HAL
archives-ouvertes.fr

Modelling the Extremely Low Frequencies Magnetic Fields Times Series Exposure by Segmentation

Hanany Tolba, Laurent Le Brusquet, Marta Parazzini, Serena Fiocchi, Emma Chiaramello, Paolo Ravazzani, Martin Röösli, Isabelle Magne, Martine Souques

► **To cite this version:**

Hanany Tolba, Laurent Le Brusquet, Marta Parazzini, Serena Fiocchi, Emma Chiaramello, et al.. Modelling the Extremely Low Frequencies Magnetic Fields Times Series Exposure by Segmentation. 11ème Congrès National de Radioprotection, Jun 2017, Lille, France. hal-01475042

HAL Id: hal-01475042
https://hal.archives-ouvertes.fr/hal-01475042

Submitted on 23 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Modelling the Extremely Low Frequencies Magnetic Fields Times Series Exposure by Segmentation

Tolba H[1], Le Brusquet L[1], Parazzini M[2], Fiocchi S[2], Chiaramello E[2], Ravazzani P[2], Röösli M[3,4], Magne I[5] and Souques M[5]

[1]Laboratoire des Signaux et Systèmes (UMR CentraleSupélec, CNRS, Université Paris-Sud), Gif-sur-Yvette, France
[2]CNR Consiglio Nazionale delle Ricerche, Istituto di Elettronica e di Ingegneria dell'Informazione e delle Telecomunicazioni IEIIT, Milano, Italy
[3]Epidemiology and Public Health, Swiss Tropical and Public Health Institute, Basel, Switzerland
[4]University of Basel, Basel, Switzerland
[5]EDF, Electricite de France, France

Introduction

ELFSTAT project, founded by the French ANSES (2015-2019, Grant agreement n. 2015/1/202), aims at characterizing children's exposure to Extremely Low Frequency Magnetic Fields (ELF-MF) in real exposure scenarios using stochastic approaches.
The present paper gives details about the first step of the project: this step aims at developing stochastic models to model personal exposure from a dataset of recorded ELF-MF signals. These recordings are coming from the ARIMMORA project [1]: 331 children has worn an EMDEX and their exposition were measured during about 3 days. The stochastic models will then be used to construct realistic simulations of ELF-MF time series.
Figure 1 gives an example of ELF-MF 24 hours-signal. The majority of ELF-MF time series are characterized by abrupt changes in structure, such as sudden jumps in mean level or in dispersion around a mean level. The developed model consists in detecting these changes and in modeling the signal between two consecutive changes by a stationary process. These stationary processes are described by parametric models. We chose Auto-Regressive model because of the possibility of characterizing mean effects, variance effects and time-correlation effects with a weak number of parameters. The full-model is obtained by modeling the distribution of the parameters from the whole dataset of 331 recordings. Figure 2 gives a schematic of the full approach.

Data sources

The 331 individuals of the database are distributed according the geographical location (Switzerland or Italy) and the time season (Winter or Summer), as shown in the following Table:

	Switzerland population	Italy population
Winter	79 individuals	86 individuals
Summer	80 individuals	86 individuals

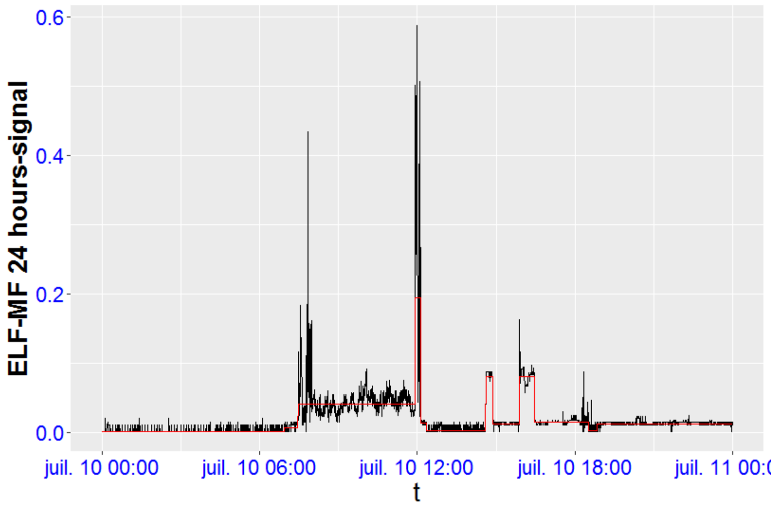

Figure 1: Example of ELF-MF recording (black points) and its obtained segmentation (in red).

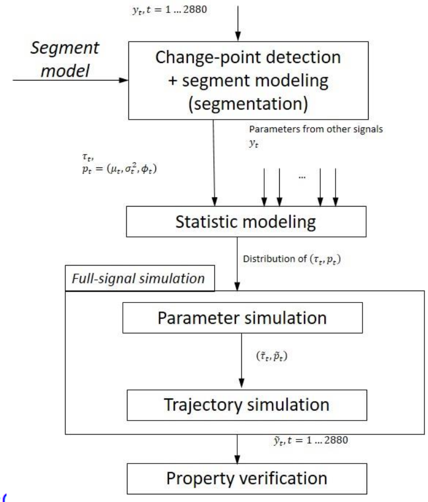

Figure 2: Schematic of the full approach

For each individual, we consider a full day of measurements (0h-24h). As the interval between two successive observations is 30 seconds, 2880 values per recording are considered.

Change-point detection and segment modeling

We consider changepoints (or equivalently breakpoints) to be the time points that divide a data set into distinct homogeneous segments (or equivalently stationary segments); the boundaries of segments can be interpreted as changes in the physical system. The problem of changepoints estimation has attracted significant attention and is required in a number of applications including financial data [2], climate data [3], biomedical data [4] and signal processing [5]. Different authors propose various approaches to the problem of changepoints detection or segmentation of time series. This issue is thoroughly surveyed in [6], where different methods are proposed and an exhaustive list of references is given. Some authors study the estimation of a single changepoint problem [7], while others extend it to multiple changepoints problem. In this last case, the number of changepoints is unknown and it is a challenge to jointly estimate the number of changepoints, their location, and also provide an estimation of the model representing each segment. In this work we consider the [2] procedure because of its ability to jointly estimate the number of changepoints, their location and the AR model parameter of each segment. To resume, the segmentation aims to:
(1) find the periods of stability and homogeneity in the behaviour of the time series;
(2) identify the locations of change, called changepoints;
(3) represent the regularities and features of each segment (estimate the model of each segment by a parameter set like changepoints location, segment amplitude, segment duration, segment regularity).
Figure 1 gives the segmentation results for the given example.

Statistical characterization of model parameters

In order to produce a completely model of the time series, the marginal and joint probability distributions of all model parameters are required. The set of parameters related to piecewise autoregressive model are, at time t: duration τ_t and coefficient of the AR model (amplitude or mean μ_t, noise variance ${\sigma_t}^2$ and the autoregressive coefficients ϕ_t).
These 4 parameters can be modeled separately (whether by a parametric probability distribution or not) as well as jointly (rather in a nonparametric way because of the difficulty to find a multidimensional probability distribution that efficiently models the parameters set).
Figure 3 shows that parameters are correlated and thus cannot be modeled separately. We used the Multivariate Kernel Density Estimation [9] that is a nonparametric and classical method. The advantage is that once we have an estimate of this density, we can simulate realistic realizations of these parameters. The figure 3 also gives the Multivariate KDE contours

and 5000 simulations of these parameters via the Multivariate KDE. Concerning the fitted distributions, first result shows no significant difference between the summer database and the winter database.
Finally, in the figure 4, we present an example simulated of ELF-MF exposure time series based in piecewise AR model.

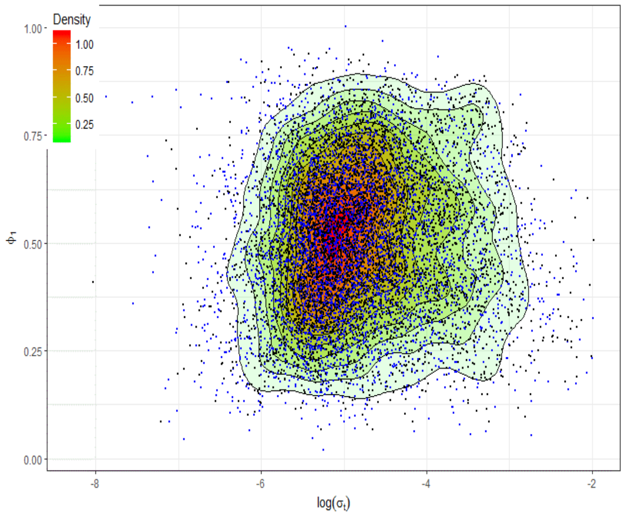

Figure 3: Scatter plot of calculated model parameters (ϕ_t, σ_t) (black points), the confidence areas of the fitted distribution and simulated points (blue points).

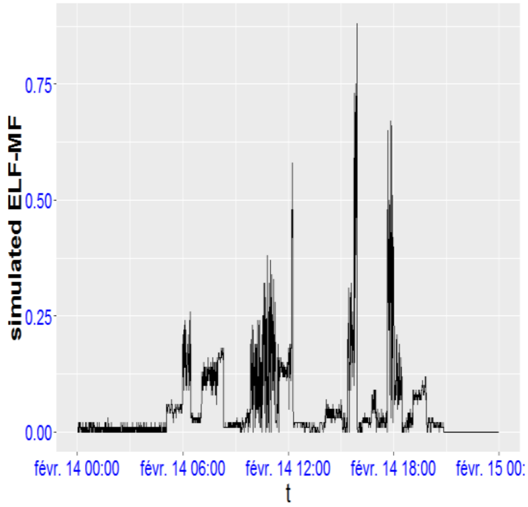

Figure 4: Example of simulated ELF-MF time series.

Aknowledgement
The ELFSTAT Project is supported by The French National Program for Environmental and Occupational Health of Anses (2015/1/202). The French data come from the EXPERS study database, subsidized by the French Ministry of Health, EDF and RTE, and carried out by Supélec, EDF and RTE.

References
[1] Struchen B, Liorni I, Parazzini M, Gaengler S, Ravazzani P, Röösli M. 2015. Analysis of children's personal and bedroom exposure to ELF-MF in Italy and Switzerland. J Expo Sci Environ Epidemiol. doi:10.1038/jes.2015.80.
[2] Killick, R., Fearnhead, P., & Eckley, I. A. (2012). Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, *107*(500), 1590-1598.
[3] Davis, R. A., Lee, T. C. M., & Rodriguez-Yam, G. A. (2006). Structural break estimation for nonstationary time series models. *Journal of the American Statistical Association*, *101*(473), 223-239.
[4] Fryzlewicz, P. (2014). Wild binary segmentation for multiple change-point detection. *The Annals of Statistics*, *42*(6), 2243-2281.
[5] Korkas, K., & Fryzlewicz, P. (2014). Multiple change-point detection for non-stationary time series using Wild Binary Segmentation. *Preprint*.
[6] Basseville, M., & Nikiforov, I. V. (1993). *Detection of abrupt changes: theory and application* (Vol. 104). Englewood Cliffs: Prentice Hall.
[7] Davis, R. A., Huang, D., & Yao, Y. C. (1995). Testing for a change in the parameter values and order of an autoregressive model. *The Annals of Statistics*, 282-304.
[8] Lavielle, M. (2005). Using penalized contrasts for the change-point problem. *Signal processing*, *85*(8), 1501-1510.
[9] Scott, D. W. (2015). *Multivariate density estimation: theory, practice, and visualization*. John Wiley & Sons.