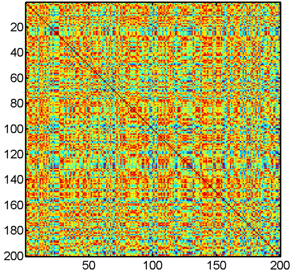
Stochastic Local Interaction Models for Processing Spatiotemporal Datasets

Dionissios T. Hristopulos* Vasiliki D. Agou* Andreas Pavlides* Panagiota Gkafa*

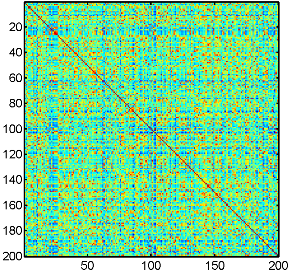
*Geostatistics Laboratory, School of Mineral Resources Engineering
Technical University of Crete, Chania 73100, Greece

Presented at *the EGU 2020 General Assembly* (Virtually in Vienna)

Motivation: Geostatistics and Gaussian Process Regression face challenges for $N \gg 1$



Dense Covariance matrix



Dense Precision matrix

Problem: Inversion of large (for $N \gg 1$), dense covariance matrices
Problem: Covariance functions are mathematical constructions typically lacking physical motivation

Motivation: Geostatistics and Gaussian Process Regression face challenges for $N \gg 1$

- Best Linear Unbiased Estimator (BLUE) (a.k.a. Ordinary Kriging)

$$\hat{X}(\mathbf{s}) = \sum_{j=1}^{m} \lambda_j X(\mathbf{s}_j), \quad \mathbf{s}_j \in B(\mathbf{s}; r_c), \; B(\cdot;\cdot) : \text{search neighborhood}$$

- *Best*: minimum mean square estimation error

$$\{\lambda_1, \lambda_2, \ldots, \lambda_m\} = \arg \min_{\lambda_1,\lambda_2,\ldots,\lambda_m} \left(\mathbb{E}\left\{ \hat{X}(\mathbf{s}) - X(\mathbf{s}) \right\}^2 \Big| \; \lambda_1 + \lambda_2 + \ldots \lambda_m = 1 \right)$$

- Spatial weights follow from the linear system: $\mathbf{C}_{N,N}\, \boldsymbol{\Lambda} = \mathbf{C}_{N,0}$

$$\begin{bmatrix} c(\mathbf{s}_1-\mathbf{s}_1) & c(\mathbf{s}_1-\mathbf{s}_2) & \ldots & c(\mathbf{s}_1-\mathbf{s}_m) & 1 \\ c(\mathbf{s}_2-\mathbf{s}_1) & c(\mathbf{s}_2-\mathbf{s}_2) & \ldots & c(\mathbf{s}_2-\mathbf{s}_m) & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c(\mathbf{s}_m-\mathbf{s}_1) & c(\mathbf{s}_m-\mathbf{s}_2) & \vdots & c(\mathbf{s}_m-\mathbf{s}_m) & 1 \\ 1 & 1 & \ldots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \\ \mu \end{bmatrix} = \begin{bmatrix} c(\mathbf{s}_1-\mathbf{s}) \\ c(\mathbf{s}_2-\mathbf{s}) \\ \vdots \\ c(\mathbf{s}_m-\mathbf{s}) \\ 1 \end{bmatrix}$$

Introduction

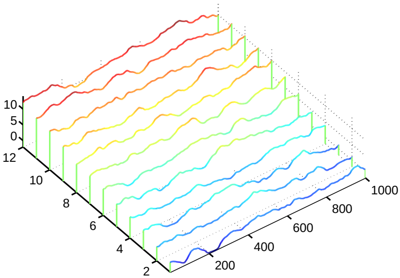
- Ideas inspired from Statistical Physics can help to model efficiently spatiotemporal data
- There is a long history of interplay between statistical physics and statistical data analysis (*Ising model, Markov Chain Monte Carlo, Metropolis updating, variational Gaussian approximations in machine learning, etc.*)
- The main themes for this presentation are:

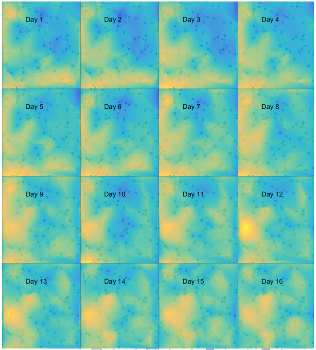
  (i) Construction of *stochastic local interaction* computationally efficient models for space-time data analysis

  (ii) Application to natural resources and environmental data

Representations of Space-Time Random Fields

- In simple terms, an STRF $X(\mathbf{s}, t; \omega)$ is *an ensemble* (collection) of functions $x(\mathbf{s}, t)$ where $\mathbf{s} \in \mathbb{R}^d$ and $t \in T \subset [0, \infty)$.
- The index ω distinguishes different states (functions). In practice it is often suppressed.





Representations of Space-Time Random Fields

- The classical case: Joint probability density function of Gaussian data in terms of the covariance matrix **C**

$$f_{\mathrm{X}} \propto \exp\left(-\frac{1}{2}\sum_{i,j=1}^{N} x_i\, C^{-1}_{x;\,i,j}\, x_j\right) \; ; \quad C_{x;\,i,j} : \text{ Covariance matrix}$$

- The Boltzmann-Gibbs representation: Joint probability density function based on an "energy function" with controlled space-time structure

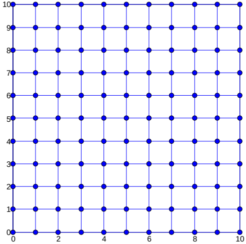
$$f[x(\mathbf{s})] = Z(\boldsymbol{\theta})^{-1}\, \mathrm{e}^{-\mathcal{H}[x(\mathbf{s};\boldsymbol{\theta})]}, \; \mathcal{H}[x(\mathbf{s})] : \text{ energy functional,}$$

- $Z(\boldsymbol{\theta})$: partition function is a normalizing factor

Local Interactions on Rectangular Grids

Grid structure: SSRFs ⇒ Gauss-Markov random fields

$$\mathcal{H}[x(\mathbf{s})] = \lambda \sum_{n=1}^{N} \left\{ [x(\mathbf{s}_n) - m_x]^2 + c_1 \sum_{i=1}^{d} \left[\frac{x(\mathbf{s}_n + a_i\,\hat{\mathbf{e}}_i) - x(\mathbf{s}_n)}{a_i} \right]^2 + c_2 \sum_{i=1}^{d} \left[\frac{x(\mathbf{s}_n + a_i\,\hat{\mathbf{e}}_i) - 2x(\mathbf{s}_n) + x(\mathbf{s}_n - a_i\,\hat{\mathbf{e}}_i)}{a_i^2} \right]^2 \right\}$$



$\hat{\mathbf{e}}_i,\ i = 1, \ldots, d$: unit vectors in lattice directions

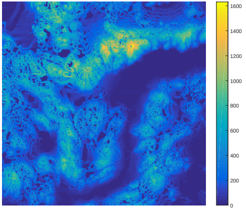
a_i : lattice steps

$$\mathcal{H}_{fgc}[X(\mathbf{s})] = \lambda S_0 + c_1\, \lambda\, S_G + c_2\, \lambda\, S_c$$

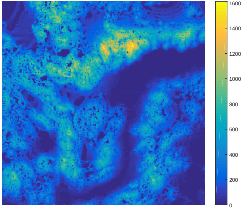
Rue and Held, *Gaussian Markov Random Fields: Theory and Applications*, Chapman and Hall/CRC, 2005

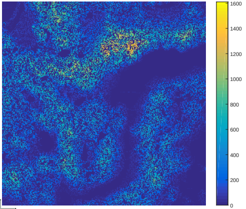
Reconstruction of Walker Lake Data (260 × 300 grid)

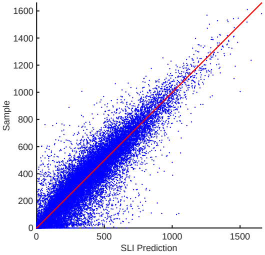
Anisotropic model based on 39 000 sampling points (50%)

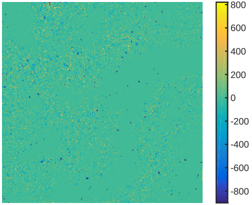


Left top: Full data set
Left bottom: Sample
Right top: SLI-based map
Right bottom: SLI spatial error



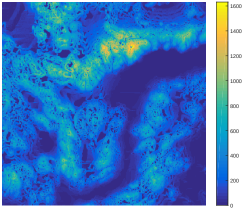


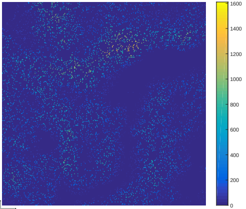




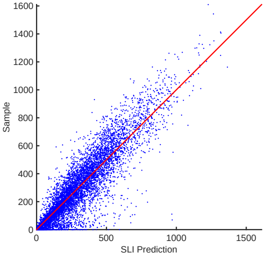
Reconstruction of Walker Lake Data (260 × 300 grid)

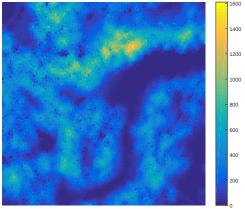
Anisotropic SLI model – 7 800 sampling points (10%)

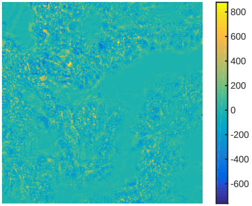




Left top: Full data set
Left bottom: Sample
Right top: SLI-based map
Right bottom: SLI spatial error







SLI Model for space-time data

- The vector $\mathbf{x}_{\mathbb{S}} \equiv (x_1, x_2, \ldots, x_N)^\top$ comprises the *observations* at the space-time point set $\mathbb{S}_N = \{(\mathbf{s}_1, t_1), (\mathbf{s}_2, t_2), \ldots (\mathbf{s}_N, t_N)\}$
- Energy function for *stochastic local interaction (SLI) model*

$$\mathcal{H}(\mathbf{x}_{\mathbb{S}}; \boldsymbol{\theta}) = \frac{1}{2\,\lambda} \left[\frac{{\mathbf{x}'_{\mathbb{S}}}^\top \mathbf{x}'_{\mathbb{S}}}{N} + c_1\, \mathcal{S}_1(\mathbf{x}'_{\mathbb{S}}; \mathbf{h}_s, \mathbf{h}_t) \right], \quad \mathbf{x}'_{\mathbb{S}} = \mathbf{x}_{\mathbb{S}} - \mathbf{m}_{\mathrm{x}}$$

- Local interactions are implemented by means of simple expressions that involve kernel-weighted averages [Hristopulos, 2015, Hristopulos & Agou, 2019]
- A crucial idea is that the kernel bandwidth is automatically adjusted to account for the local sampling density

SLI Model for space-time data

$$\mathcal{H}(\mathbf{x}_\mathrm{S};\boldsymbol{\theta}) = \frac{1}{2\,\lambda}\left[\frac{{\mathbf{x}'_\mathrm{S}}^\top \mathbf{x}'_\mathrm{S}}{N} + c_1\,\mathcal{S}_1(\mathbf{x}'_\mathrm{S};\mathbf{h}_s,\mathbf{h}_t)\right], \quad \mathbf{x}'_\mathrm{S} = \mathbf{x}_\mathrm{S} - \mathbf{m}_\mathrm{x}$$

SLI model parameters

- λ: Overall scale parameter
- c_1: Rigidity coefficient
- μ_s, K_s: Spatial kernel bandwidth: $h_{s,n} = \mu_s\, D_{n,[K_s]}(\mathbb{S}_N)$
- $D_{n,[K_s]}(\mathbb{S}_N)$: distance between $\mathbf{s}_n$ and K_s-nearest neighbor in $\mathbb{S}_N$
- μ_t, K_t: Temporal kernel bandwidth: $h_{t,n} = \mu_t\,(K_t - 1)\delta t$
- $\{\beta_k\}_{k=1}^{K}$: Coefficients of trend model $m_\mathrm{x}(\mathbf{s},t) = \sum_{k=1}^{K}\beta_k f_k(\mathbf{s},t)$

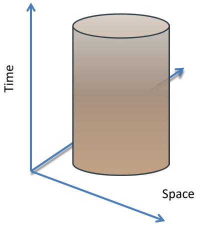
Different approaches for kernel weight selection

Separable space-time structure

$$w_{n,k} = K\left(\frac{\|\mathbf{r}_{n,k}\|}{h_{s,n}}\right) K\left(\frac{|\tau_{n,k}|}{h_{t,n}}\right)$$

This does not imply *separable covariance*

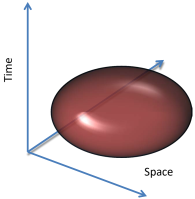
Currently we use this approach



Composite space-time metric

$$w_{n,k} = K\left(\frac{\sqrt{\mathbf{r}^2_{n,k} + \tau^2_{n,k}/\alpha^2}}{h_{s,n}}\right)$$

Relation of S-T bandwidths: $h_{t,n} = \alpha h_{s,n}$



SLI precision matrix formulation

Quadratic Form of SLI Energy: Sparse by Construction

$$\mathcal{H}(\mathbf{x}'_{\mathrm{S}};\boldsymbol{\theta}) = \frac{1}{2}\mathbf{x}'^{\top}_{\mathrm{S}}\,\mathbf{J}(\boldsymbol{\theta}')\,\mathbf{x}'_{\mathrm{S}}, \quad \boldsymbol{\theta} = \boldsymbol{\theta}' \cup \{\beta_k\}_{k=1}^{K}$$

Precision matrix (sparse, explicit) using only the first two terms

$$\mathbf{J}(\boldsymbol{\theta}') = \frac{1}{\lambda}\left\{\frac{\mathbf{I}}{N} + c_1\,\mathbf{J}_1(\boldsymbol{\theta}'')\right\}, \; \boldsymbol{\theta}'' = (\mu_s, \mu_t, K_t, K_s)^{\top}$$

Gradient Precision sub-matrix based on the normalized-weight matrix

$$[\mathbf{J}_1(\mathbf{h}_s,\mathbf{h}_t)]_{i,j} = -u_{i,j}(\mathbf{h}_s,\mathbf{h}_t) - u_{j,i}(\mathbf{h}_s,\mathbf{h}_t) + \delta_{i,j}\sum_{k=1}^{N}\left[u_{i,k}(\mathbf{h}_s,\mathbf{h}_t) + u_{k,i}(\mathbf{h}_s,\mathbf{h}_t)\right]$$

Normalized kernel weights: $u_{i,j}(\mathbf{h}_s,\mathbf{h}_t) = [\mathbf{U}]_{i,j}$

SLI Mode Predictor for Interpolation of Missing Data

- Modified energy functional in the presence of an unknown value

$$\mathcal{H}(\mathbf{x}_{\mathrm{S}};\boldsymbol{\theta}) = \frac{1}{2}\begin{bmatrix} \mathbf{x}'^{\top}_{\mathrm{S}} & x'_p \end{bmatrix}\begin{bmatrix} \mathbf{J}_{S,S} & \mathbf{J}_{S,P} \\ \mathbf{J}_{P,S} & J_{p,p} \end{bmatrix}\begin{bmatrix} \mathbf{x}'_{\mathrm{S}} \\ x'_p \end{bmatrix}$$

- x_p is the value at the *prediction point*
- $\hat{x}_p = \arg\min_{x_p} \hat{\mathcal{H}}(\mathbf{x}_{\mathrm{S}}, x_p; \boldsymbol{\theta}^*)$.
- The predictor is a linear equation (reminiscent of GMRFs) that can be evaluated with $\mathcal{O}(N)$ computational time complexity:

Conditional mean | Conditional variance

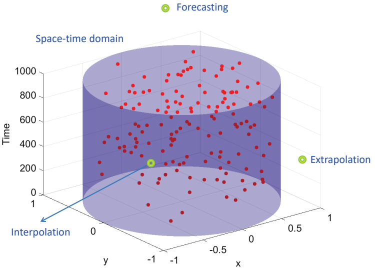
$$\hat{x}_p = m_{\mathrm{x}} - \frac{1}{J_{p,p}(\boldsymbol{\theta}'^*)}\sum_{i=1}^{N} J_{p,i}(\boldsymbol{\theta}'^*)\,(x_i - m_{\mathrm{x}}), \quad \sigma_p^2 = \frac{1}{J_{p,p}(\boldsymbol{\theta}'^*)}.$$

 [Hristopulos, 2015]

Space-Time data and parameter estimation

- Synthetic S-T data generated by the R package RandomFields (Schlather M, *et al.* (2019). RandomFields: Simulation and Analysis of Random Fields. R package version 3.3.6)
- ERA5 reanalysis data (ERA5 is the latest climate reanalysis produced by the European Centre for Medium-Range Weather Forecasts, providing hourly data on many atmospheric, land-surface and sea-state parameters together with estimates of uncertainty).

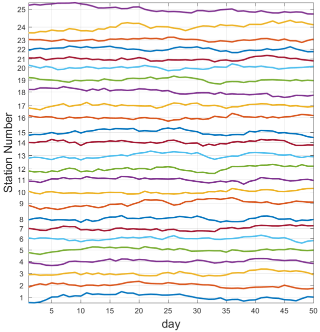
- Parameter vector: $\boldsymbol{\theta} = (\mu_s, K_s, \mu_t, K_t, \lambda, c_1, \beta_1, \ldots, \beta_K)^\top$
- K_s and K_t are pre-selected
- *Maximum Likelihood Estimation:* The log-determinant of the sparse precision matrix is calculated using LU decomposition

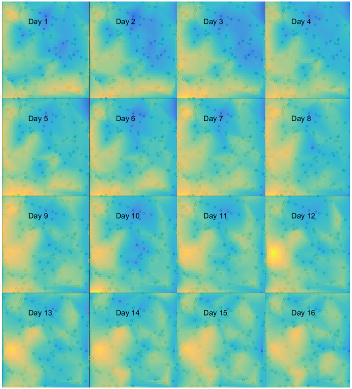


For more details see [Hristopulos & Agou, 2019].

Simulated S-T data from stationary random field

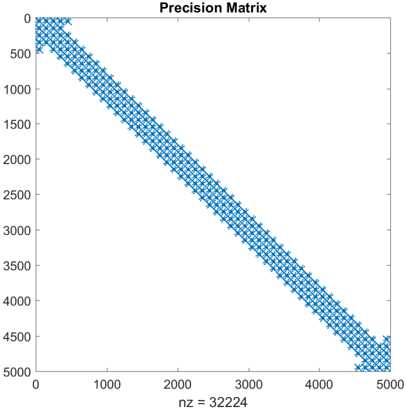
100 random locations and 50 time slices from a Gaussian RF with separable S-T exponential covariance

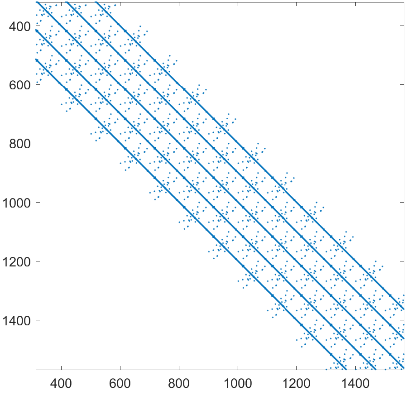




Simulated S-T data from stationary random field

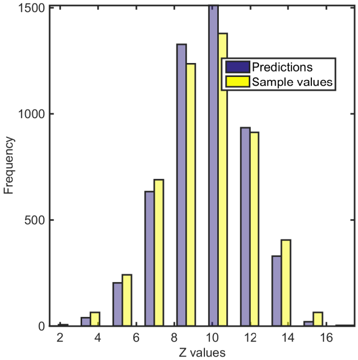
SLI Precision Matrix – Sparsity index ≈ 0.13%

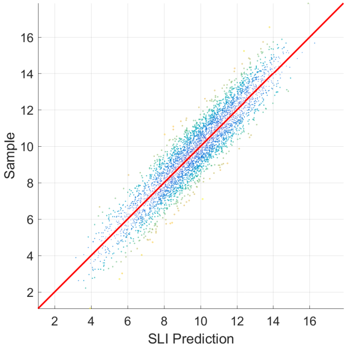




Simulated S-T data from stationary random field

Leave-one-time-slice-out Cross validation



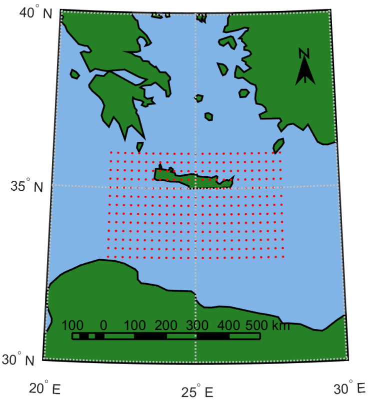


Reanalysis ERA5 Temperature data around Crete

Study design

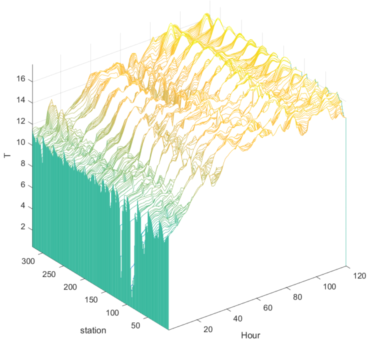
- Hourly temperature (degrees) for five days from 01-01-2017 00:00:00 to 05-01-2017 23:00:00
- Total of 39 000 data points
- SLI parameter estimation using MLE (quadratic kernel, ($K_s = K_t = 3$)
- Cross validation: Remove one time slice each time and predict using SLI

Spatial grid 13×25

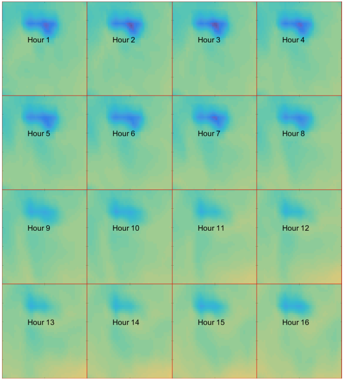


Reanalysis ERA5 Temperature data around Crete

Time series - all sites

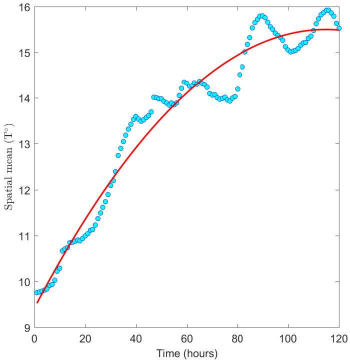


First few time slices



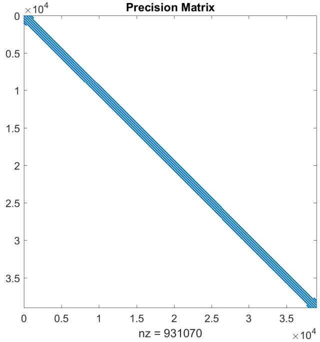
Reanalysis ERA5 Temperature data around Crete

Trend Modeling

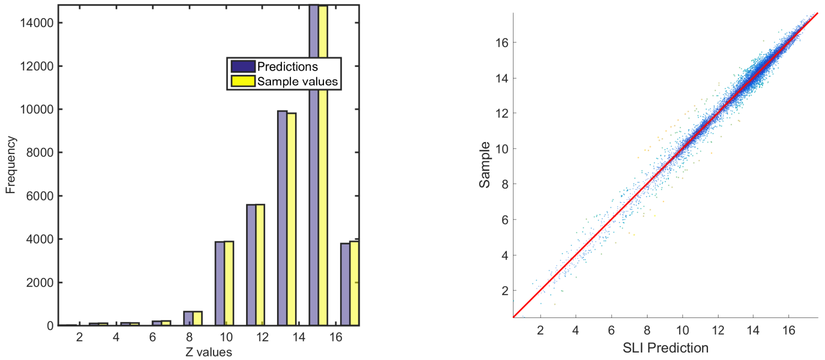


$$m_x(t) = \beta_1 + \beta_2 t + \beta_3 t^2$$

Sparsity of SLI Precision Matrix



Reanalysis ERA5 Temperature data around Crete



SLI precision matrix *sparsity index* $\approx$ 0.06%, i.e., $\approx$ 123 000 non-zero entries out of 1.521×10^9 matrix entries

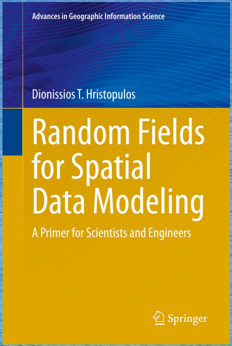
Conclusions and Future Directions

- Statistical mechanics is a useful theoretical framework for developing local methods of S-T data analysis
- The stochastic local interaction model (SLI) extends GMRF ideas to scattered data without requiring the solution of an SPDE
- SLI employs compactly supported kernel functions with adaptive kernel bandwidths to construct an explicit precision matrix. This structure leads to semi-explicit prediction equations and *computationally efficient implementations*
- SLI prediction performance is comparable to Kriging (BLUE)

Acknowledgments

Parts of this research were funded by the project *DES2iRES* ERANETMED NEXUS-14-049 and the “ARISTEIA” Action (co-funded by the European Social Fund and National Resources) under project *SPARTA*.

TUC team: Dr. E. Varouchakis, Dr. S. Elogne, Dr. M. Žukovič, Dr. A. Chorti, Dr. Ivi Tsantili, Dr. A. Pavlides, Mrs. V. Agou, MSc, Mrs. P. Gkafa, and Mr. M. Petrakis, MSc



Random Fields for Spatial Data Modeling

For more information ...

D. T. Hristopulos (2003). Spartan Gibbs Random Field Models for Geostatistical Applications. *SIAM J. Sci. Comput.*, 24(6), 2125-2162.

D. T. Hristopulos and S. Elogne (2007). Analytic properties and covariance functions of a new class of generalized Gibbs random fields. *IEEE Trans. Infor. Theory*, 53(12), 4667 - 4679.

D. T. Hristopulos and S. N. Elogne (2009). Computationally efficient spatial interpolators based on Spartan spatial random fields. *IEEE Trans. Signal Proc.*, 57(9), 3475–3487.

D. T. Hristopulos (2015). Covariance functions motivated by spatial random field models with local interactions. *Stoch. Envir. Res. Risk A.*, 29(3), 739–754.

D. T. Hristopulos (2015). Stochastic local interaction (SLI) model for incomplete data in $d-$dimensional metric spaces. *Computers & Geosciences*, 85, 26–37.

D. T. Hristopulos and I. Tsantili (2017). Space-time covariance functions based on linear response theory and the turning bands method. *Spatial Statistics*, 22(2):321–337.

D. T. Hristopulos and V. Agou (2019). Stochastic local interaction model for space-time data. *Spatial Statistics*, 2019, Art. no. 100403.

E. A. Varouchakis and D. T. Hristopulos (2019). Comparison of spatiotemporal variogram functions based on a sparse dataset of groundwater level variations. *Spatial Statistics*, 34, December, Art. no. 100245.