

Spatiotemporal Models Inspired from Statistical Physics

Dionisis Hristopulos^{1,2}

¹Geostatistics Laboratory, School of Mineral Resources Engineering
Technical University of Crete, Chania, Greece

²Telecommunication Systems Institute, Chania, Greece

Modèles Spatio-temporels en Météorologie et Océanographie
November 27-30, 2018, INRIA, Campus de Beaulieu, Rennes, France

Joint work with Vasiliki Agou, Ivi Tsantili, Emmanouil Varouchakis



HELLENIC REPUBLIC
MINISTRY OF
ECONOMY & DEVELOPMENT
SPECIAL SECRETARY FOR EDF & CEF
MANAGING AUTHORITY OF EPAnEK

EPAnEK 2014-2020
OPERATIONAL PROGRAMME
COMPETITIVENESS
ENTREPRENEURSHIP
INNOVATION

ΕΣΠΑ
2014-2020
εποχή - πράξη - καλύτερη
Partnership Agreement
2014 - 2020



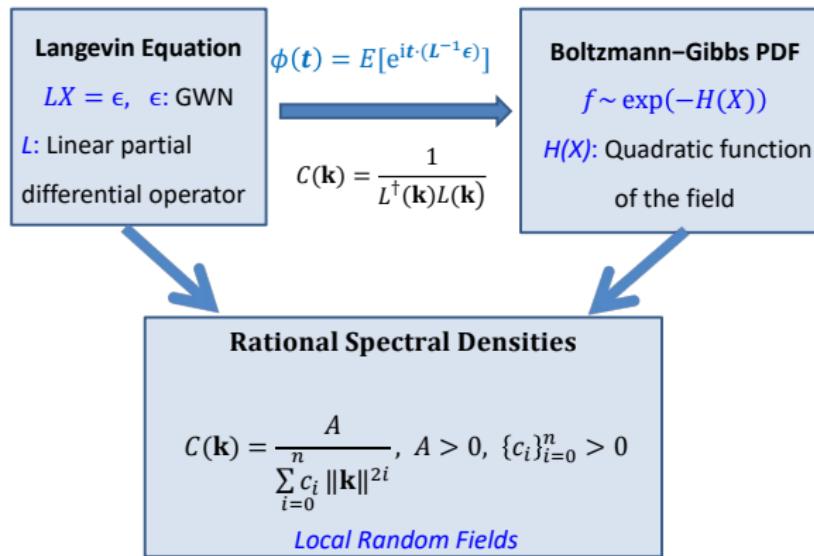
Overview of the Presentation

- ▶ Introduction
- ▶ Boltzmann-Gibbs representations of random fields
- ▶ Spartan composite-metric covariance functions
- ▶ Space-time covariances from linear response theory
- ▶ Space-time stochastic local interaction model (SLI)
- ▶ Examples of SLI data application (synthetic, reanalysis)
- ▶ Conclusions

Introduction

- ▶ Statistical physics uses **equations of motion** (leading to partial differential equations) and Boltzmann-Gibbs (exponential) probability density functions
- ▶ There is a long history of interplay between statistical physics and statistical data analysis (Ising model, Metropolis updating, Markov Chain Monte Carlo, variational Gaussian approximations in machine learning, etc.)
- ▶ For the purpose of this talk, the **main themes** are how to use:
 - (i) Boltzmann-Gibbs formulations with *local* energies to derive computationally efficient models
 - (ii) the framework of *equations of motion* to obtain new forms of spatiotemporal dependence

Local Random Fields: SPDEs & Boltzmann-Gibbs PDFs



Whittle-Matérn case: L is a pseudo-differential operator. The spectral density is a rational function only if $\nu + \frac{d}{2} = n \in \mathbb{N}$

Spartan Spatial Random Fields (SSRFs) are Boltzmann-Gibbs RFs with “local” structure

- ▶ Boltzmann-Gibbs probability density function (PDF)

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

- ▶ Z : partition function $Z = \int \mathcal{D}x(\mathbf{s}) e^{-\mathcal{H}[x(\mathbf{s})]}$

- ▶ $\mathcal{H}_{\text{fgc}}[x(\mathbf{s})] = \int \frac{d\mathbf{s}}{2\eta_0\xi^d} \left\{ [x(\mathbf{s})]^2 + \eta_1 \xi^2 [\nabla x(\mathbf{s})]^2 + \xi^4 [\nabla^2 x(\mathbf{s})]^2 \right\}$

How to extend the above continuum representation to discrete data?

- ▶ Regular grids: Use finite differences \Rightarrow GMRFs
- ▶ Scattered data: Use kernel functions \Rightarrow SLI model

D. T. Hristopulos, 2003. *SIAM J. Sci. Comput.*, 24(6), 2125–2162

Spartan Spatial Random Fields (SSRFs) are Boltzmann-Gibbs RFs with “local” structure

Langevin Equation

$\mathcal{L} X(\mathbf{s}) = \varepsilon(\mathbf{s})$, where $\varepsilon(\cdot)$ is GWN (Gaussian white noise)

$$\mathcal{L} = \frac{1}{\sqrt{\eta_0 \xi^d}} \left[1 - \left(\sqrt{2 + \eta_1} \right) \xi (-\Delta)^{1/2} + \xi^2 \Delta \right], \quad \Delta : \text{Laplacian}$$

Spectral density

$$\tilde{C}(\mathbf{k}) = \frac{1}{|\tilde{\mathcal{L}}(\mathbf{k})|^2} = \frac{\eta_0 \xi^d}{1 + \eta_1 (\|\mathbf{k}\| \xi)^2 + (\|\mathbf{k}\| \xi)^4}.$$

Covariance function PDE

$$\left[1 - \eta_1 \xi^2 \Delta_{\mathbf{s}} + \xi^4 \Delta_{\mathbf{s}}^2 \right] C(\mathbf{s} - \mathbf{s}') = -\eta_0 \xi^d \delta(\mathbf{s} - \mathbf{s}')$$

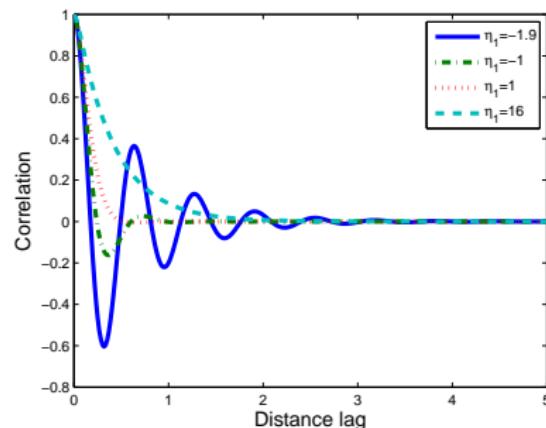
SSRF covariance for time series

$$C(h) = \frac{\eta_0}{4} e^{-h\beta_2} \left[\frac{\cos(h\beta_1)}{\beta_2} + \frac{\sin(h\beta_1)}{\beta_1} \right], \quad |\eta_1| < 2$$

$$C(h) = \eta_0 \frac{(1+h)}{4e^h}, \quad \eta_1 = 2$$

$$C(h) = \frac{\eta_0}{2\Delta} \left(\frac{e^{-h\omega_1}}{\omega_1} - \frac{e^{-h\omega_2}}{\omega_2} \right), \quad \eta_1 > 2$$

d=1



Notation

- ▶ $h = |r|/\xi$: normalized lag
- ▶ $\beta_{1,2} = \left(\frac{|2 \mp \eta_1|}{4} \right)^{1/2}$
- ▶ $\omega_{1,2} = \left(\frac{|\eta_1 \mp \Delta|}{2} \right)^{1/2}$
- ▶ $\Delta = |\eta_1|^2 - 4|^{\frac{1}{2}}$

Hristopulos & Elogne (2007), *IEEE Trans. Inform. Theory*, 53(12), 4667–4679

Damped harmonic oscillator in a heat bath is a one-dimensional SSRF

Langevin equation: $\ddot{x}(t) + \Gamma \dot{x}(t) + \omega_0^2 x(t) = \epsilon(t)$

$$C(h) = \frac{\eta_0}{4} e^{-h\beta_2} \left[\frac{\cos(h\beta_1)}{\beta_2} + \frac{\sin(h\beta_1)}{\beta_1} \right], \quad |\eta_1| < 2,$$

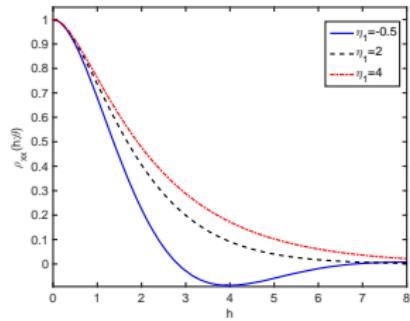
$$C(h) = \eta_0 \frac{(1+h)}{4 e^h}, \quad \eta_1 = 2,$$

$$C(h) = \frac{\eta_0}{2\Delta} \left(\frac{e^{-h\omega_1}}{\omega_1} - \frac{e^{-h\omega_2}}{\omega_2} \right), \quad \eta_1 > 2$$

■ Underdamping

■ Critical damping

■ Overdamping



$$\Gamma^2 = \frac{\eta_1 + 2}{\xi^2}, \quad 2k_B T = \frac{\eta_0 m}{\xi^2 \sqrt{e + 2}}$$

⇐ Oscillator displacement correlation function

Nørrelykke, S.F., Flyvbjerg, H. *Phys. Rev. E* **83**, 041103 (2011)

SSRF covariance functions for planar processes

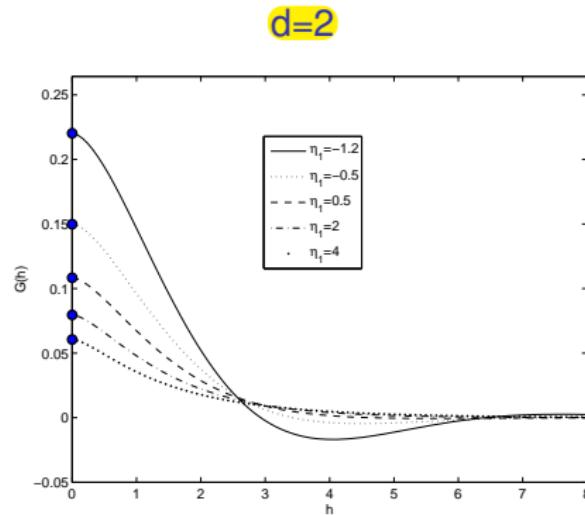
$$C(h) = \frac{\eta_0 \Im [K_0(h z_+)]}{\pi \sqrt{4 - \eta_1^2}}, \quad |\eta_1| < 2$$

$$C(h) = \left(\frac{\eta_0 h}{4\pi} \right) K_{-1}(h), \quad \eta_1 = 2$$

$$C(h) = \frac{\eta_0 [K_0(h z_+) - K_0(h z_-)]}{2\pi \sqrt{\eta_1^2 - 4}}, \quad \eta_1 > 2$$

Notation

- ▶ \Im : Imaginary part
- ▶ $z_{\pm} = \sqrt{-t_{\pm}^*}$
- ▶ $t_{\pm}^* = (-\eta_1 \pm \sqrt{\eta_1^2 - 4})/2$
- ▶ $K_{\nu}(z)$: modified Bessel function of the second kind and order ν



Hristopulos (2015), *Stoch. Environ. Res. Risk Assessss.*, 29(3), 739–754.

SSRF Covariance functions in \mathbb{R}^3 or in $\mathbb{R}^2 \times T$

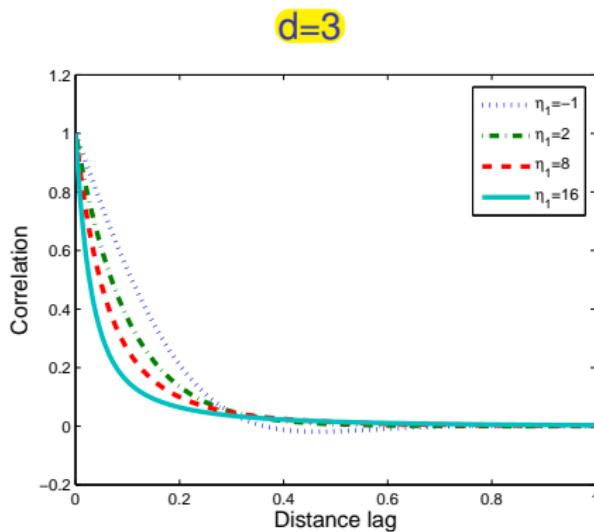
$$C(h) = \eta_0 \frac{e^{-h\beta_2}}{2\pi\Delta} \left[\frac{\sin(h\beta_1)}{h} \right], \quad |\eta_1| < 2$$

$$C(h) = \frac{\eta_0}{8\pi} e^{-h}, \quad \eta_1 = 2$$

$$C(h) = \frac{1}{4\pi\Delta} \left[\frac{e^{-h\omega_1} - e^{-h\omega_2}}{h(\omega_2 - \omega_1)} \right], \quad \eta_1 > 2$$

Notation

- ▶ $h = \|\mathbf{r}\|/\xi$,
- ▶ $\beta_{1,2} = \left(\frac{|2 \mp \eta_1|}{4} \right)^{1/2}$
- ▶ $\omega_{1,2} = \left(\frac{|\eta_1 \mp \Delta|}{2} \right)^{1/2}$
- ▶ $\Delta = |\eta_1|^2 - 4|^{\frac{1}{2}}$



Hristopulos and Elogne, 2007, *IEEE Trans. Inform. Theory*, 53(12), 4667–4679

SSRF composite space-time covariance in $\mathbb{R}^2 \times T$

$$C(h) = \eta_0 \frac{e^{-h\beta_2}}{2\pi\Delta} \left[\frac{\sin(h\beta_1)}{h} \right], \quad |\eta_1| < 2$$

$$C(h) = \frac{\eta_0}{8\pi} e^{-h}, \quad \eta_1 = 2$$

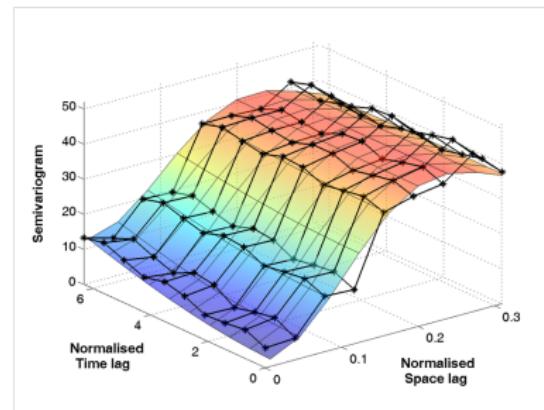
$$C(h) = \frac{1}{4\pi\Delta} \left[\frac{e^{-h\omega_1} - e^{-h\omega_2}}{h(\omega_2 - \omega_1)} \right], \quad \eta_1 > 2$$

Notation

- ▶ $h = \|\mathbf{r}\|/\xi$,
- ▶ $\beta_{1,2} = \left(\frac{|2 \mp \eta_1|}{4} \right)^{1/2}$
- ▶ $\omega_{1,2} = \left(\frac{|\eta_1 \mp \Delta|}{2} \right)^{1/2}$
- ▶ $\Delta = |\eta_1|^2 - 4|^{\frac{1}{2}}$

Composite space-time metric

$$h = \sqrt{\|\mathbf{r}\|^2/\xi^2 + \alpha\tau^2/\tau_c^2}$$



Varouchakis & Hristopulos, 2017. *Spatial Statistics*,
<https://doi.org/10.1016/j.spasta.2017.07.003>.

Space-Time models using Linear Response Theory

- ▶ Linear response theory describes the non-equilibrium response due to small deviation from equilibrium, leading to S-T Langevin equations

$$\frac{\partial x(\mathbf{s}, t)}{\partial t} = -D \left. \frac{\delta \mathcal{H}[x(\mathbf{s})]}{\delta x(\mathbf{s})} \right|_{x(\mathbf{s})=x(\mathbf{s},t)} + \zeta(\mathbf{s}, t) = V[x(\mathbf{s}, t)] + \zeta(\mathbf{s}, t)$$

- ▶ Equilibrium-restoring rate for the ST-SRRF random field

$$V[x(\mathbf{s}, t)] = -\frac{1}{2\xi^d \eta_0} \left(1 - \eta_1 \xi^2 \nabla^2 + \mu \xi^4 \nabla^4 \right) x(\mathbf{s}) \Big|_{x(\mathbf{s})=x(\mathbf{s},t)}$$

- ▶ D is a diffusion coefficient
- ▶ $\zeta(\mathbf{s}, t)$ is the random perturbation (e.g., Gaussian white noise)

Hohenberg and Halperin (1977), *Reviews of Modern Physics*, 49(3), 435–479



Equations of motion for ST-SSRF covariance can be derived using Linear Response Theory

$$\text{EOM: } \frac{\partial C(\mathbf{r}, \tau)}{\partial \tau} = -\frac{\text{sign}(\tau)}{\tau_c} (1 - \eta_1 \xi^2 \nabla^2 + \mu \xi^4 \nabla^4) C(\mathbf{r}, \tau),$$

where $\tau_c^{-1} = D/(2\xi^d \eta_0)$, and the *initial condition* is the SSRF spectral density

$$\tilde{C}(k, \tau = 0) = \frac{\eta_0 \xi^d}{1 + \eta_1(k\xi)^2 + \mu(k\xi)^4}.$$

Zero- μ solution in $d = 1$ dimension

$$C_1(h, u) = \frac{\eta_0 \lambda}{4} \left[e^{-\lambda h} \operatorname{erfc} \left(\sqrt{u} - \frac{\lambda h}{2\sqrt{u}} \right) + e^{\lambda h} \operatorname{erfc} \left(\sqrt{u} + \frac{\lambda h}{2\sqrt{u}} \right) \right],$$

$\lambda = 1/\sqrt{\eta_1}$, $h = r/\xi$, $u = |\tau|/\tau_c$, and $\operatorname{erfc}(\cdot)$ is the complementary error function.

$C_1(h, u)$ is equivalent (except for different parametrization) to the covariance derived from a parabolic SPDE in 1+1 dimensions (Heine (1955); Biometrika, 42(1-2), 170–178.)

Taming the variance divergence with space transforms

- In 3D the variance of the zero- μ solution diverges
- Space transforms are mathematical operations that can generate covariance functions based on lower-dimensional projections
[\[Mantoglou and Wilson \(1982\). Water Resources Research, 18\(5\), 1379–1394\]](#)

$$C_3(r, \tau) = \frac{1}{r} \int_0^r dx C_1(x, \tau) = \frac{1}{h} \int_0^h dy C_1(y, u)$$

- Applying the above to $C_1(h, u)$ we obtain

Space-transformed ST-SSRF covariance

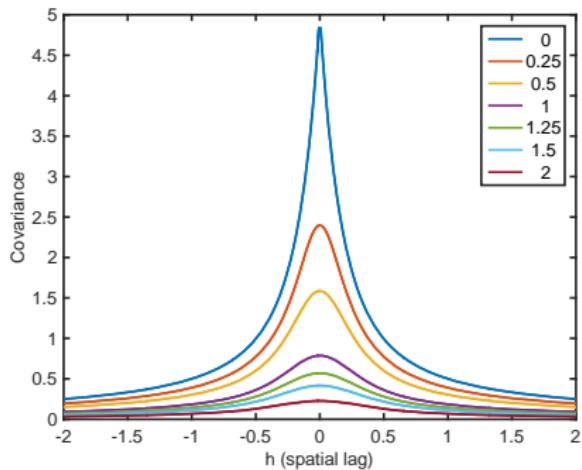
$$C_3(h, u) = \frac{\eta_0}{4h} \left[2e^{-u} \operatorname{erf}\left(\frac{\lambda h}{2\sqrt{u}}\right) + e^{\lambda h} \operatorname{erfc}\left(\sqrt{u} + \frac{\lambda h}{2\sqrt{u}}\right) - e^{-\lambda h} \operatorname{erfc}\left(\sqrt{u} - \frac{\lambda h}{2\sqrt{u}}\right) \right]$$

[Hristopulos & Tsantili, 2017. *Spatial Statistics*, 22(2):321–337]

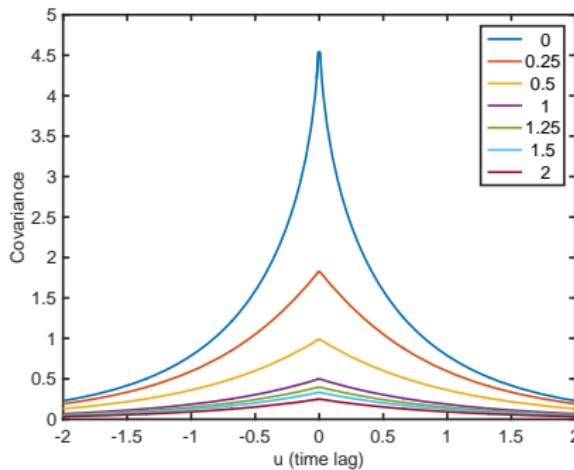


Visualization of the ST-SSRF covariance

The ST-SSRF covariance is fully symmetric, continuous, non-differentiable at $u = 0$, and *is not* in the Gneiting class (smoothness at origin does not determine smoothness along both the space and time axes)



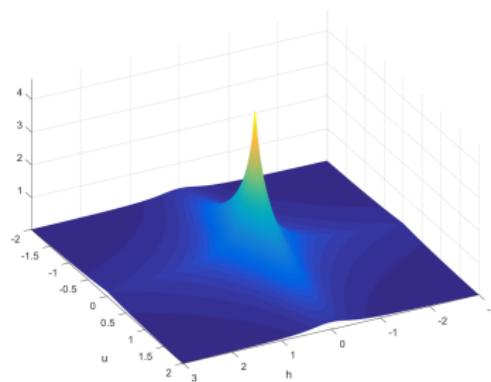
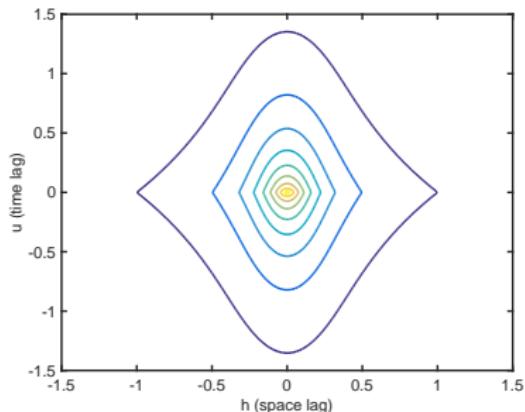
Dependence of covariance function on space lag for different temporal lags.



Dependence of covariance function on time lag for different spatial lags.

Visualization of the ST-SSRF covariance

The ST-SSRF covariance is fully symmetric, continuous, non-differentiable at $u = 0$, and *is not* in the Gneiting class (smoothness at origin does not determine smoothness along both the space and time axes)



Stochastic Local Interaction (SLI) model for space-time data

- ▶ The vector $\mathbf{x}_S \equiv (x_1, x_2, \dots, x_N)^\top$ comprises the field values at the space-time point set $\{(\mathbf{s}_1, t_1), (\mathbf{s}_2, t_2), \dots, (\mathbf{s}_N, t_N)\}$
- ▶ Space-time energy function

$$\mathcal{H}(\mathbf{x}_S; \theta) = \frac{1}{2\lambda} \left[\frac{(\mathbf{x}_S - m_x \mathbf{I})^2}{N} + c_1 \mathcal{S}_1(\mathbf{x}_S; \mathbf{h}_s, \mathbf{h}_t) \right],$$

- ▶ Kernel function based approximation of the square gradient

$$\mathcal{S}_1(\mathbf{x}_S; \mathbf{h}_s, \mathbf{h}_t) = \langle (x_n - x_k)^2 \rangle,$$

$$\langle (x_n - x_k)^2 \rangle = \frac{\sum_{n=1}^N \sum_{k=1}^N w_{n,k}(\mathbf{r}_{n,k}, \tau_{n,k}; \mathbf{h}_{s,n}, \mathbf{h}_{t,n}) (x_n - x_k)^2}{\sum_{n=1}^N \sum_{k=1}^N w_{n,k}(\mathbf{r}_{n,k}, \tau_{n,k}; \mathbf{h}_{s,n}, \mathbf{h}_{t,n})}.$$

- ▶ Parameters: $\lambda, c_1, m_x, \mu_s, \mu_t$
- ▶ Kernel based weights: $w_{n,k}(\mathbf{r}_{n,k}, \tau_{n,k}; \mathbf{h}_{s,n}, \mathbf{h}_{t,n})$

Different approaches for kernel weight selection

Separable space-time structure

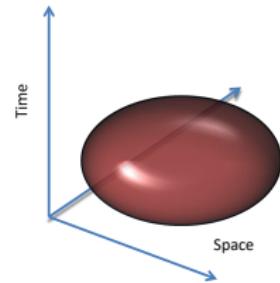
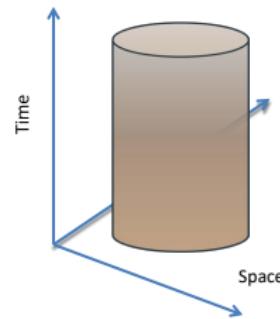
$$w_{n,k} = K\left(\frac{\|\mathbf{r}_{n,k}\|}{h_{s,n}}\right) K\left(\frac{|\tau_{n,k}|}{h_{t,n}}\right)$$

This **does not** imply *separable covariance*

Composite space-time metric

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

Relation of S-T bandwidths: $h_{t,n} = \alpha h_{s,n}$
Currently we use this approach



Different approaches for kernel weight selection

Weight matrix for separable space-time kernel structure

Sampling pattern: N_s locations and N_t times

$$\mathbf{K}_s = \begin{bmatrix} K\left(\frac{\|\mathbf{r}_{1,1}\|}{h_{s,1}}\right) & \dots & K\left(\frac{\|\mathbf{r}_{1,N_s}\|}{h_{s,1}}\right) \\ \vdots & \vdots & \vdots \\ K\left(\frac{\|\mathbf{r}_{N_s,1}\|}{h_{s,N_s}}\right) & \dots & K\left(\frac{\|\mathbf{r}_{N_s,N_s}\|}{h_{s,N_s}}\right) \end{bmatrix}, \mathbf{K}_t = \begin{bmatrix} K\left(\frac{\|\tau_{1,1}\|}{h_{t,1}}\right) & \dots & K\left(\frac{\|\tau_{1,N_t}\|}{h_{t,1}}\right) \\ \vdots & \vdots & \vdots \\ K\left(\frac{\|\tau_{N_t,1}\|}{h_{t,N_t}}\right) & \dots & K\left(\frac{\|\tau_{N_t,N_t}\|}{h_{t,N_t}}\right) \end{bmatrix}$$

Kernel weight matrix: $\mathbf{W} = \mathbf{K}_s \otimes \mathbf{K}_t$

For positive-valued kernel functions the *normalized-weight matrix* is

$\mathbf{U} = \frac{\mathbf{W}}{\|\mathbf{W}\|_1}$, where $\|\mathbf{W}\|_1 = \sum_{k=1}^N \sum_{l=1}^N |W_{k,l}|$, is the *entry-wise L_1 norm*

SLI precision matrix formulation

Quadratic Form of SLI Energy Functional

$$\mathcal{H}(\mathbf{x}_S; \boldsymbol{\theta}) = \frac{1}{2} (\mathbf{x}_S - m_x \mathbf{I})^\top \mathbf{J}(\boldsymbol{\theta}) (\mathbf{x}_S - m_x \mathbf{I})$$

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

$$\mathbf{J}(\boldsymbol{\theta}) = \frac{1}{\lambda} \left\{ \frac{\mathbf{I}}{N} + c_1 \mathbf{J}_1(\mathbf{h}_s, \mathbf{h}_t) \right\}, \quad \boldsymbol{\theta} = (\lambda, c_1, \mathbf{h}_s^\top, \mathbf{h}_t^\top)^\top$$

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

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

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

SLI Mode Predictor for Interpolation of Missing Data

- ▶ The prime (') indicates a fluctuation around m_x
- ▶ Modified energy functional in the presence of an unknown value

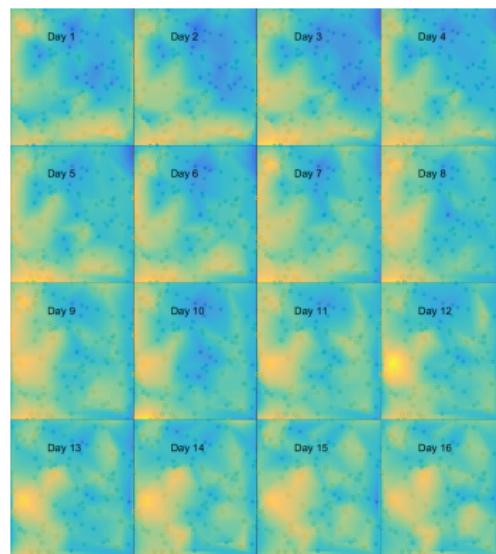
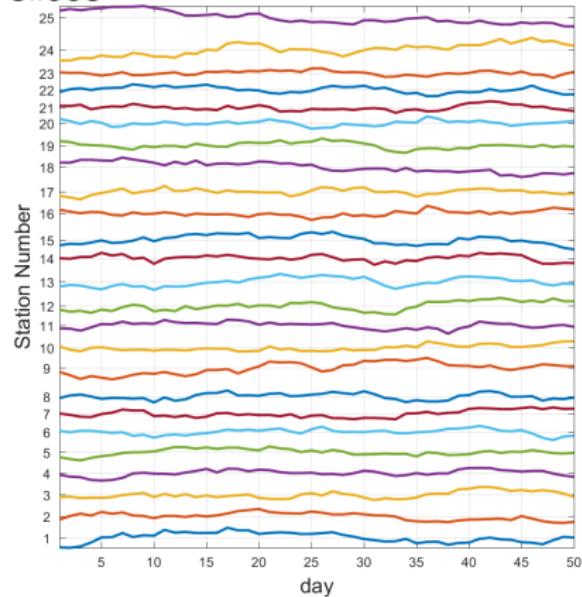
$$\mathcal{H}(\mathbf{x}_S; \boldsymbol{\theta}) = \frac{1}{2} \begin{bmatrix} \mathbf{x}'_S^\top & x'_p \end{bmatrix} \begin{bmatrix} \mathbf{J}_{S,S} & \mathbf{J}_{S,P} \\ \mathbf{J}_{P,S} & J_{p,p} \end{bmatrix} \begin{bmatrix} \mathbf{x}'_S \\ x'_p \end{bmatrix}$$

- ▶ $\hat{x}_p = \arg \min_{x_p} \hat{\mathcal{H}}(\mathbf{x}_S, x_p; \boldsymbol{\theta}^*)$
- ▶ x_p is value at the *prediction point*
- ▶ The predictor is a linear equation that can be evaluated in $\mathcal{O}(N)$:

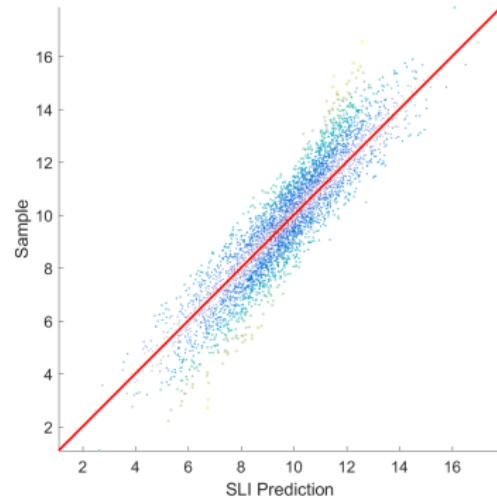
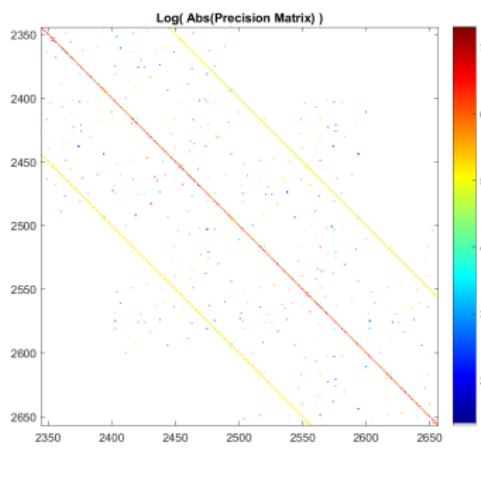
$$\hat{x}_p = m_x - \frac{1}{J_{p,p}(\boldsymbol{\theta}^*)} \sum_{i=1}^N J_{p,i}(\boldsymbol{\theta}^*) (x_i - m_x)$$

Simulated S-T data from stationary random field

100 random locations and 50 time slices



Simulated S-T data from stationary random field

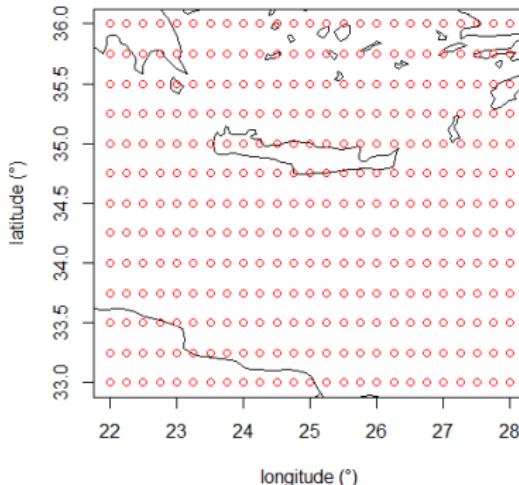


Reanalysis temperature data around Crete

Study design

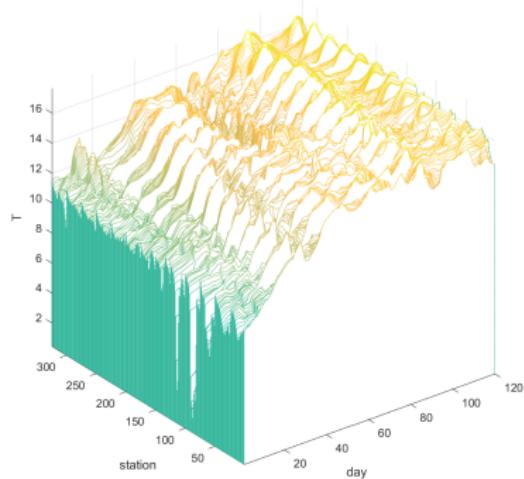
- ▶ Hourly temperature data (degrees) for five days from 01-01-2017 00:00:00 to 05-01-2017 23:00:00
- ▶ Total of 39 000 data points
- ▶ SLI parameter estimation with LOO-CV based on MAE
- ▶ Remove one time slice each time and predict using SLI

Spatial grid 13×25

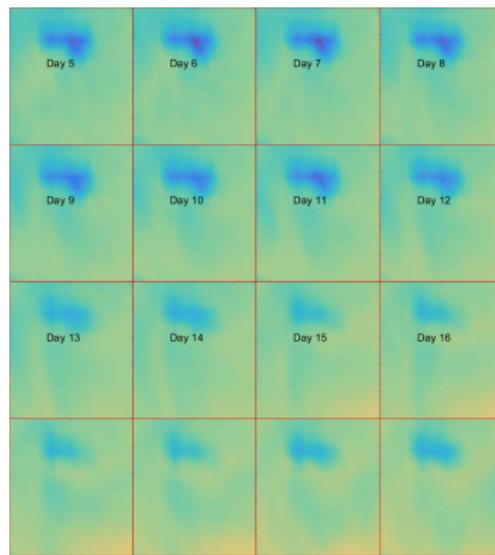


Reanalysis temperature data around Crete

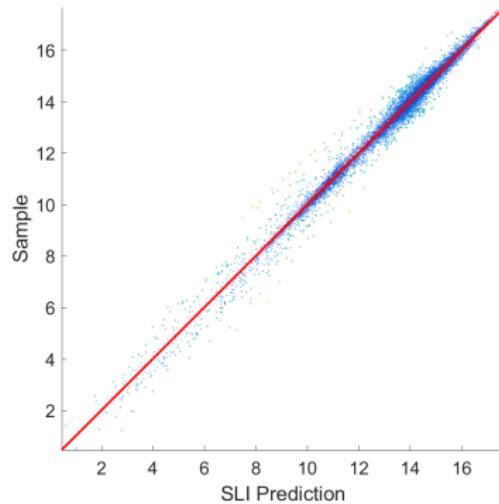
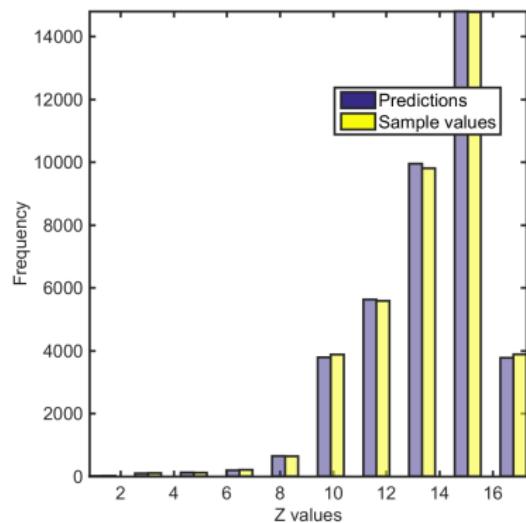
Time series - all sites



First few time slices



Reanalysis temperature data around Crete



SLI precision matrix *sparsity index* $\approx 0.08\%$, i.e., $\approx 123\,000$ non-zero entries out of 1.521×10^9

Conclusions

- ▶ Statistical mechanics is a useful framework for developing **local S-T methods**
- ▶ Linear Langevin equations \Rightarrow Rational spectral densities & Boltzmann-Gibbs probability density functions
- ▶ Linear Langevin equations \Rightarrow Linear covariance PDE
- ▶ The **stochastic local interaction** model (SLI) extends GMRF ideas to scattered data without solving an SPDE
- ▶ SLI employs **compactly supported** kernel functions with **adaptive kernel bandwidths** to build an explicit precision matrix. This structure leads to **semi-explicit prediction equations**

Thank you for your attention!



Acknowledgments: This work was funded by the Operational Program “Competitiveness, Entrepreneurship and Innovation 2014-2020” (co-funded by the European Regional Development Fund) and managed by the General Secretariat of Research and Technology, Ministry Of Education, Research and Religious Affairs under the project DES2iRES (T3EPA-00017) of the ERAnet, ERANETMED_NEXUS-14-049. This support is gratefully acknowledged.

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 (2006). "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. Environ. 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 V. Agou (2018). "Stochastic local interaction model for space-time data," in preparation.