Fast Maximum Likelihood for Random Fields

What is a Random Field?
You might often find yourselves in the same position as us: You have experimental data, which in some way, you would like to fuse with available scientific information. In many cases, we find we can use Bayesian Data Assimilation (DA) for this fusion. In Bayesian DA, the measured quantity of interest (QoI) is described by a random field: We do not exactly know what the QoI interest is, but by fusing experimental data and scientific information we can make a pretty good prediction.

Problem
To make a DA prediction of a random field, we usually have to estimate the smoothness of the field. The conventional approach is to make a Maximum Likelihood Estimate, however this can be very expensive for a large number N of data: the computational cost of a MLE scales as N².

Solution
For large data sets we replace the MLE with Frequency-domain MLE (FMLE) or Frequency-domain Sample Variogram (FSV), which use a NUFFT and scale as Βlog N.

Example 1: Oxygen data
Let us have a look at an example of a random field. Consider the Atlantic transect made by the research vessel "Pelagia".

At each station, the crew made a vertical profile of the oxygen content, using a small number of accurate bottle titrations and a large number of less accurate sensor readings. So we now have a low-fidelity and a high-fidelity set of data, along with the scientific information that they are observations of the same QoI and should be more or less correlated.

In the above plot we have fused all this information. The use of FMLE to estimate the smoothness of the random field reduces CPU time from 10⁷ to 10⁻¹ seconds.

Example 2: Terrain Elevation
In a second example we consider the measurement of terrain elevation by SRTM radar data acquired during a Space Shuttle mission and stereo matching of RecceLite aerial photographs acquired during a F-16 mission.

Again, we use Bayesian DA to fuse the information and predict a (normalized) contour plot of the terrain elevation.

We find that using FMLE reduces the CPU time significantly, in fact conventional MLE for the full data set is impractical.

Selection of current projects
FSV proved useful for the regression of the Particle Imagery Velocimetry (PIV) data of a windtunnel test of the DelFly II.

Random Fields can be used as surrogate models for Uncertainty Quantification and also for efficient Optimization. Bayesian DA is extremely suitable to incorporate adjoint-based gradient information, as in the shape optimization of a two-stage high-pressure turbine, a Msc project in cooperation with Rolls-Royce Berlin.

Random Fields enable mass conserving interpolation of flow data. We are investigating possibilities of introducing even more scientific knowledge.

CONTACT INFORMATION
Jouke de Baar, Room NB 2.38
j.h.s.debaar@tudelft.nl
+31 (6) 42703414

Publications
- Jouke H. S. de Baar, Mustafa Percol, Richard P. Dwight, Bas W. van Oudheusden, Hester Bijl, "Kriging regression of PIV data using a local error estimate," submitted 2012
- J. de Baar, R. Dwight, and H. Bijl, "Speeding up Kriging through fast estimation of the hyperparameters in the frequency-domain," Computers & Geosciences, Volume 54, April 2013, Pages 99-106
- J. de Baar, R. Dwight, and H. Bijl, "Fast maximum likelihood estimate of the Kriging correlation range in the frequency domain," IAMG Conference Salzburg 2011