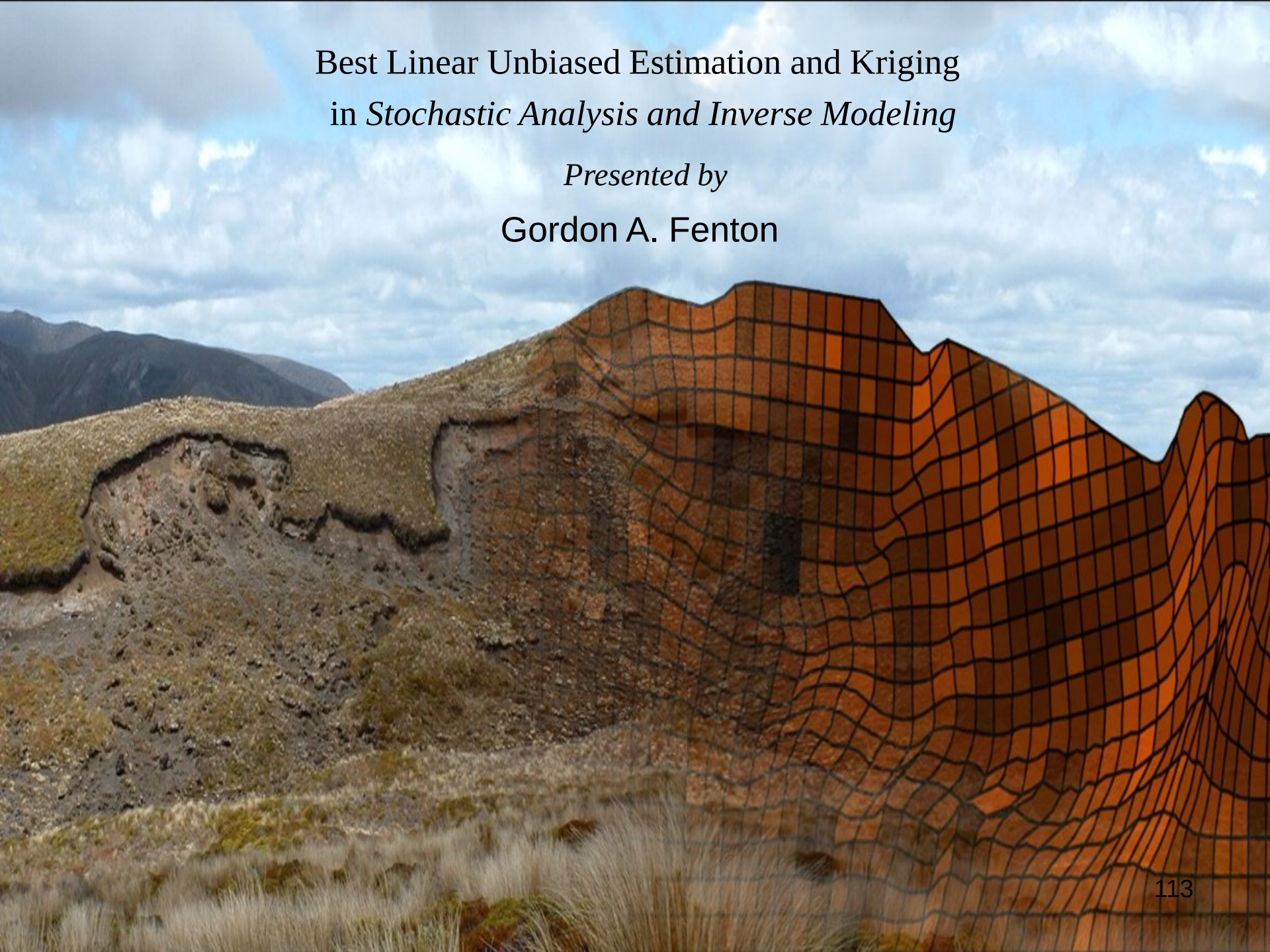
Best Linear Unbiased Estimation and Kriging in *Stochastic Analysis and Inverse Modeling*

Presented by

Gordon A. Fenton



Introduction

We often want some way of estimating what is happening at unobserved locations given nearby observations. Common ways of doing so are as follows;

1. Linear Regression: we fit a curve to the data and use that curve to interpolate/extrapolate. Uses only geometry and the data values to determine the curve (by minimizing the sum of squared errors).
2. Best Linear Unbiased Estimation (BLUE): similar to Linear Regression, except that correlation is used, instead of geometry, to obtain the best fit.
3. Kriging: similar to BLUE except the mean is estimated rather than being assumed known.

Best Linear Unbiased Estimation (BLUE)

We express our estimate at the unobserved location X_{n+1}, as a linear combination of our observations, $X_1, X_2, \ldots, X_n$;

$$\hat{X}_{n+1} = \mu_{n+1} + \sum_{k=1}^{n} \beta_k \left(X_k - \mu_k \right)$$

Notice that position does not enter into this estimate at all – we will determine the unknown coefficients, β_k , using 1st and 2nd moment information only (i.e. means and covariances).

To determine the "Best" estimator, we look at the *estimator error*, defined as the difference between our estimate and the actual value;

$$E = X_{n+1} - \hat{X}_{n+1} = X_{n+1} - \mu_{n+1} - \sum_{k=1}^{n} \beta_k \left(X_k - \mu_k \right)$$

Best Linear Unbiased Estimation (BLUE)

Aside: If the covariances are known, then they include information both about distances between observation points, but also about the effects of differing geological units.

Linear regression considers only distance between points. Thus, regression cannot properly account for two observations which are close together, but lie in largely independent layers. The covariance between these two points will be small, so BLUE will properly reflect their effective influence on one another.

Best Linear Unbiased Estimation (BLUE)

To make the estimator error as small as possible, its mean should be zero and its variance minimal. The mean is automatically zero by the selected form of the estimator;

$$\begin{aligned}
\mathrm{E}\left[X_{n+1} - \hat{X}_{n+1} \right] &= \mathrm{E}\left[X_{n+1} - \mu_{n+1} - \sum_{k=1}^{n} \beta_k \left(X_k - \mu_k \right) \right] \\
&= \mu_{n+1} - \mu_{n+1} - \sum_{k=1}^{n} \beta_k \mathrm{E}\left[X_k - \mu_k \right] \\
&= -\sum_{k=1}^{n} \beta_k \left(\mu_k - \mu_k \right) \\
&= 0
\end{aligned}$$

In other words, this estimator is *unbiased*.

Best Linear Unbiased Estimation (BLUE)

Now we want to minimize the variance of the estimator error:

$$\begin{aligned}\mathrm{Var}\left[X_{n+1}-\hat{X}_{n+1}\right] &= \mathrm{E}\left[\left(X_{n+1}-\hat{X}_{n+1}\right)^2\right] \\ &= \mathrm{E}\left[X_{n+1}^2 - 2X_{n+1}\hat{X}_{n+1} + \hat{X}_{n+1}^2\right] \\ &= \mathrm{E}\left[X_{n+1}^2\right] - 2\mathrm{E}\left[X_{n+1}\hat{X}_{n+1}\right] + \mathrm{E}\left[\hat{X}_{n+1}^2\right]\end{aligned}$$

To simplify the algebra, we will assume that $\mu = 0$ everywhere (this is no loss in generality). In this case, our estimator simplifies to

$$\hat{X}_{n+1} = \sum_{k=1}^{n} \beta_k X_k$$

Best Linear Unbiased Estimation (BLUE)

Our estimator error variance becomes

$$\mathrm{Var}\left[X_{n+1} - \hat{X}_{n+1} \right] = \mathrm{E}\left[X_{n+1}^2 \right] - 2\sum_{k=1}^{n} \beta_k \mathrm{E}\left[X_{n+1} X_k \right] + \sum_{k=1}^{n}\sum_{j=1}^{n} \beta_k \beta_j \mathrm{E}\left[X_k X_j \right]$$

$$= \mathrm{Var}\left[X_{n+1}^2 \right] - 2\sum_{k=1}^{n} \beta_k \mathrm{Cov}\left[X_{n+1}, X_k \right] + \sum_{k=1}^{n}\sum_{j=1}^{n} \beta_k \beta_j \mathrm{Cov}\left[X_k, X_j \right]$$

We minimize this with respect to our unknown coefficients, $\beta_1, \beta_2, \ldots, \beta_n$, by setting derivatives to zero;

$$\frac{\partial}{\partial \beta_l} \mathrm{Var}\left[X_{n+1} - \hat{X}_{n+1} \right] = 0 \qquad \text{for } l = 1, 2, \ldots, n$$

Best Linear Unbiased Estimation (BLUE)

$$\frac{\partial}{\partial \beta_l} \mathrm{Var}\left[X_{n+1} \right] = 0$$

$$\frac{\partial}{\partial \beta_l} \sum_{k=1}^{n} \beta_k \mathrm{Cov}\left[X_{n+1}, X_k \right] = \mathrm{Cov}\left[X_{n+1}, X_l \right] = b_l$$

$$\frac{\partial}{\partial \beta_l} \sum_{k=1}^{n} \sum_{j=1}^{n} \beta_j \beta_k \mathrm{Cov}\left[X_j, X_k \right] = 2 \sum_{k=1}^{n} \beta_k \mathrm{Cov}\left[X_l, X_k \right] = 2 \sum_{k=1}^{n} \beta_k C_{lk}$$

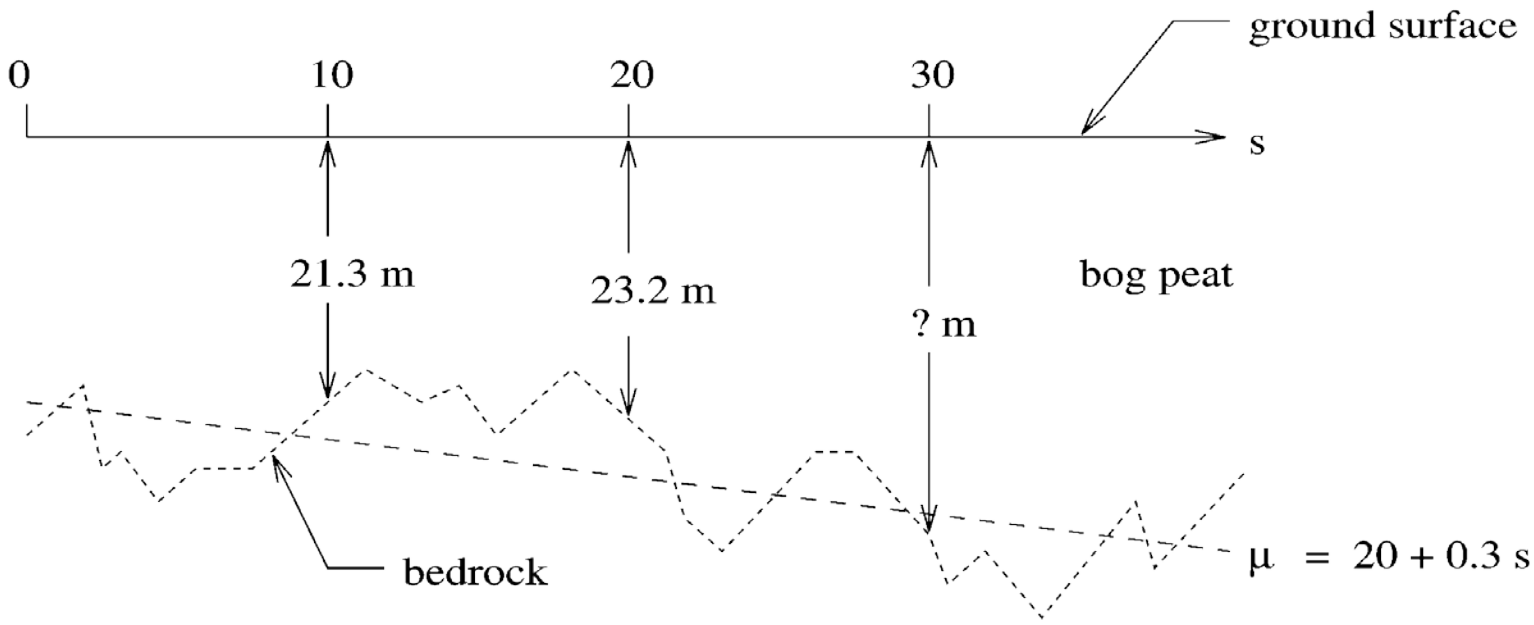
so that $\frac{\partial}{\partial \beta_l} \mathrm{Var}\left[X_{n+1} - \hat{X}_{n+1} \right] = -2b_l + 2 \sum_{k=1}^{n} \beta_k C_{lk} = 0$

in other words, β is the solution to

$$\mathbf{C}\beta = \mathbf{b} \quad \rightarrow \quad \beta = \mathbf{C}^{-1}\mathbf{b}$$

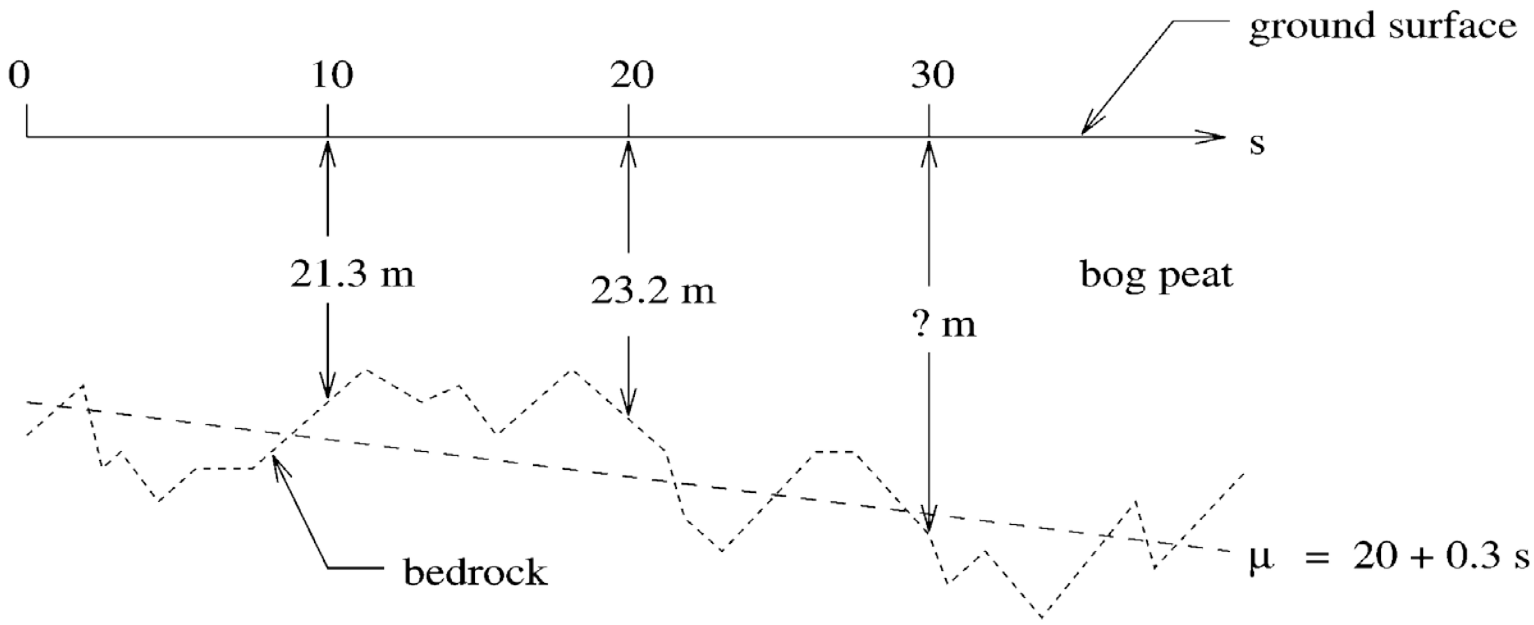
BLUE: Example

Suppose that ground penetrating radar suggests that the mean depth to bedrock, μ, in metres, shows a slow increase with distance, s, in metres, along the proposed line of a roadway.



BLUE: Example

However, because of various impediments, we are unable to measure beyond $s = 20$ m. The best estimate of the depth to bedrock at $s = 30$ m is desired.



BLUE: Example

Suppose that the following covariance function has been established regarding the variation of bedrock depth with distance

$$C(\tau) = \sigma_X^2 \exp\left\{-\frac{|\tau|}{40}\right\}$$

where σ_X = 5 m and τ is the separation distance between points.

We want to estimate the bedrock depth, X_3, at s = 30 m, given the following observations of X_1 and X_2 at s = 10 m and 20 m, respectively :

at s = 10 m, x_1 = 21.3 m

at s = 20 m, x_2 = 23.2 m

BLUE: Example

Solution:

We start by finding the components of the covariance matrix and vector:

$$\mathbf{b} = \begin{Bmatrix} \mathrm{Cov}\left[X_1, X_3\right] \\ \mathrm{Cov}\left[X_2, X_3\right] \end{Bmatrix} = \sigma_X^2 \begin{Bmatrix} e^{-20/40} \\ e^{-10/40} \end{Bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} \mathrm{Cov}\left[X_1, X_1\right] & \mathrm{Cov}\left[X_1, X_2\right] \\ \mathrm{Cov}\left[X_2, X_1\right] & \mathrm{Cov}\left[X_2, X_2\right] \end{bmatrix} = \sigma_X^2 \begin{bmatrix} 1 & e^{-10/40} \\ e^{-10/40} & 1 \end{bmatrix}$$

Note that **b** contains the covariances between the observation and prediction points, while **C** contains the covariances between observation points only. This means that it is simple to compute predictions at other points **(C** only needs to be inverted once**).**

BLUE: Example

Now we want to solve $\mathbf{C}\beta = \mathbf{b}$ for the unknown coefficients β

$$\sigma_X^2 \begin{bmatrix} 1 & e^{-10/40} \\ e^{-10/40} & 1 \end{bmatrix} \begin{Bmatrix} \beta_1 \\ \beta_2 \end{Bmatrix} = \sigma_X^2 \begin{Bmatrix} e^{-20/40} \\ e^{-10/40} \end{Bmatrix}$$

Notice that the variance cancels out – this is typical of stationary processes. Solving, gives us

$$\begin{Bmatrix} \beta_1 \\ \beta_2 \end{Bmatrix} = \begin{bmatrix} 1 & e^{-10/40} \\ e^{-10/40} & 1 \end{bmatrix}^{-1} \begin{Bmatrix} e^{-20/40} \\ e^{-10/40} \end{Bmatrix} = \begin{Bmatrix} 0 \\ e^{-10/40} \end{Bmatrix}$$

Thus, $\beta_1 = 0$ and $\beta_2 = e^{-10/40}$.

Note the Markov property: the 'future' (X_3) depends only on the most recent past (X_2) and not on the more distance past (X_1).

BLUE: Example

The optimal linear estimate of X_3 is thus

$$\begin{aligned}\hat{x}_3 &= \mu(30) + e^{-10/40}\left(x_2 - \mu(20)\right) \\ &= \left[20 + 0.3(30)\right] + e^{-1/4}\left(23.2 - 20 - 0.3(20)\right) \\ &= 29.0 - 2.8e^{-1/4} \\ &= 26.8 \text{ m}\end{aligned}$$

BLUE: Estimator Error

Once the best linear unbiased estimate has been determined, it is of interest to ask how confident are we in our estimate?

If we reconsider our zero mean process, then our estimator is given by

$$\hat{X}_{n+1} = \sum_{k=1}^{n} \beta_k X_k$$

which has variance

$$\mathrm{Var}\left[\hat{X}_{n+1}\right] = \sigma^2_{\hat{X}} = \mathrm{Var}\left[\sum_{k=1}^{n} \beta_k X_k\right]$$

$$= \sum_{k=1}^{n}\sum_{j=1}^{n} \beta_k \beta_j \mathrm{Cov}\left[X_k, X_j\right]$$

$$= \beta^T \mathbf{C} \beta = \beta^T \mathbf{b}$$

BLUE: Estimator Error

The estimator variance is often more of academic interest. We are typically more interested in asking questions such as: What is the probability that the true value of X_{n+1} exceeds our estimate by a certain amount? For example, we may want to compute

$$\mathrm{P}\left[X_{n+1} > \hat{X}_{n+1} + b \right] = \mathrm{P}\left[X_{n+1} - \hat{X}_{n+1} > b \right]$$

where b is some constant. To determine this, we need to know the distribution of the *estimator error* $(X_{n+1} - \hat{X}_{n+1})$

If X is normally distributed, then our estimator error is also normally distributed with mean zero, since unbiased,

$$\mu_E = \mathrm{E}\left[X_{n+1} - \hat{X}_{n+1} \right] = 0$$

BLUE: Estimator Error

The variance of the estimator error is

$$\sigma_E^2 = \mathrm{Var}\left[X_{n+1}^2\right] - 2\sum_{k=1}^{n}\beta_k \mathrm{Cov}\left[X_{n+1}, X_k\right] + \sum_{k=1}^{n}\sum_{j=1}^{n}\beta_k\beta_j \mathrm{Cov}\left[X_k, X_j\right]$$

$$= \sigma_X^2 + \beta^T\mathbf{C}\beta - 2\beta^T\mathbf{b}$$

$$= \sigma_X^2 + \beta^T\mathbf{C}\beta - \beta^T\mathbf{b} - \beta^T\mathbf{b}$$

$$= \sigma_X^2 + \beta^T(\mathbf{C}\beta - \mathbf{b}) - \beta^T\mathbf{b}$$

$$= \sigma_X^2 - \beta^T\mathbf{b}$$

where we made use of the fact that β is the solution to $\mathbf{C}\beta = \mathbf{b}$, or, equivalently, $\mathbf{C}\beta - \mathbf{b} = \mathbf{0}$

BLUE: Conditional Distribution

The estimator $\hat{X}_{n+1}$ is also the conditional mean of X_{n+1} given the observations. That is,

$$\mathrm{E}\left[X_{n+1} \mid X_1, X_2, \ldots, X_n\right] = \hat{X}_{n+1}$$

The condition variance of X_{n+1} is just the variance of the estimator error;

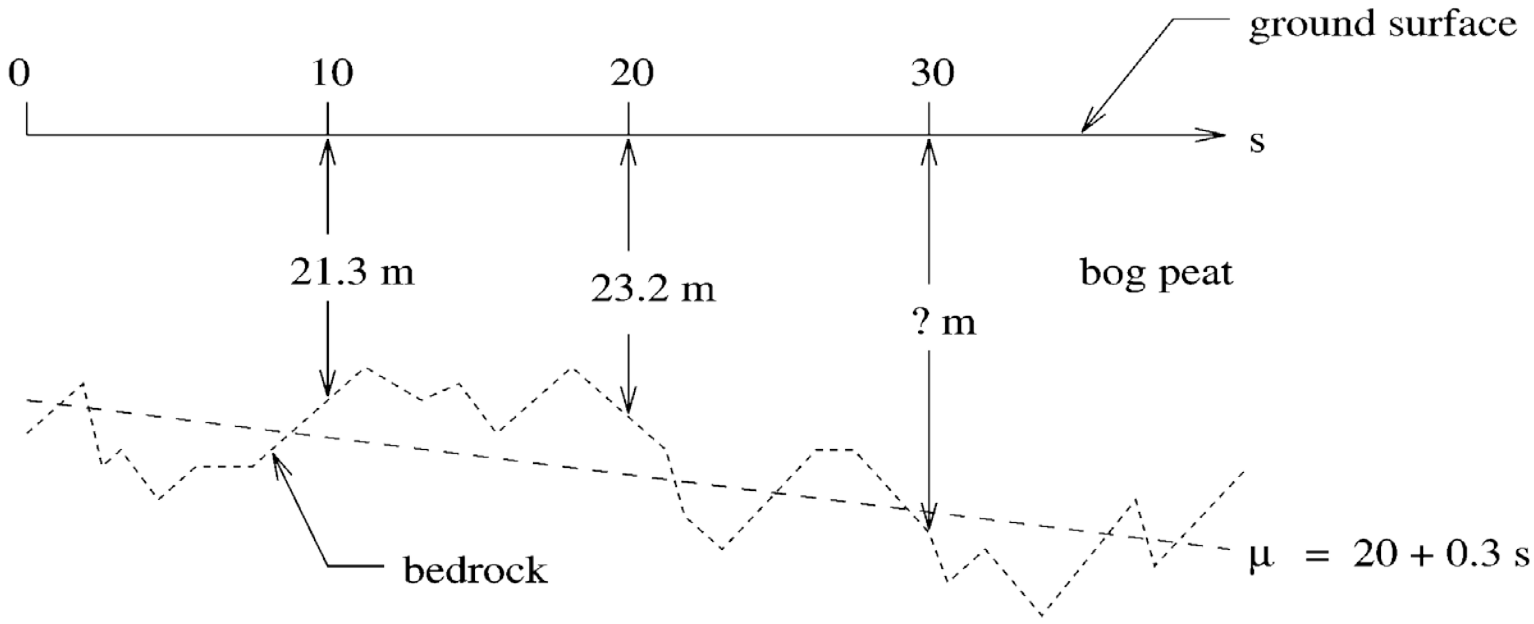
$$\mathrm{Var}\left[X_{n+1} \mid X_1, X_2, \ldots, X_n\right] = \sigma_E^2$$

In general, questions regarding the probability that X_{n+1} lies in some region should employ the conditional mean and variance of X_{n+1}, since this makes use of all of the information.

BLUE: Example

Consider again the previous example.

1. compute the variance of the estimator and the estimator error
2. estimate the probability that X_3 exceeds $\hat{X}_3$ by more than 4 m.



BLUE: Example

Solution:

We had $\mathbf{C} = \sigma_X^2 \begin{bmatrix} 1 & e^{-1/4} \\ e^{-1/4} & 1 \end{bmatrix}$

and $\beta = \begin{Bmatrix} 0 \\ e^{-10/40} \end{Bmatrix}, \qquad \mathbf{b} = \sigma_X^2 \begin{Bmatrix} e^{-2/4} \\ e^{-1/4} \end{Bmatrix}$

so that $\sigma_{\hat{X}}^2 = \mathrm{Var}\left[\hat{X}_3\right] = \beta^T \mathbf{b} = 5^2 \begin{Bmatrix} 0 & e^{-1/4} \end{Bmatrix} \begin{Bmatrix} e^{-2/4} \\ e^{-1/4} \end{Bmatrix} = 5^2 e^{-2/4}$

which gives $\sigma_{\hat{X}} = 5e^{-1/4} = 3.894 \text{ m}$

BLUE: Example

The covariance vector found previously was $\mathbf{b} = \sigma_X^2 \begin{Bmatrix} e^{-2/4} \\ e^{-1/4} \end{Bmatrix}$

The variance of the estimator error is then

$$\sigma_E^2 = \mathrm{Var}\left[X_3 - \hat{X}_3 \right] = \sigma_X^2 - \beta^T \mathbf{b}$$

$$= \sigma_X^2 - \sigma_X^2 \begin{Bmatrix} 0 & e^{-1/4} \end{Bmatrix} \begin{Bmatrix} e^{-2/4} \\ e^{-1/4} \end{Bmatrix}$$

$$= 5^2 \left(1 - e^{-2/4} \right)$$

The standard deviation is thus $\sigma_E = 5\sqrt{1 - e^{-2/4}} = 3.136$ m

This is less than the estimator variability and significantly less than the variability of X ($\sigma_X = 5$). This is due to the restraining effect of correlation between points.

BLUE: Example

We wish to compute the probability $\mathrm{P}\left[X_3 - \hat{X}_3 > 4 \right]$

We first need to assume a distribution for $\left(X_3 - \hat{X}_3 \right)$

Let us assume that X is normally distributed. Then since the estimator $\hat{X}_3$ is simply a sum of X's, it too must be normally distributed. This, in turn implies that the quantity $\left(X_3 - \hat{X}_3 \right)$ is normally distributed.

Since $\hat{X}_3$ is an unbiased estimate of X_3, $\mu_E = \mathrm{E}\left[X_3 - \hat{X}_3 \right] = 0$ and we have just computed σ_E = 3.136 m. Thus,

$$\mathrm{P}\left[X_3 - \hat{X}_3 > 4 \right] = \mathrm{P}\left[Z > \frac{4 - \mu_E}{\sigma_E} \right] = 1 - \Phi\left(\frac{4 - 0}{3.136} \right) = 0.1003$$

Geostatistics: Kriging

Kriging (named after Danie Krige, South Africa, 1951) is basically BLUE with the added ability to estimate the mean. The purpose is to provide a best estimate of the random field at unobserved points. The kriged estimate is, again, a linear combination of the observations,

$$\hat{X}(\mathbf{x}) = \sum_{k=1}^{n} \beta_k X_k$$

where **x** is the spatial position of the unobserved value being estimated. The unknown coefficients, $\boldsymbol{\beta}$, are determined by considering the covariance between the observations and the prediction point.

Geostatistics: Kriging

In Kriging, the mean is expressed as a regression

$$\mu_X(\mathbf{x}) = \sum_{i=1}^{m} a_i g_i(\mathbf{x})$$

where $g_i(\mathbf{x})$ is a specified function of spatial position **x**. Usually $g_1(x) = 1$, $g_2(x) = x$, $g_3(x) = x^2$, and so on in 1-D. Similarly in higher dimensions. As in a regression analysis, the $g_i(\mathbf{x})$ functions should be (largely) linearly independent over the domain of the regression (i.e. the site).

The unknown Kriging weights, $\boldsymbol{\beta}$, are obtained as the solution to the matrix equation

$$\mathbf{K}\beta = \mathbf{M}$$

Geostatistics: Kriging

where **K** and **M** depend on the mean and covariance structure. In detail, **K** has the form

$$\mathbf{K} = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} & g_1(\mathbf{x}_1) & g_2(\mathbf{x}_1) & \cdots & g_m(\mathbf{x}_1) \\ C_{21} & C_{22} & \cdots & C_{2n} & g_1(\mathbf{x}_2) & g_2(\mathbf{x}_2) & \cdots & g_m(\mathbf{x}_2) \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} & g_1(\mathbf{x}_n) & g_2(\mathbf{x}_n) & \cdots & g_m(\mathbf{x}_n) \\ g_1(\mathbf{x}_1) & g_1(\mathbf{x}_2) & \cdots & g_1(\mathbf{x}_n) & 0 & 0 & \cdots & 0 \\ g_2(\mathbf{x}_1) & g_2(\mathbf{x}_2) & \cdots & g_2(\mathbf{x}_n) & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ g_m(\mathbf{x}_1) & g_m(\mathbf{x}_2) & \cdots & g_m(\mathbf{x}_n) & 0 & 0 & \cdots & 0 \end{bmatrix}$$

where C_{ij} is the covariance between X_i and X_j.

Geostatistics: Kriging

The vectors $\boldsymbol{\beta}$ and $\mathbf{M}$ have the form

$$\beta = \begin{Bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_n \\ -\eta_1 \\ -\eta_2 \\ \cdot \\ \cdot \\ \cdot \\ -\eta_m \end{Bmatrix} \qquad \mathbf{M} = \begin{Bmatrix} C_{1\mathbf{x}} \\ C_{2\mathbf{x}} \\ \cdot \\ \cdot \\ \cdot \\ C_{n\mathbf{x}} \\ g_1(\mathbf{x}) \\ g_2(\mathbf{x}) \\ \cdot \\ \cdot \\ \cdot \\ g_m(\mathbf{x}) \end{Bmatrix}$$

where η_i are Lagrangian parameters used to solve the variance minimization problem subject to the non-bias conditions and $C_{i\mathbf{x}}$ are the covariances between the ith observation point and the prediction location, **x**.

Geostatistics: Kriging

The matrix **K** is purely a function of the observation point locations and their covariances – it can be inverted once and then used repeatedly to produce a field of best estimates (at each prediction point, only the RHS vector **M** changes).

The Kriging method depends on

1. knowledge of how the mean varies functionally with position (i.e., g_1, g_2, … need to be specified), and
2. knowledge of the covariance structure of the field.

Usually, assuming a mean which is constant ($m = 1$, $g_1(\mathbf{x}) = 1$, $a_1 = \mu_X$), or linearly varying is sufficient.

Estimator Error:

The estimator error is the difference between the true (random) value $X(\mathbf{x})$ and its estimate $\hat{X}(\mathbf{x})$. The estimator is unbiased, so that

$$\mu_E = \mathrm{E}\left[X(\mathbf{x}) - \hat{X}(\mathbf{x}) \right] = 0$$

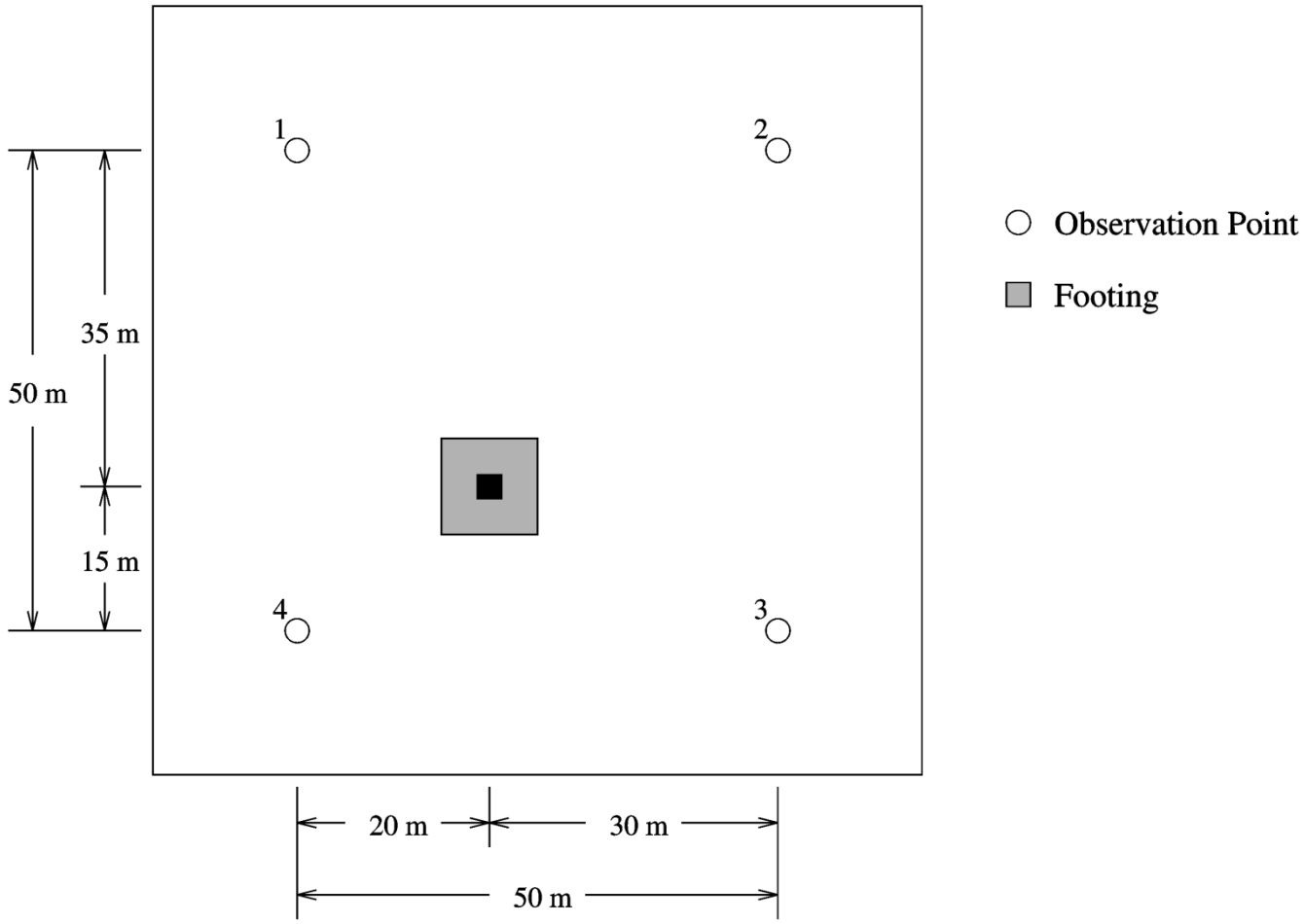
and its variance is given by

$$\sigma_E^2 = \mathrm{E}\left[\left(X(\mathbf{x}) - \hat{X}(\mathbf{x}) \right)^2 \right] = \sigma_X^2 + \beta_n^T \left(\mathbf{K}_{n\times n} \beta_n - 2\mathbf{M}_n \right)$$

where $\boldsymbol{\beta}_n$ and $\mathbf{M}_n$ are the first n elements of the vectors $\boldsymbol{\beta}$ and $\mathbf{M}$, and $\mathbf{K}_{n\times \mathrm{n}}$ is the $n \times n$ upper left submatrix of $\mathbf{K}$ containing the covariances. As with BLUE, $\hat{X}(\mathbf{x})$ is the conditional mean of $X(\mathbf{x})$. The conditional variance of $X(\mathbf{x})$ is σ_E^2.

Example: Foundation Consolidation Settlement

Consider the estimation of consolidation settlement under a footing. Assume that soil samples/tests have been obtained at 4 nearby locations.



Example: Foundation Consolidation Settlement

The samples and local stratigraphy are used to estimate the soil parameters C_c, e_o, H, and p_o appearing in the consolidation settlement equation

$$S = N\left(\frac{C_c}{1+e_o}\right)H\log_{10}\left(\frac{p_o+\Delta p}{p_o}\right)$$

where S = settlement

N = model error random variable ($\mu_N = 1.0$, $\sigma_N = 0.1$)

e_o = initial void ratio

C_c = compression index

p_o = initial effective overburden stress

Δp = mid-depth stress increase due to applied footing load

H = depth to bedrock

Example: Foundation Consolidation Settlement

We will assume that the estimation error in obtaining the soil parameters from the samples is negligible compared to field variability, so this source of error will be ignored.

We do, however, include model error through a multiplicative random variable, *N*, which is assumed to have mean 1.0 (i.e. the model correctly predicts the settlement *on average*) and standard deviation 0.1 (i.e. the model has a standard error of 10%).

The mid-depth stress increase due to the footing load, Δp, is assumed to be random with

$$\mathrm{E}\left[\Delta p\right] = 25 \text{ kPa}$$

$$\sigma_{\Delta \mathrm{p}} = \ \ 5 \text{ kPa}$$

Example: Foundation Consolidation Settlement

Table 1 Derived soil sample settlement properties.

Sample Point	C_c	e_o	H (m)	p_o (kPa)
1	0.473	1.42	4.19	186.7
2	0.328	1.08	4.04	181.0
3	0.489	1.02	4.55	165.7
4	0.295	1.24	4.29	179.1
μ	0.396	1.19	4.27	178.1
v	0.25	0.15	0.05	0.05
σ^2	0.0098	0.0318	0.04558	79.3

Assume that all four random fields (C_c, e_o, H, and p_o) are stationary and that the correlation function is estimated from similar sites to be

$$\rho\left(\mathbf{x}_i, \mathbf{x}_j\right) = \exp\left\{-\frac{2\left|\mathbf{x}_i - \mathbf{x}_j\right|}{60}\right\}$$

Example: Foundation Consolidation Settlement

Using the correlation function, we get the following correlation matrix between *sample locations*:

$$\rho = \begin{bmatrix} 1.000 & 0.189 & 0.095 & 0.189 \\ 0.189 & 1.000 & 0.189 & 0.095 \\ 0.095 & 0.189 & 1.000 & 0.189 \\ 0.189 & 0.095 & 0.189 & 1.000 \end{bmatrix}$$

We will assume that the same correlation length (θ = 60 m) applies to all 4 soil parameters. Thus, for example, the covariance matrix between sample points for C_c is $\sigma^2_{C_c}[\rho]$

We will obtain Kriging estimates from each of the four random fields independently – if cross-correlations between the parameters are known, then the method of *co-Kriging* can be applied (this is essentially the same, except with a much larger cross-covariance matrix).

Example: Foundation Consolidation Settlement

The Kriging matrix associated with the depth to bedrock, H, is

$$\mathbf{K}_H = \begin{bmatrix} 0.04558 & 0.00861 & 0.00432 & 0.00861 & 1 \\ 0.00861 & 0.04558 & 0.00861 & 0.00432 & 1 \\ 0.00432 & 0.00861 & 0.04558 & 0.00861 & 1 \\ 0.00861 & 0.00432 & 0.00861 & 0.04558 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

where we assumed *stationarity*, with $m = 1$ and $g_1(\mathbf{x}) = 1$.
If we place the coordinate axis origin at sample location 4, the footing has coordinates $\mathbf{x} = (20, 15)$ m. The RHS vector for H is

$$\mathbf{M}_H = \begin{Bmatrix} \sigma_H^2 \rho(\underset{\sim}{x}_1, \underset{\sim}{x}) \\ \sigma_H^2 \rho(\underset{\sim}{x}_2, \underset{\sim}{x}) \\ \sigma_H^2 \rho(\underset{\sim}{x}_3, \underset{\sim}{x}) \\ \sigma_H^2 \rho(\underset{\sim}{x}_4, \underset{\sim}{x}) \\ 1 \end{Bmatrix} = \begin{Bmatrix} (0.04558)(0.2609) \\ (0.04558)(0.2151) \\ (0.04558)(0.3269) \\ (0.04558)(0.4346) \\ 1 \end{Bmatrix} = \begin{Bmatrix} 0.01189 \\ 0.00981 \\ 0.01490 \\ 0.01981 \\ 1 \end{Bmatrix}$$

Example: Foundation Consolidation Settlement

Solving the matrix equation $\mathbf{K}_H \beta_H = \mathbf{M}_H$ gives the following four weights;

$$\beta_H = \begin{Bmatrix} 0.192 \\ 0.150 \\ 0.265 \\ 0.393 \end{Bmatrix}$$

We can see from this that samples closest (most highly correlated) to the footing are weighted the most heavily. (i.e. sample 4 is closest to the footing)

All four soil parameters will have identical weights.

Example: Foundation Consolidation Settlement

The best estimates of the soil parameters at the footing location are thus,

$$\begin{aligned}
\hat{C}_c &= (0.192)(0.473) + (0.150)(0.328) + (0.265)(0.489) + (0.393)(0.295) = 0.386 \\
\hat{e}_o &= (0.192)(1.42) + (0.150)(1.08) + (0.265)(1.02) + (0.393)(1.24) = 1.19 \\
\hat{H} &= (0.192)(4.19) + (0.150)(4.04) + (0.265)(4.55) + (0.393)(4.29) = 4.30 \\
\hat{p}_o &= (0.192)(186.7) + (0.150)(181.0) + (0.265)(165.7) + (0.393)(179.1) = 177.3
\end{aligned}$$

The estimation errors are given by

$$\sigma_E^2 = \sigma_X^2 + \beta_n^T \left(\mathbf{K}_{n\times n} \beta_n - 2\mathbf{M}_n \right)$$

Since $\mathbf{K}_{n\times n}$ is just the correlation matrix, $\boldsymbol{\rho}$, times the appropriate soil parameter variance (which replaces σ_X^2), and similarly $\mathbf{M}_n$ is just the correlation vector times the appropriate variance, the variance can be factored out

$$\sigma_E^2 = \sigma_X^2 \left[1 + \beta_n^T \left(\rho \beta_n - 2\rho_{\mathbf{x}} \right) \right]$$

Example: Foundation Consolidation Settlement

$$\sigma_E^2 = \sigma_X^2 \left[1 + \beta_n^T \left(\rho \beta_n - 2 \rho_{\mathbf{x}} \right) \right]$$

where $\boldsymbol{\rho}_{\mathbf{x}}$ is the vector of correlation coefficients between the samples and the footing. For the Kriging weights and given correlation function, this gives

$$\sigma_E^2 = 0.719 \sigma_X^2$$

The individual parameter estimation errors are thus

$$\begin{aligned}
\sigma_{Cc}^2 &= (0.009801)(0.719) = 0.00705 \quad \rightarrow \quad \sigma_{Cc} = 0.0839 \\
\sigma_{eo}^2 &= (0.03204)(0.719) = 0.0230 \quad \rightarrow \quad \sigma_{eo} = 0.152 \\
\sigma_H^2 &= (0.04558)(0.719) = 0.0328 \quad \rightarrow \quad \sigma_H = 0.181 \\
\sigma_{po}^2 &= (79.31)(0.719) = 57.02 \quad \rightarrow \quad \sigma_{po} = 7.55
\end{aligned}$$

Example: Foundation Consolidation Settlement

In summary, the variables entering the consolidation settlement formula have the following statistics based on the Kriging analysis:

Variable	Mean	SD	v
N	1.0	0.1	0.1
C_c	0.386	0.0839	0.217
e_o	1.19	0.152	0.128
H (m)	4.30	0.181	0.042
p_o (kPa)	177.3	7.55	0.043
Δp (kPa)	25.0	5.0	0.20

where v is the coefficient of variation (σ/μ).

Example: Foundation Consolidation Settlement

A first-order approximation to the mean settlement is obtained by substituting the mean parameters into the settlement equation;

$$\mu_S = (1.0)\left(\frac{0.386}{1+1.19}\right)(4.30)\log_{10}\left(\frac{177.3+25}{177.3}\right) = 0.0434 \text{ m}$$

Variable	Mean	SD	v
N	1.0	0.1	0.1
C_c	0.386	0.0839	0.217
e_o	1.19	0.152	0.128
H (m)	4.30	0.181	0.042
p_o (kPa)	177.3	7.55	0.043
Δp (kPa)	25.0	5.0	0.20

Example: Foundation Consolidation Settlement

A first-order approximation to the variance of settlement is given by

$$\sigma_S^2 = \sum_{j=1}^{6} \left(\frac{\partial S}{\partial X_j} \sigma_{X_j} \right)_{\mu}^2$$

where X_j is replaced by the 6 random variables, N, C_c, etc, in turn. The subscript μ means that the derivative is evaluated at the mean of all random variables.

Example: Foundation Consolidation Settlement

Evaluation of the partial derivatives of *S* with respect to each random variable gives

X_j	μ_{X_j}	$\left(\frac{\partial S}{\partial X_j}\right)_\mu$	σ_{X_j}	$\left(\frac{\partial S}{\partial X_j}\sigma_{X_j}\right)^2_\mu$
N	1.000	0.04342	0.1000	1.885×10^{-5}
C_c	0.386	0.11248	0.0889	8.906×10^{-5}
e_o	1.19	−0.01983	0.1520	0.908×10^{-5}
H	4.30	0.01010	0.1810	0.334×10^{-5}
p_o	177.3	−0.00023	7.5500	0.300×10^{-5}
Δp	25.0	0.00163	5.0000	6.618×10^{-5}

so that

$$\sigma_S^2 = \sum_{j=1}^{6}\left(\frac{\partial S}{\partial X_j}\sigma_{X_j}\right)^2_\mu = 18.952\times10^{-5} \ \text{m}^2$$

$$\sigma_S = 0.0138 \ \text{m}$$

Example: Foundation Consolidation Settlement

We can use these results to estimate the probability of settlement failure. If the maximum settlement is 0.075 m, and we assume that settlement, *S*, is normally distributed with mean 0.0434 m and standard deviation 0.0138, then the probability of settlement failure is

$$\mathrm{P}\left[S > 0.075\right] = 1 - \Phi\left(\frac{0.075 - 0.0434}{0.0138}\right) = 1 - \Phi\left(2.29\right) = 0.01$$

Conclusions

- Best Linear Unbiased Estimation requires prior knowledge of the mean and the covariance structure of the random field(s)
- if the mean and covariance structure are known, then BLUE is superior to simple linear regression since it more accurately reflects the dependence between random variables than does simply distance
- Kriging is a variant of BLUE in which the mean is estimated from the observations as part of the unbiased estimation process.