Environmental Research 126 (2013) 184–191

Contents lists available at ScienceDirect

Environmental Research

journal homepage: www.elsevier.com/locate/envres

Assessment of outdoor radiofrequency electromagnetic field exposure through hotspot localization using kriging-based sequential sampling ☆

Sam Aerts *, Dirk Deschrijver, Leen Verloock, Tom Dhaene, Luc Martens, Wout Joseph

Department of Information Technology, Ghent University/iMinds, Gaston Crommenlaan 8, Box 201, B-9050 Ghent, Belgium

ARTICLE INFO

Article history:
Received 21 February 2013
Received in revised form
19 April 2013
Accepted 13 May 2013
Available online 5 June 2013

Keywords:
Radiofrequency electromagnetic fields (RF-EMF)
Human exposure
Exposure model
Hotspots
Surrogate modeling

ABSTRACT

In this study, a novel methodology is proposed to create heat maps that accurately pinpoint the outdoor locations with elevated exposure to radiofrequency electromagnetic fields (RF-EMF) in an extensive urban region (or, hotspots), and that would allow local authorities and epidemiologists to efficiently assess the locations and spectral composition of these hotspots, while at the same time developing a global picture of the exposure in the area. Moreover, no prior knowledge about the presence of radiofrequency radiation sources (e.g., base station parameters) is required. After building a surrogate model from the available data using kriging, the proposed method makes use of an iterative sampling strategy that selects new measurement locations at spots which are deemed to contain the most valuable information—inside hotspots or in search of them—based on the prediction uncertainty of the model. The method was tested and validated in an urban subarea of Ghent, Belgium with a size of approximately 1 km^2. In total, 600 input and 50 validation measurements were performed using a broadband probe. Five hotspots were discovered and assessed, with maximum total electric-field strengths ranging from 1.3 to 3.1 V/m, satisfying the reference levels issued by the International Commission on Non-Ionizing Radiation Protection for exposure of the general public to RF-EMF. Spectrum analyzer measurements in these hotspots revealed five radiofrequency signals with a relevant contribution to the exposure. The radiofrequency radiation emitted by 900 MHz Global System for Mobile Communications (GSM) base stations was always dominant, with contributions ranging from 45% to 100%. Finally, validation of the subsequent surrogate models shows high prediction accuracy, with the final model featuring an average relative error of less than 2 dB (factor 1.26 in electric-field strength), a correlation coefficient of 0.7, and a specificity of 0.96.

1. Introduction

Public concerns about possible health effects due to the everyday exposure to radiofrequency electromagnetic fields (RF-EMF) are increasing. The general public is not familiar enough with the typical average and maximum levels of RF-EMF radiation they are exposed to in their everyday environment, although a number of studies have been performed on the matter using personal exposimeters (Bolte and Eikelboom, 2012; Frei et al., 2009a, 2010; Joseph et al., 2008, 2010; Rowley and Joyner, 2012; Röösli et al., 2010; Thuróczy et al., 2008; Viel et al., 2009a,b), e.g., by defining a large number of microenvironments, based on time of the day, activity and place, and assessing therein the average magnitude of the RF-EMF one is exposed to. Another, more visual way of filling the public information gap would be the use of a *heat map*, an easily comprehensible graphical representation of the magnitude of the RF-EMF exposure over an urban, suburban, or rural area. Naturally, heat maps can also be constructed using RF-EMF simulators, or from non-measurement-based models like the ones by Beekhuizen et al. (2013), Breckenkamp et al. (2008), Bürgi et al. (2008, 2010), Elliott et al. (2010), Frei et al. (2009b), and Neitzke et al. (2007), by calculating the exposure at arbitrary locations, probably using a uniform grid with a resolution of choice. However, these approaches are heavily dependent on accurate data, e.g., base station parameters, building coordinates, building heights, etc., data which is seldom readily available. Measurement-based models, as found in Aerts et al. (2013), Anglesio et al. (2001), Azpurua and Dos Ramos (2010), Isselmou et al. (2008),

☆This work was supported by the Interuniversity Attraction Poles Programme BESTCOM initiated by the Belgian Science Policy Office, and iMinds 'Green Wireless Efficient City Access Networks (GreenWeCan)' project, co-funded by iMinds (previously IBBT), a research institute founded by the Flemish Government in 2004, and the involved companies and institutions.

* Corresponding author. Fax: +32 9 331 48 99.
E-mail address: sam.aerts@intec.ugent.be (S. Aerts).

0013-9351/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.envres.2013.05.005

Paniagua et al. (2013), on the other hand, encompass an accuracy that depends both on the number as well as the specific locations of the measurements. The importance of the latter can be seen from the fact that, in order to attain a similar accuracy, an area featuring a rapidly changing field distribution requires a denser sampling than an area of the same size featuring a more evenly field distribution. Nonetheless, studies involving measurement-based models typically select all of their measurement locations at once, in a uniform or random grid, completely covering the area of the interest, hence disregarding the dependency of the accuracy of their model on the specific locations. The approach proposed by Aerts et al. (2013), however, tackles this problem by introducing an iterative sampling algorithm composed by Crombecq et al. (2011), in which repeatedly new batches of measurement spots are selected after analyzing the data from preceding measurements. In order to attain a globally accurate model that characterizes the overall exposure in the area, the algorithm selects measurement locations in such a way that they are both spread out as evenly as possible, as well as fine grained in those regions that are harder to interpolate (i.e., regions in which the electric field strength changes more rapidly). This methodology has also proven to be successful in various other research disciplines (Deschrijver et al., 2011, 2012). Experimental results in Aerts et al. (2013) confirm that the method is able to give an accurate prediction of the global RF-EMF exposure in the area. Using this approach, however, implies that several measurements must also be performed in regions with very low electric-field strength. In practice, however, such regions are often less interesting in terms of exposure assessment and risk evaluation, and in some cases, the electric-field strength is even immeasurable because of the measurement device's sensitivity.

In this article, a new urban RF-EMF exposure assessment method is proposed that focuses on the detection and characterization of regions with elevated or high RF-EMF exposure (*hotspots*), regions which are of particular interest for concerned citizens and epidemiologists, and applied to a large urban area. Moreover, accurate exposure measurements are performed in the identified hotspots to distinguish the contributions of various radiofrequency sources, and to check compliance with international exposure guidelines. Finally, our method is validated globally, to assess the overall prediction accuracy of the resulting model, as well as locally, to assess the hotspot prediction accuracy of the model and the performance of our method.

2. Methods

2.1. Area

The area of study, shown in Fig. 1, is urban with an approximate area of 1 km^2, comprising the city center of Ghent, Belgium. This area contains various schools (from kindergarten to University), multi-dwelling residences (mainly student dormitories), shops, restaurants, and other leisure spots (e.g., parks, concert venues, bars, etc.). There are multiple radiofrequency transmit antennas inside the area as well as close by, including frequency modulation (FM) radio, digital radio and television broadcasting (e.g., terrestrial digital audio broadcasting, or T-DAB), emergency service communication networks, mobile telephony base stations (Global System for Mobile Communications (GSM) at 900 and 1800 MHz, and Universal Mobile Telecommunications System (UMTS) at 2100 MHz), etc.

2.2. Sequential surrogate modeling

As in previous RF-EMF exposure assessment studies, the exposure metric of choice is the electric-field strength, E, in V/m (Volts per meter). Assuming we have no prior knowledge about the electric-field strength in the area under study, we can consider it as an unknown function of the location. An exact evaluation of this function is essentially impossible. However, it can be approximated by a so-called *surrogate model*, which can be defined as an approximation model for a computationally expensive simulation or a physical experiment, built from generally time-expensive samples at well-chosen locations (Crombecq et al., 2011). This choice of sample locations is called the *design of experiments* and is critical for the model's reliability. In measurement-based RF-EMF modeling studies, the design of experiments is usually a uniform or random grid, which is fixed before any measurements are performed. However, a design of experiments can also be built sequentially by founding the choice of new locations on previously chosen locations and their respective measurement outcomes, as in the study by Aerts et al. (2013), in which a surrogate model was built to accurately approximate the electric-field distribution from GSM base station radiation at 900 MHz in a small urban area, and subsequently used for analysis and mapping of the exposure. The advantage of sequential sampling consists both in performing relatively more measurements in regions that are potentially interesting (e.g., highly-varying electric-field strength in (Aerts et al., 2013), or hotspot regions where the electric-field strength is high or elevated, on which this study focuses), and in performing only as many measurements as is needed to obtain the desired accuracy.

Another reason for focusing on hotspots lies in the fact that *specificity* (a measure for correctly identifying what is unexposed) is highly important in exposure assessment. Because only a small percentage of people are exposed to higher levels of RF-EMF, one should make sure that those who are modeled to be exposed, are in fact exposed (Neubauer et al., 2007). Therefore, it is of importance for a measurement-based model that when hotspots are found, they are densely sampled and hence accurately modeled, and regions of high exposure are accurately delineated.

As this study focuses on hotspots, a different approach from the one presented in Aerts et al. (2013) is proposed. More specifically, we implemented a different search strategy for our sample locations. The iterative sampling method used in this study is based on the kriging surrogate modeling technique by Couckuyt et al. (2012). The use of kriging as interpolation technique has some distinct advantages. It takes into account the spatial structure of the interpolated variable (here, the electric-field strength), determines the best estimator of the variable (the error is minimized at all points), and it gives us information about the accuracy of the interpolation, by calculating an error estimate, called *kriging variance* (Matheron, 1963). Because of this, kriging is an often used interpolation technique in environmental research (e.g., Liu and Rossini, 1996; Paniagua et al., 2013; Sanders et al., 2012; Zirschky, 1985). The kriging variance can be used to quantify the model uncertainty, and to assist the sample search strategy in identifying potentially interesting regions in the study area based on a given condition. In this study, that condition is defined as "electric-field strength is higher than x V/m", with x a certain value to be determined. The sample search strategy enables both the efficient and dense sampling of identified interesting regions, as well as the efficient search of additional interesting regions, using a space-filling search pattern.

The considered search strategy consists in a weighted combination of two criteria. The first criterion is called the *generalized probability of improvement* criterion, defined as "the probability that the electric-field strength at a certain location lies within a certain output range", with the output range corresponding to the stated condition. In this case, the output range is defined as "all values of the electric-field strength higher than x V/m". This criterion ensures that interesting regions where hotspots are located are sampled more densely. The second criterion is called the *minimum distance* criterion, which calculates the distance to the closest measurement location. Maximizing the minimum distance criterion ensures the research area is properly searched and samples are widely spread (called *space-filling*). The mathematical breakdown of the two criteria is given in Couckuyt et al. (2012).

2.3. Measurement equipment

For model input and validation measurements, we used an NBM-550 broadband field meter with an EF-0391 isotropic electric field probe (Narda Safety Test Solutions, Pfullingen, Germany). The probe has a frequency range of 100 kHz–3 GHz, covering all relevant radiofrequency signals (e.g., (digital) radio, digital television, wireless telecommunications, etc.), and a measurement range of 0.2–320 V/m. During the measurements, which were performed at outdoor places that were accessible to the general public (e.g., streets, pavements, parking spots), the device was held at a height of 1.5 m, a typical height to characterize human exposure (ECC, 2004), as far as possible from the body, and carefully moved over an area of approximately 1 m^2, keeping the same height, but varying the orientation, to account for small-scale fading and possible shadowing of the body. The measurements were taken as the temporal averages over 30 s using the root-mean-square mode of the device. All measurements were performed between March and August 2012, on weekdays and during the daytime (avoiding the busy hours at noon and at 4 pm).

The spectrum analyzer setup, used for narrow-band measurements at the revealed hotspots, consists of a Precision Conical Dipole PCD 8250 antenna (Austrian Research Centers Seibersdorf Research GmbH, Seibersdorf, Austria), with a dynamic range of 1.1 mV/m—100 V/m and a frequency range of 80 MHz–3 GHz, in combination with a spectrum analyzer of type Rohde & Schwarz FSL6 with frequency range 9 kHz–6 GHz (Rohde & Schwarz, Zaventem, Belgium). The measurement uncertainty (the expanded uncertainty evaluated using a confidence interval of 95%) for the considered setup is ± 3 dB (CENELEC, 2008; Joseph et al., 2012). Optimal spectrum

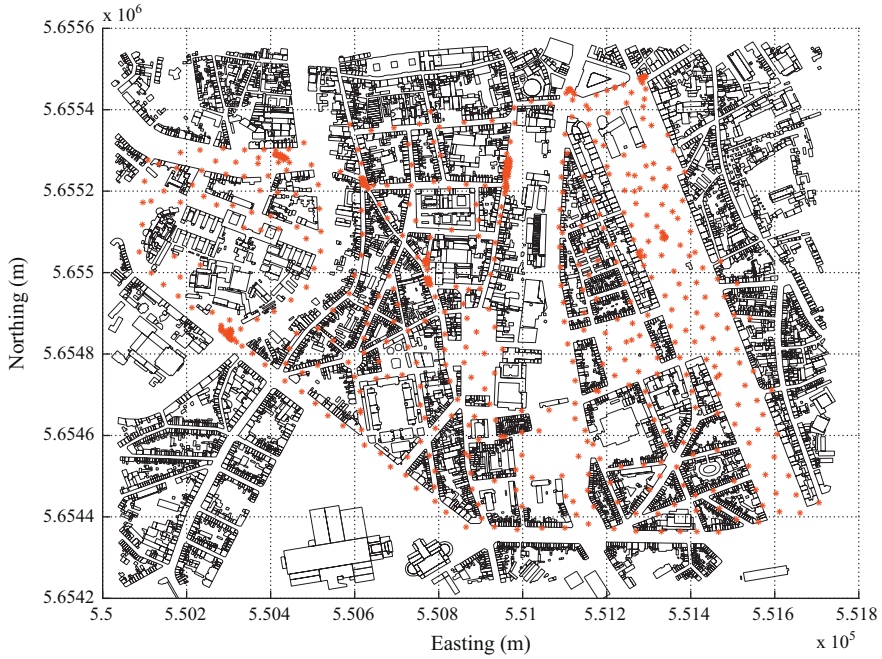

Fig. 1. Area under study of about 1 km^2 in Ghent, Belgium, with indication of the input measurement locations (red dots). The (approximately triangular) area is demarcated by the outer measurement locations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

analyzer settings for both measurements are discussed in Joseph et al. (2002, 2012). During the measurements the tri-axial probe was positioned at a height of 1.5 m. After performing an overview measurement, which consists in scanning the wireless communication frequency bands for existing signals, the frequency bands corresponding to those signals were measured in detail. The total duration of a single measurement depended on the number of dominating signals present, but was typically 30 min per location.

2.4. Measurement and modeling procedure

We developed a measurement and modeling procedure specifically for electromagnetic-field modeling in an outdoor environment, comprising an iterative choice of measurement locations. Apart from the actual measurements, the procedure is fully automated, using the Matlab "surrogate-model toolbox" (Gorissen et al., 2010). The followed procedure can be broken down in a series of steps.

Step 1*: Characterization of the area.* In order to select measurement spots only at accessible locations, the coordinates of the building blocks inside the area should be known. For this purpose, we use an online Google maps tool (http://www.birdtheme.org/useful/googletool.html) to draw and export the building block polygons.

Step 2*: Initial design.* The initial design is the distribution of the first batch of measurement locations (i.e., batch 0). Since there is no knowledge available yet on which to base our choice, we are free to choose any distribution. However, we opt for an optimized *Latin Hypercube Design* (Joseph and Hung, 2008), a space-filling design that distributes the positions in such a way that the area of interest is covered as evenly as possible. Because of the size of the research area, we chose an initial batch of 100 locations.

Step 3*: Measurements.* Broadband measurements are performed at the chosen locations. After the first batch, we decided to take the 75th percentile of the initial measurements as the x-value in the condition "all values of the electric-field strength higher than x V/m". It should be noted that this value is not updated after additional measurements are performed, as doing so could result in overlooking hotspots (if the 75th percentile would increase after additional batches), or designating too many less exposed regions as hotspots (if the 75th percentile would decrease after additional batches).

Step 4*: Modeling and sampling.* The kriging interpolation technique is used to model the measurement data, and the sample search strategy is used to determine the locations where additional measurements should be performed (see search algorithm described in Section 2.2 and the work of Couckuyt et al. (2012)).

Repeat steps 3 *and* 4. As new batches of locations are chosen, and new measurements are performed, more and more information about the electric-field strength and the hotspots is gained, and the surrogate model is subsequently updated, going from state K_0 to K_n, with n the total number of iterations. At a certain moment, however, this gain is outweighed by the effort to perform the measurements. Hence, we insert a *stopping criterion*, a certain condition that takes into account the information gain per additional measurement and when met, halts the procedure. In this study, the stopping criterion is defined as "the relative change of the surrogate model, K_i, compared to its previous state, K_{i-1}, is 2% or lower". This relative change of the model compared to its previous state is the mean of the relative change in electric-field strength calculated over the whole grid of the model (excluding indoor areas), with a resolution of 1×1 m^2, and is a measure for the amount of information added to the model by performing more measurements (Aerts et al., 2013).

Step 5*: Final surrogate model and analysis of the hotspots.* When the procedure is finished, the result is a model that outlines the RF-EMF exposure hotspots in the streets of the area under study. However, only the total electric-field strength has been measured, and no information about the contribution of individual radiofrequency signals has been obtained yet. In order to identify the signals bearing a relevant contribution to the total field in the discovered hotspots, accurate narrow-band measurements are performed with a spectrum analyzer setup (see Section 2.3, and Joseph et al. (2012)).

2.5. Validation

In order to assess the overall accuracy of our surrogate model as well as the performance of our iterative method, we applied two kinds of validations to our models.

The first kind of validation was a *global* validation, in which the models' predictions were tested against 50 (broadband) measurements, performed throughout the area under study. The locations of these validation measurements were randomly chosen, but such that the distance between any pair of them was at least 100 m, and the distance from any (model input) measurement location at least 10 m. As such, the global prediction accuracy of the subsequent states of our surrogate model is assessed.

The second kind of validation was a two-fold *local hotspot* validation, in which the predictions of models K_0–K_4 (i.e., the models built from measurement batches 0 to 4) were tested against the measurements of the last batch (i.e., batch 5). On the one hand, we assessed whether a batch 5 measurement location X exhibiting a measured electric-field strength E_{meas} higher than 0.7 V/m, was indeed predicted by models K_0–K_4 to lie inside a hotspot; or, in other words, whether the modeled electric-field strength E_{model} at location X was also higher than 0.7 V/m. On the other hand, we assessed whether the locations of batch 5 that were predicted to lie inside a hotspot ($E_{model} > 0.7$ V/m) indeed exhibited an elevated electric-field strength, $E_{meas} > 0.7$ V/m.

Performing this local validation allowed us to assess the prediction accuracy of the subsequent model states in the hotspot regions, as this could not be done in the global validation due to the distance constraints, as well as determine the overall

performance of our method and the evolution of the hotspot prediction accuracy as the model was updated.

In our analysis of both kinds of validations, we distinguish between correlation and error metrics. The correlation parameters (including coefficients of agreement) of choice are Pearson's correlation coefficient, r, the Spearman rank correlation coefficient, ρ, Cohen's kappa, κ, sensitivity, and specificity. Cohen's kappa is a statistical measure of the agreement between two data sets, taking into account the agreement occurring by chance. It represents the fraction of samples that were expected not to be in agreement (as in 'fall in the same exposure category') when only chance agreement would be present, but, in fact, are in agreement. For the calculation of this value, we use the 50th and 90th percentiles of the predicted and measured electric-field values as cut-offs (Frei et al., 2009b). The sensitivity is the ratio of the number of correctly identified "exposed" samples to the total number of measured "exposed" samples. The specificity is the ratio of the number of correctly identified "unexposed" samples to the total number of measured "unexposed" samples. A certain sample is classified as "exposed" when it lies above a certain percentile or a fixed field value, while "unexposed" means that the sample lies below a certain percentile. In this paper, we used the 90th percentiles as cut-off values. In our analysis, we assume the broadband measurements are the gold standard against which the model predictions are tested.

In case of the local hotspot validation, we introduce an additional metric, namely the prediction accuracy, i.e., the percentage of correctly predicting either a measurement result of $E_{meas} > 0.7$ V/m or whether a measurement location lies inside a hotspot.

3. Results and discussion

3.1. Broadband measurements

Altogether, 650 broadband measurements were performed during this study; six batches of 100 measurements used as input for our sequential modeling method, the locations of which are portrayed on Fig. 1, and 50 measurements for the global validation of the resulting surrogate models. The electric-field parameters of these broadband measurements are listed in Table 1.

The 75th percentile of the first batch, 0.70 V/m, is thereafter selected as threshold value for a hotspot, and the *generalized probability of improvement* criterion (Section 2.2) is then defined as "the probability that the electric-field strength at a certain location is higher than 0.70 V/m". It should be noted that this value is retained through the course of the study, even though the 75th percentile of later batches is an increasingly higher value (Table 1).

The subsequent input measurement batches show a steady increase in average electric-field strength (from 0.56 to 0.85 V/m), 75th (from 0.70 to 1.20 V/m) and 95th (from 0.96 to 2.29 V/m) percentiles, and in the observed standard deviation (0.23–0.68 V/m). Moreover, the minimum–maximum electric-field strength range is mostly expanded, going from 0.30 - 1.60 V/m (batch 0) to 0.12–3.10 V/m (batch 5), while the median electric-field strength stays relatively constant, around 0.49 V/m.

Table 1
Summary of the electric-field parameters of the broadband probe measurements for input (per batch of 100 measurements, and in total) and validation (50 measurements).

	# Measurements	Electric-field parameters (V/m)					
		E_{min}–E_{max}	E_{avg}	E_{median}	E_{p75}	E_{p95}	*STD*
Sequential design (input)							
Batch 0	100	0.30–1.60	0.56	0.48	0.70	0.96	0.23
Batch 1	100	0.15–2.83	0.64	0.50	0.73	1.46	0.44
Batch 2	100	0.10–2.77	0.63	0.46	0.87	1.66	0.50
Batch 3	100	0.12–2.52	0.76	0.45	1.13	1.96	0.61
Batch 4	100	0.04–3.06	0.75	0.49	1.01	2.07	0.63
Batch 5	100	0.12–3.10	0.85	0.56	1.20	2.29	0.68
Total	600	0.04–3.10	0.70	0.49	0.93	1.90	0.54
Validation	50	0.16–1.18	0.49	0.41	0.48	0.93	0.52

Electric-field parameters: E_{min}–E_{max} is the minimum–maximum interval, E_{avg} is the average value, E_{median} is the median value, E_{p75} and E_{p95} are the 75th and 95th percentiles of the electric-field distribution, and *STD* its standard deviation.

Remember that the combination of the two criteria in our search strategy, namely *generalized probability of improvement* and *minimum distance*, ensure that both the interesting regions (hotspots) are sampled more densely, and the research area is properly searched and samples are widely spread. The behavior of the electric-field parameters of the subsequent measurement batches perfectly reflects this search strategy (Table 1). On the one hand, generally higher field values are measured by focusing part of the sampling on the hotspots (increase in E_{avg}, E_{p75}, and E_{p95}), while on the other hand, "randomly" distributed electric-field strengths are measured in regions which had not been properly investigated, including both low electric-field values (hence sometimes, lower E_{min} are obtained in later batches), and high electric-field values (e.g., when a new hotspot is discovered). However, the majority of the electric-field values measured this way are situated around the region's average electric-field strength, which is why E_{median} barely changes.

In total, the input measurements vary between 0.04 and 3.10 V/m, with an average electric-field strength of 0.70 V/m and a median of 0.49 V/m. The validation measurements, on the other hand, vary between 0.16 and 1.18 V/m with an average electric-field strength of 0.49 V/m and a median of 0.41 V/m, respectively. Being performed at randomly chosen locations, the validation set offers us a better estimation of the average electric-field strength in the area (0.49 V/m). The standard deviation of the two measurement sets are comparable (0.54 vs. 0.52 V/m).

All measured electric-field strengths (with a maximum of 3.10 V/m) are well below the reference levels issued by the International Commission on Non-Ionizing Radiation Protection for the various contributing frequencies (e.g., 41 V/m for 900 MHz, which is the dominating frequency in our area, see Section 3.3) (ICNIRP, 1998).

3.2. Modeling

Steps 3 and 4 in Section 2.4 were repeated six times; six batches of 100 measurements resulted in six successive surrogate models of the total radiofrequency electromagnetic field exposure in the area under study. The electric-field parameters of these models, calculated with a grid resolution of 1 m × 1 m, are listed in Table 2. It should be noted that these parameters were calculated considering only those grid points located in the streets of the area under study.

The surrogate model's average electric-field strength steadily decreases from 0.57 to 0.49 V/m over the subsequent model states, K_0 to K_5. From model K_4 on, it settles at 0.49 V/m, which is in excellent agreement with the average value of the validation measurements (Section 3.1). A similar trend is seen in the

Table 2
Summary of the electric-field parameters of the subsequent interpolation models.

Model parameters					
Model #	E_{min}–E_{max} (V/m)	E_{avg} (V/m)	E_{p95} (V/m)	*STD* (V/m)	Change (%)
K_0	0.25–1.59	0.57	0.74	0.36	–
K_1	0.15–2.81	0.57	0.77	0.48	7.62
K_2	0.10–2.82	0.53	0.68	0.43	8.41
K_3	0.10–2.81	0.51	0.68	0.42	4.69
K_4	0.05–3.00	0.49	0.65	0.42	3.48
K_5	0.05–3.02	0.49	0.64	0.42	1.98

Electric-field parameters: E_{min}–E_{max} is the minimum–maximum interval, E_{avg} is the average value, E_{p95} is the 95th percentile of the electric-field distribution, and *STD* its standard deviation.
Change is the relative change of the model (in percent) compared to its previous state (e.g., the change of the kriging model K_4 compared to its previous state K_3 is 3.48%).

evolution of the 95th percentile, which decreases from 0.75 to 0.64 V/m. This behavior indicates that, although higher electric-field strengths are measured (Section 3.1), they represent a smaller area than the lower measured electric-field strengths.

The minimum–maximum range of the electric-field strength widens as the model is updated, from 0.25–1.59 V/m to 0.05–3.02 V/m, closely following the expansion of the minimum–maximum electric-field strength range of the measurement batches. The slight difference observed can be attributed to the finite grid size (1 m × 1 m) of the analyzed models.

The metric introduced as a measure for the stopping criterion, the relative change of the current model, K_i, compared to its previous state, K_{i-1}, is also listed in Table 2. After a slight rise in the change going from K_0 to K_2 (7.62–8.41%), the change drops below 2% after the sixth iteration, at model K_5, and we stop the algorithm. This parameter was also calculated considering only the grid points located in the streets.

Fig. 2 shows (a) the heat map constructed from the final surrogate model, K_5, along with (b) its associated kriging variance, a measure for the prediction error. A number of regions with electric-field strengths higher than 0.7 V/m, both large and small, are identified: five of these hotspots of reasonable size and highest maximum measured electric-field strengths are indicated on Fig. 2 (a), numbered 1 to 5, corresponding to the spectrum analyzer measurement results in Section 3.3. The greater part of the area under study exhibits electric-field strengths between 0.35 and

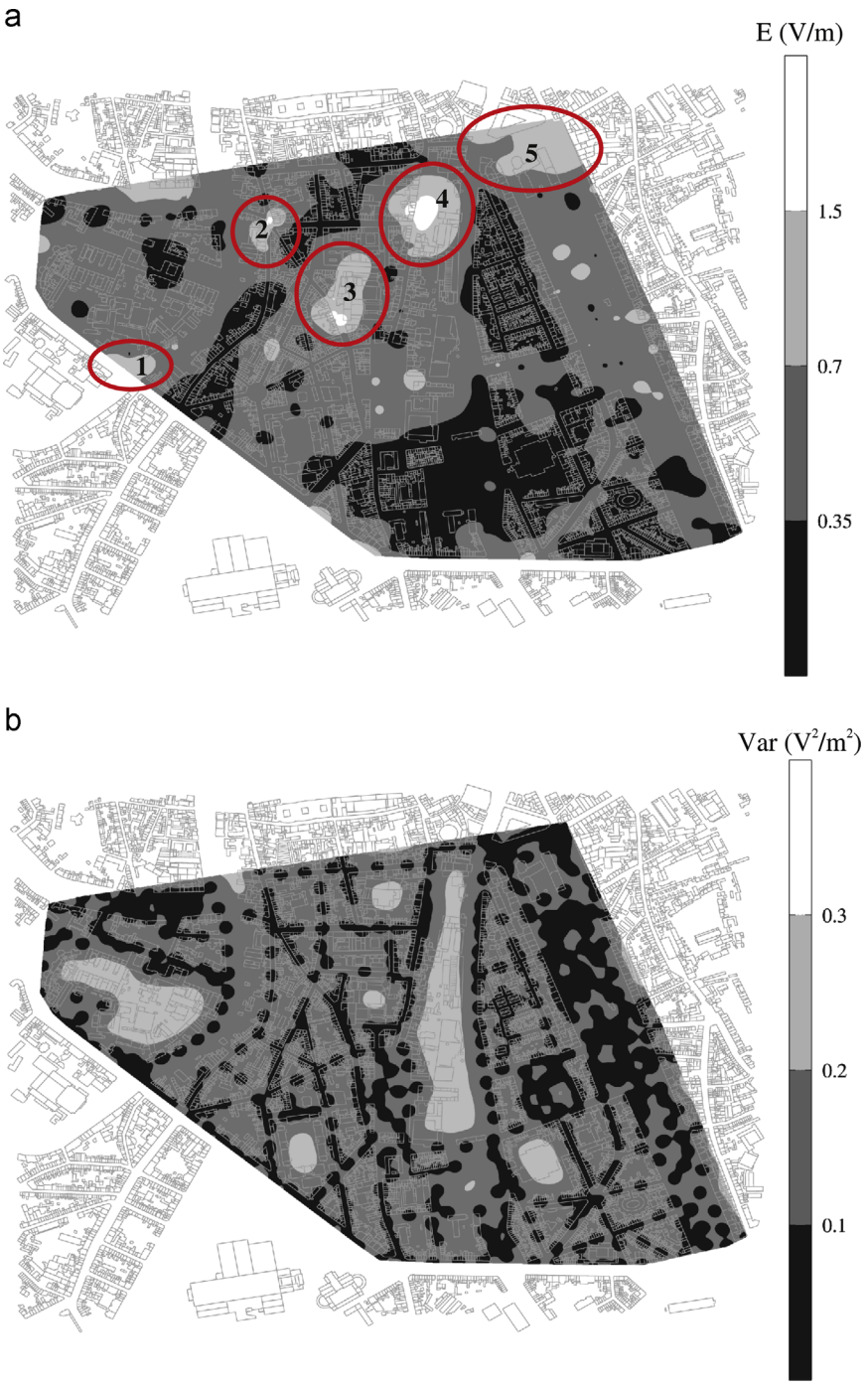

Fig. 2. (a) Heat map of the RF-EMF exposure (in V/m); (b) map of the kriging variance (Var, in V^2/m^2). Locations of the five hotspots are indicated, with the numbers corresponding to Fig. 3.

0.70 V/m, while large regions feature a relatively low exposure, with electric-field strengths below 0.35 V/m. The variance ranges from about 0.4 V^2/m^2, in areas which have not been and/or could not be surveyed (mainly due the presence of large building blocks and canals), to less than 0.1 V^2/m^2, with minimum variance at the specific measurement locations. As the hotspots were densely sampled, they exhibit a very low kriging variance. The global minimization of this variance is inherent to kriging interpolation.

3.3. Contributions to the exposure

Following the construction of the RF-EMF exposure heat map of Fig. 2(a), accurate narrow-band spectrum analyzer measurements were performed inside the five identified hotspots of reasonable size and highest maximum electric-field strengths (ranging from 1.30 to 3.10 V/m), allowing us to identify the relative contributions of the radiofrequency signals to the total exposure therein. After an initial spectral overview measurement, only the signals showing an electric-field strength of 0.05 V/m or higher were considered relevant and subsequently measured more accurately. Then, the individual, relative contributions to the total exposure (defined here as the percentual contribution to the total power density) were calculated, the results of which are shown in Fig. 3. Altogether, five signals were found to contribute to the exposure in the identified hotspots. A radio signal (FM, at approximately 100 MHz) was present in two hotspots, with relative contributions of 1% and 8%; a digital radio signal (T-DAB, at 224 MHz) was present in four hotspots, with relative contributions ranging from 0.1% to 11%; GSM base station signals at 900 MHz were present in all five hotspots, with relative contributions ranging from 45% to 100%; GSM base station signals at 900 MHz were present in two hotspots, with relative contributions of 9% and 44%; and UMTS base station signals (at 2100 MHz) were present in two hotspots, with relative contributions of 9% and 29%. Of the five contributing sources, only GSM base station signals at 900 MHz were always present, and always represent the dominant source, which is consistent with the findings of Joseph et al. (2008) and Aerts et al. (2013). At one location, GSM base station signals at 1800 MHz, however, were a close contender (hotspot 1, with 44% and 45% for GSM base station signals at 1800 and 900 MHz, respectively). A digital radio signal was also often present, but contributed, on average, the least. FM radio signals, GSM base station signals at 1800 MHz, and UMTS base station signals were all present in only two hotspots, with GSM base station signals at 1800 MHz having, on average, a higher contribution.

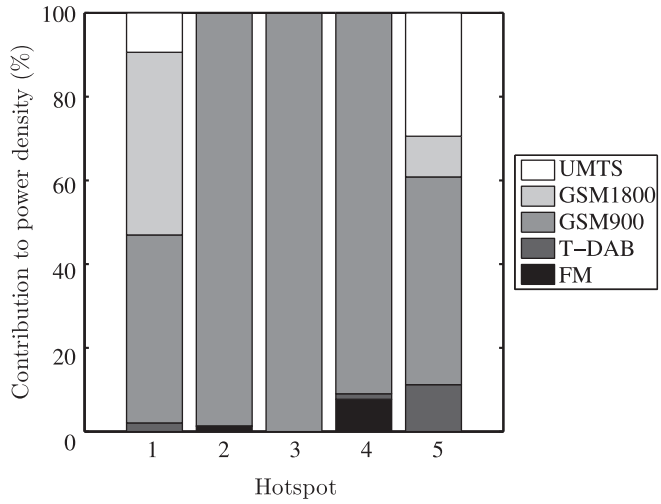

Fig. 3. Relevant radiofrequency signals and their contributions (in %) to the total exposure (power flux density, shortly noted as power density) in the five hotspots. While electric-field strength is used as the exposure metric, only the signals' power densities (in W/m^2, or Watts per square meter) can be added linearly. The relation between power density (S) and electric-field strength (E) is given by $S=E^2/377$. The numbers on the Hotspot-axis correspond to Fig. 2(a). (UMTS=Universal Mobile Telecommunications System, GSMx=Global System for Mobile Communications at x MHz, T-DAB=Terrestrial Digital Audio Broadcasting, FM=frequency modulation).

3.4. Validation

3.4.1. Global validation

The results of the global validation analysis (correlation and error metrics) of the subsequent models are listed in Table 3. Pearson's correlation coefficient, r, shows an overall increasing trend, namely from 0.55 (confidence interval (CI) 95% 0.31–0.72) for the first model, K_0, to 0.73 (CI 95% 0.56–0.84) for the last, K_5, which is an excellent value in this research (Frei et al., 2009b). A similar trend is seen for the Spearman rank correlation coefficient, ρ, which evolves from 0.58 to 0.72. Cohen's kappa, κ, on the other hand, shows a less linear evolution, with values ranging between 0.28 (CI 95% 0.04–0.51) to 0.55 (CI 95% 0.34–0.76), before settling at 0.41 (CI 95% 0.19–0.64) for K_5. The specificity ranges between 0.93 and 0.96, and the sensitivity between 0.40 and 0.60, settling at respectively 0.96 and 0.60 for the final model, K_5.

Both error metrics listed in Table 3, the average relative error, RE_{avg}, (in dB) and the percentage of relative errors above 3 dB, develop more linearly, showing a near-constant decrease, or, in other words, improvement. RE_{avg} decreases from 3.14 to 1.96 dB, while the percentage of relative errors above 3 dB decreases from 44 to 22%. In terms of correlation, the second model, K_1, surprisingly, features the best results with slightly better correlation coefficients, and the same sensitivity and specificity as K_5. In terms of accuracy, K_5 is the best model.

The very low average error, good correlation and very good specificity—all indispensable traits—of our final surrogate model indicate the usefulness of our methodology. Although the sensitivity is moderate, this should not be an obstacle, since only a small fraction of the much larger unexposed area will be, in fact, exposed (Neubauer et al., 2007).

3.4.2. Hotspot validation

The results of the local hotspot validation analysis are listed in Table 4. The first model, K_0, constructed from batch 0 (which locations are distributed in a uniform, space-filling grid) gives a very poor prediction of the hotspot locations; only about half of the locations of batch 5 with a measured electric-field strength E_{meas} higher than 0.7 V/m were, in fact, predicted to lie inside a hotspot. K_0 is, however, better at demarcating the hotspots it did find; 77% of the batch 5 locations inside its predicted hotspots indeed yield electric-field values above 0.7 V/m. Over the course of the sequential design (models K_0–K_4), the results improve considerably. For K_4, RE_{avg} is about 1.7 dB, only approximately 20% of the errors are larger than 3 dB, and the prediction accuracy has increased to 90%, while the correlation coefficient r is about 0.75, and κ is larger than 0.60. Since the error metrics, as well as the correlation coefficients, are even superior to the respective results from the global validation, we can also conclude that the functioning of our methodology is sufficiently demonstrated; the hotspots are well-defined and accurately modeled.

3.5. Strengths and limitations

The exposure modeling approach proposed in the study is based upon a new sequential design to select its sample locations by focusing on the regions that are particularly of interest, i.e., regions with elevated exposure to RF-EMF (hotspots). As such, an efficient sampling scheme is constructed that ultimately results

Table 3
Analysis of the global validation of the subsequent surrogate models, K_0–K_5.

Global validation analysis							
Model #	***r* (CI)**	**ρ**	**κ (CI)**	**Spec.**	**Sens.**	***RE*$_{avg}$ (dB)**	**Errors > 3 dB (%)**
K_0	0.55 (0.31, 0.72)	0.58	0.34 (0.11, 0.58)	0.93	0.40	3.14	44
K_1	0.78 (0.64, 0.87)	0.77	0.55 (0.34, 0.76)	0.96	0.60	2.51	32
K_2	0.68 (0.49, 0.80)	0.69	0.34 (0.11, 0.58)	0.93	0.40	2.33	34
K_3	0.70 (0.52, 0.82)	0.69	0.28 (0.04, 0.51)	0.93	0.40	2.20	30
K_4	0.71 (0.54, 0.83)	0.70	0.28 (0.04, 0.51)	0.93	0.40	2.10	26
K_5	0.73 (0.56, 0.84)	0.72	0.41 (0.19, 0.64)	0.96	0.60	1.96	22

Correlation parameters: *r* is Pearson's correlation coefficient, ρ is the Spearman rank correlation coefficient, κ is Cohen's kappa coefficient, spec. is the specificity, and sens. is the sensitivity. Cut-off percentiles for the calculation of κ are the 50th and 90th percentiles.; the latter is also used for the calculation of the sensitivity and specificity. *Error metrics*: RE_{avg} is the average deviation or relative error in dB, Errors > 3 dB is the percentage of relative errors above 3 dB (factor $\sqrt{2}$ in electric-field strength).

Table 4
Analysis of the two-fold local hotspot validation of the subsequent surrogate models, K_0–K_4. E_{meas} (batch 5) > 0.7 V/m: it is assessed whether the batch 5 locations with E_{meas} > 0.7 V/m are predicted by models K_0–K_4 to lie inside a hotspot. E_{model} (batch 5) > 0.7 V/m: it is assessed whether the batch 5 locations that are predicted to lie inside a hotspot by models K_0–K_4 (E_{model} > 0.7 V/m) exhibit an elevated electric-field strength, E_{meas} > 0.7 V/m.

Local validation analysis					
Model #	***r* (CI)**	**κ (CI)**	***RE*$_{avg}$ (dB)**	**Errors > 3 dB (%)**	**Pred. acc. (%)**
E_{meas} (batch 5) > 0.7 V/m					
K_0	−0.37 (−0.60, −0.09)	−0.32 (−0.73, 0.10)	5.72	58.7	56.5
K_1	0.51 (0.25, 0.69)	0.19 (−0.19, 0.56)	2.56	32.6	91.3
K_2	0.58 (0.34, 0.74)	0.44 (0.11, 0.77)	2.09	26.1	91.3
K_3	0.71 (0.54, 0.83)	0.50 (0.18, 0.81)	1.76	21.7	91.3
K_4	0.74 (0.58, 0.85)	0.62 (0.34, 0.90)	1.63	19.6	91.3
E_{model} (batch 5) > 0.7 V/m					
K_0	0.08 (−0.27, 0.41)	−0.07 (−0.55, 0.42)	3.07	35.3	76.5
K_1	0.52 (0.28, 0.70)	0.21 (−0.16, 0.58)	2.81	35.4	87.5
K_2	0.61 (0.40, 0.76)	0.39 (0.06, 0.73)	2.36	29.2	87.5
K_3	0.74 (0.57, 0.85)	0.51 (0.20, 0.82)	1.85	23.4	89.4
K_4	0.77 (0.61, 0.86)	0.63 (0.35, 0.91)	1.70	21.3	89.4

r is Pearson's correlation coefficient, κ is Cohen's kappa coefficient, both with CI the 95% confidence interval. Cut-off percentiles for the calculation of κ are the 50th and 90th percentiles. RE_{avg} is the average deviation or relative error in dB, Errors > 3 dB is the percentage of relative errors above 3 dB (factor $\sqrt{2}$ in electric-field strength). Pred. acc. is the hotspot prediction accuracy. E_{meas} is the $_{meas}$ured electric-field strength, E_{model} is the modeled electric-field strength.

in a heat map of the outdoor RF-EMF exposure in a large area that features well-defined hotspots, representing important graphical information for risk communication. Using classical sampling methods (e.g., Joseph et al., 2012; Paniagua et al., 2013), it is not possible to identify and characterize hotspots except coincidentally. Validation shows excellent agreement between model and measurements, both in terms of error and validation metrics as well as the overall average electric-field strength. Our model, however, is not valid indoors, and has not been validated in indoor environments.

Due to the large amount of measurements necessary in a study covering an area of this size, we used a broadband probe as measurement device, despite its inherent inaccuracy compared to other available devices, such as the spectrum analyzer. However, its portability and measurement speed are essential in a measurement-based exposure assessment of this scope, and we believe that the purpose of this study—assessment of the total, outdoor RF-EMF exposure—validates its use. The "total RF-EMF exposure" in this study encompasses only RF-EMF emanating from base station (for mobile telecommunication) or transmitter (e.g., television) antennas. Thus, we do not consider signals from personal devices (e.g., mobile phones, cordless telephones, etc.) here. And while no distinction can be made between different radiofrequency sources when using a broadband probe, performing accurate narrow-band spectrum analyzer measurements in the revealed hotspots permits us to identify the sources that are present in the different areas of elevated exposure, and their respective, relative contributions. The influence of the buildings and the topography is inherent in the measurements, however, it was not considered during the interpolation. Also, temporal variations are not accounted for. However, we assume that the locations of the hotspots do not change during the measurement campaign, unless infrastructural changes would be applied to base station or transmitter antennas.

It should be mentioned that electric-field values below the broadband measurement device's sensitivity of 0.2 V/m were nonetheless measured by the device, and hence retained, although accuracy could not be ensured. However, no relevant errors are introduced in the models, because the values and possible associated absolute errors are small.

A model input batch of 100 measurements might have been too extensive, so as to improve the efficiency of our methodology, we will investigate in following studies making use of our sequential sampling method for exposure assessment, if smaller measurement batches could be used, reducing the required total number of measurements.

4. Conclusions

Our approach results in the relatively fast construction of an accurate *heat map* of the outdoor exposure to radiofrequency electromagnetic fields that characterizes and outlines the hotspot regions, using kriging as interpolation technique. As such, it supplies an accurate, graphical representation of the exposure, which can be easily understood by laymen, and where the aim is

to identify regions of relatively high exposure (hotspots). Analysis of the validation shows a good correlation (0.7), low average relative error (below 2 dB), and near-perfect specificity (0.96). The constructed surrogate model can serve as input, optimization, or validation to more sophisticated epidemiological exposure models. Future research will consist of accounting for temporal variations as well as exposure to personal and indoor devices. Also indoor exposure prediction is a further step in this research.

Acknowledgment

This work has been carried out with the financial support of the iMinds project 'Green Wireless Efficient City Access Networks (GreenWeCan)' and the Interuniversity Attraction Poles Programme BESTCOM initiated by the Belgian Science Policy Office. D. Deschrijver and W. Joseph are Post-Doctoral Fellows of the FWO-V (Research Foundation—Flanders).

References

Aerts, S., Deschrijver, D., Joseph, W., Verloock, L., Goeminne, F., Martens, L., Dhaene, T., 2013. Exposure assessment of mobile phone base station radiation in an outdoor environment using sequential surrogate modeling. Bioelectromagnetics 34, 300–311.

Anglesio, L., Benedetto, A., Bonino, A., Colla, D., Martire, F., Saudino Fusette, S., D'Amore, G., 2001. Population exposure to electromagnetic fields generated by radio base stations: evaluation of the urban background by using provisional model and instrumental measurements. Radiat. Prot. Dosimetry 97, 355–358.

Azpurua, M.A., Dos Ramos, K., 2010. A comparison of spatial interpolation methods for estimation of average electromagnetic field magnitude. PIER M 14, 135–145.

Beekhuizen, J., Vermeulen, R., Kromhout, H., Bürgi, A., Huss, A., 2013. Geospatial modelling of electromagnetic fields from mobile phone base stations. Sci. Total Environ. 445–446C, 202–209.

Bolte, J.F.B., Eikelboom, T., 2012. Personal radiofrequency electromagnetic field measurements in the Netherlands: Exposure level and variability for everyday activities, times of day and types of area. Environ. Int. 48C, 133–142.

Breckenkamp, J., Neitzke, H.P., Bornkessel, C., Berg-Beckhoff, G., 2008. Applicability of an exposure model for the determination of emissions from mobile phone base stations. Radiat. Prot. Dosimetry 131, 474–481.

Bürgi, A., Theis, G., Siegenthaler, A., Röösli, M., 2008. Exposure modeling of high-frequency electromagnetic fields. J. Expo. Sci. Environ. Epidemiol. 18, 183–191.

Bürgi, A., Frei, P., Theis, G., Mohler, E., Braun-Fahrländer, C., Fröhlich, J., Neubauer, G., Egger, M., Röösli, M., 2010. A model for radiofrequency electromagnetic field predictions at outdoor and indoor locations in the context of epidemiological research. Bioelectromagnetics 31, 226–236.

CENELEC (European Committee for Electrotechnical Standardization), 2008. TC 106x WG1 EN 50492 in situ. Basic standard for the in-situ measurement of electromagnetic field strength related to human exposure in the vicinity of base stations. Brussels, Belgium.

Couckuyt, I., Aernouts, J., Deschrijver, D., Turck, F., Dhaene, T., 2012. Identification of quasi-optimal regions in the design space using surrogate modeling. Eng. Comput. Springer-Verlag, London Limited.

Crombecq, K., Gorissen, D., Deschrijver, D., Dhaene, T., 2011. A novel hybrid sequential design strategy for global surrogate modeling of computer experiments. SIAM J. Sci. Comput. 33, 1948–1974.

Deschrijver, D., Crombecq, K., Nguyen, H.M., Dhaene, T., 2011. Adaptive sampling algorithm for macromodeling of parameterized S-parameter responses. IEEE Trans. Microw. Theory Techn. 59, 39–45.

Deschrijver, D., Vanhee, F., Pissoort, D., Dhaene, T., 2012. Automated near-field scanning algorithm for the EMC analysis of electronic devices. IEEE Trans. Electromagn. Compat. 54, 502–510.

Elliott, P., Toledano, M.B., Bennett, J., Beale, L., De Hoogh, K., Best, N., Briggs, D.J., 2010. Mobile phone base stations and early childhood cancers: case-control study. BMJ 340, c3077.

Frei, P., Mohler, E., Neubauer, G., Theis, G., Bürgi, A., Fröhlich, J., Braun-Fahrländer, C., Bolte, J.F.B., Egger, M., Röösli, M., 2009a. Temporal and spatial variability of personal exposure to radio frequency electromagnetic fields. Environ. Res. 109, 779–785.

Frei, P., Mohler, E., Bürgi, A., Fröhlich, J., Neubauer, G., Braun-Fahrländer, C., Röösli, M., 2009b. the Qualifex Team, 2009. A prediction model for personal radio frequency electromagnetic field exposure. Sci. Total Environ. 408, 102–108.

Frei, P., Mohler, E., Bürgi, A., Fröhlich, J., Neubauer, G., Braun-Fahrländer, C., Röösli, M., 2010. Classification of personal exposure to radio frequency electromagnetic fields (RF-EMF) for epidemiological research: Evaluation of different exposure assessment methods. Environ. Int. 36, 714–720.

Gorissen, D., Couckuyt, I., Demeester, P., Dhaene, T., Crombecq, K., 2010. A surrogate modeling and adaptive sampling toolbox for computer based design. J. Mach. Learn. Res. 11, 2051–2055.

ICNIRP, 1998. Guidelines for limiting exposure to time-varying electric, magnetic, and electromagnetic fields (up to 300 GHz). Health. Phys. 74, 494–522.

Isselmou, Y.O., Wackernagel, H., Tabbara, W., Wiart, J., 2008. Geostatistical estimation of electromagnetic exposure. Quant. Geo. G 15, 59–70.

Joseph, V.R., Hung, Y., 2008. Orthogonal-maximin Latin hypercube design. Stat. Sinica 18, 171–186.

Joseph, W., Olivier, C., Martens, L., 2002. A robust, fast and accurate deconvolution algorithm for EM-field measurements around GSM and UMTS base stations with a spectrum analyser. IEEE Trans. Instr. Meas. 51, 1163–1169.

Joseph, W., Vermeeren, G., Verloock, L., Heredia, M.M., Martens, L., 2008. Characterization of personal RF electromagnetic field exposure and actual absorption for the general public. Health Phys. 95, 317–330.

Joseph, W., Frei, P., Röösli, M., Thuróczy, G., Gajsek, P., Trcek, T., Bolte, J.F.B., Vermeeren, G., Mohler, E., Juhász, P., Finta, V., Martens, L., 2010. Comparison of personal radio frequency electromagnetic field exposure in different urban areas across Europe. Environ. Res. 110, 658–663.

Joseph, W., Verloock, L., Goeminne, F., Vermeeren, G.G., Martens, L., 2012. Assessment of RF exposures from emerging wireless communication technologies in different environments. Health Phys. 102, 161–172.

Liu, L.J.S., Rossini, A.J., 1996. Use of kriging models to predict 12-hour mean ozone concentrations in Metropolitan Toronto—A pilot study. Environ. Int. 22, 677–692.

Matheron, G., 1963. Principles of geostatistics. Econ. Geol. 58, 1246–1266.

Neitzke, H.P., Osterhoff, J., Peklo, K., Voigt, H., 2007. Determination of exposure due to mobile phone base stations in an epidemiological study. Radiat. Prot. Dosimetry 124, 35–39.

Neubauer, G., Feychting, M., Hamnerius, Y., Kheifets, L., Kuster, N., Ruiz, I., Schüz, J., Überbacher, R., Wiart, J., Röösli, M., 2007. Feasibility of future epidemiological studies on possible health effects of mobile phone base stations. Bioelectromagnetics 28, 224–230.

Paniagua, J.M., Rufo, M., Jimenez, A., Antolin, A., 2013. The spatial statistics formalism applied to mapping electromagnetic radiation in urban areas. Environ. Monit. Assess. 185, 311–322.

Rowley, J.T., Joyner, K.H., 2012. Comparative international analysis of radiofrequency exposure surveys of mobile communication radio base stations. J. Expo. Sci. Environ. Epidemiol. 22, 304–315.

Röösli, M., Frei, P., Bolte, J., Neubauer, G., Cardis, E., Feychting, M., Gajsek, P., Heinrich, S., Joseph, W., Mann, S., Martens, L., Mohler, E., Parslow, R.C., Poulsen, A.H., Radon, K., Schüz, J., Thuróczy, G., Viel, J.F., Vrijheid, M., 2010. Conduct of a personal radiofrequency electromagnetic field measurement study: proposed study protocol. Environ. Health 9, 23.

Sanders, A.P., Messier, K.P., Shehee, M., Rudo, K., Serre, M.L., Fry, R.C., 2012. Arsenic in North Carolina: public health implications. Environ. Int. 38, 10–16.

Thuróczy, G., Molnar, F., Janossy, G., Nagy, N., Kubinyi, G., Bakos, J., Szabo, J., Molnár, F., Jánossy, G., Szabó, J., 2008. Personal RF exposimetry in urban area. Ann. Telecommun. 63, 87–96.

Viel, J.F., Clerc, S., Barrera, C., Rymzhanova, R., Moissonnier, M., Hours, M., Cardis, E., 2009a. Residential exposure to radiofrequency fields from mobile phone base stations, and broadcast transmitters: a population-based survey with personal meter. Occup. Environ. Med. 66, 550–556.

Viel, J.F., Cardis, E., Moissonnier, M., De Seze, R., Hours, M., 2009b. Radiofrequency exposure in the French general population: band, time, location and activity variability. Environ. Int. 35, 1150–1154.

Zirschky, J., 1985. Geostatistics for environmental monitoring and survey design. Environ. Int. 11, 515–524.