A shiny web application for disease mapping. Making easy the fit of spatio-temporal models.

Aritz Adin
aritz.adin@unavarra.es
(joint work with Tomás Goicoa and Lola Ugarte)

Department of Statistics, Computer Science and Mathematics
Public University of Navarre

UseR! 2019

Toulouse, 9-12 July 2019

Spatial Statistics
Public University of Navarre

upna
Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

Introduction

- Spatial and spatio-temporal analyses of count data are crucial in epidemiology and other fields to
  1. provide accurate estimates of mortality and/or incidence risks.
  2. unveil the underlying spatial and spatio-temporal patterns.
- However, fitting spatial and spatio-temporal models is not easy for non-expert users.
- Here, we present the interactive web application SSTCDapp for the analysis of spatial and spatio-temporal mortality (or incidence) count data, which is addressed at

  https://emi-sstcdapp.unavarra.es/

The SSTCDapp application

SSTCDapp[1] is an interactive and user-friendly web application designed for the following purposes:

- To perform descriptive analyses in space and time of mortality/incidence risks or rates.
- To fit an extensive range of fairly complex spatial and spatio-temporal models for areal data commonly used in disease mapping.

It is built with Shiny and relies on the well-founded integrated nested Laplace approximation (INLA) technique for Bayesian inference[2] is used for model fitting through the `R-INLA` package.

[1] Adin, A., Goicoa, T., and Ugarte, M.D. (2019). Online relative risks/rates estimation in spatial and spatio-temporal disease mapping. *Computer Methods and Programs in Biomedicine*, 172, 103-116

[2] Rue, H., Martino, S., and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71, 319–392

The SSTCDapp application

SSTCDapp[1] is an interactive and user-friendly web application designed for the following purposes:

- To perform descriptive analyses in space and time of mortality/incidence risks or rates.

- To fit an extensive range of fairly complex spatial and spatio-temporal models for areal data commonly used in disease mapping.

It is built with Shiny and relies on the well-founded integrated nested Laplace approximation (INLA) technique for Bayesian inference[2] is used for model fitting through the `R-INLA` package.

[1] Adin, A., Goicoa, T., and Ugarte, M.D. (2019). Online relative risks/rates estimation in spatial and spatio-temporal disease mapping. *Computer Methods and Programs in Biomedicine*, 172, 103-116

[2] Rue, H., Martino, S., and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71, 319–392

The SSTCDapp application

Main characteristics:

- Users can access the application directly from the web browser without installing any software in their computers.
- All the analysis and computations are made in a remote server.
- The users can submit a model to the remote server and collect the results when the computations are finished.
- A desktop version is also available to run the application locally in those cases in which data confidentiality could be a serious issue.

The SSTCDapp application

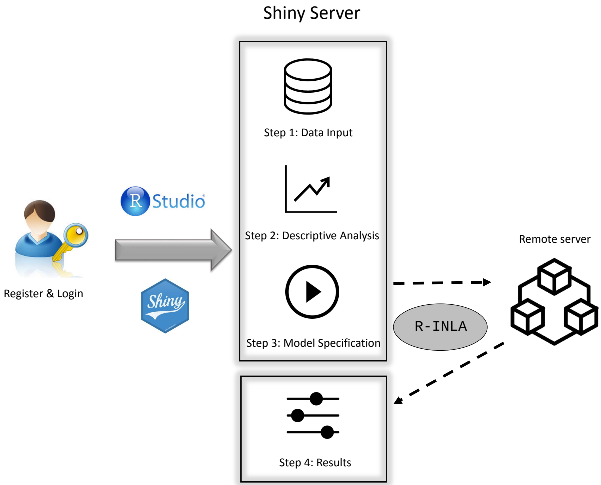


Figure 1: Workflow of the SSTCDapp application.

1. **Data Input**:
   - The data and the associated cartography are uploaded by the user, and automatically previewed on the screen.
   - Several formats for both data and cartography are supported.

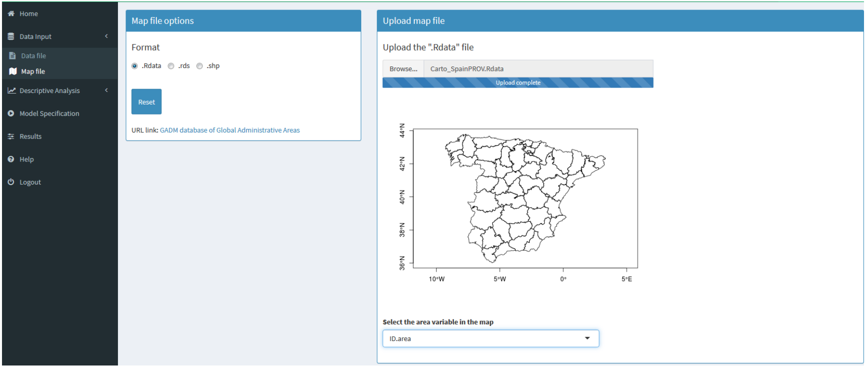
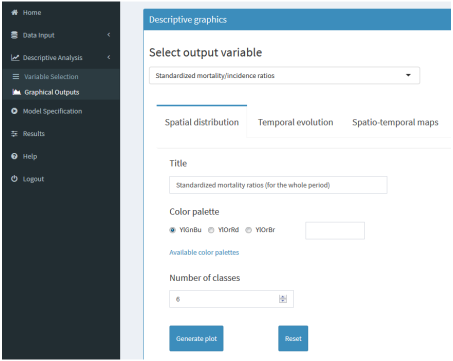


Figure 2: *Map file* input tab in SSTCDapp.

2. **Descriptive Analysis**:
   - The target variables are selected and standardized mortality ratios (SMR) or standardized rates (SR) are calculated.
   - Descriptive graphs of the spatial, temporal, and spatio-temporal distribution for the variables of interest (crude rates, SMR or SR) are generated.



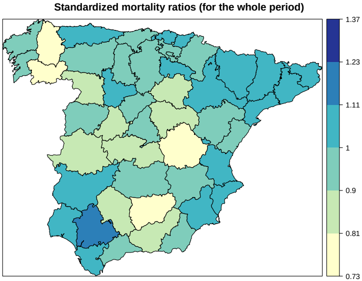


Figure 3: Plot generated from the *Graphical Outputs* tab in SSTCDapp.

3. **Model Specification**:
   - A wide variety of spatial or spatio-temporal models commonly used in disease mapping can be fitted using the `R-INLA` package.
   - The model is submitted to a remote server. Once the calculations are finished the user will receive a notification by email.

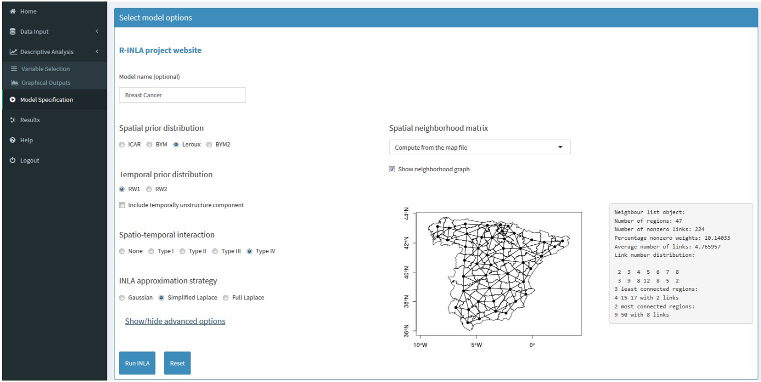


Figure 4: *Model Specification* tab in SSTCDapp.

4. **Results**:

- Summary measures are provided for the posterior distribution of model hyperparameters; relative mortality/incidence risks (or rates); and spatial, temporal, and spatio-temporal patterns.

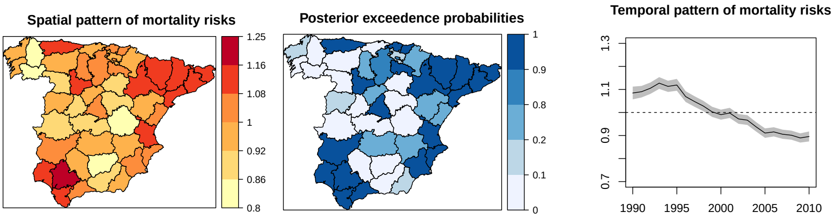


Figure 5: Posterior mean estimates of spatial and temporal patterns of breast cancer mortality risks.

Example: Breast cancer mortality data in Spanish provinces

4. **Results**:

   - Maps with the geographical distribution of the disease risks and area-specific temporal evolutions are also generated.

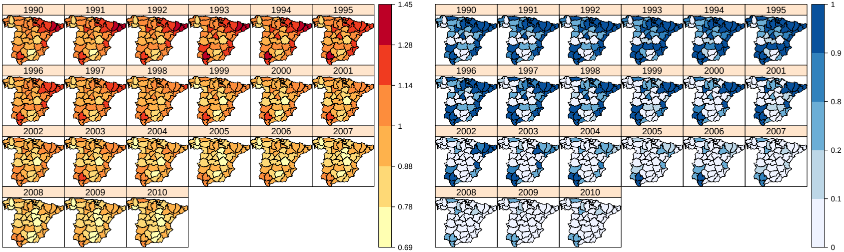


Figure 6: Posterior means of relative risks (left) and posterior exceedence probabilities (right) for breast cancer mortality in Spanish provinces.

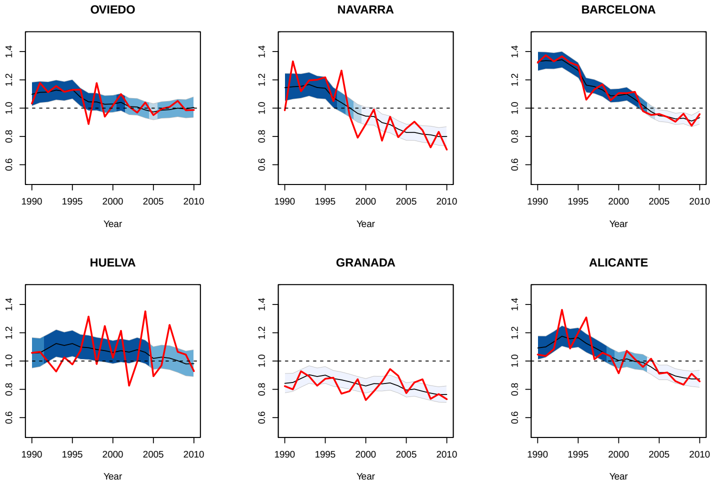


Figure 7: Temporal evolution of breast cancer mortality relative risks for six selected Spanish provinces and 95% two-sided credible intervals. The colors used in the bands are associated to the posterior exceedence probabilities of relative risks being greater than one.

Conclusions

- The SSTCDapp was mainly developed to estimate relative risks using spatial and spatio-temporal disease mapping models.
- It provides separate spatial, temporal, and spatio-temporal patterns together with the corresponding exceedence probabilities and/or credibility intervals.
- The key advantage of this application in comparison with other software commonly used in disease mapping is that it provides an easy-to-use interface that facilitate the fit of fairly complex models without installing any software in user's computer.

Future development of the application

- Integration of `sf` (simple feature) objects as cartography files to generate maps.
- To include interactive data visualization graphs using the `R` packages `leaflet` and `tmap`.
- To implement other spatio-temporal model proposals such as
  - B-spline models accounting for both spatial and temporal correlation.[3]
  - Models for age-specific mortality/incidence patterns.[4]
  - Models to estimate disease risks in the presence of local discontinuities and clusters.[5]

---

[3] Ugarte, M.D., Adin, A., and Goicoa, T. (2017). One-dimensional, two-dimensional, and three dimensional B-splines to specify space-time interactions in Bayesian disease mapping: model fitting and model identifiability. *Spatial Statistics*, 22, 451-468

[4] Goicoa, T., Adin, A., Etxeberria, J., Militino, A.F., and Ugarte, M.D. (2019). Flexible Bayesian P-splines for smoothing age-specific spatio-temporal mortality patterns. *Statistical Methods in Medical Research*, 28, 384–403

[5] Adin, A., Lee, D., Goicoa, T., and Ugarte, M.D. (2019). A two-stage approach to estimate spatial and spatio-temporal disease risks in the presence of local discontinuities and clusters. *Statistical Methods in Medical Research (in press)*

Developers and mantainers

The **SSTCDapp application** was developed by Aritz Adin together with the Spatial Statistics Group of the Public University of Navarre (Spain) under the following grants:

- Spanish Ministry of Economy and Competitiveness (Project MTM2014-51992-R).
- Health Department of the Navarre Government (Project 113, Res.2186/2014).
- AEI/FEDER, UE (Project MTM2017-82553-R).

For bug reports and support, please use the following email account:

app.estadistica@unavarra.es

Spatial Statistics Public University of Navarre

upna Universidad Pública de Navarra Nafarroako Unibertsitate Publikoa