Spatiotemporal BME Analysis and Mapping of Mortality Distribution in California

Kyung-mee Choi
University of North Carolina at Chapel Hill, Dept. of Env. Sci. and Eng.
C.B. 7400, Chapel Hill, NC, USA
kmchoi@email.unc.edu

Marc L. Serre
University of North Carolina at Chapel Hill, Dept. of Env. Sic. and Eng.
marc_serre@unc.edu

George Christakos
University of North Carolina at Chapel Hill, Dept. of Env. Sci. and Eng.
george_christakos@unc.edu

1. Introduction

As one of the most readily available epidemiologic variables across space and time, mortality offers a useful health indicator. Due to the inherent variability of death records, it is often necessary to use random field models to represent mortality rate distributions and the associated environmental exposures [Christakos and Serre, 2000]. In this work, mortality rate is represented as a spatiotemporal random field and rate estimates across space and time are derived using the Bayesian Maximum Entropy (BME) approach [Christakos, 1990; 2000; Christakos *et al.*, 2001]. We then demonstrate the effectiveness of the BME approach by analyzing mortality data from 58 California counties (USA) and obtaining meaningful spatiotemporal maps using uncertain mortality data of the interval type.

2. The BME Method

Let $X(\boldsymbol{p})$, $\boldsymbol{p}=(\boldsymbol{s},t)$, denote a spatiotemporal random field (S/TRF) varying within a space/time domain. The BME approach provides a rigorous framework to process various forms of data (hard and soft), and yields estimates $\boldsymbol{c}_k$ of the field value $\boldsymbol{c}_k$ at any point $\boldsymbol{p}_k$. The BME approach consists of three stages, as follows:

Prior stage: The general knowledge base G about the values $c_{\underline{\text{map}}}$ of the random field at the mapping points $\boldsymbol{p}_{\text{map}}$ is expressed as: $\int d\boldsymbol{c}_{\text{map}} g_{\boldsymbol{a}}(\boldsymbol{c}_{\text{map}}) f_G(\boldsymbol{c}_{\text{map}}) = \bar{h}_{\boldsymbol{a}}(\boldsymbol{p}_{\text{map}})$, $\boldsymbol{a}=0,1,\ldots,N_c$, where f_G is the prior pdf, and the $\bar{h}_{\boldsymbol{a}}$ is known for each function $g_{\boldsymbol{a}}$.

Meta-prior stage: The specificatory knowledge base S includes exact measurements c_{hard} and a variety of soft data $c_{\text{soft}}(S)$,such as interval values and probabilistic functions. Then, $c_{\text{map}}=(c_{\text{hard}}, c_{\text{soft}}, \boldsymbol{c}_k)$.

Posterior (or integration) stage: On the basis of the total knowledge base $K=G\cup S$ the posterior pdf is derived $f_G(\boldsymbol{c}_k)=B\int_{\boldsymbol{D}} d\Xi(c_{\text{soft}})\, f_G(c_{\text{map}})$ at the estimation point, where the constant B, integration domain $\boldsymbol{D}$ and operator Ξ assume various forms depending on the available data. Vectorial BME representations are, also, studied in Choi *et al.* (1998).

3. The Mortality Estimates in California

Using the daily death data collected by the Statistics Health Department during 1989 for all 58 California counties, we extracted only death records having an ICD (international classification of death) code below 800 (corresponding to non-accidental death according to the WHO 9th revision), and only for residents of California. The resulting total number of records selected for our study was 219,182. Using the death count d(i,t) for each county i (i=1,...,58) and day of the year, t=1,...,365, we obtained an interval range for the mortality rate in county i and day t as follow, [d(i,t)-1,d(i,t)+1]/population(i), where population(i) denotes the population of county i (Fig 1.). The spatiotemporal mean trend of the mortality rate was filtered out, thus leading to a S/TRF of residual mortality rate, $X(s,t)$, which is spatially homogeneous/temporally stationary. The corresponding covariance was modelled in terms of 3 nested exponential functions. Using the BMElib software [Christakos *et al.*, 2001] with a general knowledge G consisting of the covariance model above and specificatory knowledge S consisting of the soft interval data also described above, we obtained space/time maps of the mortality rate. For illustration, the mortality map shown in Fig. 2 was obtained for January 1, 1989 using interval soft data for that day as well as adjacent days.

File Name : isiFig1.eps
Creator : MATLAB, The Mathworks, Inc.
Pages : 1

Fig. 1: Mortality rate (deaths per 10^5 people per day) for county 1 (top) and county 3 (bottom).

Fig. 2: S/T estimate of mortality rate (deaths per 10^5 people per day) for January 1, 1989.

4. Conclusion

S/TRF theory provides a powerful conceptual framework to represent the distribution of mortality rate in space and time. Death counts are uncertain observables at the county scale which are rigorously processed by the BME method in an integrated space/time domain. Hence, the resulting maps are more realistic and informative than those obtained by traditional methods which do not account for composite space/time variations, soft data, etc.

REFERENCES

Choi, K. M., G. Christakos, and M.L. Serre, 1998. "Recent developments in vectorial and multi-point BME analysis", *Proc. 4th Conf. Int. Assoc. Math. Geology*, 91-96, Naples, Italy.

Christakos, G., 1990. "A Bayesian/maximum-entropy view to the spatial estimation problem". *Mathematical Geology*, 22(7), 763-776.

Christakos, G., 2000. Modern Spatiotemporal Geostatistics: Oxford Un. Pr, New York, NY.

Christakos, G. and M.L. Serre, 2000. "A spatiotemporal study of environmental exposure-health effect associations", *Jour. of Exposure Analysis & Env. Epidemiology*, 10(2), 168-187.

Christakos, G., P. Bogaert and M.L. Serre, 2001, Advanced Functions of Temporal GIS: Springer, New York, N.Y. [CD Rom included].

RESUME

Kyung-mee Choi, Ph.D. (expected May, 2001).

Marc Serre, Research Assistant Professor.

George Christakos, Professor.

All co-authors are from Dept. of Environmental Science and Engineering, School of Public Health, University of North Carolina at Chapel Hill, NC, USA