BTP

Monday, February 8, 2021 3:49 PM

1. Mobile Nutrients -
   - NO3 lacks in spatial structures

2. For Diggle and Ribeiro (2007)-
   1. model estimation
   2.prediction
   3. Testing

- Interpolation within the spatial Range
- Extrapolation outside the range where we do not have enough evidence to support out assumption.

Universal Model of Variation
From Information theory perspective an environmental variable can be defined as -

$$Z(\mathbf{s}) = Z^*(\mathbf{s}) + \varepsilon'(\mathbf{s}) + \varepsilon''$$

Where Z(s) is a probabilistic model depends upon the location s.
Z*(s) is a deterministic factor, ε' is the model correlated random component and ε'' is the pure noise in the data.

Mapping process -
1. Understand the source of variability in the data.

Mapping Methods -

Spatial prediction models (algorithms) can be classified according to the amount of statistical analysis i.e. amount of expert knowledge included in the analysis:

(1.) **MECHANICAL (DETERMINISTIC) MODELS** — These are models where arbitrary or empirical model parameters are used. No estimate of the model error is available and usually no strict assumptions about the variability of a feature exist. The most common techniques that belong to this group are:

- *Thiessen polygons*;
- *Inverse distance interpolation*;
- *Regression on coordinates*;
- *Natural neighbors*;
- *Splines*;
- …

(2.) **LINEAR STATISTICAL (PROBABILITY) MODELS** — In the case of statistical models, the model parameters are commonly estimated in an objective way, following probability theory. The predictions are accompanied with an estimate of the prediction error. A drawback is that the input data set usually need to satisfy strict statistical assumptions. There are at least four groups of linear statistical models:

- *kriging* (plain geostatistics);
- *environmental correlation* (e.g. regression-based);
- *Bayesian-based models* (e.g. Bayesian Maximum Entropy);
- *hybrid models* (e.g. regression-kriging);
- …

(3.) **EXPERT-BASED SYSTEMS** — These models can be completely subjective (*ergo* irreproducible) or completely based on data; predictions are typically different for each run. Expert systems can also largely be based on probability theory (especially Bayesian statistics), however, it is good to put them in a different group because they are conceptually different from standard linear statistical techniques. There are at least three groups of expert based systems:

- *mainly knowledge-driven expert system* (e.g. hand-drawn maps);
- *mainly data-driven expert system* (e.g. based on neural networks);
- *machine learning* algorithms (purely data-driven);

Spatial prediction models can also be classified based on the:

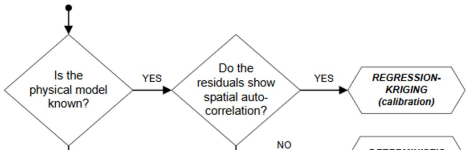
There are two kinds of models -
1. Deterministic
2. Stochastic

One can possibly argue that If there is a physical process behind, one should go for Deterministic model rather than stochastic one.
But the variability in spatial coordinates arises due to Random process called **Diffusion.** And Diffusion is a Random Process so there is meta-statistical argument that It is a stochastic process.

Model Evaluation -
1. **ME**
2. **RMSE**
3. **RMNSE**

(If The RMSE is close to 40%, then satisfactory result. (R squared is close to 80%)
For this we need independent samples which is quote expensive to collect.
So we will use **K-fold** cross validation with k = 10.

How to decide which spatial prediction model is suitable ?
Consider the decision tree given below .



Ordinary Kriging -

A standard version of kriging is called **ordinary kriging (OK)**. Here the predictions are based on the model:

$$Z(\mathbf{s}) = \mu + \varepsilon'(\mathbf{s}) \tag{1.3.1}$$

where μ is the constant *stationary* function (global mean) and $\varepsilon'(\mathbf{s})$ is the spatially correlated stochastic part of variation. The predictions are made as in Eq.(1.2.1):

$$\hat{z}_{0K}(\mathbf{s}_0) = \sum_{i=1}^{n} w_i(\mathbf{s}_0) \cdot z(\mathbf{s}_i) = \boldsymbol{\lambda}_0^{\mathrm{T}} \cdot \mathbf{z} \tag{1.3.2}$$

where $\boldsymbol{\lambda}_0$ is the vector of kriging weights (w_i), $\mathbf{z}$ is the vector of n observations at primary locations. In a way, kriging can be seen as a sophistication of the inverse distance interpolation. Recall from §1.2.1 that the key problem of inverse distance interpolation is to determine how much importance should be given to each neighbor. Intuitively thinking, there should be a way to estimate the weights in an objective way, so the weights reflect the true spatial autocorrelation structure. The novelty that Matheron (1962) and colleagues introduced to the analysis of point data is the derivation and plotting of the so-called **semivariances** — differences between the neighboring values:

$$\gamma(\mathbf{h}) = \frac{1}{2} E\left[\left(z(\mathbf{s}_i) - z(\mathbf{s}_i + \mathbf{h}) \right)^2 \right] \tag{1.3.3}$$

where $z(\mathbf{s}_i)$ is the value of a target variable at some sampled location and $z(\mathbf{s}_i + \mathbf{h})$ is the value of the neighbor at distance $\mathbf{s}_i + \mathbf{h}$. Suppose that there are n point observations, this yields $n \cdot (n-1)/2$ pairs for which a semivariance can be calculated. We can then plot all semivariances versus their separation distances, which will produce a variogram cloud as shown in Fig. 1.9b. Such clouds are not easy to describe visually, so the values are commonly averaged for a standard distance called the "*lag*". If we display such averaged data, then we get a standard **experimental or sample variogram** as shown in Fig. 1.9c. What we usually expect to see is that semivariances are smaller at shorter distance and then they stabilize at some distance within the extent of a study area. This can be interpreted as follows: the values of a target variable are more similar at shorter distance, up to a certain distance where the differences between the pairs are more less equal to the global variance[18].

How to estimate the weight (w) ?
We can estimate the weight parameter using either of the **semivariance** or **Covariance** .

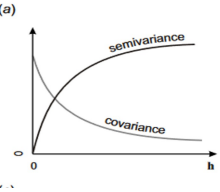
Relation between semivariance and Covariance -

$$C(\mathbf{h}) = C_0 + C_1 - \gamma(\mathbf{h})$$

Where C(h) is covariance at lag distance *h* and

$\gamma(h)$ is semivariance.

C_0 = nugget
C_1 = sill
$C_0 + C_1$ = sill variance



solve the kriging weights. The kriging OK weights are solved by multiplying the covariances:

$$\boldsymbol{\lambda}_0 = \mathbf{C}^{-1} \cdot \mathbf{c}_0; \qquad C(|\mathbf{h}| = 0) = C_0 + C_1 \tag{1.3.4}$$

where $\mathbf{C}$ is the covariance matrix derived for $n \times n$ observations and $\mathbf{c}_0$ is the vector of covariances at a new location. Note that the $\mathbf{C}$ is in fact $(n+1) \times (n+1)$ matrix if it is used to derive kriging weights. One extra row and column are used to ensure that the sum of weights is equal to one:

$$\begin{bmatrix} C(\mathbf{s}_1,\mathbf{s}_1) & \cdots & C(\mathbf{s}_1,\mathbf{s}_n) & 1 \\ \vdots & & \vdots & \vdots \\ C(\mathbf{s}_n,\mathbf{s}_1) & \cdots & C(\mathbf{s}_n,\mathbf{s}_n) & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} C(\mathbf{s}_0,\mathbf{s}_1) \\ \vdots \\ C(\mathbf{s}_0,\mathbf{s}_n) \\ 1 \end{bmatrix} = \begin{bmatrix} w_1(\mathbf{s}_0) \\ \vdots \\ w_n(\mathbf{s}_0) \\ \varphi \end{bmatrix} \tag{1.3.5}$$

where φ is the so-called *Lagrange multiplier*.

In addition to estimation of values at new locations, a statistical spatial prediction technique produces a measure of associated uncertainty of making predictions by using a given model. In geostatistics, this is often referred to as the **prediction variance**, i.e. the estimated variance of the prediction error. OK variance is defined as the weighted average of covariances from the new point ($\mathbf{s}_0$) to all calibration points ($\mathbf{s}_1, .., \mathbf{s}_n$), plus the Lagrange multiplier (Webster and Oliver, 2001, p.183):

$$\hat{\sigma}^2_{0K}(\mathbf{s}_0) = (C_0 + C_1) - \mathbf{c}_0^{\mathrm{T}} \cdot \boldsymbol{\lambda}_0 = C_0 + C_1 - \sum_{i=1}^{n} w_i(\mathbf{s}_0) \cdot C(\mathbf{s}_0, \mathbf{s}_i) + \varphi \tag{1.3.6}$$

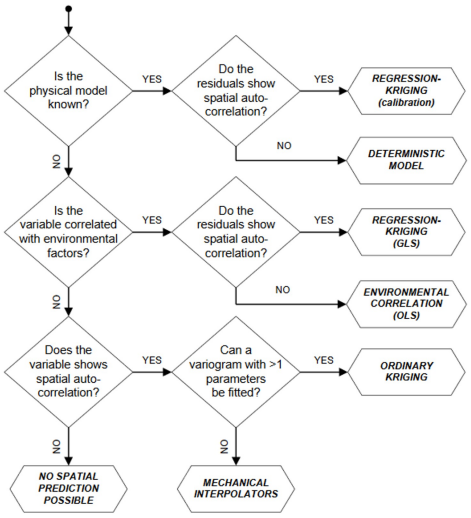
Weight W is calculated in such a way that the error is minimized (OLS).
Distance metric used depends upon the data points as well as geographical coordinates.

Residual check of model -

Residuals are useful in checking whether a model has adequately captured the information in the data. A good forecasting method will yield residuals with the following properties:

1. The residuals are uncorrelated. If there are correlations between residuals, then there is information left in the residuals which should be used in computing forecasts.
2. The residuals have zero mean. If the residuals have a mean other than zero, then the forecasts are biased

From <https://otexts.com/fpp2/residuals.html>


Is the physical model known?
YES
Do the residuals show spatial auto-correlation?
YES
REGRESSION-KRIGING (calibration)
NO
DETERMINISTIC MODEL
NO
Is the variable correlated with environmental factors?
YES
Do the residuals show spatial auto-correlation?
YES
REGRESSION-KRIGING (GLS)
NO
ENVIRONMENTAL CORRELATION (OLS)
NO
Does the variable shows spatial auto-correlation?
YES
Can a variogram with >1 parameters be fitted?
YES
ORDINARY KRIGING
NO
NO SPATIAL PREDICTION POSSIBLE
NO
MECHANICAL INTERPOLATORS