URSI GASS 2020, Rome, Italy, 29 August - 5 September 2020

Stochastic Dosimetry and Machine Learning: Innovative Approaches for Facing Challenges in Exposure Assessment in Realistic Scenarios

Emma Chiaramello(1), Gabriella Tognola*(1), Marta Bonato(1,2), Silvia Gallucci(1), Isabelle Magne(3), Martine Souques(3), Serena Fiocchi(1), Marta Parazzini(1), and Paolo Ravazzani(1)
(1) Istituto di Elettronica e di Ingegneria dell'Informazione e delle Telecomunicazioni, CNR, Milano, Italy
(2) DEIB, Politecnico di Milano, Italy, 20133
(3) EDF Electricité de France, Paris, France, 75017

Abstract

Innovative approaches, such as stochastic dosimetry and Machine Learning, can be complementary to traditional methods for electromagnetic field (EMF) exposure assessment, overcoming limitations and allowing extraction of new/deeper information. In this study, two examples of innovative EMF exposure assessment approaches are presented: (i) a stochastic approach based on low rank tensor approximations to assess indoor exposure to WLAN access point with unknown location and (ii) an application of Machine Learning to characterize indoor residential exposures to ELF magnetic field in children by considering the type of electric networks near the child home, the age and type of the child home, the type of heating and the family size.

1 Introduction

Appropriate health risk assessment to identify possible risks correlated to the exposure to Electromagnetic fields (EMF) both at extremely low frequencies (ELF, 50/60 Hz) and radio frequencies (RF) could not exist without a proper evaluation of the level of exposure in realistic exposure scenarios.

Previous works have shown that EMF human exposure depends on a great degree on several of parameters (e.g., morphology, anatomy, and postures of human bodies, tissue dielectric properties, locations and number of the sources, frequency bands, and many others) and that each of these parameters is intrinsically affected by variability and uncertainty [1]. The characterization of such complex scenarios would be impossible using standard approaches.

New techniques are therefore required both to understand which features influence the exposure scenarios, identifying recurrent and common patterns of exposure, and to characterize the uncertainty and variability of the real exposure levels. In this study we describe two recently proposed approaches based on stochastic dosimetry [1-2] and Machine Learning [3-4] that are used to extract deeper information and overcome limitations of traditional exposure assessment methods.

As an example of application of stochastic dosimetry, we present low rank tensor approximation to estimate indoor children exposure to WLAN source. As an example of Machine Learning for the assessment of human exposure to EMF, we present the identification of recurrent patterns of indoor magnetic field exposure in children by means of *K*-means clustering, an unsupervised Machine Learning approach.

2 Materials and Methods

2.1 Stochastic dosimetry

Stochastic dosimetry uses advanced statistics to build surrogate models able to estimate the distribution of the EMF exposure quantities of interest, considering the variability of the exposure scenarios. The problem of EMF exposure can be represented as a model M, such as:

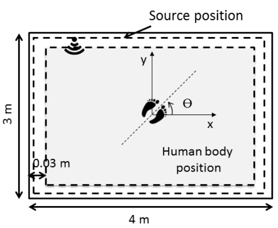
$$Y = M(X) \tag{1}$$

where $X = \{X_1, \ldots, X_M\}$ denote the M-dimensional input vector representing the parameters that influence the exposure scenario and Y denote the EMF exposure quantity of interest (e.g., the Specific Absorption Rate SAR). A surrogate model is an analytical function that shows similar statistical properties with Y, with significantly lower computational cost than the original computational model M [5]. Among all the different approaches that can be used to build surrogate models, the non-intrusive ones, i.e. approaches in which the phenomenon to approximate is seen as a "black box", were found to be the more suitable to be used in stochastic dosimetry framework (see, e.g. [6-7]). A general stochastic dosimetry approach follows three steps. First, a probabilistic model of input parameters has to defined, i.e. the input parameters X are modelled by a random vector. Second, an experimental design, i.e. a set of realizations of the input vector $X_{exp} = \{X^{(1)}, \ldots, X^{(N)}\}$ has to be randomly chosen, and the corresponding model evaluations $Y_{exp} = \{M(X^{(1)}), \ldots, M(X^{(N)})\}$ has to be estimated by computational electromagnetic methods. The third step focused on the development of the surrogate model using a proper statistical method. Considering the need of taking into account a high number of input parameters, i.e. those parameters that influence the exposure scenario, and the need of reducing as much as possible the size of the experimental design to minimize the computational cost, one of the methods found to be suitable for stochastic dosimetry is the Low Rank Tensor Approximations (LRA) approach [8]. Details about the LRA method and an example of application are following.

2.2 Low Rank Tensor Approximation to estimate indoor children exposure to WLAN source

Fig. 1 shows a schematic view of the exposure scenarios: the Whole-Body SAR (WB SAR) induced by a WLAN source (operating at 2.4 GHz) has been assessed in child tissues when varying the position of the source on the wall and the position of the child in a 3x4 m^2 room by using surrogate models based on Low Rank Tensor Approximation (LRA). The position of the source was described by two coordinates, i.e., the horizontal location L of the source and its height *z*, while the child position was described by three coordinates, i.e. the position on the floor, defined by the coordinates *x* and *y*, and the angle of rotation θ along the vertical axis. All the input parameters X_i were supposed to be uniformly distributed with ranges of variability reported in fig. 1.

350 values of WB SAR corresponding to 350 different positions of the source and the child were evaluated: 300 values were used to build the surrogate LRA-models, while 50 values were used to validate them. Because of the possible proximity of the WLAN source to the child, the field emitted could not be assessed with the plane wave hypothesis, thus a procedure based on the combined use of Spherical Wave Expansion and Finite-Difference-Time-Domain method was applied [9]. The simulations were carried out using a high resolution 8-years child model "Eartha" from the Virtual Classroom [10]; the dielectric properties in child tissues were assigned according to the data available in literature [11]. The input power of the source was set equal to 100 mW.



Input parameters	Range of variation
L	[0, 13.59] m
z	[0.25, 2] m
x	[0.3, 3.7] m
y	[0.3, 2.7] m
θ	[0°, 359°]

Figure 1. Schematic view of the exposure scenario.

LRA [8] is a non-intrusive method for developing surrogate models as a finite sum of rank-one functions:

$$Y_{LRA} = \sum_{l=1}^{R} b_l w_l = \sum_{l=1}^{R} b_l \left(\prod_{i=1}^{M} v_l^{(i)}(X_i) \right) \qquad (2)$$

where w_l is the *l*-th rank-one function obtained as product of univariate functions of the components of X_i, v_l (i) denotes a univariate function of the components of X_i in the *l*-th rank-one component, *M* is the number of input variables, b_l ($l = 1,\ldots, R$) are scalars that can be viewed as normalizing constants and *R* is the rank of the decomposition.

In order to estimate the unknown parameters, i.e. the polynomial coefficients $z_{k,l}^{(i)}$ and the normalizing coefficients b_l ($l = 1,\ldots, R$) of the surrogate model, a greedy algorithm, based on Alternated Least-Squares (ALS) minimization [8]. The employed algorithm involves a sequence of pairs of "correction step" and "updating step". In the *l*-th "correction step", the rank-one tensor w_l is built, while in the *l*-th "updating step" the set of normalizing coefficients $\{b_1,\ldots, b_l\}$ is determined.

Once all the LRA-models have been built, they were exploited to obtain a "room exposure assessment" selecting 100,000 possible positions of source and child, and using the LRA models to assess the corresponding SAR values. A statistical analysis has been performed to assess the variability of the exposure due to the position of the WLAN source and the child, in terms of Quartile Coefficient of Dispersion, calculated as QCD = $(Q_3-Q_1)/(Q_3+Q_1)$, where Q_1 and Q_3 are, respectively the first and the third percentiles of the distribution.

2.3 Clustering analysis

For the analysis of ELF residential exposure, we applied the *K*-means clustering algorithm (Matlab, ver. R2018a, MatWorks Inc., Natick MA, USA) [12]. In particular, we used *K*-means clustering to identify similar patterns of exposure (i.e., to identify the exposure clusters) based on the type and number of electric networks close to the child home. Cluster analysis is used here to perform a so-called "exploratory analysis", that is to discover possible hidden patterns in the observed data. By using a Euclidean distance metric, the method calculates the similarities/dissimilarities between the observed data and generates a number of mutually-exclusive partitions – the clusters – that contain data that share more similarities among them if compared with the observation assigned to the other partitions.

We then studied the possible effect of house heating, residence age, residence type and family size on the clusters identified with *K*-means. To this purpose we applied an association effect analysis performed with the Chi-square (χ^2) association test and the Kruskal-Wallis test. For all tests, we set the significance level at $p<0.05$. More details on the procedure we used to perform clustering and the association analysis can be found in [3-4].

2.4 ELF residential exposure – the dataset

For the characterization of residential exposure to ELF MF, we analyzed the magnetic flux density B (50 Hz component) recorded during 24h from 884 children in

France. Measurements were performed with a personal exposimeter. The analyzed dataset comes from the EXPERS study [13-14]. For each child, the dataset reports the number of electric lines near the child home according to the following classes: 400 V underground cables ≤40 m; 20 kV underground cables ≤40 m; 63-150 kV underground cables ≤20 m; 225 kV underground cables ≤20 m; 400 V overhead lines ≤40 m; 20 kV overhead lines ≤40 m; 150 or 90 or 63 kV overhead lines ≤100 m (for 150 kV lines) or 70 m (for 63 and 90 kV lines); 225 kV overhead lines ≤120 m; 400 kV overhead lines ≤200 m; 20 kV/400 V substations inside the building ≤40 m (Substation). For each child, the dataset also reports the age and type of the house, the type of house heating, and the family size.

3 Results

3.1 Low Rank Tensor Approximation to estimate indoor children exposure to WLAN source

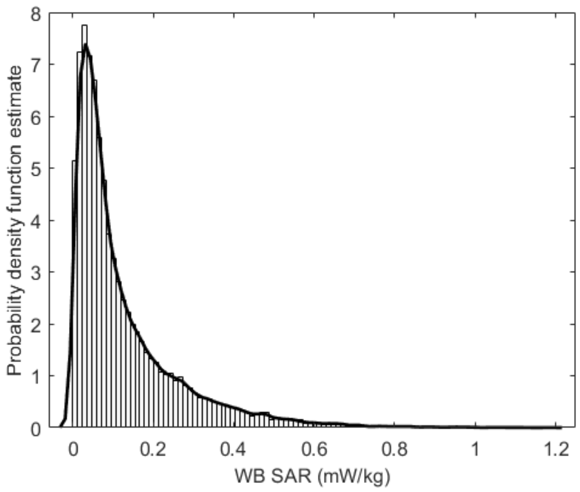


Figure 2. Histograms and probability density function of the WB SAR values for 100.000 random positions of the WLAN source and the child.

Fig. 2 shows histogram and probability density function of the WB SAR values obtained in 100,000 random positions of the WLAN source and the child by the LRA model. The histogram shows a positive-skewed shape, thus indicating that for most of the 100,000 evaluated positions of the source and the child the levels of exposure were lower that the median value of the distribution. WB SAR values showed mean, median and max values equal to 0.13 mW/kg, 0.07 mW/kg, and 1.40 mW/kg, respectively. High variability in WB SAR has been found as a function of the relative positions between the WLAN source and the child, resulting in QCD value equal to 65%. The probability density function could be approximated by a Gamma distribution with parameters a = 1.04, and b = 0.12, (with $R^2 = 0.97$).

3.3 ELF MF exposure

As to the characterization of residential ELF MF exposure, clustering identified three recurrent patterns: one of children with the highest exposure (average: 0.126 µT, 1st quartile Q1: 0.045 µT; 3rd quartile Q3: 0.225 µT) who lived close to overhead lines of high (63-150 kV), extra-high (225 kV) and ultra-high voltage (400 kV); another of children with mid exposure levels (average: 0.036 µT, Q1: 0.010 µT; Q3: 0.050 µT) who lived close to underground networks of low (400V) and mid voltage (20 kV) and substations (20kV/400V); and the last one of children with the lowest level of exposure (average: 0.025 µT, Q1: 0; Q3: 0.020 µT) who lived far from electric networks. Cluster analysis also revealed that underground lines of high (63-150 kV) and extra-high voltage (225 kV) and overhead lines of low voltage (400 V) were not relevant in differentiating the exposure patterns.

The association analysis revealed no statistical significant association of the residence age with the exposure patterns; instead, the type of residence (individual vs. apartments in bug buildings), the type of heating (electric vs. mixed and non-electric), and the family size were significantly associated to the exposure patterns. Specifically, children with mid exposure levels typically lived in houses with electric heating appliances and in big buildings with more than 10 residential units; vice versa, children with the lowest exposure levels typically lived in homes with mixed or non-electric heating appliances and in individual or terraced houses. Finally, we found that children with mid exposure levels lived in families of bigger size than those with the lowest exposures (Dunn's post-hoc test, $p<0.005$).

4 Conclusions

We presented a number of applications of advanced statistics (stochastic dosimetry) and Machine Learning approaches for the assessment of exposure to EMF. Stochastic dosimetry based on LRA gave a complete description of the exposure scenario, without limiting the assessment to few "worst case" conditions. Although not eliminating the need of simulations, the LRA approach used about 0.35% of the time that would be needed using only computational methods. As to Machine Learning, cluster analysis was useful to perform exploratory analysis on exposure data and infer new and deeper knowledge on the variables that are more important on the exposure patterns.

5 Acknowledgements

The study on WLAN exposure assessment was supported by The French National Research Program for Environmental and Occupational Health of ANSES (EST-2016-2RF-04) Project AMPERE: Advanced MaPpingof residential ExposuRE to Rf- Emf sources.

The study on ELF MF exposure assessment was supported by Project ELFSTAT-In depth evaluation of children's exposure to ELF (40–800 Hz) magnetic fields and implications for health risk of new technologies (ANSES (2015/1/202)).
Data of ELF MF exposure come from the EXPERS study database, subsidized by the French Ministry of Health, EDF (Electricite De France) and RTE (Réseau de Transport d'Électricité), and carried out by Supélec, EDF and RTE.

6 References

1. J. Wiart, Radio-Frequency Human Exposure Assessment: From Deterministic to Stochastic Methods. John Wiley & Sons, ISTE, 2016.

2. E. Chiaramello, S. Fiocchi, M. Parazzini, P. Ravazzani, and J. Wiart, "Stochastic Dosimetry for Radio-Frequency Exposure Assessment in Realistic Scenarios". In *Uncertainty Modeling for Engineering Applications* (pp. 89-102). Springer, Cham, 2019.

3. G. Tognola, M. Bonato, E. Chiaramello, S. Fiocchi, I. Magne, M. Souques, M. Parazzini, and P. Ravazzani, "Use of Machine Learning in the analysis of indoor ELF MF exposure in children", *Int J Environ Res Public Health*, **16**, 2019, pp. 1230-1243.

4. G. Tognola, E. Chiaramello, M. Bonato, I. Magne, M. Souques, S. Fiocchi, M. Parazzini, P. Ravazzani, "Cluster analysis of residential personal exposure to ELF magnetic field in children: effect of environmental variables," *Int J Environ Res Public Health*, **16**, 2019, pp. 4363-4376.

5. B. Sudret, S. Marelli, and J. Wiart. "Surrogate models for uncertainty quantification: An overview." Antennas and Propagation (EUCAP), 11th European Conference on. IEEE, 2017.

6. E. Chiaramello, M. Parazzini, S. Fiocchi, P. Ravazzani, and J. Wiart, "Stochastic dosimetry based on low rank tensor approximations for the assessment of children exposure to WLAN source," *IEEE J. Electromagn. RF Microwaves Med. Biol.*, **1**, April 2019, pp. 131-137

7. E. Chiaramello, M. Parazzini, S. Fiocchi, M. Bonato, P. Ravazzani, and J. Wiart, "Children exposure to femtocell in indoor environments estimated by sparse low-rank tensor approximations," *Annals of Telecommunications,* **74(1-2)**, 2018, pp. 113-121.

8. K. Konakli and B. Sudret, "Polynomial meta-models with canonical low-rank approximations: Numerical insights and comparison to sparse polynomial chaos expansions," *J. Comput. Phys*., **321**, Sep 2016, pp. 1144–1169.

9. Y. Pinto and J. Wiart, "Statistical Analysis and Surrogate Modeling of Indoor Exposure Induced from a WLAN Source," *Antennas and Propagation (EUCAP), 2017 11th European Conference on*, 19-24 March 2017.

10. M.-C. Gosselin, E. Neufeld, H. Moser, E. Huber, S. Farcito, L. Gerber, M. Jedensjö, I. Hilber, F. Di Gennaro, B. Lloyd, E. Cherubini, D. Szczerba, W. Kainz, and N. Kuster, "Development of a new generation of high-resolution anatomical models for medical device evaluation: the Virtual Population 3.0," *Phys. Med. Biol*., **59**, 18, Sep 2014, pp. 5287–5303.

11. C. Gabriel, S. Gabriel, and E. Corthout, "The dielectric properties of biological tissues .1. Literature survey," *Phys. Med. Biol*., **41**, 11, November 1996, pp. 2231–2249.

12. L. Kaufman, PJ. Rousseeuw, "*Finding Groups in Data: An Introduction to Cluster Analysis*," John Wiley & Sons, Inc.: Hoboken (NJ), United States, 1990.

13. M. Bedja, I., Magne, M. Souques, J. Lambrozo, L. Le Brusquet, G. Fleury et al., "Methodology of a study on the French population exposure to 50 Hz magnetic fields," *Radiat Prot Dosimetry*, **142**, 2010, pp. 146-152.

14. I. Magne, M. Souques, I., Bureau, A. Duburcq, E. Remy, J. Lambrozo, "Exposure of children to extremely low frequency magnetic fields in France: Results of the EXPERS study," *J Expo Sci Environ Epidemiol*, **27**, 2017, pp. 505–512.