



**HAL**  
open science

# Sensor-Aided EMF Exposure Assessments in an Urban Environment Using Artificial Neural Networks

Shanshan Wang, Joe Wiart

► **To cite this version:**

Shanshan Wang, Joe Wiart. Sensor-Aided EMF Exposure Assessments in an Urban Environment Using Artificial Neural Networks. *International Journal of Environmental Research and Public Health*, 2020, 10.3390/ijerph17093052 . hal-02887136

**HAL Id: hal-02887136**

**<https://telecom-paris.hal.science/hal-02887136>**

Submitted on 2 Jul 2020

**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.



Article

# Sensor-Aided EMF Exposure Assessments in an Urban Environment Using Artificial Neural Networks

Shanshan Wang \*  and Joe Wiart

Chaire C2M, LTCl, Télécom Paris, Institut Polytechnique de Paris, 91120 Palaiseau, France;

joe.wiart@telecom-paris.fr

\* Correspondence: shanshan.wang@telecom-paris.fr

Received: 2 April 2020; Accepted: 24 April 2020; Published: 28 April 2020



**Abstract:** This paper studies the time and space mapping of the electromagnetic field (EMF) exposure induced by cellular base station antennas (BSA) using artificial neural networks (ANN). The reconstructed EMF exposure map (EEM) in urban environment is obtained by using data from EMF sensor networks, drive testing and information accessible in a public database, e.g., locations and orientations of BSA. The performance of EEM is compared with Exposure Reference Map (ERM) based on simulations, in which parametric path loss models are used to reflect the complexity of urban cities. Then, a new hybrid ANN, which has the advantage of sorting and utilizing inputs from simulations efficiently, is proposed. Using both hybrid ANN and conventional regression ANN, the EEM is reconstructed and compared to the ERM first by the reconstruction approach considering only EMF exposure assessed from sensor networks, where the required number of sensors towards good reconstruction is explored; then, a new reconstruction approach using the sensors information combined with EMF along few streets from drive testing. Both reconstruction approaches use simulations to mimic measurements. The influence of city architecture on EMF exposure reconstruction is analyzed and the addition of noise is considered to test the robustness of ANN as well.

**Keywords:** EMF exposure; radiofrequency; artificial neural network; locally-connected layers; EMF sensor network

## 1. Introduction

The risk perception of electronic-magnetic field (EMF) exposure [1] is nowadays a hot issue with the fast development of wireless communications. The actual deployment of 5G infrastructures and base station antennas (BSA) strengthen such concern and reinforce the needs of EMF exposure assessment [2] and monitoring. In Europe, the EMF monitoring is often carried out using “one-time” measurement campaigns as described in [3,4]. Projects, like LEXNET [5], also studied population radiofrequency (RF) exposure. The monitoring can also be done using fixed sensors [6]. Recently, autonomic in energy and wide-band sensors have been proposed [7]. These sensors are installed and tested in cities such as Paris and Toulouse ( Observatoire des Ondes [8]). The sensor networks, as well as on-site measurements, are of great interest, but both methods are localized, which means EMF exposure is monitored at limited locations. Therefore, it is necessary to build an EMF Exposure Map (EEM) from these measurements.

In the present paper, the time and space mapping of outdoor EMF exposure induced by BSA in the 4G cellular networks is studied using artificial neural networks (ANN). The EEM in urban environment is reconstructed from sensor networks, drive testing, and information related to BSA that are accessible on the public database, including spatial locations and orientations of BSA. This information related to BSA is often used to design and optimize wireless networks but rarely to deal

with exposure assessment and monitoring. Other EMF sources, e.g., high-voltage transmission lines, substation equipment, are not considered in the present paper.

However, there are several challenges, which prevent us from reconstructing an accurate spatial map of EMF exposure using artificial intelligence (AI). Due to the complexity of building structure and city architecture, the propagation encounters multiple reflection, refraction and diffraction, which induces highly varying EMF in space. The challenge is to convert and process information into readable input variables of ANN. It is also important to find the significant features in RF exposure instead of feeding redundant inputs into ANN.

The mentioned challenges bring significant uncertainty and hinder us from reconstructing the EEM. Conventional methods, e.g., ray-based simulators [9], network-based measurements [10], and recent approaches like Kriging [11,12], are used in assessing EMF exposure. However, these approaches are either too complicated in terms of costs or not able to take into account all the meaningful features. Considering the complexity, accuracy and the ability to deal with high-dimension data, the mentioned methods are not feasible to cover all the aspects.

Due to the limitation in the diversity of inputs and the flexibility of the methods, we are strongly motivated to use ANN, a promising and novel approach, to reconstruct the spatial map of EMF exposure. ANN is one of the trending approaches nowadays, because of its efficiency and universality. It has been widely applied in solving environmental and public health issues, e.g., air pollution monitoring [13] and water quality modeling [14]. Recently it was applied in wireless communications [15,16] to quantify interference or coverage performance and propagation path loss (PL) prediction. In [17–19], PL is assessed using ANN with consideration of distance, antenna and terrain information. However, unlike PL, EMF exposure involves much more BSA, and thus it is linked to the aggregation of all possible PL plus noise. If we adopt the approach used in PL assessment, which considers all active BSA as inputs of ANN, it would bring the high-dimension inputs problem.

Therefore, our team started to assess the EMF exposure using a simple ANN model [20]. However, the simple and general ANN model is not sufficient to handle the large variability of the exposure and complexity of the city.

In the present paper, we present the reconstruction of EEM using a proposed hybrid ANN model based on drive testing assisted by sensor networks (both are extracted from simulations) and information on BSA. To analyze the accuracy and the robustness of different ANN and reconstruction methods, various EMF Exposure Reference Maps (ERM) are calculated using simulations, where a parametric propagation PL, able to reflect the city complexity, is used. To be more specific, we have the following contributions:

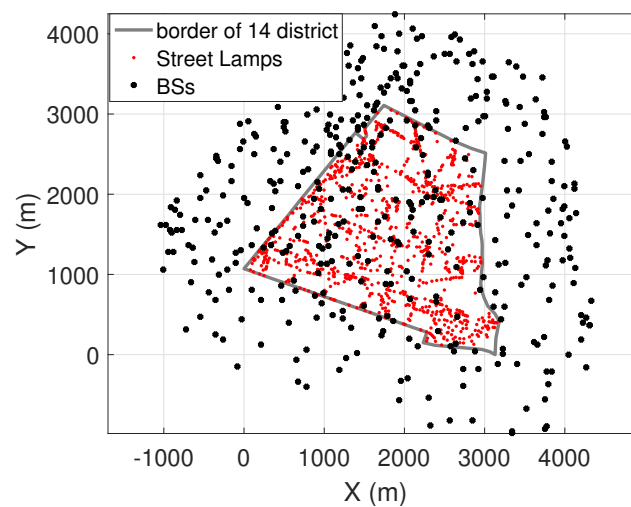
- We propose a new hybrid ANN model, and compare its performance with conventional fully-connected ANN in reconstructing EMF exposure. In order to make better use of key inputs of ANN, including distance to the source, direction of the transmitting antennas, blockage in the surrounding environment, time variation and background noise, we propose a hybrid ANN. The idea is to process information from one BSA locally, then concentrate the data from different BSA together.
- we propose a new reconstruction method based on the one-time drive testing with the help of sensor networks. We compare its performance against the reconstruction method based on sensor networks only. The new reconstruction method has the advantage of achieving good predictions with lower cost.

The paper is organized as follows: Section 2 introduces the system model and settings, Section 3 introduces the general set up of the neural network and how it can be applied to solve regression problems. Section 4 proposes a new type of hybrid neural network. Section 5 describes two main approaches used to reconstruct RF exposure. Section 6 validates the propositions in the previous sections and compares the performance between different ANN models with and without consideration of noise. Section 8 concludes the paper.

## 2. System Model

### 2.1. Exposure Reference Map Construction and Setting

In this paper, a 4G cellular network with BSA operating at 2600 MHz is considered. The real geographical locations of BSA used in the simulations are available on ANFR's databases [21]. The outdoor EMF exposure in the 14th district of Paris is of interest in the present paper, which can be interpreted as the EMF exposure received at the soil level of the open streets. The map of 14th district is shown in Figure 1, with real spatial locations of BSA and street lamps displayed in black and red dots, respectively. Sensor networks are deployed on selected street lamps and actively monitoring the real-time EMF exposure. In total, there are 3516 street lamps in the 14th district. The specifications of the sensor networks can be found in [7].



**Figure 1.** Base station antennas (BSA) and street lamps inside and around the 14th district in Paris.

Considering that each BSA is equipped with one-directional antenna operating at 2600 MHz, it is assumed that:

- the direction of each antenna is uniformly distributed;
- the transmit power for each BSA is equal.

In order to capture the diverse and varying blockages, the following PL model is proposed.

### 2.2. Parametric Exposure Reference Map

The EMF emitted from a source can be approximated as a spherical wave when the observer is in the far field region and free space. Therefore the received signal strength [22] is inversely proportional to the distance between the source and the receiver. The simple propagation PL model is widely-used [23]:

$$E = B + 10\alpha\log_{10}(d) + s; \quad (1)$$

where  $E$  is the electromagnetic field,  $B$  is a constant,  $d$  is the distance to the source and  $s$  is the shadow fading variation.  $\alpha$  is the Path Loss Exponent (PLE).

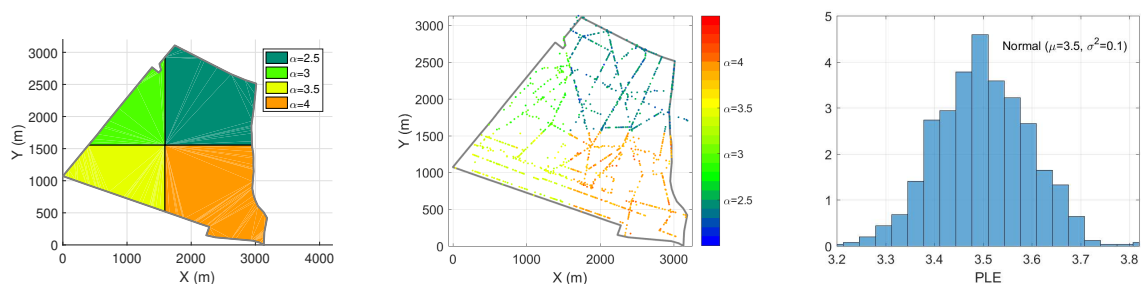
PLE is used to take into account the environment in the propagation. EMF are reflected, refracted and diffracted by obstacles on the propagation path. Each link between the receiver and the BSA experiences different attenuation caused by obstacles in the propagation, which can be represented by PLE. However, a simple PL model with a fixed value of PLE has limitations in representing complicated urban environment. In order to better mimic the varying propagation channel, we use

the following block-based stochastic PL model [9]. In this case, PLE is not a deterministic value anymore, but a random variable governed by the statistical law.

### 2.2.1. Stochastic Block-Based Path Loss Model

Since the BSA are usually located at high positions, e.g., rooftops or tower tops, the receivers' surroundings are more complicated and important. Therefore, we use a block-based PL model. In this approach PLE  $\alpha_{x_j}$  is dependent on the location of receiver  $x_j$ , which is further determined by surrounding blockages around  $x_j$  in the environment.

Here, block-based means different regions may have different surrounding environment, e.g., locations around open space, like a green land or a square, would have a higher Line-of-Sight (LoS) probability and therefore the value of PLE is close to 2. On the contrary, locations among buildings or skyscrapers would be more likely to have high PLE values. In Figure 2 (left), the 14th district is divided into four blocks, where each block represents a different averaged PLE denoted by a different color. In the same block, the receivers may experience slightly different structures, and simultaneous mobile objects. Therefore, Figure 2 (middle) shows all PLE  $\alpha_x$  for locations of receivers  $x$ , which are generated from normal distribution  $\mathcal{N}(\mu, \sigma^2)$ , where  $\mu$  is given by the mean PLE of the block. Figure 2 (right) shows PLE distribution for receivers located in  $\alpha = 3.5$  block.



**Figure 2.** Block-based path loss (PL) model. **(Left):** Example of deterministic four-block based PLE model in the 14th district. **(Middle):** Map of the Block-based stochastic Path Loss Exponent (PLE) model. **(Right):** Receivers lie inside each block have normally-distributed PLE.

**Remark 1.** This division of the region into different blocks is not obtained from empirical measurements. This assumption provides a solution towards reproducing diverse PLE level, which depends on the obstacles in the environment. Other irregular blocks or clusters can be used as well, based on the empirical environments. Figure 2 shows a non-empirically-based example of a parametric setup for PLE. None of the PLE values are obtained from empirical values nor real city structure. The goal is to demonstrate how the parametric path loss model works.

**Remark 2.** Here in the present work, we do not consider the shadowing effect in the Equation (1), however, the block-based PL model is able to take it into account. The variations of PLE for receivers located in the same block can be interpreted as shadowing. The level of ‘shadowing’ can be controlled by tuning  $\sigma^2$  in normal distribution.

The received EMF follows the PL model denoted in Equation (1). This block-based PLE model follows the rule: no matter where the BSA is located, if the receiver is located at the block with certain PLE, this PLE is adopted in the PL model.

### 2.3. Time Variation

Due to the time variation of the traffic load, the total power transmitted by all BSA varies as well, which leads to the varying exposure levels at different times of day, especially in urban environments. Usually, the peak value of exposure is experienced around noon while the bottom is reached around

dawn every day. A typical wide-band time variation pattern can be found in Figure 3, from one sensor in an assay deployed by EXEM [7] in the 14th district of Paris. This periodical phenomenon shows a significant peak and trough around noon and dawn in a day, respectively.

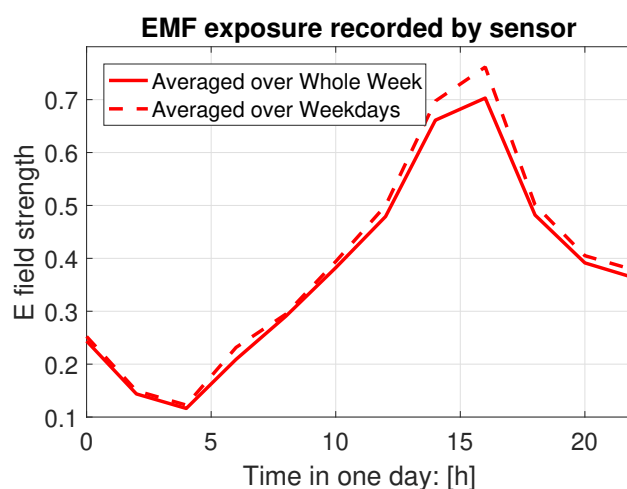


Figure 3. Time variation recorded by sensor network.

In the present paper, we use a trigonometric function to model the time variation pattern in the selected frequency band, since it can capture the most important features in the periodic pattern. Equation (2) gives the formula of exposure in terms of EMF after considering time variation.

$$\tilde{E}_{rx} = E_{rx} * f_t(t), \quad (2)$$

where  $f_t(t) = -0.4 * \sin(t) + 1$ .

**Remark 3.** Here, we assume the same time variation pattern for all BSA. Then the total EMF exposure with time variation can be simplified as previous exposure multiplied by the time variation function as seen in Equation (2). The sensor networks can record exposure for a whole day while the drive testing is carried out during working hours.

#### 2.4. Adding Noise to Exposure

As a vital part of wireless communication, the existence of background noise cannot be neglected. It includes thermal noise from electronic devices and signals originated from unlicensed frequency bands or distanced channels [24].

The additive white Gaussian noise (AWGN) is taken into account when we simulate the outdoor exposure. In addition, the noise is added to test the robustness of the neural network. Results from noiseless and with noise cases are compared and evaluated in Section 6.

### 3. Regression Neural Network Set Up

An ANN aimed at solving regression problems can be constructed as shown in Figure 4. Inputs  $x_1, x_2, \dots, x_N$  are selected from possible influential factor resulting in EMF exposure. Here,  $x_0$  is added as bias. The biases are added as an additional “weight” in each hidden layer, so that to adjust the output along with the weighted sum of the inputs to the neuron. The concept is similar but not limited to the intercept added in a linear regression. It should be noticed that this term is different from the idea of a statistical bias, in which a statistical estimation of the algorithm’s expected estimate of a quantity is not equal to the true quantity.

On the ANN in this paper, many factors contribute to the received signal strength at a certain location, e.g., the distance between the emitting and receiving antennas, their directions as well as

the city structure itself. After the inputs are fed to the ANN, it processes the information extracted from the inputs through hidden layers, where hidden layers are composed of neurons. Different hidden layers are connected by weights, e.g., neuron  $a_i^{(j)}$  from the  $j$ -th layer can be obtained by the following formula:

$$a_i^{(j)} = g \left( \theta_{i0}^{(j-1)} a_0^{(j-1)} + \theta_{i1}^{(j-1)} a_1^{(j-1)} + \dots + \theta_{in}^{(j-1)} a_n^{(j-1)} \right), \tag{3}$$

where  $g(x)$  represents the activation function. In the present paper, the activation function for all hidden layers is refined linear unit (ReLU), except for the output layer, where a pure linear function is used.  $n$  is the number of neurons in  $(j - 1)$ -th hidden layer.

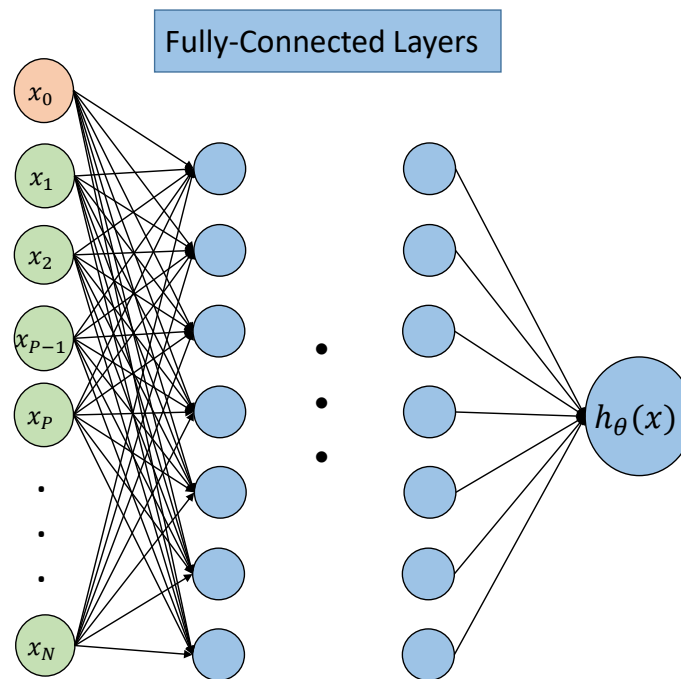


Figure 4. Structure of regression artificial neural networks (ANN).

After the forward propagation is done, the cost function is constructed as:

$$J(\theta) = -\frac{1}{m} \left[ \sum_{i=1}^m \sum_{k=1}^K y_k^{(i)} \log \left( h_\theta \left( x^{(i)} \right)_k \right) + \left( 1 - y_k^{(i)} \right) \log \left( 1 - h_\theta \left( x^{(i)} \right)_k \right) \right], \tag{4}$$

where  $K$  represents the total number of outputs that  $y^{(i)}$  has,  $m$  is the total number of training examples.

Then, to minimize the cost function in Equation (4), back propagation is performed by applying gradient descend method. To better evaluate the performance of the ANN, two metrics are used in the present paper, mean square error (MSE) and  $R^2$  [25]. MSE is used to minimize the residual sum of squares (RSS) and  $R^2$  indicates how close two sets of data are in terms of distribution. To be more specific,  $R^2$  is defined as:

$$R^2 = 1 - \frac{RSS}{TSS}, \tag{5}$$

where  $TSS = \sum (y_i - \bar{y})^2$  is the total sum of squares, and RSS is residual sum of squares, defined as  $RSS = \sum_{i=1}^m (y_i - \hat{y}_i)^2$ . TSS measures the total variance in the response  $Y$ . In contrast, RSS measures the amount of variability that is left unexplained after performing the regression. Consequently, when  $R^2 \rightarrow 1$ , a large proportion of the variability in the response has been explained by the regression [26].



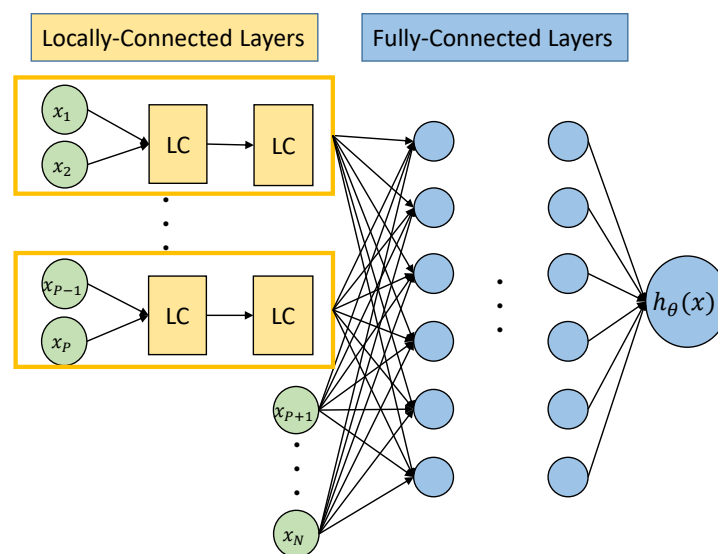
### Tuning Hyper-Parameters

While the ANN is set up, to make the most use of the ANN, tuning the hyper-parameters becomes a vital problem. Usually, this tuning process requires not only the experience and skills of the user, but also the mastery of features of data. The grid search method is used in the present paper to select the optimal combination of hyper-parameters with help from [27].

### 4. New Hybrid-Connected Neural Networks

In this section, we propose a new ANN structure with not-fully connected layers. After inspecting the physical nature of EMF exposure, we found that the exposure is aggregated from all effective BSA in the receiving range. Each BSA has its specifications, e.g., the azimuth, the location and the elevation. Therefore, to imitate and reproduce the real propagation, locally-connected layers are constructed in parallel to process the inputs from the same BSA. Then, fully-connected layers are constructed to process the concentrated outputs from locally-connected layers.

Figure 5 shows the structure of the hybrid ANN, inputs  $\{x_1, x_2\}$  and  $\{x_{p-1}, x_p\}$  are processed in independent locally-connected blocks and there are no single mutual inputs in different blocks. Then,  $\{x_{p+1}, \dots, x_N\}$  as added raw inputs to fully-connected layers, together with outputs from locally-connected layers. Here, the physical meaning of  $\{x_{p+1}, \dots, x_N\}$  can be interpreted as information, like time, location of the receiver, which remains same for different BSA in one training sample.



**Figure 5.** Structure of hybrid ANN.

**Remark 4.** This new structure of ANN is proposed to suit to the specific scenario, where inputs are clustered and organized by each BSA. The output is summed from all effective BSA. This structure may also be applied to cases, where inputs and outputs have similar relationships. In [28], a similar semi-parallel structure is used in convolutional neural networks. The advantage of hybrid connected ANN is that, it reduces unnecessary interactions between some neurons, which reduces the complexity of the network and improves the computation efficiency.

The proposed locally-connected ANN is not only capable of reproducing the predictions of exposure just as the conventional ANN, it can also do the job more efficiently. Validations and comparisons can be found in Section 6.



## 5. Reconstruction Methods

In this section, we explore two different reconstruction methods, in order to find balance between the cost of deploying sensors and the accuracy of predicting outdoor RF exposure.

### 5.1. Case 1: Through Sensor Networks

Since the sensor networks can continuously record EMF exposure for long time, the time variation pattern is fully available in the data recorded by the sensors. Reconstructing the spatial map of EMF exposure by sensor networks is first considered. Here, inputs of ANN consist of distances to 10 nearest BSA, azimuth of antennas, locations of receivers and time of the measurement.

In Figure 6, an example with 200 sensors selected from 3516 street lamps is given. The sensor based reconstruction method may bring a high cost problem in deploying real sensors. Therefore, in Section 6, the trade-off between required number of sensors and performance of the reconstruction is explored.

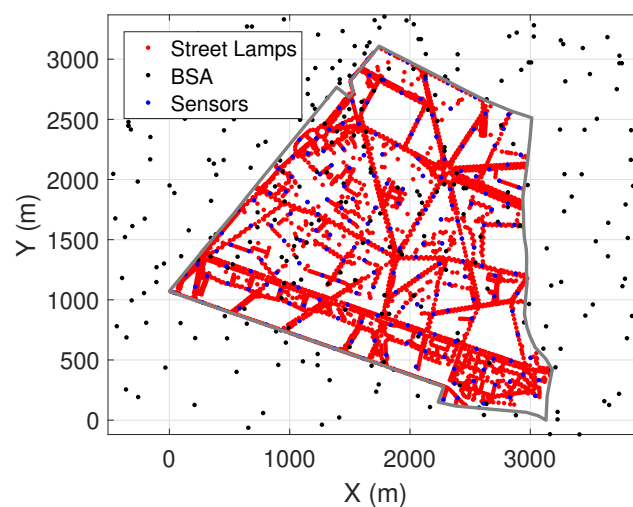


Figure 6. Sensor networks in the 14th district of Paris.

### 5.2. Case 2: Through Combined Drive Testing and Sensor Networks

Bearing in mind the potential high-cost limitation in sensor deployments, we propose a new reconstruction approach, which requires a lower number of sensors while maintaining a good reconstruction performance. We combine the one time drive testing and sensor networks together. Since the drive testing is done at different time of the day or several days and the peak of the periodic time variation usually occurs during noon in a typical day. This means that the drive testing may not be able to provide enough insights of the full-time variation pattern of the EMF exposure.

To be more specific, we keep the same BSA information available publicly as used in case 1, then, additional inputs extracted from sensor networks are considered as inputs to assist the reconstruction: the distance to nearest sensor and EMF exposure record by that sensor. The novelty of this reconstruction approach is the combination of all available information together, including sensors, drive testing and BSA information while less cost is required. The results comparing both methods can be found in Section 6.

## 6. Results

In this section, sustainable results obtained from ANN are presented to illustrate that the ANN is able to capture the main features in EMF exposure emitted from BSA. EMF exposure is reconstructed under the following two scenarios: (1) using sensor networks only, (2) combining drive testing and sensor networks.

The system model described in Section 2 is used in the simulation settings. Two scenarios with and without noise are compared in the results. Other system set up can be found in Table 1.

**Table 1.** System setup.

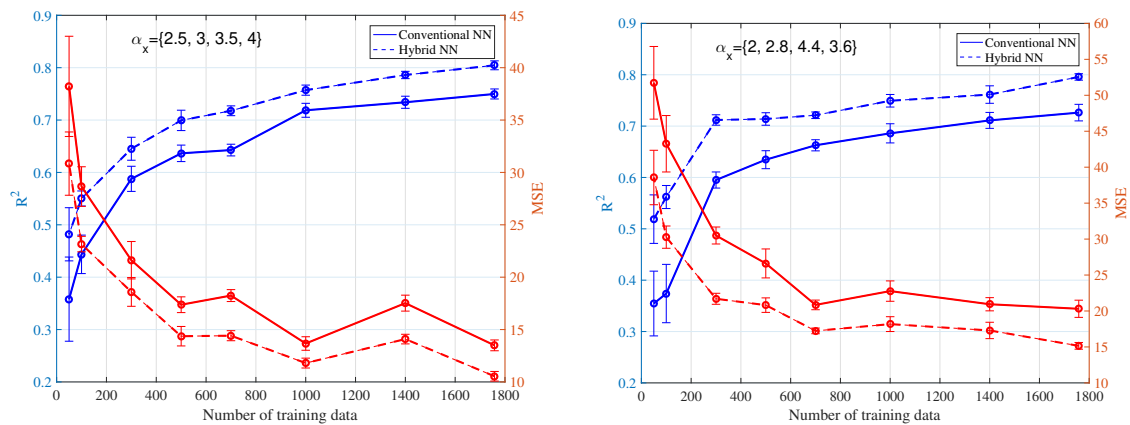
Total number of sensors in case 1	3516
Number of sensors selected in case 2	50
Additive Noise level	SNR = 15dB
PLE in Block-based PL Model	$\alpha_x = \{2.5, 3, 3.5, 4\}$ and $\{2, 2.8, 4.4, 3.6\}$
$\sigma^2$ in normal distribution	$\sigma^2 = \{0.1, 0.15\}$
$N_{training}$ , $N_{validation}$ and $N_{testing}$ in case 2	469, 201 and 670

The construction and training of ANN is carried out using Python version 3.7.3 on a PC under Tensorflow environment (GPU implemented on Nvidia Quadro RTX 8000 seed-Architecture Turin [29] will be used in the future work). Main packages used in the present paper are Keras and Sklearn. The hyper-parameters used in the present paper are shown in Table 2. The initial weight is determined from Glorot uniform initializer [30], which draws samples from a uniform distribution within [-limit, limit]. Here,  $limit = \sqrt{6 / (fan_{in} + fan_{out})}$ , where  $fan_{in}$  and  $fan_{out}$  represent the number of input units and output units in the weight tensor respectively. Biases are initialized to be zeros. Standardization is used to pre-process inputs. LC and FC represent locally-connected and fully-connected (layers).

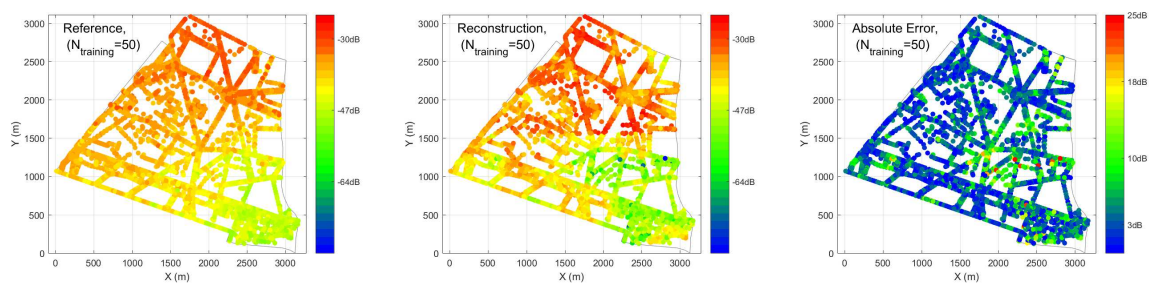
**Table 2.** Hyper-parameters in ANN.

Hyper-Parameters	Conv.-Case 1	Hybrid-Case 1	Conv.-Case 2	Hybrid-Case 2
Pre-processing	Standardization			
Num. of hidden layers	4	2(LC), 1(FC)	4	2(LC), 3(FC)
Num. of neurons	50	20(LC), 30(FC)	50	10(LC), 40(FC)
Learning rate	$1 \times 10^{-4}$			
Batch size	10			
Patience in early stopping	30			

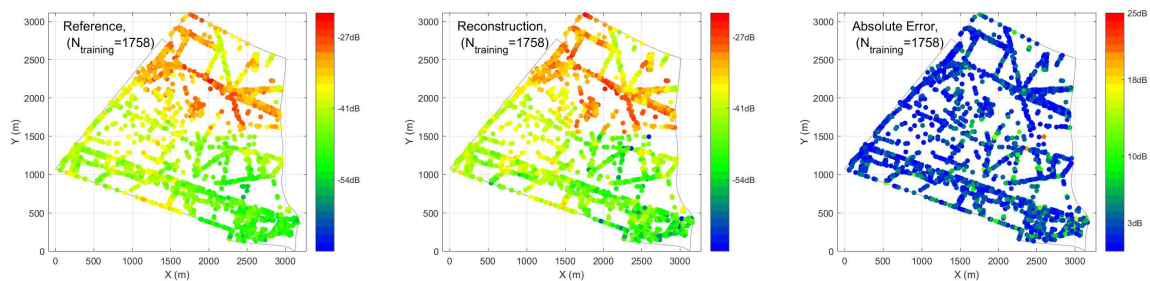
In case 1, different percentages of training and testing are explored. We have inputs of the ANN as: (1) distance to 10 nearest BSA, (2) location of the receiver, (3) angle between receiver and direction of maximum gain of antenna, (4) time of the sensor measurements. In Figure 7,  $R^2$  and MSE performance are compared under conventional and hybrid ANN, validated by two different sets of PLE. The hybrid ANN always outperforms the conventional one with different percentages of training and testing. With different percentage of training and testing, we are able to achieve at most  $R^2 = 0.80$  with 50% sensors (1758) required. When the number of training data decreases, the performance of the prediction also decreases as expected. In the worst case, where 50 sensors (1.4%) are used as training, the performance of  $R^2$  drops to less than 0.5 with high variation. Figures 8 and 9 illustrate reconstructed EEM, ERM and absolute error maps under  $N_{training} = \{50, 1758\}$  cases, it is clear that decreasing training examples would degrade the prediction of ANN significantly. However, a deployment of a larger number of sensors would imply a higher cost for installation and maintenance, thus leading to an impractical situation.



**Figure 7.** Averaged mean square error (MSE) and  $R^2$  with standard deviation (STD) based on sensor network inputs.



**Figure 8.** Exposure Reference Map (ERM), reconstructed electromagnetic field exposure map (EEM) and absolute error maps obtained from hybrid ANN for  $N_{training} = 50$  cases.



**Figure 9.** ERM, reconstructed EEM and absolute error maps obtained from hybrid ANN for  $N_{training} = 1758$  cases.

In case 2, sensor information is used to assist prediction done by drive testing. In addition to the inputs we considered in case 1, we add two more information feeds as inputs for the ANN: distance to the nearest sensor and EMF exposure information recorded by nearest sensor. We assume 50 sensors are installed on selected street lamps denoted in green markers in Figure 10. 50% of the drive testing are selected as training data of the ANN, the remaining 50% are used for testing, which is used to produce results in Table 3.

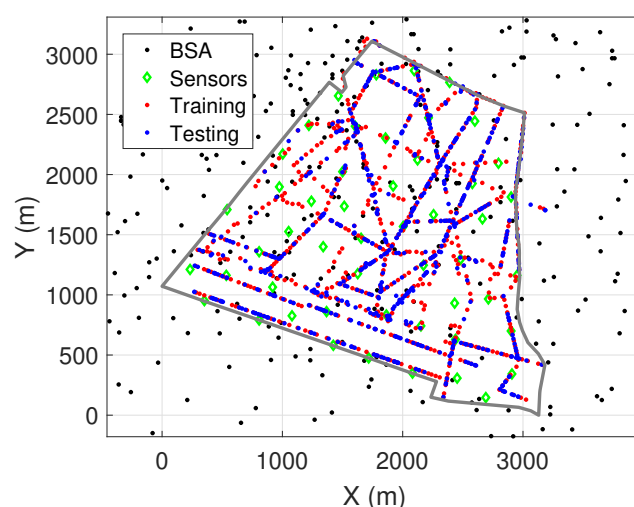


Figure 10. Selection of training and testing data.

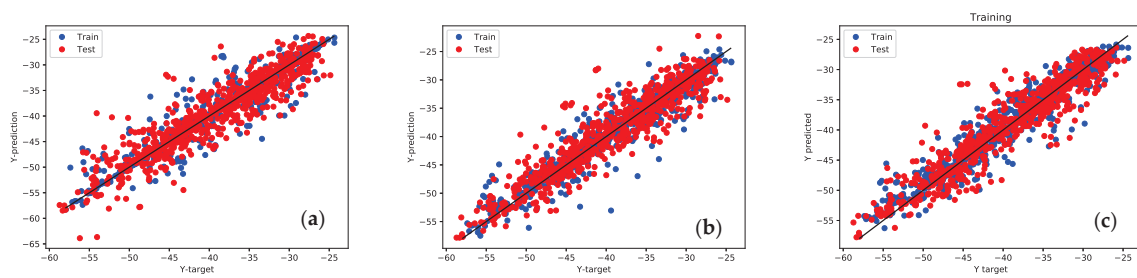
In Table 3, MSE and  $R^2$  are used as the main metrics to evaluate the performance of ANN from different scenarios. Both MSE and  $R^2$  are obtained by averaging 10 times. Results produced from conventional and hybrid ANN are compared, where a significant improvement in terms of  $R^2$  can be found between conventional and hybrid ANN. Furthermore, the number of trainable parameters in Table 3 represents the parameters that could learn, which is proportional to execution time.

Table 3. Performance of different ANN in case 2

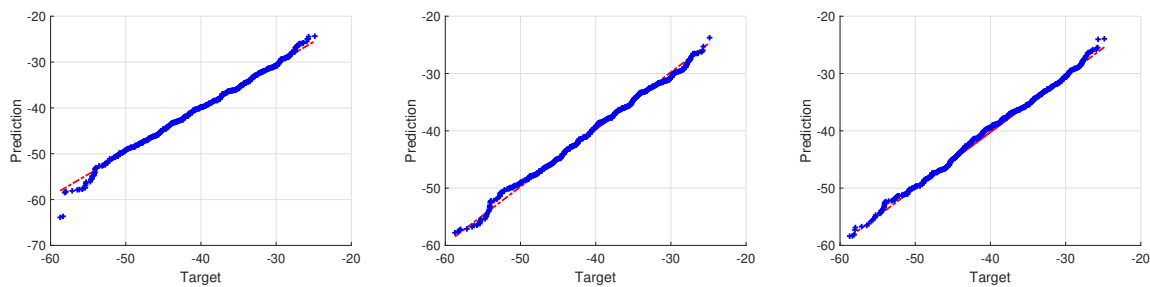
Scenarios	Num. of Para.	MSE ( $\pm$ STD)	$R^2$ ( $\pm$ STD)
Conventional ANN without considering sensors	8851	11.2 ( $\pm$ 0.91)	0.81 ( $\pm$ 0.02)
Conventional ANN considering sensors	8951	9.2 ( $\pm$ 0.73)	0.85 ( $\pm$ 0.01)
Hybrid ANN with considering sensors	5431	7.9 ( $\pm$ 1.06)	0.87 ( $\pm$ 0.02)
Conventional ANN without considering sensors with noise	8851	16.3 ( $\pm$ 0.85)	0.72 ( $\pm$ 0.01)
Conventional ANN considering sensors with noise	8951	14.6 ( $\pm$ 0.76)	0.75 ( $\pm$ 0.01)
Hybrid ANN without considering sensors with noise	5431	12.8 ( $\pm$ 0.96)	0.78 ( $\pm$ 0.02)

Figure 11 shows the performance of upper three scenarios in case 2 mentioned in Table 3, where we compare conventional ANN with proposed hybrid ANN. The closer blue and red dots get to the black diagonal, the more accurate predictions are. The training data and testing data have a similar degree of performance, which means there are no over-fitting problems. It is clear that dots are more concentrated in the figure on the right, which is from hybrid ANN. Figure 12 showing a Quantile-to-Quantile (Q-Q) plot from conventional ANN with proposed hybrid ANN as well. If targets and predictions are closer to each other, a linear relationship can be observed from the plotted quantiles. From all three figures, an overall linear relationship is observed with details improved around minimum values region in the figures from left to right. It aligns with our results in Table 3 and Figure 11.

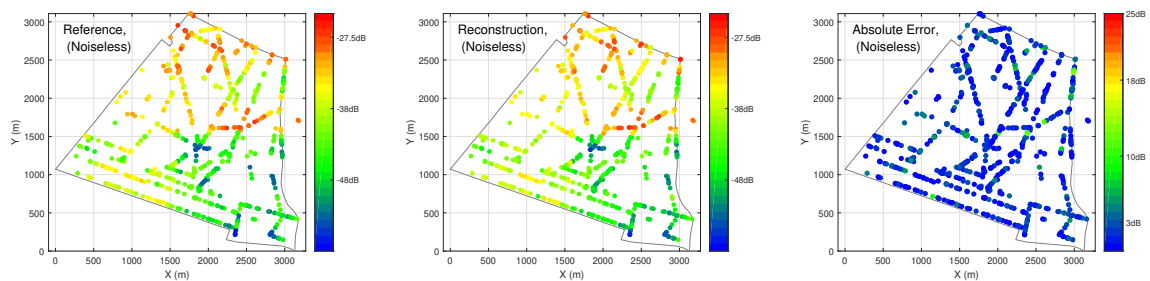
Figure 13 shows a noiseless case, EMF exposure can be well reconstructed by using hybrid ANN, where a large majority of dots in blue are observed in the figures, showing relatively small error from predictions. Figure 14 tests the robustness of ANN in dealing with noisy data. The results are obtained using hybrid ANN as well, which corresponds to the last scenario in Table 3. After adding noise to data, the hybrid ANN is still able to predict EMF exposure with  $R^2 = 0.78$ .



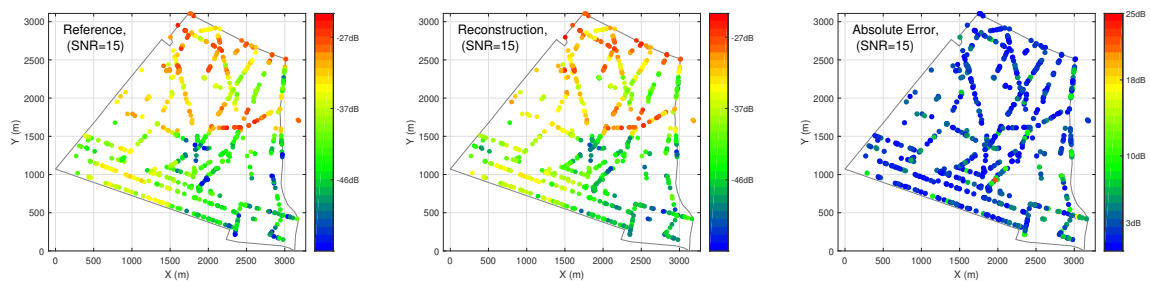
**Figure 11.** Scattering plot between predictions and targets from ANN. From left to right: (a) conventional ANN without considering sensor information as inputs of ANN; (b) conventional ANN considering sensor information as inputs of ANN; (c) hybrid ANN considering sensor information as inputs.



**Figure 12.** Quantile-to-Quantile (Q-Q) plot between predictions and targets from ANN. From left to right: (1) conventional ANN without considering sensor information as inputs of ANN; (2) conventional ANN considering sensor information as inputs of ANN; (3) hybrid ANN considering sensor information as inputs.



**Figure 13.** Spatial maps of EMF exposure obtained from hybrid ANN for noiseless case.



**Figure 14.** Spatial maps of EMF exposure obtained from hybrid ANN for noisy case (SNR = 15).

## 7. Discussion

There are two cases considered in Section 6, where case 1 only uses information from sensor networks and case 2 combines drive testing with aid from sensor networks. It is shown that a sufficient number of sensors are required to reconstruct spatial map of EMF exposure with good

prediction. However, due to the cost of deploying real sensors, it is not ideal to adopt this method. While in the second case, only 50 sensors are required as extra information to assist the assessment of EMF exposure. The results show maximum  $R^2 = 0.87$  and  $R^2 = 0.78$  can be achieved with one-time drive testing in the noiseless and noisy scenarios. It shows the robustness of the ANN, that both cases with and without presence of noise can be assessed by ANN with good predictions.

Second, we prove that with new proposed hybrid ANN, better prediction of EMF exposure can be obtained for both cases. In case 2, we are able to assess the EMF exposure from certain frequency with  $R^2$  up to 0.87 under noiseless scenarios. It represents more than 87% of the targets can be explained using predictions from ANN. Furthermore, the hybrid ANN has better performance than the conventional ANN while less trainable parameters are required.

In the future work, real drive testing and sensor measurements will be used in the ANN to train and predict EMF exposure. Furthermore, an empirical based block-based PL model will be introduced, based on the real city structure, e.g., building height, width, road width. Since the current ANN work is based on the Tensorflow and Python environment, where GPU can boost the efficiency of computing, we will use GPU implemented on Nvidia Quadro RTX 8000 seed-Architecture Turin [29] in the future work. We believe that the methodology proposed in this paper is also applicable to 5G networks. The 5G network system model can be very different from normal 4G network, with new features, like massive multiple-input multiple-output (MIMO) and different frequency spectrum. However, the aggregating nature of EMF exposure from all effective BSA nearby will not change.

## 8. Conclusions

The time and spatial mapping of EMF exposure induced by BSA using ANN is studied in this paper. Sensor networks installed on street lamps in the 14th district of Paris, drive testing and information of BSA which is publicly available are used to aid the reconstruction. The reference EMF exposure is built through simulations by considering a block-based stochastic PL model and time variation. We proposed a new efficient hybrid ANN structure and results show it outperforms the conventional ANN. First, reconstruction through sensor networks only containing different number of sensors are explored. We found that a quite large number of sensors is required for good reconstruction of EMF exposure map, which brings problems of high-cost in deployment and maintenance. Second, the reconstruction method through a combination of EMF sensors measurements and one-time drive testing has been employed. The results show that using a combination method, good prediction can be achieved with much fewer sensors required. Additional noise is considered to test the robustness of ANN.

**Author Contributions:** S.W., J.W. contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Acknowledgments:** This study is partially financed by the French Agency ANSES (Food, Environmental and Occupational Health and Safety) (grant EST-2016-2RF-04) and has contributed to the AMPERE project. The authors would like to thank EXEM for sensor deployments in the 14th district of Paris. We also would like to thank Professor Masao Taki from NICT Japan, Professor Laurent Clavier from IMT Lille Douai and Philippe De Doncker from Université libre de Bruxelles for the fruitful discussions.

**Conflicts of Interest:** The authors declare no conflict of interest.



## Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
ANN	Artificial Neural Network
AWGN	Added White Gaussian Noise
BSA	Base Station Antennas
EEM	EMF Exposure Map
EMF	Electromagnetic Field
ERM	Exposure Reference Map
FC	Fully-Connected
LC	Locally-Connected
LoS	Line-of-Sight
MSE	Mean Square Errors
NLoS	Non Line-of-Sight
PL	Path Loss
PLE	Path Loss Exponent
Q-Q Plot	Quantile-to-Quantile Plot
reLU	Refined Linear Unit
RF	Radiofrequency
RSS	Residual Sum of Squares
STD	Standard Deviation

## References

1. Electromagnetic Fields Report. Available online: [https://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs\\_347\\_en.pdf](https://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_347_en.pdf) (accessed on 25 April 2020).
2. Wiart, J. *Radio-Frequency Human Exposure Assessment: From Deterministic to Stochastic Methods*; John Wiley & Sons: Hoboken, NJ, USA, 2016.
3. Etude de l'exposition du Public aux Ondes Radioélectriques. Available online: <https://www.anfr.fr/fileadmin/mediatheque/documents/expace/20180919-Analyse-mesures-2017.pdf> (accessed on 25 April 2020).
4. Gajšek, P.; Ravazzani, P.; Wiart, J.; Grellier, J.; Samaras, T.; Thuróczy, G. Electromagnetic field exposure assessment in Europe radiofrequency fields (10 MHz–6 GHz). *J. Expo. Sci. Environ. Epidemiol.* **2015**; *25*, 37–44.
5. Tesanovic, M.; Conil, E.; De Domenico, A.; Aguero, R.; Freudenstein, F.; Correia, L.M.; Bories, S.; Martens, L.; Wiedemann, P.M.; Wiart, J. The LEXNET project: Wireless networks and EMF: Paving the way for low-EMF networks of the future. *IEEE Veh. Technol. Mag.* **2014**; *9*, 20–28.
6. Diez, L.; Agüero, R.; Muñoz, L. Electromagnetic Field Assessment as a Smart City Service: The Smartsantander Use-Case. *Sensors* **2017**, *17*, 1250. [CrossRef] [PubMed]
7. Philippe, P.; Pascal, T.; Yannick, P.; Lamine, O. Observatoire des Ondes, une Réponse au Débat Sociétal EMF Observatory, an Answer to the Societal Debate. JS URSI France. 2020. Available online: <https://ursifr-2020.sciencesconf.org/305807/document> (accessed on 25 April 2020).
8. Observatoire des Ondes. Available online: <https://observatoiredesondes.fr/> (accessed on 25 April 2020).
9. Huang, Y.; Wiart, J. Simplified assessment method for population RF exposure induced by a 4G network. *IEEE J. Electromagn. Microwaves Med. Biol.* **2017**, 34–40, doi:10.1109/JERM.2017.2751751. [CrossRef]
10. Joshi, P.; Colombi, D.; Thors, B.; Larsson, Lars, E.; Törnevik, C. Output power levels of 4G user equipment and implications on realistic RF EMF exposure assessments. *IEEE Access* **2017**, *5*, 4545–4550. [CrossRef]
11. Aerts, S.; Deschrijver, D.; Verloock, L.; Dhaene, T.; Martens, L.; Joseph, W. Assessment of outdoor radiofrequency electromagnetic field exposure through hotspot localization using kriging-based sequential sampling. *Environ. Res.* **2013**, *126*, 184–191. [CrossRef] [PubMed]
12. Lemaire, T.; Wiart, J.; De Doncker, P. Variographic Analysis of Public Exposure to Electromagnetic Radiation Due to Cellular base Stations. *Bioelectromagnetics* **2016**, *37*, 557–562. [CrossRef] [PubMed]
13. Prachi, K.N.; Matta, G. Artificial neural network applications in air quality monitoring and management. *Int. J. Environ. Rehabil. Conserv.* **2011**, *2*, 30–64.



14. Singh, K.P.; Basant, A.; Malik, A.; Jain, G. Artificial neural network modeling of the river water quality—A case study. *Ecol. Model.* **2009**, *220*, 888–895. [[CrossRef](#)]
15. Rozenblit, O.; Haddad, Y.; Mirsky, Y.; Azoulay, R. Machine learning methods for SIR prediction in cellular networks. *Phys. Commun.* **2018**, *31*, 239–253. [[CrossRef](#)]
16. Zappone, A.; Di Renzo, M.; Debbah, M.; Lam, T.T.; Qian, X. Model-aided wireless artificial intelligence: Embedding expert knowledge in deep neural networks for wireless system optimization. *IEEE Veh. Technol. Mag.* **2019**, *14*, 60–69. [[CrossRef](#)]
17. Ostlin, E.; Zepernick, H.-J.; Suzuki, H. Macrocell path-loss prediction using artificial neural networks. *IEEE Trans. Veh. Technol.* **2010**, *59*, 2735–2747. [[CrossRef](#)]
18. Popoola, S.I.; Adetiba, E.; Atayero, A.a.; Faruk, N.; Calafate, C.T. Optimal model for path loss predictions using feed-forward neural networks. *Cogent Eng.* **2018**, *5*, 1444345. [[CrossRef](#)]
19. Mom, J.M.; Mgbe, C.O.; Igwe, G.A. Application of artificial neural network for path loss prediction in urban macrocellular environment. *Am. J. Eng. Res.* **2014**, *3*, 270–275.
20. Aerts, S.; Huang, Y.; Martens, L.; Joseph, W.; Wiart, J., Use of artificial intelligence to model exposure to radiofrequency electromagnetic fields based on sensor network measurements. In Proceedings of the 4th Workshop on Uncertainty Modeling for Engineering Applications (UMEMA 2018), Split, Croatia, 10–11 December 2018.
21. ANFR-Cartoradio. Available online: <https://www.cartoradio.fr> (accessed on 25 April 2020).
22. Friis, H.T. A note on a simple transmission formula. *Proc. IRE IEEE* **1946**, *34*, 254–256. [[CrossRef](#)]
23. Erceg, V.; Greenstein, L.J.; Tjandra, S.Y.; Parkoff, S.R.; Gupta, A.; Kulic, B.; Julius, A. A and Bianchi, Renee, An empirically based path loss model for wireless channels in suburban environments. *IEEE J. Sel. Areas Commun.* **1999**, *17*, 1205–1211. [[CrossRef](#)]
24. Fu, B.; Bernáth, G.; Steichen, B.; Weber, S. Wireless background noise in the Wi-Fi spectrum. In Proceedings of the IEEE 4th International Conference on Wireless Communications, Networking and Mobile Computing, Dalian, China, 12–17 October 2008; pp. 1–7.
25. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
26. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning*; Springer: Berlin, Germany, 2013.
27. Bengio, Y. *Practical Recommendations for Gradient-Based Training of Deep Architectures*. *Neural Networks: Tricks of the Trade*; Springer: Berlin, Germany, 2012; pp. 437–478.
28. Bazrafkan, S.; Javidnia, H.; Lemley, J.; Corcoran, P. Semi-Parallel Deep Neural Network (SPDNN) Hybrid Architecture, First Application on Depth from Monocular Camera. *J. Electron. Imaging* **2018**. [[CrossRef](#)]
29. CARRI Systems. Available online: <https://www.carri.com> (accessed on 25 April 2020).
30. Glorot, X.; Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the 13th International Conference on Artificial Intelligence and Statistics, Sardinia, Italy, 13–15 May 2010; pp. 249–256.

