Data-driven Warping of Gaussian Processes for Spatial Interpolation of Skewed Data

Dionissios T. Hristopulos [1]
with Vasiliki Agou and Andreas Pavlides [2]

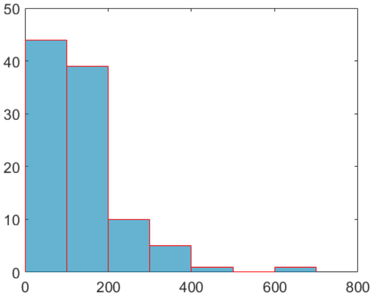
[1]School of Electrical and Electronic Engineering, Technical University of Crete, Chania, Greece

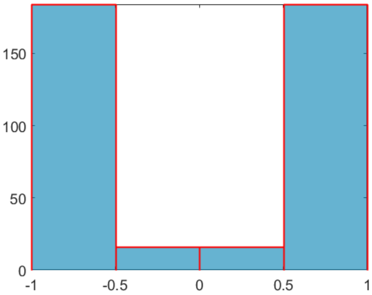
[2]School of Mineral Resources Engineering, Technical University of Crete, Chania, Greece

Presented at the *EGU General Assembly 2022*

Motivation

- Generating accurate estimates and maps from sparse, non-Gaussian data remains a challenge.
- Observed data distributions do not always comply with explicit mathematical models.
- Machine learning is in fashion; how does geostatistics fit in?





What are Gaussian Process Regression (GPR) and Data-driven Warping?

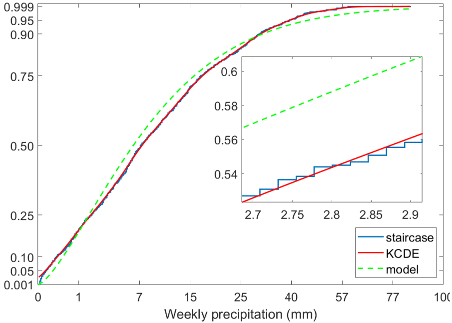
- Gaussian processes are Gaussian random fields defined in feature spaces.
- GPR prediction is “equivalent” to kriging in feature space.
- “Warping” the process means applying a nonlinear transform to normalize the data before the regression step [Snelson et al., 2004].
- Data-driven warping means that the warping transform is learned from the data using a non-parametric, kernel-based estimator.

(1) Agou, Pavlides, Hristopulos. “Spatial Modeling of Precipitation Based on Data-Driven Warping of Gaussian Processes.” *Entropy* 2022, 24, 321.

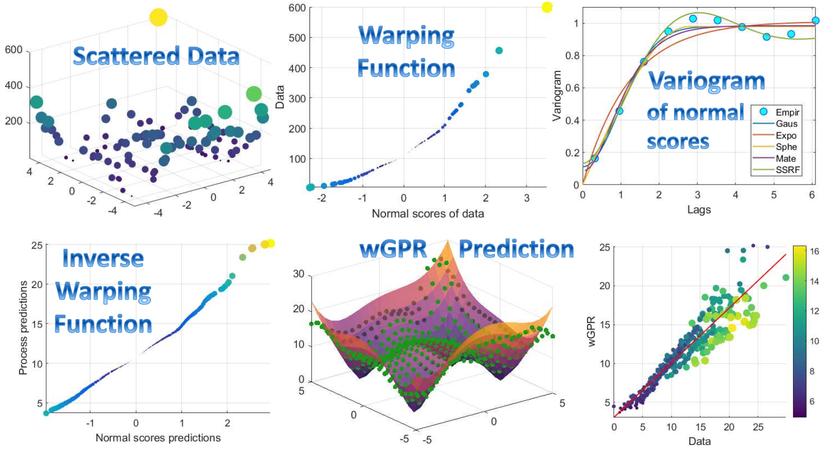
(2) Pavlides, Agou, Hristopulos. “Non-parametric Kernel-Based Estimation of Probability Distributions for Precipitation Modeling.” *arXiv preprint,* arXiv:2109.09961 (2021).

Learning the Cumulative Distribution Function using Kernels

- Kernel functions are used to estimate the CDF.
- CDF steps are smoother than the staircase estimate.
- A theoretical model of the probability distribution is not necessary.



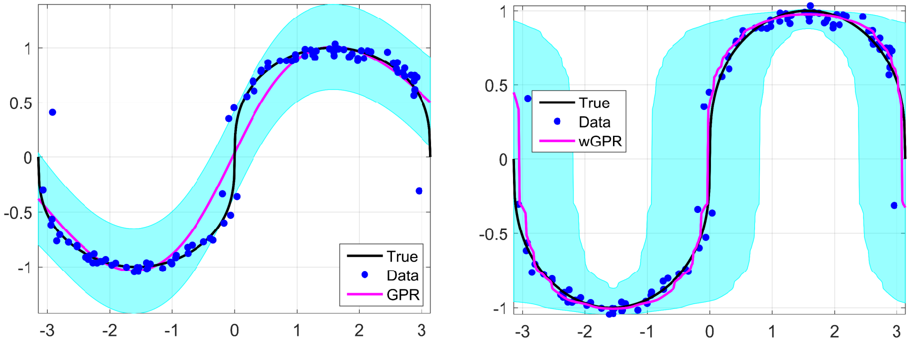
Warped Gaussian Process Regression in a Nutshell



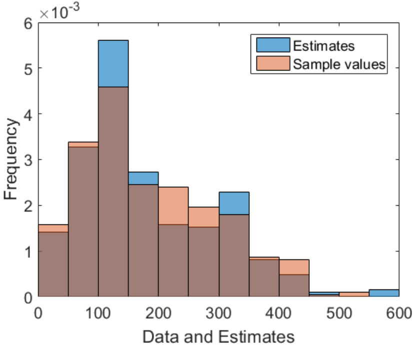
Gaussian Process Regression (GPR) compared to wGPR for Test Function

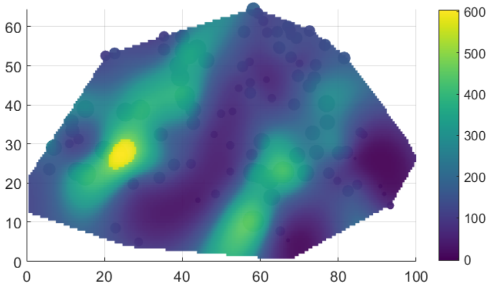
$$x(s) = \left[\sin(\pi s) + \sigma_\epsilon\,\epsilon(s)\right]^{1/3}, \quad s \in [-1, 1]$$

$$\sigma_\epsilon = 0.1,\ \epsilon(s) \sim \mathcal{N}(0, 1)$$



Example: SIC 1997 Swiss Rainfall Data



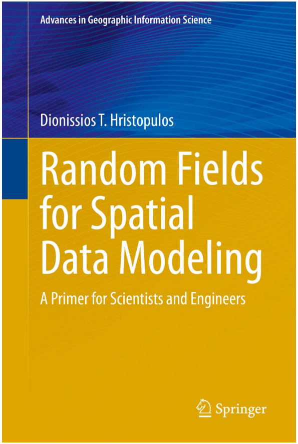


Training: 100 points. Validation: 367 points. Map Grid: 6251 point inside convex hull. Optimal variogram: Spartan variogram (Boltzmann-Gibbs with gradient and curvature terms). Kernel for warping: Epanechnikov.

Contributions

- We introduce data-driven warping which is a flexible, non-parametric approach for normalizing non-Gaussian data.
- Our approach differs from others because the normalizing transform is expressed in terms of kernel functions and the data values.
- In data-driven wGPR we combined data-driven warping with Gaussian process regression, leading to a more flexible spatial prediction method than GPR.
- wGPR allows us to use commonly known geostatistical methods in the broader framework of Gaussian processes.

This research has been funded by the Operational Programme “Human Resources Development Education and Lifelong Learning 2014-2020” co-financed by Greece and the European Union within the framework PA6 - Improving the quality and effectiveness of the educational system in the context of the project “Gaussian Anamorphosis with Kernel Estimators for Spatially Distributed Data and Time Series and Applications in the Analysis of Precipitation” (MIS 5052133).



European Union
European Social Fund

Operational Programme
Human Resources Development, Education and Lifelong Learning

Co-financed by Greece and the European Union

ΕΣΠΑ 2014-2020
ανάπτυξη - εργασία - αλληλεγγύη