\ECCOMAS Congress 2016
VII European Congress on Computational Methods in Applied Sciences and Engineering
M. Papadrakakis, V. Papadopoulos, G. Stefanou, V. Plevris (eds.)
Crete Island, Greece, 5–10 June 2016

# SURROGATE MODELING OF ANTENNA RADIATION CHARACTERISTICS BY GAUSSIAN PROCESSES

**J. Pieter Jacobs[1], and Dirk I.L. de Villiers[2]**

[1] Department of Electrical, Electronic and Computer Engineering
University of Pretoria, Pretoria, 0002, South Africa
e-mail: jpjacobs@up.ac.za

[2] Department of Electrical and Electronic Engineering
Stellenbosch University, Stellenbosch, South Africa
ddv@sun.ac.za

**Keywords:** antennas, Gaussian processes, modeling.

**Abstract.** *An overview is given on recent developments in the modeling of radiation properties of antennas by means of Gaussian process (GP) regression. Two widely differing applications are considered, namely modeling of the (possibly detrimental) effects of using a finite substrate/ground plane on the gain pattern of a microstrip patch antenna; and characterizing the highly oscillatory aperture efficiency responses of an offset Gregorian reflector system. It is shown that GP modeling performs well with respect to both problems, with interesting generalization behaviour observed in the case of the reflector system.*

## 1 INTRODUCTION

Gaussian process (GP) regression [1] has in recent years been shown to be suitable for the modelling of highly nonlinear input characteristics of microwave antennas, notably input reflection coefficient as a function of design variables and frequency [2]. The robustness of these models has been verified through incorporating them into optimization frameworks that included space mapping [3] and genetic algorithms [4]**.**

A Gaussian process is a stochastic process that can be viewed as a generalization of the Gaussian probability distribution to functions, and is characterized in full by its mean and covariance functions. During GP regression, the calculations required for learning and inference can be carried out using standard Gaussian probability distribution mathematics. Advantages over neural networks – a widely used modeling approach in electromagnetics – include easier implementation and interpretation since the optimization of far fewer parameters (in the order of the dimension of the input vectors) is required compared to the number of weights in a multi-layer perceptron. Furthermore, GP regression can automatically take into account the relative importance of input features when making predictions (a degree of flexibility not available under kernel methods such as standard support vector regression).

The present paper aims to give an overview of developments in the relatively recent area of GP modeling of antenna radiation characteristics. Two problem types – significantly different in nature – are considered. In the first instance, GP regression is applied to modeling the effects of finite substrate/ground plane size on the gain patterns of a microstrip patch antenna (finite substrate effects are often of practical interest when dealing with microstrip antennas). The operating frequency is fixed, and gain is predicted as a function of spatial direction. It is found that GP regression employing a suitable covariance function is well suited to this task, and important ripples that arise in principal plane patterns can be accurately accounted for. The second problem involves modeling the aperture efficiency responses of offset Gregorian reflector systems against frequency. These responses are of a quasi-periodic nature and highly oscillatory, and it would be expected that many training points would be needed to yield models with good predictive accuracy. It is shown that GP regression using a suitably constructed composite covariance function can significantly lessen the expected training data burden. Furthermore, it is demonstrated that a disproportionally large part of the training data can be concentrated in a section of the input space where data are relatively inexpensive to obtain (i.e., at lower frequencies). The resulting GP model can capture the underlying response structure to such an extent that accurate predictions can be made in (higher-frequency) sections of the response containing multiple adjacent cycles and no training points.

## 2 BACKGROUND TO MODELING WITH GAUSSIAN PROCESSES

### 2.1 Standard Gaussian Process Regression

A Gaussian process (GP) is a set that consists of an infinite number of random variables of which any finite subset has a jointly Gaussian distribution; it can be viewed as a distribution over functions. It is a natural extension of the jointly Gaussian distribution to the case where the mean vector is infinitely long and the covariance matrix is of infinite dimension. Notationally, $f(\mathbf{x}) \sim GP(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}'))$, where $f(\mathbf{x})$ is a sample function, $\mathbf{x}$ and $\mathbf{x}'$ are positions in $R^D$ space, and $m(\mathbf{x})$ and $k(\mathbf{x}, \mathbf{x}')$ are mean and covariance functions defining the GP [1].

Hyperparameters determine the properties of the mean and covariance functions (this is the only parameterisation that is involved). The covariance function $k(\mathbf{x}, \mathbf{x}')$ gives the covariance between the random output variables $f(\mathbf{x})$ and $f(\mathbf{x}')$. Standard covariance functions include the squared-exponential (SE) covariance function

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

and the rational quadratic covariance function [1]:

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

In (1)—(2),

$$r=\left(\sum_{k=1}^{D}\frac{(x_k-x'_k)^2}{\tau_k^2}\right)^{1/2} \tag{3}$$

where $x_k$ and $x'_k$ are the $k$th components of input vectors $\mathbf{x}$ and $\mathbf{x}'$ of dimension $D$, and $\{\tau_k| k = 1, ..., D\}$ are indicative of how quickly change occurs along the corresponding dimensions of the input space. $\sigma_f^2$ governs the overall variance of the process and $\alpha$ is a scaling parameter. $\sigma_f^2$, $\alpha$, and $\{\tau_k| k = 1, ..., D\}$ are referred to as hyperparameters; they can be optimized through a process which finds the hyperparameters for which the negative log marginal likelihood is a minimum. The log marginal likelihood is given by [1]

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

where $K = (X, X)$ is the $n\times n$ matrix of covariances evaluated between all possible pairs of $n$ training outputs using the covariance function, $X$ is the $D\times n$ matrix of training input vectors $\mathbf{x}_i$, $|K|$ is the determinant of $K$, and $\mathbf{y}$ is the training target vector.

Making predictions requires constructing the *prior* distribution, which is a jointly Gaussian distribution over the $n$ random variables that represent the training outputs (in column vector $\mathbf{f}$), and the $n_*$ random variables representing the test outputs (in $\mathbf{f}_*$); in this work the distribution is assumed to have zero mean:

$$\begin{bmatrix}\mathbf{f}\\ \mathbf{f}_*\end{bmatrix}\sim N\left(\mathbf{0},\begin{bmatrix}K(X,X) & K(X,X_*)\\ K(X_*,X) & K(X_*,X_*)\end{bmatrix}\right) \tag{5}$$

In (5), $N(\mathbf{a},\mathbf{b})$ denotes a multivariate normal distribution with mean vector $\mathbf{a}$ and covariance matrix $\mathbf{b}$; $K(X,X_*)$ is the $n\times n_*$ matrix of covariances evaluated between all possible pairs of $n$ training and $n_*$ test outputs; and $X_*$ is a matrix containing the test input vectors (other submatrices have similar definitions). The *posterior* distribution (also multivariate Gaussian) is then obtained by conditioning the test outputs on the known training outputs $\mathbf{y}$. Hence $\mathbf{f}_*\,|\,X_*,X,\mathbf{y}\sim N(\mathbf{m},\Sigma)$, with mean vector $\mathbf{m}$ and covariance matrix $\Sigma$ [1]

$$\mathbf{m}=K(X_*,X)K(X,X)^{-1}\mathbf{y} \tag{6}$$

$$\Sigma=K(X_*,X_*)-K(X_*,X)K(X,X)^{-1}K(X,X_*) \tag{7}$$

The mean vector $\mathbf{m}$ contains the most likely output predictions associated with the test input vectors in $X_*$, while the diagonal of the covariance matrix $\Sigma$ gives the predictive variances.

## 2.2 Gaussian Process Modeling of Quasi-Periodic Responses

Standard covariance functions, e.g., (1)—(2), are not sufficiently expressive to readily model highly oscillatory responses of the sort considered in Section 3.2. However, it was demonstrated [5] that a composite covariance function $k_C(x,x')$ defined as

$$k_C(x,x') = k_{SE}(x,x') \times k_{PER}(x,x') \tag{8}$$

had the requisite flexibility. In (8) – where a one-dimensional input space is assumed – $k_{SE}(x,x')$ is the squared-exponential function (1), and $k_{PER}(x,x')$ is a periodic covariation function, suitable for modeling functions that are comprised of exact repetitions of a basic function [1]:

$$k_{PER}(x,x') = \sigma_f^2 \exp\left( -\frac{2\sin^2[\pi|x-x'|/\lambda]}{\theta^2} \right) \tag{9}$$

In (9), the hyperparameter $\lambda$ determines the intervals between repetitions, $\theta$ is a length-scale parameter, and $\sigma_f^2$ is defined as above.

## 3 MODELING EXAMPLES

### 3.1 Modeling of finite substrate size effects on gain patterns of microstrip patch antennas

Fig. 1 illustrates the geometry of a probe-fed rectangular patch antenna on a single-layer dielectric substrate backed by a ground plane of the same planar footprint as the dielectric substrate. The patch had length $L$ and width $W$, and the planar dimensions of the dielectric substrate/ground plane were $D_x$ and $D_y$ in the $x$ and $y$ directions respectively. The substrate height was fixed at $h$ = 0.8 mm, its dielectric constant $\varepsilon_r$ was 4.34, and its loss tangent was 0.02. The probe feed was positioned at $y_f$ = 3.6 mm.

The aim was to establish whether $E$ ($\varphi$ = 90°) and $H$ ($\varphi$ = 0°) plane frontal gain patterns (corresponding to the half-space $z$ > 0) could be accurately modeled as a function of four variables: substrate/ground plane dimensions $D_x$ and $D_y$, and patch size $L$ and $W$; hence the design vector was $\mathbf{u} = [D_x\, D_y\, L\, W]^T$. The design space was specified by the variable ranges $0.5\lambda_0 \leq D_x \leq 3\lambda_0$, $0.5\lambda_0 \leq D_y \leq 3\lambda_0$, 12.6 mm $\leq L \leq$ 15.4 mm, and 8.64 mm $\leq W \leq$ 10.56 mm ($\lambda_0$ was the free-space wavelength at the operating frequency of 5.02 GHz). Since GP regression allows for single model outputs only, it was necessary to set up separate models for the gain in each of the principal planes. Of particular interest was the GP models' ability to predict the pattern distortion that is well known to occur for finite substrate/ground planes (as opposed to the ideal case of the patch being supported by an infinite substrate).

In order to construct a training data set, 70 design vectors $\boldsymbol{u}$ were selected from the design space using Latin hypercube sampling (LHS). For each principal plane model, 10 elevation angles (i.e., $\theta$ values) per geometry were randomly selected from the range -90° $\leq \theta \leq$ 90° such that elevation angles in general differed from geometry to geometry. This yielded $n$ = 700 training input vectors $\{\mathbf{x}_i = [\mathbf{u}_i^T\ \theta_i]^T = [D_{xi}\ D_{yi}\ L_i\ W_i\ \theta_i]^T \mid i = 1,\ldots,n\}$. The corresponding target outputs $y_i$ were gain values in the $E$ or $H$ plane – denoted below as $G_E$ and $G_H$; these were obtained by means of simulation with the time-domain solver in CST Microwave Studio [7]. Test data consisted of 20 new LHS-selected geometries with 61 equally spaced elevation angles per geometry, resulting in $n^*$ = 1220 test data points. (In order to verify the simulation setup, the radiation patterns of a related case study [8] were reproduced using the CST solver.)

Next, four candidate Gaussian process models were trained for each of the $G_E$ and $G_H$ patterns, each model using a different covariance function that included the squared exponential (1), rational quadratic (2), and two Matérn type covariance functions [1]. A Gaussian likelihood function and a prior with a zero mean function was assumed throughout. From each set of four models, the model which produced the lowest negative log marginal likelihood (4) was selected as the optimal model, and used to make predictions on the test data. In both the

$G_E$ and $G_H$ cases, the rational quadratic covariance function (2) models produced the lowest negative log marginal likelihood.

The models' predictive accuracies are summarized in Table 1 in terms of percentage root-mean-square-errors (RMSEs) normalized to the test target ranges, and linear correlation coefficients $R$ [6]. In spite of the significant variations in planar dimensions and consequently shape of the substrate, high accuracies were obtained, with RMSEs of about 1.75% for each of $G_E$ and $G_H$. Figure 2 gives simulated and predicted patterns ($E$ and $H$ plane) for a geometry from the test data set, namely $D_x$ = 2.57$\lambda_0$, $D_y$ = 1.85$\lambda_0$, $L$ = 13.52 mm, $W$ = 9.16 mm. The agreement between simulated and predicted patterns is good, and the $E$-plane pattern ripples are well represented.

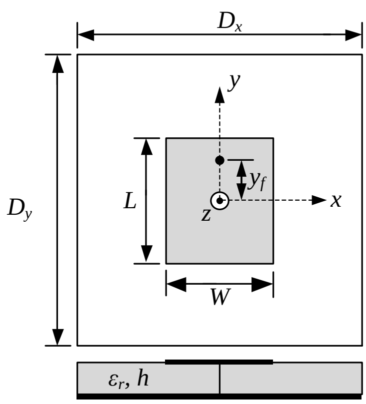
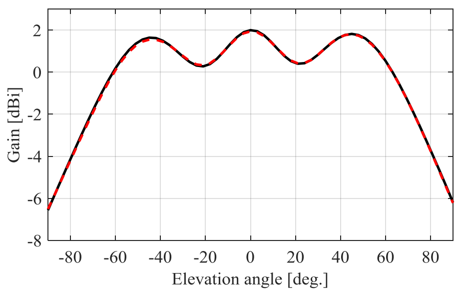


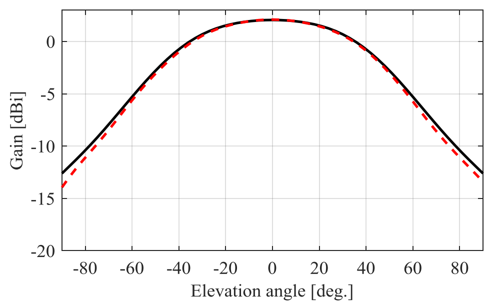

Figure 1: Probe-fed microstrip patch antenna on ground-plane-backed single-layer dielectric substrate.

| | $G_E$ ($n$ = 700, $n^*$ = 1220) | $G_H$ ($n$ = 700, $n^*$ = 1220) |
|---|---|---|
| %RMSE | 1.751 | 1.750 |
| $R$ | 0.9977 | 0.9958 |

Table 1. Predictive errors of GP models for gain patterns on test data ($n$: number of training points; $n^*$: number of test points)




(a)




(b)

Figure 2: Comparison of simulated (——) and predicted (- - -) radiation patterns for microstrip patch antenna test geometry with $D_x$ = 2.57$\lambda_0$, $D_y$ = 1.85$\lambda_0$, $L$ = 13.52 mm, $W$ = 9.16 mm.
(a) $E$-plane pattern (b) $H$-plane pattern.

### 3.2 Modeling of aperture efficiency ripple in reflector antennas

Consider the offset Gregorian antenna with geometry shown in Fig. 3, which may be designed using the equations given in [9].

Using physical optics and physical theory of diffraction in the commercial code GRASP [6], we simulated the aperture efficiency of a system with $D_s$ = 3.8 m, $D_m$ = 15.5 m, $d_g$ = 1.0 m, $\theta_e$ = 50°, fed by an ideal Gaussian feed pattern with a -10 dB edge taper at $\theta_e$, over the frequency range 0.7325 GHz to 1.5475 GHz. The efficiency responses, as function of frequency, typically display rapidly varying oscillations modulated by a slowly varying envelope. For example, the response in Fig. 4 shows four 'periods' of the slowly-varying envelope. The oscillatory response of the aperture efficiency (or directivity anywhere in the vicinity of the main beam) of the system is caused by interference between the direct reflected rays (shown in Fig. 3 at the edges of the reflectors as $\rho_{S1,2} + \rho_{M1,2} + \rho_{A1,2}$) and the diffracted rays from the sub-reflector (shown in Fig. 3 at the edges of the reflectors as $\rho_{D1,2} + \rho_{D1,2}$). The period of the oscillation is dependent on the specific geometry and separation between the dishes. A full description of this mechanism is provided in [9], [10].

A training data set was generated by selecting 28 equally-spaced points within the interval spanning the first 'period' of the slowly-varying envelope, i.e., 0.7325 GHz $\leq f \leq$ 0.95 GHz (Interval 1), as well as 26 points randomly selected from the interval spanning the remaining three periods, i.e., 0.95 GHz $< f \leq$ 1.5475 GHz (Interval 2). We postulated that thoroughly describing the first period by means of training points would enable the regression – via use of the composite covariance function (8) – to capture most of the essential rapidly varying structure of the response and that fewer data points would therefore be necessary for subsequent repetitions along the remainder of the input space (i.e., frequency).

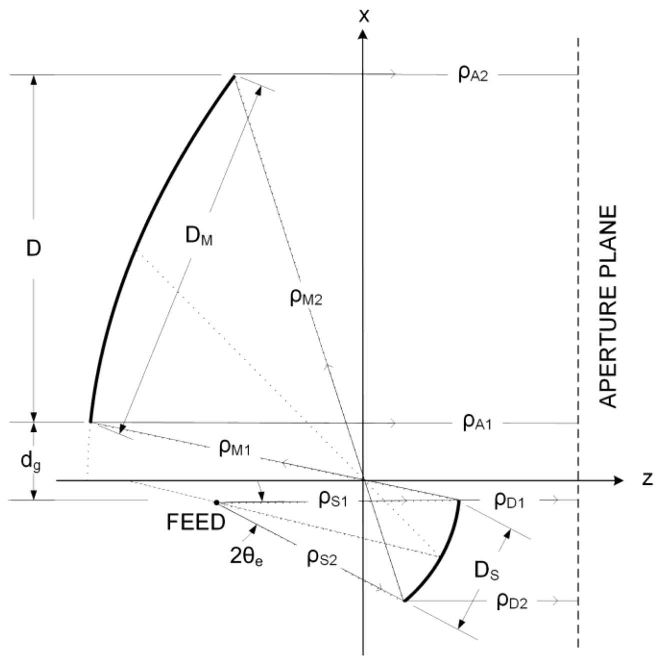



Figure 3. Symmetry plane cut view of a typical offset Gregorian reflector system.

Fig. 4 shows predictive results for GP models trained as described above. In spite of the generally non-uniform distribution of points in Interval 2, with significant stretches of the responses receiving minimal training points in some cases (e.g., the third period in Fig. 4), predictive results were excellent, with normalized root-mean-square-errors (RMSEs) of 0.74%

achieved (RMSEs were normalized to the target range). Similar results were also found for a variety of other distributions of training points, with more details given in [5].

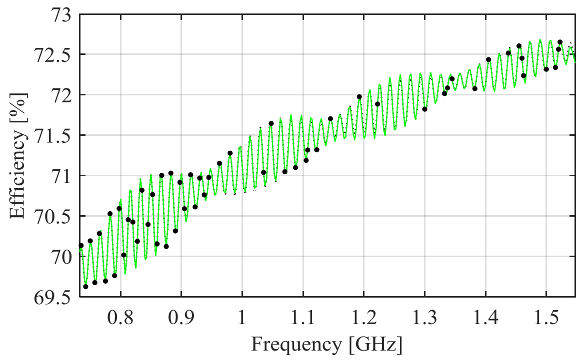



Figure 4. Target (- - -) and predicted responses (—), and training data points (•) for a randomly selected training data set. There were 28 training data points in the interval 0.7325 GHz ≤ $f$ ≤ 0.95 GHz, and 26 randomly selected training data points in the interval 0.95 GHz ≤ $f$ ≤ 1.5475 GHz. The test prediction normalized root-mean-square error was 0.74%.

## 4 CONCLUSIONS

GP regression has been shown to yield very accurate predictions for two diverse antenna modeling problems, namely gain patterns of a microstrip patch antenna subject to finite substrate/ground plane effects; and highly oscillatory aperture efficiency responses of an offset Gregorian reflector system. In the reflector system case, computations were extremely efficient as far as training data requirements are concerned. With respect to the reflector system, results are shown only for a 1-D case here, but work is ongoing on extending the method to larger dimensional spaces, and testing it on more reflector system examples. Applications may include global optimization as well as sensitivity analysis with respect to variations in the mechanical structure of the system – an important problem in high performance systems such as radio telescopes and space antennas which are required to operate in very harsh environmental conditions.

## REFERENCES

[1] C. E. Rasmussen, C. K. I. Williams, *Gaussian Processes for Machine Learning.* MIT Press, 2006.

[2] J. P. De Villiers, J. P. Jacobs, Gaussian process modeling of CPW-fed slot antennas. *Progress in Electromagnetics Research,* **98**, 233-249, 2009.

[3] J.P. Jacobs, S. Koziel, Two-Stage Framework for Efficient Gaussian Process Modeling of Antenna Input Characteristics. *IEEE Transactions on Antennas and Propagation*, **62**, 706-713, 2014.

[4] J. P. Jacobs, J. P. De Villiers, Gaussian-process-regression-based design of ultrawide-band and dual-band CPW-fed slot antennas. *Journal of Electromagnetic Waves and Applications,* **24**, 1763-1772, 2010.

[5] J.P. Jacobs, D.I.L. de Villiers, Gaussian process modeling of aperture efficiency ripple in reflector antennas. *Loughborough Antennas & Propagation Conference (LAPC 2015)*, Loughborough, UK, November 2015.

[6] J.P. Jacobs, Characterization by Gaussian processes of finite substrate size effects on radiation patterns of microstrip antennas. Submitted for publication.

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

[8] Bokhari, S.A., Mosig, J.R., and Gardiol, F.E, Radiation pattern computation of microstrip antennas on finite size ground planes. *IEE Proceedings H: Microwaves, Antennas and Propagation*, **139**, 278 – 286, 1992.

[9] C. Granet, Designing classical offset Cassegrain or Gregorian dualreflector antennas from combinations of prescribed geometric parameters. *IEEE Antennas and Propagation Magazine*, **44**, 114–123, 2002.

[10] D.I.L. de Villiers, Prediction of Aperture Efficiency Ripple in Clear Aperture Offset Gregorian Antennas. *IEEE Transactions on Antennas and Propagation*, **61**, 2457 - 2465, 2013.

[11] TICRA, GRASP10, Version 10.0.3, Copenhagen, Denmark. [Online]. Available: http://www.ticra.com.

[12] D.I.L. de Villiers, Gain ripple in small offset Gregorian antennas. *IEEE Int. Symp. Antennas Propag. (APSURSI 2011)*, July 2011.