

Efficient Resonant Frequency Modeling for Dual-Band Microstrip Antennas by Gaussian Process Regression

J. P. Jacobs

Abstract—A methodology based on Gaussian process regression (GPR) for accurately modeling the resonant frequencies of dual-band microstrip antennas is presented. Two kinds of dual-band antennas were considered, namely a U-slot patch and a patch with a center square slot. Predictive results of high accuracy were achieved (normalized root-mean-square errors of below 0.6% in all cases), even for the square-slot patch modeling problem where all antenna dimensions and parameters were allowed to vary, resulting in a seven-dimensional input space. Training data requirements for achieving these accuracies were relatively modest. Furthermore, the automatic relevance determination property of GPR provided (at no additional cost) a mechanism for enhancing qualitative understanding of the antennas’ resonance characteristics—a facility not offered by neural network-based strategies used in related studies.

Index Terms—Antennas, Gaussian processes, modeling.

I. INTRODUCTION

Microstrip antennas form a ubiquitous part of contemporary microwave engineering. Crucial for their design is knowledge about their resonant frequency behavior. The present study focuses on the modeling—by means of Gaussian process regression (GPR)—of the resonant frequencies of in particular dual-band microstrip antennas as a function of the antenna’s design variables (e.g., geometry dimensions, substrate dielectric constant). Such models are useful when fast, accurate estimates of resonant frequencies need to be obtained for multiple different configurations of design variables (i.e., antenna instances).

For example, an antenna engineer might need a dual-frequency antenna to function within an environment where surface area is limited and/or where the choice of dielectric laminates is restricted. A model of the sort presented here would enable the engineer to quickly investigate various combinations of design variables (e.g., patch dimensions, substrate height, dielectric constant)—selected to adhere to the given constraints—in order to identify a configuration that might best meet the resonant frequency requirements (or alternatively confirm that the requirements might not be realizable given the constraints). To run such a model would be orders of magnitude

faster than a direct simulation approach, which would entail obtaining each candidate antenna’s $|S_{11}|$ -against-frequency response via full-wave simulations (e.g., moment-method-based) and then extracting the resonance frequencies. For narrowband antennas, computational costs of the direct approach could be even higher given that more simulations might be needed in the vicinity of resonances to accurately account for rapid changes in $|S_{11}|$.

Previous studies concerned with antenna resonant frequency modeling have mostly focused on microstrip patch antennas intended for single-frequency operation [1]–[4]. Dual resonant frequency modeling has only recently been considered, namely for a slotted patch antenna with five design variables [5]. In [1] and [2], the modeling method was neural networks; in [3] and [4], neural networks combined with fuzzy logic techniques; and in [5], knowledge-based neural networks.

In this letter, Gaussian process regression (GPR) [6] is demonstrated for the first time to be a highly accurate and efficient technique for modeling the resonant frequencies of dual-band microstrip antennas with multiple design variables (up to seven, which exceeds the number of design variables handled in [5]). Compared to neural networks as used in [1]–[5], Gaussian processes are substantially easier to implement and interpret. For instance, GPR requires training of only a handful of parameters (in the order of the input vector dimension), as opposed to the potentially large number of weights in a multilayer perceptron (MLP) neural network (determined by the number of layers and nodes). Because of its Bayesian roots, GPR requires virtually no manual tuning. In contrast, neural networks require multiple design decisions pertaining to, e.g., the number of hidden layers and hidden units to use and learning rate and momentum (often determined based on trial-and-error and/or experience). During neural network training, an independent validation data set might be required (in addition to training and test sets) to check for overfitting; this could be problematic if data are few or expensive to obtain. A key GPR strength is that it can give very good predictive results even when trained on relatively small training data sets [7]. Furthermore, inherent to GPR is the particularly noteworthy property of automatic relevance determination (ARD) [6], which can be viewed as a weighting by the model of the respective input features (i.e., design variables) according to their relative importance. It is shown that ARD can lead to increased insight into an antenna’s resonance frequency behavior—in general, this would be especially useful when

The author is with the Centre for Electromagnetism, Department of Electrical, Electronic and Computer Engineering, University of Pretoria, Pretoria 0002, South Africa (e-mail: jppjacobs@postino.up.ac.za).

the design space is of relatively high dimensionality and/or when the model user does not have experience using the kind of antenna under consideration. The neural networks used previously for resonant frequency modeling, e.g., [1] and [2], do not provide a comparable facility.

Previous studies concerned with GPR and antenna modeling have shown that GPR is effective in modeling highly nonlinear antenna input characteristics such as $|S_{11}|$ as a function of design parameters and frequency over a band of interest, e.g., [8] and [9]. The models here differ from those in [8] and [9] in that only design variables are used as input, with a “higher-level” response feature—resonant frequency—as model output.

The GPR approach to resonant frequency modeling is illustrated using two examples of probe-fed microstrip antennas intended for dual-band operation: a U-slot patch antenna [10], [11] with four design variables, and a patch antenna with center square slot [12] with seven design variables. Section II provides an overview of the main points of GPR following [6]. Section III gives details of the proposed methodology and presents numerical verification. In Section IV, automatic relevance determination is discussed with respect to the example antennas, and conclusions are presented in Section V.

II. BRIEF OVERVIEW OF GAUSSIAN PROCESS REGRESSION

A Gaussian process (GP) is a set consisting of an infinite number of random variables of which any finite subset has a jointly Gaussian distribution. It can be considered a natural extension of the Gaussian multivariate distribution to functions, i.e., the case where the mean vector is infinitely long and the covariance matrix is of infinite dimension. A Gaussian process $f(\mathbf{x})$ is completely specified by its mean function $m(\mathbf{x})$ and covariance function $k(\mathbf{x}, \mathbf{x}')$, with \mathbf{x} and \mathbf{x}' points in input space. The mean function gives the expectation of the process, and the covariance function defines the covariance between the output random variables $f(\mathbf{x})$ and $f(\mathbf{x}')$ [6, Eq. (2.13)]. Once the mean and covariance functions are defined, Gaussian processes can be handled following rules of probability as applied to multivariate Gaussian distributions. (It is interesting to note that Gaussian processes can be considered equivalent to multilayer-perceptron neural networks with infinitely many hidden units and weights governed by Gaussian priors [13].)

Consider a finite training data set of n input–output-pair observations, $\{(\mathbf{x}_i, y_i) | i = 1, \dots, n\}$. The inputs \mathbf{x}_i are vectors of dimension D , while the output targets y_i are scalars. The corresponding Gaussian process $f(\mathbf{x})$ 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 $\{\mathbf{x}_i\}$.

In order to carry out predictions, a jointly Gaussian *prior* distribution of zero mean is assumed over the n random variables that represent the training outputs (contained in a vector \mathbf{f}), and the n_* random variables corresponding to the test outputs contained in \mathbf{f}_* .

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

where $K(X, X_*)$ is the $n \times n_*$ matrix of covariances evaluated between all pairs of n training and n_* test outputs, and X/X_* are matrices containing the training/test input vectors (other submatrices are similarly defined). The distribution of the test outputs conditioned on the known training outputs \mathbf{y} , i.e., the *posterior* distribution (again multivariate Gaussian) has mean vector \mathbf{m} and covariance matrix Σ [6]

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

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

where \mathbf{y} is the vector of training targets. 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 Σ gives the corresponding predictive variances.

The only parameterization that takes place involves the hyperparameters that determine the properties of the mean and covariance functions, and hence the matrices K in (1)–(3). The present study considers four well-known covariance functions, namely the squared-exponential (SE) covariance function

$$k_{\text{SE}}(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left(-\frac{1}{2}r^2\right) \quad (4)$$

the rational quadratic (RQ) covariance function

$$k_{\text{RQ}}(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \left(1 + \frac{1}{2\alpha}r^2\right)^{-\alpha} \quad (5)$$

and two Matérn class covariance functions

$$k_{\nu=3/2}(\mathbf{x}, \mathbf{x}') = \sigma_f^2 (1 + \sqrt{3}r) \exp(-\sqrt{3}r) \quad (6)$$

and

$$k_{\nu=5/2}(\mathbf{x}, \mathbf{x}') = \sigma_f^2 (1 + \sqrt{5}r + 5r^2/3) \exp(-\sqrt{5}r). \quad (7)$$

In (4)–(7)

$$r = \left(\sum_{k=1}^D \frac{(x_k - x'_k)^2}{\tau_k^2} \right)^{1/2} \quad (8)$$

where x_k and x'_k are the k th components of input vectors \mathbf{x} and \mathbf{x}' , and the positive length-scale parameters $\{\tau_k | k = 1, \dots, D\}$ are indicative of how quickly significant change occurs in the underlying function along the corresponding dimensions of the input space. σ_f^2 governs the overall variance of the process. The squared-exponential covariance function gives rise to very smooth Gaussian processes that are infinitely differentiable; Matérn covariance functions are less smooth, with the function corresponding to $\nu = 3/2$ the least so. The rational quadratic function can be viewed as a scale mixture of squared exponential functions, and becomes the squared exponential function in the limit when $\alpha \rightarrow \infty$ [6].

The hyperparameters σ_f^2 and $\{\tau_k | k = 1, \dots, D\}$ [as well as α in (5)] may be found by minimizing the negative of the log marginal likelihood [6], which is given by

$$\log p(\mathbf{y}|X) = -\frac{1}{2}\mathbf{y}^T K^{-1}\mathbf{y} - \frac{1}{2} \log |K| - \frac{n}{2} \log 2\pi \quad (9)$$

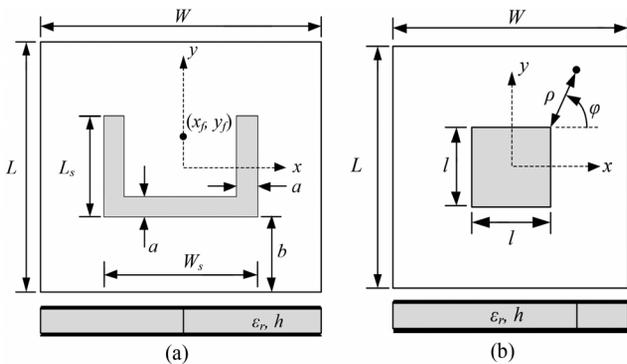


Fig. 1. Top and side views of (a) probe-fed patch antenna with U-shaped slot (patch centered with respect to origin) and (b) probe-fed patch antenna with square slot (patch and slot centered with respect to the origin). In the top views, probe positions are indicated by black dots, and slot apertures are shaded. Ground planes and dielectrics were assumed to be of laterally infinite extent.

where $|K|$ is the determinant of $K(X, X)$. This usually gradient-based optimization procedure can be considered the “training” of the Gaussian process model, and the value of the minimized negative log marginal likelihood can be interpreted as an indication of the suitability of the user-specified mean and covariance functions (i.e., the model) in view of the training data. Optimizing the length scale parameters equates to ARD [6] (cf. Section I): Input vector elements (features) with relatively long-length scales will have less influence on calculating the covariance than those with short-length scales, and therefore will have less influence on the model [cf. (4)–(8)]. Hence, ARD may be considered an implicit form of feature selection.

The computational requirement for GP regression is $O(n^3)$ due to the inversion of $K(X, X)$, which is of dimension $n \times n$ [cf. (2) and (3)].

III. VERIFICATION EXAMPLES

A. Probe-Fed U-Slot Patch Antenna (Antenna 1)

Fig. 1(a) shows the geometry of a probe-fed U-slot patch antenna implemented with air as substrate [10], [11]. The design vector was $\mathbf{x} = [W_s L_s a b]^T$, with individual variable ranges given by $22 \leq W_s \leq 34$ mm, $10 \leq L_s \leq 20$ mm, $0.4 \leq a \leq 1$ mm, and $3 \leq b \leq 10$ mm. Other dimensions (fixed) were $W = 62$ mm, $L = 47$ mm, and $h = 4$ mm, with the probe positioned at $(x_f, y_f) = (0, 0)$ mm.

Twenty training data input vectors \mathbf{x} were randomly selected from the design (input) space using Latin hypercube sampling (LHS), and the 20 geometry instances that they specified were simulated in FEKO [14]. First and second resonant frequencies $f_{\text{res},1}$ and $f_{\text{res},2}$ were determined from the minima of the $|S_{11}|$ -against-frequency responses obtained from the simulations (frequencies where the imaginary part of the input impedance becomes zero were not used as reference as this does not necessarily happen in the vicinity of minima [10]). Next, eight separate Gaussian process models were trained according to the procedure described in Section II: four models, each utilizing a different covariance function from (4)–(7), with the $f_{\text{res},1}$ values as training targets \mathbf{y} ; and four models

TABLE I
PREDICTIVE ERRORS OF GP MODELS ON TEST DATA

	ANTENNA 1 ($n = 20, n^* = 75$) ¹		ANTENNA 2 ($n = 80, n^* = 100$)	
	$f_{\text{res},1}$	$f_{\text{res},2}$	$f_{\text{res},1}$	$f_{\text{res},2}$
%RMSE	0.45	0.53	0.56	0.41
R	0.9971	0.9972	0.9993	0.9996
Covariance function	SE	RQ	Matérn, $\nu = 5/2$	Matérn, $\nu = 5/2$

¹ n : number of training points (input–output pairs)

n^* : number of test points

with the $f_{\text{res},2}$ values as training targets. A zero mean function was assumed in all cases. The total training time for the eight models was about 50 s on a Dell PowerEdge 1955 blade with dual Intel Xeon 5355 processors.

Model selection was performed by choosing from the $f_{\text{res},1}$ models the model that produced the lowest negative log marginal likelihood [cf. (9)], and likewise for the $f_{\text{res},2}$ models. In the case of f_1 , the model implementing the squared exponential covariance function produced the lowest negative log marginal likelihood; in the $f_{\text{res},2}$ case, the rational quadratic covariance function fared best.

Following the above, each of the selected models was presented with previously unseen test data obtained by randomly selecting 75 input vectors from the design space using LHS, simulating the corresponding geometries in FEKO, and identifying the resonant frequencies, i.e., test targets. Table I gives the models’ predictive accuracies [percentage root-mean-square errors (RMSEs) normalized to the center frequencies of the test target ranges, which were $2.03 \leq f_{\text{res},1} \leq 2.64$ GHz and $2.91 \leq f_{\text{res},2} \leq 3.95$ GHz]. Table I also provides the linear correlation coefficient R for predictions and targets. Highly accurate results were achieved, with normalized RMSEs $< 0.6\%$ in both cases and correspondingly high correlation coefficients.

B. Probe-Fed Patch Antenna With Center Square Slot (Antenna 2)

Fig. 1(b) shows the geometry of a probe-fed patch antenna with centered square slot implemented on a single-layer dielectric substrate [12]. The seven-dimensional design vector was $\mathbf{x} = [W L \rho \varphi \epsilon_r h]^T$, i.e., all dimensions and parameters were allowed to vary; with individual variable ranges given by $24 \leq W \leq 29$ mm, $32 \leq L \leq 40$ mm, $1 \leq l \leq 13$ mm, $0.8 \leq \rho \leq 5$ mm, $10^\circ \leq \varphi \leq 80^\circ$, $2.2 \leq \epsilon_r \leq 6.15$ mm, $0.5 \leq h \leq 6$ mm.

Data collection and model training and testing was carried out similarly to the Antenna-1 case. Training data consisted of $n = 80$ input–output pairs, while the number of test data points was $n^* = 100$. As before, eight separate Gaussian process models were trained: four $f_{\text{res},1}$ models and four $f_{\text{res},2}$ models, each utilizing one of the covariance functions of (4)–(7); in all cases, a zero mean function was assumed. The total training time for the eight models was less than 2 min. Based on the negative marginal log likelihood, the models utilizing the Matérn covariance function with $\nu = 5/2$ were selected in both the

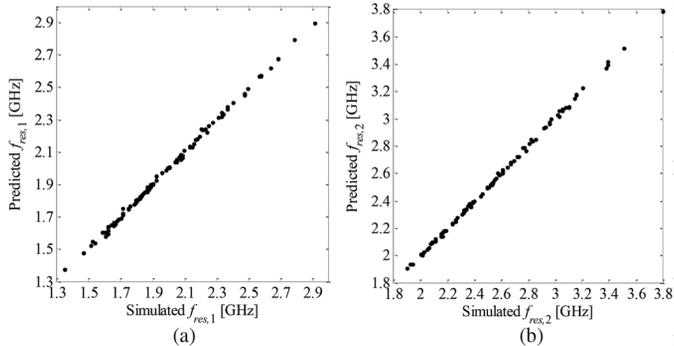


Fig. 2. Scatter plots of test predictions against test target values for Antenna 2: (a) GP model for $f_{\text{res},1}$; (b) GP model for $f_{\text{res},2}$.

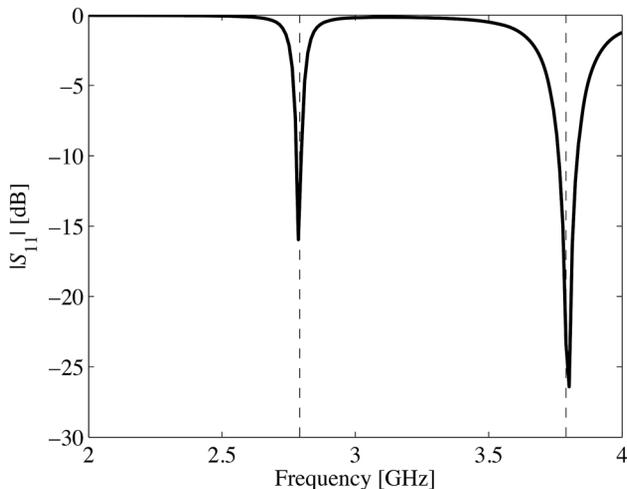


Fig. 3. $|S_{11}|$ against frequency for an Antenna-2 test geometry: $\mathbf{x} = [W \ L \ \rho \ \varphi \ \epsilon_r \ h]^T = [24.82 \text{ mm}, 34.71 \text{ mm}, 2.14 \text{ mm}, 4.03 \text{ mm}, 0.50 \text{ rad}, 2.30 \text{ mm}, 1.23 \text{ mm}]^T$. The vertical dashed lines give the positions of the predicted resonant frequencies (*left*) $f_{\text{res},1}$ and (*right*) $f_{\text{res},2}$.

$f_{\text{res},1}$ and $f_{\text{res},2}$ cases. Table I gives the models' predictive accuracies (percentage root-mean-square errors normalized to the center frequencies of the test target ranges, which were $1.35 \leq f_{\text{res},1} \leq 2.92$ GHz and $1.90 \leq f_{\text{res},2} \leq 3.80$ GHz). Predictive accuracies again were high—especially given the relatively high dimensionality of the input space—with normalized RMSEs $< 0.6\%$ in both cases and correlation coefficients close to one. Fig. 2 gives scatter plots illustrating the close agreement between simulated and predicted test targets for the $f_{\text{res},1}$ and $f_{\text{res},2}$ models; Fig. 3 shows the simulated $|S_{11}|$ -against-frequency response for a specific test geometry and the positions of the predicted resonant frequencies. It may be noted that the variable bounds on ϵ_r and h in conjunction with the target frequency ranges implied operating dielectric substrate heights that varied between about $0.003\lambda_d$ and $0.19\lambda_d$, where λ_d is the wavelength in the dielectric.

IV. AUTOMATIC RELEVANCE DETERMINATION

It was noted in Section II that the optimized length scale parameters $\{\tau_k | k = 1, \dots, D\}$ (obtained by minimizing the negative log marginal likelihood) provide an indication of the relative importance of input (i.e., design) vector features in the sense that the shorter a length scale parameter associated

TABLE II
LENGTH SCALE PARAMETERS FOR GPR MODELS OF $f_{\text{res},1}$ and $f_{\text{res},2}$
(ANTENNA 2)

MODEL	LENGTH SCALE PARAMETERS τ_k						
	τ_1 (W)	τ_2 (L)	τ_3 (l)	τ_4 (ρ)	τ_5 (φ)	τ_6 (ϵ_r)	τ_7 (h)
$f_{\text{res},1}$	144.0	27.42	15.00	118.7	111.9	9.11	11.67
$f_{\text{res},2}$	27.50	64.39	17.53	99.43	64.65	9.13	12.83

¹Input variable corresponding to a particular length scale parameter is indicated in brackets.

with a particular dimension of input space, the more quickly change occurs in the underlying function along that dimension. Table II gives the optimized length scale parameters for each of the $f_{\text{res},1}$ and $f_{\text{res},2}$ models of Antenna 2, along with the corresponding input variables. These numbers indicate that dielectric constant and height were the most influential variables for both models, followed by l and L for the $f_{\text{res},1}$ model, and l and W for $f_{\text{res},2}$. The finding about l , and the importance of L for the $f_{\text{res},1}$ model and W for the $f_{\text{res},2}$ model, is consistent with physics/mode-based explanations of the antenna's resonant frequency behavior [12], [15]. On the other hand, changes in feed position ρ and φ were indicated to have relatively little effect on $f_{\text{res},1}$ and $f_{\text{res},2}$ by comparison.

Inspection of the length scale parameters of the Antenna-1 models revealed that $f_{\text{res},1}$ and $f_{\text{res},2}$ were most sensitive to variations in the slot-length-related dimensions W_s and L_s , followed by b , the separation from the patch edge. Slot width a had relatively little effect.

V. CONCLUSION

The results presented here show that Gaussian process regression, which is much more straightforward to implement than neural networks, is eminently suitable for modeling the resonant frequencies of dual-band microstrip antennas. A particular strength of GPR is that it can give very good predictive results for relatively small training data sets, even in the presence of multiple design variables; this was confirmed by the two presented antenna examples. A unique property inherent to GPR (neural networks do not provide a comparable feature) is automatic relevance determination, which in the present context provided indications of the relative importance of the design variables for the modeled quantities $f_{\text{res},1}$ and $f_{\text{res},2}$. These indications were consistent with physics-based explanations of antenna behavior (of which GPR of course had no explicit knowledge). ARD is a highly desirable feature, particularly when the design space is of high dimensionality and/or when the model user is inexperienced with the antenna in question. Finally, the speed with which predictions can be carried out would render models of this sort suitable for use within microwave computer-aided design (CAD) applications (e.g., [3]).

REFERENCES

- [1] D. Karaboga, K. Guney, S. Sagiroglu, and M. Erler, "Neural computation of resonant frequency of electrically thin and thick rectangular microstrip antennas," *IEE Proc., Microw., Antennas Propag.*, vol. 146, no. 2, pp. 155–159, 1999.

- [2] K. Guney, S. Sagiroglu, and M. Erler, "Generalized neural method to determine resonant frequencies of various microstrip antennas," *Int. J. RF Microw. Comput. Aided Eng.*, vol. 12, pp. 131–139, 2002.
- [3] G. Angiulli and M. Versaci, "Resonant frequency evaluation of microstrip antennas using a neural-fuzzy approach," *IEEE Trans. Magn.*, vol. 39, no. 3, pp. 1333–1336, May 2003.
- [4] K. Guney and N. Sarikaya, "A hybrid method based on combining artificial neural network and fuzzy inference system for simultaneous computation of resonant frequencies of rectangular, circular, and triangular microstrip antennas," *IEEE Trans. Antennas Propag.*, vol. 55, no. 3, pp. 659–668, Mar. 2007.
- [5] T. Khan and A. De, "Estimation of radiation characteristics of different slotted microstrip antennas using a knowledge-based neural networks model," *Int. J. RF Microw. Comput. Aided Eng.*, vol. 24, no. 6, pp. 673–680, 2014.
- [6] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, 2006.
- [7] A. G. Wilson, E. Gilboa, A. Nehorai, and J. P. Cunningham, "GPatt: Fast multidimensional pattern extrapolation with Gaussian processes," arXiv preprint arXiv:1310.5288. 148, 2013.
- [8] J. P. Jacobs and S. Koziel, "Two-stage framework for efficient Gaussian process modeling of antenna input characteristics," *IEEE Trans. Antennas Propag.*, vol. 62, no. 2, pp. 706–713, Feb. 2014.
- [9] J. P. Jacobs and J. P. De Villiers, "Gaussian-process-regression-based design of ultrawide-band and dual-band CPW-fed slot antennas," *J. Electromagn. Waves Appl.*, vol. 24, pp. 1763–1772, 2010.
- [10] K.-F. Lee, S. L. S. Yang, A. A. Kishk, and K. M. Luk, "The Versatile U-Slot Patch Antenna," *IEEE Antennas Propag. Mag.*, vol. 52, no. 1, pp. 71–88, Feb. 2010.
- [11] K.-F. Lee, S. Yang, and A. Kishk, "The versatile U-slot patch antenna," in *Proc. 3rd EuCAP*, 2009, pp. 3312–3314.
- [12] W.-S. Chen, "Single-feed dual-frequency rectangular microstrip antenna with square slot," *Electron. Lett.*, vol. 34, no. 3, pp. 231–232, 1998.
- [13] R. M. Neal, *Bayesian Learning for Neural Networks*. New York, NY, USA: Springer, 1996.
- [14] "FEKO User Manual, Suite 7.0," EM Software & Systems—S. A. (Pty) Ltd., Stellenbosch, South Africa, 2014.
- [15] J.-S. Chen and K.-L. Wong, "A single-layer dual-frequency rectangular microstrip patch antenna using a single probe feed," *Microw. Opt. Technol. Lett.*, vol. 11, no. 2, pp. 83–84, 1996.