

Non-Stationary Covariance Models White Paper --- Report for the Analysis Brain-Storming Group

Jony Pinto Jr - UFPR, Brazil
Huijing Jiang -- IBM, USA
Carlos Tadeu – UFRJ, Brazil
Geir-Arne Fuglstad – NTNU, Norway
Wendy Meiring – UCSB, USA (**reporter**)
Unn Dahlen- Lund, Sweden
Raquel Vaz – UFRJ, Brazil
Estelina Capistrano- UFRJ, Brazil
Rui Hu – University of Alberta, Canada

The topics discussed by our group include substantial overlap with other groups.

Key Discussion Points included:

- 1) In comparisons, we need to simulate from stochastic processes with carefully chosen settings to **study the strengths and limitations of different models.**

Simulation settings should include different:

- Sample sizes – to evaluate the scalability of methods.
- Signal to Noise Ratios
- Smoothnesses of the random field
- Levels of non-stationarity (including regions with strong levels of non-stationarity in the covariance). Possibly also simulations with abrupt changes in covariances – such as due to rocks affecting underground pollutant contaminated water flow.
- Designs of monitoring networks – including regularly and irregularly spaced locations (grid and non-grid) since certain methods may perform better under one of these types of monitoring network designs.
- In the irregular location simulations, there should be a large variety of intersite distances, including sites relatively close together to aid estimation of the nugget (which may be constant or spatially varying).

- If possible, simulate using spatially varying parameter values (such as spatially varying nugget, range and smoothness) in a carefully designed manner.

2) **Evaluate both over-fit and under-fit tendencies under multiple settings:**

There is a need for simulations both from covariance-stationary fields (possibly isotropic) and from non-stationary covariance structures to evaluate the benefit of non-stationarity covariance models and also to assess if methods underfit/over-smooth in certain situations.

- 3) Simulations could be run both using sparse precision matrices as the truth, and also using sparse covariance structures as the truth.
- 4) Analysis should include evaluation of the ability to estimate covariances between pairs of locations where one does not have data, as well as covariances between pairs of sites where data is only available at one member in the pair.

Coverage of confidence intervals or credible intervals could be compared across methods via simulating multiple realizations where one knows the truth.

- 5) The simulation studies need to assess the **impact of including non-stationarity in the covariance models on the resulting uncertainty estimates** for the spatial field and on the **uncertainty measures** of the pairwise covariance estimates mentioned above.
- 6) **Simulations also could be included to study the improvement by including covariates in the covariance structure.**
- 7) **Our group also raised points that were earlier raised in the metrics discussion – related to how effectively to compare different spatial covariance models both in the frequentist and in the Bayesian framework.** Continuing discussion on metrics will be valuable going forward. We refer to the earlier Metrics Discussion Group Report.

8) A member of our group, Huijing Jiang, mentioned that her research group has a paper recently published, that demonstrated improvements when using a non-stationary covariance model in an engineering setting. She may be able to provide the data to us to add to our collection of real data sets.