5th International Workshop on Climate Informatics
September 24-25, 2015
Hosted by the National Center for Atmospheric Research in Boulder, CO

SEA-LEVEL ESTIMATION USING THE RIEMANNIAN MANIFOLD AND A NON-STATIONARY COVARIANCE FUNCTION

Chintan A. Dalal[1], Vladimir Pavlovic[1], Robert E. Kopp[2]

***Abstract*—Analyzing datasets, such as sea-level records, pose a challenging statistical problem for reasons including non-stationarity, non-uniformly smooth spatial boundaries, and sparsity in the data. In this paper, we propose a framework to estimate the non-stationary covariance function by employing intrinsic statistics on the local covariates. These local covariates represent the underlying local correlation in the measurements, and they are assumed to lie on a Riemannian manifold of positive definite matrices. Additionally, we provide a technique for data-assimilation of correlated natural processes in order to improve the regression estimates arising from spatially sparse datasets. Experiments on a synthetic and real dataset of relative sea-level changes across the world demonstrate improvements in the error metrics for the regression estimates using our newly proposed approach.**

I. MOTIVATION

Analyzing complex natural processes, such as the global distribution of sea-level changes, poses many statistical challenges. One such challenge is modeling the non-stationarity of the underlying stochastic process, which has been addressed by [1], [2], [3], and [4] in the context of climate related variables. [1] and [3] use the spatially evolving smooth kernels to represent the characteristic length scale (local covariates) of the covariance function, which, in turn, model the local correlation structure of the stochastic process. In our experiments, we found that the approach of [1] was the most promising as its model is entirely in the non-parametric Gaussian Process (GP) framework. Even so, [1] uses a homogeneous set of hyperparameters to ensure smoothness of the local covariates in the input space. We show in our experiments that using a homogeneous set of hyperparameters is inefficient for global-scale sea-level changes. Additionally, none of the aforementioned approaches address the issues of non-uniformly smooth spatial boundaries and sparsely distributed measurements in the datasets.

In this work, we propose a method to model the global scale non-stationary stochastic process, as well as a method to non-parametrically model the boundaries of the regional geophysical variability (i.e. non-uniformly smooth spatial boundary) and the irregular geospatial measurements of the response variable (i.e. spatially sparse dataset). Addressing these issues can be useful in capturing the region-wise correlation structure and in reducing biases in the regression estimates.

II. METHOD

To address the aforementioned issues at a global scale, we propose a model of the covariance function (as described in [5]) that considers the inherent geometric structure of the local covariates to lie on a Riemannian manifold of positive definite matrices [6], [7]. This covariance function, which we call the intrinsic non-stationary covariance function, can be used in the standard krigging methods (such as [8]) to give the final regression estimates for the sea-level datasets.

Figures 1 and 2 depict the general framework of our approach, and following are the general steps for designing such a covariance function:

1) Initialize the local covariate estimates by computing the spectral decomposition of the input space (latitude, longitude) training data points (as described in [1]).
2) Improve the local covariate estimates by computing the second-order intrinsic statistics on the manifold of the initialized local covariates (as described in [5]).
3) If a separate source of the dataset is available and is geophysically correlated, then apply the data-assimilation technique. For this step, compute the above steps for the additional dataset, and then again apply the second-order intrinsic statistics on the manifold of all of the collected datasets' local covariates.
4) Finally, the local covariates are spatially convolved to give the estimates of the characteristic length scale

Corresponding author: C. Dalal, Rutgers, The State University of New Jersey, Piscataway, NJ, chintan.dalal@rutgers.edu,[1] Department of Computer Science, Rutgers, NJ, [2] Department of Earth & Planetary Sciences, Rutgers, NJ.

(CLS). This CLS is then used to form the non-stationary covariance function (as described in [5]).

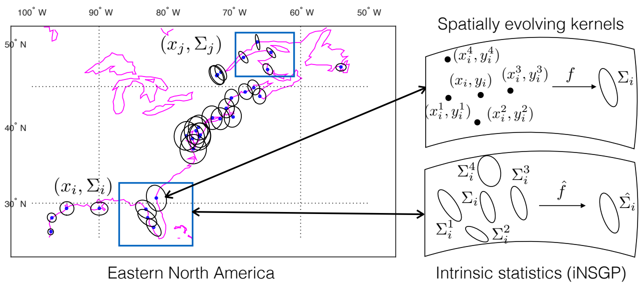

Fig. 1: Framework for the intrinsic non-stationary covariance function.

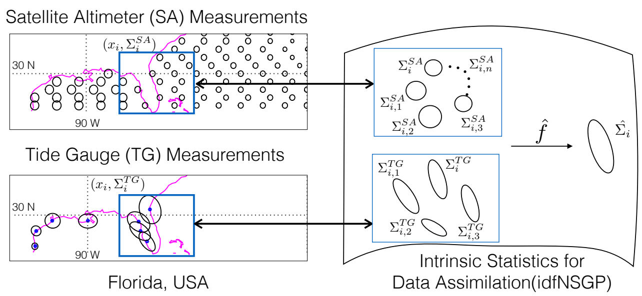

Fig. 2: Framework for data assimilation using the intrinsic non-stationary covariance function.

III. Evaluation

To gain insight into the applicability of our proposed covariance function, we implemented and compared kriging with four different covariance functions: the widely used stationary anisotropic Matérn covariance function (statGP) of [9], the baseline non-stationary covariance function (NSGP) of [1], and the intrinsic non-stationary covariance function (that we propose in this paper) using a single source (iNSGP) and multiple sources (idfNSGP). These methods were evaluated using two standard performance measures for kriging: the standardized mean squared error (sMSE) and the negative log predictive density (nLPD).

The two datasets used are: the geophysics driven synthetic data of the relative sea level (Glacio-Isostatic Adjustment - GIA), as modeled in [10], and the global-scale real dataset (Tide Gauge - TG) collected in [11]. To perform data-assimilation experiments, two additional datasets were used: the GIA vertical land motion, as modeled in [10], and the satellite altimeter [12].

Table I summarizes the performance of the four covariance functions (statGP, NSGP, iNSGP, and idfNSGP) when implemented on the two datasets (GIA and TG). The table also shows the improvements in the evaluation metric using iNSGP, as well as the best error metric value achieved using idfNSGP. These results suggest that our proposed covariance function improves the statistical estimates of the sea-level datasets. Future work will explore additional sources of correlated datasets that could further aid in sea-level estimation across the world.

TABLE I: Evaluations of the synthetic (GIA) and real (TG) datasets. Datasets' results are in units of mm/year.

DATASET	Metric	StatGP	NSGP	iNSGP	idfNSGP
GIA	sMSE	0.58	0.56	0.54	**0.52**
	nLPD	3.08	2.00	1.94	**1.91**
TG	sMSE	0.85	0.75	0.71	**0.66**
	nLPD	2.81	2.82	2.78	**2.75**

Acknowledgments

This work is supported by the U.S. Department of Homeland Security under Grant Award Number 2012-ST-104-000044. The authors would like to thank Netta Gurari, Chao Chen, Danna Gurari, and the anonymous reviewers for their feedback.

References

[1] C. Paciorek and M. Schervish, "Nonstationary covariance functions for gaussian process regression," *Advances in neural information processing systems*, vol. 16, pp. 273–280, 2004.
[2] D. Nychka, C. Wikle, and J. A. Royle, "Multiresolution models for nonstationary spatial covariance functions," *Statistical modelling*, vol. 2, no. 4, pp. 315–331, 2002.
[3] D. Higdon, J. Swall, and J. Kern, "Non-stationary spatial modeling," *Bayesian statistics*, vol. 6, no. 1, pp. 761–768, 1999.
[4] M. Stein, *Interpolation of spatial data: some theory for kriging*. N.Y.: Springer, 1999.
[5] C. A. Dalal, V. Pavlovic, and R. E. Kopp, "Intrinsic non-stationary covariance function for climate modeling," *ArXiv e-prints*, July 2015.
[6] S. Amari, *Differential geometry in statistical inference*, vol. 10. Inst. of mathematical statistic, 1987.
[7] X. Pennec, "Intrinsic statistics on riemannian manifolds: Basic tools for geometric measurements," *Journal of mathematical imaging and vision*, vol. 25, no. 1, pp. 127–154, 2006.
[8] C. E. Rasmussen and C. Williams, "Gaussian processes for machine learning. 2006," *The MIT press, Cambridge, MA, USA*, vol. 38, pp. 715–719, 2006.
[9] B. Matérn *et al.*, "Spatial variation. stochastic models and their application to some problems in forest surveys and other sampling investigations.," *Meddelanden fran statens skogsforskningsinstitut*, vol. 49, no. 5, 1960.
[10] W. Peltier, "Global glacial isostasy and the surface of the ice-age earth: the ice-5g (vm2) model and grace," *Annu. rev. earth planet. sci.*, vol. 32, pp. 111–149, 2004.
[11] S. J. Holgate, A. Matthews, P. L. Woodworth, L. J. Rickards, M. E. Tamisiea, E. Bradshaw, P. R. Foden, K. M. Gordon, S. Jevrejeva, and J. Pugh, "New data systems and products at the permanent service for mean sea level," *Journal of coastal research*, vol. 29, no. 3, pp. 493–504, 2012.
[12] C. M. Domingues, J. A. Church, N. J. White, P. J. Gleckler, S. E. Wijffels, P. M. Barker, and J. R. Dunn, "Improved estimates of upper-ocean warming and multi-decadal sea-level rise," *Nature*, vol. 453, no. 7198, pp. 1090–1093, 2008.