

Discussion of “analysis of spatio-temporal mobile phone data: a case study in the metropolitan area of Milan”

by P. Secchi, S. Vantini, and V. Vitelli

Helle Sørensen¹  · Bo Markussen¹ · Anders Tolver¹

Accepted: 23 April 2015 / Published online: 1 May 2015
© Springer-Verlag Berlin Heidelberg 2015

1 Introduction

We congratulate the authors on an interesting paper, and we are delighted to have the opportunity to comment on the paper.

The authors have successfully analyzed a large dataset with a complicated time and space structure and have obtained a dimension reduction via a spatio-temporal decomposition with clear interpretations. In order to achieve this, two algorithms from the recent literature, treelet analysis and Voronoi tessellations, were combined in an ingenious way and accommodated to the present problem.

2 Smoothing, type of time-decomposition, and sampling

We would like to emphasize that many decisions must be taken during complex analyses as the present one. In the following we take the opportunity to discuss some of the choices that have been made. Some of them, but not all, are already mentioned by the authors.

Smoothing The first part of the analysis consists of data smoothing. This is common practice when working with functional data, but (a) it is not always necessary, (b) it introduces an extra step and in that sense complicates the analysis, and (c) the effect of smoothing is often ignored later in the evaluation of the results. In the present application a simpler solution would be to use the raw mean over the two weeks.

✉ Helle Sørensen
helle@math.ku.dk

¹ Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen Ø, Denmark

In the end, smoothing would implicitly be obtained in the bagging step where the compositions from different tessellations are combined. The authors mention missing data as an argument for smoothing the data, but the treelet algorithm would work on incomplete data as well since it requires only computation of pairwise correlations and PCA in two directions.

Basis for time-decomposition The treelet algorithm which is used for the decomposition in time, is developed for unordered variables (multivariate data). Hence, the time-ordering of the data is not used for the computations but only for the interpretation of the components. This may be considered an advantage as no explicit assumptions are needed on the structure over time; smoothness in the components are driven exclusively by correlation in the data. However, one may also argue that taking advantage of the inherent time-structure in the data could strengthen the analysis. Moreover, the data-driven nature of treelets introduces the need for an extra “matching” step in the bagging algorithm since the order of components is not comparable between different bootstrap samples.

Functional principal component analysis (FPCA) is discarded by the authors because it does not yield sparse linear combinations of the variables and therefore tends to give components that are more difficult to interpret. However, FPCA can be accommodated to achieve sparse representations. One option is rotation of the selected principal components. Another is to introduce penalty terms that force PC components to be localized. An example is the work on fused loadings (Guo et al. 2010). Here, a so-called fusion penalty is used to create blocks of highly correlated variables and force loadings for variables in the same block to be identical. The block structure is determined by the correlation structure of the data; no ordering of the variables is assumed.

With functional data the natural ordering in time could be utilized by explicitly penalizing changes in the principal components. More precisely, we suggest to minimize

$$\sum_{j=1}^J \sum_{\mathbf{x} \in S_0} \left\| E_{\mathbf{x}}(t_j) - \sum_{k=1}^K d_k(\mathbf{x}) \psi_k(t_j) \right\|^2 + \lambda_1 \sum_{k=1}^K \sum_{j=1}^J |\psi_k(t_j)| + \lambda_2 \sum_{k=1}^K \sum_{j=1}^{J-1} |\psi_k(t_{j+1}) - \psi_k(t_j)|$$

subject to orthogonality constraints in order to choose components ψ_k . The first penalty term drives the components to be sparse, whereas the second penalty term drives the components to be constant on intervals. In other words this approach is data driven, gives sparse solutions, and takes the time-ordering in the data into account.

Distribution of Voronoi cells The Voronoi cells are sampled according to the uniform distribution on the lattice, such that all areas of the lattice are represented equally often. We wonder if it would be more efficient to let the sampling intensity vary over the grid. For example, there is a high mobile activity at the railway station which is

covered by a single site in the grid. Supposedly, such a local feature would be easier to catch if cells around the railway station were selected more often as Voronoi cells.

In relation to this, it could be interesting to study the bootstrap variance of the sample $\{\tilde{d}_j^b(x)\}_{b=1}^B$ site by site, i.e. consider

$$TV_n(\mathbf{x}) = \sum_{j=1}^J \text{Var}_b\{\tilde{d}_j^b(\mathbf{x})\}.$$

The *TAV* criterion is the average over sites of $TV_n(\mathbf{x})$, and the number n of Voronoi cells is selected as to minimize *TAV*. Site-wise minimization of a smoothed version of $TV_n(\mathbf{x})$ wrt. n would suggest regions where the Voronoi cells should preferably be dense and regions where they should preferably be sparse. The analysis would thus indicate how Voronoi cells could preferably be distributed over the grid.

A related comment is concerned with the computation of local representatives. In the paper each site is allocated to the closest Voronoi cell and used for the computation of exactly one local representative. An alternative would be to use a kernel smoother on all sites in the vicinity of the Voronoi cell as this allows each site to contribute to several (or none) local representatives.

3 Perspectives and impact

A substantial amount of research on spatially dependent functional data has emerged recently [see e.g. (Delicado et al. 2010) and references therein]. Two problems have received much attention: smoothing taking into account the spatial dependence and prediction of functional signals at unobserved spatial locations. Dimension reduction for dependent functional data has been studied in various scenarios: time series of functional data (Hörmann et al. 2015), spatially correlated multilevel functional data (Staicu et al. 2010), and spatial functional data (Hörmann and Kokoszka 2013; Liu et al. 2014). These papers mainly rely on techniques from FPCA. Although we acknowledge that the above methods may not solve the present research question, it would be interesting to know if parts of the approaches from the literature could be accommodated for an analysis of the mobile data (or vice versa).

We believe that the suggested method may find application in other areas of applied statistics. One particular example is temporal recording of neural activity in the brain as response to various stimuli (Harvey and Roland 2013). Strong spatial dependence between signals from co-working areas must be expected, and a spatio-temporal decomposition of the data may reveal interesting patterns about communication dynamics in the brain.

This example also demonstrates the perspectives of realizing that the role of space and time may be interchanged in the suggested method. The goal of the analysis may as well be a decomposition of the form

$$E_t(\mathbf{x}) = \sum_{k=1}^K D_k(t)\psi_k(\mathbf{x})$$

where we look for spatially localized components ψ_k (in the neuroscience application representing co-working areas of the brain). The spatial components may be estimated using treelet analysis in a two-dimensional grid, and the tessellation approach may be applied along the one-dimension time argument to adapt to local smoothness of the signals in the time direction.

In summary, we are impressed of the analysis with its extraction of interpretable information from a large spatio-temporal data set. The paper contains new methodology for dependent functional data, and future research will contribute to the understanding of advantages and disadvantages of the various approaches. We see a great potential for new data applications and look forward to following the development in the field.

References

- Delicado P, Giraldo R, Comas C, Mateu J (2010) Statistics for spatial functional data: some recent contributions. *Environmetrics* 21:224–239
- Guo J, James G, Levina E, Michailidis G, Zhu J (2010) Principal component analysis with sparse fused loadings. *J Comput Graph Stat* 19:930–946
- Harvey MA, Roland, PE (2013). Laminar firing and membrane dynamics in four visual areas exposed to two objects moving to occlusion. *Front Syst Neurosci* 7. doi:[10.3389/fnsys.2013.00023](https://doi.org/10.3389/fnsys.2013.00023)
- Hörmann S, Kokoszka P (2013) Consistency of the mean and the principal components of spatially distributed functional data. *Bernoulli* 19:1535–1558
- Hörmann S, Lukasz K, Hallin M (2015) Dynamic functional principal components. *J R Stat Soc Ser B Stat Methodol* 77:319–348
- Liu C, Ray S, Hooker G (2014) Functional principal components analysis of spatially correlated data. Tech. rep., [arXiv:1411.4681](https://arxiv.org/abs/1411.4681)
- Staicu AM, Crainiceanu CM, Carroll JR (2010) Fast methods for spatially correlated multilevel functional data. *Biostatistics* 11:177–194