Modelling the exposure of children to extremely low frequency magnetic fields by segmentation

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

Modelling the exposure of children to extremely low frequency magnetic fields by segmentation

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

[1]*Istituto di Elettronica e di Ingegneria dell'Informazione e delle Telecomunicazioni IEIIT CNR, Milano, Italy*
[2]*Département Signal et Statistiques Centrale Supélec Paris France*
[3]*EDF Electricite de France, Paris France*
[4]*Epidemiology and Public Health, Swiss Tropical and Public Health Institute, Basel, Switzerland*

***Abstract*—In this paper, the exposure of children to extremely low frequency magnetic field (ELF-MF, 40-800Hz) is investigated. In particular, a stochastic approach for extracting information from time series registrations is used. The aim is to obtain a better characterization and description of the phenomenon and to investigate on possible correlations or different features between various data subsets.**

***Keywords*— Children exposure, ELF-MF, stochastic model, Kernel Density Estimation.**

I. INTRODUCTION

CHILDREN exposure to extremely low frequency magnetic fields (ELF-MF, 40-800 Hz) is a topic of high interest, due to the possible correlation between the exposure and the onset of children leukaemia [1] [2]. Measurement campaigns have been therefore carried out to investigate on the real level of children exposure in real life, particularly in Europe [3] [4] [5].
The aim of this paper is to characterize the children's exposure to ELF-MF in real exposure scenarios using a stochastic approach based on segmentation starting from personal measurements of children exposure [6].

II. MATERIALS AND METHODS

In this study, two databases of ELF magnetic field children personal exposure measurements have been considered and summed together: the first is the dataset of the ARIMMORA project [3] [4], which includes a total of 331 children registrations in Italy and Switzerland during winter and summer seasons and the second one is the dataset of the EXPERS project [5], for a total of 977 children registrations in France during cold season. For each subject, full days of personal measurements were extracted (from 00:00 to 24:00), which means 2880 values per recording (the time step is 30 s). From the two datasets, the complete days of recorded registrations are 682 for ARIMMORA and 767 for EXPERS, for a total of 1449 full days. In Fig. 1 an example of 24 h personal measurement is shown.

It can be notice that ELF-MF time series were characterized by abrupt changes, like sudden jumps of the value of the mean or in dispersion around the mean. Therefore, the first step of our approach consisted in detecting these change-points and dividing the 24 h personal measurements in segments between two consecutive jumps. There are different approaches to detect change-points, like Binary Segmentation [7], Circular Binary Segmentation [8], Wild Binary Segmentation [9] and Pruned Exact Linear Time method (PELT) [10] [11] [12]. In this work, the PELT method was used, due its ability to estimate the number of change-points, their locations and the Auto-Regressive (AR) model parameter of each segment. After the process of segmentation, the number of change-points detected for the whole dataset was of 11510 points in total.

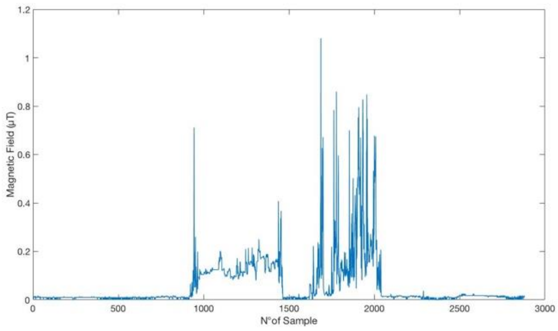

Fig. 1: An example of 24 hours ELF-magnetic field personal measurement.

The second step of our approach consisted in modeling every obtained segment between 2 jumps with a stationary process, in particular with an AR model. At the end, every segment was described by 4 parameters: i) the mean of the stationary process (μ); ii) the variance of the stationary process (σ^2); iii) the Auto-Regressive coefficient of the AR process of order 1 (φ); iv) the duration of the stationary process (T). Every segment is so identified by $\mathrm{p} = (\mu, \sigma^2, \varphi, \mathrm{T})$ and the whole dataset includes $(\mathrm{p}_1, \ldots, \mathrm{p}_{11510})$.

To carry out a first analysis, the whole dataset was divided in two different groups, the first one includes the points regarding segments during the day time (from 07:00 to 21:00 h) and the second one the points regarding the night time (from 21:01 to 06:59 h).

In order to compare the two groups and to extract information, the Kernel Density Estimation (KDE) of the 4 parameters was estimated. Kernel density estimation is a nonparametric technique for probability density function estimation [13]. Let $f(\boldsymbol{p})$ the unknown probability density function, the KDE of $f(\boldsymbol{p})$ is given by (1):

$$\hat{f}(\boldsymbol{p}) = \frac{1}{n}\sum_{i=1}^{n} K_h\,(\boldsymbol{p} - \boldsymbol{p}_i) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{(\boldsymbol{p}-\boldsymbol{p}_i)}{h}\right) \quad (1)$$

where n is the number of data of the groups, K is the kernel function, which in this case is a Gaussian kernel and h is the bandwidth of the kernel. In this case h was calculated from the empirical rule given in [14].

This study reported the first results of the kernel density estimations of the 4 parameters divided in day and night segments.

III. Results

The preliminary results on the analysis of the union of the two datasets (ARIMMORA and EXPERS) are shown in the following figure.

Fig. 2 represents, as example, the evaluated kernel density estimation of the 4 variables $(\mu, \sigma^2, \varphi, \mathrm{T})$ for the segments regarding the day time compared with the segments regarding the night time.

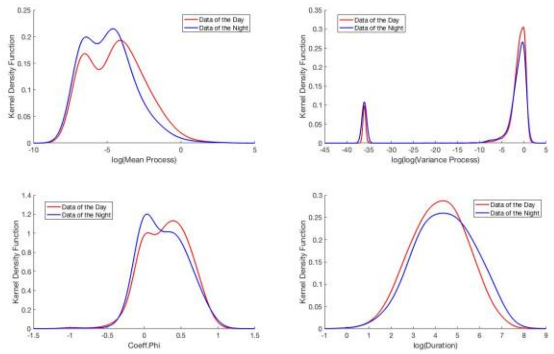

Fig. 2: Estimated kernel density function of the four variables of the segments of the day (in red) comparing with estimated kernel density function of the four variables of the segments of the night (in blue), from above to down, left to right: mean μ, variance σ^2, coefficient φ and duration T.

IV. Conclusion

The aim of this paper was to obtain a better description and a more accurate analysis about the exposure of children to ELF magnetic field. Preliminary analysis seems to show that the mean value (μ) of the exposure to ELF-MF is higher during the day than during the night. Moreover, the segments during the night show a longer duration (T) and less variance (σ^2) compared with the segments of the day. This resulted also in differences in KDE of the coefficient φ of the two groups. A more systematic search of the features that are the most influent will be conducted. This will be performed splitting the dataset into several sub-datasets according to a feature: the features that will lead to the biggest difference between models will be considered as the most discriminant.

Acknowledgement

This research is supported by the French National Research Program for Environmental and Occupational Health of ANSES (2015/1/202): Project ELFSTAT - In depth evaluation of children's exposure to ELF (40 – 800 Hz) magnetic fields and implications for health risk of new technologies, 2015-2019. 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] A. Ahlbom, N. Day, and M.Feytching, "A pooled analysis of magnetic fields and childhood leukemia", *British Journal of Cancer*, vol. 83(5), pp. 692-698, 2000.

[2] L. Kheifets, A. Ahlbom, C.M. Crespi et al., "Pooled analysis on recent studies on magnetic fields and childhood leukemia", *British Journal of Cancer*, vol. 103(7), pp. 1128-1135, 2010.

[3] I. Liorni, M. Parazzini, B. Struchen, S. Fiocchi, M. Röösli, and P. Ravazzani, "Children's Personal Exposure Measurements to Extremely Low Frequency Magnetic Fields in Italy", *International journal of environmental research and public health*, vol. *13*(6), pp. 549, 2016.

[4] B. Struchen, I. Liorni, M. Parazzini, S. Gängler, P. Ravazzani, and M. Röösli, "Analysis of personal and bedroom exposure to ELF-MFs in children in Italy and Switzerland", *Journal of Exposure Science and Environmental Epidemiology*, vol. *26*(6), pp. 586-596, 2016.

[5] I. Magne, M. Souques, I. Bureau, A. Duburcq, and E. Remy, J. Lambrozo "Exposure of children to extremely low frequency magnetic fields in France: Results of the EXPERS study." JESEE, 2017, 27 (5) : 505-512.

[6] H. Tolba, L. Le Brusquet, M. Parazzini, S. Fiocchi, E. Chiaramello, P. Ravazzani, M. Röösli, I. Magne, and M. Souques. "Modelling the Extremely Low Frequencies Magnetic Fields Times Series Exposure by Segmentation." In *11ème Congrès National de Radioprotection*, 2017.

[7] A. J. Scott, M. Knott, "A cluster analysis method for grouping means in the analysis of variance", *Biometrics*, pp.507-512, 1974.

[8] A.B. Olshen, E. Venkatraman, R. Lucito, and M. Wigler, "Circular binary segmentation for the analysis of array-based dna copy number data", *Biostatistics*, vol. 5, pp. 557-572, 2004.

[9] P. Fryzlewicz et al. "Wild binary segmentation for multiple change-point detection", *The Annals of Statistics*, vol. 42, pp. 2243-2281, 2014.

[10] R. Killick, P. Fearnhead, I. Eckley, " Optimal detection of changepoints with a linear computational cost", *Journal of the Americn Statistical Association*, vol. 107, pp. 1590-1598, 2012.

[11] B. Jackson, J. D. Scargle, D. Barnes, S. Arabhi, A. Alt, P. Gioumousis, E. Gwin, P. Sangtrakulcharoen, L. Tan, and T. T. Tsai, "An algorithm for optimal partitioning of data on an interval", *IEEE Signal Processing Letters*, vol. 12, pp. 105-108, 2005.

[12] R. Killick, I. Eckley, "Changepoint: An r package for changepoint analysis", *Journal of Statistical Software*, vol. 58, pp. 1-19, 2014.

[13] D.W. Scott, S.R. Sain. "9-Multidimensional Density Estimation", Handbook of statistics 24, pp. 229-261, 2005

[14] D.W. Scott, "Multivariate density estimation: theory, practice and visualization", John Wiley & Sons, 2015