Estimating Mean of Maximum Fields Inside Reverberation Chambers Using Deep Neural Networks

Neda Nourshamsi, *Member, IEEE*, Pedro Uría Rodríguez, Amir H. Jafari, and Charles F. Bunting, *Fellow, IEEE*

***Abstract*—Recently, there has been great interest in estimating the mean of the maximum field strength in a nested reverberation chamber in such conditions that the field coupled inside the equipment under test (EUT) deviates from the parent distribution, thus generating a non-Rayleigh distribution. For this purpose, a new regression model with the deep feedforward neural network is proposed to predict the mean of the maximum field inside a nested reverberation chamber configuration. In our proposed method, a frequency range that comprises the EUT in the overmoded regime is treated as an input of the network, and the mean of the maximum field is treated as the output of the network. Several networks with different numbers of hidden layers are trained, while adaptive learning rates and early stopping techniques are used to improve the network training process, subsequently reducing the uncertainties. After training, the networks are verified using a test set that is not implicitly employed during the training process. The testing and training mean-squared errors (625e-5 and 325e-5) with the network with five layers have a good agreement for a considered configuration, demonstrating a novel regression model that is able to rigorously extrapolate the mean of the maximum field in the other frequency steps that are not used in the training set.**

***Index Terms*—Deep neural networks, external and internal stirring, mean of maximum fields, reverberation chamber, small box.**

I. Introduction

REVERBERATION chambers are well-known test facilities that can be used as an alternative to outdoor testing or anechoic chamber for a variety of applications in electromagnetic compatibility (EMC)/vulnerability testing. High-conductivity walls inside a reverberation chamber, in association with mechanical or frequency stirring, can create a uniform and isotropic multipath environment, which can reproduce different EMC problems.

In the framework of immunity testing, statistical modeling of the maximum distribution and the associated mean of the maxima within the equipment under test (EUT) inside a reverberation chamber is a topic of interest, since it can lead to repeatable electronic system failures inside the shielding box. Traditionally, statistics of the maximum field and the associated mean have been found from an ideal parent Rayleigh distribution under specific conditions for which it applies [1]. However, the parent distribution can deviate from the Rayleigh distribution when the chamber is in one of the following conditions: 1) it is not stirred sufficiently (due to tuner size, placement, or configuration for a particular frequency range) and 2) the EUT configuration is sufficiently changed during testing to affect the field-level coupled into the EUT shielding enclosure [2], [3], and therefore, the parent distribution will not be known. In these cases, the generalized extreme value (GEV) distribution has been introduced and applied [4]–[7]. The GEV distribution has three parameters, i.e., location (m), scale (s), and shape (k), that need to be estimated [3], [8].

Nourshamsi *et al.* [9] have experimentally validated the application of the GEV distribution, while the parent distribution is derived from a Rayleigh distribution by placing a tuner inside the EUT box. The GEV's parameters were then estimated by using a maximum likelihood estimator (MLE). This method will be a successful approach depending on the algorithm to estimate the parameters and high-quality measurements, by including a large number of samples to decrease measurement or uncertainties. However, since the GEV distribution is not symmetric, the mean of the maximum field must be calculated [9] at each frequency point using all three estimated parameters. Additionally, it has been shown in [2] and [9], and peak-to-null ratios are large at low frequencies and get smaller at high frequencies independent of the aperture shape. This demonstrates a nonlinear relation between frequency steps and the associated mean of maximum fields, to statistically model the underlying function between these two variables.

In recent years, there has been a great development of machine learning in a variety of EMC applications [10]–[14]. Therefore, machine learning can provide an alternative approach to overcoming the mentioned difficulties, especially with decreasing residual uncertainties from other estimation techniques [15], which is a major concern in EMC testing, where the maximum

Manuscript received September 24, 2019; revised November 23, 2019 and January 22, 2020; accepted February 17, 2020. *(Corresponding author: Neda Nourshamsi.)*

Neda Nourshamsi was with the Department of Electrical and Computer Engineering, Oklahoma State University, Stillwater, OK 74078 USA. She is now with the Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824 USA (e-mail: n.noorshamsi@ieee.org).

Pedro Uría Rodríguez and Amir H. Jafari are with the Data Science Program, The George Washington University, Washington, DC 20052 USA (e-mail: pedrouriar@gwmail.gwu.edu; ajafari@email.gwu.edu).

Charles F. Bunting is with the Department of Electrical and Computer Engineering, Oklahoma State University, Stillwater, OK 74078 USA (e-mail: reverb@okstate.edu).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TEMC.2020.2975131

field exposure is of interest [7], [16]. In this article, we present a machine learning technique to predict the mean of maximum fields within the EUT inside a reverberation chamber, while the EUT front face is covered by a circular aperture. For this purpose, a multilayer perceptron (MLP) will be used, which is well known as a nonlinear regression method, to estimate the mean of maximum fields within the EUT inside a reverberation chamber [15], [17]. There is some theoretical [18], [19] and experimental [11] evidence that shows how an MLP can learn the underlying complex nonlinear function between inputs and outputs by having enough hidden layers. In this article, any layer that is not an output layer is considered a "hidden layer," while the output layer is the last layer of the network.

To use the MLP, the measurement was performed inside a nested chamber configuration, while mechanical stirring was applied in both stirrings inside the reverberation chamber (referred to external stirring) and the small cavity (referred to as internal stirring). This provides an unknown parent distribution compared to only external stirring, which provides a Rayleigh distribution at each frequency [2], [20]. The Anderson–Darling test [21] is applied to a set of measurement samples to ensure that the parent distribution deviates from Rayleigh. The universal approximation theorem states that an MLP (with a single hidden layer) can approximate continuous nonlinear functions on compact subsets of R^n [11], [15]. Since the frequency steps and the associate mean of maximum fields within the EUT inside a reverberation chamber meets the mentioned condition, we train an MLP to predict the mean of the maximum field at each frequency step on the data collected in our experiments. We split the data into training and testing and show that the MLP is indeed able to learn this function by minimizing the mean square error (MSE) between the network output (estimated mean of the maximum field) and target outputs (measured mean of maximum field). Overfitting is avoided by techniques such as early stopping, which improves the results significantly and enhances the training algorithm [22]. Also, by extending the MLP to the deep feedforward neural network (DFFN), we improve the extrapolation for testing results. Since the DFFN (defines as a MLP with more than two layers) is a more complex network with a very large number of parameters, we use more advanced optimization algorithms like *Adam* [23], [24], which uses the average gradient to find better learning rates over period iterations. The overall network is tested with a random sample test that is not implicitly used in the training set and is validated through the comparison with the MSE of the training set.

In previous work [9], we employed traditional maximum likelihood estimation to estimate the GEV's parameters and the associated mean of the maximum field at each frequency step individually. Here, the ultimate aim is to introduce a novel model with applying DFFN, with the feasibility of predicting the mean of maximum field at the frequency steps that are not used in the training set of the network.

II. Measurement Setup

The measurement was performed by placing a sample EUT shielding enclosure within an ETS-Lindgren SMART-80 rectangular reverberation chamber, as shown in Fig. 1. The SMART-80

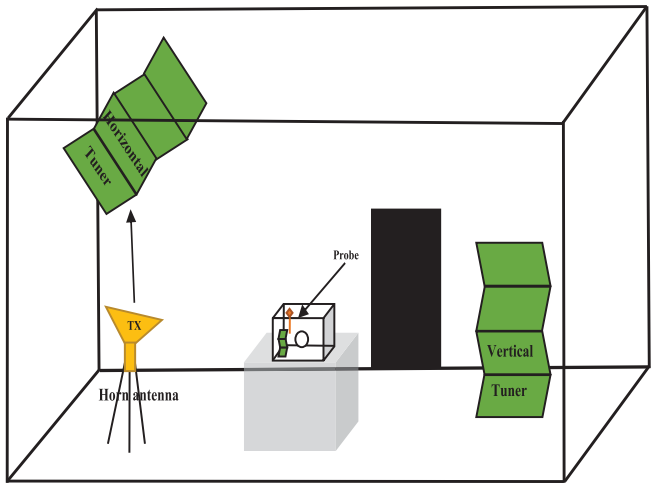

Fig. 1. Measurement setup in SMART-80 ETS-Lindgren.

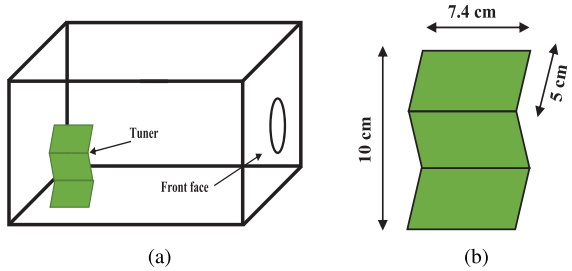

Fig. 2. EUT configuration. (a) Cavity information. (b) Tuner dimension.

has dimensions of 13.2 × 6.15 × 4.95 m. The reverberation chamber includes two large mechanical tuners: one oriented vertically and one oriented horizontally. Previous publication has shown that a uniform and isotropic random field can be produced using a single mechanical tuner at the frequencies considered here [9]. The vertical tuner was, therefore, rotated to provide the external field stirring and horizontal tuner held stationary during testing. The rectangular enclosure was fabricated from aluminum with dimensions of 30 cm × 30 cm × 12 cm, as shown in Fig. 2(a). A small internal z-fold tuner was placed 9 cm from the back wall of the enclosure and 6.5 cm from the left side wall. The z-fold tuner had a total height of 10 cm and a width of 7.4 cm, as shown in Fig. 2(b). The lower and upper "arms" of the z-shape were 5 cm in length. The field within the enclosure was sampled by using a straight-wire monopole probe of 10.4 cm length with a radius of 0.41 mm. The probe was fed through the enclosure wall by using an n-type connector, and it was located at four different sides of the box to increase the number of samples (spatial sampling). The transmitting antenna was a dual-ridge horn antenna placed 1 m from the EUT aperture face, as shown in Fig. 1. The transmitter antenna radiated toward the vertical tuner to provide a normal reverberant environment (nondirect illumination). The EUT enclosure was placed on a large Styrofoam box within the working volume of the reverberation chamber during the test. Since the dimensions of the SMART-80 are large, the chamber operates in the overmoded regime during the test, and therefore, individual spatial components of the magnitude field theoretically follow the Rayleigh distribution

under only external tuning [20]. The EUT was derived from the one considered in [6] and [9], but with the face containing only a circular aperture with a radius of 1.5 cm. Additionally, the front face of the EUT was fabricated from a 5-gauge aluminum sheet metal, giving a thickness of approximately 0.01 λ, where λ is the electromagnetic wavelength at the highest frequency (6 GHz) here.

The experiment was performed by rotating the SMART-80 tuner one step and then rotating the EUT tuner 50 times with the angle of 7.2°, while the S_{21} magnitude field coupled in the measurement probe was collected from 1–6 GHz over 1601 steps. This process was repeated for 50 positions (7.2°) within the SMART-80, which provide totally 2500 external and internal position pairs, while the EUT front face is covered by the circular aperture.

It has been suggested [25] to use the Weyl equation in order to determine an indication of frequency at which the enclosure is overmoded

$$N_{\text{modes}} \cong \frac{8\pi}{3} abd \frac{f^3}{c^3} - (a+b+d)\frac{f}{c} + \frac{1}{2} \tag{1}$$

where a, b, and d are the dimensions of the enclosure, N shows the total number of resonant modes, and c is the speed of light. The cavity is expected to be in the overmoded regime when N exceeds 60. Therefore, we expected the enclosure to be in the overmoded regime at frequencies above 2.7 GHz. Notably, statistic of the maximum field in the undermoded regime is not considered here and will be presented in a subsequence work.

III. Theory

A. Basic Network Architecture

In this section, we provide a mathematical overview of MLPs as well as a method that can be used to train them. The basic unit that makes up all artificial neural networks (ANNs) is called a (artificial) neuron. The input vector $\mathbf{p}$ to this unit is multiplied by a parameter $\mathbf{W}$ called the weight matrix, and the result is added to another parameter $\mathbf{b}$ called the bias vector. Their sum ($\mathbf{n}$, the net input) goes into an activation function (transfer function) $f(n)$ to get the neuron output $\mathbf{a}$. We can have this artificial neuron learn the line that fits any two-dimensional data the best, by updating the values of $\mathbf{W}$ and $\mathbf{b}$ through an iterative optimization algorithm called gradient descent (GD) [26]. By assigning a different weight to each input dimension, a neuron can be extended to model N-dimensional data. In this case, both $\mathbf{p}$ and $\mathbf{W}$ are N-dimensional vectors, and the neuron output equation is

$$\mathbf{a} = f(\mathbf{W} \cdot \mathbf{p} + \mathbf{b}). \tag{2}$$

By combining multiple neurons and using suitable transfer functions, this mathematical entity can model complex data (i.e., data that follows a high-order nonlinear function or pattern). However, this methodology is only true when we stack at least two layers of neurons sequentially, because it allows for the computation of many different transformations by combining together the transformations at each layer, together with the weights and biases.

According to the universal approximation theorem, an MLP with two layers (i.e, one hidden layer and one output layer) can approximate any continuous function on compact subsets of R^n [15]. However, we need a mathematical method to find the correct parameters that will result in such approximation. Most supervised ANNs are optimized by minimizing a performance index (a measure of the network's performance), which is obtained by comparing the network's output a with the expected output (usually called target). In our case, the target $\mathbf{y}$ is the mean of the maximum fields, and the performance index is just the MSE $\mathbf{F}(\mathbf{x}) = (\mathbf{y} - \mathbf{a})^2$, where $\mathbf{x}$ is a vector containing all the parameters in the network. The most widely used algorithm is called GD [15] and consists of an iterative process over k epochs, where the parameters are updated using the following equation [15]:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha \nabla F|_{\mathbf{x}_k}. \tag{3}$$

Because the gradient of a function always points to the direction of steepest ascent, this operation, given a fitting value of α (learning rate), will move $\mathbf{x}$ closer to the minimum. However, finding a suitable α in advance is rarely possible, and other problems may arise such as the algorithm getting stuck in a local minima and/or saddle points, which becomes more likely as the network's complexity increases. Furthermore, while for a single-layer network, the performance index depends directly on all the network's parameters, and thus, finding the gradients is straightforward, for multilayer networks, we need to start from the last layer, apply the chain rule, and propagate the gradients on the backward direction by applying a succession of matrix operations, a process called backpropagation [15].

B. Advanced Optimization

A key drawback of GD is that it only uses the slope of $F(\mathbf{x})$ to determine the search direction. As with the GD algorithm, *Adam* starts with the gradient $\mathbf{g}_k = \nabla F(\mathbf{x}_k)$. In addition, it computes a smoothed gradient and a smoothed squared gradient

$$\mathbf{m}_{k+1} = \beta_1 \mathbf{m}_k + (1-\beta_1)\mathbf{g}_k \tag{4}$$

$$\mathbf{v}_{k+1} = \beta_2 \mathbf{v}_k + (1-\beta_2)\mathbf{g}_k \circ \mathbf{g}_k \tag{5}$$

where $\circ$ is the Hadamard (element-by-element) product. These vectors are initialized with zeros as $\mathbf{m}_0 = 0$ and $\mathbf{v}_0 = 0$. The smoothed squared gradient $\mathbf{v}_{k+1}$ measures the second moment of the gradient, which is helpful in getting curvature information. The smoothed values $\mathbf{m}_k$ and $\mathbf{v}_k$ are biased toward zero, because of their initialization. To correct this, a second stage is performed

$$\hat{\mathbf{m}}_{k+1} = \frac{\mathbf{m}_{k+1}}{(1-\beta_1^{k+1})} \tag{6}$$

$$\hat{\mathbf{v}}_{k+1} = \frac{\mathbf{v}_{k+1}}{(1-\beta_2)} \tag{7}$$

where $0 < \beta_1$ and $\beta_2 < 1$ are constant exponential decay rates. The adjusted smoothed gradient $\hat{\mathbf{m}}_{k+1}$ is the search direction, and the adjusted smoothed squared gradient $\hat{\mathbf{v}}_{k+1}$ is used to adjust the learning rate

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha \hat{\mathbf{m}}_{k+1} (\oslash \sqrt{\hat{\mathbf{v}_{k+1}}} + \epsilon) \tag{8}$$

where α, $\oslash$, and ϵ represent the learning rate, the Hadamard (element-by-element) division, and a small positive constant to prevent division by zero, respectively.

Earlier work has been reported that using more than two layers (i.e, a DFFN) makes it fairly easier to approximate a function, as opposed to increasing the number of neurons on a single hidden layer indiscriminately (shallow neural network) [27]. However, from a practical point of view, increasing the number of layers leads to a harder training process. We will, therefore, experiment with different number of layers to find the best deep network for this particular problem, which is trained via the *Adam* optimization technique.

The behavior of hidden layers depends very much on the choice of the transfer function. The hyperbolic tangent function and the logsigmoid are two well-known transfer functions that we can use in hidden layers. Although the logsigmoid was used extensively in the past, it has been demonstrated theoretically [28] and experimentally [29] that the hyperbolic tangent function is preferred over the logsigmoid. This is due to the fact that tanh is symmetric around zero, leading to higher derivatives (faster training) and less parameter bias, while logsigmoid is symmetric around 0.5. Therefore, the hyperbolic tangent was selected as the hidden transfer functions. The function can be defined as [15]

$$\tanh(n) = \frac{e^n - e^{-n}}{e^n + e^{-n}} \tag{9}$$

where n represents the net input.

IV. Results and Analysis

In this article, the SMART-80 reverberation chamber was operating at frequency of 1–6 GHz; therefore, the chamber was in the overmoded regime by using (1). In this case, the measured field samples under only external stirring follow the Rayleigh distribution at each frequency [20]. The statistics of the samples from external stirring and internal stirrings were examined by using the Anderson–Darling test [21]. The test is performed with 99% confidence level at each frequency, and the rejection rate with the external stirring is 8% as expected [9]. Additionally, the rejection rate with internal stirring while the external stirring is fixed at one position is more than 90% at all frequency ranges, which shows the distribution deviating significantly from the Rayleigh distribution. This is likely due to either small dimensions of the tuner/aperture or sparsely moded characteristics of the cavity.

We considered the maximum field from internal stirring here. The measurement provides a total of 2500 unique external/internal tuner pair samples, yielding 50 sample maxima from the internal stirring. It has been suggested [30] that spatial sampling can be used to increase the number of internal field samples, which decrease the uncertainties in statistical analysis. Therefore, the probe was located on four different sides of the box at least $\lambda/2$ from the walls and other locations, where λ is the longest wavelength and 200 samples maxima were collected at each frequency. However, the internal maximum fields when the box is expected to be overmoded (from 2.7 GHz to 6 GHz) are considered here. Since the fields did not follow

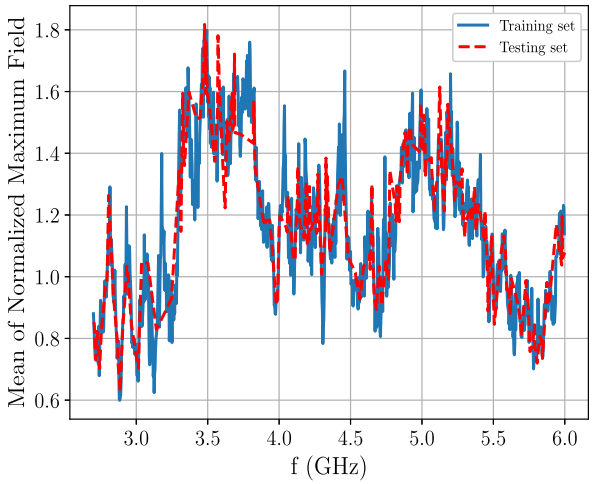

Fig. 3. Mean of normalized maximum field versus the frequency, split into training and testing set.

Rayleigh distribution, the GEV parameters were obtained, using an MLE, and the mean associated with the statistical maximum distribution was determined at each frequency step [31], [32]. Then, we used a supervised machine learning problem, in which the goal is to fit a function to these two-dimensional data (mean of maxima as a function of the frequency). The maximum of fields is considered to be in the 2.7–6-GHz frequency range. In other words, we train the MLP to minimize the MSE between the network outputs and the corresponding targets, which are the mean of maximum fields corresponding to the different input frequencies.

We randomly divide our data into two separate sets: training and testing, as shown in Fig. 3. The training data consist of 845 samples, while the testing data consist of 212 samples. Then, we proceeded to train multiple DFFNs with different number of layers and parameters to fit the training data and use the testing data as an early stopping condition for each of them. That is, while the network is training, we compute the MSE on the testing data and stop training when this error starts to increase, meaning that the particular epoch in which we halt the training process will differ for the different networks. This is one of the heuristic and most effective ways to avoid overfitting in deep learning [15].

During our experiments, there are a few things that remain the same, such as the number of input dimensions is just one, so our first weight matrix will have shape S^1x1, where S^1 is the number of neurons on the first layer. Because our target dimension is again only 1, S^M, the number of neurons on our output layer, must also be equal to 1. The output function is just a linear function, and for the rest of the layers, we chose the hyperbolic tangent as our activation function and *Adam* as our optimizer, and we performed a mix of batch and mini-batch GD, as well as try several learning rates ranging from e^{-3} to e^{-7}, which are updated during the training process for each model. The rest of the hyperparameters are chosen to be the default ones. We implement the networks[1] and run the training process on a single graphics processing unit (GPU) NVIDIA

[1]Online. [Available]: https://pytorch.org/PyTorch

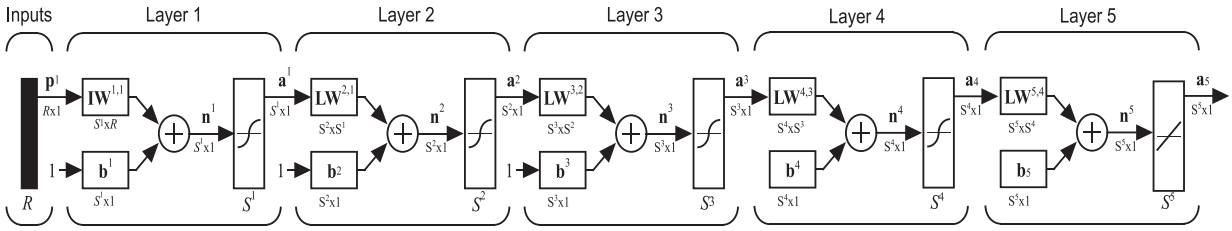

Fig. 4. Five-layer deep feedforward network for an equivalent regression model.

TABLE I
MSE FOR DIFFERENT MLPS

Layers	Parameters	Epochs	Training MSE	Testing MSE
2	1,501	89,570	0.04273	0.04348
2	180,001	506,950	0.04235	0.04117
3	20,601	282,745	0.04235	0.04352
5	181,200	93,270	0.00325	0.00643
7	662,400	513,602	0.00229	0.00670
9	1,623,600	450,982	0.00451	0.00773

Quadro P1000. PyTorch is a popular open-source deep learning framework released in 2016 that provides high flexibility for implementing and training neural networks. Notably, the use of a dynamic computational graph and CUDA programming enable the training of very deep networks in GPUs.

Table I shows some of our results for networks with different number of layers. Notably, a two-layer network is theoretically enough to approximate any function. However, it becomes much more difficult to learn a complicated function in practice. In our case, the second MLP has only one hidden layer, but almost the same number of parameters as the forth MLP, which has four hidden layers. The error for the latter is one order of magnitude less than the former, and the same goes for the number of epochs. These results show that it is very efficient to use deeper networks to approximate such a complex function. On the other hand, if we look at the testing errors, using too many layers can also mean a decrease in performance. This is due to the fact that the more complex the network, the more quickly it tends to overfitting. Thus, even with early stopping and other more advanced regularization methods, it is usually best to increase the number of layers and parameters gradually in order to find the optimal number of layers that estimates the target function with the lowest uncertainty. Therefore, in this article, a network with four hidden layers and one output layer, the hyperbolic tangent function (9) as the hidden transfer function, and the linear function as the output transfer function was chosen. Fig. 4 shows the structure of this network. The number of neurons on each layer are $S^1 = 100$, $S^2 = 200$, $S^3 = 400$, $S^4 = 200$, and $S^5 = 1$. For training of the network, *Adam* was used as the optimizer with $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = e^{-8}$. We started with $\alpha = 5e^{-4}$, switched to $\alpha = 5e^{-5}$ after the first 1000 epochs, and to $\alpha = 2.5e^{-5}$ at epoch 1500, and to $\alpha = 5e^{-6}$ after epoch 50 000. Both the training and testing errors during the training process are shown for this network in Fig. 5. In this particular

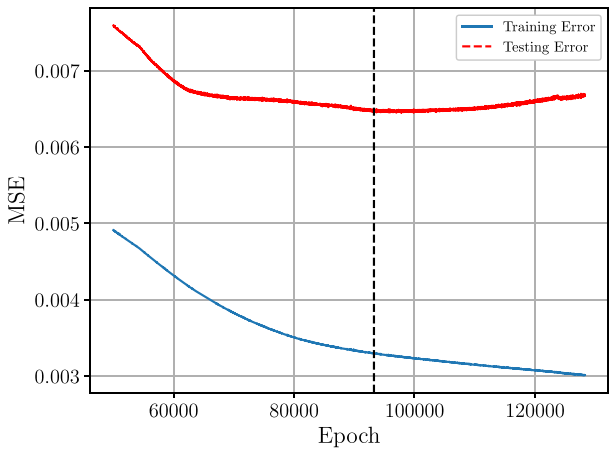

Fig. 5. Training and testing errors with the five-layer network. Early stopping is performed at epoch 93 270. Black dashed line shows the epoch 93 270.

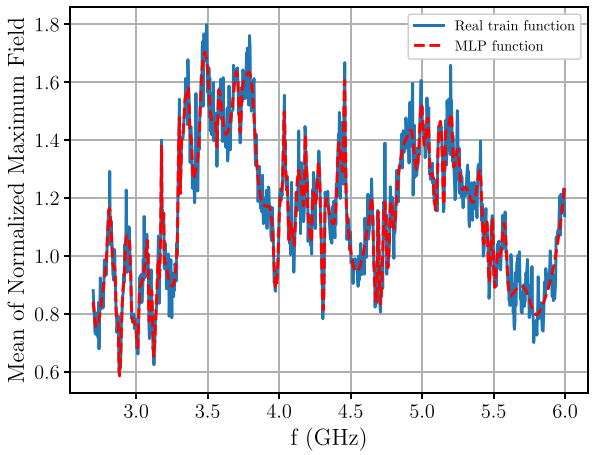

Fig. 6. Real train target (from MLE) and trained DFFN output versus the frequency. MLP with five layers and 181 200 parameters and MSE = 0.00325.

case, we stop at epoch number 93 270: if we leave the network training longer, it will overfit the training data and begin to fit the noise, which will result in a poorer performance on the testing data, as can be seen in Fig. 5, where the testing error (in red) begins to increase after this epoch, while the training error keeps decreasing.

Fig. 6 shows the trained dataset compare to the MLE traditional method. Overall, the five-layer network was trained accurate enough in most of the frequency steps.

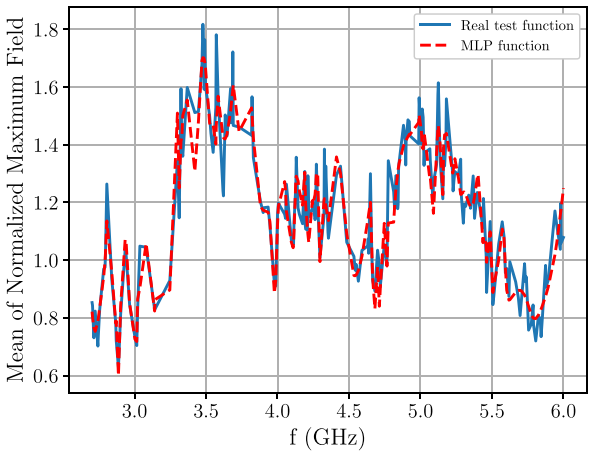

Fig. 7. Real test target (from MLE) and DFFN output versus frequency testing set. MLP with five layers and 181 200 parameters and MSE = 0.00645.

TABLE II
PREDICTED AND REAL TARGETS FOR RANDOM TEST SAMPLES

Test frequency (GHz)	MLE output	DFFN output
2.90	0.81	0.84
3.03	1.04	0.90
3.48	1.59	1.69
4.10	1.05	1.04
4.74	1.05	1.12
5.74	0.94	0.86

Finally, we validated the regression function by using the testing set that was selected randomly, as shown in Fig. 3. The responses of the network to the testing set are shown in Fig. 7, and Table II contains some test frequencies randomly sampled from the test set, together with their associated mean of maximum field (targets) and the values predicted by the best MLP (the one with five layers). Overall, very good agreement is achieved in all testing datasets using the DFFN and the results from the traditional MLE method. The approximated function from the training network can predict the mean of the maximum field at frequency steps considered in the testing set, and no retraining is required to use the model, while the MLE estimator (traditional method) needs to be performed individually at each frequency step [9]. Therefore, the approximated function from the DFFN can provide the results faster (approximately ten times) than the MLE method.

V. CONCLUSION

In this article, we proposed a deep neural network technique to predict the mean of the maximum field strength at each frequency in an overmoded regime within a cavity configuration testing, while the parent distribution is non-Rayleigh. Both internal and external stirrings were used to generate enough field samples and a non-Rayleigh distribution. Then, an MLP was developed and tested with the maximum magnitude of field samples within the inner cavity inside the reverberation chamber. Different numbers of layers were trained, and the MSE from training sets was demonstrated. Effects of overfitting on the regression model were shown, and early stopping was presented to prevent it. *Adam* optimization was used, since it converged faster than other algorithms. Testing MSE with a five-layer perceptron is low enough compared to other layers' results, which provide a significant reduction in uncertainties compared to using other estimators [3]. However, it has been shown that, in general, the performance of the network can still be improved by applying a committee of networks and using a Monte Carlo cross-validation, which is deferred to later work [15]. Finally, we note that the trained deep neural network can be used to determine the mean of the maximum fields that is not contained in the measured data, while the traditional MLE is not capable of. Additionally, prediction of the mean of the maximum field strength at another location of the aperture and extensions of the technique in order to predict the mean of maximum fields within other aperture shapes will be a topic of future work.

REFERENCES

[1] T. H. Lehman and G. J. Freyer, "Characterization of the maximum test level in a reverberation chamber," in *Proc. IEEE Int. Symp. Electromagn. Compat.*, Austin, TX, USA, Aug. 1997, pp. 44–47.
[2] J. C. West, R. Bakore, and C. F. Bunting, "Statistics of the current induced within a partially shielded enclosure in a reverberation chamber," *IEEE Trans. Electromagn. Compat.*, vol. 59, no. 6, pp. 2014–2022, Dec. 2017.
[3] N. Nourshamsi, "Statistical analysis of electromagnetic complex cavities," Ph.D. dissertation, Dept. Elect. Comput. Eng., Oklahoma State Univ., Stillwater, OK, USA, 2018.
[4] G. Orjubin, "Maximum field inside a reverberation chamber modeled by the generalized extreme value distribution," *IEEE Trans. Electromagn. Compat.*, vol. 98, no. 1, pp. 104–113, Feb. 2007.
[5] G. Gradoni and L. R. Arnaut, "Generalized extreme-value distributions of power near a boundary inside electromagnetic reverberation chambers," *IEEE Trans. Electromagn. Compat.*, vol. 52, no. 3, pp. 506–515, Aug. 2010.
[6] N. Nourshamsi, J. C. West, and C. F. Bunting, "Investigation of electromagnetic complex cavities by applying the generalized extreme value distribution," in *Proc. IEEE Int. Symp. Electromagn. Compat.*, Long Beach, CA, USA, Jul. 30–Aug. 4, 2018, pp. 622–627.
[7] C. E. Hager, J. D. Rison, and G. B. Tait, "Electromagnetic probability-of-effect assessment tool for high-power susceptibility testing," *IEEE Trans. Electromagn. Compat.*, vol. 58, no. 4, pp. 1306–1313, Aug. 2016.
[8] J. M. Mendel, *Estimation Theory for Signal Processing, Communication and Control*, 2nd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 1995.
[9] N. Nourshamsi, J. C. West, C. Hager, and C. Bunting, "Generalized extreme value distributions of fields in nested electromagnetic cavities," *IEEE Trans. Electromagn. Compat.*, vol. 61, no. 4, pp. 1337–1344, Aug. 2019.
[10] Y.-F. Shu, X.-C. Wei, J. Fan, R. Yang, and Y.-B. Yang, "An equivalent dipole model hybrid with artificial neural network for electromagnetic interference prediction," *IEEE Trans. Microw. Theory, Techn.*, vol. 67, no. 5, pp. 1790–1797, May 2019.
[11] N. Nourshamsi, M. T. Hagan, and C. F. Bunting, "Estimation of required absorbing material dimensions inside metal cavities using neural networks," in *Proc. IEEE Int. Symp. Electromagn. Compat.*, Washington, DC, USA, Aug. 2017, pp. 69–74.
[12] G. Yang, K. Chen, Z. Yu, Y. Zhang, F. Zhou, and J. He, "An inversion method for evaluating lightning current waveform based on time series neural network," *IEEE Trans. Electromagn. Compat.*, vol. 59, no. 3, pp. 887–893, Jun. 2017.
[13] T. Lu, J. Sun, K. Wu, and Z. Yang, "High-speed channel modeling with machine learning methods for signal integrity analysis," *IEEE Trans. Electromagn. Compat.*, vol. 60, no. 6, pp. 1957–1964, Dec. 2018.
[14] V. Devabhaktuni, C. F. Bunting, D. Green, D. Kvale, L. Mareddy, and V. Rajamani, "A new ANN-based modeling approach for rapid EMI/EMC analysis of PCB and shielding enclosures," *IEEE Trans. Electromagn. Compat.*, vol. 55, no. 2, pp. 385–394, Nov. 2012.

[15] M. T. Hagan, H. B. Demuth, M. H. Beale, and O. D. Jesús, *Neural Network Design*, 2nd ed. Boston, MA, USA: PWS, 2014.
[16] N. Nourshamsi, J. C. West, and C. F. Bunting, "Effects of aperture dimension on maximum field level inside nested chambers by applying the generalized extreme value distribution," in *Proc. IEEE Int. Symp. Electromagn. Compat. Signal Integrity Power Integrity*, Long Beach, CA, USA, Jul. 30–Aug. 4, 2018, pp. 628–633.
[17] V. Kurkova, "Kolmogorov's theorem and multilayer neural networks," *Neural Netw.*, vol. 5, no. 3, pp. 501–506, 1992.
[18] J. Xu, L. Zhang, M. Sapozhnikov, and J. Fan, "Application of deep learning for high-speed differential via TDR impedance fast prediction," in *Proc. IEEE Int. Symp. Electromagn. Compat. Signal Integrity Power Integrity*, Long Beach, CA, USA, Jul. 30–Aug. 4, 2018, pp. 645–649.
[19] W. Na, F. Feng, C. Zhang, and Q. J. Zhang, "A unified automated parametric modeling algorithm using knowledge-based neural network and l1 optimization," *IEEE Trans. Microw. Theory Techn.*, vol. 65, no. 3, pp. 729–745, Mar. 2017.
[20] D. A. Hill, "Electromagnetic theory of reverberation chamber," Nat. Inst. Standards Technol., Boulder, CO, USA, Tech. Rep. 1506, 1998.
[21] M. A. Stephens, "EDF statistics for goodness of fit and some comparisons," *J. Amer. Statist. Assoc.*, vol. 69, no. 347, pp. 730–737, Sep. 1974.
[22] Y. Bengio, "Practical recommendations for gradient-based training of deep architectures," in *Neural Networks: Tricks of the Trade*. Berlin, Germany: Springer, 2012.
[23] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," version 9, Jan. 2017, *arXiv:1412.6980* [cs.LG].
[24] A. Jafari and M. Hagan, "Application of new training methods for neural model reference control," *Eng. Appl. Artif. Intell.*, vol. 74, pp. 312–321, Sep. 2018.
[25] G. Tait, C. Hager, M. B. Slocum, and M. O. Hatfield, "On measuring shielding effectiveness of sparsely moded enclosures in a reverberation chamber," *IEEE Trans. Electromag. Compat.*, vol. 55, no. 2, pp. 231–240, Apr. 2013.
[26] H. Robbins and S. Monro, "A stochastic approximation method," *Ann. Math. Statist.*, vol. 22, no. 3, Sep. 1951.
[27] H. Mhaskar and T. Poggio, "Deep vs. shallow networks: An approximation theory perspective," *Anal. Appl.*, vol. 14, pp. 829–848, 2016.
[28] Y. LeCun, L. Bottou, G. B. Orr, and K. R. Muller, "Efficient backprop," in *Neural Networks: Tricks of the Trade, This Book is an Outgrowth of a 1996 NIPS Workshop*. London, U.K.: Springer-Verlag, 1998, pp. 9–50.
[29] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Performance analysis of various activation functions in generalized MLP architectures of neural networks," *Int. J. Artif. Intell. Exp. Syst.*, vol. 1, no. 4, pp. 111–122, Dec. 2010.
[30] L. R. Arnaut, "Compound exponential distributions for undermoded reverberation chambers," *IEEE Trans. Electromagn. Compat.*, vol. 44, no. 3, pp. 442–457, Aug. 2002.
[31] P. Embrechts, C. Kluppelberg, and T. Mikosch, *Modelling Extremal Events for Insurance and Finance*. Berlin, Germany: Springer, 1996.
[32] S. Kotz and S. Nadarajah, *Extreme Value Distributions: Theory and Applications*. Singapore: World Scientific, 2000.

Neda Nourshamsi (Member, IEEE) received the M.S. degree from the University of Strasbourg, Strasbourg, France, in 2014, and the Ph.D. degree from the Oklahoma State University, Stillwater, OK, USA, in 2018, both in electrical engineering.

She is currently a Postdoctoral Research Associate with the School of Electrical and Computer Engineering, Michigan State University, East Lansing, MI, USA. Her current research interests include statistical electromagnetics in complex cavities, antennas and propagation, and microwave wireless techniques for motion detection and classification.

Dr. Nourshamsi is reviewer of IEEE TRANSACTIONS ON ELECTROMAGNETIC COMPATIBILIY and Springer Neural Computing and Applications. Dr. Nourshamsi is a member of the IEEE Electromagnetic Compatibility (EMC) Society and involved with several technical committees and educational activities through the IEEE EMC Society.

Pedro Uría Rodríguez is working toward the master's degree with the Data Science Program, George Washington University (GWU), Washington, DC, USA. He has a background in physics and decided to make the leap to data science when he realized its many applications and relevance in today's world. During his first year at GWU, he quickly developed a strong interest in machine learning and neural networks, which was solidified through his professional and academic experience in the fields of natural language processing and computer vision.

Amir H. Jafari received the M.S. degree from the American University of Sharjah, Sharjah, United Arab Emirates, in 2011, and the Ph.D. degree from the Oklahoma State University, Stillwater, OK, USA, in 2016, both in electrical engineering.

He is currently an Adjunct Assistant Professor with the Data Science Program, The George Washington University, Washington, DC, USA. He is also a Google Cloud Faculty Expert. He has been conducting research in the areas of statistical modeling, neural network design, artificial intelligence, and predictive modeling for the last ten years. His research has encompassed a variety of application areas: signal processing, dynamic systems, neural network training, optimization, machine learning, and deep learning. In recent years, his research has been focused on the use of neural networks for data science applications, time-series predictions, clustering and modeling, deep learning network architectures, and training big data on GPU. Dr. Jafari is a Gust Editor of Springer Neural Computing and Applications.

Charles F. Bunting (Fellow, IEEE) received the A.A.S. degree in electronics technology from Tidewater Community College, Virginia Beach, VA, USA, in 1985, the B.S. degree in engineering technology from Old Dominion University, Norfolk, VA, in 1989, and the M.S. and Ph.D. degrees in electrical engineering from Virginia Tech, Blacksburg, VA, in 1992 and 1994, respectively.

He was with the Naval Aviation Depot, Norfolk, as an Apprentice, an Electronics Mechanic, and an Electronics Measurement Equipment Mechanic, from 1981 to 1989. From 1994 to 2001, he was an Assistant/Associate Professor with the Department of Engineering Technology, Old Dominion University, where he worked closely with the NASA Langley Research Center on electromagnetic field penetration in aircraft structures and reverberation chamber simulation using finite-element techniques. In 2001, he joined the faculty of Oklahoma State University (OSU), Stillwater, OK, USA, as an Associate Professor and became a Full Professor in 2011. In 2012, he was named as the Halliburton Professor of Engineering and in 2016 the Bellmon Chair. In late 2012, he became an Associate Dean of Research for the College of Architecture and Technology. He is the Director of the Robust Electromagnetic Field Testing and Simulation Laboratory, OSU, where he is a Professor of Electrical and Computer Engineering. Under his leadership, OSU has become an important center of electromagnetic compatibility (EMC) continuing education for the application of test methodologies for radiated emissions and susceptibility. The reverberation chamber course has been held annually in Stillwater since 2007, drawing from industry and government, providing engineers with practical training in the application of statistical methods required for proper reverberation chamber operation (over 200 attendees total). His research interests include electromagnetic field penetration in aircraft structures and reverberation chamber simulation using numerical and experimental methods, engineering education, applied computational electromagnetics, statistical electromagnetics, and electromagnetic characterization and application of reverberation chambers for communication systems, including radio frequency identification and unmanned aircraft systems. He has also performed research in the development of near-infrared tomography and microwave acoustic tomography for prostate cancer detection.

Dr. Bunting was a Distinguished Lecturer for the IEEE EMC Society from 2011 to 2012, the IEEE EMC Society Board of Directors from 2011 to 2016, the Technical Activities Committee Vice Chair from 2014 to 2016, and the Vice-President of Technical Services from 2017 to 2021.