DISCUSSION

Comments on: Comparing and selecting spatial predictors using local criteria

Stefano Castruccio · Marc G. Genton

1 Introduction

We would like to congratulate the authors for producing this interesting and novel methodology for comparison of spatial fields. It is important that this problem be discussed and expanded, especially in the context of large data sets, which is undoubtedly one of the main challenges that any statistician working with environmental space–time data has to face in contemporary research. We have divided our comments in three sections that cover separate (albeit intertwined) topics. In Sect. 2, we link this work with previous literature on hypothesis testing with multiple predictors and discuss the consequences of not accounting for predictor uncertainty. In Sect. 3, we discuss the interpretability of Locally Selected Predictors (LSP) and Globally Selected Predictors (GSP) estimates and potential challenges around this issue, while in Sect. 4, we expand some points on oversmoothing and scalability for large data sets that have arisen in this work.

2 Uncertainty in spatial predictors

In the case of two spatial predictors, the previous literature has developed hypothesis tests with the null hypothesis that there is no difference between the predictors either for every location (Sheng et al. 2002; Sendur et al. 2005; Pavlicová et al. 2008) or

---

This comment refers to the invited paper available at doi:10.1007/s11749-014-0415-1.

S. Castruccio
School of Mathematics & Statistics, Newcastle University, Newcastle upon Tyne NE1 7RU, UK
e-mail: stefano.castruccio@ncl.ac.uk

M. G. Genton (✉)
CEMSE Division, King Abdullah University of Science and Technology,
Thuwal 23955-6900, Saudi Arabia
e-mail: marc.genton@kaust.edu.sa

with a single hypothesis that on average the predictors are the same (Hering and Genton 2011). Bradley, Cressie and Shi define a methodology for selecting predictors at unsampled locations with an estimate of the square prediction error, thereby allowing for maps such as the ones in Figure 8 of the paper that show where different predictors perform better (more in the next sections). This methodology also allows for comparison of multiple predictors. This approach however does not allow for formal testing as in the aforementioned papers, since not all spatial predictors come with a measure on uncertainty (and consequently spatial correlation), as acknowledged in the final discussion. This makes the interpretation of the patterns in Figure 8 potentially problematic, because it is not possible to understand if a method is better in one area simply by chance. Besides, there could be patterns of correlations among predictors that could lead to simultaneous good or bad prediction in some locations.

3 Interpretability of spatial predictors comparisons

An issue that arises in the use of LSP instead of a GSP is interpretability. The use of a global criterion and consequently of a single predictor for all domain D does not acknowledge that some predictors are locally better than others, but it does allow a reconstructed map that retains more physical interpretability to be chosen. A similar issue arises in climate science when we need to determine what is the "best" global or regional circulation model in a multi-model ensemble, and how to combine information to gain better predictions (see for example Jun et al. 2008b, a; Knutti 2010; Tebaldi et al. 2004, 2005). In that context, the field output, be it any physical variable, represents the solution of systems of partial differential equations (PDEs): it therefore has an important physical meaning. On the other hand, while reconstructing a field by merging information from multiple fields could lead to a problematic interpretation, a map of the selected predictor for different locations can be helpful to discuss where some methods fail, and why. We found the maps in the left column of Figure 8 of the paper particularly interesting, because there are clearly some spatial patterns, especially for the moving-window predictor selection (MWS) and nearest-neighborhood predictor selection (NNS), that could help to understand what is the best interpolation method, and why. Are these patterns due to the geometry of the retrieval locations? Moreover, it would be interesting to see the percentages of success of different methods adjusted for regions of interest for impact assessments (urban land areas).

4 Smoothness in spatial predictors and the big data challenge

Another interesting point of discussion is the use of scalable algorithms for large data sets such as the AIRS retrievals in the work, comprising approximately 75,000 locations. The authors have given a detailed review of techniques to overcome these challenges in Sect. 3.1, and we would like to further expand on some of their remarks. One of the issues that arises with spatial interpolation techniques that reduce the rank of the covariance matrix is oversmoothing (Stein 2014), as is noticeable in the second panel in Figure 7. [Regarding the Stochastic PDE approach (SPD), we were quite

surprised to see such a smooth reconstructed field from this, and we guess the authors have used a stochastic PDE with constant coefficients, which implies an isotropic Matérn field as a reference for the Markov Random Field approximation. It would be interesting to see how spatially varying parameters could change the field, if they could add some structure, and how this would impact Figure 8.] A possible, alternative solution could be to divide the world into multiple regions and perform a localized kriging in each one of them. This would imply a subjective choice of the regions and suboptimal performances near the boundaries, but this method is scalable and could be a useful approach if the scientific questions arising from the reconstructed field are geographically constrained. Another, more computationally demanding approach would be to perform a global kriging. With the data and the computer at hand, it is not an impossible task: 75,000 retrievals imply a matrix of approximately 42Gb (in double precision, ignoring matrix symmetry) and the matrix inversion required for kriging can be performed using iterative methods. This approach will not be directly applied if a time window larger than the nine days is considered. In this regard, we believe that the development of scalable methods for large space–time data sets it still in its infancy, and this is evident just from the referral of data sets of the size of less than 10^5 as "very large" or even "massive", when truly massive data sets [most noticeably from climate model output; for example, the latest CMIP5 multi-model ensemble (Taylor et al. 2012)] can be easily several orders of magnitude larger. In some cases, especially when data are gridded, it is possible to formulate scalable algorithms for tens of millions of data points using large memories and a large pool of processors (Castruccio and Stein 2013; Castruccio and Genton 2014), but in cases where the geometry is not gridded, its extension is not immediate.

References

Castruccio S, Stein ML (2013) Global space–time models for climate ensembles. Ann Appl Stat 7(3):1593–1611

Castruccio S, Genton MG (2014) Beyond axial symmetry: an improved class of models for global data. Stat 3(1):48–55

Hering A, Genton MG (2011) Comparing spatial predictions. Technometrics 53(4):414–425

Jun M, Knutti R, Nychka DW (2008a) Local eigenvalue analysis of CMIP3 climate model errors. Tellus A 60:992–1000

Jun M, Knutti R, Nychka DW (2008b) Spatial analysis to quantify numerical model bias and dependence: how many climate models are there? J Am Stat Assoc 103(483):934–947

Knutti R (2010) The end of model democracy? Climatic Change 102(3–4):395–404. doi:10.1007/s10584-010-9800-2

Pavlicová M, Santer TJ, Cressie N (2008) Detecting signals in FMRI data using powerful FDR procedures. Statist Its Interf 1:23–32

Sendur L, Maxim V, Whitcher B, Bullmore E (2005) Multiple hypothesis mapping of functional MRI data in orthogonal and complex wavelet domains. IEEE Trans Signal Process 53(9):3413–3426

Sheng X, Huang HC, Cressie N (2002) Nonparametric hypothesis testing for a spatial signal. J Am Stat Assoc 97:1122–1140

Stein ML (2014) Limitations on low rank approximations for covariance matrices of spatial data. Spatial Stat 8:1–19

Taylor K, Stouffer R, Meehl G (2012) An overview of CMIP5 and the experiment design. Bull Am Meteorol Soc 93:485–498. doi:10.1175/BAMS-D-11-00094.1

Tebaldi C, Mearns LO, Nychka DW, Smith RL (2004) Regional probabilities of precipitation change: a Bayesian analysis of multimodel simulations. Geophys Res Lett 31(24). doi:10.1029/2004GL021276

Tebaldi C, Smith R, Nychka DW, Mearns LO (2005) Quantifying uncertainty in projections of regional climate change: a Bayesian approach to the analysis of multimodel ensembles. J Climate 18:1524–1540