Characterisation by Gaussian processes of finite substrate size effects on gain patterns of microstrip antennas

J. Pieter Jacobs

Department of Electrical, Electronic and Computer Engineering, University of Pretoria, Pretoria, South Africa
E-mail: jpjacobs@up.ac.za

Abstract: A procedure is presented for characterising the effects of varying finite substrate/ground plane size on the gain properties of microstrip antennas by means of Gaussian process regression (GPR). Two kinds of microstrip antenna were considered, namely a probe-fed patch antenna on both thin and thick dielectric substrates, and an *L*-probe-fed patch on a thick air substrate. CST Microwave Studio was used to generate training and test data for the GPR models. Frontal *E* and *H*-plane gain patterns could be predicted with normalised root-mean-square errors (RMSEs) of <1.8% for the thin-substrate probe-fed patch and the *L*-probe-fed patch; for the thick-substrate probe-fed patch, RMSEs were 2.1 and 2.8% for the two principal plane gain patterns, respectively. Furthermore, the GPR models could predict patterns at least two orders of magnitude faster than it took to obtain them via direct simulation in CST. Such models are expected to be useful in CAD-based environments for rapidly obtaining estimates of substrate/ground-plane size effects on gain characteristics in lieu of time-consuming full-wave simulations.

1 Introduction

Full-wave electromagnetic simulation (e.g. method of moments, finite element method) is widely used in microwave antenna analysis and design. High-fidelity simulations can characterise antenna performance very accurately, but are usually computationally costly. This fact becomes very important for tasks that require multiple analyses of the antenna under consideration, such as design optimisation and statistical analyses. Noteworthy is the case of global optimisation by means of metaheuristics (e.g. genetic algorithms [1]) that can require many thousands of high-fidelity simulations of candidate geometries of the antenna being designed. In view of these considerations, fast but nevertheless accurate models that serve as surrogates for full-wave simulations of antenna performance characteristics have become an essential part of procedures that involve multiple analyses. After training on suitably selected input–output data samples, properly constructed surrogate models exhibit good generalisation capability, i.e. they permit reliable prediction of antenna behaviour for input variables that were not presented to the model during training.

In the recent past, surrogate modelling of especially the input characteristics of antennas (e.g. S_{11} as a function of geometry dimensions and frequency) has received attention; modelling techniques have included neural networks [2], adaptive neuro-fuzzy inference systems [3], Gaussian process regression (GPR) [4, 5], and Bayesian support vector regression [6]. Higher-level properties of input characteristics have also been modelled. Such studies have dealt with resonant frequency modelling by neural-network-based methods [7, 8] and GPR [9]; resonant input impedance modelling using support vector machines [10]; and impedance bandwidth modelling by neural networks (e.g. [11]). Fewer studies have attempted to model antenna radiation properties, especially radiation patterns. Examples involving microstrip antennas (the focus of the present study) include the modelling of resonant frequency gains (single-valued) as a function of dual-band antenna geometry by means of neural networks [12], and modelling of the co-polar radiation pattern of a circular microstrip antenna by means of adaptive neuro-fuzzy inference systems [3].

In the present work, GPR [13] – which previously was shown to be effective in modelling highly non-linear antenna S_{11} responses [3, 4] – is demonstrated to be an accurate and effective tool for modelling the effects of finite substrate/ground-plane size on the far-field gain characteristics of microstrip antennas, in particular frontal *E*- and *H*-plane gain patterns; front-to-back ratios are also considered. Several studies (e.g. [14–16]) have found significant effects of finite substrate size on antenna radiation patterns – perhaps most notable ripples – and in certain cases on input characteristics [16]. Even so, it appears as if there have been hardly any attempts to formally construct surrogate models that can account for such effects (e.g. none of [2–12] included substrate/ground-plane dimensions as model input variables; in [17] however, a neural network was used to model self- and mutual-admittances of a monopole array as they varied with ground-plane dimensions, amongst other variables).

The emphasis here is on global surrogate models that aim to provide reliable predictions over all of the design space (as opposed to local models that only aim to make accurate predictions in subregions of the design space as required by iterative optimisation algorithms). Application areas of global models include statistical analyses, parametric optimisation, and rapid evaluation within CAD frameworks. Consider for example an antenna engineer who needs to design a microstrip antenna subject to certain surface area constraints. A fast model of the sort developed here would provide the engineer with a swift estimate of the effect on radiation performance of the size of the substrate/ground plane permitted by the currently available real estate, and assess detrimental effects (if any) that would result if the original substrate/ground plane were to be adapted in order to meet externally imposed restrictions on its shape (for instance, it might be required that the length of the substrate along one axis in the plane be longer than the length along the orthogonal axis). To obtain such estimates from a model would be significantly faster than a direct simulation approach; it has been observed that even for time-domain-based solvers (e.g. [18]), the accurate simulation of antennas on finite ground planes can be time-consuming.

The paper is organised as follows. Section 2 briefly describes the theory of GPR based on [13]. Section 3 gives details of the construction and evaluation of GPR models for characterising the effect of finite substrate/ground-plane sizes on the frontal gain patterns of microstrip antennas; gain front-to-back ratios are also

considered. Verification examples involving two antennas are presented, namely a probe-fed microstrip patch antenna on both thin and thick dielectric substrates backed by a ground plane [14], and an L-probe-fed microstrip patch antenna on a thick ground-plane-backed air substrate [19]. Results are given that confirm the soundness of the present modelling approach. Conclusions are presented in Section 4.

2 Background to GPR

Gaussian processes (GPs) are stochastic processes that can be described in terms of a generalisation of the multivariate Gaussian probability distribution [20] to functions. Formally, a GP is a set consisting of an infinite number of random variables of which any finite subset has a joint Gaussian distribution (as follows from the properties of the multivariate Gaussian distribution [13]).

In practice, regression with GPs can be carried out via standard operations performed on multivariate Gaussian distributions. Generally, GPR is significantly easier to apply to modelling problems than artificial neural networks since it requires optimisation of much less parameters than the number of weights in a neural network that typically needs to be learned (in GPR the number of parameters to optimise is of the order of the input vector length).

In the present work, it is of interest to learn a mapping between a set of variables that include the geometry of an antenna and a certain elevation angle (the azimuth angle is assumed to be fixed), and the gain of the antenna in that direction. Hence, the modelling problem requires that a training data set of n input–output pairs, $\{(\boldsymbol{x}_i, y_i)|i = 1, \ldots, n\}$ be set up, where input vectors $\boldsymbol{x}_i$ are comprised of geometry dimensions and an angle, and scalar target responses y_i are gain values. A test data set consisting of n^* input–output pairs $\{(\boldsymbol{x}_i^*, y_i^*)|i = 1, \ldots, n^*\}$ that are not included amongst the training data is compiled in a similar manner.

The GPR framework [13] specifies that the n training outputs (one for each input vector $\boldsymbol{x}_i$) be modelled by random variables $[f_1 \ldots f_n]^{\mathrm{T}} = [f(\boldsymbol{x}_1) \ldots f(\boldsymbol{x}_n)]^{\mathrm{T}}$, and the n^* test outputs by random variables $[f_1^* \ldots f_n^*]^{\mathrm{T}} = [f(\boldsymbol{x}_1^*) \ldots f(\boldsymbol{x}_n^*)]^{\mathrm{T}}$, with $f(\bullet)$ a GP. As noted above, a GP can be interpreted as a generalisation of the Gaussian distribution to functions. This implies a distribution with an infinitely long mean vector $\boldsymbol{m}$ that is represented by a mean function $m(\boldsymbol{x})$; and an infinitely large covariance matrix with entries given by a covariance function $k(\boldsymbol{x}, \boldsymbol{x}')$; $f(\boldsymbol{x})$ can be viewed as a point under this multivariate distribution. The mean function is defined as $m(\boldsymbol{x}) = E[f(\boldsymbol{x})]$; the covariance function gives the covariance between outputs $f(\boldsymbol{x})$ and $f(\boldsymbol{x}')$ in terms of inputs $\boldsymbol{x}$ and $\boldsymbol{x}'$, and is defined as $k(\boldsymbol{x}, \boldsymbol{x}') = E[(f(\boldsymbol{x}) - m(\boldsymbol{x}))(f(\boldsymbol{x}') - m(\boldsymbol{x}'))]$, where $E(X)$ is the expected value of the random variable X [13] [in practice $k(\boldsymbol{x}, \boldsymbol{x}')$ is calculated from covariance functions as discussed below].

The first step in making predictions is to assume a zero-mean Gaussian distribution over the n training outputs and n^* test outputs [13]. This gives the prior distribution

$$\begin{bmatrix} \boldsymbol{f} \\ \boldsymbol{f}^* \end{bmatrix} \sim N\left(\boldsymbol{0}, \begin{bmatrix} \boldsymbol{K}(\boldsymbol{X}, \boldsymbol{X}) & \boldsymbol{K}(\boldsymbol{X}, \boldsymbol{X}^*) \\ \boldsymbol{K}(\boldsymbol{X}^*, \boldsymbol{X}) & \boldsymbol{K}(\boldsymbol{X}^*, \boldsymbol{X}^*) \end{bmatrix}\right) \tag{1}$$

where $N(\boldsymbol{a}, \boldsymbol{b})$ is a Gaussian (normal) distribution with mean vector $\boldsymbol{a}$ and covariance matrix $\boldsymbol{b}$ (the notation indicates that the multivariate point on the left is sampled from the normal distribution on the right); $\boldsymbol{X}$ and $\boldsymbol{X}^*$ are matrices containing the training and test input vectors; and $\boldsymbol{K}(\boldsymbol{X}, \boldsymbol{X}^*)$ is a $n \times n^*$ sub-matrix of covariances calculated between all pairs of training and test outputs (other sub-matrices are defined in a comparable manner). Given that the training outputs $\boldsymbol{y}$ are known, the (multivariate Gaussian) posterior distribution can be computed, i.e. the distribution of the test outputs $\boldsymbol{f}^*$ conditioned on $\boldsymbol{y}$. The mean vector $\boldsymbol{p}$ and covariance matrix Σ of this distribution is given by [13]

$$\boldsymbol{p} = \boldsymbol{K}(\boldsymbol{X}^*, \boldsymbol{X})\boldsymbol{K}(\boldsymbol{X}, \boldsymbol{X})^{-1}\boldsymbol{y} \tag{2}$$

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

Referring to the above antenna modelling problem, $\boldsymbol{p}$ contains the gain predictions, in other words the most likely values of the test outputs associated with the test input vectors. The diagonal of Σ contains the variances of the predictions that are indicative of the confidence in the predictions [13].

Since the covariance functions determine the prior and posterior distributions' covariance matrices, they specify the broad class of GPs that will be favoured by the distributions. In the present work, four standard covariance functions are considered, namely the rational quadratic covariance function, and three Matérn class covariance functions [13]. The rational quadratic is given by the following equation

$$k_{RQ}(\boldsymbol{x}, \boldsymbol{x}') = \sigma_f^2\left(1 + \frac{1}{2\alpha}r^2\right)^{-\alpha} \tag{4}$$

while the three Matérn class covariance functions are

$$k_{\nu=1/2}(\boldsymbol{x}, \boldsymbol{x}') = \sigma_f^2 \exp(-r) \tag{5}$$

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

$$k_{\nu=5/2}(\boldsymbol{x}, \boldsymbol{x}') = \sigma_f^2\left(\frac{1 + \sqrt{5}r + 5r^2}{3}\right)\exp\left(-\sqrt{5}r\right) \tag{7}$$

In (4)–(7)

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

where x_k and x_k' are the kth components of input vectors $\boldsymbol{x}$ and $\boldsymbol{x}'$ in the D-dimensional input space, and $\{\tau_k|k = 1, \ldots, D\}$, σ_f, and α are hyperparameters to be optimised during training. In particular, training involves minimisation of the negative log marginal likelihood (the Gaussian likelihood [13] is used here)

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

where $\boldsymbol{K}$ is the $n \times n$ matrix $\boldsymbol{K}(\boldsymbol{X}, \boldsymbol{X})$. Following training, the size of any hyperparameter τ_k will be inversely proportional to the influence on the regression of component k of the input vectors [cf. (8)]. This property is referred to as automatic relevance determination, and can be viewed as a form of automatic feature selection.

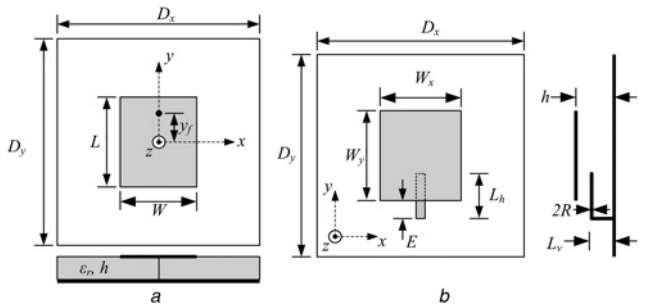

Fig. 1 *Top and side views of antennas*

a Probe-fed microstrip patch antenna on ground-plane-backed single-layer dielectric substrate (Antenna 1)

b L-probe-fed microstrip patch antenna on ground-plane-backed air substrate (Antenna 2)

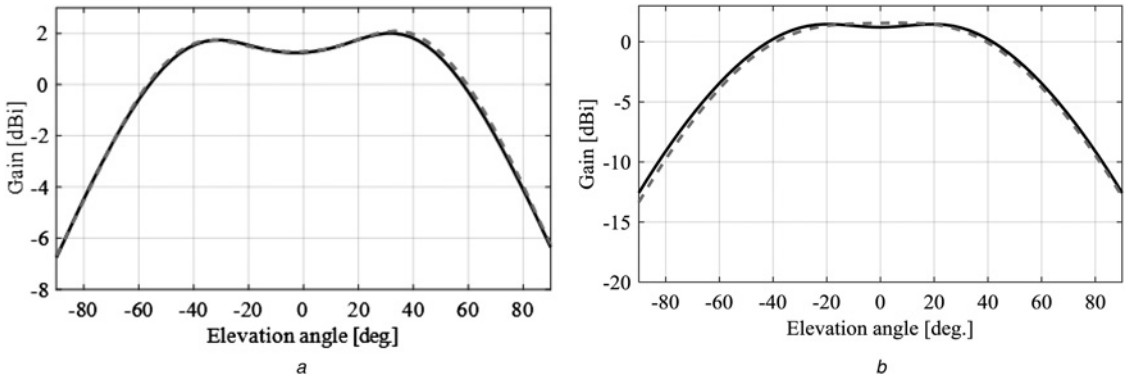

Fig. 2 *Comparison of simulated (——) and predicted (- - -) radiation patterns for Antenna 1 test geometry ($h = 0.028\lambda_d$): $D_x = 83.14$ mm $= 1.39\lambda_0$, $D_y = 118.70$ mm $= 1.99\lambda_0$, $L = 14.13$ mm, and $W = 10.55$ mm*

a *E*-plane pattern
b *H*-plane pattern

3 Implementation and results

3.1 Probe-fed microstrip patch antenna (Antenna 1)

Fig. 1*a* shows the geometry of a probe-fed rectangular patch antenna on a single-layer dielectric substrate backed by a ground plane with the same footprint. The patch length and width was L and W, and the ground plane/dielectric substrate dimensions were D_x and D_y in the x and y directions, respectively. The substrate had a dielectric constant ε_r of 4.34, and a loss tangent of 0.02 [14]. The probe feed was positioned at $y_f = 3.6$ mm. Two substrate heights were considered, representing electrically thin and thick dielectrics, respectively: $h = 0.8$ mm $= 0.028\lambda_d$, and $h = 4.02$ mm $= 0.14\lambda_d$ (λ_d was the wavelength in the dielectric at the operating frequency of 5.02 GHz).

Given a particular substrate thickness, the aim was to establish whether E ($\varphi = 90°$) and H ($\varphi = 0°$)-plane frontal gain patterns (i.e. corresponding to the half-space $z > 0$) could be accurately modelled as a function of varying ground-plane dimensions D_x and D_y, and patch size L and W. Hence, the design vector was $\boldsymbol{u} = [D_x \; D_y \; L \; W]^{\mathrm{T}}$, and the design space was specified by the variable ranges $0.5\lambda_0 \le D_x \le 3\lambda_0$, $0.5\lambda_0 \le D_y \le 3\lambda_0$, 12.6 mm $\le L \le$ 15.4 mm, and 8.64 mm $\le W \le$ 10.56 mm (λ_0 was the free-space wavelength at the operating frequency). Since GPR allows for single model outputs only, separate models had to be set up for the gain in each of the principal planes.

Training data were set up as follows. For each substrate thickness, design vectors $\boldsymbol{u}$ were selected from the design space using latin hypercube sampling (LHS); 70 design vectors were selected in the thin-substrate case, and 100 in the thick substrate case. Given a particular principal plane model, ten elevation angles (i.e. θ values) per geometry were randomly selected from the range $-90° \le \theta \le 90°$ such that elevation angles in general differed from geometry to geometry. This resulted in n training input vectors of the form $\{\boldsymbol{x}_i = [\boldsymbol{u}_i^{\mathcal{T}} \; \theta_i]^{\mathcal{T}} = [D_{xi} D_{yi} L_i W_i \theta_i]^{\mathcal{T}} \mid i = 1, \ldots, n\}$, with $n = 700$ and 1000 for the thin and thick substrate cases, respectively. The corresponding target outputs y_i were gain values in the E or H planes – denoted below as G_E and G_H – that were obtained via simulation at the operating frequency by using the time-domain solver in CST Microwave Studio [18]. Test data for each substrate consisted of 20 new LHS-selected geometries with 61 equally spaced elevation angles per geometry, giving $n^* = 1220$ test data points. (In order to verify the simulation setup, the radiation patterns of a case study [14] involving a probe-fed rectangular patch antenna with $L = 14$ mm and $W = 9.6$ mm on three different sizes of the above thin substrate – $\lambda_0 \times \lambda_0$, $1.5\lambda_0 \times 1.5\lambda_0$, and $3\lambda_0 \times 3\lambda_0$ – were simulated using the CST time-domain solver. The radiation patterns obtained agreed well with the measured patterns in [14, Figs. 10–12].)

Next, four candidate Gaussian process models were trained for each of the G_E and G_H patterns associated with the two substrate thicknesses, each model using a different covariance function from (4)–(7) (i.e. for $h = 0.8$ mm, four models were trained for the G_E pattern and four for the G_H pattern; and likewise for $h = 4.02$ mm). Throughout, a Gaussian likelihood function and a prior with a zero-mean function was assumed. From each group of four models, the model which produced the lowest negative log marginal likelihood (9) was selected as the optimal model, and used to make predictions on the test data. In both the G_E and G_H cases for each substrate thickness, the models employing the rational quadratic covariance function produced the lowest negative log marginal likelihood.

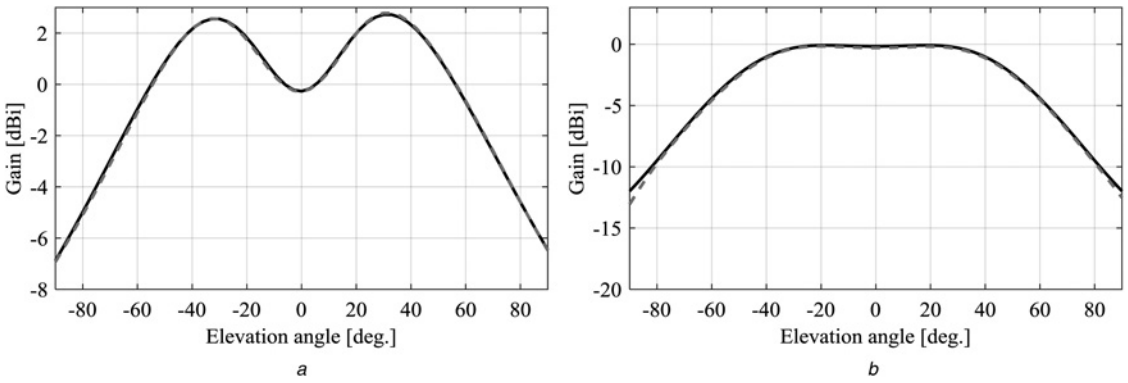

Fig. 3 *Comparison of simulated (——) and predicted (- - -) radiation patterns for Antenna 1 test geometry ($h = 0.028\lambda_d$): $D_x = 114.04$ mm $= 1.91\lambda_0$, $D_y = 132.78$ mm $= 2.22\lambda_0$, $L = 14.39$ mm, and $W = 10.20$ mm*

a *E*-plane pattern
b *H*-plane pattern

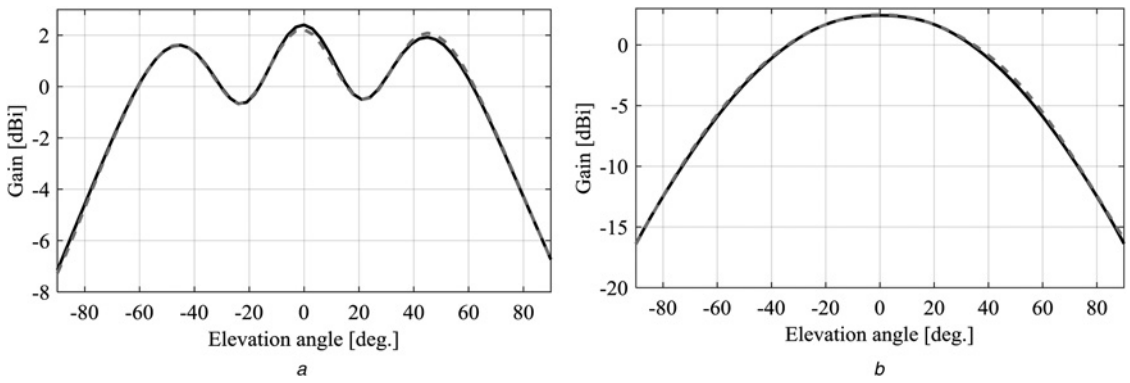

Fig. 4 *Comparison of simulated (——) and predicted (- - -) radiation patterns for Antenna 1 test geometry ($h = 0.028\lambda_d$): $D_x = 164.62$ mm $= 2.76\lambda_0$, $D_y = 161.15$ mm $= 2.70\lambda_0$, $L = 14.92$ mm, and $W = 9.31$ mm*

a *E*-plane pattern
b *H*-plane pattern

Table 1 summarises the models' predictive accuracies in terms of percentage root-mean-square-errors (RMSEs) normalised to the test target ranges, and linear correlation coefficients R (calculated from test targets and predictions). Accuracies were high given the large variations in planar dimensions (and consequently shape) of the substrate, with RMSEs of 1.751 and 1.750% for G_E and G_H, respectively, in the thin-substrate case, and RMSEs of 2.130 and 2.811% in the thick-substrate case. The somewhat lower accuracies obtained in the latter case – viewed in conjunction with the fact that more training points were used compared with the thin-substrate case, as described above – suggest that gain patterns become more difficult to predict as substrate height increases. In this regard it is notable that E-plane patterns as a function of elevation angle exhibit greater asymmetry with increased substrate height, as can be seen from a comparison of Figs. 2–4 with Figs. 5–7 (described next).

Figs. 2–4 give simulated and predicted patterns (E and H planes) for three geometries from the test data set for the thin-substrate case, namely $D_x = 83.14$ mm $= 1.39\lambda_0$, $D_y = 118.70$ mm $= 1.99\lambda_0$, $L = 14.13$ mm, $W = 10.55$ mm; $D_x = 114.04$ mm $= 1.91\lambda_0$, $D_y = 132.78$ mm $= 2.22\lambda_0$, $L = 14.39$ mm, and $W = 10.20$ mm; and $D_x = 164.62$ mm $= 2.76\lambda_0$, $D_y = 161.15$ mm $= 2.70\lambda_0$, $L = 14.92$ mm, and $W = 9.31$ mm. The agreement between simulated and predicted patterns is quite good, and E-plane pattern ripples are well accounted for (recall that none of the above geometry configurations were included amongst the training data). Likewise, results for three thick-substrate test geometries are shown in Figs. 5–7, again confirming the goodness of the models' fit to the data.

A secondary investigative aim was to model gain front-to-back ratio as a function of geometry. In view of the complexity of the variation of back-radiated patterns with geometry and direction, the number of training geometries were increased to 120 for each substrate thickness; as before they were generated via LHS and simulated in CST to obtain the training outputs (i.e. ratios of gains in the directions $\theta = 0°$ and $\theta = 180°$). Test data were the 20 test geometries considered previously and their front-to-back ratios. For each substrate, four Gaussian process models were trained, each using a different covariance function (4)–(7) (input vectors were of the form $\boldsymbol{x}_i = \boldsymbol{u}_i^{\mathrm{T}} = [D_{xi}\ D_{yi}\ L_i\ W_i]^T$). In the thin-substrate case, the Matérn covariance function with $\nu = 5/2$ produced the lowest negative log marginal likelihood and a predictive error of 2.69% ($R = 0.9962$) on the test front-to-back ratios. For the $h = 0.14\lambda_d$ substrate, the rational quadratic covariance function performed best, yielding a front-to-back ratio of 6.5% ($R = 0.9861$). The lower accuracies of front-to-back ratio predictions (compared with frontal pattern predictive accuracies) suggest that front-to-back ratio varies in a highly non-linear manner with antenna and substrate dimensions, and that difficulty of modelling becomes more pronounced as substrate height increases.

Table 1 Predictive errors of GP models on test antenna gain patterns

	Antenna 1				Antenna 2	
	$h = 0.028\lambda_d$		$h = 0.14\lambda_d$		$h = 0.1\lambda_0$	
	G_E	G_H	G_E	G_H	G_E	G_H
RMSE, %	1.751	1.750	2.130	2.811	1.755	1.088
R	0.9977	0.9958	0.9942	0.9920	0.9973	0.9991

3.2 L-probe-fed microstrip patch antenna on thick air substrate (Antenna 2)

The geometry of a rectangular patch antenna fed by an L-probe on a finite ground plane with an air substrate [19] is shown in Fig. 1*b*. The patch dimensions were W_x and W_y, and the ground-plane dimensions were D_x and D_y. The height h of the patch above the ground plane was 7 mm, while the vertical and horizontal segments L_v and L_h of the probe were 5 and 10 mm, respectively. The radius of the probe was $R = 0.5$ mm.

As before, the aim was to construct Gaussian process models of the E- and H-plane frontal ($z > 0$) gain patterns with ground plane and patch dimensions as geometry variables, as well as a model of the front-to-back ratio. The design vector was $\boldsymbol{u} = [D_x\ D_y\ W_x\ W_y]^{\mathrm{T}}$; the design space was defined by the variable ranges $0.75\lambda_0 \le D_x \le 3\lambda_0$, $0.75\lambda_0 \le D_y \le 3\lambda_0$, 27 mm $\le W_x \le$ 33 mm, and 23.4 mm $\le W_y \le$ 28.6 mm, where λ_0 was the free-space wavelength at the 4.3 GHz operating frequency.

Training data input vectors were based on 100 design vectors selected from the design space using LHS, and 8 randomly selected elevation angle values θ per geometry randomly picked from the range $-90° \le \theta \le 90°$ for the G_H model. Given the greater complexity of the G_E patterns (see Fig. 3), ten random θ values per geometry were selected. Hence, the total number of training input vectors of the form $\boldsymbol{x}_i = [\boldsymbol{u}_i^{\mathrm{T}}\ \theta_i]^{\mathrm{T}} = [D_{xi}\ D_{yi}\ W_{xi}\ W_{yi}\ \theta_i]^T$ were 1000 and 800 for the G_E and G_H models, respectively. Training outputs were the corresponding gain values G_E or G_H simulated using the CST time-domain solver. Test data were 25 new LHS-selected geometries with 181 equally spaced elevation angles per geometry ($n^* = 4525$). To verify the simulation setup, a case study [19] involving a rectangular patch antenna with $W_x = 30$ mm, $W_y = 26$ mm, and $E = 2$ mm on a $\lambda_0 \times \lambda_0$ ground plane was simulated using the CST time-domain solver (other geometry variables were the same as described above). Good agreement was observed between the CST radiation patterns and the measured patterns [19, Fig. 3].

As before, four candidate Gaussian process models were trained for each of the G_E and G_H patterns corresponding to the covariance functions (4)–(7). In both the G_E and G_H cases, the rational quadratic covariance function produced the lowest negative log marginal likelihood.

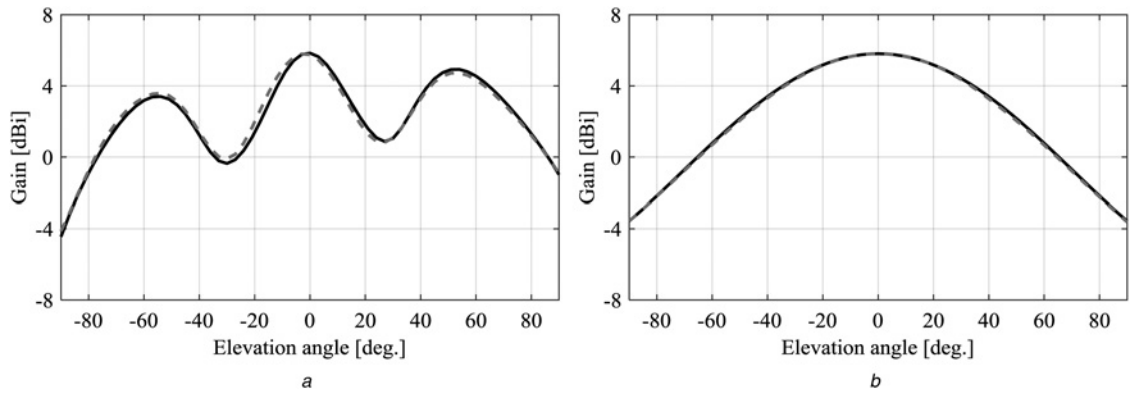

Fig. 5 *Comparison of simulated (——) and predicted (- - -) radiation patterns for Antenna 1 test geometry ($h = 0.14\lambda_d$): $D_x = 147.38$ mm = $2.47\lambda_0$, $D_y = 35.95$ mm = $0.60\lambda_0$, $L = 13.14$ mm, and $W = 8.74$ mm*

a *E*-plane pattern
b *H*-plane pattern

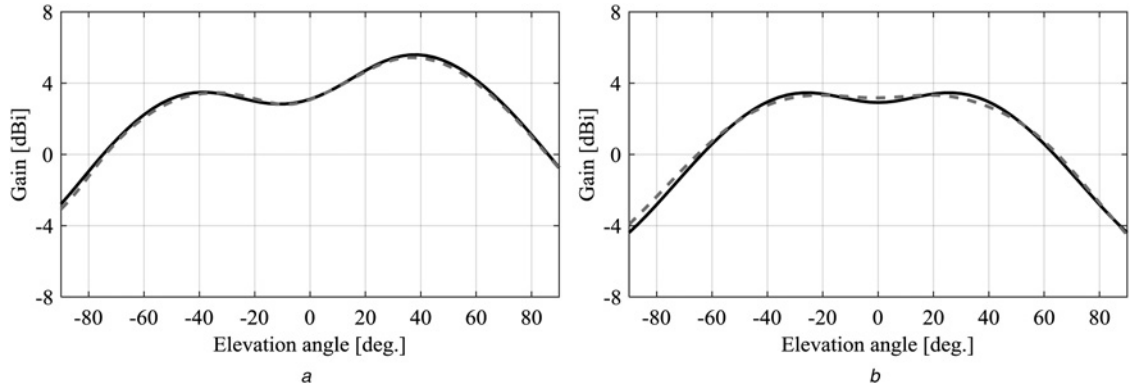

Fig. 6 *Comparison of simulated (——) and predicted (- - -) radiation patterns for Antenna 1 test geometry ($h = 0.14\lambda_d$): $D_x = 73.58$ mm = $1.23\lambda_0$, $D_y = 85.56$ mm = $1.43\lambda_0$, $L = 13.68$ mm, and $W = 9.61$ mm*

a *E*-plane pattern
b *H*-plane pattern

Table 1 presents the models' predictive results. Normalised RMSEs were 1.755 and 1.088% for the G_E and G_H models, respectively, with correlation coefficients of 0.9973 and 0.9991. In Figs. 8–10, CST-simulated and GP-predicted patterns (E and H planes) are given for three test geometries, namely $D_x = 192.63$ mm = $2.76\lambda_0$, $D_y = 206.87$ mm = $2.97\lambda_0$, $W_x = 27.72$ mm, $W_y = 25.39$ mm; $D_x = 203.82$ mm = $2.92\lambda_0$, $D_y = 60.80$ mm = $0.87\lambda_0$, $W_x = 28.90$ mm, and $W_y = 25.26$ mm; and $D_x = 128.89$ mm = $1.85\lambda_0$, $D_y = 88.50$ mm = $1.27\lambda_0$, $W_x = 28.22$ mm, and $W_y = 28.26$ mm. Agreement between simulated and predicted patterns is seen to be good.

For front-to-back ratio modelling, training data were 175 geometries generated by LHS serving as input vectors, with training targets the associated ratios of gains. Test data were the front-to-back ratios of the above 25 test geometries. Four Gaussian process models were trained, each using a different covariance function (4)–(7). The rational quadratic covariance function yielded the lowest negative log

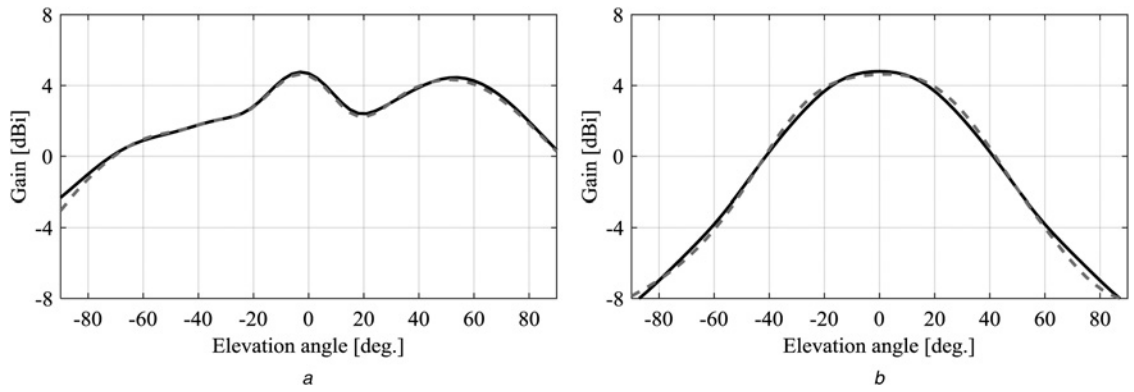

Fig. 7 *Comparison of simulated (——) and predicted (- - -) radiation patterns for Antenna 1 test geometry ($h = 0.14\lambda_d$): $D_x = 124.31$ mm = $2.08\lambda_0$, $D_y = 140.91$ mm = $2.36\lambda_0$, $L = 13.93$ mm, and $W = 9.56$ mm*

a *E*-plane pattern
b *H*-plane pattern

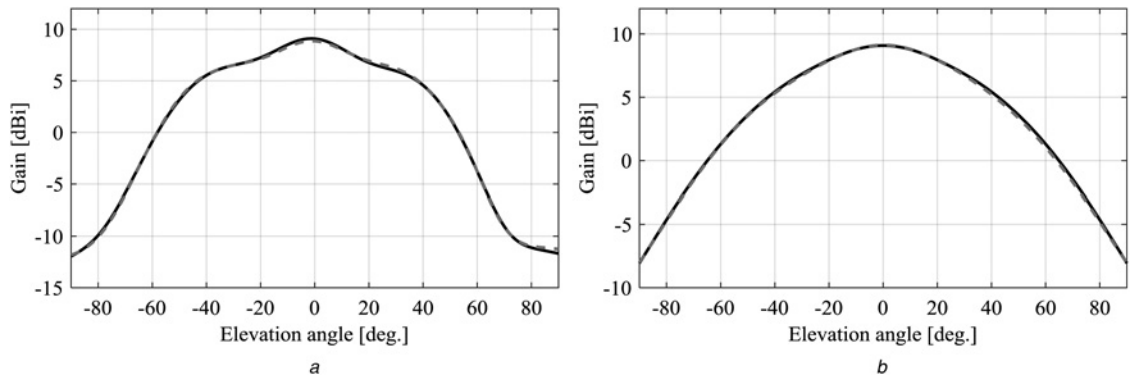

Fig. 8 *Comparison of simulated (——) and predicted (- - -) radiation patterns for Antenna 2 test geometry with $D_x = 192.63$ mm $= 2.76\lambda_0$, $D_y = 206.87$ mm $= 2.97\lambda_0$, $W_x = 27.72$ mm, and $W_y = 25.39$ mm*

a E-plane pattern
b H-plane pattern

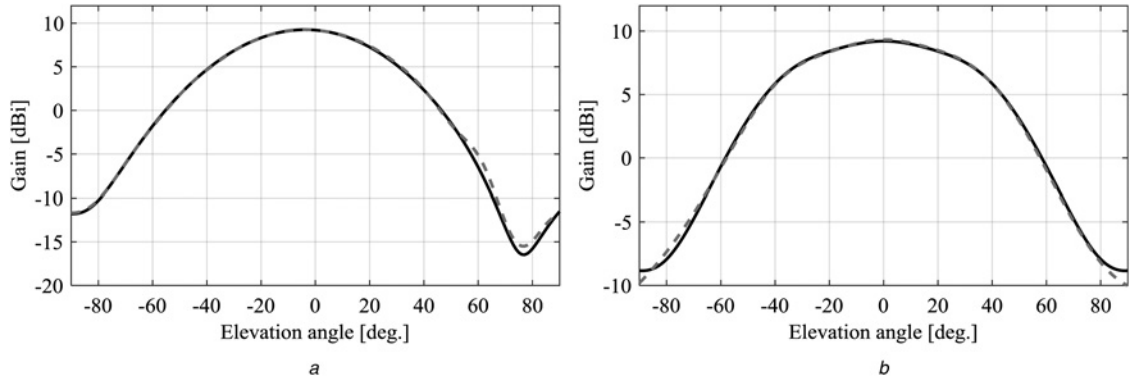

Fig. 9 *Comparison of simulated (——) and predicted (- - -) radiation patterns for Antenna 2 test geometry with $D_x = 203.82$ mm $= 2.92\lambda_0$, $D_y = 60.80$ mm $= 0.87\lambda_0$, $W_x = 28.90$ mm, and $W_y = 25.26$ mm*

a E-plane pattern
b H-plane pattern

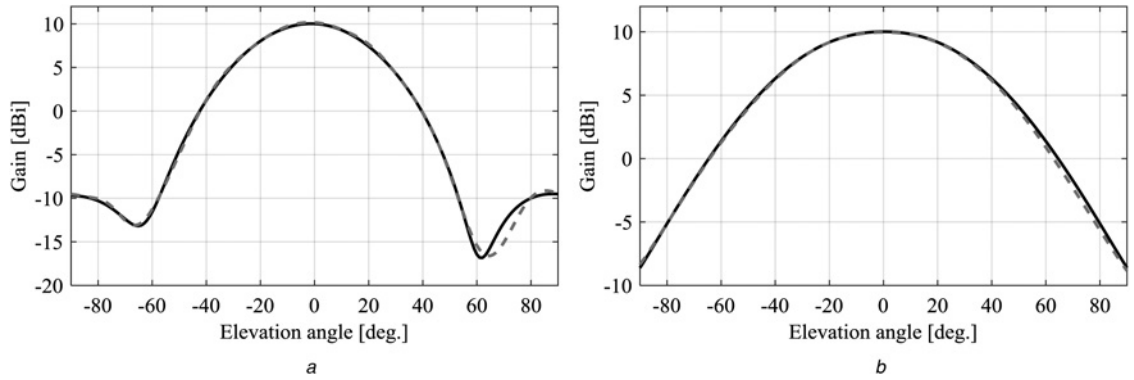

Fig. 10 *Comparison of simulated (——) and predicted (- - -) radiation patterns for Antenna 2 test geometry with $D_x = 128.89$ mm $= 1.85\lambda_0$, $D_y = 88.50$ mm $= 1.27\lambda_0$, $W_x = 28.22$ mm, and $W_y = 28.26$ mm*

a E-plane pattern
b H-plane pattern

marginal likelihood on this data, and a predictive error of 1.724% (R = 0.9976) on the test front-to-back ratios was achieved.

4 Conclusions

A notable deficiency in contemporary antenna modelling literature has been a lack of studies aimed at modelling the effects of finite substrate/ground-plane size on the gain patterns of microstrip antennas, even though the importance of such effects is well known. The present work has shown that GPR can accurately model such effects as manifested in the frontal *E*- and *H*-plane gain patterns of microstrip patch antennas on electrically thin and thick substrates, allowing for substantial variation in the planar dimensions of the substrate/ground plane (from 0.5 to $3\lambda_0$ along each dimension). After training (a once-off process), a Gaussian process model can predict these patterns a great deal faster than they were to be obtained by direct simulation using a full-wave simulator, such as the CST time-domain solver. For example, while CST simulation of test cases such as those represented in

Figs. 2–10 took between 80 s and 7 min each (depending on ground-plane size), the corresponding GP model predictions took a fraction of a second (on a 3.412 GHz Intel Core i7-3770 CPU with 8 GB RAM). Modelling of front-to-back ratios was also considered; best results were achieved for the air-substrate *L*-probe-fed patch, and for the thin-substrate probe-fed patch. It is anticipated that Gaussian process models of the sort described above could be used to good effect in CAD environments where rapid estimates of substrate/ground-plane size on gain patterns are required.

5 References

1 Rahmat-Samii, Y., Michielssen, E.: 'Electromagnetic optimization by genetic algorithms' (John Wiley & Sons, New York, NY, 1999)

2 Kim, Y., Keely, S., Ghosh, J., *et al.*: 'Application of artificial neural networks to broadband antenna design based on a parametric frequency model', *IEEE Trans. Antennas Propag.*, 2007, **55**, (3), pp. 669–674

3 Dadgarnia, D., Heidari, A.A.: 'A fast systematic approach for microstrip antenna design and optimization using ANFIS and GA', *J. Electromagn. Waves Appl.*, 2010, **24**, pp. 2207–2221

4 Jacobs, J.P., De Villiers, J.P.: 'Gaussian-process-regression-based design of ultrawide-band and dual-band CPW-fed slot antennas', *J. Electromagn. Waves Appl.*, 2010, **24**, pp. 1763–1772

5 Jacobs, J.P., Koziel, S.: 'Two-stage framework for efficient Gaussian process modeling of antenna input characteristics', *IEEE Trans. Antennas Propag.*, 2014, **62**, (2), pp. 706–713

6 Jacobs, J.P.: 'Bayesian support vector regression with automatic relevance determination Kernel for modeling of antenna input characteristics', *IEEE Trans. Antennas Propag.*, 2012, **60**, (4), pp. 2114–2118

7 Karaboga, D., Guney, K., Sagiroglu, S., *et al.*: 'Neural computation of resonant frequency of electrically thin and thick rectangular microstrip antennas', *IEE Proc. Microw. Antennas Propag.*, 1999, **146**, (2), pp. 155–159

8 Guney, K., Sagiroglu, S., Erler, M.: 'Generalized neural method to determine resonant frequencies of various microstrip antennas', *Int. J. RF Microw. Comput. Aided Eng.*, 2002, **12**, pp. 131–139

9 Jacobs, J.P.: 'Efficient resonant frequency modeling for dual-band microstrip antennas by Gaussian process regression', *IEEE Antennas Wirel. Propag. Lett.*, 2015, **14**, pp. 337–341

10 Angiulli, G., Cacciola, M., Versaci, M.: 'Microwave devices and antennas modelling by support vector regression machines', *IEEE Trans. Magn.*, 2007, **43**, pp. 1589–1592

11 Sağiroğlu, Ş., Güney, K., Erler, M.: 'Calculation of bandwidth for electrically thin and thick rectangular microstrip antennas with the use of multilayered perceptrons', *Int. J. RF Microw. Comput. Aided Eng.*, 1999, **9**, (3), pp. 277–286

12 Khan, T., De, A.: 'Estimation of radiation characteristics of different slotted microstrip antennas using a knowledge-based neural networks model', *Int. J. RF Microw. Comput. Aided Eng.*, 2004, **24**, pp. 673–680

13 Rasmussen, C.E., Williams, C.K.I.: 'Gaussian processes for machine learning' (MIT Press, Cambridge, MA, 2006)

14 Bokhari, S.A., Mosig, J.R., Gardiol, F.E.: 'Radiation pattern computation of microstrip antennas on finite size ground planes', *IEE Proc. H, Microw. Antennas Propag.*, 1992, **139**, (3), pp. 278–286

15 Huang, J.: 'The finite ground plane effect on the microstrip antenna radiation patterns', *IEEE Trans. Antennas Propag.*, 1983, **31**, (4), pp. 649–653

16 Bhattacharyya, A.K.: 'Effect of finite ground plane on the radiation characteristics of a circular patch antenna', *IEEE Trans. Antennas Propag.*, 1990, **38**, (2), pp. 152–159

17 Yazdanbakhsh, P., Solbach, K.: 'Optimization of monopole four-square array antenna using a decoupling network and a neural network to model ground plane effects'. Proc. Third European Conf. on Antennas and Propagation (EuCAP), March 2009, pp. 3014–3018

18 CST Microwave Studio, ver. 2015, CST AG, Bad Nauheimer Str. 19, D-64289 Darmstadt, Germany, 2015

19 Guo, Y.-X., Luk, K.-M., Lee, K.-F.: 'L-probe fed thick-substrate patch antenna mounted on a finite ground plane', *IEEE Trans. Antennas Propag.*, 2003, **51**, (8), pp. 1955–1963

20 Johnson, R.A., Wichern, D.W.: 'Applied multivariate statistical analysis' (Pearson, Upper Saddle River, NJ, 2007, 6th edn.)