



## General Considerations in Using Geospatial Methods for Optimization

This section provides a brief overview of some important concepts in geospatial analysis, which are discussed more fully in the [fundamental concepts](#) section. Keeping these concepts in mind while planning or reviewing geospatial analyses will help the user to avoid [common misapplications](#) of geospatial analysis. These important concepts include:

- variable measurements and spatial correlation
- interpolation and mapmaking
- assessing uncertainty by geospatial methods
- spatial and temporal optimization.

See the [Methods](#) section for further discussion of the methods mentioned in these descriptions.

### Spatial Correlation and Measurement Error

#### ▼[Read more](#)

The first step in a geospatial analysis is to perform exploratory data analysis (EDA). One of the main goals of EDA is to determine whether the sample data exhibit spatial correlation. Spatial correlation is the property that samples that are closer to one another are more similar than samples that are not close to one another. If no spatial correlation is evident, then a geospatial analysis may not be needed or may be counterproductive compared to more traditional statistical methods. The presence or absence of spatial correlation should be noted in the conceptual site model (CSM). At an appropriate sampling density, there should be some spatial (and temporal) correlation for almost all environmental sampling results, particularly for groundwater, surface water, or sediment data.

Some types of data may exhibit measurement error or variability at small scales, such as when duplicate samples have somewhat different measured results. Certain geospatial methods, such as [inverse distance weighting](#) (IDW), [natural neighbor interpolation](#), and most varieties of [kriging](#) require only one data value at a sampled location. Other methods, such as [nonparametric regression](#) (also referred to as local spatial regression) do not require a single value, which makes it easier to analyze field or laboratory replicates or locations with multiple samples collected over a window of time. When using kriging methods, the use of a variogram with a nugget representative of measurement variability can be an option for addressing variable measurements determined by field duplicate samples. Note that if the goal is to include the measurement uncertainty in the kriging estimated (interpolated) values, then the variograms should include the measurement uncertainties; see [Spatial Correlation Models for Advanced Methods](#).

Since geospatial methods estimate the value at an unsampled location by constructing a weighted average of samples near the unsampled location (by different weighting schemes), all geospatial methods implicitly or explicitly assume a positive spatial correlation between nearby locations. That is, it is assumed that the value (for example soil concentration) at an unsampled location is likely to be similar to the closest known values. In some hydrogeologic environments, for example highly fractured bedrock or karst, this assumption may be violated and it may be difficult to accurately apply geospatial methods.

### Interpolation and Mapmaking

#### ▼[Read more](#)

Interpolation in geospatial methods creates estimates at unsampled locations by weighting and averaging the sample data closest to that location. Different methods lead to different estimates, depending on how the data are weighted. For geospatial interpolation from sampled to unsampled locations, kriging may be the best-known and most widely available method. Other methods, including inverse distance weighting, natural neighbor interpolation, and parametric and nonparametric regression, are also used in environmental applications.

[Advanced methods](#) such as kriging have both advantages and disadvantages when compared to other geospatial methods. While kriging is widely used, and generally gives good results, it requires fitting a spatial correlation model (such as a variogram) to the data. Another potential disadvantage of kriging is that it is computationally demanding for large data sets. On the other hand, simple methods such as [inverse distance weighting](#) and [natural neighbor interpolation](#) do not require a

model or variogram and can easily be used with large data sets. Implementing more complex methods, such as local polynomial regression, require selecting parameters that control the methods, but do not require estimating a spatial correlation model. These methods are less computationally demanding than kriging, but more so than the simple methods. Local regression methods are flexible and can be used to incorporate other predictor variables in the estimation, or to interpolate data over time and space simultaneously.

If desired, kriging methods can be forced to 'honor' or exactly interpolate every known data value. This means that the kriging estimate at each sampled location precisely equals the reported concentration (the same is true for inverse distance weighting and natural neighbor interpolation). When those data are known with precision, methods providing exact interpolation perform well.

Environmental samples, however, often have significant uncertainty associated with the sampling and measurement process. The reported sample values can be quite different from the 'actual' or 'true' values. In this case, it may be better to use methods such as spatial regression and kriging with a nugget, which do not exactly interpolate the sample data. Instead, these methods attempt to estimate the underlying spatial distribution while assuming that the data may be in error to some degree. This approach is similar to using linear regression to estimate a trend and noting that the fitted curve does not pass through any or many of the sample values. In a variogram, the use of a nugget (which represents measurement uncertainty and variability smaller than the scale of observation) is a means to account for the uncertainty in the measured values; see [Exact versus Inexact Interpolation](#).

## Assessing Uncertainty by Geospatial Methods

▼ [Read more](#)

Most geospatial methods not only can make maps, but also assess the uncertainty around those estimates. Unlike kriging and the more complex methods, inverse distance weighting and natural neighbor interpolation offer no obvious way to estimate uncertainty.

Although the kriging variance can be used to estimate spatial uncertainty, it is also problematic. For example, the kriging variance at any given unsampled location almost completely depends on the sampling design and arrangement of the sample data. The kriging variance is less influenced by the actual (local) data values, except through the variogram (which necessarily averages data variation across the site). The variance of local spatial regression, by contrast, depends highly on the observed nearby sample values; see also a further discussion on how [kriging variance](#) provides a relative, but not absolute, measurement of local uncertainty.

## Spatial and Temporal Optimization

▼ [Read more](#)

Environmental monitoring optimization often incorporates geospatial methods to search for statistical redundancy in network sampling locations, sampling frequencies, or both over time. Spatial optimization involves selecting the optimal number and arrangement of sampling locations (for example, groundwater monitoring wells, soil, vapor, or sediment sampling points). Temporal optimization involves selecting the optimal sampling frequency. Both spatial and temporal analysis should be considered when optimizing the monitoring network. For additional information, see the discussion on [using the results](#) of the geospatial methods in monitoring optimization and an example of [sampling redundancy](#) analysis.