Journal of Radiological
Protection

PAPER • **OPEN ACCESS**

Generic neural network model for estimating exposure levels in industrial environments

To cite this article: David Plets *et al* 2026 *J. Radiol. Prot.* **46** 021518

View the article online for updates and enhancements.

You may also like

- Estimation Downward Longwave Radiation Using Neural Networks in Baghdad City
Weaam M. Al-Sibahee and Ali M. AL-Salihi

- An improved procedure to accurately assess the variability of the exposure to electromagnetic radiation emitted by GSM base station antennas
Paul Bechet and Simona Miclaus

- Solution of the Applied Task to Determine Total Exposure for Mechanical Engineering Workers
O V Medvedeva, K N Orlova and I F Borovikov

This content was downloaded from IP address 146.103.254.11 on 02/09/2026 at 09:13

IOP Publishing J. Radiol. Prot. 46 (2026) 021518 https://doi.org/10.1088/1361-6498/ae6c32

Journal of Radiological Protection

Check for updates

OPEN ACCESS

RECEIVED
10 February 2026

REVISED
5 May 2026

ACCEPTED FOR PUBLICATION
12 May 2026

PUBLISHED
22 May 2026

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence.

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

PAPER

Generic neural network model for estimating exposure levels in industrial environments

David Plets[1,*], Christos Apostolidis[2,*], Shanshan Wang[3], Blaž Valič[4], Luc Martens[1], Theodoros Samaras[2] and Peter Gajšek[4]

[1] Ghent University—imec-WAVES, Tech Lane Science Park 126, 9052 Ghent, Belgium
[2] Center for Interdisciplinary Research and Innovation, Aristotle University of Thessaloniki, 57001 Thermi, Greece
[3] Chaire C2M, LTCI, Télécom Paris, Institut Polytechnique de Paris, 91120 Palaiseau, France
[4] Institute of Non-Ionizing Radiation (INIS), 1000 Ljubljana, Slovenia
[*] Authors to whom any correspondence should be addressed.

E-mail: david.plets@ugent.be and capos@physics.auth.gr

Keywords: EMF RF exposure, industrial environments, validation, industry 4.0, neural networks, modelling

Abstract

This study describes a neural network-based method for estimating exposure levels in industrial environments, without requiring detailed technical inputs, allowing usage of the model by layman people or by workers active in these areas. A pipeline based on Blender environments and MATLAB ray-tracing simulations is created and after defining a set of 11 candidate input parameters for the model, more than 20 000 different wireless configurations are simulated, varying the different environmental and wireless input parameters. A correlation analysis shows that main inputs influencing the exposure levels in the industrial area are the transmit power of the antennas, the density of clutter in the area, the density of transmitters in the area, and the height and location of the transmitters. A multi-layer fully connected neural network regression model is developed to predict median (E_{50}) and 95th percentile (E_{95}) exposure levels in industrial areas. Testing the obtained model on an unseen dataset of environments with E_{50} values between 0 and 3.25 V m^{-1} and E_{95} values between 0 and 7 V m^{-1}, demonstrates the good prediction performance of the model: root-mean-square error values below 0.173 V m^{-1} and R^2 values above 95% are obtained. Subsequently, the model is validated with measurement data collected in three distinct realistic industrial environments. The average absolute deviation of the model predictions with respect to the measurements is limited to 20.4%. This novel and broadly accessible approach demonstrates that it is possible to reliably estimate exposure levels in realistic environments without having to rely on external experts or on dedicated complex software.

1. Introduction

The increased adoption of reliable wireless technologies in recent years, especially under the umbrella of Industry 4.0 and 5.0 [Islam], has revolutionised industrial operations. In modern factories, smart sensors, industrial Internet-of-Things devices, and autonomous systems increasingly rely on Wi-Fi, 5G, and other radio-frequency (RF) communications [Artetxe] to enable real-time data exchange, predictive maintenance, flexible automation, or human-centric wireless support applications such as augmented reality [Majid]. While these advances improve productivity and efficiency, they often require a dense wireless infrastructure, raising important concerns about RF electromagnetic field (EMF) exposure in indoor industrial environments, where workers may be persistently exposed to elevated field strengths. In recent years, most exposure-assessment research has focused on residential or office environments; comprehensive studies addressing industrial scenarios remain relatively scarce. For instance, [Joseph] conducted *in-situ* measurements across diverse environments, including industrial sites, observing median exposure in industrial areas around 0.49 V m^{-1}. A private standalone (SA) 5G network operating at 3.7 GHz in a real industrial environment was numerically modelled in [Gajšek] and compared with

© 2026 The Author(s). Published on behalf of the Society for Radiological Protection by IOP Publishing Ltd

in-situ RF EMF measurements. More recent work highlights that confined industrial spaces—such as tunnels, subways, or factory floors—present challenges in accurate exposure estimation due to reflections, metallic structures, and spatial heterogeneity [Ashraf]. Other efforts combine sensor networks with spatial interpolation techniques (e.g. Kriging) to generate high-resolution exposure maps in complex indoor environments [Jabeur]. Beyond classical measurement and interpolation approaches, the use of machine learning and neural network approaches has emerged for estimating RF exposure from small-size data sources. For instance, in [Wang], the authors used artificial neural networks to reconstruct exposure maps in urban settings by leveraging sparse sensor data, drive-testing, and public base-station databases. The work in [Zhang] uses a.o., building height maps, antenna radiation maps and land cover maps as inputs to a deep learning scheme for exposure estimation. In [Mallik], a comparative study was presented, using finite- and infinite-width convolutional neural networks to estimate realistic EMF exposure from a network of low-density sensors. Kiouvrekis *et al* attempted to construct electric-field maps of urban areas. They used explainable AI based on the calculation of SHAP values, revealing that the total volume of buildings in the vicinity of antennas had a large influence on exposure levels [Kiouvrekis].

Due to the growing interest in wireless industrial applications and increasing concerns among both the general public and workers in industrial environments, research efforts are increasingly focused on communicating exposure assessment results. However, such information is necessarily generalised and provides limited insight into the actual RF exposure levels at the specific industrial sites. As shown in the literature overview, current methods rely either on expensive *in-situ* measurement campaigns or on simulations using specialised electromagnetic modelling software with detailed input parameters [Hirata]. In most cases, industrial site owners are not willing to commission measurement campaigns, nor do they have the required information or expertise to perform site-specific simulations. Therefore, the goal of this paper is to investigate to what extent a limited set of input parameters—understandable and readily accessible to a layman user or employer—can be used to obtain a reliable estimate of the electric-field strength in a given industrial environment. State-of-the-art numerical modelling and simulation tools were used in conjunction with machine learning approaches. Trained and tested on a large number of simulation scenarios and a selected set of experimental datasets, it provides a foundation for informing industrial workers on expected exposure levels, without requiring complex software simulations or technical knowledge of the wireless setup. Specific contributions are the following (1) identification of configuration parameters that have most predictive power for exposure levels in the area, (2) creation of more than 20 000 wireless industrial environments that are used to train, validate, and test a neural network model, (3) experimental validation of the created model on measurement datasets collected in different countries. The created datasets are made publicly available for the exposure research community to add datasets, finetune the model or create new models.

2. Materials and methods

This section will describe the approach that was followed to create and validate the generic model. First, we select a list of candidate input parameters that are expected to relate to the exposure levels in the area. Next, a large set of simulations of wireless configurations is conducted using a pipeline of Blender [Blender] and MATLAB software. To find a good balance between calculation complexity and calculation time, we determine optimal simulation settings. Based on first simulation results, we determined the correlation values of the various input and output parameters. After retaining the input parameters with the largest predictive power, we extended the exploration of the input parameter space and built, validated, and tested a machine learning-based prediction model. Finally, we experimentally validated the obtained model. Below, we elaborate on the different steps of the approach and their building blocks. Figure 1 shows a high-level overview of the flow graph for creating the generic model.

2.1. Model input and output parameters

First, we define candidate input parameters, as well as the output parameters. Note that we only consider input variables that should be relatively easy for laymen to understand and obtain, without assuming prior technical knowledge on antennas, propagation, or exposure. Each prediction instance is represented by [$\boldsymbol{X}$ $\boldsymbol{Y}$], with $\boldsymbol{X}$ the vector of input parameters and $\boldsymbol{Y}$ the vector of output parameters. Initially, $\boldsymbol{X}$ is defined as a 11-dimensional vector ($X1, X2, \ldots, X11$), and $\boldsymbol{Y}$ as a 2-dimensional vector ($Y1, Y2$). The **input parameters** were the following:

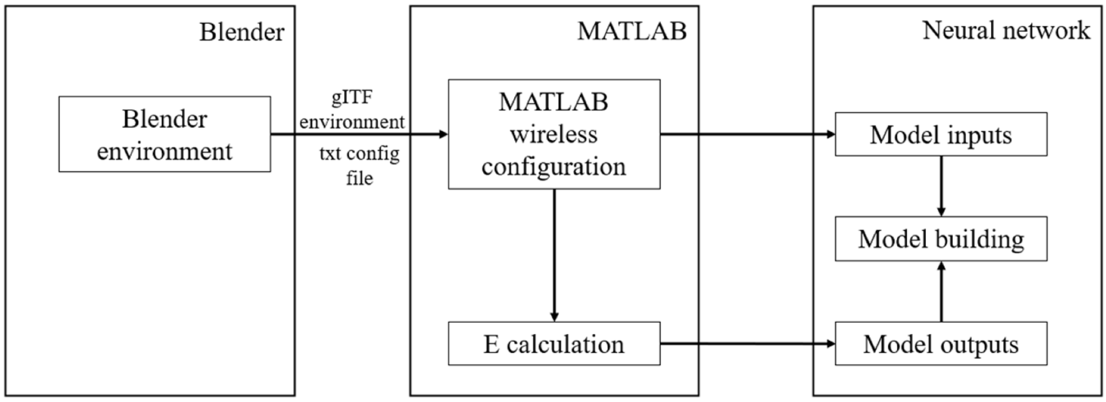


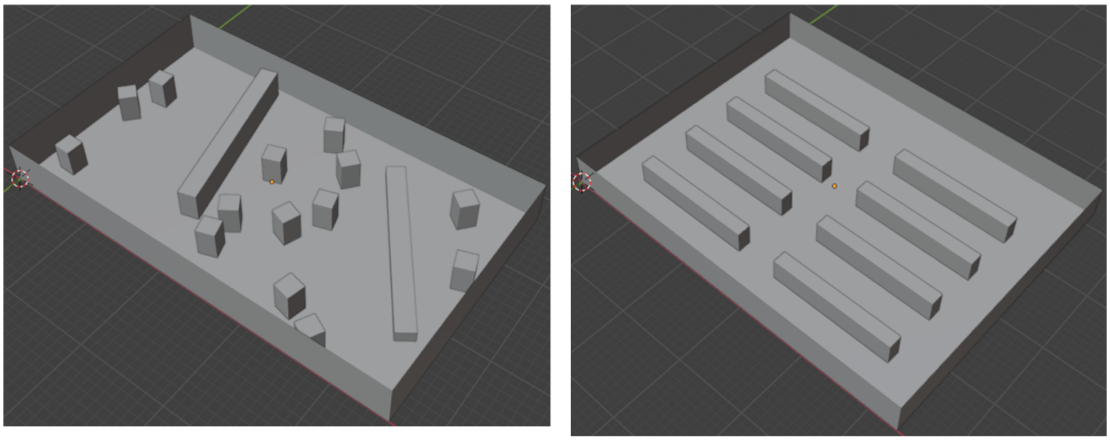
Figure 1. Flow graph for creating the generic model.

- $X1$ = Area: the total wirelessly covered evaluation area of the factory (in m^2, real value $\epsilon]0,+\infty[$)
- $X2$ = No_Tx: the number of transmit sources or base stations (BS) in the evaluation area (unitless, integer value ϵ {1,2,3,…}).
- $X3$ = Den_Tx: the density of transmitters, defined as the number of square meters of area per deployed transmitter, i.e. $X1/X2$ (m^2, real value $\epsilon]0,+\infty[$).
- $X4$ = Freq: the operation frequency used by the base stations (GHz, real value ϵ {3.5, 2.4, 5.0, …}).
- $X5$ = H_Tx: the height of the base station antennas above the ground level (m, real value $\epsilon]0,+\infty[$).
- $X6$ = P_Tx: the input power to the antennas of the base stations (dBW, real value $\epsilon]-\infty,+\infty[$).
- $X7$ = Type_Ant: the antenna type, either on the ceiling (i.e. omnidirectional, 2.15 dBi gain) or on the wall (i.e. half-omnidirectional antenna covering a 180° sector, 4.53 dBi gain), internally converted to the corresponding antenna gain (dBi, real value ϵ {2.15, 4.53, …}).
- $X8$ = Den_Clutter: the percentage of floor area covered by clutter/objects (real value real value ϵ $[0,1[$).
- $X9$ = Mat_Clutter: the clutter material type that is dominant in the area (wood, metal, water, concrete) internally defined as the material's coefficient of reflection (unitless, real value ϵ $[0,1]$)
- $X10$ = Mat_Walls: the dominant outer wall material of the area (glass, metal, concrete, wood, brick, drywall), internally defined as the material's coefficient of reflection (unitless, real value ϵ $[0,1]$)
- $X11$ = H_Ceiling: the height of the ceiling (m, real value $\epsilon]3,+\infty[$)

As **output parameters**, we assumed the following two parameters, allowing to assess the exposure distribution in the area:

- $Y1 = E_{50}$: the median electric field value over the considered area where no clutter is present.
- $Y2 = E_{95}$: the maximal electric field value, defined as the 95th percentile E value over the considered area where no clutter is present.

In an actual usage scenario, it might be advisable to translate the median and/or E_{95} value to a textual/descriptive formulation for the employer (negligible, low, high,…). We opted for median and 95th percentile E values as output instead of mean E values, as exceptionally high (simulated) E values close to the antennas could (wrongly) draw the mean E value to unrealistically high values. Note that a constant base station height and transmit power is considered over all involved base stations. Electric-field strength was preferred over individual dose metrics [El Habachi, Jiang], as the latter require detailed information on body morphology, posture, and orientation, which is typically not available for generic industrial scenarios. In addition, electric-field values can be directly compared to established (inter)national exposure limits and guidelines, making them more suitable for practical assessment and



(a) (b)

Figure 2. Environments created with python script in Blender: (a) randomly positioned clutter, (b) regularly positioned clutter.

communication purposes. While individual dose metrics may provide more subject-specific information, their estimation generally requires significantly more detailed modelling and input parameters, which falls outside the scope of this work.

2.2. Simulations

To investigate the correlation between inputs and outputs and to eventually build the model, simulation configurations were executed, exploring the input parameter space. These simulations were executed in MATLAB, based on environments created in Blender. This section describes how this dataset was built.

2.2.1. Blender environment creation

Blender v4.4 was used to generate the 3D indoor environments. Blender allows the creation of detailed geometric models and the assignment of material properties to individual objects, enabling an accurate representation of the geometry and electromagnetic characteristics required for the simulations. The environments were exported in glTF format, which can be imported directly into MATLAB, where the electromagnetic simulations were performed.

Using Blender's Python API, a Python script was developed to automatically create the indoor environments. Each indoor space's length and width were randomly selected from a specified range, while its height was chosen from a set of predefined values (see section 2.2.3 Simulation scenarios). Non-overlapping clutter objects with a generic rectangular form were generated and positioned both at random locations (see figure 2(a)) or in a typical aisle-like fashion to resemble typical industrial environments with racks (see figure 2(b)). The number and dimensions of the clutter objects that were generated were determined by the indoor space's dimensions and were randomly set in a way that fulfilled the requirement to create environments with a clutter density ranging from 0% to over 50%. Materials were subsequently assigned to the clutter objects and the outer walls.

Each completed environment was defined by the dimensions and materials of its floor ($X1$), ceiling ($X11$), outer walls ($X10$), and clutter objects ($X8$, $X9$). Along with the 3D model, a text file containing the geometric characteristics of the environment (e.g. dimensions, clutter positions) was also produced. This file was subsequently used to define parameters such as the number and placement of the base stations, and the positions of the receiver grid points in the electromagnetic simulations.

2.2.2. MATLAB electric-field calculation

For the electromagnetic simulations, a numerical tool was developed in MATLAB in which the wireless configuration parameters were defined and dynamically adjusted to cover the full range of the model's input variables (see section 2.2.3 Simulation scenarios). The values for the number of base stations ($X2$) and their frequency ($X4$), height ($X5$), transmit power ($X6$), and antenna type ($X7$) were set accordingly. The density of sources ($X3$) was calculated from the size of the environment ($X1$) and the number of base stations ($X2$). The transmitter positions were dynamically assigned based on their number, the environment's dimensions, and the placement of clutter, to obtain realistic deployments. Finally, the grid points at which the electric field was calculated were determined by the environment's area and the

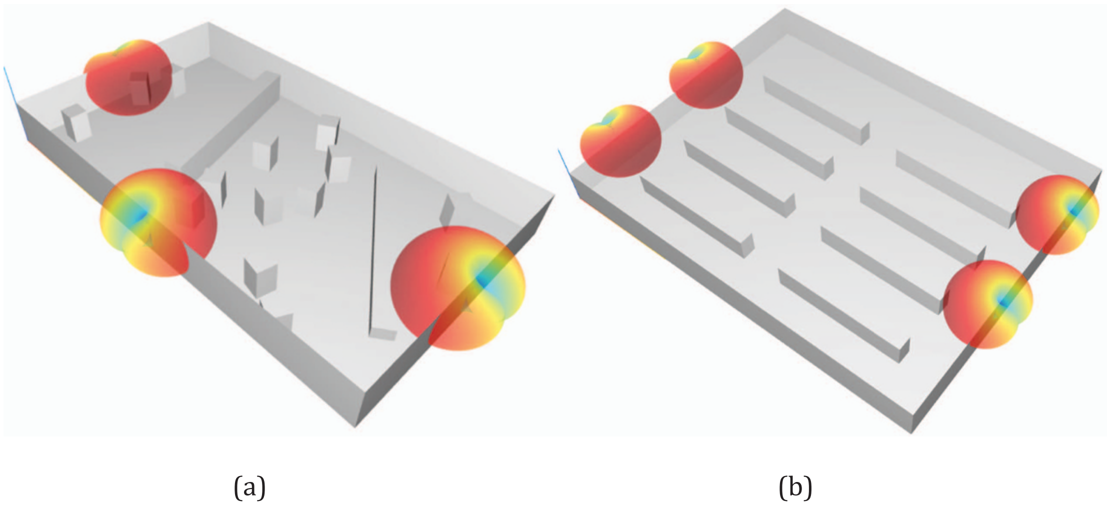


Figure 3. (a) Environment with randomly positioned clutter and 3 base stations, (b) Environment with regularly positioned clutter and 4 base stations. All base stations are wall-mounted in these cases.

regions occupied by clutter. Figure 3 shows the same two environments as in figure 2, with the added base stations.

For the electric-field calculation, the numerical tool utilised MATLAB's built-in ray tracing propagation model to identify the line-of-sight and non-line-of-sight (NLOS) propagation paths between the transmitters and the receiver locations within the 3D model of the environment. For NLOS paths, the propagation model implements the shooting and bouncing rays method, where a number of uniformly spaced rays are launched from the transmitter and interact with surrounding objects. If a ray strikes a flat surface, it reflects and if it strikes an edge, it generates several diffracted rays in accordance with the law of diffraction. For each ray that reaches a receiver, the phase shift and the total path loss are calculated, which includes the free-space, reflection and diffraction losses. Reflection and diffraction losses are calculated using the Fresnel equations and the uniform theory of diffraction, respectively, and depend on the geometry and the electromagnetic properties of the materials defined in the environment's 3D model [Iskandar, Bertoni]. The total electric field (RMS) caused by a transmitter k is obtained as:

$$E_{k,\mathrm{tot}} = \sqrt{\frac{4\pi Z_0 P_{k,tx}}{\lambda^2}} \cdot \left| \sum_{ray=1}^{N_{ray}} \frac{10^{G_{k,tx,ray}/20}}{10^{L_{tot,ray}/20}} \cdot e^{i\theta_{ray}} \right| \tag{1}$$

where λ is the wavelength (in meters), Z_0 is the impedance of free space (in ohms), $P_{k,tx}$ is transmitter k's root-mean-square output power (in Watts), $G_{k,tx,ray}$ is transmitter k's gain (dBi) in the direction of the *ray*, $L_{\mathrm{tot},ray}$ is the *ray's* total path loss (in dB), θ_{ray} ($=2\pi d/\lambda$, d being the ray's propagation distance) is the *ray's* phase shift, i is the imaginary unit, and N_{ray} is the number of rays considered. In case more than one transmitter or base station is used, the total electric field is calculated as

$$E_{tot} = \sqrt{\sum_{k=1}^{K} E_{k,tot}^2} \tag{2}$$

where $E_{k,\mathrm{tot}}$ is the electric field caused by transmitter k, and K is the total number of transmitters or base stations. In all simulations, we assume an evaluation height of 1.5 m above ground level. For simplicity, we always assumed a worst-case duty cycle of 100%, meaning that we do not account for duty cycle/traffic load/data usage as model input.

2.2.3. Simulation scenarios

To build and evaluate the generic model, many different $[\boldsymbol{X}\,\boldsymbol{Y}]$ configurations were required in which the 11 input parameters $\boldsymbol{X}$ were varied. This section describes the total number of simulation scenarios and which input parameters were varied.

In a first phase, 12 800 different configurations were evaluated. 200 random factory layouts were created with a random length and width ($X1$) as described in table 1, with a ceiling height of 5 m ($X11$) and with a random percentage of clutter ($X8$). These layouts were divided into 4 groups of 50, each assigned a unique combination of clutter ($X9$) and wall ($X10$) material. Within each group, we ensured

Table 1. Input parameter ranges for the two sets of simulations.

Inputs	Explored parameter range	
	1st set of simulations	2nd set of simulations
Area ($X1$)	440–2025 m^2 Random Length × Width (Length: 22–45 m Width: 20—Length)	462–2970 m^2 Random Length × Width (Length: 22–55 m Width: 20—Length)
No_Tx ($X2$)	1, 2, 3, 4 1 and 2 (for 440–750 m^2 areas) 2 and 3 (for 750–1250 m^2 areas) 3 and 4 (for areas > 1250 m^2)	1, 2, 3, 4 1 and 2 (for 440–750 m^2 areas) 2 and 3 (for 750–1250 m^2 areas) 3 and 4 (for areas > 1250 m^2)
Den_Tx ($X3$)	220–750 m^2/BS derived from $X1$ and $X2$	231–990 m^2/BS derived from $X1$ and $X2$
Freq ($X4$)	3.5, 26 GHz	3.5 GHz
H_Tx ($X5$)	3, 4, 5 m (varied only for wall antennas; for ceiling antennas, H_Tx is equal to the ceiling height)	3, 4, 5, 6 m (varied only for wall antennas; for ceiling antennas, H_Tx is equal to the ceiling height)
P_Tx ($X6$)	−20, −10, 0, 10 dBW (i.e. 0.01, 0.1, 1, 10 W)	−7, −3, 0, 3 dBW (i.e. 0.2, 0.5, 1, 2 W)
Type_Ant ($X7$)	2.15, 4.53 dBi	2.15, 4.53 dBi
Den_Clutter ($X8$)	0%–25%	2.2–62.5%
Mat_Clutter ($X9$)	0, 0.146, 1 (reflectivity values of 0 for no-clutter, 0.146 for wooden clutter, 1 for metal clutter)	1 (only metal clutter)
Mat_Walls ($X10$)	0.046, 0.348 (reflectivity values of 0.046 for glass, 0.348 for concrete)	0.348 (only concrete walls)
H_Ceiling ($X11$)	5	5, 6

that there were 10 configurations with 0%–5% clutter, 10 with 5%–10% clutter, … and 10 with 20%–25% clutter. These 200 unique models were then imported into MATLAB, where the BS were added, with their number ($X2$) varying among four values, depending on the area of the environment. $X3$ was automatically calculated from $X1$ and $X2$. The BS frequency ($X4$) was set to either 3.5 or 26 GHz. The BS height ($X5$) varied across three values up to the factory ceiling height of 5 m, depending on the antenna type. The Base Station power ($X6$) was varied among four values. The BS antenna type ($X7$) was selected from two options (wall or ceiling), which determined the corresponding radiation pattern (half-omnidirectional with a gain of 4.53 dBi and omnidirectional with a gain of 2.15 dBi respectively).

In realistic deployments, the wireless installation will be configured to provide a realistic coverage percentage throughout the area. Therefore, if more than 10% of the values received a signal level below −81 dBm (3.5 GHz) [Castellanos] (bandwidth of 120 MHz, with noise floor of −93 dBm, noise figure of 7 dB and SNR of 5 dB for MCS 7) or below −74.9 dBm (26 GHz) [Santana], the simulation was not withheld for further processing. In the first phase, this led to rejection of 2604 simulations or 20% of the total.

Based on the obtained results, we performed a second round of simulations, considering larger areas ($X1$) also including more clutter ($X8$), refining the BS power ($X6$) input parameter evaluations. 240 new environments are generated ($X1$,$X8$,$X9$,$X10$), with 2 numbers of base stations ($X2$) and 4 BS powers ($X6$). These were further varied with different BS heights ($X5$) up to the ceiling height ($X11$, 5 and 6 m) for wall and ceiling base stations ($X7$). This resulted in a set of 8640 simulations. In this second phase, 929 simulations or 11% were rejected for inclusion in the model.

Table 2. Hyperparameters for generic neural network model.

Hyper-parameter	Value
Number of neurons in each layer	20, 20, 20, 3
Activation function	Relu
Optimiser	Adam
Initial learning rate	0.005
Loss function	mean_squared_error
Train and test ratio	8:2

2.2.4. Optimal simulation settings

Since training a reliable machine learning model requires a large amount of data and, consequently, many simulations, an appropriate balance between computation time per simulation and simulation accuracy must be found. As more **reflections** and/or **diffractions** are allowed in the simulation, accuracy is expected to increase, but calculation time as well. Therefore, for a set of four largely different simulation scenarios, the output parameters E_{50} and E_{95} were determined, as well as the calculation time. This allowed us to determine simulation settings that strike a good balance between accuracy and calculation time.

Similarly to the number and type of interactions taken into account, the number of receiver points in the area at which the E value is calculated has an impact on calculation time. Therefore, we investigated the difference in $E_{50/95}$ values for a **receiver grid** with spatial **resolutions** of 10, 20, 50, 100, and 200 cm. Although a sparse grid reduces calculation time, it might not be able to catch all local differences in E values, having an impact on the representativeness of the E_{50} and (particularly) E_{95} outputs. Therefore, we determined the $E_{50/95}$ values as a function of the grid size resolution for the four scenarios. For a 'worst-case' environment (lot of clutter), we calculated the output (median E), and determined to what extent coarser grids still gave the same output.

2.3. Machine learning approach

As a generic model, we used a fully connected neural network, where each neuron in one layer is connected to every neuron in the next layer. These networks consist of an input layer, one or more hidden layers, and an output layer. Each neuron applies a weighted sum followed by a non-linear activation function to capture complex patterns in the data. We opted for a neural network because of the high dimensionality of the input vectors and the non-linear relationships between the inputs and output. The hyper-parameters used are shown in table 2 below.

The model was developed in Python using standard libraries. The relevant libraries used include TensorFlow/Keras (model definition and training), scikit-learn (data splitting, normalisation, and performance metrics), NumPy (numerical operations), and Pandas (data handling).

3. Results

3.1. Optimal simulation settings

Figure 4 shows E_{50}, E_{95}, and the calculation time as a function of the number of interactions (reflections/diffractions) allowed for a typical wireless configuration with a grid resolution of 1 m. As the number of interactions increased, calculation time increases exponentially. The obtained E values remained relatively stable though. Therefore, as a suitable tradeoff, we allowed 2 reflections for the remainder of the analysis.

Figure 5 shows E_{50} and E_{95} as a function of the **grid size** resolution (in m) for four simulation scenarios. For the worst-case scenario (scenario 1), the differences remained below roughly 4% when choosing a grid size of 1 m, compared to smaller grid sizes. Therefore, we opted for a 1 m grid size as a good tradeoff between accuracy and calculation time.

3.2. Correlation analysis

To increase the model's genericity and usability, we aim to reduce the model complexity and retain a limited set of physically meaningful and easily obtainable input parameters. Thereto, we investigate the predictive power of the various inputs, via calculation of the Pearson correlation coefficient between all inputs and outputs. It measures the strength and direction of the linear relationship between two variables. The correlation analysis between all input variables and all potential output variables are shown in table 3. Input-output correlations with an absolute value larger than 0.2 are underlined. We observe

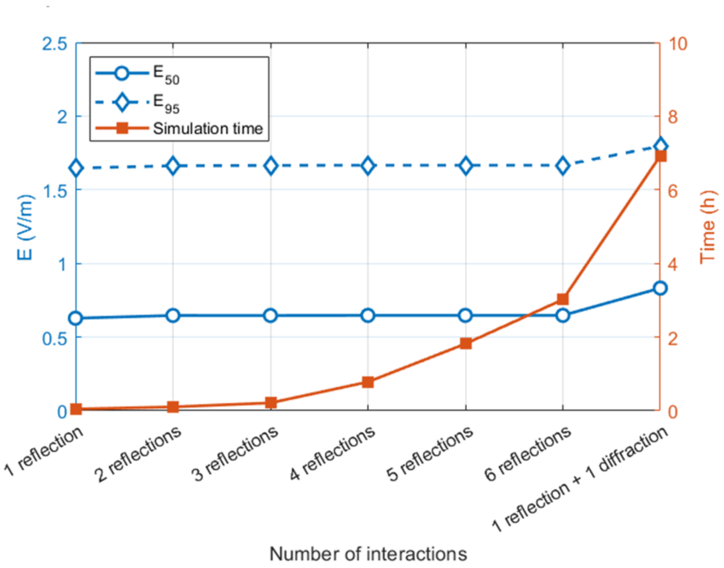


Figure 4. Comparison of E_{50}, E_{95}, and calculation time for different numbers of interactions allowed.

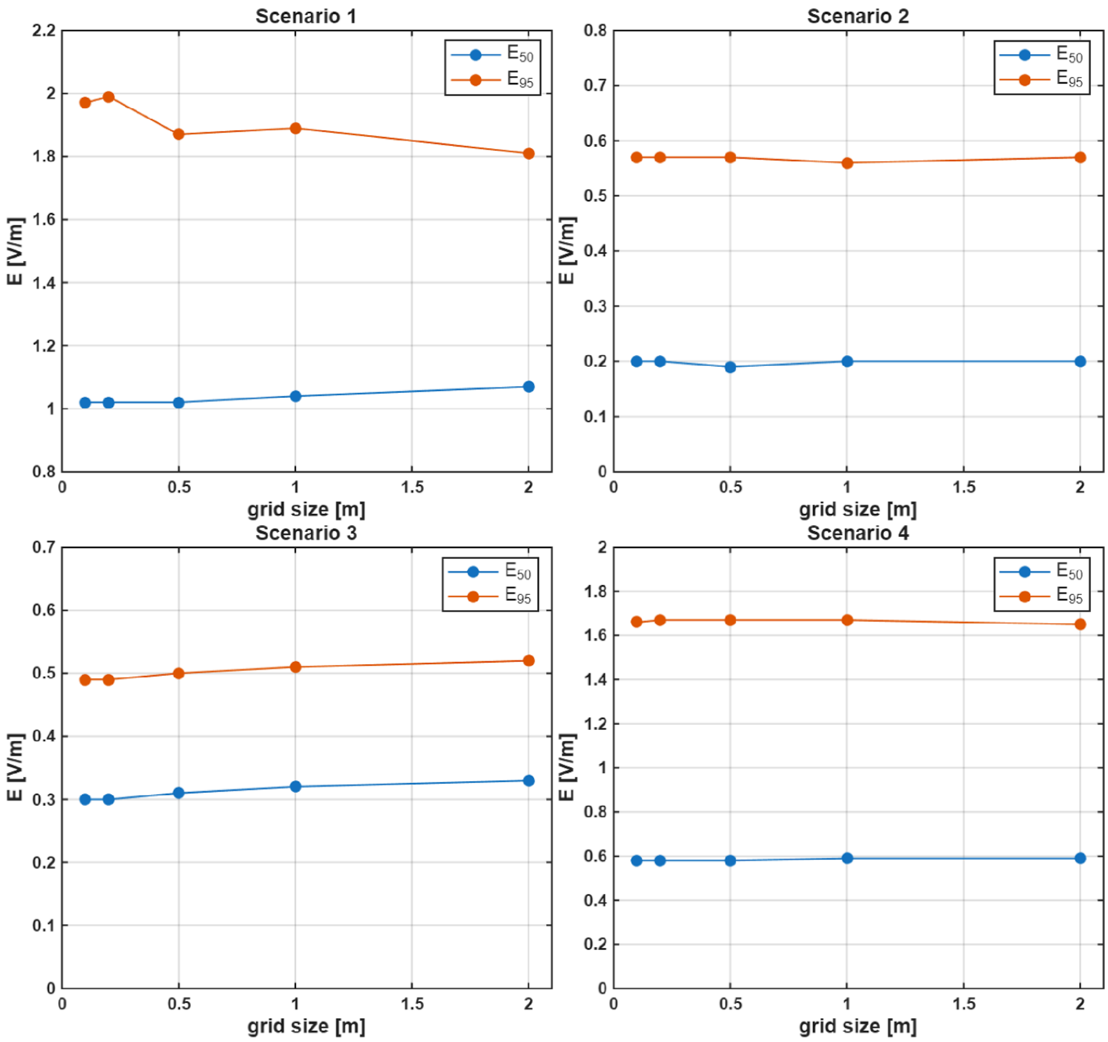


Figure 5. Comparison of E_{50} and E_{95} for different receiver grid sizes.

Table 3. Correlation table between all inputs and outputs.

	Area	No_Tx	Den_Tx	Freq	H_Tx	P_Tx	Type_Ant	Den_Clutter	Mat_Clutter	Mat_Wall	H_Ceiling	E_{50}	E_{95}
Area	1.000	0.709	0.596	−0.228	0.065	0.074	0.029	0.344	0.250	0.236	0.296	−0.086	−0.080
No_Tx	0.709	1.000	−0.104	−0.134	0.039	0.044	0.017	0.247	0.153	0.127	0.173	0.084	0.001
Den_Tx	0.596	−0.104	1.000	−0.152	0.043	0.050	0.019	0.181	0.174	0.175	0.199	−0.224	−0.116
Freq	−0.228	−0.134	−0.152	1.000	−0.082	−0.089	−0.026	−0.226	−0.281	−0.273	−0.357	0.068	0.065
H_Tx	0.065	0.039	0.043	−0.082	1.000	0.027	−0.443	0.060	0.084	0.082	0.242	−0.021	−0.212
P_Tx	0.074	0.044	0.050	−0.089	0.027	1.000	0.009	0.074	0.092	0.089	0.116	0.721	0.827
Type_Ant	0.029	0.017	0.019	−0.026	−0.443	0.009	1.000	0.048	0.027	0.026	0.055	−0.023	0.091
Den_Clutter	0.344	0.247	0.181	−0.226	0.060	0.074	0.048	1.000	0.270	0.227	0.346	−0.248	−0.054
Mat_Clutter	0.250	0.153	0.174	−0.281	0.084	0.092	0.027	0.270	1.000	0.280	0.367	−0.099	−0.074
Mat_Wall	0.236	0.127	0.175	−0.273	0.082	0.089	0.026	0.227	0.280	1.000	0.357	−0.073	−0.073
H_Ceiling	0.296	0.173	0.199	−0.357	0.242	0.116	0.055	0.346	0.367	0.357	1.000	−0.088	−0.105
E_{50}	−0.086	0.084	−0.224	0.068	−0.021	0.721	−0.023	−0.248	−0.099	−0.073	−0.088	1.000	0.835
E_{95}	−0.080	0.001	−0.116	0.065	−0.212	0.827	0.091	−0.054	−0.074	−0.073	−0.105	0.835	1.000

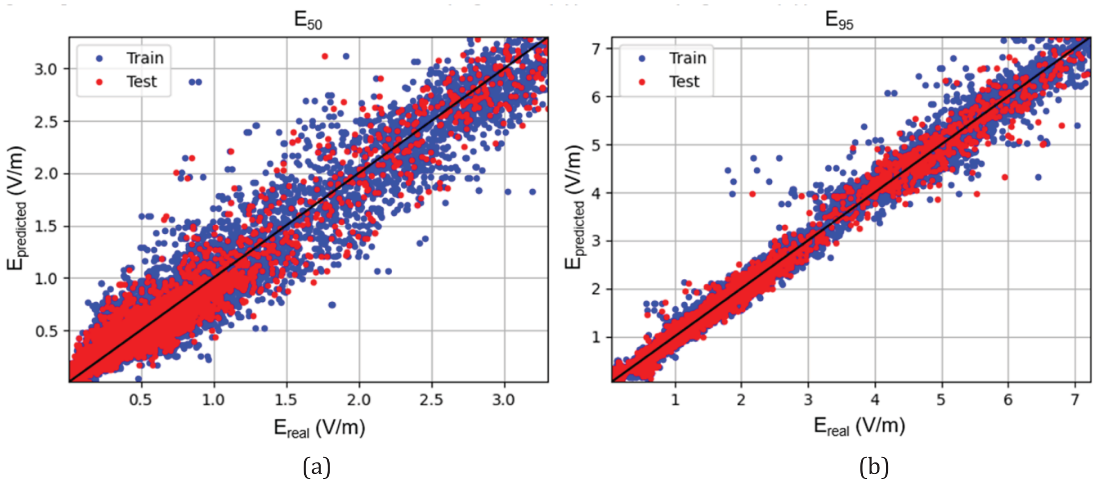


Figure 6. Scatter plots of predicted values vs ground truth values for the three output variables.

Table 4. RMSE and R^2 for training and testing datasets.

	E_{50}	E_{95}
RMSE-Training	0.164 V m^{-1}	0.173 V m^{-1}
RMSE-Testing	0.161 V m^{-1}	0.164 V m^{-1}
R^2-Training	0.950	0.990
R^2-Testing	0.949	0.990

strong correlations between transmit power ($X6$) and E_{50} and E_{95} in the outputs. Some significant correlations were also observed from density of transmitters ($X3$), transmitter height ($X5$), and density of clutter ($X8$). The height of the transmitter mainly impacts the maximal exposure values with an inverse relationship, as can be expected. Clutter density has a larger impact on the median exposure values, rather than on the maximal exposure values (which are found at unobstructed locations near the transmitters). As could be expected, the correlation analysis confirmed that $X3$ (density of transmitters) provides the same information as the evaluation area ($X1$) and the number of transmitters ($X2$). It was observed that the antenna type ($X7$) impacts the E-field distribution shape (kurtosis), and inherently has a link with $X5$ in the model. The correlation between inputs and outputs revealed the essential parameters that could be used in machine learning model building. As a conclusion to this analysis, we retained the parameters in grey as the essential input features for building the model: $X3$, $X5$, $X6$, $X7$, $X8$.

3.3. Model performance

Based on the correlation analysis and the approach described in section 2.3, a generic ML model was built, with input parameters $X3$, $X5$, $X6$, $X7$, $X8$ and output parameters $Y1$, $Y2$. For the fully connected neural network model, the training and testing performance are shown in the scatter plots shown in figure 6. A good prediction can be observed from E_{50} (figure 6(a)) and E_{95} (figure 6(b)).

In terms of training and testing performance metrics, table 4 shows the root mean squared error (RMSE) and R^2 for training and testing. RMSE measures the root of the average squared difference between predicted values and actual target values. R^2 measures how well the model explains the variance in the target variable. It ranges from 0 to 1, where 1 means the model perfectly explains the variance, 0 means it performs no better than predicting the mean, and negative values indicate worse performance than the mean predictor.

The model demonstrates strong predictive performance for both E_{50} and E_{95}, with low RMSE and high R^2. For E_{50}, the R^2 is around 0.95 for both training and testing, indicating the model explains 95% of the variance with consistent generalisation. These results confirm that the model captures the underlying trends very accurately.

The loss graph of figure 7 shows the training and validation loss across 100 epochs. This indicates that the model is learning effectively, with no strong signs of overfitting since the training and validation losses remain close throughout training.

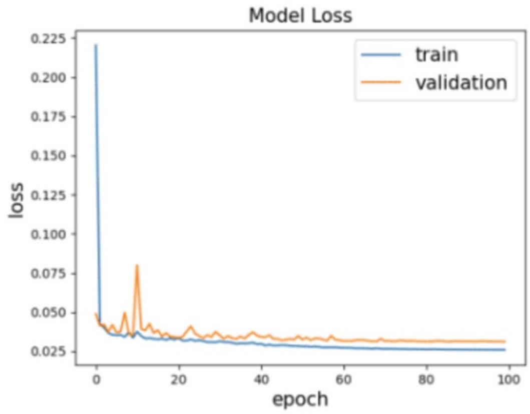


Figure 7. Training and validation loss as a function of number of epochs.

Table 5. Descriptive parameters of the three experimental datasets. Parameters in italic & bold ($X3$, $X5$, $X6$, $X7$, $X8$) are retained as inputs to the generic model (*2 antennas were mounted at a height of 3.1 m).

Input parameter	Value		
	Warehouse	Production hall	Wireless networks test facility
$X1$—environment size (m^2)	3398.6	2116.8	399
$X2$—number of transmitters (-)	4	6	1
$X3$—transmitter density (/100 m^2)	0.118	0.283	0.251
$X4$—frequency (GHz)	3.75	3.75	3.47
$X5$—antenna height (m)	5	5.9*	3
$X6$—transmit power (dBW)	0	0	13
$X7$—antenna type/location	Ceiling (omni)	Ceiling (omni)	Wall (half-omni)
$X8$—clutter area percentage (%)	40	20	0
$X9$—dominant clutter material	Metal	Metal	—
$X10$—outer wall material	Concrete	Concrete	Concrete
$X11$—ceiling height	5.5	6.9	5.6

3.4. Experimental validation

The obtained model is then tested on experimental data collected in three real-world environments, where 5 G private networks are deployed, hereafter referred to as 'Warehouse', 'Production hall', and 'Wireless networks test facility'. The environments of the first two are described in detail in [Gajšek]. The 'Wireless Networks Test Facility' covered a space of approximately 400 m^2, primarily constructed of concrete, with a metal roof at a height of approximately 6 m. One side featured metal doors approximately 6 m wide, extending the full height of the wall, while the opposite side bordered offices, with metallic wall panels and several small windows. The facility did not contain any clutter. The private SA 5G network operated in the 3420–3520 MHz frequency range. A single Ericsson 6524 antenna, with an output power of 20 W and a maximum gain of 12 dBi, was mounted on the wall at a height of 3 m. Since the outputs of the generic model are the E_{50} and E_{95} values representing the entire occupational environment, raster measurements were used for the validation to ensure uniform coverage of the area of interest. However, because measurements cannot provide detailed coverage over large areas such as those in these environments, the generic model results were also compared with corresponding values obtained from calculations using the commercial software Narda EFC 400 [Narda] over a denser set of points. The Narda tool serves as an additional independent validation of the generic model since its simulations were not used to train the generic model.

Table 5 lists the parameters for characterising these three environments, that were initially considered as inputs to the generic model. As described in the correlation analysis, only the parameters $X3$, $X5$, $X6$, $X7$, and $X8$ were eventually used as inputs to the generic model. In the 'Production hall' environment, base station antennas were installed at two heights, but since the current generic model requires a single input value, a height of 5.9 m was used as 4 out of 6 antennas are mounted at that height.

Table 6 presents the outputs of the generic model for each environment in which it was tested, along with the simulated output parameters using the Narda EFC 400 software, and the experimental output parameters based on the measurement campaigns.

Table 6. Comparative output of simulation and measurements with generic model estimations.

Output parameter	Method	Warehouse	Production hall	Wireless networks test facility
E_{50} (V m^{-1})	Simulations NARDA	0.34	0.74	4.01
	Measurements	0.39	0.67	5.05
	Generic model	0.24	0.75	4.10
E_{95} (V m^{-1})	Simulations NARDA	1.06	1.14	11.31
	Measurements	1.01	1.22	10.74
	Generic model	1.21	1.34	8.23

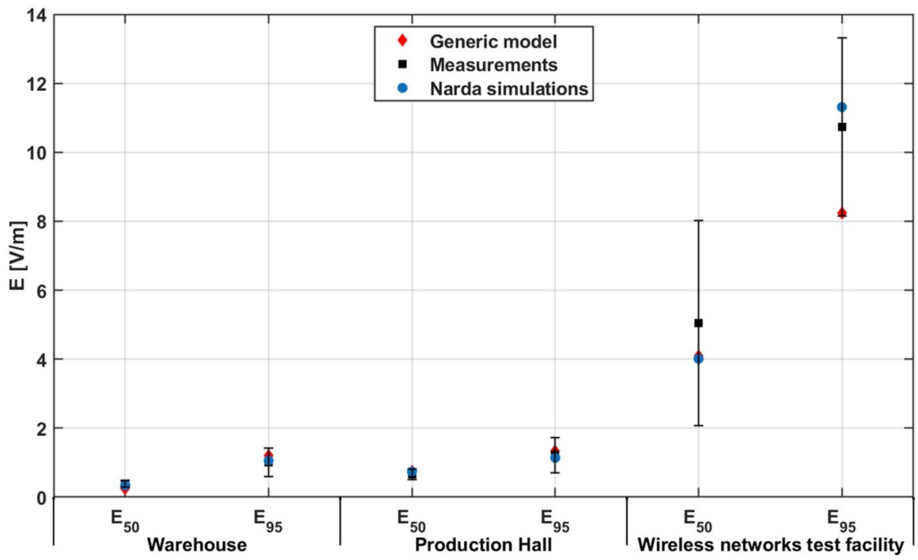


Figure 8. Comparison of results obtained by generic model (red diamonds), Narda simulations (blue circles) and measurements (black square).

In figure 8, the E_{50} and E_{95} values obtained from the generic model, the measurements and Narda simulations are plotted. The error bars represent the expanded uncertainty (95% confidence interval) of the E_{50} and E_{95} statistics obtained from the measurement dataset using the bootstrap method [Efron]. Its detailed calculation was performed in accordance with the guidelines in [NIST].

Overall, table 6 and figure 8 show a good agreement between the generic model results and the measurements. Only one value falls outside the uncertainty range: E_{50} in the 'Warehouse' environment, where the generic model's output was 0.24 V m^{-1}, which is lower than the uncertainty range's lower limit of 0.29 V m^{-1} (the actual measured value was 0.39 V m^{-1}). For E_{95} in the 'Wireless networks test facility' environment, the model's value of 8.23 V m^{-1} was just within the bounds of the uncertainty range of the measured value of 10.74 V m^{-1}. For all other cases, the generic model agreed very well with the measurements, with relative differences not exceeding 20%. e.g. for the 'Production hall' environment, both median values (measured value of 0.67 V m^{-1} vs generic model prediction of 0.75 V m^{-1}) and 95th percentile values (measured value of 1.22 V m^{-1} vs generic model prediction of 1.34 V m^{-1}) correspond very well.

A similar level of agreement is observed when comparing the generic model results with the simulations. For the median value of the 'Warehouse' case, the relative difference between the two models was 29% (0.24 and 0.34 V m^{-1} respectively). Although this percentage appears relatively high, it is mainly due to the low absolute field strengths. The absolute difference of 0.1 V m^{-1} is generally not considered significant in the context of RF human exposure assessment. For the 95th percentile value of the 'Wireless Networks Test Facility' case the relative difference was similar, at 27% (8.23 and 11.31 V m^{-1}). This relatively high difference is attributed primarily to the high antenna directivity (12 dBi) as well as its output power (13 dBW) which were outside the range of the data used for the training of the generic model. It is to be noted that neural networks mainly perform well for input data within their training range.

4. Conclusion

In this paper, a neural network model was created and validated to enable generic estimation of exposure levels in industrial environments, without requiring technically complex inputs or technical expertise. First, a set of eleven parameters was defined, relating to the physical environment and to the wireless deployment of the network. These parameters were selected to be accessible to non-experts and expected to correlate with exposure levels. A Blender-MATLAB pipeline was created for simulating an extensive set of different wireless configurations. From over 12 000 simulations, five main input parameters were retained that correlate well with the exposure level distribution in an area: the height, density, location and transmit power of the transmitter, and the density of clutter in the area. Based on these input parameters, a neural network was trained on over 16 000 simulations. Subsequently it was successfully tested on more than 4 000 simulated new wireless deployments, achieving R^2 values of over 95% for the median and maximal exposure levels in the area. Three experimental datasets from industrial warehousing and production environments were used as real-life test cases for the generic exposure estimation model. The results show that the model provides good approximations of electric-field values measured in real environments, with most predictions falling within the measurements' uncertainty bounds. The model's accuracy decreases when input parameters lie outside the range represented in the training data. Although the findings indicate that machine-learning models can approximate exposure levels without prior electromagnetic expertise or dedicated simulation tools, they also suggest that performance can be improved—and applicability broadened—by incorporating additional and more complex training environments. All simulated and experimental datasets are made publicly available, enabling researchers to add new experimental data, refine the existing model, or explore alternative modelling approaches.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://doi.org/10.5281/zenodo.18598701 [Apostolidis].

Funding

This research was funded by the European Union's Horizon Europe Framework Programme under Grant Agreement number 101057622 (SEAWave Project). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the Health and Digital Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

ORCID iDs

David Plets 0000-0002-8879-5076
Christos Apostolidis 0000-0001-5799-6919
Luc Martens 0000-0001-9948-9157
Theodoros Samaras 0000-0003-0170-4520

References

[Artetxe] Artetxe E, Barambones O, Calvo I, Fernández-Bustamante P, Martin I and Uralde J 2023 Wireless technologies for industry 4.0 applications *Energies* **16** 1349

[Ashraf] Ashraf M A and Celik T 2024 Evaluating radiofrequency electromagnetic field exposure in confined spaces: a systematic review of recent studies and future directions *Radiat. Prot. Dosim.* **200** 598–616

[Bertoni] Bertoni H L 2000 Radio Propagation for Modern Wireless Systems (Prentice Hall PTR)

[Blender] Blender Online Community 2018 Blender—a 3D modelling and rendering package (available at: www.blender.org)

[Castellanos] Castellanos G, De Gheselle S, Martens L, Kuster N, Joseph W, Deruyck M and Kuehn S 2022 Multi-objective optimisation of human exposure for various 5G network topologies in Switzerland *Comput. Netw.* **216** 109255

[Efron] Efron B and Gong G 1983 A leisurely look at the bootstrap, the jackknife, and cross-validation *Am. Stat.* **37** 36–48

[El Habachi] El Habachi A, Conil E, Hadjem A, Vazquez E, Wong M F, Gati A, Fleury G and Wiart J 2010 Statistical analysis of whole-body absorption depending on anatomical human characteristics at a frequency of 2.1 GHz *Phys. Med. Biol.* **55** 1875–87

[Islam] Islam M T, Sepanloo K, Woo S, Woo S H and Son Y-J 2025 A review of the industry 4.0–5.0 transition: exploring the intersection, challenges, and opportunities of technology and human–machine collaboration *Machines* **13** 267

[Jabeur] Jabeur R and Alaerjan A 2024 Improving monitoring of indoor RF-EMF exposure using IoT-embedded sensors and kriging techniques *Sensors* **24** 7849

[Jiang] Jiang Y, Wang H, Sun X, Li C and Wu T 2023 Evaluation of Chinese populational exposure to environmental electromagnetic field based on stochastic dosimetry and parametric human modelling *Environ. Sci. Pollut. Res.* **30** 40445–60

[Joseph] Joseph W, Verloock L, Goeminne F, Vermeeren G and Martens L 2012 Assessment of RF exposures from emerging wireless communication technologies in different environments *Health Phys.* **102** 161–72

[Gajšek] Gajšek P, Apostolidis C, Plets D, Samaras T and Valič B 2025 EMF exposure of workers due to 5G private networks in smart industries *Electronics* **14** 2662

[Hirata] Hirata A *et al* 2021 Assessment of human exposure to electromagnetic fields: review and future directions *IEEE Trans. Electromagn. Compat.* **63** 1619–30

[Iskandar] Iskandar N and Shimamoto S Prediction of propagation path loss for stratospheric platforms mobile communications in urban site LOS/NLOS environment *2006 IEEE Int. Conf. on Communications* pp 5643–8

[Kiouvrekis] Kiouvrekis Y, Psomadakis I, Vavouranakis K, Zikas S, Katis I, Tsilikas I, Panagiotakopoulos T and Filippopoulos I 2025 Explainable machine learning-based electric field strength mapping for urban environmental monitoring: a case study in paris integrating geographical features and explainable AI *Electronics* **14** 254

[Majid] Majid M, Habib S, Javed A R, Rizwan M, Srivastava G, Gadekallu T R and Lin J C 2022 Applications of wireless sensor networks and internet of things frameworks in the industry revolution 4.0: a systematic literature review *Sensors* **22** 2087

[Mallik] Mallik M, Clavier L and Gaillot D P 2025 Realistic EMF exposure estimation from low density sensor network by finite & infinite neural networks *2025 19th European Conf. on Antennas and Propagation (Eucap)* (*Stockholm, Sweden*) pp 1–5

[Narda] Narda n.d. EFC-400 Software (available at: www.narda-sts.com/en/products/emf-broadband-measuring-devices/efc-400-software/)

[NIST] NIST/SEMATECH 2012 *Bootstrap Plot. e-Handbook of Statistical Methods* (https://doi.org/10.18434/M32189)

[Santana] Santana Y H, Martinez Alonso R, Guillen Nieto G, Martens L, Joseph W and Plets D 2024 5g mmwave network planning using machine learning for path loss estimation *IEEE Open J. Commun. Soc.* **5** 3451–67

[Wang] Wang S and Wiart J 2020 Sensor-aided EMF exposure assessments in an urban environment using artificial neural networks *Int. J. Environ. Res. Public Health* **17** 3052

[Apostolidis] Apostolidis C, Plets D, Wang S, Valič B, Gajšek P, Samaras T and Martens L 2026 RF Exposure generic model dataset 2026 (https://doi.org/10.5281/zenodo.18598701)

[Zhang] Zhang Y, Wang S and Wiart J 2025 ExposNet: A Deep Learning Framework for EMF Exposure Prediction in Complex Urban Environments (https://doi.org/10.48550/arXiv.2503.02966)