Progress In Electromagnetics Research, PIER 98, 233–249, 2009

GAUSSIAN PROCESS MODELING OF CPW-FED SLOT ANTENNAS

J. P. De Villiers and J. P. Jacobs

Department of Electrical, Electronic and Computer Engineering
University of Pretoria
Pretoria 0002, South Africa

Abstract—Gaussian process (GP) regression is proposed as a structured supervised learning alternative to neural networks for the modeling of CPW-fed slot antenna input characteristics. A Gaussian process is a stochastic process and entails the generalization of the Gaussian probability distribution to functions. Standard GP regression is applied to modeling S_{11} against frequency of a CPW-fed second-resonant slot dipole, while an approximate method for large datasets is applied to an ultrawideband (UWB) slot with U-shaped tuning stub — A challenging problem given the highly non-linear underlying function that maps tunable geometry variables and frequency to S_{11}/ input impedance. Predictions using large test data sets yielded results of an accuracy comparable to the target moment-method-based full-wave simulations, with normalized root mean squared errors of 0.50% for the slot dipole, and below 1.8% for the UWB antenna. The GP methodology has various inherent benefits, including the need to learn only a handful of (hyper) parameters, and training errors that are effectively zero for noise-free observations. GP regression would be eminently suitable for integration in antenna design algorithms as a fast substitute for computationally intensive full-wave analysis.

1. INTRODUCTION

The design of antennas on layered media, such as those employed in microstrip and coplanar waveguide (CPW) technologies, typically requires finding the (optimal) geometry that would yield the desired performance characteristics (e.g., impedance bandwidth, gain, axial ratio, etc.). Sophisticated and accurate full-wave analyses software such as moment-method-based codes are available for analyzing,

Corresponding author: J. P. De Villiers (pieter.devilliers@up.ac.za).

in general, any particular geometry instance of the antenna under consideration. However, optimization tools such as genetic algorithms might require thousands of such full-wave analysis of different geometries of the antenna to be optimized [1]. The ensuing computational cost could render the optimization process prohibitively cumbersome if not ultimately infeasible. In recent years this problem has been widely addressed through the use of artificial neural networks (ANNs); in the EM community at large, ANNs have in fact been applied to both antenna modeling [2, 3] and microstrip circuit modeling and design [4] including transmission line structures [5, 6] (other applications of neural networks have included array synthesis [7], direction-of-arrival estimation [8], radar target identification [9], and identification of complex Bragg gratings [10]). After supervised training of an ANN with a training set consisting of a limited number of input-output pairs, e.g., particular geometries and their concomitant performance characteristics obtained from full-wave simulations, the ability of the network to generalize over the input space makes it possible to quickly obtain the desired performance characteristics for inputs not previously presented to the network. Hence great savings in computational effort can been achieved compared to when all input cases have to be analyzed by means of full-wave simulations.

In practice, neural networks are however not always easy or straightforward to implement due to the lack of a "principled framework" [11] for making certain decisions regarding setup and implementation, such as the type of activation function and learning rate to choose, and whether momentum should be used. Even for the multilayer perceptron (MLP), which is ubiquitous in its application to supervised learning problems, the number of hidden units that should be chosen often is an open question, relying for a solution on a trial-and-error approach or experience [12, 13]. While adaptive techniques have been proposed for deleting or adding hidden units to a neural network during training [14], this adds further complexity to the modeling process. Furthermore, mappings involving highly non-linear underlying functions require more hidden units — However, ANNs with too large a number of hidden units may not be feasible [1]. Large ANNs can also have problems with overfitting [11].

In this article, we introduce Gaussian process (GP) regression as a supervised learning alternative to neural networks for the modeling of antenna characteristics. A Gaussian process is a stochastic process that entails the generalization of the Gaussian probability distribution to functions. The Gaussian nature of the distribution leads to tractable — even relatively simple — calculations when learning and inference need to be performed. Under appropriate conditions, Gaussian processes

can be considered equivalent to large neural networks [11]; however, Gaussian processes are generally easier to implement and interpret. One reason is that a Gaussian process model requires training of far fewer parameters (in the order of the dimension of the input vectors) than, e.g., a MLP with one hidden layer, where the number of weights to be learned typically is in the order of $N_i \times N_h + N_h \times N_o$, with N_i, N_h, and N_o the number of input, hidden, and output nodes respectively. A Gaussian process is an instance of a so-called kernel machine; it is differentiated by its probablistic basis from the support vector regression machine (SVRM), another kernel machine that has very recently found application in antenna-related problems (e.g., [13]).

The layout of the article is as follows. Section 2 gives key theoretical background to standard regression with Gaussian processes following [11]. In Section 3, standard GP regression is applied to the modeling of the input reflection coefficient S_{11} against frequency of a CPW-fed second-resonant slot dipole antenna. Then an approximate regression technique for large datasets, the subset of datapoints method, is applied to the challenging problem of modeling S_{11} against frequency of an ultrawideband (UWB) CPW-fed slot antenna with U-shaped tuning stub. Conclusions and suggestions for further work are presented in Section 4.

2. THEORETICAL BACKGROUND

A Gaussian process describes a distribution over functions. More specifically, it is a mathematical set consisting of an infinite number of random variables, of which any subset is jointly Gaussian [11, 15] (it is a natural extension of a jointly Gaussian distribution to the case where the mean vector is infinitely long and the covariance matrix is of infinite by infinite dimension). A GP can be notated as $f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}'))$, with $\mathbf{x}$ a position in $\mathbb{R}^D$-dimensional space, and $m(\mathbf{x})$ and $k(\mathbf{x}, \mathbf{x}')$ its mean and covariance functions respectively, defined as [11, (2.13)]. Note that the GP encapsulates *all* possible functions in the vast space of functions that subscribe to $m(\mathbf{x})$ and $k(\mathbf{x}, \mathbf{x}')$. The model is semi-parametric in the sense that any sample function is not specified in terms of a finite number of parameters (such as weights in the case of a linear model), but rather directly in the space of functions. Consider for example a finite (practical) training data set of n observations, $\mathcal{D} = \{(\mathbf{x}_i, y_i) \mid i = 1, \ldots, n\}$. The inputs $\mathbf{x}_i$ are column vectors of dimension D, while the corresponding output targets y_i are scalars. The corresponding Gaussian process $f(\mathbf{x})$ in this case would be implemented as the collection of random variables $f_i = f(\mathbf{x}_i)$, with any n-dimensional point under their jointly Gaussian

distribution representing n values of a sample function with index set the set of inputs $\mathbf{x}_i$.

The only parameterisation that takes place is the specification of *hyperparameters* which determine the properties of the mean and covariance functions. The present study uses the squared-exponential covariance function $k(\mathbf{x}_p, \mathbf{x}_q) = \sigma_f^2 \exp(-\frac{1}{2}(\mathbf{x}_p - \mathbf{x}_q)^\top M(\mathbf{x}_p - \mathbf{x}_q))$, which gives the covariance between the output random variables $f(\mathbf{x}_p)$ and $f(\mathbf{x}_q)$. Here, the matrix $M = \mathrm{diag}(\boldsymbol{\ell})$, with $\boldsymbol{\ell}$ the vector of positive characteristic length-scale parameters corresponding to the components of the input vectors, and σ_f^2 is the signal variance (length-scale parameters are indicative of how quickly change occurs along the corresponding dimensions of the input space). Together, $\boldsymbol{\ell}$ and σ_f constitute the hyperparameters of the covariance function.

In order to carry out predictions, a jointly Gaussian distribution (usually of zero mean) is assumed over the n random variables which represent the training outputs and are contained in column vector $\mathbf{f}$, and the n_* random variables representing the test outputs contained in $\mathbf{f}_*$ — This is the prior distribution:

$$\begin{bmatrix} \mathbf{f} \\ \mathbf{f}_* \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K(X,X) & K(X,X_*) \\ K(X_*,X) & K(X_*,X_*) \end{bmatrix}\right) \tag{1}$$

In (1), $K(X, X_*)$ is the $n \times n_*$ matrix of covariances evaluated between all possible pairs of n training and n_* test outputs, where the columns of the $D \times n$ matrix X are the training input column vectors $\{\mathbf{x}_i \mid i = 1, \ldots, n\}$, and X_* likewise contains the test input vectors; the other submatrices of the covariance matrix are similarly defined. (Training of the Gaussian process prior to inference is described below.)

The distribution of the test outputs conditioned on the known training outputs $\mathbf{y}$, or the posterior distribution, can then be expressed as $\mathbf{f}_*|X_*, X, \mathbf{y} \sim \mathcal{N}(\mathbf{m}, \boldsymbol{\Sigma})$ (cf. [11, Eq. (2.19)]), with mean vector $\mathbf{m}$ and covariance matrix Σ given by

$$\mathbf{m} = K(X_*, X)K(X, X)^{-1}\mathbf{y} \tag{2}$$

$$\Sigma = K(X_*, X_*) - K(X_*, X)K(X, X)^{-1}K(X, X_*) \tag{3}$$

In the above, the predictive mean $\mathbf{m}$ contains the most likely values of the test outputs associated with the test input vectors in X_*, while the diagonal of the covariance matrix $\boldsymbol{\Sigma}$ gives the corresponding predictive variances. Conditioning on the known training data can be interpreted as retaining in the posterior distribution only functions that pass through the training data points.

The computational requirement for GP regression is $\mathcal{O}(n^3)$ due to the required inversion of $K(X, X)$ which is of dimension $n \times n$. A number of methods of dealing with large datasets have been

proposed [11]; here a subset of datapoints approach has been adopted (see Section 3.2).

The hyperparameters may be found through a structured methodology which involves a process similar to Bayesian model selection [11]. It entails finding the hyperparameters for which the negative log marginal likelihood, or error function, is a minimum. The log marginal likelihood in the noise-free case is given by (cf. [11, (2.30)])

$$\log p(\mathbf{y}|X) = -\frac{1}{2}\mathbf{y}^\top K^{-1}\mathbf{y} - \frac{1}{2}\log|K| - \frac{n}{2}\log 2\pi \tag{4}$$

with $K = K(X, X)$, $|K|$ the determinant of K, $\mathbf{y}$ the training target column vector, and X the matrix of input column vectors.

In what follows, training (optimization of hyperparameters) and inference were carried out using existing computer implementations of general algorithms for GP regression and negative log likelihood minimization [11]; the regression algorithm employs numerically efficient Cholesky decomposition for inverting $K(X, X)$.

3. MODELING OF INPUT CHARACTERISTICS OF CPW-FED SLOT ANTENNAS

3.1. Second-resonant Slot Dipole Antenna

Figure 1 shows a slot dipole antenna fed by CPW on a single-layer dielectric substrate. When operated in the vicinity of its second resonance, the antenna exhibits significantly wider impedance bandwidth than a microstrip patch antenna on a comparable substrate; other advantages include easy integration with microwave circuit components due to the uniplanar configuration of the CPW feed line and radiating slot. CPW-fed slots have frequently been used in array applications, e.g., [16, 17]. An example of a practical situation that requires many full-wave analyses of different slot geometries over a range of frequencies is the generation of isolated slot self-admittance data for use in a non-uniform linear array iterative design algorithm such as [17], with the additional requirement that off-resonance evaluation of array performance must be possible.

GP regression was used to model the input reflection coefficient S_{11} of a single CPW-fed slot antenna over an input space spanned by slot dimensions L and W, and frequency f. The parameter ranges were $(1 \leq W \leq 5)$ mm, $(20 \leq L \leq 29)$ mm, and $(4.5 \leq f \leq 5.5)$ GHz; it was established that the extremes of the geometry range correspond to significantly different S_{11}-against-frequency characteristics. Independent GP regression models were set up for $\Re[S_{11}]$ and $\Im[S_{11}]$, the real and imaginary parts of S_{11}.

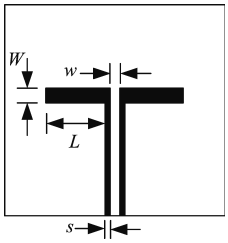

Figure 1. Top view of CPW-fed slot dipole antenna.

Training and test data were generated using IE3D [18], a full-wave moment-method-based simulator that uses magnetic currents to model the slots while assuming laterally infinite ground planes and dielectrics. A fixed substrate was used with dielectric constant $\epsilon_r = 3.38$ and thickness $h = 0.508\,\text{mm}$, and the feed line was a $50\,\Omega$ CPW with $w = 4\,\text{mm}$ and $s = 0.2\,\text{mm}$. A database of potential training data was generated by computing S_{11} against frequency for 30 antenna geometries corresponding to pairs of values of W and L that were obtained by uniform random sampling from the two-dimensional geometry space; the frequencies were 4.5 to 5.5 GHz in steps of 0.1 GHz. A set of $n = 84$ training input vectors $\{\mathbf{x}_i = [W_i\ L_i\ f_i]^\top \mid i = 1, \ldots, n\}$ with corresponding output target scalars $y_i = \Re[S_{11i}]$ or $\Im[S_{11i}]$ were randomly selected from this database. Test data involving 15 new pairs of W and L were similarly generated; the total number of test points was $n_* = 165$.

For each regression, the training set was used to "learn" (optimize) the hyperparameters $\boldsymbol{\Theta} = [\ell_W, \ell_L, \ell_f, \sigma_f]$ of the covariance function by minimizing the negative log marginal likelihood. Since the real and imaginary parts of S_{11} are modeled separately, each part will have its own set of hyperparameters. The hyperparameter values thus obtained were $\boldsymbol{\Theta}_\mathbf{R} = [4.4018, 1.0323, 1.9271, 0.3200]$ and $\boldsymbol{\Theta}_\mathbf{I} = [3.1802, 1.1443, 2.5290, 0.3510]$ for the real and imaginary parts respectively. Predictions were then made for the test data set. Table 1 gives the root mean square error (RMSE) and the percentage normalized RMSE (defined as in [12]) for both training and test sets (note that $-1 \leq \Re[S_{11}], \Im[S_{11}] \leq 1$). For the test points, the normalized RMSE was 0.50% for both $\Re[S_{11}]$ and $\Im[S_{11}]$ regressions — A good result confirmed by (absolute values of) normalized maximum residuals of about 2% for each model. The RMSE values for the training points were effectively zero, confirming that the squared-exponential covariance function adopted for the purposes of the present study had sufficient flexibility.

Table 1. Root mean squared error (RMSE) and percentage normalized RMSE for training points ($n = 84$) and test points ($n_* = 165$).

	$\Re[S_{11}]$		$\Im[S_{11}]$	
	RMSE	RMSE (%)	RMSE	RMSE (%)
Train	3.200e-12	3.41e-10	8.574e-12	8.59e-10
Test	4.705e-3	0.50	3.137e-3	0.50

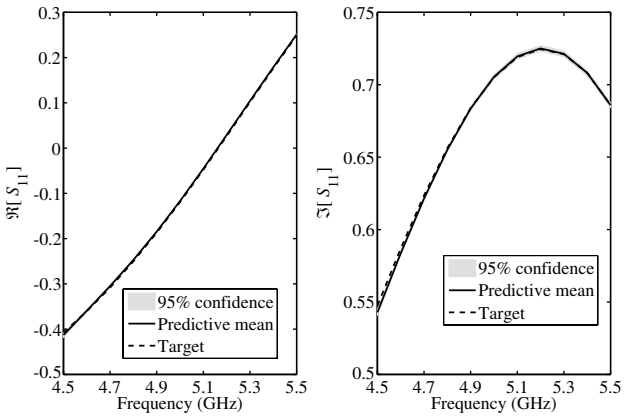

Figure 2. Regression results corresponding to the test geometry $[W, L] = [4.257, 26.78]$ mm after training with $n = 84$ data points.

Figure 2 shows, for the test geometry with $W = 4.257$ mm and $L = 26.78$ mm, the predictive mean, predictive 95% confidence region (± 2 predictive standard deviations) about the mean, and target function for the $\Re[S_{11}]$ and $\Im[S_{11}]$ regression models after training with 84 data points (this was a typical result). The predictive standard deviation obtained from (3) is the model's own estimate, based on the training and test *input* data, of the uncertainty with respect to its predictions (neither training nor test targets are required for making this estimate). The predictive mean and target functions are virtually coincident for both models, with confidence regions that are very narrow (they are barely visible on the graphs), indicating very low uncertainty.

3.2. Ultrawideband Slot Antenna with U-shaped Tuning Stub

Recent years have seen the introduction of a variety of types of CPW-fed antennas for dual-band [19, 20] and ultrawideband [21–27] applications. The focus of the present study is a CPW-fed ultrawideband rectangular slot antenna with U-shaped tuning stub [28], which is shown in Fig. 3. The modeling task, as defined here, was to describe S_{11} against frequency over the ultrawide band of frequencies 2–10 GHz as the five tuning stub dimensions a, b, c, d, and e are varied; their ranges are given in Table 2 (other dimensions and parameters were the same as in [28] for purposes of comparison later on). Such a model would be useful when it is desired to optimize the bandwidth according to a prespecified ultrawideband criterium.

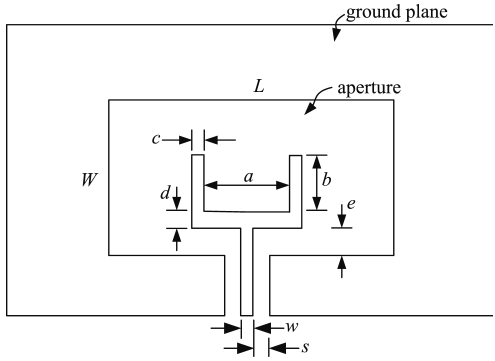

Figure 3. Top view of CPW-fed rectangular slot antenna with U-shaped tuning stub. $W = 32.2\,\text{mm}$; $L = 21.1\,\text{mm}$; $w = 1.88\,\text{mm}$ and $s = 0.125\,\text{mm}$ ($50\,\Omega$ CPW); $\epsilon_r = 3.38$; $h = 0.813\,\text{mm}$ [28].

Table 2. Parameter ranges for ultrawideband slot antenna.

Parameter	Min.	Max.
a (mm)	6	14
b (mm)	6	12
c (mm)	0.5	2.5
d (mm)	0.5	2.5
e (mm)	0.5	2.5
f (GHz)	2	10

The modeling of the input impedance of ultrawideband antennas is challenging given that it is highly nonlinear, containing multiple resonances within the band with steep zero-crossings in the imaginary part and sharp peaks in the corresponding portions of the resistance curves. Many data points are required for an accurate representation, resulting in a large neural network if this is the regression means of choice (e.g., 250 hidden units and 6769 weights in [1]). As in the case of the second-resonant slot dipole antenna, we chose to model the reflection coefficient S_{11}, which at least has the advantage of contained extreme values since $|S_{11}| \leq 1$. Training and test data were generated using IE3D in a manner similar to the case of the slot dipole. In the present case, a total of $n = 50625$ potential training data points were generated, consisting of 625 randomly chosen geometry points, each evaluated at 81 frequencies. The test set consisted of 175 geometry points, also randomly selected and evaluated at 81 frequencies, resulting in a total of $n_* = 14175$ test data points.

Since the computational complexity of the GP regression method scales according to $\mathcal{O}(n^3)$, using the complete training data set is computationally prohibitive; this is predominantly due to the inversion of $K(X, X)$ required in (2) and (3). To address this problem, the subset of datapoints method [11] was utilised — It entails selecting an "active" set of m training points from the total set of n training points, where $m < n$. Only the active set is then used in the regression.

The first step was to optimize the hyperparameters using a random subset of 2048 training points, which can be handled with computational ease. It is important that enough training points be used, since the hyperparameters are determined once only, and the performance of all subsequent regressions depend on their choice. The hyperparameters $\mathbf{\Theta} = [\ell_a, \ell_b, \ell_c, \ell_d, \ell_e, \ell_f, \sigma_f]$ thus determined were $\mathbf{\Theta_R} = [1.1831, 1.0347, 3.6484, 2.5633, 3.0156, 0.1885, 0.2300]$ and $\mathbf{\Theta_I} = [1.2317, 1.0693, 3.8164, 2.6073, 2.6112, 0.1952, 0.2439]$.

The next step was to select a subset of training data points for making predictions with respect to the test data set using (2) (this is different from standard GP regression, where the training data used for determining the hyperparameters are also used for performing the regression). The most straightforward, but also sometimes the least efficient way of implementing the selection, is to simply randomly select a subset in such a way that each point has an equal probability of being chosen. However, since this might be wasteful of data, it was deemed appropriate to also consider jointly selecting a subset of datapoints according to some optimal criterion. The optimal method of selecting a subset of data points involves combinatorics and for most large datasets this becomes infeasible. A more practical method is

"greedy" sequential selection. This involves starting with a small base set of randomly selected training data and effecting GP regression at the same positions as the remaining *unused* data points. The predictive variance (PV) criterion selects the point where the predictive variance is maximised to be included in the training set. This effectively searches the remaining unused data point positions where the model is the most uncertain, and then reduces the uncertainty at that point to zero by including that data point in the training set. On the other hand, when using the squared error (SE) criterion, remaining unused data point positions are searched to find where the SE is maximised. The SE at that point is then reduced to zero by including the relevant data point. Ideally, one would like to cycle through all remaining unused data points at each step. However, for the present problem, that would entail performing regression at 50625 points in order to include only a single point in the training set. This is computationally infeasible, and hence a new subset of 4096 randomly selected unused data points was considered at each inclusion step. This means that each unused point would have been visited approximately once in every ten steps.

For the size of training subsets considered here, and given the above set of hyperparameters, performance of the three methods was fairly similar, with the PV and SE greedy selection methods somewhat outperforming the random selection method in terms of root mean squared error. In particular, for $m = 1024$ data points the percentage RMSE for the test data set was 3.49%, 3.13%, and 2.93% for the random (RND), PV, and SE methods respectively (training errors were effectively zero). This is already very good given the great variability of S_{11} against frequency over the geometry input space. For $m = 4096$ training points, the percentage RMSE further decreased to 1.89%, 1.51%, and 1.63% for RND, PV, and SE respectively. Since computational complexity of GP regression is proportional to $\mathcal{O}(n^3)$, the smaller training set might be preferable and quite likely would be acceptable for most practical problems (in [1], a broadband antenna was successfully optimized using a neural network substitute for full-wave computations of input impedance that had a RMSE training error of 16.4%). It is however noteworthy that in order to perform regression according to (2) for a series of test points, as would be the case in an optimization context, inversion of $K(X, X)$ has to be performed *once* only, namely when doing the first regression; the stored result can be used for subsequent regressions.

Figure 4 shows the predictive mean, 95% confidence region, and target S_{11} function for a test geometry representing a typical result; the training data subset had 4096 data points and was selected according to

the SE criterion. The corresponding predictive mean and target input impedance plots are shown in Fig. 5 indicating impressive modeling of sharp resonance peaks. The impedance plots were obtained using the standard transformation of the S_{11} test points according to $Z_{in} =$

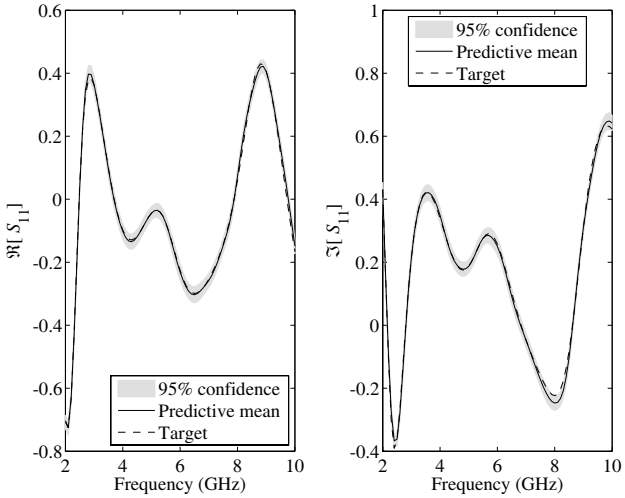

Figure 4. $\Re[S_{11}]$ and $\Im[S_{11}]$ against frequency for UWB test geometry $[a, b, c, d, e] = [9.57, 10.75, 1.54, 2.37, 1.38]$ mm.

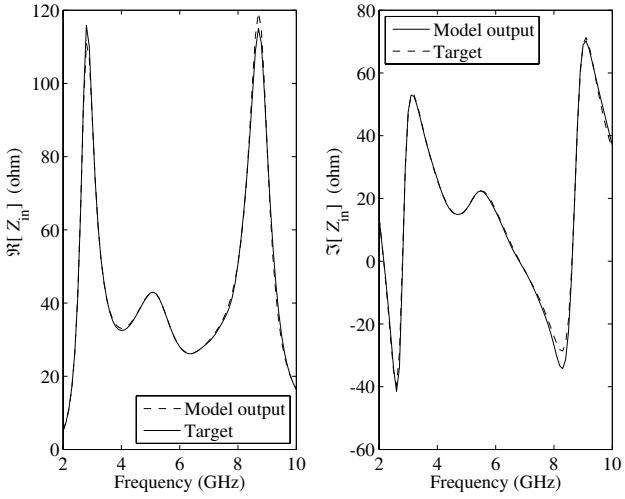

Figure 5. $\Re[Z_{in}]$ and $\Im[Z_{in}]$ against frequency corresponding to test geometry of Fig. 4.

$Z_0 \frac{1+S_{11}}{1-S_{11}}$, and thus required the results from both regression models ($Z_0 = 50\,\Omega$).

Figure 6 shows plots of S_{11} against frequency for a test geometry designed for optimal bandwith given in [28] (this test geometry was not part of the original randomly generated test set). The corresponding $|S_{11}|$ plot is shown in Fig. 7 for comparison with [28, Fig. 2]. Again, the regression results are very good, and of comparable quality with respect to those of the test geometry of Figs. 4 and 5.

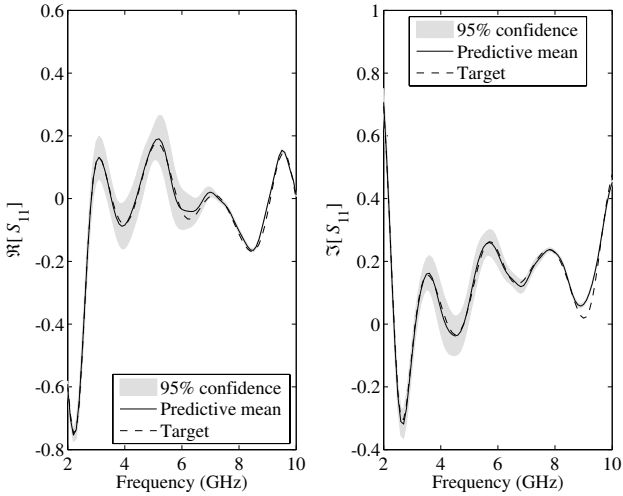

Figure 6. $\Re[S_{11}]$ and $\Im[S_{11}]$ against frequency for test geometry optimized for bandwidth obtained from [28]: $[a, b, c, d, e] = [12, 9, 2, 1, 1]$ mm.

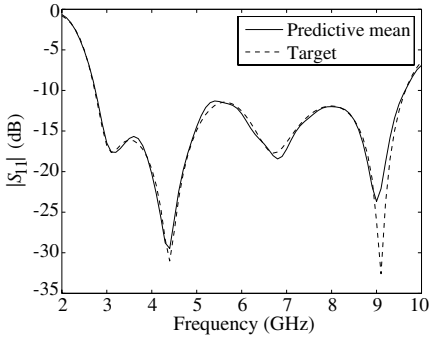

Figure 7. $|S_{11}|$ against frequency for UWB test geometry corresponding to Fig. 6.

Figure 8 shows the scatter plot of model predictions plotted against the corresponding target values for the complete test set (14175 points) utilizing the above training set of 4096 points selected using the SE criterion. The intercepts and slopes obtained by linear regression were 0.996 and -0.000762 ($\Re[S_{11}]$) and 0.997 and -0.000482 ($\Im[S_{11}]$). The models' performances are very satisfactory, with small intercepts indicating negligible model bias. The slopes are very close to one, indicating negligible multiplicative factor error. There are a few significant outliers, but they number by the tens within a test set consisting of 14175 points.

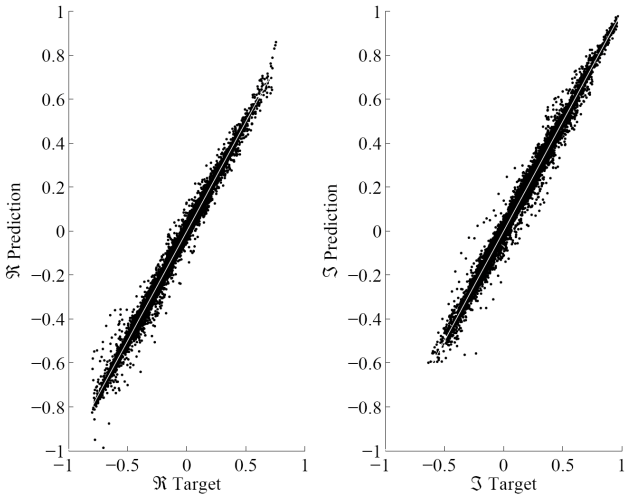

Figure 8. Model output against target value scatter plot for 14175 uniformly random selected test points.

4. CONCLUSION

GP regression presents an attractive structured alternative methodology to neural networks for supervised learning and prediction of antenna characteristics that might otherwise only be attainable through computationally intensive, repetitive full-wave calculations. It requires far less tuning, validation and training than neural network strategies. In the present study, GP regression has been applied to the modeling of S_{11} (and/or input impedance) against frequency of CPW-fed slots. The second-resonant slot dipole, which has well-behaved characteristics, required a fairly small set of training samples for accurate modeling of S_{11} as a function of two tunable geometry parameters and frequency

using basic GP regression. Normalized RMSEs of 0.50% were achieved for both the $\Re[S_{11}]$ and $\Im[S_{11}]$ regressions. The more difficult problem involved a CPW-fed UWB slot antenna with U-shaped tuning stub, where five tunable stub dimensions and frequency map to S_{11} through a highly non-linear underlying function. Since tens of thousands of potential training points were available and computational complexity scales with size of the data set, direct application of GP regression was not feasible, and an approximate technique — the subset of data points method — was used. Excellent predictions were demonstrated using a training set of 4096 points selected using the SE method for data selection (a further 2048 randomly selected data points were required for initial training of the hyperparameters), with normalized RMSEs of 1.63% and 1.75% for the $\Re[S_{11}]$ and $\Im[S_{11}]$ regressions. Various benefits of the GP methodology were discussed, including the fact that the only tunable parameters are the hyperparameters of the covariance function, which can be determined in a structured manner based on Bayesian model selection. Also, in modeling applications which have noiseless observations such as the ones described in this paper, training errors will effectively be zero. This is in contrast with neural network models which can retain a fairly significant error at these points when regression involving highly-nonlinear functions is carried out. Computational complexity does not scale significantly with input dimensionality — This makes GP regression particularly suitable for modeling problems which have a high number of input parameters (extension of the model of the ultrawideband slot to include the length and width of the outer rectangular slot as tunable parameters, if desired, would be straightforward and only add two length-scale hyperparameters to the model). Modeling of the real and imaginary parts of S_{11} were achieved using two separate GP models. However, there exists strong correlation between these components, and a model taking these correlations into account is sure to maintain performance levels with the use of fewer training data points (cf. [29]). Selection and sampling of training points might be further explored along the lines of [30].

REFERENCES

1. Kim, Y., S. Keely, J. Ghosh, and H. Ling, "Application of artificial neural networks to broadband antenna design based on a parametric frequency model," *IEEE Trans. Antennas Propagat.*, Vol. 55, No. 3, 669–674, 2007.
2. Patnaik, A., D. E. Anagnostou, R. K. Mishra, C. G. Christodoulou, and J. C. Lyke, "Applications of neu-

ral networks in wireless communications," *IEEE Antennas Propagat. Mag.*, Vol. 46, No. 3, 130–137, 2004.

3. He, Q. Q., Q. Wang, and B. Z. Wang, "Conformal array based on pattern reconfigurable antenna and its artificial neural model," *Journal of Electromagnetic Waves and Applications*, Vol. 22, No. 1, 99–110, 2008.
4. Rayas-Sanchez, J. E., "EM-based optimization of microwave circuits using artificial neural networks: The state-of-the-art," *IEEE Trans. Microw. Theory Tech.*, Vol. 52, No. 1, 420–435, 2004.
5. Kaya, S., M. Turkmen, K. Guney, and C. Yildiz, "Neural models for the elliptic- and circular-shaped microshield lines," *Progress In Electromagnetics Research B*, Vol. 6, 169–181, 2008.
6. Yildiz, C. and M. Turkmen, "Quasi-static models based on artificial neural neworks for calculating the characteristic parameters of multilayer cylindrical coplanar waveguide and strip line," *Progress In Electromagnetics Research B*, Vol. 3, 1–22, 2008.
7. Ayestarán, R. G., F. Las-Heras, and J. A. Martinez, "Non uniform-antenna array synthesis using neural networks," *Journal of Electromagnetic Waves and Applications*, Vol. 21, No. 8, 1001–1011, 2007.
8. Zainud-Deen, S. H., H. A. El-Azem Malhat, K. H. Awadalla, and E. S. El-Hadad, "Direction of arrival and state of polarization estimation using radial basis function neural network (RBFNN)," *Progress In Electromagnetics Research B*, Vol. 2, 137–150, 2008.
9. Kizilay, A. and S. Makal, "A neural network solution for identification and classification of cylindrical targets above perfectly conducting flat surfaces," *Journal of Electromagnetic Waves and Applications*, Vol. 21, No. 14, 2147–2156, 2007.
10. Rostami, A. and A. Yazdanpanah-Goharrizi, "Hybridization of neural networks and genetic algorithms for identification of complex Bragg gratings," *Journal of Electromagnetic Waves and Applications*, Vol. 22, No. 5–6, 643–664, 2008.
11. Rasmussen, C. E. and C. K. I. Williams, *Gaussian Processes for Machine Learning*, MIT Press, Cambridge, Massachussets, 2006.
12. Zhang, Q.-J., K. C. Gupta, and V. K. Devabhaktuni, "Artificial neural networks for RF and microwave design — From theory to practice," *IEEE Trans. Microw. Theory Tech.*, Vol. 51, No. 4, 1339–1350, 2003.
13. Angiulli, G., M. Cacciola, and M. Versaci, "Microwave devices and antennas modelling by support vector regression machines," *IEEE Trans. Magnetics*, Vol. 43, No. 4, 1589–1592, 2007.

14. Devabhaktuni, V. K., M. C. E. Yagoub, and Q.-J. Zhang, "A robust algorithm for automatic development of neural-network models for microwave applications," *IEEE Trans. Microw. Theory Tech.*, Vol. 49, No. 12, 2282–2291, 2001.
15. MacKay, D. J. C., *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, 2003.
16. Qiu, M., M. Simcoe, and G. V. Eleftheriades, "High-gain meanderless slot arrays on electrically thick substrates at millimeter-wave frequencies," *IEEE Trans. Microw. Theory Tech.*, Vol. 50, No. 2, 517–528, 2002.
17. Jacobs, J. P. and J. Joubert, "Design of a linear nonuniform CPW-fed slot array with reduced sidelobe levels," *Microw. Opt. Tech. Lett.*, Vol. 51, No. 9, 2175–2178, 2009.
18. Zeland Software, *IE3D Users Manual*, Release 14, 2007.
19. Zhang, L., Y. C. Jiao, Y. L. Zhao, G. Zhao, Y. Song, Z. B. Wong, and F. S. Zhang, "Dual-band CPW-fed double H-shaped slot antenna for RFID application," *Journal of Electromagnetic Waves and Applications*, Vol. 22, No. 8–9, 1050–1055, 2008.
20. Zhang, T. L., Z. H. Yan, L. Chen, and Y. Song, "A compact dual-band CPW-fed planar monopole antenna for WLAN applications," *Journal of Electromagnetic Waves and Applications*, Vol. 22, No. 14–15, 2097–2104, 2008.
21. Zhang, G. M., J. S. Hong, B. Z. Wang, Q. Y. Qin, J. B. Mo, and D. M. Wan, "A novel multi-folded UBW antenna fed by CPW," *Journal of Electromagnetic Waves and Applications*, Vol. 21, No. 14, 2109–2119, 2007.
22. Chen, Y.-I., C.-L. Ruan, and L. Peng, "A novel ultra-wideband bow-tie slot antenna in wireless communication systems," *Progress In Electromagnetics Research Letters*, Vol. 1, 101–108, 2008.
23. Lee, S. H., J. N. Lee, J. K. Park, and H. S. Kim, "Design of the compact UWB antenna with PI-shaped matching stub," *Journal of Electromagnetic Waves and Applications*, Vol. 22, No. 10, 1440–1449, 2008.
24. Wang, X., Z. F. Yao, Z. Cui, L. Luo, and S. X. Zhang, "Band-notched characteristics for CPW-fed printed monopole antenna with E shape slot," *Journal of Electromagnetic Waves and Applications*, Vol. 22, No. 16, 2171–2178, 2008.
25. Yao, Z. F., X. Wang, S. G. Zhou, B. H. Sun, and Q. Z. Liu, "Compact ultra-wideband slot antenna with dual band-notched characteristics," *Journal of Electromagnetic Waves and Applications*, Vol. 22, No. 13, 1765–1774, 2008.

26. Yao, Z. F., S. G. Zhou, X. Wang, L. Sun, B. H. Sun, and Q. Z. Liu, "Study of the band-notched functions for CPW-fed UWB antenna," *Journal of Electromagnetic Waves and Applications*, Vol. 22, No. 17–18, 2309–2321, 2008.
27. Yin, X.-C., C.-L. Ruan, C.-Y. Ding, and J.-H. Chu, "A planar U type monopole antenna for UWB applications," *Progress In Electromagnetics Research Letters*, Vol. 2, 1–10, 2008.
28. Chair, R., A. A. Kisk, and K. F. Lee, "Ultrawide-band coplanar waveguide-fed rectangular slot antenna," *IEEE Antennas Wireless Propagat. Lett.*, Vol. 3, 227–229, 2004.
29. Boyle, P. and M. Frean, "Dependent gaussian processes," *Advances in Neural Information Processing Systems*, Vol. 17, 217–224, 2005.
30. Jones, D. R., "A taxonomy of global optimization methods based on response surfaces," *Journal of Global Optimization*, Vol. 21, 345–383, 2001.