpolator. Conditional simulation (also called stochastic imaging) is not subject to the smoothing associated with kriging (conceptually, the variation lost by kriging due to smoothing is added back) as predictions are drawn from equally probable joint realisations of the Random Variables (RVs) which make up a RF model (Deutsch and Journel, 1998). That is, simulated values are not the expected values (i.e., the mean), but are values drawn randomly from the conditional cumulative distribution function (ccdf): a function of the available observations and the modelled spatial variation (Dungan, 1999). Simulated realisations represent a possible reality whereas kriging predictions do not. Therefore SGS realisations can be used to represent possible variation for properties of a rock face or succession whereas kriging produces a smoothed interpolation of the data. The simulation is considered "conditional" if the simulated values honour the observations at their locations (Deutsch and Journel, 1998). Simulation allows the generation of many different possible realisations which encapsulate the uncertainty in spatial prediction (Journel, 1996) and thus may be used as a guide to potential errors in the characterisation of variation in rock properties (Journel, 1996).

Probably the most widely used form of conditional simulation is sequential Gaussian simulation (SGS) as used in this study. With sequential simulation, simulated values are conditional on the original data and previously simulated values (Deutsch and Journel, 1998). In SGS the ccdfs are all assumed to be Gaussian. The SGS algorithm follows several steps (Goovaerts, 1997; Deutsch, 2002) as detailed below:

1. Apply a standard normal transform to the data.
2. Go to the location $\mathbf{x}_1$.
3. Use SK (note OK is often used instead; see Deutsch and Journel 1998 about this issue), conditional on the original data, $z(\mathbf{x}_\alpha)$, to make a prediction. The SK prediction and the kriging variance are parameters (the mean and variance) of a Gaussian ccdf:

$$F(\mathbf{x}_1; z|(n) = \text{Prob}\{Z(\mathbf{x}_1) \leq z|(n)\} \qquad (8)$$

4. Using Monte Carlo simulation, draw a random residual, $z^l(\mathbf{x}_1)$, from the ccdf.

5. Add the SK prediction and the residual to give the simulated value; the simulated value is added to the data set.

6. Visit all locations in random order and predict using SK conditional on the *n* original data and the *i*-1 values, $z^l(\mathbf{x}_i)$, simulated at the previously visited locations $\mathbf{x}_j$, j=1,…,n to model the ccdf:

$$F(\mathbf{x}_1; z|(n+i-1) = \text{Prob}\{Z(\mathbf{x}_1) \leq z|(n+i-1)\} \qquad (9)$$

7. Follow the procedure in steps 4 and 5 until all locations have been visited (that is, until realisations l = 1,…, L have been obtained).

8. Back transform the data values and simulated values.

By using different random number seeds, the order of visiting locations is varied and multiple realisations can be obtained. In other words, since the simulated values are added to the data set, the values available for use in simulation are partly dependent on the locations at which simulations have already been made and, because of this, the values simulated at any one location vary as the available data vary. SGS is discussed in detail in several texts (e.g., Goovaerts, 1997; Deutsch and Journel, 1998; Chilès and Delfiner, 1999; Deutsch, 2002).

GEOGRAPHICALLY WEIGHTED REGRESSION (GWR)

This section describes the basic global linear regression model and its extension to geographically weighted regression (GWR) (local regression with spatially varying parameters).

Linear Regression

The basic linear regression model is:

$$y_i = \beta_0 + \sum_{k=1}^{m-1} \beta_k x_{ki} + \varepsilon_i \qquad (10)$$

predicting the target variable y_i at location i as a function of m parameters β_0 and β_k (k=1, …, m-1) and m-1 explanatory variables x_{ki}. The model includes an independent normally distributed error term with zero mean ε_i. Usually, the least squares method, based on matrix algebra, is used to estimate the β_ks. This is readily expressed in matrix notation as;

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{x}^t \mathbf{x}\right)^{-1} \mathbf{x}^t \mathbf{y} \qquad (11)$$

where $\hat{\boldsymbol{\beta}}$ is a single column vector of estimated parameters (i.e., regression coefficients), a first column of 1s and the m-1 explanatory variables fill the m columns of **x**, the target variable fills the single column of **y** and t indicates the transpose of a matrix. The single column of 1s in the matrix **x** allows estimation of β_0.

Equation 10 is a *stationary* model of the relations between **y** and **x**, that is, a model in which the parameters of the model are constant irrespective of geographical location. Thus, the model is appropriate only where the relations may reasonably be modelled as invariant over space. Where the relations are expected or known to change with geographical location it is necessary to apply a spatially non-stationary model: one that allows variation in the parameters of the model as one moves from place to place. Non-stationary regression modelling is not new, but Brunsdon *et al*. (1999) developed the approach and termed it geographically weighted regression (GWR). GWR offers a basis upon which to address some of the limitations of conventional global regression models described in the introduction.

Geographically Weighted Regression

GWR extends the traditional regression model of Equation 10 to the non-stationary case by allowing the $\boldsymbol{\beta}$ coefficients to vary with geographical location i:

$$y_i = \beta_{i0} + \sum_{k=1}^{m-1} \beta_{ik} x_{ik} + \varepsilon_i \qquad (12)$$

where β_{ik} is the value of the k^{th} parameter at location i. Equation 10 is a special case of the more general Equation 12, where the parameters do not vary across space.

A potential problem with Equation 5 is that there is only one observation at each i for the target variable y_i and the m-1 explanatory variables x_{ik}. To provide sufficient observations with which to fit a local regression model for each i, Brunsdon *et al*. (1999) considered the relation between **y** and **x** within a moving window around i, while acknowledging that such an approach involves approximation.

The weighted least squares (WLS) approach to fitting regression models provides a means by which to vary the influence of individual observations on the fitted model. Specifically, in WLS a weight w_i is applied to each squared difference before minimisation. The weights are usually chosen to inversely reflect some measure of uncertainty so that more uncertain observations are assigned less weight. If **w** is the diagonal matrix of the w_is then Equation 11 can be extended readily as:

$$\tilde{\boldsymbol{\beta}} = \left(\mathbf{x}^t \mathbf{w} \mathbf{x}\right)^{-1} \mathbf{x}^t \mathbf{w} \mathbf{y} \qquad (13)$$

Since the diagonal weights matrix **w** varies with location j, such that a different calibration exists for each j, it is helpful to write Equation 6 more strictly as:

$$\tilde{\boldsymbol{\beta}} = \left(\mathbf{x}^t \mathbf{w}(\mathbf{j}) \mathbf{x}\right)^{-1} \mathbf{x}^t \mathbf{w}(\mathbf{j}) \mathbf{y} \qquad (14)$$

Location has now been referenced by j rather than i because the location at which a model is fitted need not equal the location of the available data i.

Brunsdon *et al*. (1999) consider a range of possible weighting functions for use in Equation 7. These include a simple step function:

$$\begin{cases} w_{ij} = 1 & \text{if} \quad d_{ij} < d_j \\ w_{ij} = 0 & \text{otherwise} \end{cases} \qquad (15)$$

where d_{ij} is the distance between the location i of an observation used in model fitting and the location j at which the model is fitted. This weighting function is fast because the number of data used in the regression model is limited. However, its discontinuous nature can result in artefacts (spatial discontinuities) in the regression coefficients. An alternative is to specify w_{ij} as a *continuous* function of d_{ij}. The exponential function is a common choice:

$$w_{ij} = \exp\left(\frac{-d_{ij}}{a}\right) \qquad (16)$$

where, a is a non-linear parameter. As a compromise between the computational saving of Equation 15 and the continuity of Equation 16, Brunsdon *et al*. (1999) suggest using the bisquare function:

$$\begin{aligned} w_{ij} &= \left[1 - d_{ij}^2 / d^2\right]^2 \quad \text{if} \quad d_{ij} < d_j \\ w_{ij} &= 0 \qquad\qquad\quad \text{otherwise} \end{aligned} \qquad (17)$$

The essential idea of GWR, whatever weighting function is chosen, is to give more weight to observations close to the location j at which the regression model is desired than to those observations that are further away. In this sense, the method has some parallels with geostatistical techniques for local prediction, as described above.

References

Brunsdon, C., Aitkin, M., Fotheringham, S., and Charlton, M.E., 1999, A comparison of random coefficient modelling and geographically weighted regression for spatially non-stationary regression problems: Geographical and Environmental Modelling, v. 3, p. 47-62.

Chilès, J.P., and Delfiner, P., 1999, Geostatistics: Modeling Spatial Uncertainty: New York, John Wiley and Sons, 695 p.

Deutsch, C.V., 2002, Geostatistical Reservoir Modelling: New York, Oxford University Press, 376 p

Deutsch, C.V., and Journel, A.G., 1998, GSLIB: Geostatistical Software Library and User's Guide, Second Edition: New York, Oxford University Press, 369 p.

Dungan, J.L., 1999, Conditional simulation, *in* Stein A., van der Meer, F., and Gorte B., eds., Spatial Statistics for Remote Sensing, Stein: Dordrecht, Kluwer Academic Publishers, p. 135–152.

Goovaerts, P., 1997, Geostatistics for Natural Resources Evaluation: New York, Oxford University Press, 483 p.

Journel, A.G., 1996, Modelling uncertainty and spatial dependence: Stochastic imaging. International Journal of Geographical Information Systems, v. 10, p. 517-522.

Journel, A.G., and Huijbregts, C.J., 1978, Mining Geostatistics, Orlando: Academic Press,600p.

Matheron, G., 1971, The Theory of Regionalised Variables and its Applications:Fontainebleau, Centre de Morphologie Mathématique de Fontainebleau, 211 p.

McBratney A.B., Webster R., 1986, Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates: Journal of Soil Science, v. 37, p. 617–639.

McKinley, J.M., Lloyd, C.D., and Ruffell, A.H., 2004, Use of variography in permeability characterisation of visually homogeneous sandstone reservoirs with examples from outcrop studies: Mathematical Geology, v. 36, p. 761-779.