Scale up of data ensemble generation process for the ARM SGP measured surface energy fluxes

V. Rao Kotamarthi[1], Meredith Franklin[2], Allen Montz1 and Michael Stein[3]

[1]Argonne National Laboratory, [2]University of Southern California, [3]University of Chicago

In this project we describe HPC implementation of the numerical methods used to develop data ensembles of ARM SGP surface flux measurements. The observations are assumed to follow a Gaussian spatial process for which we use maximum likelihood (ML) methods to estimate the parameters. This ML estimation requires optimization of a likelihood function, which at present has been implemented through existing R libraries. Conditional simulation of a Gaussian process provides an interpolated gridded surface. The current challenge is in scaling up our methods to estimate multiple realizations of these gridded surfaces at fine spatial resolutions from the large ARM SGP data sets that span at least one decade at varying time resolutions. Multiple realizations will enable uncertainty estimation in the interpolated observations which is crucial when being used for climate model evaluation.

However, conditional simulation in a dense, two-dimensional grid can be computationally expensive, particularly for very smooth processes, as the standard very fast approach, circulant embedding, often fails. We are currently developing a variant of the circulant embedding method for simulating smooth processes. First, we create a filtered version of the process which is less smooth than the actual process. Circulant embedding is then applied to simulate this filtered process. Given the filtered process, a much smaller boundary set of the original process is conditionally simulated. From this, the entire smooth process may be recovered with minimal computation. We have proposed a new method for the simulation of smooth process using a variant of circulant embedding process. The usefulness of this algorithm is predicated on the notion that solves with large block-Toeplitz, Toeplitz-block matrices, as is the case of simulating the filtered process, are much faster than a direct simulation of the Gaussian process. Application of this method for the ARM SGP surface flux data set will be presented.