Project Title
Parallelized Kriging for Background Radiation Interpolation on Apache Spark

Project Objective
Develop a parallelized Gaussian Process Regression algorithm to be deployed on Apache Spark for background radiation interpolation.

Project Description
The purpose of this project is to leverage large quantities of radiation data collected from a mobile radiation sensor network to generate a near real time background-radiation map. The interpolation of this map will be performed using standard Geostatistics techniques such as Kriging Interpolation (Gaussian Process Regression). This work intends to apply this commonly used interpolation method on the distributed Apache Spark ecosystem for faster processing of large volumes of data. The Kriging method is set apart from other interpolation algorithms because it provides a variance associated with each predicted point. This known variance should prove very useful in outlier detection scenarios.

Project Relevance to Nuclear Nonproliferation
The ability to interpolate background radiation with associated prediction variance will make outlier detection much easier. Detecting outliers is a key task in identifying anomalous sources in a sensor network and crucial to nuclear non-proliferation efforts.
Products and Outcomes of Project

Several preliminary tests have been run on Apache Spark, and the Kriging algorithm has been applied on small datasets of one day. The Kriging algorithm has also been applied to a simulated dataset. Final version is currently in progress, scheduled to be complete this fall.

Presentations


