THE NNSA GROUND-BASED NUCLEAR EXPLOSION MONITORING MULTI-REGION KRIGING MODEL

James R. Hipp

Sandia National Laboratories

Sponsored by National Nuclear Security Administration
Office of Nonproliferation Research and Engineering
Office of Defense Nuclear Nonproliferation

Contract No. DE-AC04-94AL85000

ABSTRACT

The National Nuclear Security Administration (NNSA) Ground-Based Nuclear Explosion Monitoring (GNEM) multi-region kriging model provides a convenient mechanism for merging seismic statistical models developed specifically for bounded regions, over a much broader area in a smooth and continuous fashion. The method involves defining a set of kriging and travel-time model parameters within specific seismic provinces whose boundaries are defined by an enclosing polygon. Each polygon's boundary edges are further characterized using a set of transition blending functions that describe the reduction in influence of the polygon's statistical and model parameters as a point of interest moves from within the polygon boundary to outside the boundary. Spatial domains that are not covered by specific region-dependent polygons inherit a default set of kriging parameters and travel-time models. The entire collection of seismic provinces and the default region determine the extent of the multi-region model.

Region-based statistical and model parameters are averaged at points of interest using an assigned weight based on the amount of influence exerted by each of the regions that are near the point of interest. The weights are normalized to the total influence exerted by all regions influencing the point of interest. Transitions at the vertices of the polygons are smoothed using a patching circle to ensure that the entire bounding influence transition is continuous and differentiable everywhere. With this approach, not only statistical model parameters such as kriging, correlation, shape, and range can be regionally integrated over a broad area, but so can region-dependent model parameters such as total travel time, ellipticity corrections, model variance, and bulk static corrections. In fact, any regional based parameter can be successfully deployed over the entire multi-region domain.

Using this approach, a set of statistically distinct spatial regions can be combined into a single continuous interpolated surface for use by client applications that communicate with the NNSA Parametric Grid Library (PGL). Elements of the multi-region model are currently used by the NNSA Calibration Integration Tool (CIT) and the Client GNEM Interface (CGI) Library.

KEY WORDS: seismic kriging, polygon, multi-region

OBJECTIVE

To improve the U.S. National Data Center (USNDC) capability to perform nuclear explosion monitoring, researchers at the U.S. Department of Energy (DOE) national laboratories are collaborating to develop a collection of data and associated access tools collectively known as the Knowledge Base (KBase). The KBase will consist of many types of information, which will improve the performance of all major monitoring technologies. The number of different data sets in the KBase is expected to be huge, but viewed from an information representation perspective, there are only four basic types of data: reference event information, parametric grid information, contextual information, and supporting information. In this report parametric grid information is the primary focus, and in particular, the new Multi-Region Kriging (MRK) model.

The MRK is an extension to the Modified Bayesian Kriging (MBK) model developed by the national laboratories and described below. The MRK extends the MBK by allowing the development of a set of kriging statistical parameters for a collection of adjacent regions and providing a framework that treats the collection as an unique entity that can provide smooth residual and associated error results across the entire multi-region area. This paper will provide a brief overview of the MBK method and then give a synopsis of the MRK method before describing some of the mechanical details involved in the formulation of an MRK representation.

RESEARCH ACCOMPLISHED

The sections below give an overview of the new Multi-Region Kriging (MRK) capability developed by the national laboratories. The first section describes the Modified Bayesian Kriging (MBK) method which remains the workhorse in both mono- and multi-region kriging problems. The MBK is the primary tool for evaluating seismic path correction residuals and associated errors. References relating to the MBK are provided so that the interested reader can learn more about the method if desired. The next section describes the MRK methodology and defines its basic construction and terminology. The remaining section illustrates in greater detail the process of converting MBK parameters for use in the MRK framework.

Bayesian Kriging Overview

Researchers at the national laboratories associated with the GNEM project use kriging primarily to develop seismic travel-time residual path corrections and associated errors. Other data development efforts, however, are also beginning to employ the method to generate results and associated errors (e.g. magnitude, azimuth and slowness corrections, etc.). The kriging method provides a means of linearly interpolating values from a data set based on the spatial correlation characteristics of the data itself. The term kriging actually refers to a family of interpolation techniques, most of which are Best Linear Unbiased Estimator (BLUE) processes (e.g. Wackernagel, 1995). As with other linear interpolation methods, the kriged estimate of the actual value *Zp* at a given point *Xp* is a weighted combination of the values at the *n* known observations,

$$Z_P = \sum_{i=1}^{n} w_i \left(X_P \right) Z_i \ ,$$

where Z_i and w_i are the value and weight of the i*th* observation, respectively. Kriging has two important features that distinguish it from other linear interpolation methods, however. First, in kriging the weights are determined such that the variance of the predicted error is minimized. Second, kriging produces an estimate of the error associated with the interpolated value.

The kriging method employed by the GNEM project employs a modified version of simple kriging (zero mean kriging). We call the method Modified Bayesian Kriging, or MBK (Schultz et al., 1998). MBK is based on simple kriging where we assume that the mean has been removed from the data before employing the kriging process. The modifications concern the behavior of the observation measurement error and the influence of the observations relative to the distance to an interpolation location.

Generally, kriging imposes an observation dependent modeling error that is the same for all points. However, for most of the physical seismological data sets considered for GNEM analysis the measurement error varies from one observation to the next. For this reason, our MBK model utilizes observation dependent measurement error when performing kriging calculations.

The second modification in MBK allows for use of *a priori* range of influence information for each point. For many typical GNEM data sets the spatial samp ling is poor but the analyst may have some additional information that cannot be derived from the data due to the limited sampling (e.g. the scale length of a tectonic province). In the MBK model the influence of the observations on the kriging result at a distance can be controlled through the use of smooth neighborhood kriging. This method can reduce or eliminate the influence of points based solely on the distance between the point and the interpolation point of interest. This behavior makes MBK behave much better as an extrapolator when interpolating in regions that lack good spatial characteristics (point coverage and density).

These two modifications are the primary changes employed over simple kriging in the MBK model. Other subtle changes do exist but are not relevant to the current discussion. The interested reader who desires more detailed information concerning the MBK method and its development should consult Schultz et al.

Multi-Region Approach

As discussed in the MBK approach above, sometimes the complexity of geophysical regions can force the modification of standard analytical models so that the inherent complexity is more correctly represented. In the case of the MBK, the simple kriging models were extended to incorporate observation dependent measurement errors and distance dependent influence on the kriged results. These modifications tend to yield very good results within the confines of geographical regions where the same set of kriging statistics apply. However, as the spatial extent (range) of the data increases the properties of the statistical parameters used in performing the kriging analysis, such as background model error and the correlation range and shape, can also change. These changes, for example, can be caused by perturbations in geophysical properties, such as amorphous inhomogeneities, or by sharp discontinuities, such as a tectonic province boundaries. Generally, one simply reformulates the statistical properties of distant regions and uses those in place of the original formulation. A modeling problem, however, can occur at or near the boundaries or regions of overlap between two or more models. In those regions different results can be produced at the same location depending on which set of statistical properties are used. This causes discontinuities in the spatial solution from one model to the next. This, in turn, presents difficulties for certain analytical processes (e.g. seismic location) where smoothness guarantees solution stability. What is needed is an approach that captures the complexity sufficiently so that statistical properties can change with position but does so in a smooth and continuous fashion so that there are no solution discontinuities.

To accomplish this we have defined separate statistical regions that are modeled with discrete polygonal boundaries. We refer to these regions interchangeably as kriging or statistical regions, when describing statistical attributes, or polygon regions, when discussing geometric qualities. Regions that are not contained by a specific kriging polygon are said to reside in the default kriging polygon. Figure 1 below shows an example of a set of polygons that define specific statistical regions of influence. Notice that the polygon boundaries may or may not lie adjacent to other separate and independent statistical regions. The polygon outlined in red[1] will be used in the remainder of this discussion to illustrate the particular properties of these statistical or kriging polygons, and their impact relative to providing a global solution over all polygons in the region collection.

Aside from the specific polygon boundaries, we must define a transition region about the fixed polygon boundary of each unique statistical region that controls the amount of influence that the regions statistics confers on any arbitrary location. Further, we require that the change in influence as the position changes be a smooth and continuous process. The transition region serves to shift the influence of the polygon's statistical properties from all influencing, for points that are contained completely within the fixed polygon's inner transition boundary, to no influence, for regions that lie completely outside of the polygon's exterior, or outer, transition boundary. Figure 2a illustrates the concept of a transition region. The extent of the transition region is determined by defining vertices for both the inner and outer transition boundaries. Each inner and outer transition vertex is associated with a specific vertex of the polygon boundary. These vertices have no placement restrictions other than they must form polygons that do not have any self-crossings (edges that intersect) nor do they cross the defining kriging polygon. In other words, the inner transition boundary is completely contained within its associated polygon boundary and does not self-intersect, and the outer transition boundary lies completely exterior to its associated polygon boundary and does not self-intersect.

---

[1] Because this information is impossible to convey without the use of color, and the printed Proceedings are in black and white as a cost-saving measure, we refer the reader to one of the electronic versions of this Proceedings for interpreting the figures in this paper. The electronic versions can be found at http://www.nemre.nn.doe.gov/review2001 or on the CD ROM version of the Proceedings. Or, readers should feel free to contact the authors.

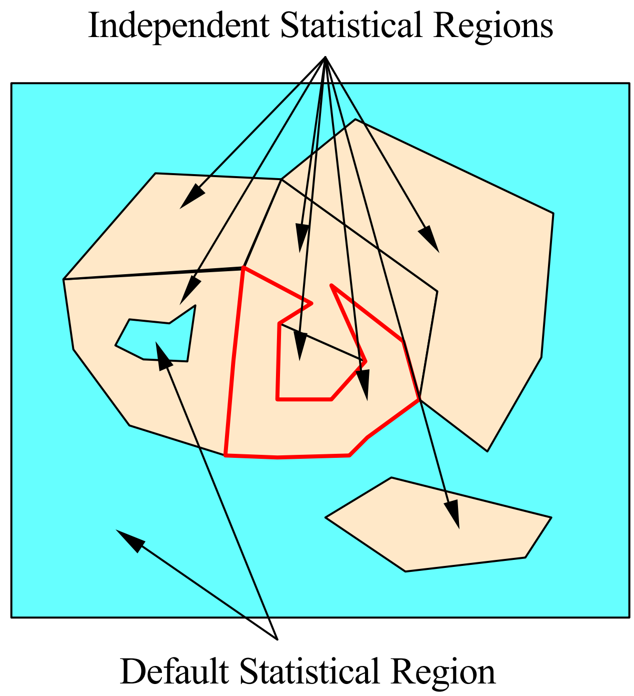

Figure 1. This depiction illustrates an example collection of polygons that represent a set of discrete areas, each with roughly constant statistical properties. The default region is considered to be any and all regions in the model that are not contained within one of the independent statistical regions. The statistical region outlined in **red** is used throughout the remainder of the paper as an example statistical region for purposes of clarifying the details of the model.

Next we map a smooth shape function between the transition boundaries that has a value of one with a slope of zero at the inner transition boundary, and a value of zero with a slope of zero at the outer transition boundary. The shape of the function is not important but it should be monotonically smooth. With this model we can assign a weight for the amount of influence provided by a statistical region relative to an arbitrary point. Points that are on or completely within the inner transition boundary have a weight of one (i.e. complete influence), while points that lie on or completely exterior to the outer boundary have a weight of zero (i.e. no influence). Points that lie between the inner and outer boundaries have an intermediate value between one and zero depending on the shape of the function and the mapping projection at those locations.

Lastly, we ensure differentiability in the transition region around the boundary vertices by specifying a patch circle that connects the edges of a vertex at an equal but arbitrary distance away from the vertex. This distance, referred as the Tangent-Point-Distance, or TPD, can be adjusted to control the amount of curvature around the vertex and must be specified for both the inner and outer transitions at each vertex. The TPD position along each edge that connects with an inner or outer boundary vertex is arbitrary but must be chosen so that the circle formed between two tangent points on either side of a boundary vertex contains the kriging polygon boundary vertex within the confines of the transition region. Figure2b depicts the transition region with the TPD patch circles connected to the inner and outer transition boundaries. Figure 2c shows the same diagram with the completed circle patch at each vertex of the inner and outer transition. Finally, Figure 2d shows the plane view of the transition surface after mapping a hermite shape function into the transition region. The dark red area has a value of one while the blue area has a value of zero. Figure 3 shows a corresponding three-dimensional view of the surface.

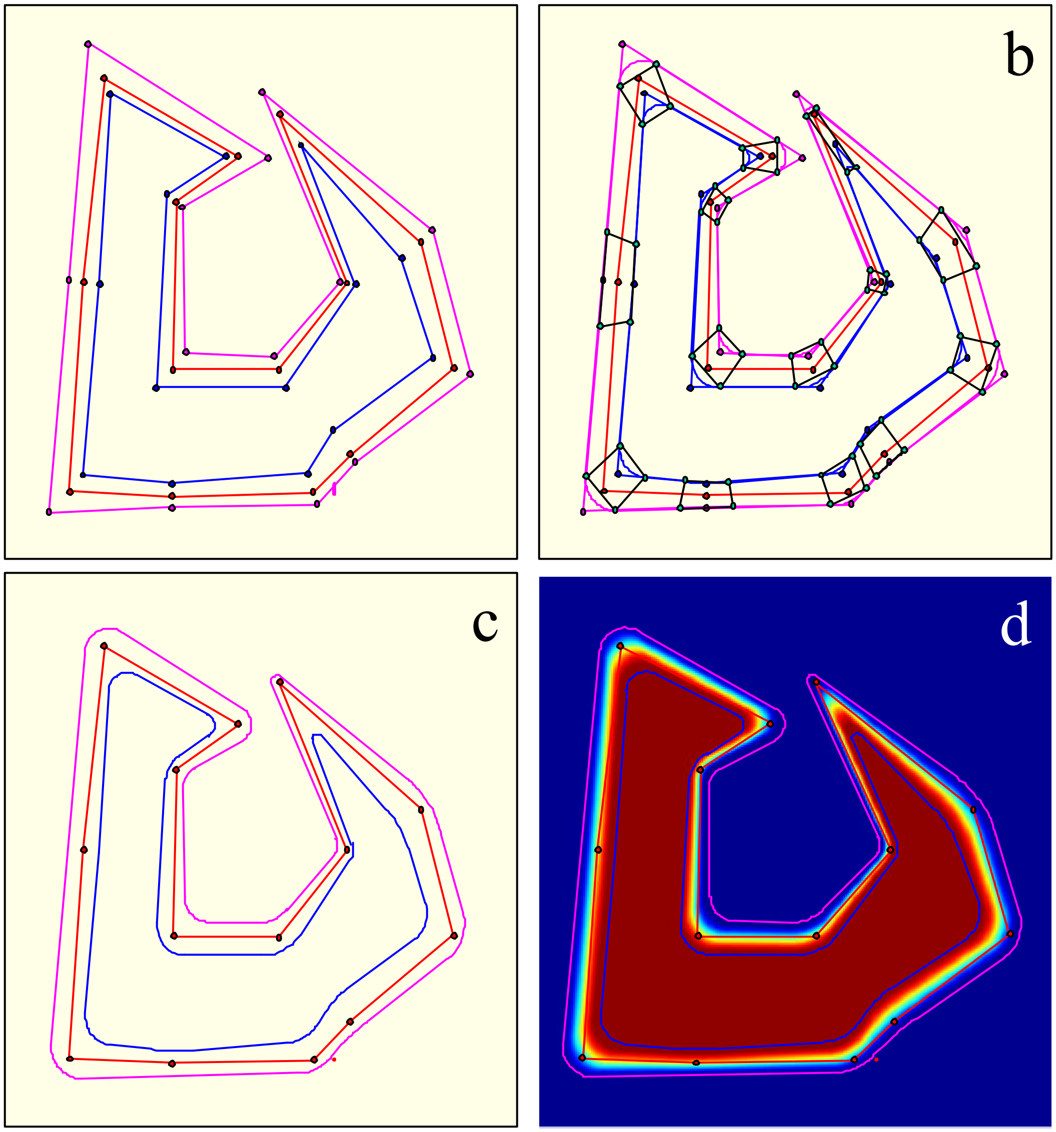

Figure 2. **a)** This panel illustrates an example polygon boundary (**red** edges and vertices) enclosed by an outer transition boundary (**magenta** edges and vertices), and containing an inner transition boundary (**blue** edges and vertices). **b)** This panel shows the construction of the TPD patch points (**green** vertices) with the existing inner and outer transition boundary edges. **c)** This panel depicts the completed smooth inner and outer transition boundaries. **d)** This panel shows a plane view of the final surface with the mapped hermite transition function that moves from a value of one (**dark red**), inside the inner transition boundary, to a value of zero (**dark blue**) outside the outer transition boundary.

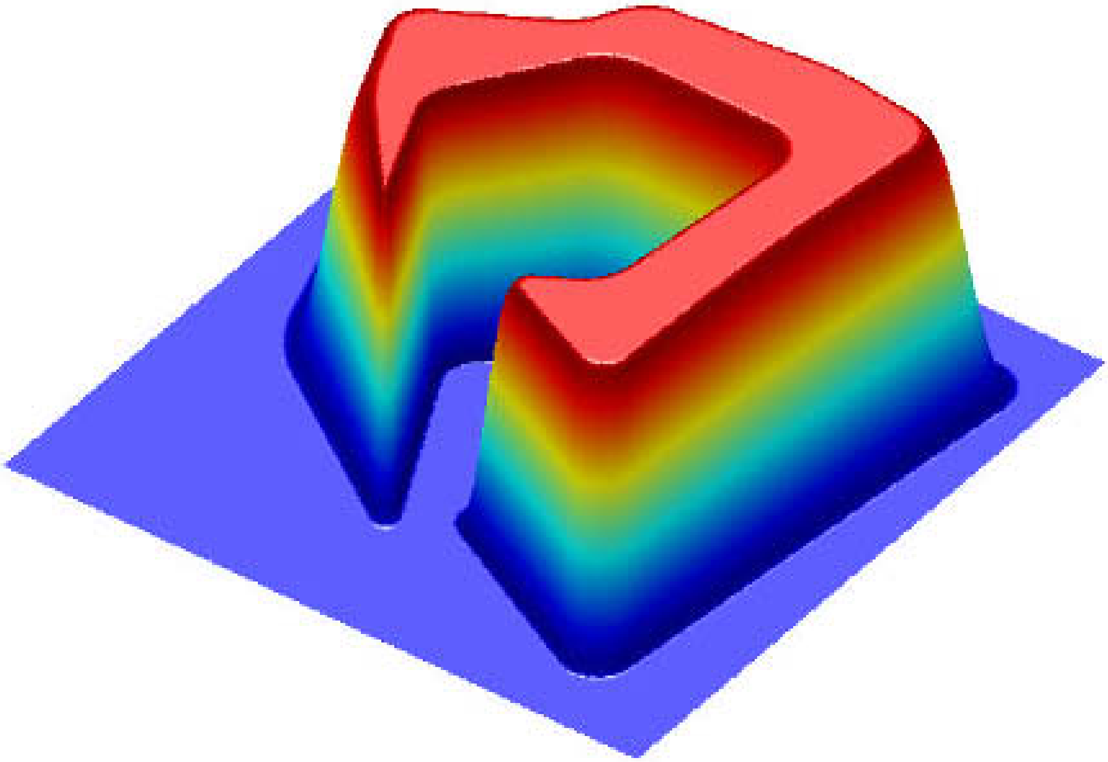

Figure 3. This Figure illustrates a three-dimensional view of the patched and mapped transition weight surface for the example polygon shown in Figure 1 and 2. Notice the smooth curvature everywhere. The shape of the monotonic hermite polynomial function is evident as we move from a region that is completely contained by the inner transition (a weight of one and colored **red**), to a region that is completely outside of the outer transition (a weight of zero and colored **blue**).

Region Based Weights

We now have a means for defining the amount of influence that a particular statistical region can have at an arbitrary point. The researcher is free to design the shape of a statistical region transition zone to be almost anything. Rapid changes in statistical influence caused by crossing over tectonic boundaries are easily modeled by making the distance between the inner and outer transition small. Gradual changes in statistical influence due to large-scale amorphous heterogeneity in the Earth are modeled by making the distance between the inner and outer transition boundaries large. Given this flexible influence weight strategy, we can consider methods for weighting the intrinsic statistical properties of each unique kriging region.

Given the design of the MRK model above, we define the amount of influence that each statistical property should receive based on position. If a position lies completely within the inner transition region of a polygon we allow that statistical property to be totally weighted ($T = 1$ where T is the transition weight of some statistical region). On the other hand, if a position lies completely outside of a polygon's outer transition boundary, we allow no influence for the statistical properties from that region ($T = 0$). Points that lie between the inner and the outer transition boundaries will have some intermediate value (0 ? T ? 1).

Given any position-dependent property of a statistical region we can define an average value accommodating all influencing regions by summing the weighted value of the parameter from each region with its associated transition weight. The transition weights are evaluated at some interpolation position X_P. The expression for this averaging technique can be written as

$$\overline{A}_P = \sum_R T^R(X_P) A^R(X_P),$$

where $A^R(X_P)$ is a regional property to be averaged in the Rth region evaluated at a point X_P, and $T^R(X_P)$ is the weight of the Rth region evaluated at X_P. The weights T^R are normalized across all regions R that influence a specific point X_P. Typical point-dependent parameters that may use this weighting scheme in the MRK model include the background model variance and the correlation range. We also use this scheme to average the base model travel time if the kriging regions defined within the MRK model utilize different base model representations.

We can perform the same weighting for properties that depend on the distance between two points such as correlation. For these cases, however, we need to make some assumptions about how points correlate across different statistical regions. First, if two points lie completely contained within a region's inner transition boundary, and further, they are not contained by any other polygon's outer transition boundary, then the two points are assumed to be correlated completely using the containing region's statistics. Second, if two points are completely contained by different regions such that both points lie completely within their respective regions inner transition boundaries, and further, neither point is contained by any other polygon's outer transition boundary, then the two points are considered to be completely uncorrelated. Figure 4 illustrates these kinds of relationships and depicts other cases that are intermediate to the two described above.

If we assume that each point for which we are trying to determine correlation contributes some weight to the process then in the first case (above) both weights are one, while in the second case at least one weight is zero. This kind of a relationship implies bilinear weighting of the correlation parameter with respect to a region. This relationship can be defined as

$$\overline{r}_{ij}^R = T^R(X_i) T^R(X_j) r_{ij}^R(h_{ij}).$$

In this expression the correlation between the two points, relative to region R's statistics, are scaled by the product of the weights in region R evaluated at both points positions (X_i, and X_j). The parameter h_{ij} is the distance between the two points. Notice that if both points lie completely within region R's inner transition boundary then both weights are 1, and the average value is simply the region R correlation using the distance between the two points. Also notice that if either point lies completely outside of region R's outer transition boundary then at least one weight is zero so that average correlation is zero between the two points. We can normalize these weights to the average value of both weights and sum the effect of other influencing regions to yield an average correlation at any point regardless of how many statistical regions influence the solution. This expression can be written as

$$\overline{r}_{ij} = \frac{\sum_R T^R(X_i) T^R(X_j) r_{ij}^R(h_{ij})}{\sum_R \left(T^R(X_i) + T^R(X_j)\right)/2}.$$

This expression utilizes all influencing statistical correlation definitions to find the average correlation between any two points in a collection of discrete statistical regions. For readers interested in how correlation, correlation range, background model variance, and other parameters are used in the MBK model please see Schultz et al. (1998).

Kriging Region Transition

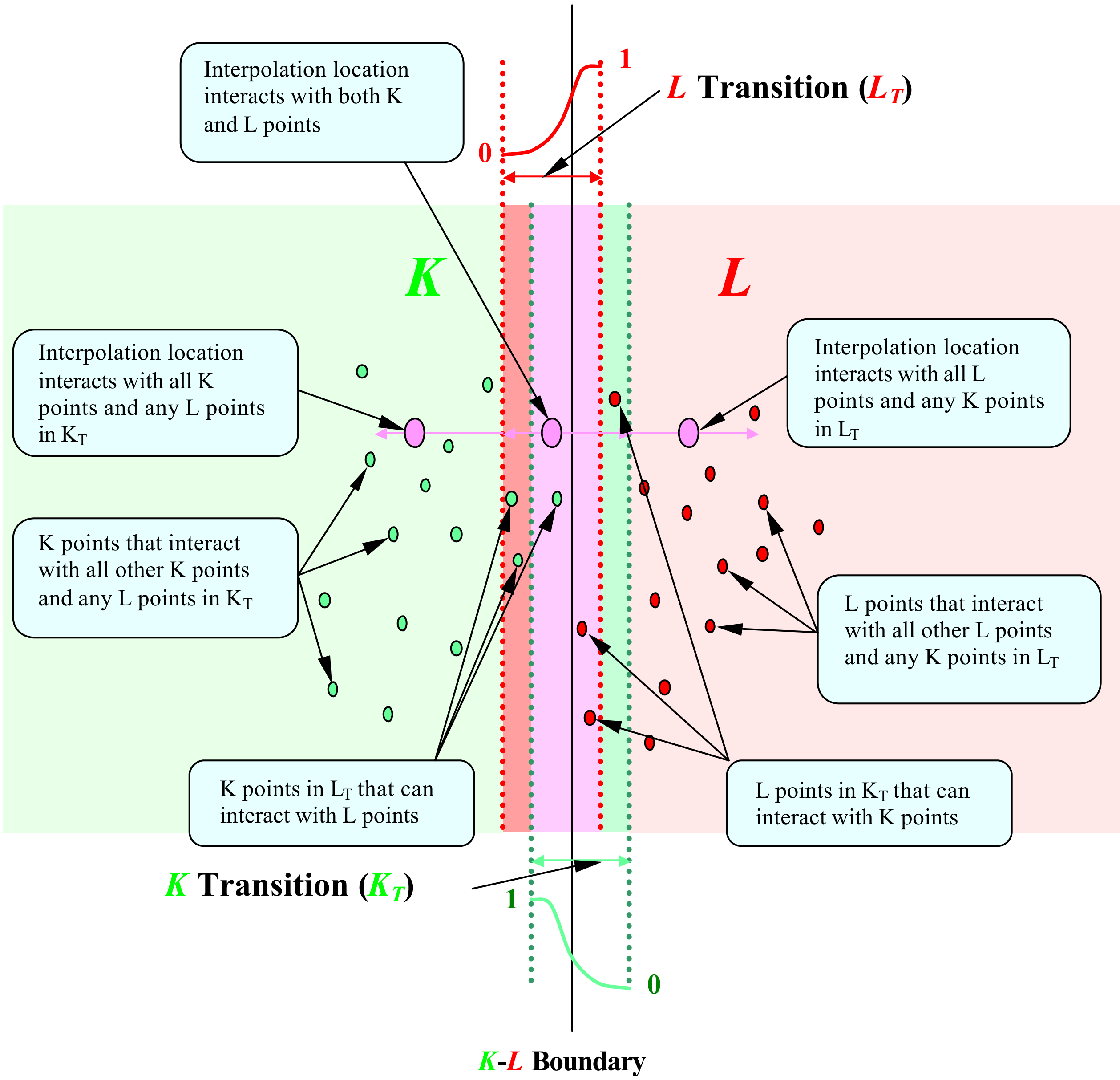

Figure 4. This Figure illustrates the possibilities for data observation correlation interaction between two adjacent kriging regions that share a common boundary. The two regions (***K*** and ***L***) both have defined transition intervals (K_T and L_T) that overlap with one another.

CONCLUSIONS AND RECOMMENDATIONS

In this paper we have presented the new Multi-Region Kriging (MRK) model used by DOE researchers to expand the Modified Bayesian Kriging (MBK) model across regional boundaries where statistical properties differ from one region to the next.

First we summarized the MBK model which uses the basic simple kriging method which attains a solution by minimizing the predicted error, and provides an associated error term at the requested interpolation location. We also defined the major modifications in the MBK model that introduce observation-dependent error and the use of a priori neighborhood range of influence information to aid the MBK model's extrapolation capability.

Next we turned our attention to the MRK model and developed the basic construction methodology and terminology necessary to describe the MRK. This discussion included the definition of a collection of statistically dissimilar spatial regions that may be disjoint from, or adjacent to, other statistical regions in the collection. Then we described the necessary requirements to define a monotonic transition region that provides for a smooth transition in a regions influence weight as a function of spatial position. We also illustrated the process to use circle patching at the vertices of the defining polygons to smoothly change direction from one polygon edge to the next.

Finally, we defined how the statistical properties in a kriging region can be spatially averaged accounting for the regions that influence any arbitrary point. In particular, we considered point-dependent properties such as background modeling variance and correlation range, and point-to-point distance-dependent properties such as correlation.

REFERENCES

Schultz, C., S. Myers, J. Hipp, and C. Young (1998), Nonstationary Bayesian Kriging: Application of Spatial Corrections to Improve Seismic Detection, Location and Identification, *Bull. Seism. Soc. Am*., 1275-1288.

Wackernagel, H. (1995), Multivariate Geostatistics, Springer, Berlin.