IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 23 March 2023, accepted 1 April 2023, date of publication 10 April 2023, date of current version 3 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3265723

RESEARCH ARTICLE

Machine Learning Modeling for Radiofrequency Electromagnetic Fields (RF-EMF) Signals From mmWave 5G Signals

ABDULMAJEED AL-JUMAILY[1,2], (Senior Member, IEEE), ADUWATI SALI[1], (Senior Member, IEEE), MOHAMMED RIYADH[1], SANGIN QAHTAN WALI[1], (Student Member, IEEE), LU LI[1], AND ANWAR FAIZD OSMAN[3]

[1]WiPNET Research Centre, Department of Computer and Communication Systems Engineering, Faculty of Engineering, Universiti Putra Malaysia (UPM), Serdang, Selangor 43400, Malaysia
[2]Department of Signal Theory and Communications, Universidad Carlos III de Madrid, 28903 Madrid, Spain
[3]Rohde & Schwarz, PAT SQUARE, Shah Alam, Selangor 40150, Malaysia

Corresponding authors: Abdulmajeed Al-Jumaily (abdulmajeed@tsc.uc3m.es) and Aduwati Sali (aduwati@upm.edu.my)

This work was supported in part by the IGNITE—Interference Modeling for 5G and FSS Coexistence at mmWave with Climate Change Considerations in the Tropical Region under Grant FRGS/1/2021/TK0/UPM/01/1, and in part by the BIDANET: Parametric Big Data Analytics Over Wireless Networks under Grant UPM.RMC.800-3/3/1/GPB/2021/9696300 and Grant 9696300.

ABSTRACT 5G is the next-generation mobile communication technology that is expected to deliver better data rates than Long-Term Evolution (LTE). It offers ultra-low latency and ultra-high dependability, enabling revolutionary services across sectors. However, 5G mmWave base stations may emit harmful radiofrequency electromagnetic fields (RF-EMF), raising questions about health and safety. Our research suggests that the RF-EMF prediction model lacks sufficient papers or publications. Therefore, this study employs IEEE and ICNIRP standards for assessment and exposure limits. The measuring campaign analyses one sector of a 5G base station (5G-BS) operating on 29.5 GHz in Cyberjaya, Malaysia. This study proposes two prediction models. The first model predicts the signal beam RF-EMF, while the second predicts the base station RF-EMF. Each model contains three machine learning techniques to forecast RF-EMF values: Approximate-RBFNN, Exact-RBFNN, and GRNN. The results are analysed and compared with the measured data, determining which algorithm is more accurate by calculating the RMSE of each algorithm. As a result, it can be observed that the Exact-RBFNN algorithm is the best algorithm to predict the RF-EMF because it shows good agreement with the measured value. Moreover, in a 1-minute duration, the difference between the predicted and measured values reached 0.2 less channels. However, in 6 minutes and 30 minutes, it can observe more accurate results since the differences between values reach 0.1 in these situations. Additionally, the ICNIRP standard was used and compared with the results and validation values of the algorithms.

INDEX TERMS 5G, RF-EMF, machine learning ML, approximate-RBFNN algorithm, exact RBFNN algorithm, GRNN algorithm.

I. INTRODUCTION

The fifth generation (5G) mobile network technology represents the next phase in the standardisation of telecommunications. It outperforms the current 4G/LTE networks in terms of speed, capacity, and reliability [1]. 5G technology will address revolutionary ideas such as linked and autonomous automobiles, smart factories, and smart cities, meeting the ever-growing expectations of mobile network consumers. The 5G network will be able to serve significantly more terminals per square kilometre, providing greater data speeds (peak rates up to 20 Gbps), extremely low latency (less than one millisecond), and high dependability (99.999%) [2], [3]. As long as the necessary infrastructure is in place, the EMF

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

level from networks and devices should remain at an average level during the 5G rollout period. However, expanding 5G networks through the use of new 5G radios and the installation of smaller cells (near sites) may increase the level of electromagnetic field (EMF) in the environment. Optimising the highest amount of power output by mobile devices when used in conjunction with 5G mobile communication technologies is necessary, as this has an immediate effect on the system's capacity and coverage. However, the maximum output power of the radio frequency (RF) electromagnetic field (EMF) [4] typically limits phased arrays. Therefore, the EMF exposure property of the mmWave phased array in mobile devices must be thoroughly investigated. To determine the current amount of smartphone EMF exposure, the specific absorption rate (SAR, measured in W/kg) must be calculated [2], [5], [6]. The degree to which electromagnetic fields can penetrate the human body is defined by the Federal Communications Commission (FCC) rules and the International Commission on Non-Ionizing Radiation Protection (ICNIRP) regulations. For frequencies above 6 GHz, the spatial peak value of power density is reviewed in the FCC standards, and an upper limit of 10 W/m is set [2], as shown in Table 1.

TABLE 1. Mandatory Standard Exposure Limits Set by MCMC.

Type of Exposure	Frequency range (Hz-GHz)	Electric field strength(V/m)	Magnetic field Strength H(A/m)	Equivalent plane Wave power Density Seq. (W/m2
Occupational	1-10MHz	610/f	1.6/f	-
	10-400Mhz	61	0.16	10
	400-2000Mhz	3f	0.008f	f/40
	2-300GHz	137	0.36	50
General public	1-10MHz	87/f	0.073/f	-
	10-400Mhz	28	0.073	2
	400-2000Mhz	1.375f	0.0037f	f/200
	2-300GHz	61	0.16	10

It is anticipated that the adoption of the 5G network, which is scheduled to take place in the following years, would result in significant growth in throughput (up to 20 gigabit/sec, according to some estimates). A reduction in data transmission (compared to the first ITU definition) and a reduction in data transmission latency are reduced to one millisecond. The requirement for a frequency band that demands such a high throughput will necessitate implementing new frequency resources, such as several of the tens of gigahertz (millimeter-wave). Furthermore, this will result in a more significant proportion of the field [7]. The radius of curvature of the micro-cell number is very small. Because their antennas will be located much closer to the users than the competition, the micro-cells and macro-cells are separated by around 200 metres [8].

The 5G network is expected to support many thousands of devices at the same time while maintaining their security. Due to the avoidance of interference, radio connections are highly dependable. 3D beamforming technology will be implemented in users' equipment devices. To support this technology, the massive MIMO method in conjunction with an antenna matrix will be needed, which may allow for simultaneous sending and receiving. Current modulation and coding techniques, with hundreds or even thousands of antennas, provide greater flexibility and the ability to shape and influence spatial relationships. Base stations will use various antenna beams to connect to individual devices, allowing for a reduction in the power sent to 5G-BS antennas in several stations [9], [10]. To implement the 5G network, new approaches are being developed, along with a rethink of the strategy to determine the EMF levels. The anticipated adjustments will apply to the measuring and testing equipment. However, creating highly specialised measurement devices is often time-consuming due to the need to validate the measuring capabilities and undergo an essential calibration process before release.

In addition, it is important to keep in mind the most important guidelines when using measurement equipment [11], [12]. New modeling approaches that enable the estimation of continuous EMF distributions while considering the site's geometry and the technical aspects of base stations will become more relevant soon. It has been noticed that this is a problem even at the planning stage for 5G-BS locations. The current approach for estimating the received RF-EMF exposure is based on the assumption that the transmitting antennas emit radiation patterns that are well-known and predictable. Additionally, it assumes that the base stations are broadcasting radio signals at their maximum theoretical strength and that the EMF distribution in and around the base station is quasi-deterministic, which is a highly conservative estimate [13].

On the other hand, both of these assumptions are incorrect in the context of the new fifth-generation technology. Unlike the second, third, and fourth-generation networks that use cutting-edge emitters such as 8-element dual-polarised passive antennas that operate in frequencies ranging from 0.7 GHz to 3.4 GHz, the active antenna systems planned for the fifth generation will be significantly different. They will feature up to 256 active antennas working at higher frequencies than those currently in use, ranging from 3.4 GHz to 6 GHz and 20 GHz to 60 GHz [10], [14].

Aside from that, 3D beamforming will allow steering of the beam in both horizontal and vertical planes, conveying radio signals precisely to the receiving terminal as opposed to the steady transmission over a wide area that is currently the case with 2nd, 3rd, and 4th generation technologies [15]. These characteristics provide a fresh perspective on the EMF measurement approaches in 5th-generation networks. They imply the inability to directly apply the methodologies that have proven successful in second, third, and fourth-generation networks. The 5G networks will be implemented gradually, and the second, third, and fourth generations of BS will not be decommissioned daily. Furthermore, while assessing EMF exposure, it is always necessary to consider the cumulative radiation released by multiple RF signal sources (for example, DVB-T broadcasting transmitters) rather than just the radiation emitted by antennas in mobile networks. Because of this, the new EMF testing methodology must consider the effects of all RF signal sources, including those produced by multiple radio communication systems that coexist in a given location. The use of higher frequencies for arranging 5G networks must also be considered. This unique and complicated radiation pattern of the 5G-BS antenna results from distinct propagation and diffraction conditions caused by such frequencies [10], [16]. The pattern should include an envelope of all radio beams that have been emitted. However, this is further enhanced by 3D beamforming. As a result, when 5G-BS are deployed, it is feasible that the current methods used for assessing exposure to EMFs will not suffice in the future. A new model proposed for 5th generation EMF exposure evaluation will require reliable verification and calibration as a consequence of the close proximity of the connecting 5G-BS, as well as the use of 3D beamforming and massive MIMO techniques in conjunction with the vast spectrum of frequencies that will be utilised [17]. The current EMF assessment requirements necessitate taking measurements assuming that the theoretical maximum transmission power is broadcast.

II. RELATED WORK

In recent years, several studies have been conducted on using machine learning algorithms for predicting and analysing radio frequency (RF) and electromagnetic field (EMF) signals. One such study proposed a method for obtaining the power density value, which is the standard for RF EMF human exposure from mmWave mobile devices, using a deep learning network. As a related work to our research, the study noted that mmWave mobile communication devices using an array antenna require a large number of phase conditions for covering a wide communication range. However, obtaining power density values for each phase condition incurs a lot of time and cost. To address this issue, the study proposed a deep learning network that can input the phase conditions of the mmWave array antenna and simultaneously output the power density results for the phase conditions of the array antenna [18].

This study proposed the EMGAN methodology, which is a deep learning-based method for estimating exposure maps in urban environments. The methodology includes a generator and a discriminator that incorporate city topology as a conditional input to improve the accuracy of exposure map estimation. Unlike traditional approaches that make biased assumptions about radio propagation, EMGAN uses radio environment information from the training process. The results showed that EMGAN outperforms traditional methods in terms of accuracy. Future research aims to expand EMGAN to estimate exposure maps over time [19].

This study proposes a machine learning-based method using neural networks (NN) to assess radiofrequency (RF) exposure generated by WiFi sources in indoor scenarios. The objective is to create an NN model that can handle the complexity and variability of real-life exposure setups, including the effects of up-link transmission from various sources (e.g., laptops, printers, tablets, and smartphones) in addition to down-link transmission from access points (APs). The NN model uses easily obtainable data such as the position and type of WiFi sources and the position and material characteristics of walls to predict RF exposure. The model is evaluated on a new layout and achieves a remarkable field prediction accuracy with a median prediction error of -0.4 to 0.6 dB and a root mean square error of $2.5-5.1$ dB, compared to the target electric field estimated by a deterministic indoor network planner. This proposed approach is effective for different layouts and can be used to assess RF exposure in indoor scenarios. This study evaluates various 5G network topologies in terms of human exposure, mobile communication quality, and sustainability. The focus is on assessing human exposure in 5G networks, specifically those that include Massive Multiple-Input Multiple-Output (MaMIMO) and comparing them to existing 4G deployments in Switzerland. The study uses a novel Exposure Ratio (ER) metric to evaluate human exposure and extrapolates data rates from mobile operators to the year 2030 to evaluate mobile communication quality and sustainability. The study employs a multi-objective optimization algorithm to design 5G network topologies, aiming to maximize user coverage while minimizing both downlink (DL) and uplink (UL) exposure [20], [21].

Moreover, there are several experimental investigations based on a massive data set collected from actual cellular networks have shown that actual 2G, 3G, and 4G base station output powers often approach close to the theoretical maximum. For instance, according to the statistics published in [20], the 90th percentile of the output power for a 2nd generation base station supplied with two or more transceivers was not larger than 65 % during times of high demand and traffic. In the case of 3G installations, the 90th percentile had an even worse performance, with a percentage that did not exceed 43 %. Furthermore, in the area of machine learning modeling, there have also been a number of new approaches to RF-EMF and 5G. Deep learning algorithms have been used to model RF-EMF signals from 5G networks. These include convolutional neural networks (CNNs), recurrent neural networks (RNNs), and deep belief networks (DBNs) [19]. The RF-EMF levels at different frequencies and distances from the source can be accurately predicted by these models. Gaussian process regression (GPR) is a non-parametric machine learning algorithm that has been used to model RF-EMF signals. GPR has been shown to be effective in predicting RF-EMF levels at different locations and frequencies and can be used to generate 3D maps of RF-EMF levels. Ensemble learning algorithms such as random forests and gradient boosting have been used to model RF-EMF signals.

These models combine multiple weak models to generate a stronger, more accurate model. Ensemble learning algorithms have been shown to be effective in predicting RF-EMF levels from 5G networks. Transfer learning is a machine learning technique that involves using a pre-trained model as a starting point for a new model. Transfer learning has been used to model RF-EMF signals from 5G networks and has been shown to improve model accuracy. Bayesian optimization is a technique used to optimise the hyperparameters of machine learning models. Bayesian optimization has been used to optimise the hyperparameters of machine learning models used to model RF-EMF signals from 5G networks.

Improved algorithms for dynamically allocating network resources have been included in the infrastructure's base stations. In addition, this enables the automated shifting of traffic from lower-load to higher-load frequency bands and vice versa and allows the automatic movement of traffic from lower-load to higher-load frequency bands, such that lower transmission powers are used throughout. Balanced use of radio resources at base stations results from this optimisation process. 5G-BS can operate over a wide range of output powers that are significantly lower than the maximum value and that depend on a number of factors that shift over time, such as the amount of traffic that needs to be transferred, signal propagation conditions, and discontinuous transmission associated with time-based medium wireless access networks [10].

There have been several recent publications that have evaluated different aspects of RF exposure and 5G network topologies. In comparison to these studies, our proposed work makes some unique contributions. In a study by [16], a 5G MIMO antenna array was evaluated for its potential human exposure risk. The authors found that the MIMO system had a lower peak spatial specific absorption rate (SAR) than traditional single antenna systems however did not assess exposure under realistic scenarios. In contrast, our proposed work uses a novel ML method based on neural networks to assess RF exposure generated by WiFi sources in indoor scenarios, taking into account the effects of both down-link and up-link transmission by different sources.

Another study by evaluated the electromagnetic radiation of 5G base stations in a suburban area. They found that the measured radiation levels were within safety limits. However, acknowledged that further studies were necessary to assess the long-term effects of exposure. Similarly, our proposed work evaluates the human exposure in 5G networks that include Massive Multiple-Input Multiple-Output (MaMIMO) and compares them with existing 4G deployments. Our study provides a novel Exposure Ratio (ER) metric to assess exposure and evaluates the quality and sustainability of mobile communication [22].

Finally, a study by used a simulation-based approach to assess the human exposure in a 5G network with small cells. They found that human exposure in a 5G network with small cells was lower than that of a 4G network but did not investigate exposure in realistic indoor scenarios. Our proposed work goes beyond small cells and investigates the exposure in indoor scenarios using a multi-objective optimization algorithm to design 5G network topologies that maximize user coverage while minimizing exposure [21].

III. REAL APPLICATIONS AND CHALLENGES IN IMPLEMENTATIONS

The real applications and implementation challenges are important aspects to consider in any research. In the case of the 5G network topologies evaluation mentioned in the previous example, one of the main real applications is to provide guidelines for the development of 5G networks that balance user coverage and exposure levels. This is crucial in ensuring the safety and well-being of individuals while also providing efficient and reliable mobile communication.

However, there are several challenges that need to be addressed in the implementation of these guidelines. For example, the implementation of Massive Multiple-Input Multiple-Output (MaMIMO) technology can be complex and expensive, requiring significant infrastructure and equipment upgrades. Additionally, achieving optimal UL and DL exposure levels may require careful tuning of network parameters, which can be challenging in real-world scenarios. Another example is the proposed machine learning method to assess RF exposure generated by WiFi sources in indoor scenarios [11].

One real application is to provide a tool for assessing RF exposure levels in indoor environments, which is useful for ensuring compliance with safety guidelines and regulations. However, there are challenges in the implementation of this method, such as the need for accurate data input and training data, as well as the need to validate the model under different scenarios and conditions. In general, real applications and implementation challenges are important considerations in any research, as they help to ensure the practicality and usefulness of the proposed solutions. It is important to address these challenges and develop practical solutions that can be easily implemented in real-world scenarios to improve the safety and well-being of individuals and provide efficient and reliable communication [7].

IV. METHODOLOGY

This paper begins by completing a literature review that demonstrates the prior work. In addition to what was mentioned earlier, this section of the paper thoroughly explains each step of the work procedure. We used MATLAB to develop two models, which were used to run a simulation that forecasts the RF-EMF ratio created by communication towers and inputs information into a processor. We can use the information gained from the previous simulation to achieve the best feasible accuracy in our acquired outcomes. The machine learning algorithm was used to help us reach our goals and meet our expectations, which is a task that must be completed.

Moreover, the focus of the study was on mmWave in 5G as our input data. After that, we turned on the model and recorded data. If the ratio data prediction resulted in

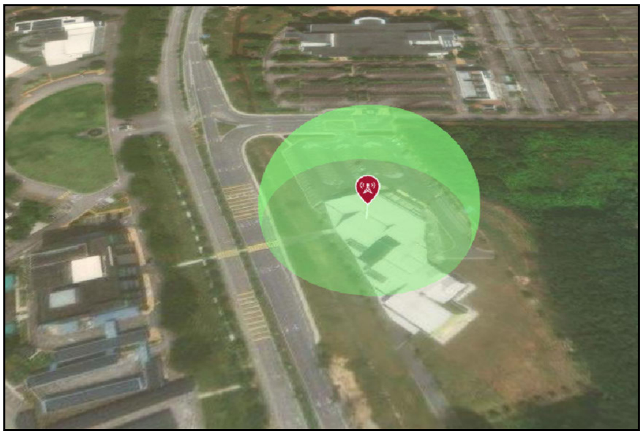

FIGURE 1. An illustrated view of the mmWave base station at Rekascape building, Cyberjaya.

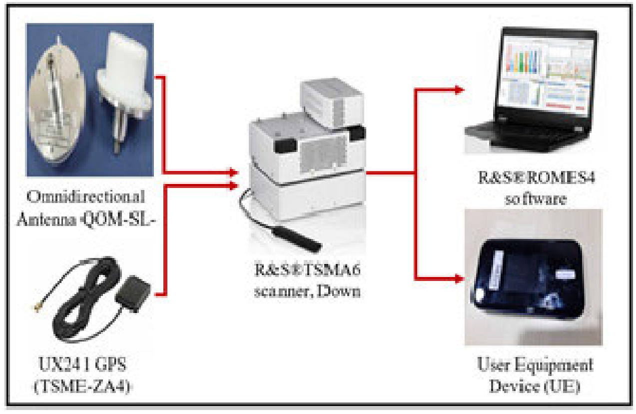

FIGURE 2. Equipment used to measure input data [28].

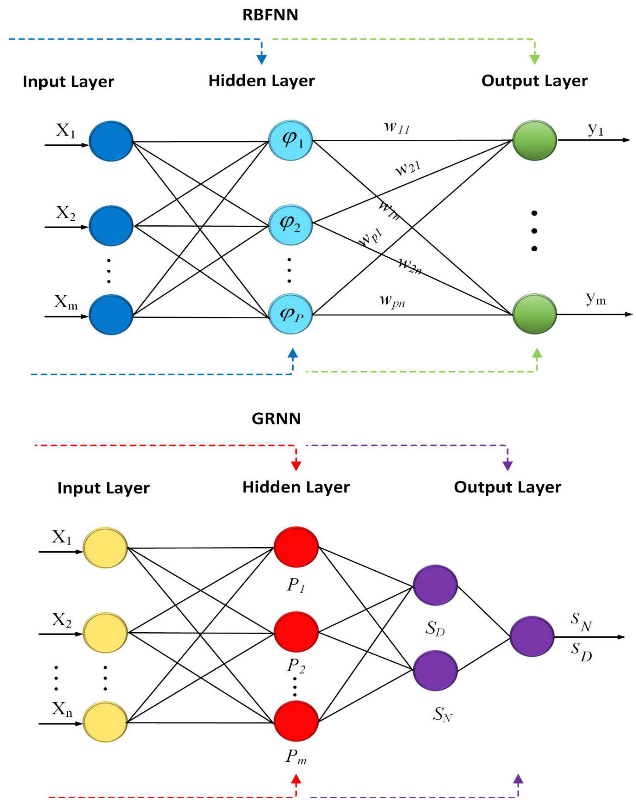

FIGURE 3. Schematic of machine learning algorithms.

high accuracy, we restarted the analysis of the data. If not, we restimulated the model to improve its development. In addition, we created Model one, which is responsible for predicting the value of each SSB beam. This model shows the equations used to analyse the actual input data and use it as input for the prediction model, as well as the algorithms used to predict the result. Besides that, we designed another model used to predict the average value of RF-EMF at the base station. This model shows the equations and the algorithm used to get the best prediction of the RF-EMF value and compare it with measured data.

A. MEASUREMENT EQUIPMENT FOR INPUT DATA

The R&S® TSMA6 scanner, along with the R&S® ROMES4 software, is designed to assist with 5G-New Radio (5G-NR) measurements below 6 GHz and in mmWave frequency ranges. The R&S® TSME44DC downconverter is used to analyze signals in the 24 GHz to 44 GHz range during the measurement process. Primary parameters, such as power levels (e.g., RSRP) and signal-to-noise ratios (e.g., SINR) of the variant signals in the 5G New Radio SSB, were obtained from the scanner. These parameters were used to determine the RF conditions at a specific location, which formed the basis for network access via 5G-NR equipment. In this project, a 26-40 GHz vertically polarized omnidirectional antenna fitted with a K type connector and Radome was used as the receiver antenna side. The antenna was connected directly to the scanner using a cable and automatically received the signal. The UX241 GPS (TSME-ZA4) was another equipment used to precisely determine the measurement's location and distance. Finally, user equipment was used by one of our teams to obtain input data.

B. ALGORITHMS

This article employed three machine learning algorithms, as shown in Fig. 3. The first algorithm used was the Approximate Radial Basis Function Network (Approximate-RBFNN), followed by the Exact Radial Basis Function Network (Exact-RBFNN). A feedforward machine learning with three layers was also utilized. The first layer of the network corresponds to the inputs, the second layer consists of several RBFNN non-linear activation units, and the last layer relates to the network's output [23].

Radial Basis Function Networks (RBFN) and Generalized Regression Neural Networks (GRNN) are both types of neural networks commonly used in machine learning. Both algorithms can be used for regression problems, where the goal is to predict a continuous output value given a set of input features. There are three layers in a RBFN: the input layer, the hidden layer, and the output layer. The hidden layer uses a set of radial basis functions to change the input data into a higher-dimensional feature space. The result of each radial basis function is based on the Euclidean distance between the input and a centre point. The output layer then makes the final output value by adding up the outputs of the hidden layers in a linear way. In order to train a RBFN, you must choose the centre points for the radial basis functions and figure out the weights for the output layer. The weights can be determined using linear regression or gradient descent.

A GRNN also consists of three layers: an input layer, a hidden layer, and an output layer. The hidden layer uses a Gaussian activation function to transform the input data into a higher-dimensional feature space. The output layer uses a kernel density estimation to generate the final output value. Training a GRNN involves selecting the standard deviation of the Gaussian activation function and the smoothing parameter for the kernel density estimation. The standard deviation can be selected using cross-validation or other methods. The smoothing parameter can be determined using linear regression or gradient descent. Both RBFN and GRNN can be used for various applications, such as time series prediction, image recognition, and speech recognition. However, they have different strengths and weaknesses. RBFN is better suited for problems with sparse data or high-dimensional input spaces, while GRNN is better suited for problems with noisy data or non-linear input-output relationships. The algorithmic notation for both RBFN and GRNN:

RBFN:

Input: - *Training set:* $[(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)]$
- *Number of radial basis functions:* k

Output: -*Set of center points:* $[c_1, c_2, \ldots, c_k]$
- *Set of weights:* $[w_1, w_2, \ldots, w_k]$

Algorithm: -*Compute the radial basis function matrix R using the formula R(i,j) = exp(-gamma∗D(i,j)^2), where gamma is a tuning parameter.*
- *Solve the linear system* $Rw = y$ *to obtain the set of weights* $[w_1, w_2, \ldots, w_k]$.
- *Output the RBFN model with the center points* $[c_1, c_2, \ldots, c_k]$ *and the weights* $[w_1, w_2, \ldots, w_k]$.

GRNN:

Input: - *Training set: [(x1, y1), (x2, y2), ..., (xn, yn)]*
- *Standard deviation of the Gaussian activation function: sigma*

Output: - *Kernel density estimation function: f(x)*

Algorithm: - *Compute the distance matrix D between each input xi and all other inputs* x_j.
- *Compute the Gaussian activation function matrix G using the formula G(i,j) = exp(-D(i,j)^2/(2∗sigma^2)).*
- *Compute the kernel density estimation function f(x) using the formula f(x) = sum(*w_i*∗yi)/ sum(wi), where wi = G (i, x) for all training inputs xi and yi.*
- *Output the GRNN model with the kernel density estimation function f(x).*

Conventionally, Gaussian functions are used as activation functions in RBFNs. Additionally, the main advantage of using this algorithm is its ability to process massive amounts of information. The third algorithm discussed in this paper is the General Regression Neural Network (GRNN), which is a highly parallel, one-pass learning process. Even with limited data in multidimensional measurement space, the technique offers smooth transitions from one observed value to the next. The algorithmic form can be used for any regression problem for which the assumption of linearity is not warranted, and the primary benefit of using this approach is its high estimation precision.

In this study, three machine learning techniques were used to predict RF-EMF values for 5G over the mmWave base station with a MIMO antenna. However, other machine learning and deep learning models could also be used for this task. Here are some potential models for comparison:

Support vector machine (SVM): SVM is a widely used classification algorithm that can be used for regression tasks as well. SVM works by finding the optimal hyperplane that separates the input data into different classes. In the context of RF-EMF prediction, SVM could be trained to predict the RF-EMF values based on input features such as antenna location and building density.

Random forest (RF): RF is an ensemble learning method that combines multiple decision trees to make predictions. Each tree in the forest is trained on a random subset of the input data, and the final prediction is made by aggregating the predictions of all the trees. RF has been used successfully in various applications, including image classification and stock price prediction.

Recurrent neural network (RNN): RNN is a type of neural network that is well-suited for modeling time series data. RNNs have a feedback mechanism that allows them to remember previous inputs and use that information to make predictions. In the context of RF-EMF prediction, RNN could be used to predict RF-EMF values based on historical data.

Deep neural network (DNN): DNN is a neural network with multiple layers between the input and output layers. DNNs can learn complex patterns in the input data and are widely used in various applications such as speech recognition and image classification. In the context of RF-EMF prediction, DNN could be trained to predict the RF-EMF values based on input features such as antenna location and building density.

Finally, there are several machine learning and deep learning models that could be used for RF-EMF prediction in the context of 5G deployment. Future work could investigate the performance of these models and compare them with the models used in this study.

C. ELECTRICAL FIELD STRENGTH CALCULATION

It is important to ensure that the data is properly sorted and organised before further analysis. One way to accomplish this is to export the data from the scanner into an Excel sheet and sort it according to specific parameters. For example, the data can be sorted based on the physical layer cell ID of the base station, with the top SSB being listed first and then working downwards. Once the data is sorted, it can be reported as a fixed column for each physical layer's cell ID. This column should include the SS-RERP power values for each SSB. It is important to note that SS-RERP power is a measure of the received power from a base station, and it can be converted to electrical field strength [V/m] using Equation (1) [22].

$$E_{field} = E_{SSB} = \frac{1}{\sqrt{20}} 10^{\frac{P+AF}{20}} \tag{1}$$

Next, the ESSB that has been determined with the help of some other parameters should be inputted into another equation to obtain the projected electrical field strength for each PCI, as shown in Equation (2) [24].

$$E_{asmt} = E_{SSB} * \sqrt{F_{extBeam} * F_{BW} * F_{PR} * F_{TCD}} \tag{2}$$

For all of the SSBs in a single Physical Layer Cell ID, they should be squared and placed under the square root to get the total SSB for that PCI. It is important to repeat this procedure for every PCI in the base station, as shown in Equation (3).

$$E_{\text{channel n}} = \sqrt{{E_{asmt}}_1^2 + {E_{asmt}}_2^2 + {E_{asmt}}_n^2} \tag{3}$$

The overall exposure for all sectors (PCI) of the base station can be computed using the following equation, as shown in Equation (4).

$$E_{BS} = E_{max} = \sqrt{E_{ch_1}^2 + E_{ch_2}^2 + E_{ch_n}^2} \tag{4}$$

V. RESULTS AND DISCUSSION

In this study, machine-learning algorithms were trained using a database from a previous study, which measured the 29 GHz data at a 5G mmWave base station located in RekaScape Cyberjaya, Selangor. The base station had one sector (PCI) consisting of four channels, and code-selective measurements were used due to changes in radiated power with data flow. The Synchronization Signal/Physical Broadcast Channel (SS/PBCH) in a 5G-BS includes the Synchronisation Signals (SS), the PBCH, and the PBCH-DMRS, which is the only "always-on" NR signal. Each channel periodically transmits SSBs, and in this investigation, the R&S®TSMA6 scanner automatically recognized the 5G-NR carriers, decoded, and measured the SSBs and PCI.

Furthermore, the 5G-BS at RekaScape has 16 SSB beams per channel and a total of 64 beams. However, this section calculates the maximum exposure, which represents the worst-case scenario, as well as the average exposure, based on data collected using a Video Streaming scanner. In addition, User Equipment (UE) was used in five subsequent tests. The UE was connected to the first 5G-BS channel (2098117), and each test was conducted for 1 minute, 6 minutes, and 30 minutes. The data collected from these tests were then used as inputs for machine learning methods to predict the RF-EMF of each SSB channel using three different algorithms, and their accuracy was compared. Finally, the average RF-EMF of the base station was predicted.

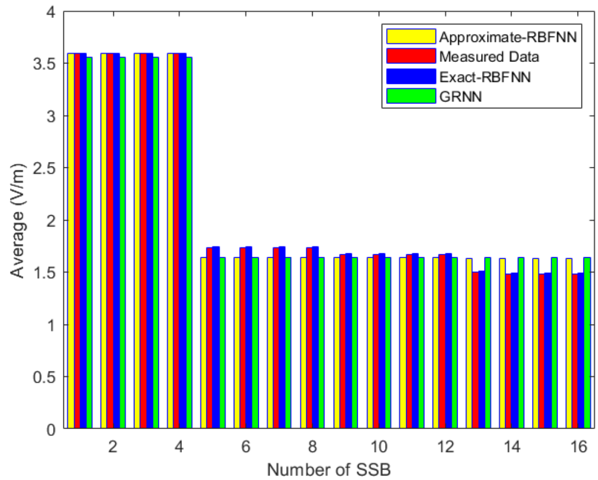

FIGURE 4. RF-EMF for each SSB channel using all three algorithms for 1-minute.

In Fig. 4, the first four channels are almost identical with only slight differences, while the next four demonstrate slightly higher predicted values compared to the measured data. The observed data for the last four SSB channels are slightly higher than the algorithm's predictions. On the other hand, the Exact RBFNN method shows the differences between the measured and predicted data, as the first eight channels are practically identical with just small variances. The following eight channels show that the predicted data are slightly lower than the measured data. The first eight channels are almost identical with only slight differences, while the next eight show similar predicted values. The values of the first four channels are higher than those of the second four channels.

The first pattern is stronger than the second. The measured data shows that the projected values are higher than the measured data. Additionally, the GRNN algorithm illustrates the differences between the measured and predicted data, as the first 12 channels show that the predicted data is slightly higher than the measured data. The ratios of various algorithms show that the predicted data is slightly lower than the measured data for the final four channels. The comparison between the expected and measured data is shown in Fig. 5.

While using the Exact-RBFNN, the outputs of the first SSB are identical in every way. In the Approximate-RBFNN, it is evident that the results of the first two SSBs are comparable, with the difference between the predicted and actual results being about 0.00152. However, the difference in results for the remaining SSBs, especially the final four, is significantly higher than this value. The difference between predicted and actual performance in SSBs is only a small 0.02611 value, but this value is exceeded by the discrepancy with results from other channels. Specifically, the difference between the results of the last SSB was 0.19199, which is also small but more significant than the difference between the results of the other SSBs. Finally, the GRNN result illustrates modest

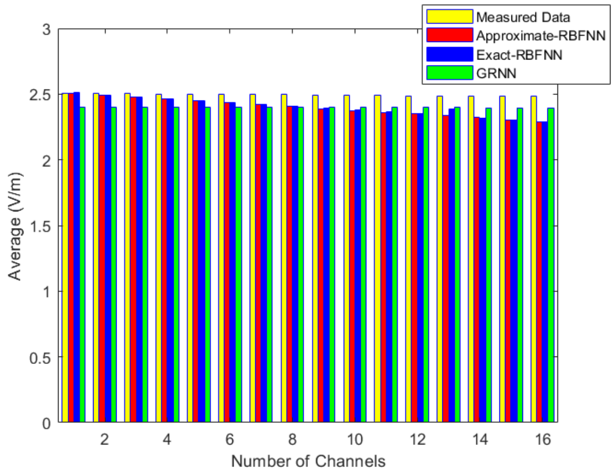

FIGURE 5. RF-EMF for each SSB channel using all three algorithms for 6-minute.

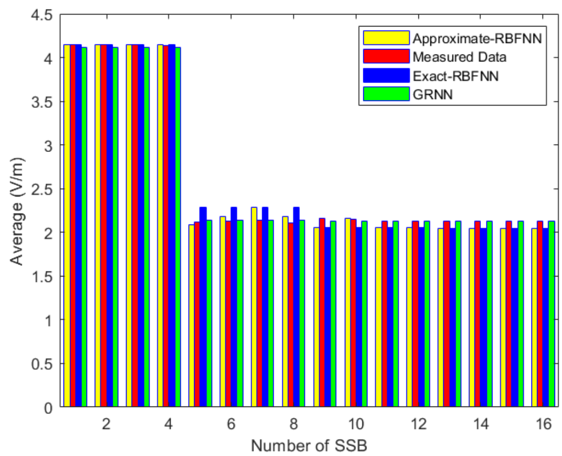

FIGURE 6. RF-EMF for each SSB channel using all three algorithms for 30-minute.

variations between the predicted and measured results for all SSBs. For example, the difference between the first and last SSB is 0.10822 and 0.08672, respectively. Additionally, there are two cases where the predicted value exceeds the measured value, as shown in Fig. 6.

In the Approximate-RBFNN algorithm, the measured and expected results of the first four SSBs are only 0.00116 apart. The results of four channels are comparable, and the difference between SSB 4 and 8 is slightly more significant, indicating that the expected result is higher than the observed result. In the last eight SSBs, the measured values were higher than predicted, but only by tiny digits. On the other hand, in the Exact-RBFNN algorithm, it is evident that the predicted outcomes and the actual results are similar, with just a few minor differences. Additionally, the values are the same in the first four SSBs, and the differences between subsequent SSBs are minimal, as shown in Fig. 7.

On the other hand, the GRNN algorithm results in Fig. 7 show that the predicted values for the first eight channels

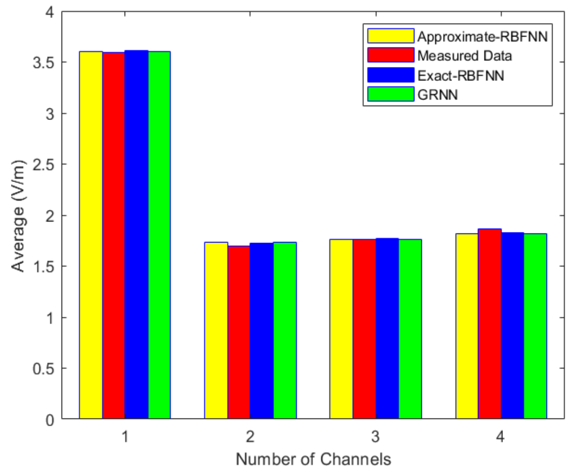

FIGURE 7. RF-EMF average using all three algorithms for 1- minute.

are higher than the measured values. The difference between the predicted and measured values for the first four channels is 0.03492, while the difference for the second four channels is 0.1442. Furthermore, in the first eight channels, the predicted values are lower than the measured values.

This scenario involves performing a numerical analysis of the data to convert the input power to RF-EMF values using equations (1) and (2). The resulting values are then used in equation (3) to calculate the RF-EMF average. Next, we can apply the model to predict the results using three algorithms and compare their accuracy. It is not surprising that the results for each channel are quite similar. In the Approximate-RBFNN algorithm, the difference between the predicted and actual results for the first channel is only 0.0036. The outcomes for the initial channel are comparable. However, the difference between channels two and three is significantly larger, indicating that the expected result is greater than the observed result. Additionally, the measured values for the final channel were higher than the predicted value, although only by a few decimal places. When the predicted value exceeds the actual value, it is considered an error.

However, the difference in findings between channels two, three, and four is somewhat smaller than this amount, indicating that the projected result is less than the actual result. Moreover, in the Exact-RBFNN in Fig. 8, it is evident that the results of all channels are remarkably similar, with a gap between predicted and measured results of only 0.0001. On the other hand, in the GRNN algorithm result, the discrepancy between the predicted and observed results for the first channel is 0.95745, as demonstrated in Fig. 8. The Approximate-RBFNN algorithm results demonstrate that the findings of all channels are comparable, with the difference between expected and measured results ranging from 0.0061 in the first channel to 0.0049 in the last channel. Moreover, the anticipated values in every channel are less than those measured. The exact RBFNN value reveals that

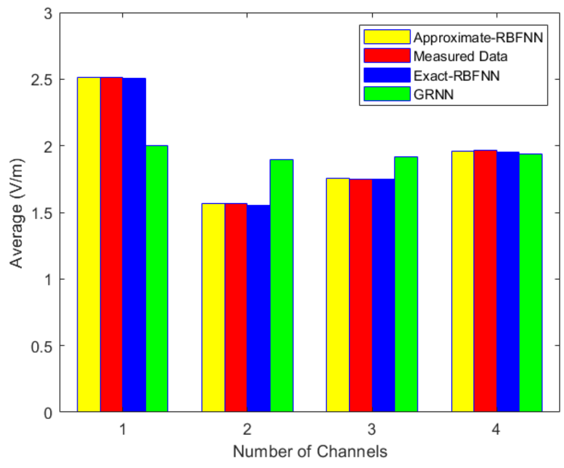

FIGURE 8. RF-EMF average using all three algorithms for 6- minute.

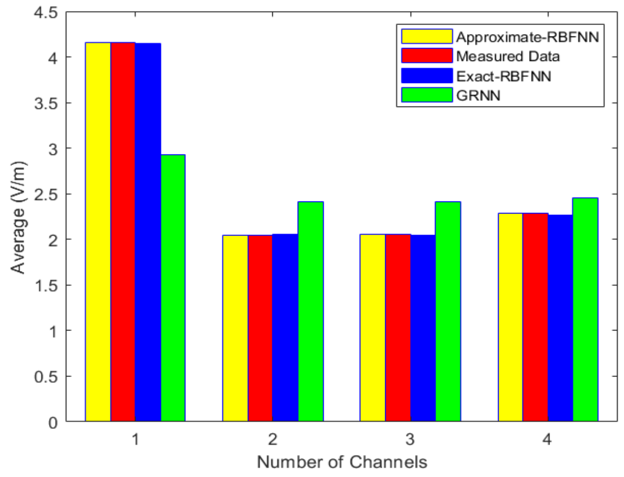

FIGURE 9. RF-EMF average using all three algorithms for 30- minute.

the results of all channels are astonishingly identical, with a difference of only 0.0003 percentage points between the anticipated and actual values. The outcomes of each channel are quite comparable. The GRNN algorithm's first channel demonstrates a discrepancy between the collected and predicted data, with this discrepancy being more significant than that of the other channels, reaching 0.51.

The analysis indicates that there are differences between the collected and predicted data across the various channels. In the Approximate-RBFNN algorithm, the discrepancy between the collected and anticipated data is small, with the first channel having the smallest difference of 0.0001. However, channels two and three have a larger discrepancy, reaching 0.0024. In the Exact-RBFNN algorithm, the discrepancy is incredibly minor, reaching 0.0001 in the worst-case scenario. In contrast, in the GRNN algorithm, the first channel has a more significant discrepancy of 1.22, while the discrepancy in the other channels is less pronounced, with the fourth channel having the largest difference of 0.1695. Additionally, the anticipated value for the first channel is greater than the measured value, as shown in Fig. 9.

The remaining channels, in contrast, have a smaller observed value than the expected value. In summary, the GRNN algorithm has the worst performance compared to the first two algorithms in terms of prediction accuracy. After observing the data for 30 minutes, it can be concluded that the Exact-RBFNN algorithm is the best algorithm for predicting RF-EMF because it has demonstrated values that are nearly identical to the actual values. On the other hand, the GRNN algorithm performs the worst among the three methods for predicting RF-EMF, primarily due to the slight difference between the predicted and measured values. Overall, the Exact-RBFNN algorithm is the most effective method for predicting the average RF-EMF for 6 minutes, as it has exhibited values that are almost identical to the actual values.

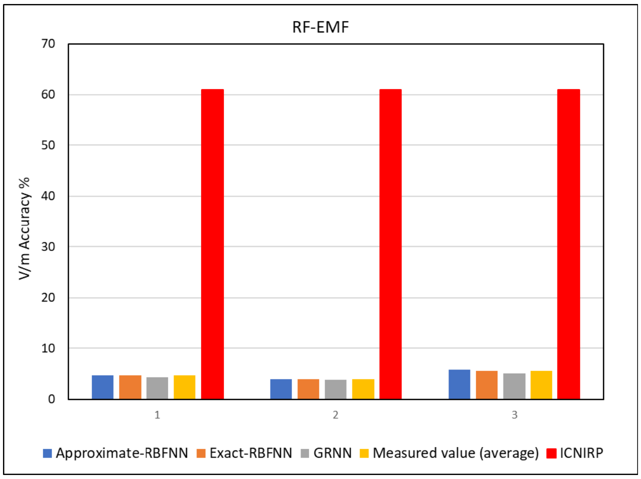

FIGURE 10. Validation result with ICNIRP organisation [21].

TABLE 2. RF-EMF Average Values for Base Station.

Algorithms	RF-EMF exposure durations		
	1-minute	6-minute	30-minute
Approximate-RBFNN	4.728 V/m	3.9061V/m	5.766 V/m
Exact-RBFNN	4.731 V/m	3.9527 V/m	5.5633 V/m
GRNN	4.366 V/m	3.877 V/m	5.127 V/m
Measured value (average)	4.733V/m	3.9627V/m	5.5533V/m
ICNIRP [21]	61V/m	61V/m	61V/m

VI. VALIDATION RESULT WITH ICNIRP ORGANISATION

Fig. 10 and Table 2 illustrate the validation of production results compared to the ICNIRP standard. The study found that human exposure to 5G via mmWave is safe in Malaysia, with the highest exposure recorded at 5.71V/m and an average exposure of 2.02V/m. These levels are well below the ICNIRP standard of 61V/m [21], indicating that they are significantly below the allowed limit.

VII. CONCLUSION

Several experiments have investigated the RF-EMF level in 5G using C-band frequencies and concluded that human deployment of 5G in the C-band is safe. However, no research on developing mmWave exposure prediction models has been found. This study aimed to determine the maximum and average exposure in a tropical 5G-BS using mmWave frequency and to model 5G over the mmWave site for the first time. The study used a unique, comprehensive technique to evaluate exposure to RF-EMF for 5G-BS and estimated and examined the maximum and average RF-EMF exposure for 5G over the mmWave base station with a MIMO antenna following international laws. The study used three machine learning techniques to estimate the RF-EMF value and developed two models to forecast RF-EMF values, which were run three times each. The results show that the Exact-RBFNN algorithm is optimal for predicting RF-EMF due to its high correlation with the measured value. The study also examined the impact of measurement time on the accuracy of the predicted values, indicating that the duration of measurement affects the prediction accuracy.

Moreover, the study successfully developed a comprehensive model approach and evaluated the exposure to RF-EMF for 5G over the mmWave base station with a MIMO antenna, which has not been previously researched. The study demonstrates that the level of RF-EMF exposure in 5G over the mmWave base station is well below the international standards set by the ICNIRP guidelines, indicating that 5G deployment in the mmWave band is safe.

The study also explored the use of other machine-learning techniques for RF-EMF prediction, and future work could be extended to investigate the potential variation in RF-EMF exposure levels in other regions and environments. Additionally, the study can be extended to investigate the effects of RF-EMF exposure on human health, an area of ongoing research and public concern.

ACKNOWLEDGMENT

This project is a collaboration with Rohde & Schwarz Malaysia for consultations on the measurement analysis.

REFERENCES

[1] M. Attaran, "The impact of 5G on the evolution of intelligent automation and industry digitization," *J. Ambient Intell. Humanized Comput.*, vol. 14, pp. 5977–5993, Feb. 2021.

[2] R. Pawlak, P. Krawiec, and J. Zurek, "On measuring electromagnetic fields in 5G technology," *IEEE Access*, vol. 7, pp. 29826–29835, 2019.

[3] S. H. Alrubaee and B. B. Burhan, "Study of 5G networks as a promising future wireless mobile," in *Proc. Int. Conf. Commun. Inf. Technol. (ICICT)*, Jun. 2021, pp. 127–132.

[4] M. Mevissen, "Effects of radiofrequency electromagnetic fields (RF EMF) on cancer in laboratory animal studies," *Environ. Int.*, vol. 161, Mar. 2022, Art. no. 107106.

[5] S. Aerts, L. Verloock, M. Van Den Bossche, D. Colombi, L. Martens, C. Törnevik, and W. Joseph, "In-situ measurement methodology for the assessment of 5G NR massive MIMO base station exposure at sub-6 GHz frequencies," *IEEE Access*, vol. 7, pp. 184658–184667, 2019.

[6] L.-M. Schilling, C. Bornkessel, and M. A. Hein, "Impact of small-cell deployment on combined uplink and downlink RF exposure compared to the status quo in mobile networks," in *Proc. 51st Eur. Microw. Conf. (EuMC)*, Apr. 2022, pp. 618–621.

[7] N. K. Mallat, M. Ishtiaq, A. Ur Rehman, and A. Iqbal, "Millimeter-wave in the face of 5G communication potential applications," *IETE J. Res.*, vol. 68, no. 4, pp. 2522–2530, Jul. 2022.

[8] L. Banda, M. Mzyece, and F. Mekuria, "5G business models for mobile network operators—A survey," *IEEE Access*, vol. 10, pp. 94851–94886, 2022.

[9] S. Patrício, L. M. Correia, and M. Gomes, "Influence of active antennas on EMF restrictions in 5G base stations deployment," in *Proc. IEEE Int. Symp. Antennas Propag. USNC-URSI Radio Sci. Meeting (AP-S/URSI)*, Jul. 2022, pp. 1280–1281.

[10] A. Al-Jumaily, A. Sali, V. P. G. Jiménez, F. P. Fontán, M. J. Singh, A. Ismail, Q. Al-Maatouk, A. M. Al-Saegh, and D. Al-Jumeily, "Evaluation of 5G coexistence and interference signals in the C-band satellite earth station," *IEEE Trans. Veh. Technol.*, vol. 71, no. 6, pp. 6189–6200, Jun. 2022.

[11] L. Chiaraviglio, C. Lodovisi, D. Franci, S. Pavoncello, T. Aureli, N. Blefari-Melazzi, and M.-S. Alouini, "Massive measurements of 5G exposure in a town: Methodology and results," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 2029–2048, 2021.

[12] I. López, N. Félix, M. Rivera, A. Alonso, and C. Maestú, "What is the radiation before 5G? A correlation study between measurements in situ and in real time and epidemiological indicators in Vallecas, Madrid," *Environ. Res.*, vol. 194, Mar. 2021, Art. no. 110734.

[13] G. Oliveri, G. Gottardi, N. Anselmi, and A. Massa, "Capacity-driven low-interference fast beam synthesis for next generation base stations," *IEEE Trans. Antennas Propag.*, vol. 70, no. 6, pp. 4472–4484, Jun. 2022.

[14] F. Samat, J. Mandeep, A. Al-Jumaily, and A. Sali, "Mobile set-top box over the Internet of Things (IoT) to secure rain-induced signal outages issue in satellite broadcasting services," Tech. Rep., 2022.

[15] H. Tsukada, K. Kumakura, S. Tang, and M. Kim, "Millimeter-wave channel model parameters for various office environments," *IEEE Access*, vol. 10, pp. 60387–60396, 2022.

[16] M. A. Salem, H. S. Lim, M. Y. Chua, S. F. Chien, C. C. Zarakovitis, C. Y. Ng, and N. Z. A. Rahman, "Investigation of EMF exposure level for uplink and downlink of 5G network using ray tracing approach," *Int. J. Technol.*, vol. 13, no. 6, p. 1298, Nov. 2022.

[17] A. Goudarzi, M. M. Honari, and R. Mirzavand, "A millimeter-wave resonant cavity antenna with multibeam and high-gain capabilities for 5G applications," *IEEE Trans. Antennas Propag.*, vol. 70, no. 10, pp. 9149–9159, Oct. 2022.

[18] C. Chountala, J.-M. Chareau, and P. Chawdhry, "Radio frequency electromagnetic field measurements in a commercial 5G network," in *Proc. IEEE 4th 5G World Forum (5GWF)*, Oct. 2021, pp. 275–280.

[19] B. A. Taha, Y. A. Mashhadany, A. H. J. Al-Jumaily, M. S. D. B. Zan, and N. Arsad, "SARS-CoV-2 morphometry analysis and prediction of real virus levels based on full recurrent neural network using TEM images," *Viruses*, vol. 14, no. 11, p. 2386, Oct. 2022.

[20] S. Wang, T. Mazloum, and J. Wiart, "Prediction of RF-EMF exposure by outdoor drive test measurements," *Telecom*, vol. 3, no. 3, pp. 396–406, Jun. 2022.

[21] S. Q. Wali, A. Sali, J. K. Allami, and A. F. Osman, "RF-EMF exposure measurement for 5G over mm-wave base station with MIMO antenna," *IEEE Access*, vol. 10, pp. 9048–9058, 2022.

[22] A. Sali, S. Q. Wali, and A. F. Osman, "Evaluation of RF-EMF exposure for sub-6 GHz 5G NR massive MIMO base station," in *Proc. IEEE 6th Int. Symp. Telecommun. Technol. (ISTT)*, Nov. 2022, pp. 16–21.

[23] J. Kaur, M. A. Khan, M. Iftikhar, M. Imran, and Q. E. U. Haq, "Machine learning techniques for 5G and beyond," *IEEE Access*, vol. 9, pp. 23472–23488, 2021.

[24] T. Kopacz, S. Schießl, A.-M. Schiffarth, and D. Heberling, "Effective SSB beam radiation pattern for RF-EMF maximum exposure assessment to 5G base stations using massive MIMO antennas," in *Proc. 15th Eur. Conf. Antennas Propag. (EuCAP)*, Mar. 2021, pp. 1–5.

ABDULMAJEED AL-JUMAILY (Senior Member, IEEE) was born in Fallujah, Iraq. He received the B.Sc. degree (Hons.) in electronics and communication engineering, in 2003, and the M.Sc. degree in wireless communications engineering and the Ph.D. degree in wireless communication and networks engineering from University Putra Malaysia (UPM), in 2014 and 2022, respectively. During the Ph.D. and master's degrees, he was a Research Assistant with the WiPNET Research Centre, Department of Computer and Communication Systems Engineering, Faculty of Engineering, UPM, from the master's days, in 2011, until the completion of the Ph.D. degree, in 2023. His work is in mobile and satellite communications. For the M.Sc. degree, he investigated the use of ku-band satellite transmission to serve communications on high-speed trains in Malaysia, whereas for the Ph.D. degree, he analyzed measurement results and proposed mitigation techniques for 5G and fixed satellite service (FSS) co-existence at C-band based on machine learning. His research interests include wireless communications, including mobile communications systems, satellite communications, radio propagation, link budget, WiFi, bluetooth, signal processing, image processing, augmented reality, digital communication, AI, and machine learning. He has research experience and hands-on experience conducting experimental applications and measurements in the mobile communications laboratory, which keeps him up to date with the latest developments in the field. He is also responsible for supervising research projects, including the Ph.D., master's theses, and undergraduate projects, and developing successful educational programmes and courses that inspire and challenge students. Additionally, he regularly publishes research articles in peer-reviewed journals to share his findings with the scientific community. He is currently a Ph.D. Research Fellow on an Erasmus + KA107 exchange as a Ph.D. Researcher with the University Charles III of Madrid, Spain, from 2022 to 2023. Since 2017, he has been a member of the International Association of Engineers (IAENG). He is also a member of Malaysia's Board of Technologists (MBOT) and the U.K.'s BCS, the Chartered Institute for IT.

ADUWATI SALI (Senior Member, IEEE) received the B.Eng. degree in electrical electronics engineering (communications) from the University of Edinburgh, U.K., in 1999, the M.Sc. degree in communications and network engineering from Universiti Putra Malaysia (UPM), Malaysia, in April 2002, and the Ph.D. degree in mobile and satellite communications from the University of Surrey, U.K., in July 2009. She was the Assistant Manager of Telekom Malaysia Bhd, from 1999 to 2000. She was the Deputy Director of the UPM Research Management Centre (RMC), where she was responsible for research planning and knowledge management, from 2016 to 2019. She has been a Professor with the Department of Computer and Communication Systems, Faculty of Engineering, UPM, since February 2019. She was a recipient of the 2018 Top Research Scientists Malaysia (TRSM) Award from the Academy of Sciences Malaysia (ASM). She is involved with IEEE as the Chair to ComSoc/VTS Malaysia, in 2017 and 2018, and Young Professionals (YP), in 2015; Young Scientists Network-Academy of Sciences Malaysia (YSN-ASM) as the Chair, in 2018; and Science Policy as the Co-Chair, in 2017. She is also a Chartered Engineer (C.Eng.) registered under the U.K. Engineering Council and a Professional Engineer (P.Eng.) under the Board of Engineers Malaysia (BEM).

MOHAMMED RIYADH was born in Baghdad, Iraq, in 1988. He received the bachelor's degree in computer and communication engineering from the AL-Mansor University College, Baghdad, in 2010, and the M.S. degree from the Department of Computer and Communication Systems Engineering, Faculty of Engineering, University Putra Malaysia (UPM), Seri Kembangan, Malaysia, in 2022. In 2011, he was with Holly Quran Radio Station, Baghdad, as a Communication Engineer. In 2017, he was with the Computer Department, Alwaquf Al Sunny, Iraq, in 2022.

SANGIN QAHTAN WALI (Student Member, IEEE) was born in Kurdistan, Iraq, in 1997. She received the bachelor's degree in electrical and computer engineering from Duhok University, Kurdistan, Iraq, in 2018, and the M.S. degree from the Department of Computer and Communication Systems Engineering, Faculty of Engineering, Universiti Putra Malaysia (UPM), Seri Kembangan, Malaysia, in 2021.

LU LI was born in China. He is currently serving as an executive committee of IEEE Malaysia ComSoc/VTS Joint Chapter and working toward a Ph.D. degree in wireless communications and networks engineering at the Wireless and Photonic Network Research Centre of Excellence (WiPNET) with the Department of Computer and Communication Systems Engineering, Faculty of Engineering, Universiti Putra Malaysia (UPM), Selangor, Malaysia. His research interests include wireless and mobile communication, optical sensors, biomedical sensors, channel estimation and propagation technology of mmWave, IoT systems, industrial application, peatland forests, fire prediction and machine learning. Additionally, he has been invited to serve as a reviewer for many international journals indexed by Scopus. He is also a collaborator for projects under international funding bodies, namely the NICT Japan-ASEAN IVO, Asia Pacific Telecommunity (APT) and SEARCA University Consortium (UC) Seed Fund for Collaborative Research Grant.

ANWAR FAIZD OSMAN was born in Georgetown, Penang. He received the B.Sc. degree in electrical engineering from Purdue University, USA, in 2003, and the M.Sc. degree in electrical and electronic engineering from Universiti Sains Malaysia (USM), in 2015. His master's thesis is on wideband low noise amplifier design. His current role is a Regional System and Application Engineer with Rohde & Schwarz. He has published multiple technical papers on LNA, RF switches, and RF filter designs. His research interests include wireless testing for mobile operators and interference hunting. He has been a Committee Member of IEEE ED/MTT/SSC Penang Chapter, since 2015, currently serves as the Chapter Auditor. He is also the Former Deputy Chair of IEEE Microwave, Electron Devices, and Solid-State Symposium IMESS 2018. He is a part of the National 5G Taskforce Member and the Chairperson of the 5G Ecosystem and Timeline Sub-Working Group, under the Spectrum Working Group of the Task Force. He is also leading the study on the 5G and FSS co-existence in the 5G taskforce and one of the authors of the taskforce 5G C-band report and 5G mmWave report.