IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 20 May 2024, accepted 3 June 2024, date of publication 5 June 2024, date of current version 12 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3410090

RESEARCH ARTICLE

Statistical Analysis of RF-EMF Exposure Induced by Cellular Wireless Networks in Public Transportation Facilities of the Paris Region

YARUI ZHANG[1], SHANSHAN WANG[2], (Member, IEEE), WASSIM BEN CHIKHA[1], (Senior Member, IEEE), JIANG LIU[1], CE ZHENG[1], THEODOROS SAMARAS[3], (Member, IEEE), AND JOE WIART[1], (Senior Member, IEEE)

[1]Chaire C2M, LTCI, Télécom Paris, Institut Polytechnique de Paris, 91120 Palaiseau, France
[2]ETIS, ENSEA, CNRS, UMR 8051, CY Cergy Paris Université, 95014 Cergy, France
[3]Department of Applied and Environmental Physics, School of Physics, Faculty of Sciences, Aristotle University of Thessaloniki, 541 24 Thessaloniki, Greece

Corresponding author: Yarui Zhang (yarzhang@telecom-paris.fr)

This work was supported in part by Europeen (EU) project SEAWAVE through the Horizon Europe Research and Innovation Programme under Grant 101057622, and in part by the project Beyond 5G (B5G) granted by bpiFrance [Banque publique d'investissement (BPIF)] and Ministère de l'économie, des finances et de l'industrie (MINEFI).

ABSTRACT Wireless communications are increasingly used today. Despite such use, there is a significant perception of risk which makes exposure monitoring a significant concern today. The work described in this article was carried out within the framework of the European SEAWave project and the French Beyond5G project. The exposure assessment was evaluated using a personal exposimeter (MVG EMF Spy) whose compactness and ease of use make it more suitable and portable than a system combining measuring probes and spectrum analyzers. Measurements were carried out on the cellular frequency bands used by 2G, 3G, 4G, and 5G, as well as that of Wi-Fi, in different modes of public transportation (RER, metro, tramway, bus, and train) circulating in the Paris region. The measurements have been analyzed by frequency band, type of public transportation, and type of environment encountered. For each set of measurements (e.g., metro lines, tramways), the mean, standard deviation, skewness, and kurtosis were evaluated and analyzed. For all exposure measurements taken in the 700, 800, 900, 1800, 2100, 2600, and 3500 MHz frequency bands, the overall average values are 0.39, 0.43, 0.30, 0.21, 0.18, 0.24 and 0.18 V m^{-1}, respectively. These measurements have, in all cases, a significant dispersion as shown by the ratios of standard deviations to mean values. The well-known K-means clustering technique was applied to these four parameters for different subsets of data. The number of clusters k = 3 has been chosen based on the analysis of the optimal value of k for the current dataset. Our analysis indicates that the first group's members display the highest mean values with moderate variance and the lowest values for the third and fourth moments. The second cluster is distinguished by points with large mean and variance, accompanied by moderate skewness and kurtosis. Conversely, the third group comprises points with the smallest mean and variance values, yet the largest measurements for the third and fourth moments.

INDEX TERMS Electromagnetic fields (EMF) exposure assessment, K-means, personal exposimeter (PEM), public transportation.

The associate editor coordinating the review of this manuscript and approving it for publication was Xiaodong Liang.

I. INTRODUCTION

Today, wireless communication utilizing radiofrequency electromagnetic fields (RF-EMF) has become increasingly prevalent. According to [1], there were in 2023 over 11.856 billion mobile connections worldwide, which were 3.799 billion more mobile connections than people worldwide. There is a widespread use of cellular wireless communication. For instance, approximately 87% of the French population owned smartphones in 2022, including over 95% of individuals within the age group of $12 - 59$ years [2]). Protection guidelines exist, such as those by the International Commission on Non-Ionizing Radiation Protection [3] and standardized methods have been designed and adopted for testing the compliance with these guidelines [4], [5], [6]. Despite this, there remains a significant perception of risk associated with exposure to RF-EMF [7] and the recent deployment of the 5G wireless communications network has further increased this perception of risk [8] and is causing wide debates. In response to growing concerns about 5G deployment, recent research has intensified efforts to analyze and understand the specific characteristics and potential impacts of 5G networks. Studies such as those documented in [9] and [10] have explored the propagation, signal behavior, and frequency-specific attenuation of 5G technologies, as well as the coverage and performance of 5G deployments across various environments, further enriching our understanding of how these new technologies perform in diverse settings.

The last decades have seen large efforts dedicated to EMF exposure assessment, standardization, and EMF monitoring [11]. EMF probes have been installed in many countries [12], [13], [14]. These devices provide the temporal variation of electric field (E-field) strength [15]. However, their measurements are constrained to the specific locations where the sensors are positioned, limiting the comprehensiveness of data collection to those fixed points. EMF spot measurements have also been carried out in response to the demands of the public and local authorities [13], [16], [17], [18]. While these measurements offer valuable insights, the majority have been conducted outdoors, despite the fact that people spend the vast majority of their time (approximately 90%) indoors, as indicated by [19]. This is particularly relevant in large urban areas like the Paris region, where time spent in public transportation is significant (e.g., over 100 minutes in 2022), according to [20]. Additionally, much of the existing literature focuses on 2G, 3G, or 4G technologies. Only a few of them address 5G, but in those cases, most focus on compliance, and few address exposure during commuting [21], [22], [23], [24], [25], [26], [27], [28]. This context emphasizes the importance of RF-EMF exposure monitoring that has been assigned to the objectives of the EU call "HORIZON-HLTH-2021-ENVHLTH-02-01" [29]. Following this EU call, several projects have been funded (SEAWave, GOLIAT, ETAIN, and nextGEMS). The SEAWave project, as cited in [30] and under which the research for this paper was conducted, has been selected in response to this call and aims to address EMF monitoring and related concerns. In this paper, we detail measurements conducted to evaluate RF-EMF exposure within Parisian public transportation systems, alongside an analysis and discussion of the findings. In the first section, we introduce the measurement equipment and protocol employed. Following this, we apply unsupervised clustering through the K-means method, as documented in [31] and [32]. The conclusion section reflects on the results and offers a comprehensive summary and conclusion of the study.

II. MEASUREMENT CAMPAIGN

A. MEASUREMENT EQUIPMENT: PERSONAL EXPOSIMETER (PEM)

The objective is to evaluate exposure levels in public transportation, especially during peak hours when the environment tends to be crowded. Under such conditions, using a complex system involving an isotropic probe, a tripod, and a spectrum analyzer was deemed impractical. Therefore, for this study, we opted for a much less bulky exposimeter, the EME Spy Evolution [33]. This system is a frequency-selective portable electromagnetic field meter equipped with dedicated hard filters. The measurement campaign was carried out with three EME Spy, each of them equipped with a tri-axial electric field probe, allowing the measurement of the 3 spatial components of the electromagnetic field. The device has the capability to capture and analyze up to 20 user-defined frequency bands (out of a total of over 70), covering a frequency range from 80 MHz to 6 GHz. In our study, we specifically chose 13 cellular frequency bands and 2 Wi-Fi bands utilized in France. This allowed to create a measurement scenario tailored to the French context. Considering the environment of moving transportations, where access to public Wi-Fi hotspots is typically unavailable, our analysis primarily focuses on cellular frequency bands. For indoor public transportation settings, such as train stations, we will also provide a brief overview of the E-field levels observed in the two Wi-Fi frequency bands. The exposimeter's minimum recording period (or maximum sampling rate) is contingent on the number of selected frequency bands, varying from 2 seconds to 255 seconds. In our scenario, we chose a minimum recording period of 4 seconds when considering 7 cellular bands. In locations where the two Wi-Fi bands are present, the recording period is extended to 6 seconds.

Table 1 presents the list of the 7 frequency bands measured using the PEM. Most of these frequency bands are allocated for use in 4G or 5G networks. The 900 MHz band is utilized by 2G (GSM) and 3G (UMTS), the 1800 MHz band is employed by LTE and partially by 3G (UMTS), and the 2100 MHz band is utilized by 5G and partially by 3G. Note that the 5G data mentioned in this article is focused on the Sub-6 GHz spectrum and does not involve the millimeter wave bands. The sensitivity, indicating the lower detection limit of the PEM, varies across different frequency bands. The upper detection limit is set at $6\ \mathrm{V\,m^{-1}}$. Consequently, if the

TABLE 1. Frequency bands configuration for measurement campaigns.

Technology	Frequency (MHz)	Sensitivity (V m^{-1})
LTE 700 DL+ 5G NR	758-803	0.045
LTE 800 DL	791-821	0.03
GSM+UMTS 900 DL	925-960	0.02
GSM+LTE 1800 DL	1805-1880	0.02
UMTS 2100 DL 5G NR	2110-2170	0.02
LTE 2600DL	2620-2690	0.03
5G NR 3500	3300-3800	0.02

E-field in one or more bands exceeds this upper limit, the result will be truncated and a smaller E-field value will be recorded.

Exposure assessment using PEM is straightforward, yet, as highlighted in [34], several challenges arise, such as the body of the person holding the PEM potentially masking the incident field, and the coupling between the dosimeter's antenna and the experimenter's body influencing the assessment. These factors can affect the accuracy of the measurements.

To tackle this issue, we carried out a comparative analysis of measurements taken with a reference system equipped with an isotropic probe and a spectrum analyzer (e.g., Narda SRM and Tektronix RSA306B [35]) versus those obtained with the PEM close to the human body. This analysis aimed to account for the interactions between the experimenter's body and the PEM, adjusting the measurements to more accurately estimate the actual EMF exposure.

During the measurement process, the angles of arrival of waves from base stations, as well as the experimenter's relative positions to these antennas, remained unknown. Without specific angular information, correcting the measurement becomes a random variable, unsuitable for deterministic correction approaches. Moreover, being in motion, the correction factor for measurements evolves as a random variable, independent of the electromagnetic waves' directions of arrival. Therefore, along a given path, the actual exposure, being the product of two independent variables, averages to the product of their averages. This means that the average real exposure equates to the product of the average of the correction factor and the average of the E-field measured by the PEM.

To determine the average correction factor, we start with a reference measurement at a central location and compare it to measurements taken with the PEM carried in one position, as shown in Fig. 1. For each frequency band, we recorded measurements using both the PEM and the Tektronix system. Figure 2 depicts the Probability Density Function (PDF) of the correction factor for each measurement point (i.e., the ratio of the value recorded by the Tektronix system to that obtained from the PEM at the current point) at 1800 MHz, providing insight into the distribution of coefficients throughout the measurement process.

The average correction factor was then determined by calculating the ratio of the Root Mean Square (RMS) values from the Tektronix measurements during the whole

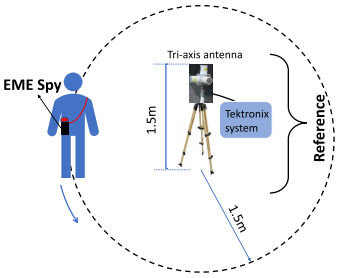


FIGURE 1. Measurement set-up to assess the impact of the body on PEM reading.

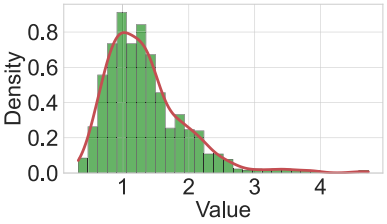


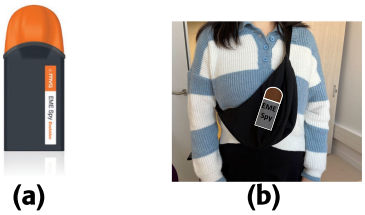
FIGURE 2. PDF of coefficients at each measurement point for 1800 MHz band.

measurements to those obtained with the PEM. For the frequencies 700 MHz, 800 MHz, 900 MHz, 1800 MHz, 2100 MHz, 2600 MHz, and 3500 MHz, the correction factors, are respectively 3.18, 3.71, 1.80, 1.17, 0.98, 1.32, and 3.54. It is important to note that while applying these correction coefficients adjusts the average values of the PEM measurements, individual values at specific local points may be amplified with some level of uncertainty.

B. SELECTED AREA OF INVESTIGATION

The Paris region, known as Île-de-France, was selected for its diverse landscape, ranging from semi-rural areas to densely populated urban zones. Home to over 12 million people, constituting around 18% of the French population, Île-de-France residents often face the challenge of commuting due to the spatial distribution of their residences and workplaces. Commuting is a routine aspect of daily life for many residents, and a significant portion relies heavily on public transportation for their daily journeys to and from work. While commuting, residents in Île-de-France engage in various wireless communication activities such as making calls, watching videos, and accessing social media. Monitoring the EMF exposure during these commuting periods is thus crucial.

To assess EMF exposure, we have selected various modes of public transportation. The Regional Express Network (a.k.a RER) and the Regional Express Train (a.k.a TER) have been specifically chosen due to their widespread use by citizens residing in both close and distant suburbs. The RER network, serving as the primary suburban transportation network, experiences substantial ridership. Comprising five RER lines labeled from A to E, the network caters to



(a) (b)

FIGURE 3. (a) EME spy evolution. (b) How to wear the PEM.

TABLE 2. Description of public transportation assessed in paris area.

Description	No. of scenarios
RER	5
Metro	14
Bus	7
Train	2
Tram	8
Train stations	9
Airport	2

a combined daily ridership exceeding 3 million users, as reported by [36].

As an illustration, the RER A, spanning 110 km (including 26 km underground), stands out as Europe's most crucial rail line, catering to over 1.2 million passengers daily. The extensive reach and daily ridership highlight its significance in the region's transportation network. Additionally, other widely used modes of transportation, including the metro, bus, and tram, were subject to exposure assessments. The interconnected metro lines alone witness the daily transit of more than 4 million passengers, emphasizing the need to comprehensively study and monitor RF-EMF exposure in these densely utilized public transportation systems.

Table 2 summarizes the various modes of public transportation, including RER, Metro, Bus, Train, Tram, etc., that have been selected for the assessment of RF-EMF exposure in the Paris region.

C. MEASUREMENT PROTOCOL

The experimenters, carrying a PEM (refer to Fig. 3) in a small bag worn close to their bodies, alternated between walking, sitting, or standing, depending on the circumstances, during their journeys on designated public transportation routes.

In moving transportations, the PEM was configured to conduct frequency-selective measurements every 4 seconds across all relevant cellular frequency bands (refer to Section II-B), spanning from 700 MHz to 3.5 GHz. For all selected environments, the minimum duration of exposure was set at 15 minutes, corresponding to 225 measurements with the chosen sampling rate of 15 measurements per minute. In indoor scenarios such as train stations and airports, the measurement interval is set to 6 seconds as we include both 2.4 GHz and 5 GHz Wi-Fi bands.

Using the measurement system and the protocol described previously, we conducted measurements across various modes of public transportation. In total, around 17000 frequency-selective measurements were carried out, offering a comprehensive overview of exposure levels in these environments.

III. RESULTS ANALYSIS

In this section, we will commence our analysis of the measured data by examining key statistical metrics such as mean value, standard deviation, and 95th percentile across various scenarios and frequencies. Illustrative figures will be given to depict the time-space variation and PDF of the data. Following this, an in-depth analysis will be conducted utilizing statistical parameters to characterize the distribution patterns of different datasets. The K-means technique will then be employed to cluster data exhibiting similar distributions, facilitating a more refined exploration based on categorization. It is important to highlight that measurements falling below the sensitivity level for each frequency band were assigned a value of zero.

A. TIME-SPACE VARIATIONS AND GLOBAL ANALYSIS

To access the total exposure level, we adopt the total E-field Equivalent 900 band ($E_{\text{Total Eq900}}$) as the metric. It is given by

$$E_{\text{Total Eq900}} = \sqrt{\sum_{i}^{N_f} \frac{L_{900}^2}{L_{fi}^2} E_{fi}^2} \tag{1}$$

where E_{fi} is the E-field strength at frequency f_i, N_f is the number of frequency bands, and L_{fi} is the reference level linked to frequency f_i [3]. Here, L_{900} is 41 V m^{-1}.

Measurements have been conducted in various modes of moving transportation, including underground locations where GPS does not work, making it challenging to determine the exact measurement location. As depicted in Fig. 4, the E-field strength exhibits variations along the journey on the metro line 4. These fluctuations are caused by the fast-changing tunnel environment and the distances to the surrounding active base stations. Therefore, in such cases, a statistical analysis, such as Cumulative Distribution Function (CDF), can be employed. As illustrated in Fig. 5a, on the metro line 6, the frequency bands 700 MHz, 800 MHz have a greater median than other bands. Additionally, exposure from 3500 MHz band is observed throughout journeys on metro line 6, which is uncommon for measurements taken in other metro lines. This is attributed to portions of metro line 6 being above ground, in contrast to the predominantly underground routes of most other metro lines. Another example, illustrated in Fig. 5b, demonstrates that a significant proportion of measurements recorded on RER lines are close to 0. Notably, the highest 99th percentile value is observed on RER D.

Here, the RMS of E-field strength is used as the primary statistical characteristic. This choice is motivated by the emphasis on exposure linked to the power density, which is proportional to the square of the E-field strength. Similarly, the standard deviation is derived from the square root of the standard deviation of the squared E-field strength values. In Table 4 and Table 5 presented in the appendix, we evaluate

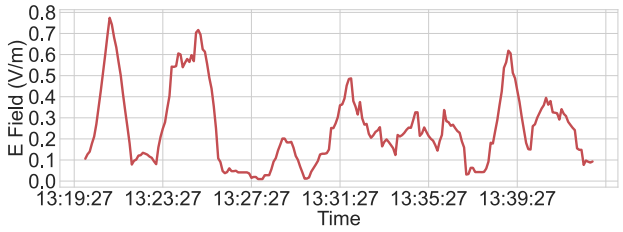


FIGURE 4. E-field variations along metro line 4 (Averaging for 1 minute).

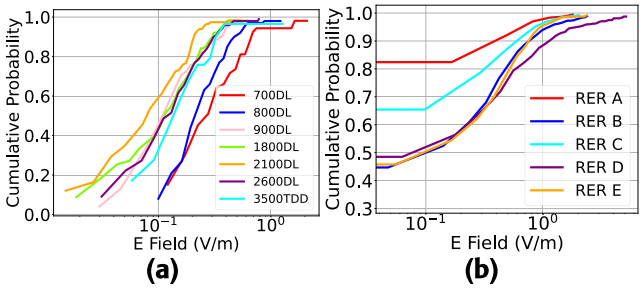


FIGURE 5. CDF in terms of (a) different frequencies on metro line 6, and (b) $E_{\text{Total Eq900}}$ in different RER lines. The x-axis is represented on a logarithmic scale, and the CDF is plotted up to the 99th percentile, rather than the maximum value.

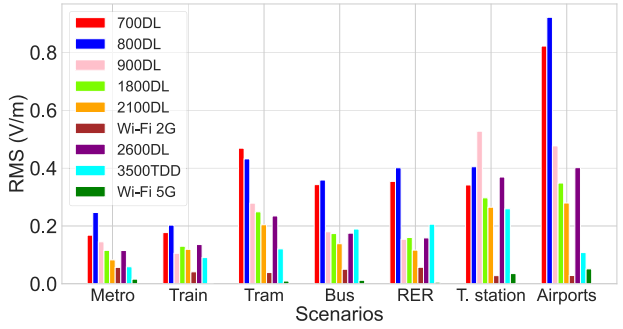


FIGURE 6. Barplot of average E-field level in different scenarios.

the RMS value, standard deviation which measures the dispersion or spread of data, and the 95th percentile, which is more robust given the short sampling time and the fact that the experimenter takes the measurements while moving on the vehicle. It should be noted that the results displayed in the tables are not raw data; the correction coefficients have been applied. Fig. 6 shows the RMS value of the E-field in different scenarios across frequencies.

The exposure measurements conducted over frequency bands of 700, 800, 900, 1800, 2100, 2600, and 3500 MHz exhibit overall mean values of 0.39, 0.43, 0.30, 0.21, 0.18, 0.24, and 0.18 V m^{-1}, respectively. A significant dispersion is evident, as indicated by the coefficient of variation (Std./Mean) for all frequency bands: 2.93, 3.38, 3.01, 2.19, 2.17, 2.50, and 3.46, respectively. This dispersion can be attributed to the substantial heterogeneity of the coverage. Tables 4 and 5 also demonstrate that exposure levels, even for the rare event such as the 95th percentile, are relatively low and well below the ICNIRP guidelines [3]. Additionally, the tables reveal that the 95th percentile of E-field levels is often less than 2 times the mean value, which implies that the measurements do not contain many extreme outliers, and outliers are not large enough to significantly impact the average value. The maximum mean exposure values are observed in train stations and airports, and regardless of the frequency, they remain below 1 V m^{-1}. Specifically, the maximum mean exposure values for the investigated frequency bands are as follows: 0.92, 0.97, 0.98, 0.42, 0.35, 0.60, and 0.42 V m^{-1}.

Dealing with 5G, Table 5 indicates that 3500 MHz is not deployed or has just begun in the metro, where most measurements are equal to 0. Meanwhile, the underground setting of the metro effectively shields against the influence of nearby base stations, resulting in comparatively low values at this band. In most public transportation settings, the overall mean value is small at 3500 MHz because 5G NR is not widely accessible, leading to a standard deviation-to-mean ratio close to 3.5. Measurements conducted mainly outside (bus, tram, train, RER) show higher exposure than those performed mainly underground, such as in the metro.

Regarding the Wi-Fi 2.4 and 5 GHz bands, the overall mean values are 0.05 and 0.02 V m^{-1}. As illustrated in Fig. 6, these bands consistently show the lowest E-field levels across various scenarios. Therefore, our subsequent analyses will predominantly concentrate on the cellular frequency bands.

B. ADVANCED CHARACTERIZATION OF DISTRIBUTION USING FOUR MOMENTS IN STATISTICS

As mentioned in the previous section, the mean values of data are relatively small, accompanied by large coefficients of variation, and the median is usually close to zero. Therefore, a majority of measurements exhibit low values. Given the challenges of accurately determining the location of each measurement and establishing connections to nearby base stations, we focus on describing the overall statistical distribution of each dataset. In this section, the data will be analyzed based on four statistical moments: Mean, Variance, Skewness, and Kurtosis. The K-means clustering technique will be applied to these four parameters for different subsets of the data.

It is important to note that in this section, the statistical parameters are computed from data that has been processed with a 1-minute moving average. Additionally, the analysis utilizes power density as the random variable, rather than the E-field strength, to better capture the variability and characteristics of the exposure.

1) SKEWNESS AND KURTOSIS PARAMETERS

In statistical analysis, the primary objective is usually to characterize the mean and variance of a dataset, representing the first and second moments. A further characterization of the dataset involves *Skewness* and *Kurtosis*, which are based on the third and fourth moments, respectively.

Skewness which defines the normalized central moment of third-order data (2), characterizes the shape of a data

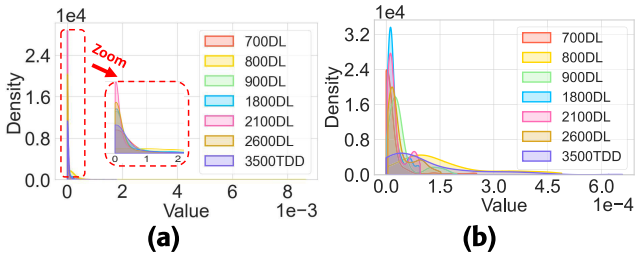


FIGURE 7. PDF of power density in (a) RER A and (b) Massy train station.

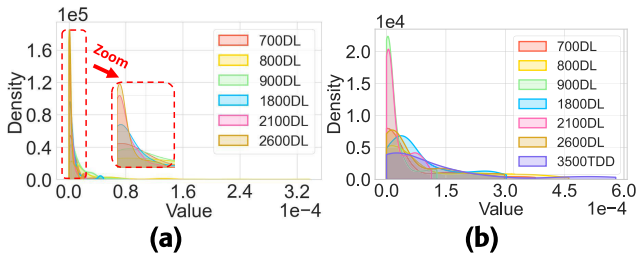


FIGURE 8. PDF of power density in (a) metro line 9 and (b) Austerlitz train station.

distribution. It quantifies [37] the degree of symmetry within a given distribution of global values.

$$S = \frac{\sum_{i=1}^{N}(X_i - \overline{X})^3/N}{\sigma^3} \tag{2}$$

where the standard deviation, σ, is the square root of the variance. This quantity can take on negative, zero, or positive values. It is unitless and often compared to the normal distribution to analyse symmetry.

The kurtosis [38] quantifies the degree of steepness or flatness in the distribution of all values within the population. Intuitively, the kurtosis provides insight into the sharpness of peaks within the distribution. It is defined in Eq. (3) as the standard fourth-order central moment of the sample.

$$K = \frac{\sum_{i=1}^{N}(X_i - \overline{X})^4/N}{\sigma^4} - 3 \tag{3}$$

2) SKEWNESS AND KURTOSIS OF MEASUREMENTS IN PUBLIC TRANSPORTATION FACILITIES

As depicted in Table 6, all measurements across different frequency bands exhibit positive skewness. The majority of the datasets demonstrate positive kurtosis, with only a few showing negative kurtosis. It is important to note that certain scenarios lack valid measurements across some frequency bands. For instance, measurements in most metro lines at the 3500 MHz band register as 0, implying that their associated variance, skewness, and kurtosis would also default to 0. However, these zero values should not be misconstrued as indicative of a distribution approximating normality. They are, instead, indicative of invalid or absent data. To address this, a preliminary filtering step will be incorporated to ascertain the validity of the dataset prior to computing skewness and kurtosis. Specifically, a dataset will only be considered valid for further statistical analysis if its 95th percentile value is non-zero. If this criterion is not met, the resulting zero values for skewness and kurtosis will be deemed invalid and disregarded.

A positive skewness indicates that the tail of the distribution of the data extends towards higher values, suggesting an excess of low values. For example, in the aforementioned table, we can find that data measured in RER A gives the almost highest skewness value for all frequencies. Fig. 7a illustrates that the shape of data distributions in RER A is extremely asymmetric with the majority of values concentrate around zero. In contrast, as shown in Fig. 7b, relatively smaller skewness values of data distributions in Massy train station lead to more symmetric distribution shape, indicating a relatively symmetric distribution with minor differences in the shapes of the left and right tails. This suggests a small disparity between the mean and median values, resulting in an overall flatter, smoother shape of the distribution.

Fig. 8 provides two examples with relatively higher (Fig. 8a) and lower (Fig. 8b) kurtosis values. What we can observe is that the distribution of data in line 9 possesses a sharper distribution with more extreme values, especially in 3500 MHz band. Conversely, the data in Austerlitz train station has a flatter distribution with a relative lack of extreme values, which also reflects a more dispersed nature of the data relative to the mean.

From the aforementioned figures, it is observed that distributions exhibiting relatively higher skewness values, as depicted in Figures 7a and 8a, also tend to possess relatively higher kurtosis values. While skewness and kurtosis may exhibit correlation in certain instances, their association is not consistently strong. Specifically, high skewness can be correlated with kurtosis, particularly in positively skewed distributions. For instance, when a distribution displays a prolonged right tail, the skewness value may rise, accompanied by an increase in kurtosis. In contrast, for a symmetric normal distribution, both skewness and kurtosis values are typically zero, indicating a lack of significant correlation. In summary, the relationship between skewness and kurtosis is not universally applicable and is influenced by the specific characteristics of the data distribution. When interpreting data distributions, skewness and kurtosis should usually be considered at the same time to fully understand the characteristics of the distribution.

By combining the four statistical moments for each dataset at specific frequencies, we can construct a Mean-Variance-Skewness-Kurtosis (*MVSK*) model to characterize the distribution.

C. ANALYSIS OF CLUSTERING RESULTS BASED ON MVSK MODEL AND K-MEANS METHODS

1) K-MEANS METHOD

The K-means clustering method [39] is an unsupervised algorithm specifically formulated to identify optimal cluster

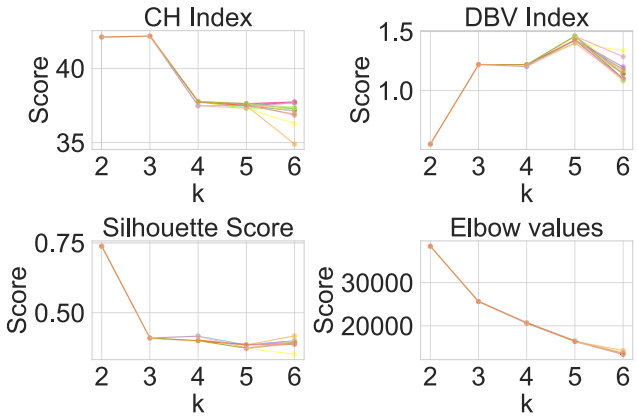


FIGURE 9. Results of the CH Index, DBV Index, silhouette score, and elbow method over 20 repetitions.

assignments by minimizing the sum of squared distances between data points and their respective cluster centroids.

Our objective is to categorize datasets with similar data distributions into the same group. To achieve this, we utilize the four moments mentioned earlier to characterize the distribution of a given dataset at specific frequencies, serving as input for the K-means method. The input comprises M vectors of statistical parameters, each vector of size $4 * N$ expressed as:

$$\begin{aligned} &\{\mu_{f_1}, \sigma^2_{f_1}, S_{f_1}, K_{f_1}, \ldots, \\ &\mu_{f_i}, \sigma^2_{f_i}, S_{f_i}, K_{f_i}, \ldots, \\ &\mu_{f_{N_f}}, \sigma^2_{f_{N_f}}, S_{f_{N_f}}, K_{f_{N_f}}\} \quad \text{for } i = 1 \ldots N_f \end{aligned} \tag{4}$$

where M is the number of selected scenarios. Therefore, each dataset has 28 parameters to characterize its distribution.

To determine the optimal number of clusters k, we employ four criteria: the Calinski-Harabasz Index (CH Index), Davies-Bouldin Variation Index (DBV Index), Silhouette Score, and the Elbow Method [40], [41], [42]. The following briefly outlines what each of these criteria measures:

1) The CH Index quantifies the ratio of between-cluster variance to within-cluster variance. Therefore, **higher** values indicate better-defined and more separated clusters.
2) The DBV Index assesses the compactness and separation of clusters based on the average similarity between each cluster and its most similar cluster. **Lower** values indicate better-defined clusters with greater separation.
3) The Silhouette Score measures how well-defined a cluster is, considering both the average distance within a cluster and the average distance between clusters. It ranges from -1 to 1, with **higher** values indicating better cluster definition.
4) The Elbow Method aids in determining the optimal K by identifying **the "elbow" point** in a plot of within-cluster sums of squares versus the number of clusters.

TABLE 3. Clustering results (MVSK values) for $k = 3$ over all frequency bands.

	Results
Cluster 1	M3-6, 8, 11-14, RER E, line P, All trams (except T1), 5 buses, All train stations (except CDG), CDG airport
Cluster 2	M1, 2, 7, 9, 10, RER B, D, T1, 3 buses, Line H, CDG train station, Orly airport
Cluster 3	RER A, C

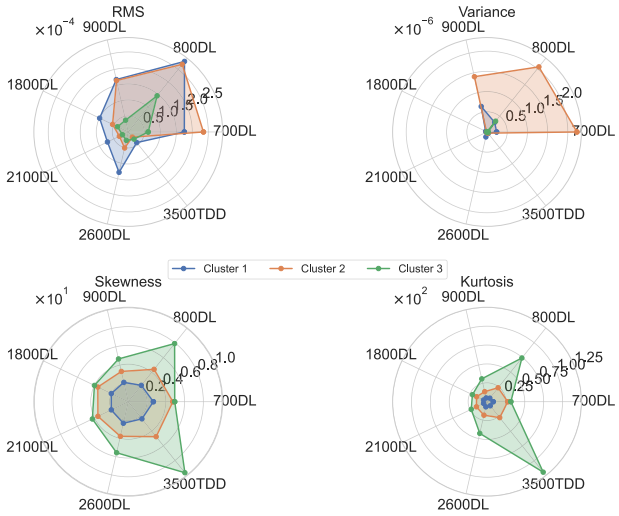


FIGURE 10. Radar charts of clustering results, for 4 moments: ***MVSK*** across all frequencies.

Figure 9 depicts the outcomes derived from applying the four specified criteria to the input dataset using the K-means approach. We conduct a comprehensive analysis of the results for each criterion, aiming to determine the optimal value of k for the current dataset. To enhance result robustness, we perform the clustering for each k for 20 times, with the outcomes of each repetition represented in different colors. After analyzing all repetitions, the optimal k values derived from these four methods are 3, 2, 2, and 3, respectively. Hereafter, we adopt $k = 3$ for further analysis.

2) CLUSTERING RESULTS

Table 3 displays the clustering results obtained using the K-means method. The table displays an overview of the assignment of each type of transportation to its respective cluster.

To better show the distinguishing characteristics of the three clusters, radar charts depicting different frequency bands and moments for members in the three clusters are provided in Fig. 10. From the figure, one can clearly observe that members in cluster 1 have the highest RMS values in most of the bands, moderate variance, lowest skewness and kurtosis values. Members in cluster 2 have high RMS and variance values, moderate skewness, and kurtosis values. Members in cluster 3 have the lowest

TABLE 4. RMS, variance, and 95th percentile.

	700			800			900			1800		
	μ	σ	P95	μ	σ	P95	μ	σ	P95	μ	σ	P95
All public transportation scenarios												
RER	0.354	1.044	0.556	0.401	1.748	0.612	0.154	0.296	0.329	0.16	0.312	0.341
Metro	0.168	0.475	0.324	0.247	0.619	0.441	0.145	0.294	0.293	0.116	0.269	0.23
Bus	0.343	0.781	0.674	0.359	0.614	0.768	0.18	0.294	0.398	0.174	0.294	0.377
Train	0.177	0.363	0.397	0.203	0.443	0.356	0.105	0.186	0.229	0.13	0.243	0.294
Tram	0.469	1.041	0.867	0.432	0.789	0.879	0.279	0.487	0.601	0.249	0.44	0.501
Train station	0.342	0.787	0.64	0.405	0.685	0.862	0.528	1.32	0.915	0.297	0.616	0.605
Airport	0.822	1.982	1.49	0.922	2.096	1.617	0.477	1.061	0.88	0.349	0.64	0.612
RER												
RERA	0.223	0.527	0.465	0.494	2.489	0.555	0.148	0.322	0.294	0.137	0.276	0.298
RERB	0.356	0.829	0.612	0.316	0.707	0.575	0.118	0.206	0.253	0.148	0.246	0.331
RERC	0.299	1.076	0.458	0.347	0.992	0.601	0.162	0.3	0.352	0.164	0.32	0.35
RERD	0.705	1.496	1.167	0.559	1.122	1.139	0.186	0.329	0.395	0.192	0.361	0.403
RERE	0.354	0.93	0.584	0.334	0.739	0.646	0.152	0.259	0.315	0.185	0.359	0.376
Metro												
M1	0.131	0.269	0.331	0.348	0.712	0.746	0.184	0.388	0.329	0.158	0.367	0.225
M2	0.288	0.696	0.453	0.402	0.873	0.683	0.176	0.331	0.323	0.16	0.345	0.252
M3	0.178	0.387	0.414	0.237	0.618	0.319	0.157	0.335	0.264	0.101	0.232	0.129
M4	0.156	0.301	0.415	0.241	0.408	0.531	0.185	0.302	0.409	0.153	0.259	0.361
M5	0.098	0.205	0.237	0.287	0.778	0.408	0.119	0.245	0.231	0.084	0.13	0.182
M6	0.338	0.772	0.649	0.281	0.489	0.526	0.179	0.274	0.389	0.154	0.224	0.322
M7	0.131	0.328	0.231	0.155	0.326	0.333	0.098	0.185	0.201	0.067	0.132	0.151
M8	0	0	0	0.103	0.193	0.264	0.054	0.084	0.119	0.015	0.027	0.036
M9	0.108	0.256	0.211	0.201	0.565	0.273	0.125	0.286	0.219	0.097	0.299	0.113
M10	0.164	0.468	0	0.257	0.808	0.147	0.132	0.191	0.281	0.152	0.34	0.298
M11	0.038	0.097	0	0.215	0.537	0.313	0.089	0.144	0.22	0.045	0.08	0.099
M12	0.098	0.206	0.237	0.184	0.401	0.313	0.19	0.344	0.407	0.138	0.26	0.293
M13	0.118	0.319	0.194	0.223	0.557	0.356	0.092	0.231	0.15	0.047	0.098	0.085
M14	0.13	0.298	0.248	0.117	0.21	0.266	0.113	0.21	0.244	0.068	0.12	0.143
Bus												
9a	0.373	0.77	0.785	0.271	0.474	0.608	0.159	0.253	0.397	0.147	0.274	0.331
9b	0.311	0.573	0.723	0.309	0.587	0.669	0.177	0.281	0.401	0.164	0.273	0.401
11	0.477	1.318	0.579	0.19	0.294	0.412	0.086	0.132	0.164	0.071	0.102	0.153
23	0.476	0.926	0.983	0.456	0.564	0.884	0.328	0.448	0.627	0.291	0.421	0.57
194	0.441	0.865	0.657	0.58	0.874	1.293	0.215	0.254	0.401	0.258	0.304	0.476
9105	0.291	0.524	0.66	0.338	0.554	0.694	0.126	0.188	0.279	0.137	0.267	0.273
9106	0.176	0.277	0.363	0.442	0.578	0.927	0.174	0.232	0.268	0.124	0.14	0.225
9110	0.312	0.694	0.551	0.323	0.629	0.674	0.18	0.305	0.392	0.178	0.285	0.359
Train												
Line H	0.064	0.176	0.146	0.141	0.302	0.267	0.066	0.12	0.131	0.038	0.071	0.086
Line P	0.263	0.44	0.611	0.266	0.53	0.531	0.143	0.221	0.293	0.196	0.295	0.47
Tram												
T1	0.255	0.526	0.472	0.251	0.403	0.465	0.174	0.315	0.338	0.156	0.229	0.307
T2	0.56	1.198	0.773	0.419	0.738	0.736	0.29	0.44	0.599	0.261	0.409	0.655
T3a	0.919	1.585	1.96	0.744	1.126	1.575	0.441	0.672	0.868	0.328	0.582	0.547
T3b	0.435	0.658	1.081	0.38	0.511	0.744	0.248	0.341	0.463	0.301	0.437	0.603
T4	0.46	0.849	0.979	0.458	0.693	0.973	0.367	0.63	0.755	0.244	0.357	0.543
T7	0.302	0.602	0.556	0.32	0.52	0.696	0.196	0.323	0.352	0.263	0.478	0.5
T8	0.436	1.066	0.674	0.278	0.382	0.49	0.271	0.421	0.636	0.227	0.373	0.464
T11	0.279	0.433	0.594	0.641	1.079	1.365	0.308	0.431	0.694	0.27	0.502	0.468
Train station												
CDG	0.34	0.981	0.529	0.297	0.673	0.555	0.825	1.827	1.572	0.157	0.323	0.308
Austerlitz	0.158	0.247	0.394	0.21	0.315	0.442	0.092	0.134	0.211	0.184	0.218	0.369
Est	0.45	0.986	0.643	0.478	0.811	0.747	0.298	0.426	0.539	0.314	0.406	0.455
Lyon	0.248	0.403	0.588	0.488	0.779	1.031	0.976	1.809	2.176	0.417	0.837	0.698
Nord	0.428	0.861	0.791	0.422	0.578	0.865	0.365	0.718	0.673	0.229	0.419	0.389
Marne-la-Vallée	0.269	0.398	0.54	0.479	0.889	0.797	0.277	0.361	0.532	0.262	0.321	0.485
Massy	0.102	0.173	0.262	0.254	0.258	0.424	0.131	0.18	0.228	0.094	0.1	0.18
Montparnasse	0.389	0.655	0.8	0.429	0.652	0.93	0.418	0.545	0.851	0.391	0.565	0.785
Saint-Lazare	0.234	0.368	0.564	0.424	0.649	1.054	0.201	0.333	0.475	0.328	0.856	0.635
Airpot												
CDG	0.562	1.658	0.902	0.801	1.533	1.579	0.392	0.645	0.802	0.306	0.729	0.484
Orly	0.903	2.072	1.637	0.965	2.228	1.661	0.506	1.14	0.88	0.364	0.592	0.66

RMS and variance values, but the highest skewness, and kurtosis.

In conclusion, the members in cluster 1 (e.g., Fig. 7b) exhibit a distribution that is relatively more flat and symmetric. Conversely, the distribution of members in cluster 3 (e.g., Fig. 7a) is asymmetric and not flat, indicating that a majority of the measurement points concentrate around 0. Notably, the distribution of members in cluster 2 (e.g., Fig. 8a) falls between the two other clusters, except the 3500 band.

TABLE 5. RMS, variance, and percentile 95.

	2100			2600			3500			Total (Eq900)		
	μ	σ	P95	μ	σ	P95	μ	σ	P95	μ	σ	P95
All public transportation scenarios												
RER	0.117	0.223	0.257	0.159	0.44	0.335	0.206	0.656	0.36	0.645	1.983	1.081
Metro	0.083	0.198	0.152	0.115	0.359	0.191	0.059	0.273	0	0.379	0.822	0.777
Bus	0.139	0.238	0.292	0.175	0.39	0.355	0.189	0.912	0.181	0.618	1.067	1.314
Train	0.119	0.273	0.219	0.136	0.304	0.278	0.091	0.254	0.187	0.353	0.625	0.848
Tram	0.204	0.351	0.445	0.235	0.429	0.472	0.121	0.306	0.244	0.807	1.388	1.550
Train station	0.265	0.509	0.556	0.369	0.823	0.704	0.259	0.539	0.566	0.884	1.606	1.670
Airport	0.279	0.524	0.503	0.402	0.679	0.74	0.108	0.226	0.254	1.492	3.196	2.791
RER												
RERA	0.105	0.212	0.223	0.128	0.319	0.239	0.193	0.584	0.319	0.632	2.650	0.935
RERB	0.113	0.198	0.24	0.151	0.28	0.316	0.346	0.858	0.666	0.609	1.153	1.181
RERC	0.113	0.207	0.256	0.144	0.269	0.335	0.164	0.625	0.23	0.563	1.507	1.016
RERD	0.144	0.229	0.343	0.262	0.766	0.407	0.114	0.267	0.237	1.044	2.047	1.959
RERE	0.136	0.302	0.275	0.178	0.359	0.364	0.114	0.24	0.287	0.597	1.240	1.036
Metro												
M1	0.101	0.204	0.183	0.092	0.192	0.181	0	0	0	0.462	0.876	1.127
M2	0.144	0.279	0.279	0.227	0.531	0.348	0.134	0.393	0	0.611	1.192	1.057
M3	0.078	0.208	0.116	0.046	0.09	0.107	0	0	0	0.371	0.782	0.793
M4	0.125	0.246	0.239	0.164	0.3	0.389	0	0	0	0.403	0.596	0.924
M5	0.058	0.105	0.145	0.091	0.169	0.22	0.113	0.371	0	0.366	0.876	0.642
M6	0.123	0.197	0.224	0.17	0.305	0.335	0.151	0.456	0.214	0.556	0.962	0.950
M7	0.037	0.07	0.088	0.053	0.115	0.114	0	0	0	0.251	0.484	0.601
M8	0.021	0.043	0.048	0.02	0.038	0.046	0	0	0	0.124	0.215	0.301
M9	0.078	0.231	0.098	0.061	0.111	0.139	0.007	0.028	0	0.292	0.741	0.455
M10	0.039	0.113	0.039	0.07	0.192	0.083	0	0	0	0.376	1.031	0.440
M11	0.038	0.072	0.06	0.037	0.068	0.111	0	0	0	0.253	0.579	0.448
M12	0.047	0.083	0.106	0.164	0.58	0.139	0	0	0	0.330	0.584	0.746
M13	0.034	0.071	0.064	0.052	0.101	0.106	0	0	0	0.292	0.628	0.516
M14	0.063	0.137	0.132	0.089	0.222	0.135	0	0	0	0.240	0.409	0.517
Bus												
9a	0.106	0.197	0.233	0.135	0.263	0.298	0.151	0.405	0.147	0.567	1.015	1.205
9b	0.12	0.213	0.266	0.189	0.547	0.315	0.091	0.283	0.157	0.550	0.880	1.353
11	0.041	0.07	0.093	0.061	0.113	0.111	0.022	0.053	0	0.588	1.514	0.797
23	0.191	0.257	0.358	0.222	0.34	0.462	0.245	0.504	0.459	0.861	1.184	1.610
194	0.222	0.301	0.404	0.267	0.414	0.466	0.061	0.111	0.149	0.874	1.271	1.702
9105	0.118	0.234	0.252	0.144	0.251	0.328	0.318	1.306	0.213	0.569	1.009	1.165
9106	0.142	0.208	0.227	0.173	0.212	0.322	0.048	0.096	0.103	0.566	0.696	1.114
9110	0.149	0.254	0.293	0.18	0.351	0.368	0.126	0.352	0.152	0.568	0.988	1.123
Train												
Line H	0.037	0.071	0.079	0.038	0.072	0.087	0.052	0.119	0.114	0.188	0.348	0.392
Line P	0.179	0.335	0.362	0.205	0.373	0.386	0.126	0.313	0.273	0.498	0.750	1.031
Tram												
T1	0.103	0.152	0.191	0.118	0.157	0.24	0.083	0.173	0.199	0.460	0.682	0.839
T2	0.229	0.373	0.624	0.213	0.437	0.397	0.086	0.154	0.215	0.875	1.467	1.867
T3a	0.33	0.507	0.761	0.366	0.615	0.649	0.228	0.489	0.437	1.446	2.071	2.906
T3b	0.267	0.384	0.538	0.258	0.319	0.543	0.184	0.327	0.405	0.768	0.956	1.554
T4	0.146	0.224	0.297	0.237	0.342	0.535	0.081	0.196	0.092	0.843	1.210	1.744
T7	0.175	0.251	0.384	0.233	0.448	0.444	0.066	0.123	0.168	0.588	0.921	1.123
T8	0.17	0.298	0.35	0.19	0.269	0.38	0.061	0.153	0.108	0.679	1.246	1.165
T11	0.233	0.308	0.544	0.303	0.469	0.654	0.082	0.191	0.148	0.873	1.324	1.871
Train station												
CDG	0.11	0.201	0.229	0.083	0.138	0.166	0.056	0.123	0.131	0.975	2.13	1.890
Austerlitz	0.104	0.155	0.221	0.174	0.243	0.391	0.269	0.511	0.459	0.398	0.515	0.846
Est	0.25	0.378	0.371	0.391	0.695	0.559	0.419	0.559	0.839	0.912	1.482	1.268
Lyon	0.319	0.465	0.75	0.594	1.033	1.257	0.387	0.754	0.829	1.289	2.007	2.798
Nord	0.33	0.739	0.578	0.567	1.192	0.943	0.059	0.122	0.14	0.890	1.424	1.584
Marne-la-Vallée	0.182	0.249	0.368	0.261	0.333	0.572	0.157	0.312	0.341	0.720	1.027	1.347
Massy	0.103	0.118	0.201	0.124	0.143	0.237	0.265	0.429	0.542	0.389	0.408	0.733
Montparnasse	0.352	0.439	0.693	0.356	0.542	0.715	0.246	0.515	0.518	0.890	1.035	1.621
Saint-Lazare	0.244	0.482	0.509	0.26	0.373	0.617	0.332	0.652	0.769	0.689	1.021	1.670
Airpot												
CDG	0.244	0.4	0.493	0.396	0.654	0.683	0.084	0.189	0.191	1.196	2.279	2.238
Orly	0.292	0.554	0.517	0.404	0.688	0.742	0.116	0.237	0.266	1.593	3.403	3.117

IV. CONCLUSION

Despite the increasing use of wireless communication, there remains a significant perception of risk that requires exposure monitoring. In this article, the exposure assessment has been carried out using a personal exposimeter in different forms of public transportation serving in the Paris area.

TABLE 6. Skewness and kurtosis values.

	700		800		900		1800		2100		2600		3500	
	Skew.	Kurt.	Skew.	Kurt.	Skew.	Kurt.	Skew.	Kurt.	Skew.	Kurt.	Skew.	Kurt.	Skew.	Kurt.
All public transportation scenarios														
RER	19.259	413.838	13.615	233.097	4.611	31.405	4.269	27.361	3.863	18.832	6.013	58.06	10.129	122.692
Metro	9.225	101.703	9.841	119.88	7.345	74.921	8.233	87.147	7.5	74.77	8.584	100.736	9.118	91.7
Bus	3.858	17.315	3.452	15.864	2.701	8.655	2.976	10.208	2.618	7.769	3.709	19.616	13.957	206.342
Train	4.145	17.896	4.417	23.606	4.576	26.423	4.916	28.679	5.443	34.225	6.278	42.468	4	16.982
Tram	6.844	60.273	6.248	52.895	3.083	11.783	3.098	11.852	4.06	21.505	5.585	43.611	4.589	25.074
Train station	6.987	63.83	3.128	12.113	7.751	68.721	4.669	31.703	3.052	12.134	6.089	49.063	7.207	71.268
Airport	5.642	38.14	6.569	51.028	4.893	31.209	3.395	16.026	3.146	13.198	2.714	8.249	6.567	59.712
RER														
RERA	4.047	19.068	10.183	107.031	4.591	29.715	4.314	25.609	4.234	22.881	6.831	59.803	8.219	88.466
RERB	7.066	58.063	5.566	37.605	3.722	15.538	3.185	11.623	3.963	20.351	3.833	20.623	5.081	28.074
RERC	5.852	44.086	5.654	41.031	4.754	32.369	3.685	17.167	4.255	23.515	4.322	26.22	11.129	150.122
RERD	6.003	37.013	5.872	36.67	2.878	9.584	4.234	21.723	2.749	7.488	4.9	28.455	3.789	21.125
RERE	2.7	8.125	4.717	25.296	1.522	1.908	1.703	2.518	2.148	4.406	2.128	5.006	4.564	23.714
Metro														
M1	4.509	20.347	4.709	22.842	4.901	24.231	4.87	23.629	4.906	24.234	4.686	22.7	N/A	N/A
M2	4.037	15.679	4.323	20.136	4.172	18.76	4.307	20.144	3.222	11.482	3.524	14.09	3.537	12.08
M3	2.694	6.299	2.893	8.641	2.306	5.006	2.4	6.706	2.657	7.597	2.844	7.516	N/A	N/A
M4	2.958	8.374	1.768	3.173	2.495	7.011	2.493	5.797	2.735	7.627	3.542	13.459	N/A	N/A
M5	1.759	2.755	3.29	9.384	2.622	5.969	2.81	7.331	2.921	7.82	3.128	9.465	3.213	8.618
M6	3.263	10.78	1.771	2.141	2.335	5.028	1.512	1.567	1.603	1.684	1.484	1.048	3.339	10.898
M7	5.025	27.734	3.708	16.232	3.29	13.968	3.036	9.952	4.289	23.758	6.71	47.946	N/A	N/A
M8	N/A	N/A	0	0	0	0	0	0	0	0	0	0	N/A	N/A
M9	5.282	34.649	3.693	18.317	1.904	3.858	3.53	12.763	2.743	7.67	2.492	6.371	N/A	N/A
M10	4.607	25.373	5.288	31.677	1.549	2.766	3.245	10.695	5.301	32.166	5.356	32.709	N/A	N/A
M11	0	0	1.942	3.297	2.626	9.013	0	0	0	0	0	0	N/A	N/A
M12	2.903	7.449	4.203	17.909	2.366	5.975	1.873	3.585	2.629	5.702	3.969	14.885	N/A	N/A
M13	3.794	14.882	3.385	11.416	1.799	3.175	2.086	4.542	0	2.77	2.76	9.782	N/A	N/A
M14	1.946	2.741	4.225	21.415	1.855	4.359	1.988	4.214	2.136	4.339	2.139	4.344	N/A	N/A
Bus														
9a	2.636	6.583	2.49	5.774	2.495	5.407	2.936	9.202	2.804	7.145	2.463	5.136	3.254	10.142
9b	3.99	16.25	2.751	8.018	2.642	7.775	3.313	11.418	4.131	18.908	3.316	11.067	4.675	24.307
11	2.445	6.09	2.324	4.63	1.594	1.574	1.458	0.93	2.152	3.65	2.124	3.707	0	0
23	2.866	8.274	0.957	0.167	0.888	0.265	1.203	1.094	1.138	0.508	1.416	1.483	2.216	5.998
194	2.643	7.945	1.929	2.864	0.999	-0.492	1.039	-0.021	1.217	1.028	2.661	7.923	0	2.093
9105	4.211	20.311	2.08	4.315	2.103	4.437	4.329	23.7	3.728	15.944	2.711	8.017	7.011	49.013
9106	2.214	4.658	2.534	6.698	1.462	2.439	0.735	-0.446	0.297	-1.082	0.935	0.145	0	0
9110	2.212	4.198	3.818	17.001	2.74	8.867	3.583	16.202	2.201	4.532	3.532	14.151	5.784	35.536
Train														
Line H	3.355	10.904	5.432	31.798	2.8	8.535	4.782	25.416	4.365	21.898	5.614	35.502	4.035	17.571
Line P	2.437	5.33	3.389	13.348	3.116	11.384	3.132	11.115	3.456	13.14	3.978	16.11	2.419	5.239
Tram														
T1	4.238	23.659	4.413	24.28	4.808	28.071	1.835	3.563	2.665	9.104	1.779	3.031	4.372	21.728
T2	1.548	1.468	2.197	4.876	1.954	4.543	4.035	16.613	4.107	17.064	4.509	21.015	1.986	3.836
T3a	2.789	8.008	2.254	6.522	2.146	5.983	1.47	2.456	2.946	10.373	3.7	14.169	1.825	2.855
T3b	3.432	12.03	0.887	-0.236	1.807	4.587	2.726	8.004	2.37	5.601	2.53	6.746	3.047	10.751
T4	2.425	5.504	1.18	0.761	1.28	0.293	1.857	3.217	1.522	1.376	2.554	6.394	3.146	9.569
T7	1.709	2.48	2.017	4.294	1.465	1.628	2.45	5.677	1.84	2.495	1.604	1.882	1.999	3.452
T8	3.459	11.93	1.904	4.703	3.844	15.881	3.933	15.258	2.953	9.123	1.404	1.482	2.114	3.39
T11	2.122	3.819	2.65	6.324	2.228	3.731	1.259	0.757	2.212	3.705	2.04	2.957	3.343	10.994
Train station														
CDG	7.258	63.381	4.486	24.869	4.395	20.757	3.063	10.802	3.136	10.641	2.128	5.98	3.353	12.123
Austerlitz	1.885	2.52	1.393	0.561	2.173	3.79	1.251	0.446	1.261	0.14	1.448	0.913	1.874	2.699
Est	2.877	8.923	1.521	2.986	1.691	2.983	0.995	2.485	2.122	8.088	3.035	13.163	0.589	-0.718
Lyon	4.126	18.296	1.888	2.508	3.009	8.123	2.429	6.479	2.164	3.996	2.499	6.101	3.477	11.571
Nord	3.797	18.946	1.678	2.201	3.429	12.065	2.568	7.339	2.243	6.163	4.405	23.982	2.909	8.984
Marne-la-Vallée	1.261	0.391	2.027	3.177	2.642	7.873	1.205	0.059	1.933	3.836	1.366	1.604	1.722	3.966
Massy	4.023	16.662	1.112	0.463	2.007	3.206	1.821	2.467	1.383	0.471	1.448	0.973	1.687	2.885
Montparnasse	5.069	28.203	2.804	9.984	2.333	6.921	2.778	12.177	1.362	2.064	1.242	1.313	2.535	5.822
Saint-Lazare	2.124	3.561	2.626	6.063	3.237	10.85	4.232	19.661	3.966	17.298	2.011	2.779	2.965	8.852
Airpot														
CDG	3.022	12.049	1.388	1.11	1.883	3.317	1.875	3.488	1.823	4.017	3.911	18.362	4.147	18.181
Orly	4.758	26.836	5.635	36.769	4.413	24.202	3.094	12.705	2.824	10.076	2.392	6.042	6.004	48.235

The cellular frequency bands of interest were those used by 2G, 3G, 4G, and 5G networks, as well as that of Wi-Fi. In public transportation, which is often heavily used, the means and measurement protocols recommended by the [6] and [4] are not always suitable. They allow relatively precise measurements to be made but in fact, are rarely used in places where there are a lot of people such as in public transportation or commercial areas. In this case, personal exposimeters are well-suitable but have the main disadvantage of having the measurement influenced by the body of the person wearing this equipment. To evaluate human exposure along a route we used the property of

independence of incident exposures (i.e., induced by base stations) and those of the correction coefficients of the values estimated by the exposimeter. In this case, the actual exposure (the average over the route of the power density) is given by the product of the average (i.e., RMS) of the correction coefficients by the average of exposures measured by the PEM. Thousands of measurements have been carried out in RER, Metro, Bus, Train, Tram, Train Stations as well as Airports. The medians of measurement distributions show that a large part of the exposures are below or close to the detection limit of the dosimeters. All measurements, across different frequency bands, exhibit positive skewness. This indicates that the tail of the distribution of the data extends towards higher values, suggesting an excess of low values. While in certain instances, skewness and kurtosis exhibit correlation their association is not consistently strong.

A Mean-Variance-Skewness-Kurtosis model has been built to characterize the distribution and perform a cluster analysis using the K-means clustering method. To enhance result robustness, the optimal k values have been studied and $k = 3$ has been used. Members of Cluster 1 exhibit the highest RMS values, moderate variance, and the lowest skewness and kurtosis, indicative of distributions that are relatively flatter and more symmetric. In contrast, Cluster 2 members have high RMS values and variance, alongside moderate skewness and kurtosis. Their distributions, unlike those in Cluster 1, are asymmetric and not flat, suggesting that a majority of the measurement points are near 0. Cluster 3 members display, across most frequency bands, the highest levels of skewness and kurtosis, pointing to a significant deviation from the normal distribution.

Overall, the RF-EMF exposures measured in this study indicate that across the 700, 800, 900, 1800, 2100, 2600, and 3500 MHz frequency bands, the average exposure values are 0.39, 0.43, 0.30, 0.21, 0.18 0.24, and 0.18 $\mathrm{V\,m^{-1}}$, respectively. For these frequency bands, the 95th percentile values are 0.64, 0.78, 0.52, 0.43, 0.38, 0.47 and 0.31 $\mathrm{V\,m^{-1}}$, respectively. Thus, both the mean and 95th percentile exposure levels are significantly below the limits set by the ICNIRP guidelines.

APPENDIX
DATA TABLES

See Tables 4–6.

REFERENCES

[1] Bankmycell. (2024). *How Many Smartphones are in the World?*. Accessed: May 13, 2024. [Online]. Available: https://www.bankmycell.com/blog/how-many-phones-are-in-the-world

[2] Ministère de l'Économie et des Finances. (2022). *Digital Barometer*. Accessed: May 13, 2024. [Online]. Available: https://www.economie.gouv.fr/cge/barometre-numerique-2022

[3] International Commission on Non-Ionizing Radiation Protection, "Guidelines for limiting exposure to electromagnetic fields (100 KHz to 300 GHz)," *Health Phys.*, vol. 118, no. 5, pp. 483–524, 2020.

[4] International Electrotechnical Commission. (2016). *Determination of RF Field Strength, Power Density and SAR in the Vicinity of Base Stations for the Purpose of Evaluating Human Exposure*. Accessed: May 13, 2024. [Online]. Available: https://webstore.iec.ch/publication/64934

[5] J. Wiart, *Radio-Frequency Human Exposure Assessment: From Deterministic to Stochastic Methods*. Hoboken, NJ, USA: Wiley, 2016, doi: 10.1002/9781119285137.

[6] Agence Nationale des Fréquences. (2019). *Measurement Protocol*. Accessed: May 13, 2024. [Online]. Available: https://www.anfr.fr/maitriser/les-installations-radioelectriques/protocole-de-mesure

[7] Eurobarometer. (2010). *Electromagnetic Fields*. Accessed: May 13, 2024. [Online]. Available: https://europa.eu/eurobarometer/surveys/detail/843

[8] T. Koh, J. Choi, M. Seo, H.-D. Choi, and K. Kim, "Factors affecting risk perception of electromagnetic waves from 5G network base stations," *Bioelectromagnetics*, vol. 41, no. 7, pp. 491–499, Aug. 2020, doi: 10.1002/bem.22290.

[9] A. Narayanan, M. I. Rochman, A. Hassan, B. S. Firmansyah, V. Sathya, M. Ghosh, F. Qian, and Z.-L. Zhang, "A comparative measurement study of commercial 5G mmWave deployments," in *Proc. IEEE Conf. Comput. Commun.*, May 2022, pp. 800–809, doi: 10.1109/INFOCOM48880.2022.9796693.

[10] M. I. Rochman, V. Sathya, D. Fernandez, N. Nunez, A. S. Ibrahim, W. Payne, and M. Ghosh, "A comprehensive analysis of the coverage and performance of 4G and 5G deployments," *Comput. Netw.*, vol. 237, Dec. 2023, Art. no. 110060, doi: 10.1016/j.comnet.2023.110060.

[11] P. Gajšek, P. Ravazzani, J. Wiart, J. Grellier, T. Samaras, and G. Thuróczy, "Electromagnetic field exposure assessment in Europe radiofrequency fields (10 MHz–6 GHz)," *J. Exposure Sci. Environ. Epidemiology*, vol. 25, no. 1, pp. 37–44, Jan. 2015.

[12] Exem. (2020). *National Wave Observatory*. Accessed: May 13, 2024. [Online]. Available: https://www.observatoiredesondes.com/fr/

[13] Regulatory Authority for Electronic Communications and Postal Services. (2024). *Ratel: Exposure Map*. Accessed: May 13, 2024. [Online]. Available: https://emf.ratel.rs/eng/rezultati-merenja/

[14] L. Diez, R. Agüero, and L. Muñoz, "Electromagnetic field assessment as a smart city service: The SmartSantander use-case," *Sensors*, vol. 17, no. 6, p. 1250, May 2017, doi: 10.3390/s17061250.

[15] S. Wang, W. B. Chikha, Y. Zhang, J. Liu, E. Conil, O. Jawad, L. Ourak, and J. Wiart, "RF electromagnetic fields exposure monitoring using drive test and sensors in a French city," in *Proc. 35th Gen. Assem. Sci. Symp. Int. Union Radio Sci. (URSI GASS)*, Aug. 2023, pp. 1–4, doi: 10.23919/URSIGASS57860.2023.10265402.

[16] Agence Nationale des Fréquences. (2021). *Cartoradio*. Accessed: May 13, 2024. [Online]. Available: https://www.cartoradio.fr/#/

[17] (2022). *Study of the Public's Exposure to Radio Waves*. Accessed: May 13, 2024. [Online]. Available: https://www.anfr.fr/fileadmin/userupload/2022Analysemesures2021vf.pdf

[18] S. Iakovidis, C. Apostolidis, A. Manassas, and T. Samaras, "Electromagnetic fields exposure assessment in Europe utilizing publicly available data," *Sensors*, vol. 22, no. 21, p. 4841, 2022, doi: 10.3390/s22218481.

[19] Espace Presse VELUX. (2018). *L'excès de Temps Passè à L'intérieur: Un Enjeu de Santé Pour la Génération 'Indoor'*. Accessed: May 13, 2024. [Online]. Available: https://presse.velux.fr/generation-indoor/

[20] Moovit Insights. (2022). *Public Transport Data Classified by Country and City*. Accessed: May 11, 2024. [Online]. Available: https://moovitapp.com/insights/fr/Indice_deTransportenCommundeMoovit-countries

[21] D. Suka, M. Simic, and P. Pejovic, "On the assessment of exposure from LTE 800-MHz downlink frequency band through the time-averaged and integral-based measure," *Radiat. Protection Dosimetry*, vol. 198, pp. 454–466, Jun. 2022, doi: 10.1093/rpd/ncac075.

[22] B. Xu, D. Colombi, C. Törnevik, F. Ghasemifard, and J. Chen, "On actual maximum exposure from 5G multicolumn radio base station antennas for electromagnetic field compliance assessment," *IEEE Trans. Electromagn. Compat.*, vol. 63, no. 5, pp. 1680–1689, Oct. 2021, doi: 10.1109/TEMC.2021.3090107.

[23] P. Baracca, A. Weber, T. Wild, and C. Grangeat, "A statistical approach for RF exposure compliance boundary assessment in massive MIMO systems," in *Proc. 22nd Int. ITG Workshop Smart Antennas*, Mar. 2018, pp. 1–6.

[24] M. Celaya-Echarri, L. Azpilicueta, P. Lopez-Iturri, E. Aguirre, S. De Miguel-Bilbao, V. Ramos, and F. Falcone, "Spatial characterization of personal RF-EMF exposure in public transportation buses," *IEEE Access*, vol. 7, pp. 33038–33054, 2019, doi: 10.1109/ACCESS.2019.2903405.

[25] H. Jalilian, M. Eeftens, M. Ziaei, and M. Röösli, "Public exposure to radiofrequency electromagnetic fields in everyday microenvironments: An updated systematic review for Europe," *Environ. Res.*, vol. 176, Sep. 2019, Art. no. 108517, doi: 10.1016/j.envres.2019.05.048.

[26] S. Aerts, D. Plets, A. Thielens, L. Martens, and W. Joseph, "Impact of a small cell on the RF-EMF exposure in a train," *Int. J. Environ. Res. Public Health*, vol. 12, no. 3, pp. 2639–2652, Feb. 2015, doi: 10.3390/ijerph120302639.

[27] K. Gryz and J. Karpowicz, "Radiofrequency electromagnetic radiation exposure inside the metro tube infrastructure in Warszawa," *Electromagn. Biol. Med.*, vol. 34, pp. 265–273, Jul. 2015, doi: 10.3109/15368378.2015.1076447.

[28] N. Loizeau, M. Zahner, J. Schindler, C. Stephan, J. Fröhlich, M. Gugler, T. Ziegler, and M. Röösli, "Comparison of ambient radiofrequency electromagnetic field (RF-EMF) levels in outdoor areas and public transport in Switzerland in 2014 and 2021," *Environ. Res.*, vol. 237, Nov. 2023, Art. no. 116921, doi: 10.1016/j.envres.2023.116921.

[29] European Commission. (2021). *The Funding & Tenders Portal*. Accessed: May 13, 2024. [Online]. Available: https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/topic-details/horizon-hlth-2021-envhlth-02-01

[30] SEAWave Project. (2022). *Scientific-based Exposure and Risk Assessment of Radiofrequency and mm-Wave Systems From Children to Elderly (5G and Beyond)*. Accessed: May 13, 2024. [Online]. Available: https://seawave-project.eu

[31] X. Wu, V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, P. S. Yu, Z.-H. Zhou, M. Steinbach, D. J. Hand, and D. Steinberg, "Top 10 algorithms in data mining," *Knowl. Inf. Syst.*, vol. 14, no. 1, pp. 1–37, 2008.

[32] W. B. Chikha, Y. Zhang, J. Liu, S. Wang, S. Sandeep, M. Guxens, A. F. Veludo, M. Röösli, W. Joseph, and J. Wiart, "Assessment of radio frequency electromagnetic field exposure induced by base stations in several micro-environments in France," *IEEE Access*, vol. 12, pp. 21610–21620, 2024, doi: 10.1109/ACCESS.2024.3363914.

[33] Microwave Vision Group. (2024). *EME Spy Evolution*. Accessed: May 13, 2024. [Online]. Available: https://www.mvg-world.com/fr/produits/securite-RF/public-RF-safety/eme-spy-evolution

[34] J. Toftgard, S. N. Hornsleth, and J. B. Andersen, "Effects on Portable Antennas of the Presence of a Person," *IEEE Trans. Antennas Propag.*, vol. 41, no. 6, pp. 739–746, 1993, doi: 10.1109/8.250451.

[35] Tektronix. (2024). *RSA306B USB Spectrum Analyzer*. Accessed: May 13, 2024. [Online]. Available: https://www.tek.com/fr/products/spectrum-analyzers/rsa306

[36] Île-de-France Mobilités. (2021). *Presentation of Rer (Réseau Express Régional)*. Accessed: May 13, 2024. [Online]. Available: https://www.iledefrance-mobilites.fr/le-reseau/services-de-mobilite/rer/presentation-plans-rer

[37] D. P. Doane and L. E. Seward, "Measuring skewness: A forgotten statistic?" *J. Stat. Educ.*, vol. 19, no. 2, pp. 1–18, 2011, doi: 10.1080/10691898.2011.11889611.

[38] D. Ruppert, "What is kurtosis? An influence function approach," *Amer. Stat.*, vol. 41, no. 5, pp. 1–5, 1987, doi: 10.1080/00031305.1987.10475431.

[39] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A K-means clustering algorithm," *Appl. Statist.*, vol. 28, no. 1, pp. 100–108, 1979.

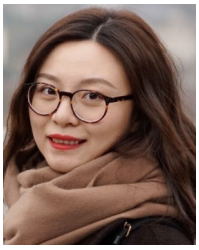
[40] T. Calinski and J. Harabasz, "A dendrite method for cluster analysis," *Commun. Statist. Theory Methods*, vol. 3, no. 1, pp. 1–27, 1974, doi: 10.1080/03610927408827101.

[41] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-1, no. 2, pp. 224–227, Apr. 1979, doi: 10.1109/TPAMI.1979.4766909.

[42] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *J. Comput. Appl. Math.*, vol. 20, pp. 53–65, Nov. 1987, doi: 10.1016/0377-0427(87)90125-7.

YARUI ZHANG received the B.E. degree from the School of Electronic Engineering, Xidian University, Xi'an, China, in 2018, and the M.S. degree in control, signal, and image processing and the Ph.D. degree in signal and image processing from Université Paris-Saclay, Gif-sur-Yvette, France, in 2019 and 2023, respectively.

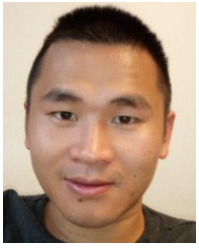
She is currently a Postdoctoral Researcher with Télécom Paris, IP Paris, France. Her research interests include inversion and imaging, electromagnetic inverse scattering problems, wavelets, deep learning, and EMF exposure.



SHANSHAN WANG (Member, IEEE) was born in Nanjing, China, in 1991. She received the Ph.D. degree in networks, computer sciences, and telecommunications from the Laboratory of Signals and Systems (L2S), Paris-Saclay University, France, in 2019.

From 2014 to 2015, she was a Research Engineer with the Toshiba Telecommunication Laboratory, Bristol, U.K. From 2015 to 2018, she was with French National Center for Scientific Research (CNRS), Paris, as an Early Stage Researcher of the European-funded Project H2020 ETN-5Gwireless. From 2019 to 2023, she was a Postdoctoral Researcher with Télécom Paris, IP Paris, France. She is currently an Assistant Professor (Enseignante Checheuse) with ETIS (UMR8051, a joint research laboratory between CY Cergy Paris University, ENSEA, and CNRS). She was the Task 1.1 Leader of the European Horizon Project SEAWave. Her research interests include stochastic geometry and system-level modeling of wireless networks, EMF exposure characterization and AI-based prediction, machine learning, AI, and low-footprint communications. She was a recipient of the 2018 INISCOM Best Paper Award.

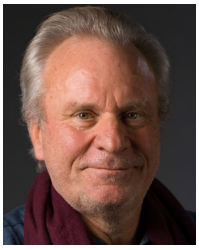
WASSIM BEN CHIKHA (Senior Member, IEEE) received the Engineering degree in computer science from the National Engineering School of Sfax, Tunisia, in 2012, and the Ph.D. degree in electronic and technology of information and communication from the Tunisia Polytechnic School, in 2017. From 2017 to 2020, he was a Postdoctoral Researcher with the University of Udine, Italy, in close collaboration with STMicroelectronics Company. His research included IoT sensors for STM32 open development environment low-power applications using the Contiki operating system. From 2020 to 2021, he worked on radio network management using artificial intelligence (AI) as a Postdoctoral Researcher with the Orange Laboratories. From 2021 to 2023, he was a Research and Development Engineer with Telecom Paris, IP Paris, France. He is currently a Ph.D. Engineer with the Expleo Group, France. His current research interest includes the application of AI in the assessment of electromagnetic field exposure.



JIANG LIU received the B.S. degree from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2014, the master's degree from the National Key Laboratory of Science and Technology on Communications, and the Ph.D. degree from the Laboratory of Signals and Systems, Paris-Saclay University, Paris, France, in 2022. His research interests include wireless communications and communication theory, with a particular focus on signal detection of wireless communication systems.

CE ZHENG received the bachelor's degree from Harbin Institute of Technology, in 2013, the master's degree from Xi'an Jiaotong University, in 2016, and the Ph.D. degree from the University of Lille, in December 2020. From 2021 to 2023, he was a Wireless Researcher with the SONY Research and Development Center, China. He is currently a Postdoctoral Researcher with Télécom Paris. His research interests include stochastic geometry, alpha-stable distribution, copula theory, federated learning, distributed learning, 3GPP networks, and EMF exposure.

THEODOROS SAMARAS (Member, IEEE) received the M.Sc. degree (Hons.) in medical physics from the University of Surrey, Guildford, U.K., and the Ph.D. degree from the Aristotle University of Thessaloniki, Thessaloniki, Greece. He joined the Swiss Federal Institute of Technology in Zurich, where he was involved in the computational dosimetry of electromagnetic fields. He later moved to the Erasmus Medical Centre of Rotterdam, where he worked on the quality assurance of superficial microwave hyperthermia with a Marie-Curie Postdoctoral Fellowship from the European Commission. In December 1999, he returned to the Aristotle University of Thessaloniki, where he is currently a Professor in bioelectromagnetics. His research interests include numerical techniques and computational multiphysics modeling with applications in biomedical technology and the safety of electromagnetic fields.



JOE WIART (Senior Member, IEEE) received the Diploma degree in telecommunication engineering, in 1992, the Ph.D. degree, in 1995, and the H.D.R. degree, in 2015. Since 2015, he has been the Chair C2M "Caractérisation, Modélisation et Maîtrise of the Institut Mines Telecom," Télécom Paris. His works gave rise to more than 150 publications in journals and more than 200 communications. His research interests include experimental, numerical methods, machine learning, statistics applied in electromagnetism, dosimetry, and exposure monitoring. He has been an Emeritus Member of The Society of Electrical Engineers (SEE), since 2008. He is also the Chairperson of the TC106x of the European Committee for Electrotechnical Standardization (CENELEC) in charge of EMF exposure standards. He is also the past Chairperson of the International Union of Radio Science (URSI) Commission K and has been the Chairperson of the French Chapter of URSI.