Modified planar rotator model for efficient gap filling in spatial data

M. Žukovič[1], D.T. Hristopulos[2]

[1]Institute of Physics, Faculty of Science, P.J. Šafárik University in Košice, Slovakia

[2]Geostatistics Research Unit, Department of Mineral Resources Engineering, Technical University of Crete, Greece

An ever increasing amount of data collected by various remote sensing technologies as well as the need for their efficient (such as near real time) processing calls for development of new techniques. For example, geostatistical methods [1], that are traditionally used to predict missing values in spatial data, suffer from computational inefficiency, restriction to Gaussian data, as well as various subjective choices in variogram modeling [2]. Therefore, they are impractical if not useless in case of massive data sets. Novel techniques are often developed based on an interdisciplinary approach, which was also behind the idea of alleviating the above limitations via modeling spatial correlations by means of interaction-based Gibbs random fields [3] and spin [4] models from statistical physics.

In the same spirit, in the present study we introduce a spatial prediction method inspired from statistical physics for the efficient estimation of missing data on partially sampled Cartesian grids. The prediction model is based on a classical planar rotator (or XY) spin model, which is modified in order to display relevant short-range correlations and to allow an appropriate one-to-one mapping between the data and spin values. Spatial correlations present in the data are captured in terms of nearest-neighbor interactions between the spin variables. The only parameter of the model is the thermodynamic temperature, which is estimated from the sample-based nearest-neighbor correlations (which are included in the energy function).

Conditional Monte Carlo simulations honoring the sample values are performed at the inferred temperature on the entire grid to bring the system into thermal equilibrium and subsequently collect prediction statistics. Since the model does not show undesirable critical slowing down, the relaxation process is rather fast.In addition, the short-range nature of the interactions allows vectorization of the algorithm. Consequently, the proposed method achieves roughly linear scaling with the system size. This scaling implies that it is significantly more efficient than the conventional geostatistical approaches and also applicable to huge spatial data sets, such as satellite and radar images. We also discuss the potential of its implementation on GPU processors.

[l] P.J. Diggle, P.J. Ribeiro, Model-based Geostatistics. Series: Springer series in statistics, (New York, 2007).

[2] H. Wackernagel, Multivariate Geostatistics (Springer, New York, 2003).

[3] D.T. Hristopulos, SIAM J. Scient. Comput. **24**, 2125 (2003).

[4] M. Žukovič, D.T. Hristopulos, Phys. Rev. E **80**, 011116 (2009).