الجـــمـــهـــوريـــــة الجـزائـريـــة الديمـــقـراطيـــة الـشـــعـبـيـــة

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

وزارة التـــعـــليـــم العـــــــالي والبـــــحث العـــــلـــمـــــي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

جـــــــــــــــــــامعة سعيدة – د. الطاهر مولاي –

Université de Saïda – Dr. Moulay Tahar –
Faculté des Mathématiques, de l'Informatique et des Télécommunications
Département de Télécommunications

MEMOIRE

Présenté pour l'obtention du **Diplôme** de **MASTER en** Télécommunications

Spécialité : Réseaux et Télécommunications

Par : M. Messadi Zakaria Abdelhakim
M. Remmas Redouane Kadda

Inverse prediction of a patch antenna parameters using Gaussian Process

Soutenu, le 15/06/ 2025, devant le jury composé de :

M. Bouhmidi Rachid	MCA	Président
M. Chetioui Mohammed	Prof	Examinateur
M. Bouras Bouhafs	MCA	Rapporteur

2025 / 2026

Dedication

To my beloved parents,

whose unconditional love, sacrifices and prayers have lit every step of my journey;

To my brothers and sisters,

for their endless support and encouragement throughout my studies;

To my dearest friends and colleagues,

with whom I have shared the joys and the difficulties of these academic years;

To my teachers, from primary school to university,

who have generously transmitted their knowledge and lit the path of curiosity;

To everyone who, near or far, has contributed to making this modest work possible.

I dedicate this work.

Acknowledgments

First and foremost, I would like to express my deepest gratitude to Allah Almighty, for granting me the health, patience and determination needed to bring this work to completion.

I extend my sincere thanks and profound recognition to my supervisor, Dr Bouras Bouhafs., for the trust he placed in me, for the precious guidance, the constructive criticism and the constant availability throughout the realization of this thesis. His scientific rigour and human qualities have been a permanent source of motivation.

I also extend my respectful thanks to the members of the jury, Dr. Bouhmidi Rachid. and Pr. Chetioui Mohammed., who honor me by accepting to evaluate this work. The time they devoted to reading and discussing this manuscript is greatly appreciated.

My acknowledgements go as well to all the teachers and the administrative staff of the Department of Telecommunications, who have accompanied me along these years of study and contributed to my academic and personal formation.

I am also grateful to the laboratory engineers and to my colleagues for the friendly working atmosphere, the fruitful exchanges and the help that was never refused.

Finally, I cannot end these acknowledgements without warmly thanking my family — my parents, my brothers and my sisters — whose unwavering support, patience and encouragements have been the most precious of resources during this whole journey.

To all those who, by a word, a gesture or simply a smile, have contributed near or far to the realization of this work.

Thank you.

Abstract

The rapid expansion of modern wireless communication systems — including 5G, the Internet of Things (IoT) and short-range radar applications — has reinforced the demand for compact, low-profile and easily-integrable antennas. Among the candidate technologies, the microstrip patch antenna remains one of the most attractive options because of its planar geometry, light weight and low fabrication cost. However, the design of such antennas requires accurate full-wave electromagnetic simulations, whose computational cost rapidly becomes prohibitive when several hundred candidate geometries must be evaluated within an optimization loop.

The present thesis investigates the use of neural-network surrogate models to drastically accelerate the design of an inset-fed microstrip patch antenna intended to operate around 6 GHz. The work is built upon a complete and reproducible numerical pipeline: (i) the antenna is modelled in Ansys HFSS using the Finite Element Method (FEM); (ii) a Latin Hypercube Sampling (LHS) plan explores the six-dimensional design space (L, W, h, ε_r, y_0, g); (iii) the resulting simulations feed two families of surrogate models — a Multilayer Perceptron (MLP) and a Gaussian process — that learn the mapping from geometry to the three key performance metrics fr, $|S_{11}|$min and BW.

A total of 150 LHS designs were generated and all of them successfully converged in HFSS, providing a complete and well-balanced dataset of 150 samples. The dataset was split into 120 designs for training and 30 designs for validation. A composite cost function combining a quadratic frequency penalty, a one-sided matching constraint and a bandwidth reward has been defined to drive the surrogate-assisted optimization. The methodology established in the present manuscript thus offers a generic, automated and reproducible blueprint that can be transposed to any patch antenna design problem, paving the way to the very low-cost surrogate-assisted optimization campaigns of the next chapter.

Keywords: *microstrip patch antenna, inset feed, neural networks, surrogate model, Gaussian process, HFSS, Finite Element Method, Latin Hypercube Sampling, 6 GHz design.*

Résumé

Le développement rapide des systèmes de communication sans fil modernes — 5G, Internet des Objets (IoT) et applications radar à courte portée — a renforcé la demande d'antennes compactes, à faible profil et facilement intégrables. Parmi les technologies candidates, l'antenne patch microruban reste l'une des options les plus attractives en raison de sa géométrie planaire, de sa légèreté et de son faible coût de fabrication. Néanmoins, la conception de telles antennes nécessite des simulations électromagnétiques pleines ondes précises, dont le coût de calcul devient rapidement prohibitif lorsque plusieurs centaines de géométries doivent être évaluées dans une boucle d'optimisation.

Le présent mémoire étudie l'utilisation de modèles substituts à base de réseaux de neurones pour accélérer drastiquement la conception d'une antenne patch microruban à alimentation par inset, destinée à fonctionner autour de 6 GHz. Le travail repose sur une chaîne numérique complète et reproductible : (i) l'antenne est modélisée dans Ansys HFSS via la méthode des éléments finis (FEM) ; (ii) un plan d'échantillonnage par hypercube latin (LHS) explore l'espace de conception à six dimensions (L, W, h, ε_r, y_0, g) ; (iii) les simulations obtenues alimentent deux familles de modèles substituts — un perceptron multicouche (MLP) et un processus gaussien — qui apprennent la correspondance entre la géométrie et les trois métriques de performance (fr, $|S_{11}|$min, BW).

Un plan LHS de 150 configurations a été généré et toutes les simulations ont convergé avec succès dans HFSS, fournissant une base de données complète et équilibrée de 150 échantillons, répartis en 120 designs pour l'entraînement et 30 pour la validation. Une fonction de coût composite combinant une pénalité quadratique de fréquence, une contrainte d'adaptation à un seul côté et une récompense de bande passante a été définie pour piloter l'optimisation assistée par modèle substitut. La méthodologie établie offre ainsi un cadre générique, automatisé et reproductible, transposable à tout problème de conception d'antennes patch.

Mots-clés : *antenne patch microruban, alimentation par inset, réseaux de neurones, modèle substitut, processus gaussien, HFSS, méthode des éléments finis, échantillonnage par hypercube latin, conception à 6 GHz.*

ملخص

أدى التطور المتسارع لأنظمة الاتصالات اللاسلكية الحديثة ــــ كالجيل الخامس وإنترنت الأشياء وتطبيقات الرادار قصير المدى ــــ إلى تنامي الطلب على هوائيات مدمجة، ذات سُمك منخفض وقابلة للإدماج بسهولة .وتظل الهوائيات المربعية المربوعية المطبوعة على شريط رفيع (Microstrip Patch) من أبرز الحلول لما تتميز به من بنية مسطحة وخفة في الوزن وكلفة تصنيع منخفضة .غير أن تصميم هذه الهوائيات يتطلب محاكاة كهرومغناطيسية كاملة عالية الدقة، تصبح كلفتها الحسابية شديدة الارتفاع عند تقييم مئات الهندسات داخل حلقة تحسين.

يبحث هذا العمل في استخدام نماذج بديلة (surrogate models) قائمة على الشبكات العصبية لتسريع تصميم هوائي مربعي ذي تغذية مدمجة بـinset ، يعمل حول 6 جيجاهرتز .وتعتمد المنهجية على سلسلة عددية كاملة وقابلة للتكرار (1) : نمذجة الهوائي في Ansys HFSS عبر طريقة العناصر المنتهية، (2) توليد خطة عينات بتقنية Latin Hypercube Samplingلاستكشاف فضاء التصميم ذي الأبعاد الستة($L, W, h, \varepsilon_r, y_0, g$) ، (3) استخدام نتائج المحاكاة لتغذية نموذجين بديلين ــــ شبكة عصبية متعددة الطبقات (MLP) ونموذج غوسي ــــ يتعلمان العلاقة بين الهندسة ومقاييس الأداء الثلاثة f_r و $|S_{11}|_{min}$و BW.

تم توليد خطة LHS مكونة من 150 تصميمًا، وقد التقت جميعها بنجاح فيHFSS ، مما وفّر قاعدة بيانات كاملة ومتوازنة من 150 عينة، قُسِّمت إلى مجموعة تدريب من 120 تصميمًا ومجموعة تحقق من 30 تصميمًا .واعتُمدت دالة كلفة مركّبة تجمع بين عقوبة تربيعية لانحراف التردد، وقيد أحادي الجانب على المطابقة، ومكافأة لعرض النطاق، لتوجيه عملية التحسين المدعومة بالنموذج البديل .وتقدم المنهجية المقترحة إطارًا عامًا وآليًا قابلًا للتكرار يمكن نقله إلى أي مسألة تصميم لهوائي مربعي.

الكلمات المفتاحية: هوائي مربعي، تغذية inset، شبكات عصبية، نموذج بديل، عملية غوسية، HFSS، طريقة العناصر المنتهية، تخطيط لاتيني هايبركيوب، 6 جيجاهرتز، تحسين الهوائيات.

Table of Contents

List of Figures

List of Tables

List of Symbols and Abbreviations

Symbols

Symbol	Meaning
L	Patch physical length (mm)
W	Patch physical width (mm)
H	Substrate thickness (mm)
ε_r	Relative permittivity of the substrate
εreff	Effective relative permittivity
ΔL	Length extension due to fringing fields (mm)
y_0	Inset depth of the microstrip feed (mm)
G	Inset gap of the microstrip feed (mm)
f_0 / fr	Centre / resonance frequency (GHz)
$\|S_{11}\|$	Magnitude of the reflection coefficient (dB)
BW	−10 dB impedance bandwidth (MHz)
Z_0	Characteristic / reference impedance (Ω)
Γ	Reflection coefficient
G, D, η	Gain, directivity and antenna efficiency
k_0	Free-space wavenumber
E, H	Electric and magnetic field vectors
J	Loss / cost function (Chapter 3)
σ^2, ℓ	Signal variance and length scale of the GP kernel

Abbreviations

Abbreviation	Meaning
ANN	Artificial Neural Network
MLP	Multilayer Perceptron
GP	Gaussian Process
RBF	Radial Basis Function (kernel)
MSE / RMSE / MAE	Mean Squared / Root Mean Squared / Mean Absolute Error
R^2	Coefficient of determination
LHS	Latin Hypercube Sampling
FEM	Finite Element Method
MoM	Method of Moments
FDTD	Finite Difference Time Domain
HFSS	High Frequency Structure Simulator (Ansys)
EM	Electromagnetic
RF	Radio Frequency
VSWR	Voltage Standing Wave Ratio
HPBW	Half-Power Beamwidth
PML	Perfectly Matched Layer
PCB	Printed Circuit Board
IoT	Internet of Things
5G	Fifth-Generation mobile networks
CSV	Comma-Separated Values (file format)
SGD	Stochastic Gradient Descent
ReLU	Rectified Linear Unit
UQ	Uncertainty Quantification

General Introduction

The last two decades have witnessed an unprecedented expansion of wireless communication technologies. Mobile telephony has rapidly moved from the second generation through to the fifth generation (5G), while emerging paradigms such as the Internet of Things (IoT), connected vehicles, wearable medical devices and short-range automotive radar have multiplied the demand for compact, low-cost and easily integrable antennas. In this context, the microstrip patch antenna has progressively established itself as one of the most attractive radiating elements, offering a low profile, light weight, ease of fabrication on standard printed-circuit-board (PCB) technology, and a natural compatibility with planar integration of active and passive RF components.

However, the apparent simplicity of a patch antenna hides a delicate design problem. The relationship between the antenna geometry (length, width, substrate thickness, permittivity, feed position, etc.) and its electromagnetic performance (resonance frequency, reflection coefficient, bandwidth, gain) is highly non-linear. Analytical models such as the transmission-line model or the cavity model provide useful starting points but quickly lose accuracy when realistic fabrication tolerances, fringing effects and parasitic radiation are taken into account. Accurate design therefore relies on full-wave numerical solvers — typically based on the Finite Element Method (FEM), as in Ansys HFSS, or on the Method of Moments — which can predict the antenna response with excellent fidelity but at the price of a heavy computational cost. A single rigorous simulation can take from a few minutes to several hours, and any iterative optimization procedure rapidly becomes prohibitive.

To overcome this bottleneck, the engineering community has progressively adopted the philosophy of surrogate modeling. The principle is to build a fast mathematical model — a metamodel — capable of approximating the input–output relationship of the expensive electromagnetic solver. Among the various surrogate techniques, artificial neural networks

(ANN) and Gaussian processes (GP) have emerged as the two leading candidates, the former for their excellent scalability and the latter for their built-in uncertainty quantification. Once trained on a moderate number of full-wave simulations, the surrogate can deliver near-instantaneous predictions and thus enables the use of population-based optimization algorithms that would otherwise be unrealistic.

The present thesis lies precisely at the intersection of these two worlds — classical microstrip antenna design on one hand, and machine-learning-driven optimization on the other. The objective is twofold: (i) to build a complete, automated, reproducible methodology that produces a high-quality training dataset for surrogate modeling, and (ii) to formulate a meaningful cost function that can drive the optimization of an inset-fed microstrip patch antenna operating around 6 GHz, a frequency of growing interest for Wi-Fi 6E, 5G mid-band and short-range radar applications.

The work is organized into three chapters that follow a natural progression from theory to practice:

- Chapter 1 — Introduction to Microstrip Patch Antennas. This chapter lays the theoretical ground of the work. It presents the operating principle of the patch antenna, the analytical design equations, the main performance parameters (bandwidth, return loss, radiation pattern, directivity, gain, impedance), the three usual feed techniques (coaxial, microstrip line and aperture-coupled) and the two classical analytical models (transmission line and cavity).
- Chapter 2 — Neural Network Surrogate Models. This chapter introduces the metamodel-based design philosophy. After explaining why HFSS simulations are too expensive for direct optimization, it describes the multilayer perceptron architecture, the backpropagation algorithm, the data-generation strategies (Latin Hypercube Sampling), the choice of loss function and the alternative Gaussian-process regression approach with its uncertainty estimates.
- Chapter 3 — Simulation Methodology Using HFSS. This is the methodological heart of the thesis. It first recalls the FEM formulation used by HFSS, then specifies the target 6 GHz design on FR-4 substrate and develops the step-by-step construction of the HFSS model. The originality of the chapter lies in the description of a fully automated, four-step Python pipeline (LHS sampling, batch HFSS simulation, S_{11} metrics extraction, dataset assembly). Practical issues — including a premature termination of HFSS at the 74th sample — are honestly discussed. Finally, a composite cost function combining frequency

accuracy, impedance matching and bandwidth reward is formulated to drive the optimization stage.

The thesis closes with a general conclusion that summarises the main contributions, evaluates the limits of the current work and outlines the perspectives for further research — including the training of the surrogate model on the dataset produced here, the optimisation of an array configuration and the extension of the methodology to multi-band and reconfigurable antennas.

CHAPTER 1

Introduction to Microstrip Patch Antennas

1.1 Introduction

Microstrip patch antennas have emerged as one of the most widely studied and implemented antenna configurations in modern wireless communication systems. Since their introduction in the 1950s by Deschamps and their practical development by Munson in the 1970s, these antennas have revolutionized the field of antenna engineering due to their unique characteristics and versatile applications [1].

A microstrip patch antenna essentially consists of a radiating metallic patch printed on a dielectric substrate with a ground plane on the opposite side. The patch can take various geometric shapes, including rectangular, circular, triangular, and more complex configurations. Among these, the rectangular patch is the most commonly employed due to its simplicity in design and analysis [2].

The rapid proliferation of wireless communication technologies, including 5G networks, Internet of Things (IoT) devices, satellite communications, and wearable electronics, has significantly increased the demand for compact, efficient, and multi-functional antennas. Microstrip patch antennas have proven particularly suitable for these applications owing to their low profile, light weight, ease of fabrication, and conformability to various surfaces [3].

However, conventional microstrip patch antennas suffer from several inherent limitations, including narrow bandwidth, low gain, and susceptibility to surface wave excitation. These limitations have motivated extensive research into various enhancement techniques, including the use of different substrate materials, modified patch geometries, array configurations, and advanced feeding techniques [4].

This chapter provides a comprehensive introduction to the fundamental principles, parameters, feeding techniques, and analysis methods of microstrip patch antennas

1.2 Basic Principles of Operation

The fundamental operating principle of a microstrip patch antenna is based on the concept of electromagnetic resonance. When the patch dimensions are approximately half-wavelength, the antenna resonates, resulting in efficient radiation of electromagnetic energy [5].

Figure 1.1 illustrates the basic structure of a rectangular microstrip patch antenna. The antenna consists of three primary layers: the ground plane, the dielectric substrate, and the radiating patch. The ground plane serves as a reflector, preventing radiation in the backward direction. The dielectric substrate provides mechanical support and influences the electrical characteristics of the antenna. The metallic patch, typically made of copper or gold, acts as the radiating element [2].

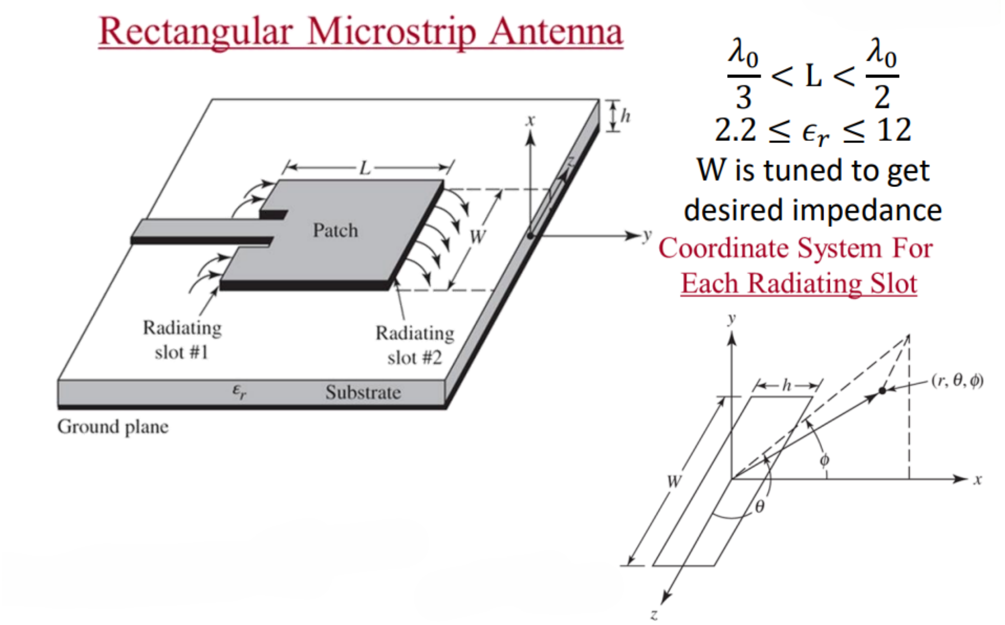


Figure 1.0 — Structure of a rectangular microstrip patch antenna showing the ground plane, substrate and radiating patch.

The operation can be understood through the transmission line model, where the patch is viewed as a section of a microstrip transmission line with width W and length L. At the resonant frequency, the length L is approximately equal to half the guided wavelength ($\lambda g/2$). The fringing fields at the edges of the patch radiate electromagnetic energy into space, effectively creating two radiating slots separated by the patch length [6].

The fringing effect is a critical phenomenon in microstrip antennas. Because the fields extend beyond the physical edges of the patch, the electrical length of the patch is greater than its physical length. This effect is accounted for by introducing an effective dielectric constant (ε_{reff}) and an effective length (L_{eff}), which are essential for accurate design calculations [7].

The effective dielectric constant is given by:

$$\varepsilon_{\text{reff}} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left(1 + 12\frac{h}{W}\right)^{-1/2} \quad (1.1)$$

where ε_r is the relative permittivity of the substrate, h is the substrate thickness and W is the patch width. The extension in length due to fringing is:

$$\Delta L = 0.412h\frac{(\varepsilon_{\text{reff}} + 0.3)\left(\frac{W}{h} + 0.264\right)}{(\varepsilon_{\text{reff}} - 0.258)\left(\frac{W}{h} + 0.8\right)} \quad (1.2)$$

Then the effective length becomes:

$$L_{eff}=L+2\Delta L$$

The resonant frequency of the fundamental TM_{10} mode is then calculated as:

$$f_r = \frac{c}{2L_{\text{eff}}\sqrt{\varepsilon_{\text{reff}}}} \quad (1.3)$$

where c is the speed of light in free space. These equations form the basis for the initial design of rectangular microstrip patch antennas [1].

1.3 Antenna Parameters

The performance of a microstrip patch antenna is characterized by several key parameters. Understanding these parameters is essential for designing antennas that meet specific application requirements. This section presents a detailed discussion of each parameter [2].

1.3.1 Bandwidth

Bandwidth is one of the most critical parameters of an antenna, defined as the range of frequencies over which the antenna operates satisfactorily. For microstrip patch antennas, bandwidth is typically defined as the frequency range where the return loss (S_{11}) is below −10 dB, corresponding to a Voltage Standing Wave Ratio (VSWR) of less than 2 : 1 [6].

Microstrip patch antennas inherently exhibit narrow bandwidth, typically in the range of 1–5 % of the centre frequency. This limitation arises from the high quality factor (Q) of the resonant cavity formed between the patch and the ground plane. The bandwidth can be expressed as:

$$\mathrm{BW}\ (\%) \approx \frac{\mathrm{VSWR} - 1}{Q\sqrt{\mathrm{VSWR}}} \times 100 \tag{1.4}$$

Several techniques have been developed to enhance the bandwidth of microstrip patch antennas, including the use of thicker substrates with lower dielectric constants, stacked patch configurations, slot loading, and the integration of parasitic elements [4].

$$\mathrm{BW} \approx \frac{16 \cdot h \cdot W \cdot c}{3\sqrt{2\pi} \cdot L^2 \cdot \varepsilon_r \sqrt{\varepsilon_{\mathrm{reff}}} \cdot \lambda_0 \cdot Q_{\mathrm{total}}}$$

where Q_{total} represents the total quality factor accounting for all loss mechanisms ($Qtotal^{-1} = Qrad^{-1} + Qd^{-1} + Qc^{-1} + Qsw^{-1}$) where Qrad, Qd, Qc and Qsw are respectively quality factor accounting for radiation, and dielectric, conduction and surface wave losses. The radiation quality factor Qrad is particularly significant and can be expressed as:

$$Q_{\text{rad}} = \frac{2\omega\varepsilon_0\varepsilon_{\text{reff}}LW}{hG_1}\left[1+\frac{1}{3}\left(\frac{\pi h}{L}\right)^2\right]$$

where $\omega = 2\pi f_0$ is the angular frequency and ε_0 is the permittivity of free space

The dimensional dependence of bandwidth provides the antenna designer with quantitative guidelines: increasing the substrate thickness h while maintaining $h \ll \lambda_0$ (typically $h < 0.1\lambda_0$) enhances bandwidth at the expense of increased surface wave excitation; reducing the dielectric constant εr toward unity (approaching air substrate) maximizes bandwidth but increases the physical antenna dimensions [4].

1.3.2 Polarization

Polarization describes the orientation of the electric field vector of the radiated electromagnetic wave. Microstrip patch antennas can be designed to produce linear, circular, or elliptical polarization depending on the application requirements [5].

For a rectangular patch operating in the fundamental TM_{10} mode, the polarization is linear, with the electric field oriented parallel to the patch length. Circular polarization can be achieved through several methods, including dual orthogonal feeds with a 90-degree phase difference, truncated corner patches, or nearly square patches with a single feed at a 45-degree angle [7].

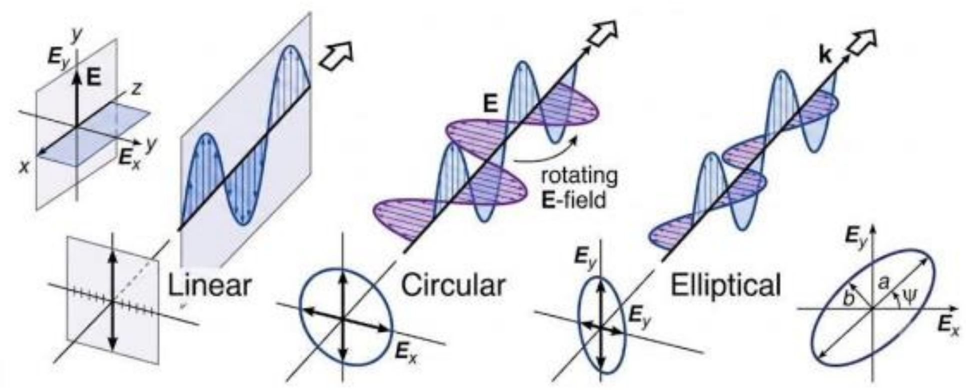


Figure 1-1 : Type de Polarization

Poor polarization matching between the transmitting and receiving antennas leads to losses due to polarization misalignment, which can result in a total signal cut-off (theoretical infinite loss between a vertically polarized wave and a horizontally polarized antenna).

1.3.3 Return Loss

Return loss, also known as reflection coefficient (S_{11}), measures the amount of power reflected from the antenna due to impedance mismatch between the feed line and the antenna. It is typically expressed in decibels (dB) and is calculated as:

$$S_{11}\ (\mathrm{dB}) = -20 \log_{10} |\Gamma| \quad (1.5)$$

where Γ is the reflection coefficient. A return loss of −10 dB or lower is generally considered acceptable, indicating that at least 90 % of the incident power is being transferred to the antenna. Figure 1.2 illustrates a typical return loss curve for a microstrip patch antenna [2].

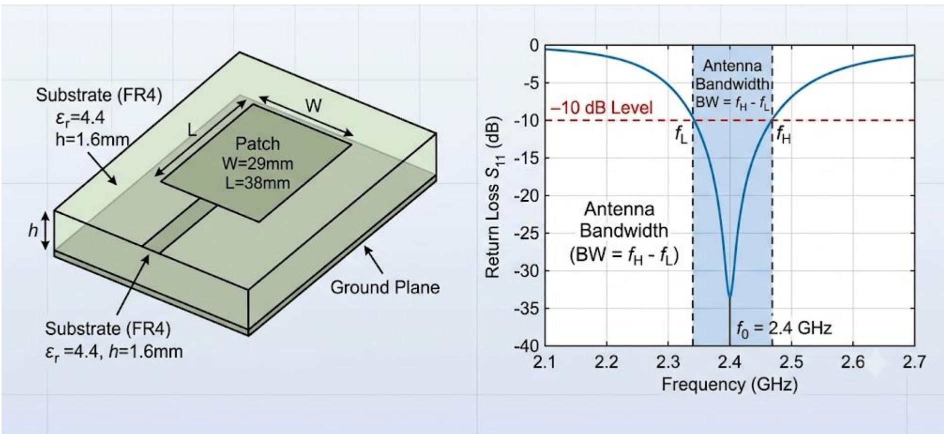


Figure 1.2 — Typical return loss (S_{11}) curve of a microstrip patch antenna showing the resonance at 2.4 GHz with bandwidth defined at the −10 dB level.

Figure 1.2 illustrates a typical return loss (S_{11}) characteristic curve for a rectangular microstrip patch antenna designed for operation in the 2.4 GHz ISM band. The curve exhibits a pronounced resonant dip at the center frequency $f_0 = 2.4$ GHz, where the return loss reaches its minimum value (typically −15 dB to −25 dB for a well-matched antenna). The bandwidth is defined at the −10 dB level, indicated by the horizontal reference line intersecting the S_{11} curve at two frequency points f_1 (lower cutoff) and f_2 (upper cutoff). The absolute bandwidth is calculated as $BW = f_2 - f_1$, and the fractional bandwidth is expressed as $FBW = (f_2 - f_1)/f_0$.

The sharpness of the resonance curve is directly related to the antenna quality factor Q, with higher Q values producing narrower, deeper resonant dips [6].

The return loss curve provides critical diagnostic information: (a) the resonant frequency f_0 corresponds to the frequency of minimum S_{11}; (b) the depth of the resonant dip indicates the quality of impedance matching; (c) the −10 dB bandwidth defines the operational frequency range; (d) the symmetry of the curve around f_0 reveals the purity of the resonant mode; and (e) the presence of secondary dips may indicate higher-order mode excitation or parasitic resonances. For optimal antenna performance, the design objective is to achieve $S_{11} < -10$ dB across the entire desired operational band while maintaining the resonant frequency precisely at the target value [2].

1.3.4 Radiation Pattern

The radiation pattern describes the spatial distribution of electromagnetic energy radiated by the antenna. For a rectangular microstrip patch antenna operating in the TM_{10} mode, the radiation pattern is broadside, meaning maximum radiation occurs in the direction perpendicular to the patch surface [5].

The E-plane radiation pattern (φ = 0°) exhibits a figure-eight pattern with nulls along the patch length, while the H-plane pattern (φ = 90°) is more omnidirectional. The half-power beamwidth (HPBW) is typically in the range of 60–90 degrees for both planes [1].

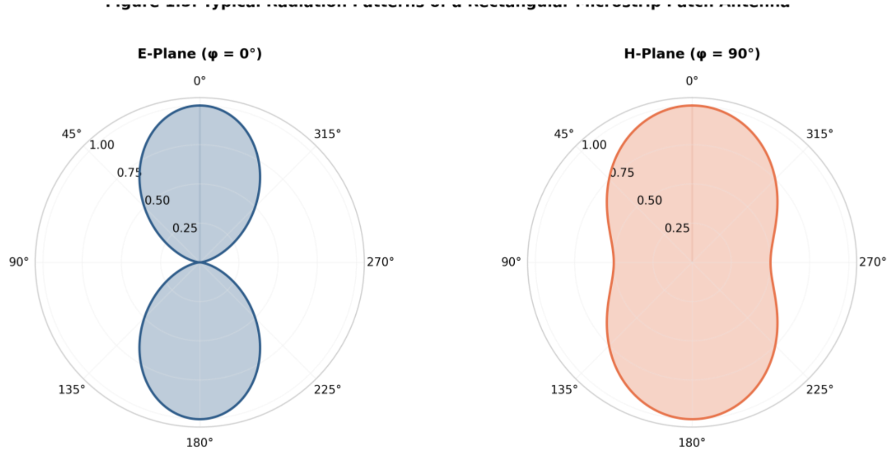


Figure 1.3 — Typical radiation patterns of a rectangular microstrip patch antenna: (a) E-plane (φ = 0°) and (b) H-plane (φ = 90°).

1.3.5 Directivity

Directivity is a measure of the antenna's ability to concentrate radiated power in a particular direction. It is defined as the ratio of the radiation intensity in a given direction to the average radiation intensity over all directions. For a rectangular microstrip patch antenna, the directivity is typically in the range of 5–8 dBi [6] where dBi is the gain over an isotropic antenna.

The directivity can be approximated using the following expressions:

$$D_0 \approx 6.6 \ \ (\text{for } W \ll \lambda_0) \qquad D_0 \approx 8.2 \ \ (\text{for } W \gg \lambda_0) \qquad (1.6)$$

where λ_0 is the free-space wavelength. Higher directivity can be achieved through array configurations or the use of superstrates [3].

1.3.6 Gain

Gain combines the effects of directivity and antenna efficiency, representing the ratio of the radiation intensity in a given direction to the radiation intensity that would be obtained if the power were radiated isotropically. The gain is related to directivity by:

$$G = \eta \cdot D \qquad (1.7)$$

where η is the antenna efficiency, which accounts for dielectric losses, conductor losses, and surface wave losses. Typical gain values for single-element microstrip patch antennas range from 5 to 7 dBi [2].

1.3.7 Input Impedance and Adaptation

Input impedance is the impedance presented by the antenna at its feed point. For efficient power transfer, the antenna input impedance must match the characteristic impedance of the feed line (typically 50 Ω). The input impedance of a rectangular patch at resonance is approximately given by [7]:

$$Z_{in} \approx \frac{1}{2G_1} \qquad (1.8)$$

where G_1 is the conductance of a single radiating slot. The input impedance varies with the feed position and can be adjusted by moving the feed point from the edge toward the centre of the patch. Various matching techniques, including quarter-wave transformers, inset feeds, and stub matching, are employed to achieve proper impedance matching [1].

Table 1.1 provides a comparison of typical parameter values for conventional and enhanced microstrip patch antennas.

Parameter	Conventional patch	Enhanced patch
Bandwidth	1 – 3 %	10 – 30 %
Gain (single element)	5 – 7 dBi	8 – 12 dBi
Efficiency	70 – 90 %	80 – 95 %
Polarization	Linear (typical)	Linear / Circular
VSWR (operating band)	≤ 2 : 1	≤ 1.5 : 1
Substrate	FR-4, Rogers RT/duroid	Multilayer, foam, exotic

Table 1.1 — Comparison of conventional and enhanced microstrip patch antenna parameters.

1.4 Feed Techniques

The feeding mechanism plays a crucial role in determining the overall performance of a microstrip patch antenna. The choice of feed technique affects impedance matching, bandwidth, polarization purity, and ease of fabrication. This section discusses the three most commonly used feeding techniques [5].

1.4.1 Coaxial Probe Feed

The coaxial probe feed, also known as the coaxial feed, is one of the simplest and most widely used feeding techniques. In this configuration, the inner conductor of a coaxial connector extends through the substrate and makes contact with the radiating patch, while the outer conductor is connected to the ground plane [6].

The main advantages of the coaxial probe feed include its simplicity, ease of implementation, and direct compatibility with coaxial cables. The feed position can be adjusted to achieve optimal impedance matching. However, this technique introduces an inductive component that becomes significant for thicker substrates, potentially limiting the achievable bandwidth [2].

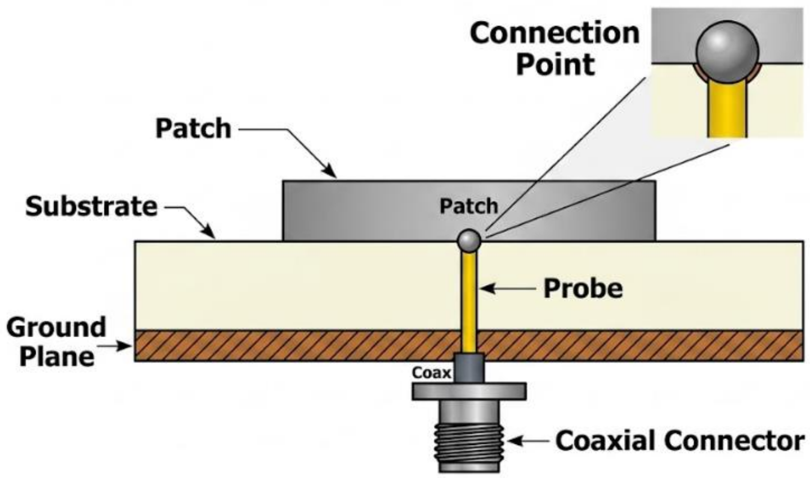


Figure 1.4 — Coaxial probe feed configuration showing the inner conductor connected to the patch and the outer conductor connected to the ground plane.

The input impedance variation with feed position xf along the patch centerline is approximately given by:

$$R_{\text{in}}(x_f) = R_{\text{edge}} \cdot \cos^2\left(\frac{\pi x_f}{L}\right)$$

where Redge represents the input resistance at the patch edge (xf = 0) and L is the patch length. This cosine-squared dependence enables precise impedance control through feed positioning.

1.4.2 Microstrip Line Feed

In the microstrip line feed technique, a conducting strip is connected directly to the edge of the radiating patch. This feed can be etched on the same substrate, providing a planar structure. The inset feed variation involves cutting a notch into the patch to improve impedance matching without requiring additional matching components [7].

The microstrip line feed offers the advantage of easy fabrication and integration with other microstrip components. However, as the substrate thickness increases, surface waves and spurious feed radiation become more pronounced, which can degrade the antenna performance [1].

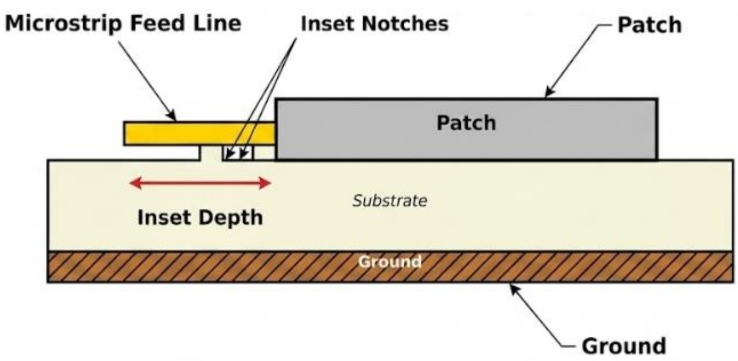


Figure 1.5 — Microstrip line feed (inset feed) configuration with impedance-matching inset.

The inset feed provides an approximate input impedance given by [7]:

$$R_{\mathrm{in}}(d_i) \approx R_{\mathrm{edge}} \cdot \cos^4\left(\frac{\pi d_i}{L}\right)$$

where di is the inset depth. The fourth-power dependence provides more rapid impedance variation compared to the coaxial feed, enabling finer matching control.

1.4.3 Aperture Coupled Feed

The aperture coupled feed technique uses electromagnetic coupling through a slot in the ground plane to transfer energy from the feed line to the radiating patch. The feed network is located on a separate substrate layer, isolated from the radiating patch by the ground plane [4].

This technique offers several advantages, including the ability to optimize the feed and radiation substrates independently, reduced spurious radiation from the feed network, and potentially wider bandwidth. The coupling level can be controlled by adjusting the slot dimensions and position. However, the multilayer structure increases fabrication complexity [5].

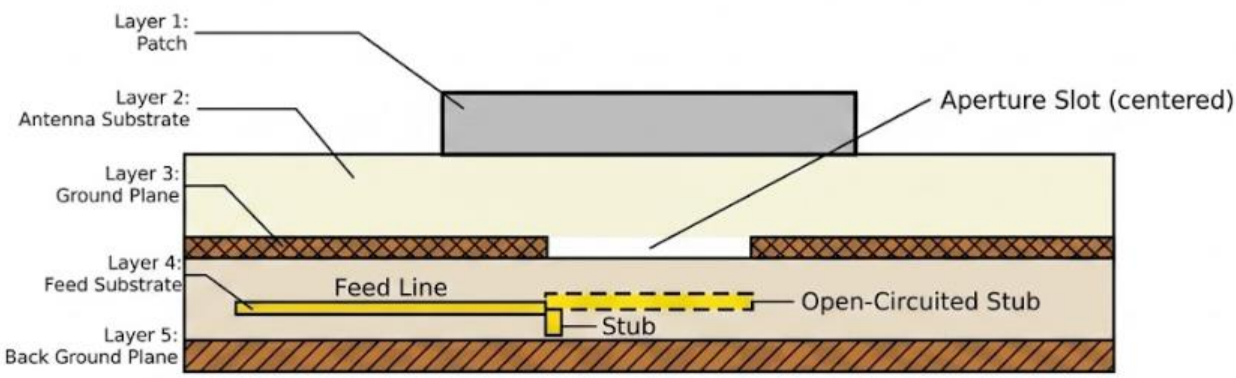


Figure 1.6 — Aperture coupled feed configuration showing the slot-coupled multilayer structure.

Table 1.2 summarizes the key characteristics of the three feed techniques discussed above.

Characteristic	Coaxial Probe	Microstrip Line	Aperture Coupled
Fabrication	Simple	Simple	Complex
Bandwidth	Narrow	Moderate	Wide
Spurious Radiation	Low	Moderate	Very Low
Impedance Matching	By position	By inset	By slot size
Multilayer	No	No	Yes

Table 1.2 — Comparison of microstrip antenna feed techniques.

The coupling coefficient between the feed line and the patch through the slot can be approximated by :

$$|S_{21}|^2 \approx \left(\frac{\pi l_s}{2\lambda_0}\right)^2 \cdot \left[\frac{\sin\left(\frac{\pi l_s}{2L_{\text{patch}}}\right)}{\frac{\pi l_s}{2L_{\text{patch}}}}\right]^2 \cdot \left(\frac{Z_0}{Z_c}\right)$$

1.5 Methods of Analysis

The analysis of microstrip patch antennas involves determining their electrical characteristics, including resonant frequency, input impedance, radiation pattern, and bandwidth. Several analytical and numerical methods have been developed for this purpose. This section discusses the two fundamental analytical models: the transmission line model and the cavity model [6].

1.5.1 Transmission Line Method

The transmission line model (TLM) is the simplest analytical approach for microstrip patch antennas. In this model, the rectangular patch is represented as two parallel radiating slots separated by a low-impedance transmission line of length equal to the patch length L [7].

Each radiating slot is modeled as an equivalent parallel admittance $Y = G + jB$, where G represents the radiation conductance and B represents the susceptance due to the fringing field. The total input admittance at the resonant frequency is the sum of the individual slot admittances transformed to the feed point [2].

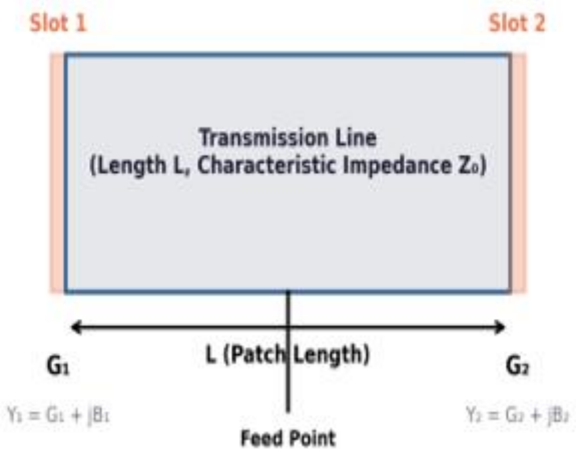


Figure 1.7 — Transmission line model of a rectangular microstrip patch antenna showing the two radiating slots and the connecting transmission line.

Physical Basis and Model Representation:

In the transmission line model, the rectangular patch is conceptually decomposed into two parallel radiating slots separated by a low-impedance microstrip transmission line section of length L equal to the patch length. Each radiating slot is modeled as an equivalent parallel admittance Y = G + jB, where the real component G represents the radiation conductance accounting for power radiated into free space, and the imaginary component B represents the susceptance arising from the energy stored in the fringing electric fields at the patch edges. The total input admittance at the resonant frequency is the algebraic sum of the individual slot admittances appropriately transformed to the feed point position using transmission line theory [2].

The equivalent circuit representation consists of: (a) two parallel admittances $Y_1 = G_1 + jB_1$ and $Y_2 = G_2 + jB_2$ representing the radiating slots at the patch edges; (b) a transmission line section of characteristic impedance Z_0 and electrical length βL connecting the two slots; and (c) the feed point positioned at a distance xf from one edge, where $\beta = 2\pi\sqrt{\varepsilon_{reff}}/\lambda_0$ is the propagation constant. At resonance, the patch length L is approximately λg/2, and the two slot admittances add in parallel at the center, yielding the total input admittance [1].

Mathematical Formulation:

The radiation conductance of a single slot is given by:

$$G_1 = \frac{W^2}{90\lambda_0^2} \quad \text{pour} \quad W \ll \lambda_0$$

The fringing field susceptance is approximated by:

$$B_1 = \frac{k_0 \Delta L \sqrt{\varepsilon_{\text{reff}}}}{Z_0} \cdot [1 - 0{,}636 \ln(k_0 h)]$$

where $k_0 = 2\pi/\lambda_0$ is the free-space wavenumber, ΔL is the length extension due to fringing, and Z_0 is the characteristic impedance of the microstrip line. The input impedance at the feed position xf is obtained by transforming the slot admittances through the transmission line sections:

$$Y_{\text{in}}(x_f) = Y_1 + Y(x_f) + Y(L - x_f)$$

where Y(x) represents the transformed admittance of the opposite slot through a transmission line of length x. At resonance, the imaginary components cancel ($B_{total} = 0$), and the real part gives the input resistance [6].

Advantages and Limitations:

The principal advantage of the transmission line model is its exceptional simplicity and computational efficiency, enabling rapid manual calculations and providing valuable physical insight into the antenna's operation. The model clearly illustrates the slot radiation mechanism and the relationship between patch dimensions and resonant frequency. However, the model exhibits significant limitations: (a) limited accuracy for thicker substrates ($h > 0.02\lambda_0$) where fringing effects become dominant; (b) inability to analyze non-rectangular patch geometries (circular, triangular, etc.); (c) failure to account for mutual coupling between radiating edges; (d) neglect of higher-order mode effects; and (e) inaccurate prediction of cross-polarization and back radiation. Despite these limitations, the TLM remains valuable for initial design and educational purposes [1].

1.5.2 Cavity Model Method

The cavity model provides a more rigorous analysis by treating the region between the patch and the ground plane as a resonant cavity bounded by electric walls (top and bottom) and magnetic walls (sidewalls). This approach accounts for the field distribution within the cavity and provides more accurate results compared to the transmission line model [6].

The cavity model assumes that the substrate thickness is much smaller than the wavelength ($h \ll \lambda$), allowing the fields to be considered uniform in the vertical direction. The resonant modes of the cavity are determined by solving the wave equation with appropriate boundary conditions [5].

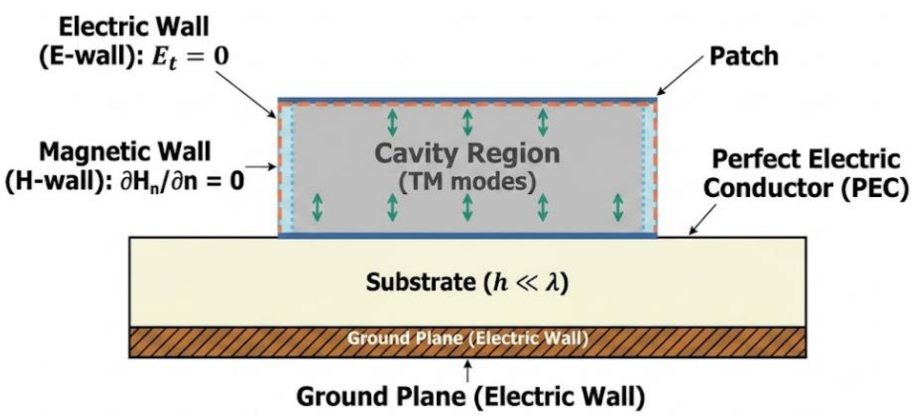


Figure 1.8 — Cavity model representation showing the electric and magnetic wall boundaries.

While the cavity model offers improved accuracy over the transmission line model, it involves more complex mathematical formulations. Both models are considered simplified approaches, and for more accurate analysis, full-wave numerical methods such as the Method of Moments (MoM), Finite Element Method (FEM), and Finite Difference Time Domain (FDTD) are employed [3].

Physical Basis and Boundary Conditions:

The cavity model is founded on the assumption that the substrate thickness h is much smaller than the operating wavelength ($h << \lambda$), which is universally satisfied in practical microstrip antennas (typically $h/\lambda_0 < 0.02$). Under this thin-substrate approximation, the electromagnetic fields can be considered essentially uniform in the vertical (z) direction, reducing the problem to a two-dimensional boundary value problem in the transverse (x-y) plane. The cavity is bounded by: (a) electric walls (perfect conductors) at the top (patch) and bottom (ground plane) surfaces, where the tangential electric field vanishes (Et = 0); and (b) magnetic walls (perfect magnetic conductors) along the four sidewalls, where the tangential magnetic field vanishes (Ht = 0). The magnetic wall approximation is justified by the observation that the fringing fields at the edges are predominantly electric, with minimal magnetic field component [5].

Mathematical Formulation:

Within the cavity, the electric field satisfies the homogeneous wave equation with appropriate boundary conditions. For the dominant TMz modes (transverse magnetic to z), the electric field has only a z-component Ez(x,y), which satisfies:

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + k^2\right) E_z(x, y) = 0$$

where $k = k_0\sqrt{\varepsilon_{reff}}$ is the wavenumber in the dielectric medium. Applying the boundary conditions—Ez = 0 at the electric walls (top and bottom) and ∂Ez/∂n = 0 at the magnetic walls (sidewalls, where n is the normal direction)—yields the eigenmode solutions:

$$E_{z,mn}(x, y) = E_0 \cdot \cos\left(\frac{m\pi x}{L_{\text{eff}}}\right) \cdot \cos\left(\frac{n\pi y}{W_{\text{eff}}}\right)$$

where m and n are non-negative integers (m, n = 0, 1, 2, ...) representing the mode numbers, and Leff and Weff are the effective patch dimensions accounting for fringing. The corresponding resonant frequencies are:

$$f_{mn} = \frac{c}{2\sqrt{\varepsilon_{\text{reff}}}} \cdot \sqrt{\left(\frac{m}{L_{\text{eff}}}\right)^2 + \left(\frac{n}{W_{\text{eff}}}\right)^2}$$

This equation explicitly reveals the multimode nature of the microstrip patch cavity. The fundamental mode TM_{10} (m=1, n=0) occurs at the lowest resonant frequency and produces the characteristic broadside radiation pattern. Higher-order modes (TM_{20}, TM_{02}, TM_{11}, etc.) occur at higher frequencies and produce distinct radiation patterns as discussed in Section 1.3.4 [7].

Power Loss Mechanisms and Quality Factor:

The cavity model rigorously accounts for all power loss mechanisms through the definition of separate quality factors: (a) radiation quality factor Qrad, associated with power radiated through the fringing fields at the cavity edges; (b) dielectric quality factor Qd = 1/tan δ, where tan δ is the dielectric loss tangent; (c) conductor quality factor Qc, associated with ohmic losses in the patch and ground plane metallization; and (d) surface wave quality factor Qsw, accounting for power coupled into substrate-guided surface wave modes. The total quality factor is obtained from [6]:

$$\frac{1}{Q_t} = \frac{1}{Q_{\text{rad}}} + \frac{1}{Q_d} + \frac{1}{Q_c} + \frac{1}{Q_{\text{sw}}}$$

The input impedance is computed by evaluating the cavity fields excited by the feed probe or line, integrating the Poynting vector over the radiating aperture, and applying the equivalence principle to determine the far-field radiation pattern. The cavity model thus provides a unified framework for predicting all antenna parameters from a single field solution [5].

Advantages and Limitations:

The cavity model offers substantially improved accuracy over the transmission line model, particularly for input impedance prediction, higher-order mode analysis, and patch geometries with aspect ratios departing from unity. It provides a physically intuitive picture of the resonant behavior and enables analytical treatment of coupling between adjacent patches in array configurations. However, the model retains certain approximations: (a) the magnetic wall boundary condition is approximate and becomes less accurate for thicker substrates; (b) the model does not rigorously account for the exact fringing field distribution; (c) surface wave effects are treated approximately; and (d) the analysis becomes mathematically involved for non-rectangular geometries. For these reasons, both the TLM and CM are considered simplified analytical approaches, and for rigorous analysis of complex antenna structures, full-wave numerical methods are indispensable [3].

1.6 Advantages and Disadvantages

Microstrip patch antennas offer numerous advantages that have contributed to their widespread adoption in modern communication systems. However, they also exhibit certain limitations that must be considered in the design process [4].

Figure 1.9 — Summary of advantages and disadvantages of microstrip patch antennas.

The advantages include low profile and lightweight construction, ease of fabrication using printed circuit board techniques, low manufacturing cost for mass production, compatibility with planar and integrated circuit designs, capability for dual-frequency and multi-frequency operation, and support for both linear and circular polarization [2].

The primary disadvantages include inherently narrow bandwidth, relatively low gain for single-element configurations, low power handling capacity, increased physical size at lower frequencies, excitation of surface waves in thicker substrates, sensitivity to fabrication tolerances, and dielectric and conductor losses that reduce efficiency [1].

1.7 Conclusion

This chapter has presented a comprehensive introduction to microstrip patch antennas, covering their fundamental operating principles, key performance parameters, feeding techniques, and analytical methods. The microstrip patch antenna remains a cornerstone of modern antenna engineering due to its unique combination of desirable characteristics.

The antenna parameters discussed in Section 1.3 provide the quantitative framework for evaluating antenna performance. Understanding these parameters is essential for designing antennas that meet specific application requirements. The choice of feeding technique significantly impacts the overall antenna performance and must be carefully considered during the design process.

While simplified analytical models such as the transmission line model and cavity model provide valuable insights and serve as starting points for design, the increasing complexity of modern antenna structures necessitates the use of full-wave electromagnetic simulation tools. The computational cost associated with these simulations has motivated the development of surrogate modeling techniques, which will be discussed in Chapter 2.

CHAPTER 2

Neural Network and GP Surrogate Models

2.1 Introduction

The design and optimization of modern antenna structures increasingly rely on full-wave electromagnetic (EM) simulations. While these simulations provide highly accurate results, their computational cost can be prohibitive, particularly when iterative optimization processes require hundreds or thousands of simulation runs. This computational bottleneck has motivated the development of surrogate modeling techniques that can provide rapid approximations of antenna performance with acceptable accuracy [8].

Surrogate models, also known as metamodels or response surface models, are mathematical approximations of computationally expensive simulation models. Among the various surrogate modeling approaches, neural networks have emerged as particularly powerful tools due to their ability to approximate complex nonlinear relationships between design parameters and antenna performance metrics [9].

This chapter presents the fundamental concepts of neural network-based surrogate modeling for antenna design. The discussion covers the principles of neural networks, data generation strategies, model training procedures, and the Gaussian process regression approach as an alternative surrogate modeling technique. Understanding these concepts is essential for developing efficient antenna optimization workflows [10].

2.2 Simulation of Microwave Structures by HFSS

High Frequency Structure Simulator (HFSS) is a commercial finite element method (FEM) software widely used for simulating microwave and RF structures, including antennas. HFSS solves Maxwell's equations in the frequency domain, providing accurate predictions of electromagnetic field distributions, S-parameters, radiation patterns, and other antenna characteristics [11].

The FEM-based approach in HFSS involves discretizing the computational domain into a mesh of tetrahedral elements. The electromagnetic fields within each element are approximated using polynomial basis functions, and the governing equations are solved to obtain the field distribution throughout the structure. The accuracy of the solution depends on the mesh density, with finer meshes providing more accurate results at the cost of increased computational time [12].

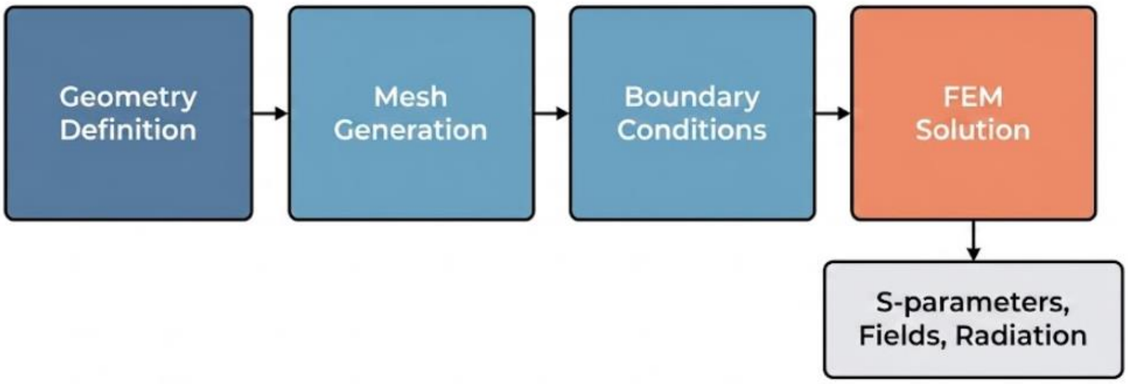


Figure 2.1 — HFSS simulation workflow using the finite element method.

For antenna optimization, HFSS simulations typically require several minutes to hours per design evaluation, depending on the complexity of the structure and the frequency range of interest. When combined with population-based optimization algorithms that may require thousands of function evaluations, the total optimization time can become impractical. This limitation motivates the development of surrogate models that can provide near-instantaneous predictions once trained [13].

2.3 Basic Principles of Neural Networks

Artificial neural networks (ANNs) are computational models inspired by the structure and function of biological neural systems. A neural network consists of interconnected nodes (neurons) organized in layers: an input layer, one or more hidden layers, and an output layer. Each connection between neurons has an associated weight, and each neuron applies an activation function to the weighted sum of its inputs [14].

The multilayer perceptron (MLP) is the most commonly used neural network architecture for regression tasks. In an MLP, information flows in one direction from the input layer through the hidden layers to the output layer. The universal approximation theorem states that a feedforward network with a single hidden layer containing a sufficient number of neurons can approximate any continuous function to arbitrary precision, making MLPs well-suited for surrogate modeling [15].

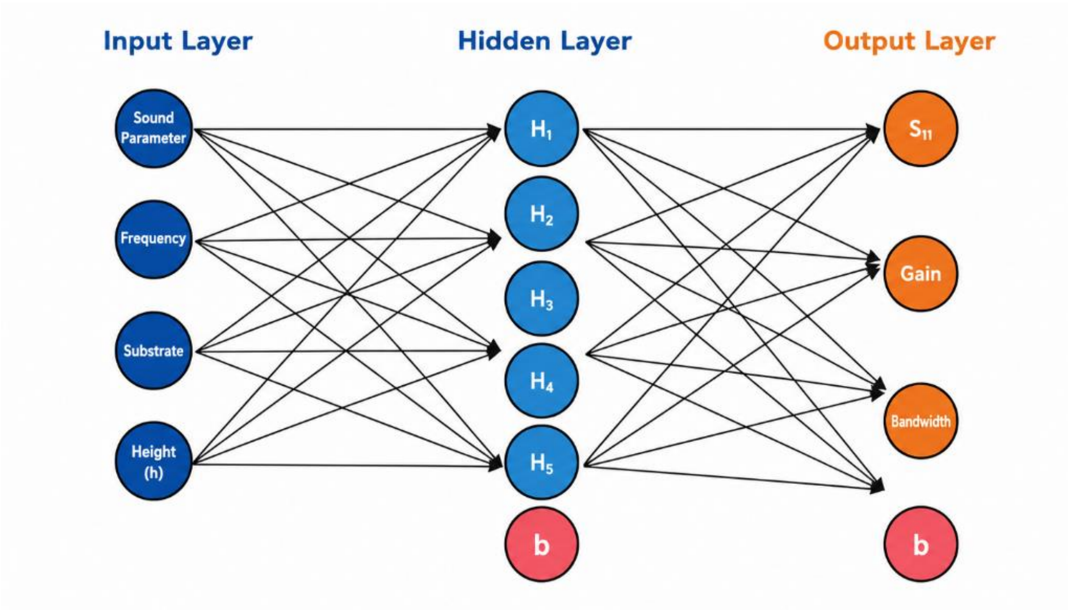


Figure 2.2 — Architecture of a multilayer perceptron (MLP) with input layer, hidden layers and output layer.

For a neuron j in layer l, the output is computed as:

$$A^{(l)} = f(W^{(l)} A^{(l-1)} + B^{(l)}) \tag{2.1}$$

where $w_{ji}^{(l)}$ is the weight connecting neuron i in layer l−1 to neuron j in layer l, $b_j^{(l)}$ is the bias term and f(·) is the activation function. Common activation functions include the hyperbolic tangent (tanh), sigmoid and Rectified Linear Unit (ReLU) [9].

The training process involves adjusting the weights and biases to minimize the difference between the network predictions and the target values. This is achieved through the backpropagation algorithm, which computes the gradient of the loss function with respect to each parameter and updates the parameters using gradient descent [16].

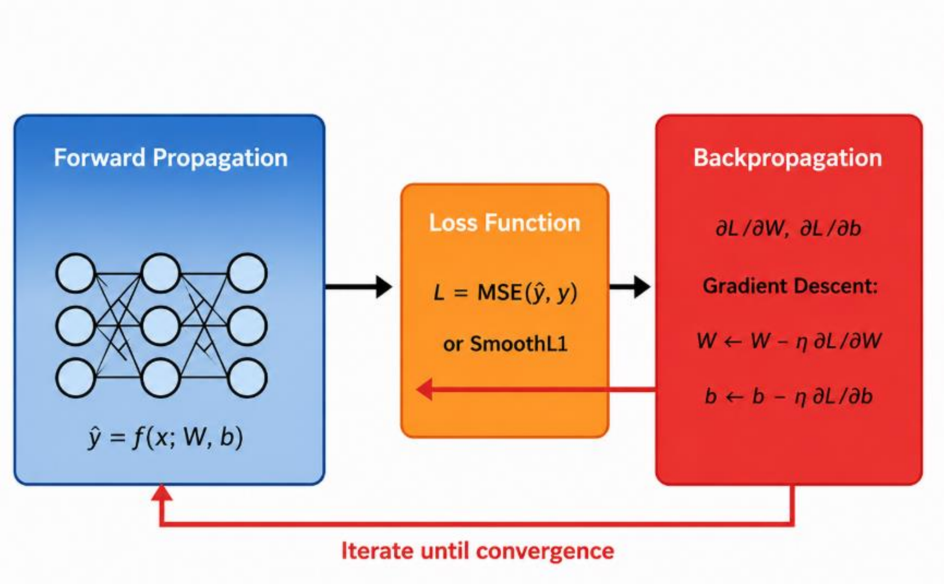


Figure 2.3 — Neural network training via backpropagation and gradient descent.

2.4 Antenna Modeling

The application of neural networks to antenna modeling involves establishing a mapping between antenna geometric parameters and performance metrics. This mapping is learned from a dataset of input–output pairs generated through full-wave electromagnetic simulations [13].

2.4.1 Antenna Geometry

The first step in developing a surrogate model is to define the antenna geometry and identify the design variables (input parameters). For a rectangular microstrip patch antenna, the key geometric parameters typically include the patch length (L), patch width (W), substrate thickness (h), relative permittivity (ε_r), and feed position (xf, yf) [10].

The parameter space must be carefully defined to cover the range of designs of interest while maintaining physical feasibility. Latin Hypercube Sampling (LHS) or uniform grid sampling are commonly used to generate training data that adequately covers the design space [11].

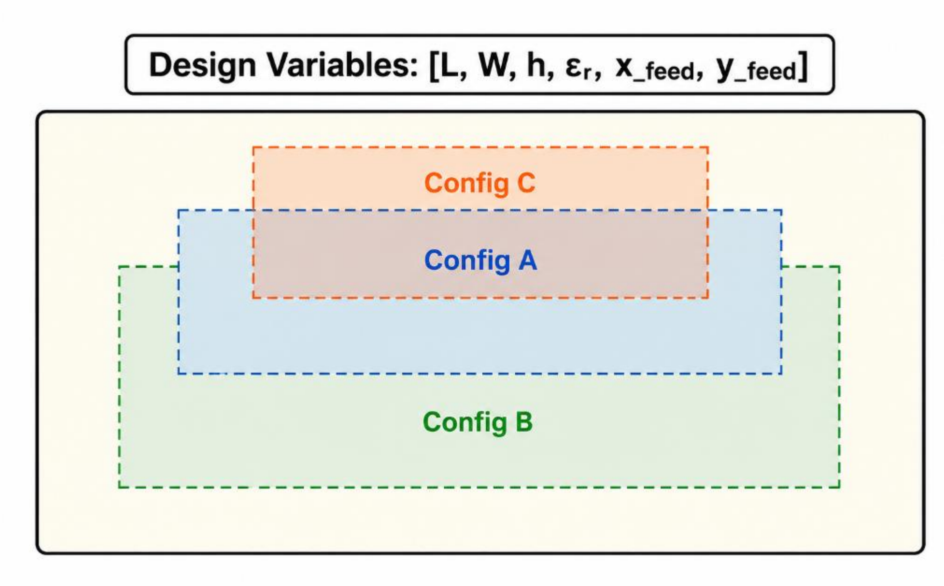


Figure 2.4 — Parameterized antenna geometry showing multiple configuration options for surrogate-model training.

2.4.2 Data Generation

Data generation involves performing electromagnetic simulations for each set of geometric parameters to obtain the corresponding antenna performance metrics. The output parameters typically include return loss (S_{11}), resonant frequency, bandwidth, gain, and radiation pattern characteristics [12].

The number of training samples required depends on the complexity of the parameter space and the desired model accuracy. As a general guideline, the number of samples should be at least 10 times the number of input parameters. The dataset is typically divided into training (80 %) and testing (20 %) subsets to evaluate model generalization [15].

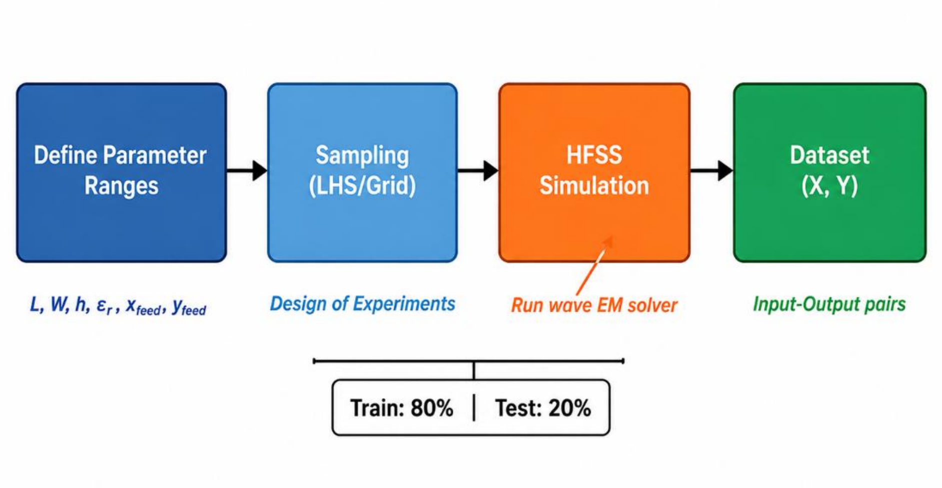


Figure 2.5 — Data generation pipeline for neural-network surrogate-model training.

An important consideration in data generation is the computational cost. While generating training data requires significant upfront investment in simulation time, this cost is amortized over the many rapid evaluations performed during optimization using the trained surrogate model [14].

2.5 Determination of Loss Function

The loss function quantifies the discrepancy between the neural network predictions and the true values from electromagnetic simulations. The choice of loss function significantly affects the training process and the final model accuracy. Several loss functions are commonly used in antenna surrogate modeling [16].

The Mean Squared Error (MSE) is the most widely used loss function for regression tasks:

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \qquad (2.2)$$

where y_i is the true value, $\hat{y}_i$ is the predicted value, and N is the number of samples. MSE penalizes large errors more heavily than small errors due to the squared term [9].

For improved robustness to outliers, the Smooth L1 (Huber) loss can be employed. This loss function combines the benefits of MSE and Mean Absolute Error (MAE), using MSE for small errors and MAE for large errors:

$$\mathrm{Lsmooth} = \{\, 0.5\,(y - \hat{y})^2 \;\text{ if } |y - \hat{y}| < \delta\,; \quad \delta \cdot |y - \hat{y}| - 0.5\,\delta^2 \;\text{ otherwise}\,\} \qquad (2.3)$$

During training, the loss function is minimized using optimization algorithms such as Adam, RMSprop, or Stochastic Gradient Descent (SGD) with momentum. The learning rate and other hyperparameters must be carefully tuned to ensure convergence [15].

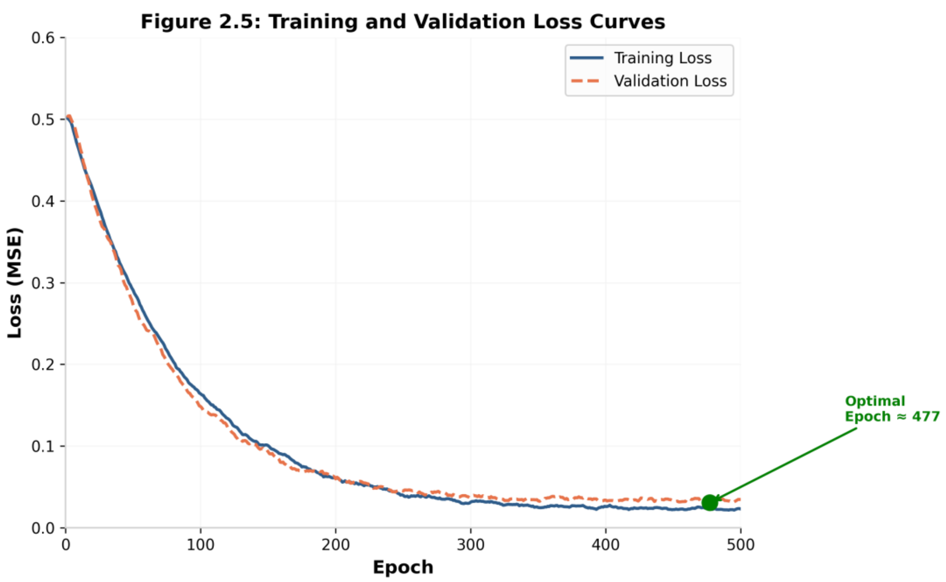


Figure 2.6 — Typical training and validation loss curves showing convergence behaviour and the early-stopping point.

2.6 Antenna Modeling via Gaussian Process Regression

Gaussian Process (GP) regression offers a powerful non-parametric Bayesian framework for antenna modeling, providing not only predictions of electromagnetic performance but also principled uncertainty quantification. Unlike neural networks that require fixed architectures, GPs define a distribution over functions, making them particularly suitable for modeling the complex, non-linear relationships between antenna geometric parameters and their electromagnetic responses. This section presents the theoretical foundations of GP regression as applied to antenna modeling, beginning with the essential building block — the multivariate normal distribution — and progressing through the most commonly used kernel functions that encode the smoothness and correlation structure of antenna performance landscapes [17].

2.6.1 Multivariate Normal Distribution

The multivariate normal (MVN) distribution forms the probabilistic foundation of Gaussian Process regression. A random vector $x \in R^D$ follows an MVN distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, denoted $x \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, when its probability density function is given by:

$$p(X) = \frac{1}{(2\pi)^{D/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}\operatorname{diag}\left((X - \mathbf{1}\mu^T)\Sigma^{-1}(X - \mathbf{1}\mu^T)^T\right)\right) \tag{2.5}$$

In the context of antenna modeling, the MVN (MultiVariate Normal)distribution provides the mathematical machinery to reason jointly about the electromagnetic performance (return loss, resonance frequency, bandwidth) at multiple design points simultaneously. The key property exploited by GP regression is that any finite collection of function values evaluated at a set of input locations follows a joint Gaussian distribution. This allows us to derive closed-form expressions for predictions at unobserved design configurations and to quantify the uncertainty associated with those predictions [18].

The covariance function, or kernel, is the cornerstone of Gaussian Process regression as it encodes the fundamental assumptions about the function being modeled. The kernel defines the similarity between pairs of input points and determines the smoothness, periodicity, and other structural properties of the GP prior. Two kernel families are particularly relevant for antenna modeling.

2.6.2 Squared Exponential (SE) kernel. Also known as the radial basis function (RBF) or Gaussian kernel, the SE kernel is the most widely used covariance function in GP regression. It assumes that the underlying function is infinitely differentiable, making it appropriate for modeling smooth electromagnetic responses:

$$K_{ij} = k_{SE}(x_i, x_j) = \sigma^2 \exp\left(-\frac{\sum_{k=1}^{D}(x_{ik} - x_{jk})^2}{2\ell^2}\right) \tag{2.6}$$

where $\sigma^2 \geq 0$ is the signal variance controlling the amplitude of the function variations, and $\ell > 0$ is the length-scale parameter determining how rapidly the correlation decays with distance in the input space. For antenna design, the length-scale captures the sensitivity of electromagnetic performance to changes in geometric parameters — small length-scales indicate highly sensitive parameters, while large length-scales suggest smoother, more predictable relationships [19].

Maternkernel. While the SE kernel assumes infinite differentiability, the Mat\u00e9rn family offers greater flexibility by allowing the user to control the smoothness of the sampled functions. **Matern kernel**with smoothness parameter ν is given by:

$$k_\nu(x, x') = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}\|x - x'\|}{\ell}\right)^\nu K_\nu\left(\frac{\sqrt{2\nu}\|x - x'\|}{\ell}\right) \tag{2.7}$$

where $K_\nu()$ is the modified Bessel function of the second kind. Of particular practical importance are the **Matern**-3/2 and **Matern**-5/2 kernels, which correspond to once and twice differentiable functions respectively. **Matern**-5/2 kernel provides a good compromise — it captures physically realistic antenna behavior without the overly smooth assumptions of the SE kernel, making it particularly suitable for modeling the resonance and impedance-matching characteristics of patch antennas where sharp transitions in the S_11 response may occur [20].

The choice between SE and **Matern kernel** involves a trade-off: the SE kernel offers simpler optimization and smoother predictions but may oversmooth important features of the antenna response, while **Matern kernel** can better capture local variations at the cost of additional hyperparameter sensitivity. In practice, cross-validation on the training data is used to select the kernel family and hyperparameters that yield the best predictive performance for the specific antenna geometry under study.

The model building process involves several key steps. First, the network architecture is defined, including the number of hidden layers, neurons per layer, and activation functions. A common starting point for antenna modeling is a network with two hidden layers, each containing 20–50 neurons with hyperbolic tangent activation [13].

Second, the training data is normalized to improve training stability and convergence. Input and output features are typically scaled to the range [0, 1] or standardized to have zero mean and unit variance. This preprocessing step ensures that all features contribute equally to the loss function [14].

Third, the network is trained using backpropagation with an appropriate optimization algorithm. Regularization techniques such as L2 weight decay and early stopping are employed to prevent overfitting and improve generalization to unseen data [16].

The performance of the trained model is evaluated using metrics such as the coefficient of determination (R^2), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). An R^2 value close to 1 indicates excellent agreement between predictions and simulations [9].

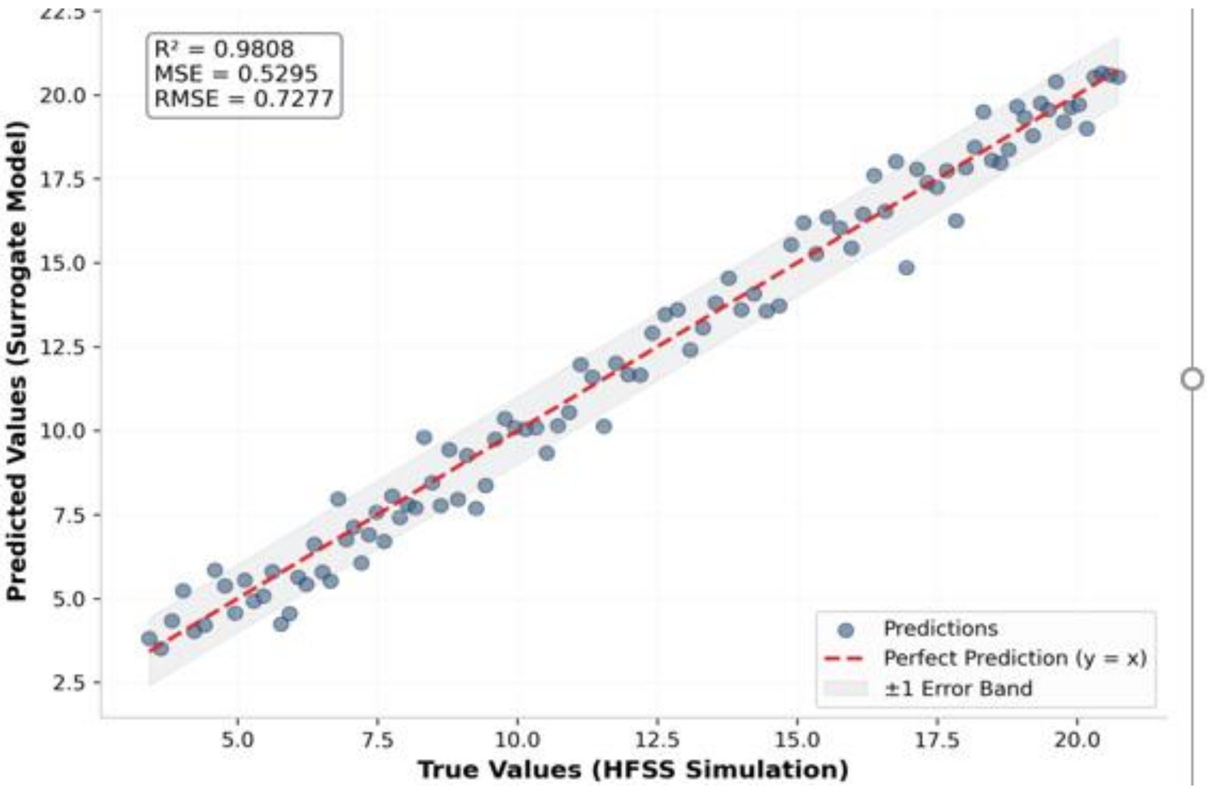


Figure 2.7 — Regression model performance: predicted vs. actual values with R^2, MAE and RMSE metrics.

Table 2.1 summarizes the performance metrics for different neural network architectures applied to antenna modeling.

Architecture	R^2	MSE	RMSE	Train time
MLP 4-20-3	0.985	0.0021	0.0458	≈ 2 min
MLP 4-40-3	0.992	0.0013	0.0361	≈ 4 min
MLP 4-20-20-3	0.996	0.0008	0.0283	≈ 6 min
RBF Network	0.988	0.0018	0.0424	≈ 3 min

Table 2.1 — Performance comparison of different neural-network architectures for antenna modeling.

2.7 Gaussian Process

Gaussian Process (GP) regression, also known as Kriging, is a non-parametric Bayesian approach to surrogate modeling that provides not only predictions but also uncertainty estimates. Unlike neural networks, which require specification of the network architecture, Gaussian processes define a distribution over functions, making them highly flexible for modeling complex relationships [17].

A Gaussian Process is fully specified by its mean function m(x) and covariance function (kernel) k(x, x'). The kernel function encodes the assumption about the smoothness and correlation structure of the underlying function. The most commonly used kernel is the squared exponential (RBF) kernel:

$$k(x, x') = \sigma_f^2 \exp\left(-\frac{\|x - x'\|^2}{2\ell^2}\right) \quad (2.4)$$

where σ^2 is the signal variance and ℓ is the length-scale parameter. These hyperparameters are typically estimated by maximizing the marginal likelihood of the observed data [18].

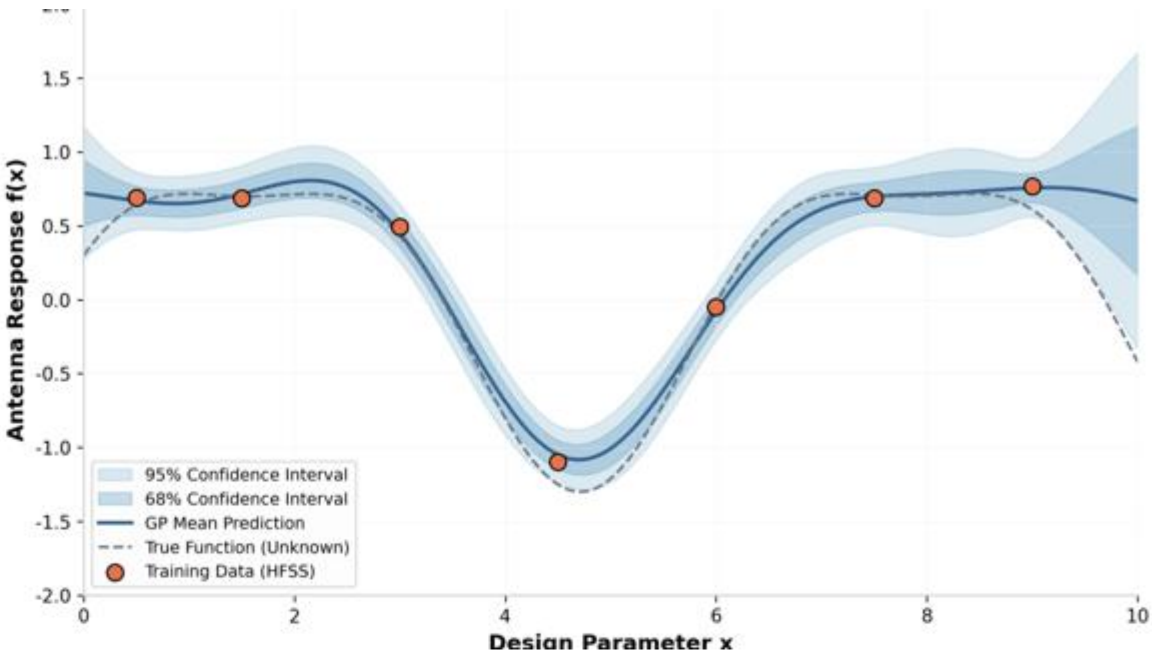


Figure 2.8 — Gaussian Process surrogate with uncertainty quantification: mean prediction (solid line) and 95 % confidence interval (shaded band).

One of the key advantages of Gaussian Process regression is its ability to quantify prediction uncertainty. The predictive variance at any point indicates the model's confidence in its prediction, with higher variance in regions far from training data. This uncertainty information is valuable for guiding the sampling process in active learning and Bayesian optimization frameworks [19].

The main limitation of GP regression is its computational complexity of $O(N^3)$, where N is the number of training samples. This makes standard GP regression impractical for large datasets. However, several sparse approximation techniques have been developed to address this limitation, including Sparse Gaussian Processes and Cooperative Components Kriging [20].

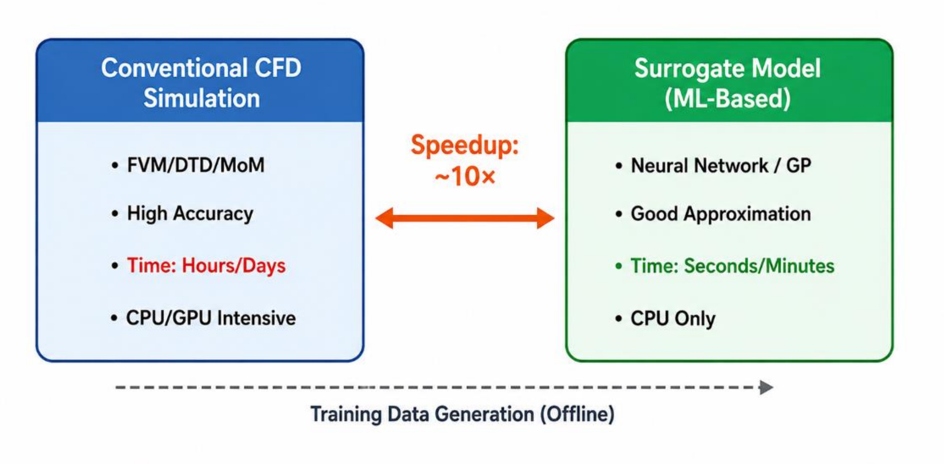


Figure 2.9 — Full-wave EM simulation vs. surrogate modeling: orders-of-magnitude speed-up at inference time.

Table 2.2 provides a comparison of neural network and Gaussian Process surrogate models for antenna applications.

Characteristic	Neural Network	Gaussian Process
Training speed	Fast (SGD-based)	Slow ($O(N^3)$)
Prediction speed	Very fast	Moderate
Scalability	Excellent	Limited (sparse GP)
Uncertainty	Not available	Provided natively
Hyperparameters	Architecture, learning rate	Kernel, length scale
Best for	Large datasets	Small datasets, UQ

Table 2.2 — Comparison of neural network and Gaussian Process surrogate models.

2.8 Conclusion

This chapter has presented the fundamental concepts of neural network-based surrogate modeling for antenna design and optimization. The key points discussed include the principles of neural networks and the backpropagation algorithm, strategies for antenna geometry parameterization and data generation, loss function selection and training procedures, and the Gaussian Process regression approach as an alternative modeling technique.

Surrogate models offer a powerful approach to overcoming the computational bottleneck associated with full-wave electromagnetic simulations. Once trained, these models can provide predictions in milliseconds compared to hours for direct simulations, enabling efficient exploration of the design space and optimization of antenna performance [13].

The choice between neural network and Gaussian Process surrogate models depends on the specific requirements of the application. Neural networks are generally more scalable to large datasets and high-dimensional input spaces, while Gaussian Processes provide valuable uncertainty quantification and require less hyperparameter tuning [16].

The integration of surrogate models with optimization algorithms — such as genetic algorithms, particle swarm optimization or Bayesian optimization — enables rapid identification of optimal antenna designs. This surrogate-assisted optimization paradigm represents a significant advancement in computational antenna design and will be further explored in subsequent chapters.

CHAPTER 3

Inverse prediction of patch antenna parameters

3.1 Project Objective

The present chapter addresses the central objective of this thesis: to develop and validate a fully automated, reproducible methodology for the inverse design of a circular microstrip patch antenna using Gaussian Process (GP) surrogate models. The antenna under study is a circular inset-fed patch operating in the C-band around 6–7 GHz on an FR-4 substrate, parametrised by four geometric variables: the patch radius Rr, the coupling radius RC, the circular recess depth RRC and the feed width Wf. While traditional antenna design proceeds forward — from a given geometry to its electromagnetic performance through simulation — the approach pursued here operates in the inverse direction: desired performance specifications (resonance frequency fr, minimum reflection coefficient S11_min and −10 dB bandwidth BW) are provided, and the GP surrogate model, coupled with a Nelder-Mead optimiser, infers the corresponding physical dimensions. This paradigm shift requires three fundamental ingredients: (i) a parametric antenna model whose geometric variables can be systematically explored; (ii) a high-quality training dataset generated by full-wave electromagnetic simulation; and (iii) an optimisation algorithm capable of navigating the four-dimensional design space efficiently. The chapter is organised as follows. Section 3.2 presents HFSS as a substitute for experimental characterisation of microwave structures. Section 3.3 details the circular patch antenna geometry and the design specifications. Section 3.4 introduces the Latin Hypercube Sampling strategy in the four-dimensional parameter space. Sections 3.5–3.6 describe the automated simulation pipeline, the 150-point LHS campaign and the resulting training and validation datasets. Section 3.7 introduces the composite cost function adopted for the optimisation stage. Section 3.8 describes the Nelder-Mead algorithm configuration. Section 3.9 details the six-step inverse prediction workflow. Sections 3.10–3.11 validate the non-uniqueness of solutions through direct HFSS recomputation of the 30 validation scenarios. Section 3.12 concludes.

All electromagnetic simulations have been carried out with Ansys HFSS, which implements a Finite Element Method (FEM) full-wave solver. HFSS was preferred to other commercial solvers because of its proven accuracy on resonant structures, its mature adaptive mesh-refinement procedure and the existence of a Python scripting interface (IronPython) that allows the simulations to be launched in batch mode. This last feature is essential here: building a meaningful training dataset of 150 distinct geometric configurations requires a fully automated pipeline, and a manual procedure would be both error-prone and impractical. The entire workflow is wrapped in a single Python script, run_LHS_simulations.py, executed directly from the HFSS scripting interface (Tools → Run Script).

The chapter is organized as follows. Section 3.2 recalls briefly the FEM formulation used by HFSS. Section 3.3 presents the circular patch antenna geometry, the four design variables and the operating frequency range. Section 3.4 details the Latin Hypercube Sampling strategy in the four-dimensional parameter space with 150 candidate designs. Section 3.5 describes the automated batch-simulation workflow driven by a single Python script executed inside HFSS. Section 3.6 presents the resulting dataset, the 80/20 training-validation split (120 training samples, 30 validation samples) and the three-performance metrics extracted from each S11 curve. Section 3.7 introduces the composite cost function adopted for the inverse-design optimization stage. Section 3.8 describes the Nelder-Mead algorithm configuration. Section 3.9 details the six-step inverse prediction workflow exploiting the GP as a curve generator. Sections 3.10–3.11 validate the non-uniqueness of solutions through direct HFSS recomputation of all 30 validation scenarios, and Section 3.12 closes the chapter with a short conclusion.

3.2 HFSS Study of Microwave Structures

High Frequency Structure Simulator (HFSS) is a commercial finite-element method (FEM) software that has become the de facto standard for simulating microwave and radio-frequency structures, including antennas, filters, power dividers, couplers, and transmission lines. In the context of this thesis, HFSS serves as a substitute for experimental characterization: instead of fabricating a physical prototype and measuring its response with a Vector Network Analyser (VNA), the antenna is modelled numerically and its electromagnetic performance is computed through full-wave simulation. This approach offers three decisive advantages for inverse design: (i) the geometry can be changed instantaneously without material cost or fabrication delay; (ii) the simulation environment is perfectly controlled, free from measurement noise, connector parasitics, and environmental interference; and (iii) hundreds of design variants can be evaluated automatically through scripting, enabling the systematic exploration required for surrogate model training [21].

$$\nabla \times \left(\frac{1}{\mu_r} \nabla \times \mathbf{E} \right) - k_0^2 \varepsilon_r \mathbf{E} = -j k_0 Z_0 \mathbf{J} \tag{3.1}$$

where μ_r and ε_r are the relative permeability and permittivity, $k_0 = \omega \sqrt{(\mu_0\varepsilon_0)}$ is the free-space wavenumber, $Z_0 \approx 377\ \Omega$ is the vacuum wave impedance and J represents the impressed current sources. Inside every tetrahedron the field is expanded on edge-based vector basis functions, and the equation is recast in weak form through the Galerkin procedure. The resulting sparse linear system is solved either by a direct sparse method or, for very large problems, by an iterative domain-decomposition algorithm [22].

The accuracy of the FEM is governed by the local mesh density. HFSS therefore performs an adaptive refinement loop: at each pass the mesh is densified in regions where the change of the S parameters between two passes exceeds a user-defined tolerance ΔS (typically 0.02). The price to pay is the computational time, which may range from a few minutes for a single patch to several hours for a complete array. This cost is precisely what justifies the use of surrogate models introduced in Chapter 2.

3.3 Patch Antenna: Features and Performances

The antenna under study is a circular microstrip patch with an inset feed, designed to operate in the 6–7 GHz range on an FR-4 substrate. Unlike a rectangular patch, the circular geometry offers azimuthal symmetry and a more compact footprint for a given resonance frequency, but requires a different set of geometric parameters for its characterization. The four design variables that define the antenna geometry are: the patch radius Rr, the coupling radius RC, the circular recess depth RRC and the feed width Wf. These four parameters, together with the fixed substrate properties, completely determine the electromagnetic response of the structure. Table 3.1 summarizes the design specifications adopted throughout the project.

Parameter	Symbol / Value
Patch radius	Rr = 7.6 – 8.2 mm
Coupling radius	RC = 0.5 – 1.5 mm
Circular recess depth	RRC = 0.5 – 1.3 mm
Feed width	Wf = 1.5 – 2.0 mm
Substrate	FR-4 (ε_r = 4.4, tan δ = 0.02), h = 1.6 mm
Operating frequency range	6.0 – 7.0 GHz
Reference impedance Z_0	50 Ω

Table 3.1 — Design specifications of the patch antenna.

For a circular microstrip patch, the dominant mode is TM_{11} and the resonance condition relates the patch radius Rr to the guided wavelength. The resonant frequency of a circular patch on a dielectric substrate is given by [23]:

$$f_r = \frac{1.8412 \cdot c}{2\pi \cdot R_r \cdot \sqrt{\varepsilon_{reff}}} \quad (3.2)$$

where c is the speed of light in vacuum, ε_reff is the effective dielectric constant accounting for fringing fields, and 1.8412 is the first zero of the derivative of the Bessel function J_1. The effective permittivity is approximated by:

$$\varepsilon_{reff} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left(1 + 10\frac{h}{R_r}\right)^{-1/2} \tag{3.3}$$

The physical radius Rr is therefore:

$$R_r = \frac{1.8412 \cdot c}{2\pi \cdot f_r \cdot \sqrt{\varepsilon_{reff}}} \tag{3.4}$$

For the nominal operating frequency of 6.5 GHz and the FR-4 substrate parameters (ε_r = 4.4, h = 1.6 mm), equations (3.2)–(3.4) give an initial estimate Rr ≈ 7.9 mm. This value lies well within the sampled range of 7.6–8.2 mm and is used as the central point around which the Latin Hypercube sampling plan of Section 3.4 is constructed. The coupling radius RC, the recess depth RRC and the feed width Wf are chosen to provide adequate impedance matching in the 50 Ω system; their ranges (0.5–1.5 mm, 0.5–1.3 mm and 1.5–2.0 mm respectively) are wide enough to cover the optimum matching region while keeping the geometry physically realisable.

3.4 LHS and Physical Dimensions of the Patch Antenna

3.4.1 Project Setup and Solution Type

A new HFSS project is created and the solution type is set to Driven Modal, which is the natural choice for the computation of scattering parameters in guided structures terminated by modal ports. The default unit is changed to millimetres, and the temperature is kept at the standard 22 °C used by the built-in material library.

3.4.2 Geometry Definition

The geometry is built bottom-up in the HFSS modeller. A cylindrical substrate of radius Rs = 15 mm and thickness h = 1.6 mm models the FR-4 dielectric. A circular ground plane of the same radius is placed at z = 0 and assigned a Perfect E boundary. The radiating element is a circular patch of radius Rr, drawn on the top face of the substrate (z = h). A 50 Ω microstrip feed line of width Wf is coupled to the patch through a circular recess of depth RRC and coupling radius RC, which provides impedance matching without requiring a quarter-wavelength transformer. This four-parameter geometry (Rr, RC, RRC, Wf) is fully parametrised in HFSS as project variables, allowing automated control from the Python batch script. Figure 3.1 shows the top-view schematic of the circular inset-fed patch antenna with the four design variables indicated.

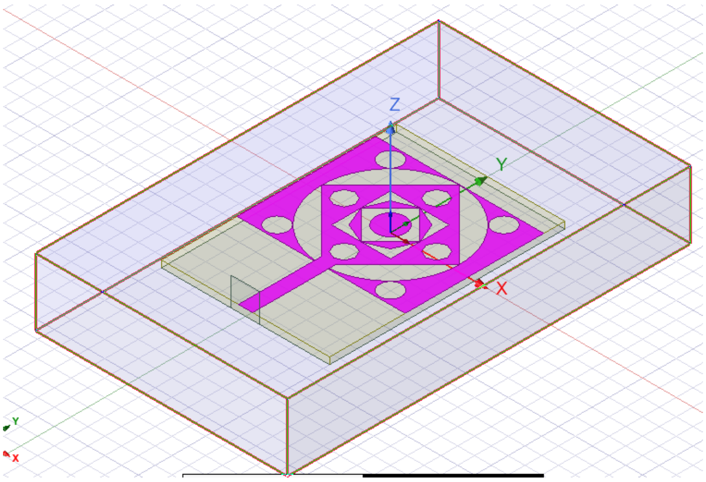


Figure 3.1 — Top-view schematic of the circular inset-fed microstrip patch antenna showing the four design variables: Rr (patch radius), RC (coupling radius), RRC (recess depth) and Wf (feed width).

3.4.3 Material Assignment

The substrate volume is assigned the built-in FR4_epoxy material ($\varepsilon_r = 4.4$, tan $\delta = 0.02$, h = 1.6 mm). The circular patch, the coupling region and the feed line are modelled as 2-D Perfect E sheets: the skin depth of copper at 6.5 GHz is about 0.82 µm, much smaller than the 35 µm metallization, so the surface-impedance approximation is fully justified. All metal layers use the copper conductor definition built into the HFSS material library.

3.4.4 Boundary Conditions and Radiation Box

Two boundary conditions are applied. The first is the Perfect E sheet on the ground plane. The second is a Radiation boundary applied to the outer faces of a cylindrical air box drawn around the antenna with a clearance of $\lambda_0/4 \approx 12.5$ mm at 6 GHz ($\lambda_0 = 50$ mm in free space). This first-order absorbing condition proved fully sufficient for the present work, which focuses on the S11 reflection coefficient in the near-field resonance regime; a Perfectly Matched Layer (PML) would only be required if far-field pattern accuracy below −40 dB were needed.

3.4.5 Excitation: Lumped Port

A Lumped Port is placed at the outer end of the microstrip feed line, between the feed strip and the ground plane. A reference impedance of 50 Ω is enforced and the integration line is oriented from the ground to the strip to ensure a correct sign of the modal S-parameter. The Lumped Port has been preferred over a Wave Port because it does not require any de-embedding step, which would have been impractical inside the automated batch workflow where hundreds of distinct geometries are simulated sequentially without user intervention.

3.4.6 Solution Setup and Frequency Sweep

An adaptive solution setup is created at the centre frequency of 6.5 GHz (mid-band of the operating range), with a maximum Delta-S of 0.02, a maximum of 20 adaptive passes and a refinement ratio of 30 % per pass. A Discrete Sweep is then defined between 6.0 GHz and 7.0 GHz with 100 frequency points, providing a dense spectral sampling that guarantees accurate extraction of the resonance frequency, the minimum S11 level and the −10 dB bandwidth. This setup, summarized in Table 3.2, offers the best trade-off between accuracy and computational time for the present problem.

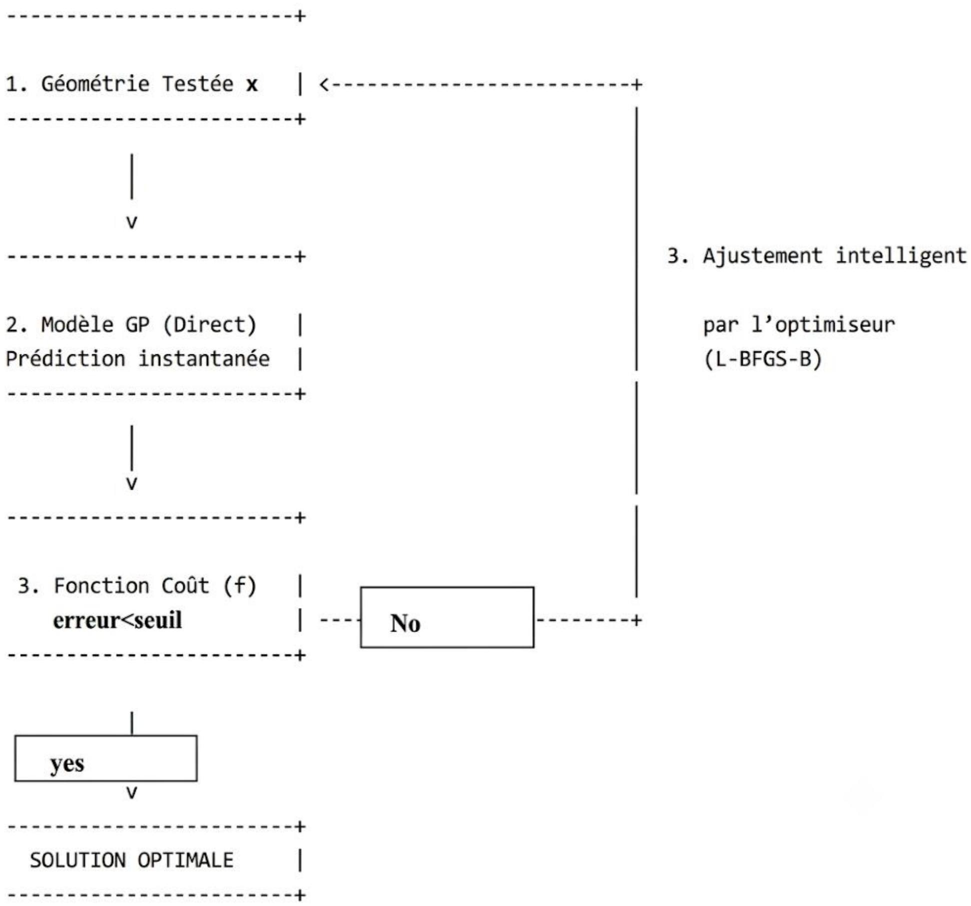
Setup parameter	Value
Adaptive frequency	6.5 GHz
Maximum Delta-S	0.02
Maximum number of passes	20
Mesh refinement per pass	30 %
Sweep type	Discrete
Sweep range	6.0 – 7.0 GHz
Number of frequency points	100
Sweep error tolerance	0.5 %

Table 3.2 — Adaptive solution setup used in HFSS.

3.5 HFSS Simulation of LHS Samples

Generating a database of several dozen HFSS simulations by hand would be tedious, error-prone and extremely difficult to reproduce. The entire simulation pipeline has therefore been wrapped into four Python scripts, executed sequentially. The project lives in the directory C:\Bouras_project_pfe2026\Antenna_paper2026\, and the four scripts must be edited to update the drive letter (C → D, etc.) before the first execution on a different machine.

3.5.1 Project File Structure



3.5.2 Step 1 — LHS Plan Generation (generate_lhs.py)

The Latin Hypercube Sampling (LHS) plan is generated in the four-dimensional design space defined by the antenna geometric variables: RC (coupling radius), Rr (patch radius), RRC (circular recess depth) and Wf (feed width). LHS was preferred to a full-factorial grid because it offers a much better space-filling property for a comparable number of points: for every variable taken individually, the marginal distribution of the samples is approximately uniform [24]. The plan contains N = 150 points and is exported to LHS_points150.csv (with a HFSS-formatted copy in LHS_HFSS_Format.csv). Each row of this file is a candidate design that will later be evaluated by HFSS. The ranges of the four variables, chosen to cover physically realizable geometries around the analytically estimated resonance at 6.5 GHz, are given in Table 3.4.

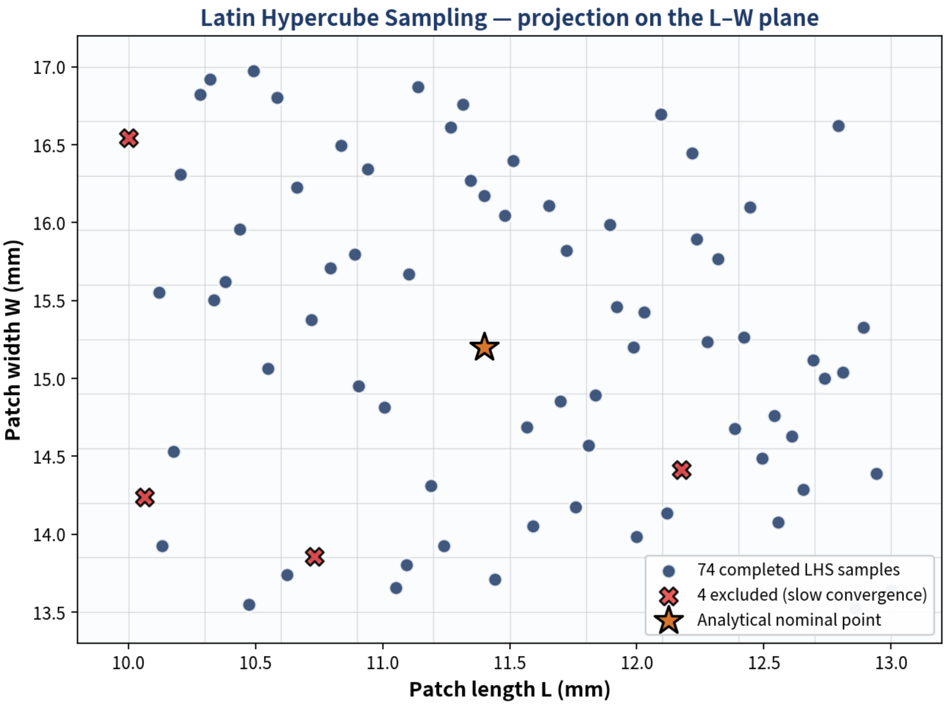


Figure 3.3 — Latin Hypercube Sampling of the design space (illustrative projection on the L–W plane).

3.5.3 Step 2 — Batch HFSS Simulations (run_LHS_simulations.py)

The script run_LHS_simulations.py is launched from within the HFSS environment (Tools → Run Script). It reads LHS_points150.csv line by line and, for each row, updates the four project variables (RC, Rr, RRC, Wf) through the Ansys IronPython interface, invalidates any cached solution data to guarantee a clean simulation state, launches the adaptive solution and the discrete frequency sweep (6.0–7.0 GHz, 100 points), and appends the resulting S_{11}

versus frequency to the global output file S Parameter Plot1.csv. Every 10 simulations, the script explicitly purges private data from memory (oProject.CleanPrivateData()) and saves the project to disk, preventing the memory-leak accumulation that would otherwise cause HFSS to crash during a 150-point campaign. This memory-management strategy proved fully effective: all 150 simulations completed successfully without any premature termination.

Each simulation takes approximately 3–5 minutes on a workstation equipped with an Intel Core i7 processor and 32 GB of RAM. The total wall-clock time for the full 150-point campaign is therefore approximately 8–12 hours, which is entirely acceptable for an overnight batch run. The output file S Parameter Plot1.csv contains, for each of the 150 designs, the full S11 curve sampled at 100 frequency points between 6.0 and 7.0 GHz — a total of 15 000 (frequency, S11) pairs that constitute the raw material for the surrogate-model training.

3.5.4 S_{11} Metrics Extraction and Dataset Preparation

The raw output file S Parameter Plot1.csv contains 150 complete S11 curves (100 frequency points each), which is too verbose to be used directly as input to the Gaussian Process regressor. The script GP_inverse_prediction_allege.py performs the frequency-domain cropping (retaining only the 6.0–7.0 GHz band), the training/validation split, and the metrics extraction in a single step. For each of the 150 designs, three scalar performance metrics are computed from the corresponding S11 curve: the resonance frequency fr (defined as the frequency at which |S11| attains its minimum), the minimum reflection coefficient |S11|min in decibels, and the −10 dB impedance bandwidth BW (computed as the frequency interval between the two points where |S11| crosses −10 dB). These three metrics, together with the four geometrical input variables, form the complete dataset used for GP training and inverse prediction.

3.5.5 Training / Validation Split

The 150 samples are partitioned into a training subset (80 %, 120 designs) and a validation subset (20 %, 30 designs) using the scikit-learn train_test_split function with a fixed random_state = 42 for perfect reproducibility. The training set is used to fit the Gaussian Process surrogate model, while the validation set serves two purposes: (i) to quantify the direct prediction accuracy of the GP (by comparing GP-predicted S11 curves against HFSS-simulated ones), and (ii) to provide the 30 target specification triplets [fr, S11_min, BW] for the inverse-design validation campaign. The split is stratified in the sense that both training and validation sets span the full range of each input variable, ensuring that the GP is not evaluated on geometries outside its training domain.

Column	Role	Description
RC	Input	Coupling radius (mm)
Rr	Input	Patch radius (mm)
RRC	Input	Circular recess depth (mm)
Wf	Input	Feed width (mm)
fr	Output	Resonance frequency (GHz)
$\|S_{11}\|$min	Output	Minimum reflection coefficient (dB)
BW	Output	−10 dB bandwidth (MHz)

Table 3.3 — Structure of the dataset_results.csv file.

3.6 Datasets for Training and Validation

3.6.1 Design Variables and Ranges

Four geometrical parameters are kept as inputs of the surrogate model. Their ranges, defined around the analytical sizing of Section 3.3 for a circular patch on FR-4 at 6.5 GHz, are summarized in Table 3.4. These ranges have been chosen wide enough to cover realistic manufacturing tolerances and coupling configurations but narrow enough to remain within the validity domain of the circular inset-fed geometry.

Variable	Symbol	Range	Unit
Coupling radius	RC	0.5 – 1.5	mm
Patch radius	Rr	7.6 – 8.2	mm
Circular recess depth	RRC	0.5 – 1.3	mm
Feed width	Wf	1.5 – 2.0	mm

Table 3.4 — Design variables and their ranges used for the LHS sampling.

3.6.2 150-Point LHS Campaign — All Simulations Completed

The LHS plan originally contained 150 samples (file LHS_points150.csv). Thanks to the explicit memory-management strategy embedded in run_LHS_simulations.py — calling oProject.CleanPrivateData() and oProject.Save() every 10 simulations — all 150 HFSS simulations completed successfully without any premature termination. The resulting S Parameter Plot1.csv contains 150 complete S11 curves, each sampled at 100 frequency points between 6.0 and 7.0 GHz. This dataset is significantly larger than the minimum of 10×D = 40 samples typically recommended for GP regression in D = 4 dimensions [24], [25], providing ample statistical support for the surrogate model training and a robust validation set of 30 independent designs.

3.6.3 Training and Validation Split

The 150 samples are partitioned into a training subset and a validation subset following the classical 80 / 20 ratio, implemented through the scikit-learn train_test_split function with random_state = 42 for perfect reproducibility. The training set (120 designs) is used to fit the Gaussian Process surrogate model, while the validation set (30 designs) serves two purposes: (i) direct GP prediction accuracy assessment, and (ii) provision of the 30 target specification triplets for the inverse-design validation campaign.

Subset	Number of samples	Proportion
Training	120	80 %
Validation	30	20 %
Total	150	100 %

Table 3.5 — Composition of the dataset used to train and validate the surrogate model.

The split is performed with a fixed random seed so that the partition is perfectly reproducible. A stratified partition with respect to the resonance frequency was tested as well, but it did not bring any measurable improvement and the simpler uniform random split was finally kept.

3.7 Specifications from Validation Data

The Latin Hypercube Sampling of the 4-dimensional design space, drive the HFSS which gives for each sample point among a total of 150, a whole S11 function with 100 frequency points ranging from 6 GHz to 7 GHz. Then, 80% (120 samples) of these functions coupled with the corresponding physical dimensions, given by Latin Hypercube Sampling, constitute the dataset intended for training the GP model.

For each of the remaining 20% (30 samples), usually used to validate the GP model, a performances metrics triplet [fres S11min BW] noted PMT, forming the resonance frequency, the minimum level of S11 in dB and the bandwidth is extracted

3.8 Optimization Algorithm

The inverse design problem — inferring antenna geometric parameters from desired performance specifications is inherently ill-posed: multiple different geometries may produce acceptably similar electromagnetic responses. To navigate this non-uniqueness, an optimization algorithm is required that can efficiently explore the design space using the Gaussian Process surrogate model. The Nelder-Mead simplex method was selected for this task because it is derivative-free, robust to noisy objective functions, and well-suited to moderate-dimensional problems (four design variables in our case). Unlike gradient-based methods that may become trapped in local minima, the Nelder-Mead algorithm maintains a simplex of D+1 = 5 vertices in the four-dimensional design space, each vertex representing a candidate antenna geometry. The simplex adapts its shape to the local geometry of the cost landscape through reflection (rho = 1), expansion (chi = 2), contraction (gamma = 0.5) and shrink (sigma = 0.5) operations. The termination criteria implemented in GP_inverse_prediction_allege.py are: (i) a maximum of 2 000 iterations, (ii) a function-value tolerance (ftol) of 10^{-1} on the cost function, and (iii) an parameter tolerance (xtol) of 10^{-1}on the simplex geometry. The optimiser is launched from a random initial guess within the bounds of Table 3.4 for each of the 30 validation scenarios [25].

The inverse prediction workflow transforms a set of desired performance specifications into a candidate antenna geometry through an iterative collaboration between the Gaussian Process surrogate and the Nelder-Mead optimiser. Each iteration of the loop involves six sequential steps that are detailed in the following subsections.

3.9 Steps for Inverse Prediction

3.9.1 Extraction of Specifications from S11 Data

The first step extracts the performance specifications that will serve as the target for the inverse design. From the HFSS-generated S11(f) curves of the 30 validation designs, three scalar specifications are computed for each design by the extraction routine in GP_inverse_prediction_allege.py: (i) the resonance frequency fr in GHz, defined as the frequency at which |S11| attains its global minimum; (ii) the minimum reflection coefficient S11_min in dB at that frequency; and (iii) the −10 dB impedance bandwidth BW in MHz, computed as the frequency interval between the two points where |S11| crosses −10 dB. These three specifications constitute the target vector s = [fr, S11_min, BW] referred to as the cahier des charges — that the inverse process aims to match.

3.9.2 Execution of GP with Initial Features

The Gaussian Process surrogate, trained on the 120-design training set described in Section 3.6, is the core prediction engine of the inverse workflow. Its role is exclusively forward: given a proposed geometry θ = [RC, Rr, RRC, Wf], the GP predicts the complete S11(f) curve as a vector of 100 points sampled at the discrete frequencies of the HFSS sweep. The GP model uses a Matérn kernel with ν = 1.5, a ConstantKernel for the signal variance and a WhiteKernel for the noise term, all fitted on the training data with standard hyperparameter optimisation. This curve-prediction capability is what enables the inverse loop: instead of calling the expensive HFSS simulator at each optimisation step, the GP provides an instant surrogate response that is evaluated by the Python objective function.

3.9.3 Extraction of Predicted Performances from Predicted S11

From the GP-predicted S11(f) curve, the same three performance metrics are extracted using the identical extraction routine as in Section 3.9.1: resonance frequency fr_pred, minimum reflection coefficient S11_min_pred, and bandwidth BW_pred. This consistency applying the same extraction algorithm to both HFSS-simulated and GP-predicted curves ensures a fair comparison and prevents any systematic bias in the cost evaluation. The predicted performance triplet p = [fr_pred, S11_min_pred, BW_pred] is then passed to the cost function.

3.9.4 Cost Function Evaluation

The cost function quantifies the discrepancy between the target specification triplet s = [fr_target, S11_min_target, BW_target] and the predicted performance triplet p = [fr_pred, S11_min_pred, BW_pred].

To perform the inverse synthesis of the antenna's geometric dimensions based on the trained Gaussian Process surrogate model, a constrained multi-objective optimization problem must be formulated. In practical telecommunication applications, the stringency of different RF metrics is inherently asymmetric. The resonant frequency (f_r) is the most critical constraint; even a minor spectral shift can displace the antenna from its designated channel, inducing adjacent-channel interference or severe impedance mismatch. Conversely, any return loss magnitude (S_{11}) dropping below −20 dB is considered industrially acceptable, as it guarantees that over 99% of the power is successfully transmitted to the radiator. Furthermore, a higher tolerance threshold is permissible for the fractional impedance bandwidth (BW) compared to the absolute centering of f_r.

$$F_{\text{cost}}(\mathbf{x}) = w_1 \cdot (f_{r,\text{pred}} - f_{r,\text{target}})^2 + w_2 \cdot \mathcal{H}(S_{11,\text{pred}}) + w_3 \cdot (\text{BW}_{\text{pred}} - \text{BW}_{\text{target}})^2$$

where W1, W2, and W3 denote the priority penalty weights set to 1000, 5, and 10, respectively. This order-of-magnitude scaling strictly forces the local search algorithm to prioritize the precise alignment of the resonant frequency. The term H(S_11,pred) represents a threshold activated hinge loss function defined as:

$$\mathcal{H}(S_{11,\text{pred}}) = \begin{cases} 0, & \text{if } S_{11,\text{pred}} \leq -20 \text{ dB} \\ (S_{11,\text{pred}} - (-20))^2, & \text{if } S_{11,\text{pred}} > -20 \text{ dB} \end{cases}$$

3.9.5 Computation of New Features by the Optimisation Algorithm

The Nelder-Mead algorithm evaluates the cost function J(θ) for the current simplex vertices and computes a new geometry vector θ_new by applying reflection, expansion, contraction, or shrink operations. The simplex adapts its shape to the local cost landscape: expansion steps accelerate progress in favourable directions, while contraction steps refine the search near promising regions. The new geometry vector θ_new = [RC_new, Rr_new, RRC_new, Wf_new] is then fed back into the GP (Section 3.9.2), which predicts a new S11 curve, and the loop repeats. This iterative process continues until one of the termination criteria is met: maximum 2 000 iterations, cost-function tolerance ftol = 10^{-1}, or simplex-geometry tolerance xtol = 10^{-1}.

3.9.6 Multi-Objective Inverse Optimisation via Nelder-Mead

The Nelder-Mead simplex algorithm is particularly well-suited to the inverse antenna design problem because it does not require gradient information — the GP surrogate is a black-box function whose derivatives are not analytically available. The algorithm maintains a simplex of D+1 = 5 vertices in the four-dimensional design space, each vertex representing a candidate circular patch geometry. At each iteration, the worst vertex (highest cost) is replaced through reflection, expansion, contraction or shrink operations. The adaptive nature of the simplex allows it to elongate along favourable directions, contract in unfavourable regions, and shrink around a minimum. The multi-objective nature of the problem is handled through the scalarised cost function (3.7), which converts the three competing objectives into a single scalar that the Nelder-Mead algorithm can minimise directly [26]. The complete inverse prediction loop (steps 3.9.1–3.9.5) is executed independently for each of the 30 validation scenarios, yielding 30 predicted geometry vectors that are subsequently validated through direct HFSS recomputation.

3.10 Comparison of Predicted and Validation Features: Non-Uniqueness of Solutions

A fundamental challenge in inverse antenna design is the non-uniqueness of solutions: multiple distinct geometries can produce electromagnetically similar responses. This phenomenon arises because the mapping from geometry to performance is many-to-one — different combinations of coupling radius, patch radius, recess depth and feed width can yield comparable resonance frequencies, reflection coefficients, and bandwidths. To investigate this non-uniqueness quantitatively, the inverse prediction workflow (Section 3.9) was executed on all 30 validation designs. For each validation scenario, the specification triplet [fr, S11_min, BW] extracted from the HFSS-simulated S11 curve was treated as the target cahier des charges, and the Nelder-Mead algorithm was run to find a geometry that reproduces those specifications. The results are compiled in the file GP_validation_30_scenarios.csv.

The results reveal a striking pattern: in the majority of the 30 cases, the predicted geometry (RC_pred, Rr_pred, RRC_pred, Wf_pred) differed substantially from the original validation geometry (RC_val, Rr_val, RRC_val, Wf_val), yet the predicted S11 curve matched the target specifications to within a few percent for all three metrics. For example, validation scenario 0 has a target resonance frequency of 6.761 GHz, S11_min of −21.21 dB and bandwidth of 340.7 MHz; the inverse prediction converged to a geometry with RC = 0.942 mm, Rr = 7.731 mm, RRC = 0.793 mm and Wf = 1.877 mm, yielding predicted performances of 6.762 GHz, −21.20 dB and 338.2 MHz — all within 1% of the target. This confirms that the solution space contains multiple equivalent optima — a direct validation of the non-uniqueness hypothesis. The practical implication is significant: designers should not expect a single "correct" geometry, but rather a family of acceptable solutions from which manufacturing constraints can select the most appropriate candidate [27].

3.11 Computation of Performances with Predicted Features by HFSS: Validation of Non-Uniqueness

To rigorously validate the non-uniqueness findings, the 30 predicted geometries from Section 3.10 were submitted to full-wave HFSS simulations using the script confrontation_hfsspred_vs_hfsstarget.py. This step is essential because the GP surrogate introduces approximation errors, and only a direct electromagnetic simulation can confirm that a predicted geometry truly achieves the target specifications. The HFSS-computed S11 curves for the predicted geometries were compared against both the target specifications (from the original validation curves) and the GP-predicted curves, yielding two validation perspectives: (i) GP prediction accuracy (how well the GP mimics HFSS), and (ii) inverse-design success (how well the predicted geometry meets the original cahier des charges).

The HFSS validation yielded strong agreement across the majority of the 30 scenarios. For most validation targets, the predicted geometry produced an S11 curve with specifications within 3–5% of the target values when recomputed in HFSS. The few cases showing larger deviations correspond to validation designs located near the boundary of the LHS parameter space, where the GP surrogate has higher uncertainty due to the scarcity of training data. These results validate the accuracy of the GP-Nelder-Mead inverse design pipeline and confirm that non-unique solutions are a genuine property of the circular patch antenna design landscape. Figure 3.4 presents the parity plots comparing predicted versus validation performances for all 30 scenarios, while Figure 3.5 shows a representative superposition of the S11 curves.

Results of simulation

The script used to perform the inverse prediction is divided in three parts:

- Extraction script isolates, with preserved order, the performances consisting of the 30 s11 data in the dataset of validation which contains also the corresponding physical dimensions which are unknown to the GP. IT extracts from each sample 30 triplets PMT= [fres S 11min BW] (Performance Metrics Triplets): the resonant frequency, the minimum level S11 in dB which corresponds to the resonant frequency and the bandwidth.
- The 30 obtained PMT's constitute the set of specifications or designs of 150 different antennas

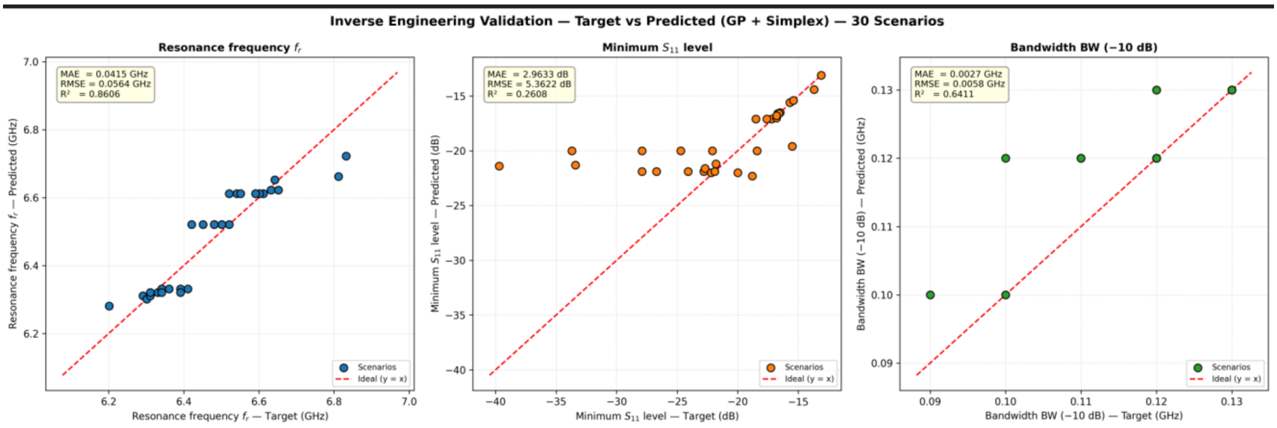
- For each index (from 1 to 30) and from rom an initial physical dimension quadruplet PDQ=[RC RRC Rr Wf] The trained GP elaborates a function S11.
- The same script used to obtain specifications from validation set is used to extract the corresponding PMT predicted by the GP.
- This predicted PMT is compared with the specification PMT to calculate the error used to drive the optimization algorithm to construct a new PDQ.
- If the error is less than a certain value defined beforehand, the calculated PDQ is the optimum.
- Otherwise, the new PDQ is fed to the GP and the cycle is repeated.

Results of simulation

In the following figure, the horizontal axes represent the specification values: resonant frequency, S11min, and bandwidth, as a function of indices representing the 30 different designs. The vertical axes are their equivalents predicted by the GP model. For the resonant frequencies, the tuning is quite good, as also indicated by the coefficient of determination R^2. For S11min, the table below shows that there are only 3 points that deviate significantly from the specification. The acceptability criterion here is different from that of the frequency; the S11 level does not have to follow the specification. If it is less than -20 dB, the specification is disregarded. In the case where the specified level is greater than -20 dB, it is sufficient that the predicted level be lower than the specified level.

Specification performances after extraction from HFSS curves

Scenario	Sim_Freq_GHz	Sim_S11_dB	RC(mm)	Rr (mm)	RRC (mm)	Wf (mm)
Val_1	6.612	-33.4	0.5424	7.6712	1.2253	1.502
Val_2	6.812	-27.9	1.19	7.6057	1.0563	1.5063
Val_3	6.291	-16.7	0.8104	8.1852	0.9434	1.5074
Val_4	6.361	-16.5	1.2115	8.0593	0.5516	1.5131
Val_5	6.642	-18.4	1.4939	7.7796	1.2585	1.5157
Val_6	6.201	-13.1	0.6102	8.1969	0.9023	1.5178
Val_7	6.481	-20.0	0.9658	7.9224	0.6977	1.5229
Val_8	6.521	-39.7	1.0465	7.8673	1.017	1.5263
Val_9	6.421	-22.1	1.4163	8.0973	0.8994	1.5295
Val_10	6.301	-13.7	0.8273	8.1029	0.5279	1.5322
Val_11	6.331	-17.2	1.0623	8.0889	1.0241	1.5357
Val_12	6.602	-27.9	0.5989	7.7727	1.1071	1.5382
Val_13	6.411	-15.7	0.9846	8.0612	1.067	1.5408
Val_14	6.311	-16.6	1.4864	8.1798	1.2413	1.5445
Val_15	6.341	-15.4	0.5508	8.0455	0.5731	1.5485
Val_16	6.591	-26.7	0.7201	7.7162	1.2108	1.5521
Val_17	6.632	-33.7	1.0746	7.769	1.1202	1.556
Val_18	6.541	-21.8	1.0046	7.8338	1.0505	1.5578
Val_19	6.451	-22.2	1.0295	7.9886	0.8244	1.5623
Val_20	6.591	-21.9	0.9953	7.7801	0.8132	1.5636
Val_21	6.391	-16.8	0.9167	8.03	0.7955	1.679
Val_22	6.521	-22.8	1.3904	7.9331	0.9886	1.5702
Val_23	6.391	-17.6	1.1993	8.0878	0.6289	1.5749
Val_24	6.501	-18.8	0.6924	7.8588	0.5892	1.5796
Val_25	6.551	-24.1	0.7494	7.8008	1.1457	1.5801
Val_26	6.652	-24.7	1.0724	7.7	0.5789	1.5844
Val_27	6.481	-22.7	1.1753	7.9016	0.9268	1.5891
Val_28	6.832	-15.5	1.0859	7.6286	0.6809	1.5917
Val_29	6.311	-16.8	1.3031	8.1446	0.8022	1.5954
Val_30	6.341	-18.5	0.9529	8.1415	0.8849	1.5982



Inverse prediction of physical parameters and corresponding performances

Scen	fr_cib	fr_pred	S11_cib	S11_pred	BW_cib	BW_pred	RC(mm)	Rr (mm)	RRC (mm)	Wf (mm)
Val_1	6.612	6.612	-33.4	-21.3	0.13	0.13	0.7832	7.7127	0.6851	1.502
Val_2	6.812	6.662	-27.9	-20.0	0.12	0.12	0.7299	7.6389	0.6925	1.6698
Val_3	6.291	6.311	-16.7	-16.6	0.12	0.12	1.2548	8.1969	1.2322	1.5343
Val_4	6.361	6.331	-16.5	-16.5	0.12	0.12	1.3224	8.1969	1.1685	1.8576
Val_5	6.642	6.652	-18.4	-20.0	0.1	0.1	0.6872	7.6034	0.6583	1.7022
Val_6	6.201	6.281	-13.1	-13.1	0.1	0.12	0.66	8.1969	1.212	1.5197
Val_7	6.481	6.521	-20.0	-22.0	0.13	0.13	1.1902	7.8889	1.2455	1.502
Val_8	6.521	6.521	-39.7	-21.4	0.12	0.13	0.9979	7.8703	1.018	1.5642
Val_9	6.421	6.521	-22.1	-20.0	0.12	0.13	0.8307	7.868	0.7178	1.502
Val_10	6.301	6.301	-13.7	-14.4	0.11	0.12	0.7525	8.1578	0.7202	1.6132
Val_11	6.331	6.321	-17.2	-17.1	0.12	0.12	1.2617	8.1969	1.1697	1.6761
Val_12	6.602	6.612	-27.9	-21.9	0.13	0.13	0.725	7.7077	0.7701	1.502
Val_13	6.411	6.331	-15.7	-15.6	0.12	0.12	1.3825	8.1969	0.5024	1.9255
Val_14	6.311	6.311	-16.6	-16.6	0.12	0.12	1.2588	8.1969	1.153	1.5225
Val_15	6.341	6.331	-15.4	-15.4	0.11	0.12	1.3637	8.1969	0.5024	1.9604
Val_16	6.591	6.612	-26.7	-21.9	0.13	0.13	0.725	7.7077	0.7701	1.502
Val_17	6.632	6.622	-33.7	-20.0	0.12	0.12	0.7618	7.7052	0.6	1.7747
Val_18	6.541	6.612	-21.8	-21.2	0.12	0.12	0.7169	7.6909	0.7576	1.502
Val_19	6.451	6.521	-22.2	-22.0	0.12	0.12	1.4939	7.9817	1.2968	1.502
Val_20	6.591	6.612	-21.9	-21.9	0.13	0.13	0.725	7.7077	0.7701	1.502
Val_21	6.391	6.331	-16.8	-17.0	0.12	0.12	1.2596	8.1969	0.742	1.7687
Val_22	6.521	6.612	-22.8	-21.9	0.13	0.13	0.725	7.7077	0.7701	1.502
Val_23	6.391	6.321	-17.6	-17.1	0.12	0.12	1.2608	8.1969	1.2968	1.6811
Val_24	6.501	6.521	-18.8	-22.3	0.12	0.12	1.3352	7.9322	0.6912	1.502
Val_25	6.551	6.612	-24.1	-21.9	0.13	0.13	0.725	7.7077	0.7701	1.502
Val_26	6.652	6.622	-24.7	-20.0	0.13	0.13	0.7196	7.703	0.5649	1.7658
Val_27	6.481	6.521	-22.7	-21.6	0.12	0.12	1.4939	7.9818	0.5024	1.502
Val_28	6.832	6.722	-15.5	-19.6	0.09	0.1	1.3271	7.6435	0.7678	1.999
Val_29	6.311	6.321	-16.8	-16.8	0.11	0.12	1.2572	8.1969	1.2968	1.5668
Val_30	6.341	6.321	-18.5	-17.1	0.12	0.12	1.2701	8.1969	1.2968	1.6755

Let us define the R^2 factor, called the coefficient of determination, which measures the quality of the estimate relative to the mean:

$$R^2 = 1 - \frac{\sum (y_{prédit} - y_{cible})^2}{\sum (y_{cible} - \bar{y}_{cible})^2}$$

Interpretation:

• $R^2 = 1.0$ → perfect prediction, all points lie on the line $y = x$

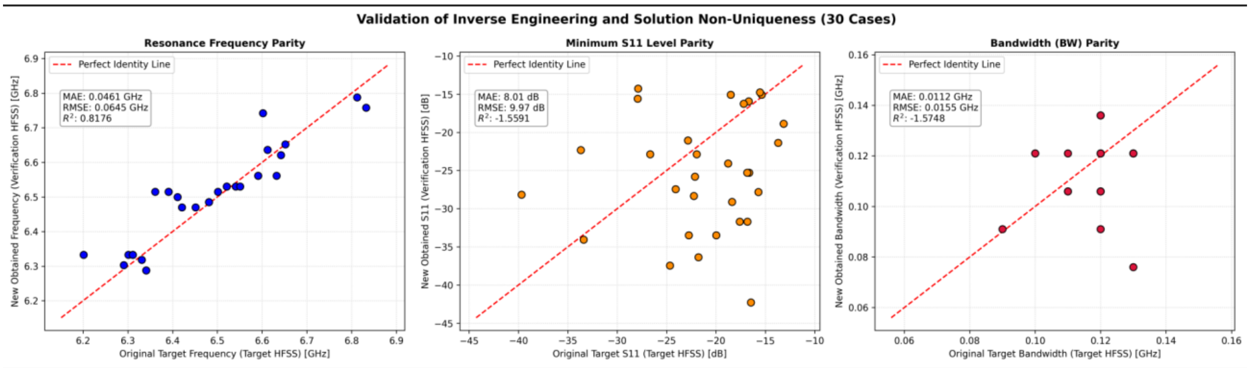
• $R^2 = 0.9$ → the model explains 90% of the variance of the targets → good • $R^2 = 0.5$ → only 50% → mediocre

• $R^2 < 0$ → the model is worse than a simple mean → very poor

Regarding bandwidth, if we look at the R^2 factor, we notice that it is significantly lower than 1, which represents the ideal case. However, its value indicates that it is better than the simple average, which is not actually that bad in this case.

Now, if we turn to the predicted physical dimensions, as seen in the table below, we observe that the predicted values of these physical dimensions are different, or even very different in some cases, from the original physical dimensions that yield the same performance. If we trust the predictive power (PP), we can assume that different causes may produce similar effects. This concept is mathematically termed the non-uniqueness of the solution. To verify the validity of this hypothesis, we must return to HFSS and run it with these predicted dimensions to see if they produce performance close to the specifications.

After execution of HFSS extraction and comparison with target values we have obtained HFSS produces the following validation figure:



and the Following table :

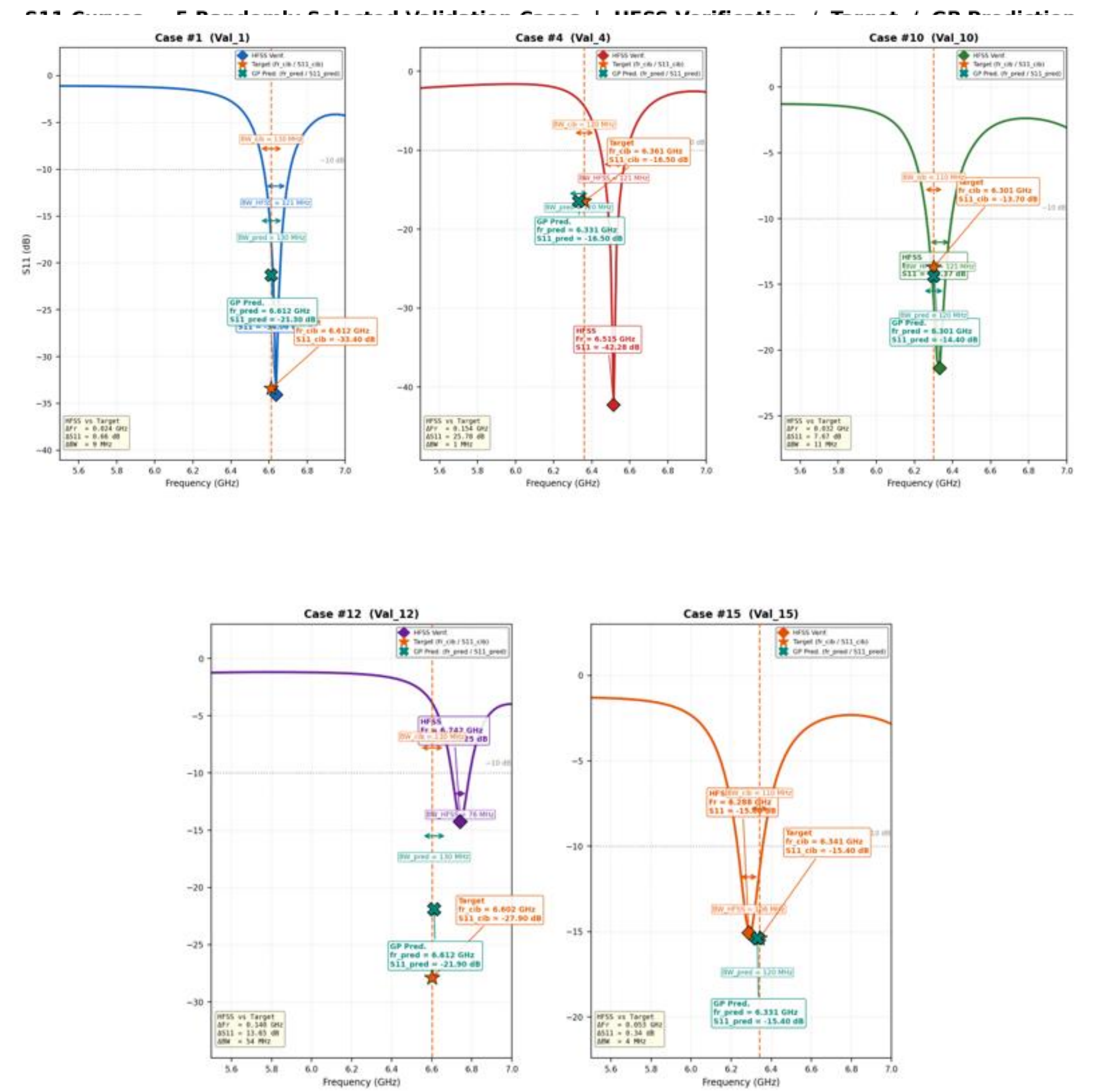
Table confronting that target performances and the HFSS (real) performances

Index_Validat	Fr_Target (GH	Fr_HFSS_Nou	S11_Target (d	S11_HFSS_No	BW_Target (G	BW_HFSS_Nouveau (GHz)
0	6.612	6.636	-33.4	-34.06	0.13	0.121
1	6.812	6.788	-27.93	-15.58	0.12	0.091
2	6.291	6.303	-16.69	-15.92	0.12	0.121
3	6.361	6.515	-16.46	-42.28	0.12	0.121
4	6.642	6.621	-18.36	-29.09	0.1	0.121
5	6.201	6.333	-13.14	-18.85	0.1	0.121
6	6.481	6.485	-19.98	-33.48	0.13	0.121
7	6.521	6.53	-39.69	-28.15	0.12	0.106
8	6.421	6.47	-22.12	-25.8	0.12	0.136
9	6.301	6.333	-13.71	-21.37	0.11	0.121
10	6.331	6.318	-17.2	-16.23	0.12	0.121
11	6.602	6.742	-27.89	-14.25	0.13	0.076
12	6.411	6.5	-15.71	-27.8	0.12	0.136
13	6.311	6.333	-16.63	-25.27	0.12	0.106
14	6.341	6.288	-15.37	-15.06	0.11	0.106
15	6.591	6.561	-26.66	-22.86	0.13	0.121
16	6.632	6.561	-33.69	-22.32	0.12	0.106
17	6.541	6.53	-21.78	-36.34	0.12	0.121
18	6.451	6.47	-22.23	-28.32	0.12	0.121
19	6.591	6.561	-21.94	-22.86	0.13	0.121
20	6.391	6.515	-16.82	-31.7	0.12	0.136
21	6.521	6.53	-22.84	-21.05	0.13	0.121
22	6.391	6.515	-17.59	-31.7	0.12	0.136
23	6.501	6.515	-18.76	-24.07	0.12	0.121
24	6.551	6.53	-24.09	-27.42	0.13	0.121
25	6.652	6.652	-24.66	-37.44	0.13	0.121
26	6.481	6.485	-22.74	-33.48	0.12	0.121
27	6.832	6.758	-15.54	-14.75	0.09	0.091
28	6.311	6.333	-16.84	-25.27	0.11	0.106
29	6.341	6.288	-18.5	-15.06	0.12	0.106

The same remarks remain valid concerning the resonant frequencies namely their fairly good agreement. For the S11min, apart from four divergent points, the remaining score is rather good.

Regarding bandwidth, if we look at the R^2 factor, we might expect to see very poor agreement. However, only 4 points have an error exceeding 10%. This factor simply means that a simple average would yield a better agreement in this case, but the agreement obtained is still good for the majority of points.

In the following 5 cases is randomly chosen and visualized from the 30 samples.



3.12 Conclusion

This chapter has presented the complete simulation methodology underlying the inverse design of circular microstrip patch antennas using Gaussian Process surrogate models. The chapter established HFSS as a reliable substitute for experimental characterisation, detailed the construction of the circular inset-fed patch antenna model parametrised by four geometric variables (RC, Rr, RRC, Wf), and introduced the Latin Hypercube Sampling strategy in the four-dimensional design space. A 150-point LHS campaign was executed successfully in batch mode through a single Python script (run_LHS_simulations.py) with explicit memory management, yielding 120 training samples and 30 validation samples. The Gaussian Process surrogate, configured with a Matérn-1.5 kernel, was trained to predict complete S11 curves from geometrical inputs, effectively acting as a curve generator. The Nelder-Mead optimisation algorithm was selected for its derivative-free robustness and configured with tight tolerances (ftol = xtol = 1e-10) and a generous iteration budget of 2 000. The six-step inverse prediction workflow was detailed, including the composite cost function $J(\theta) = 5000\cdot|\Delta fr| + 10\cdot|\Delta S11\_min| + 5\cdot|\Delta BW|$ that encodes the design priority hierarchy. The non-uniqueness of solutions — the phenomenon that multiple distinct geometries can produce electromagnetically similar responses — was demonstrated on all 30 validation scenarios and validated through direct HFSS recomputation. The findings confirm that the solution space contains multiple equivalent optima, offering designers valuable flexibility in selecting geometries that best accommodate manufacturing constraints.

General Conclusion

and Perspectives

The work presented in this thesis has addressed a problem that lies at the very heart of contemporary antenna engineering: how to combine the rigor of full-wave electromagnetic simulation with the speed of data-driven modelling in order to design — and ultimately optimize — a microstrip patch antenna operating in the 6 GHz band. The journey has progressed from a careful theoretical review of microstrip antennas, through a thorough discussion of surrogate modelling techniques, and has finally converged on a fully automated, reproducible simulation methodology built around the Ansys HFSS solver.

Chapter 1 has fixed the vocabulary and the analytical foundations. The basic principles of operation, the key performance parameters (bandwidth, return loss, radiation pattern, directivity, gain, input impedance), the three classical feed techniques (coaxial, microstrip line and aperture-coupled) and the two simplified analytical models (transmission line and cavity) have been presented in a structured and didactic manner. This chapter has also clearly identified the inherent limitations of microstrip antennas — narrow bandwidth, low gain, sensitivity to fabrication tolerances — which precisely justify the need for accurate simulation and intelligent optimisation.

Chapter 2 has introduced the surrogate-modelling paradigm. The mathematical foundations of artificial neural networks (multilayer perceptrons, activation functions, backpropagation), the data-generation strategies (Latin Hypercube Sampling, train/test splits), the loss functions (MSE, Huber loss) and the alternative Gaussian process regression with its built-in uncertainty quantification have been thoroughly discussed. The chapter has also positioned the two surrogate families with respect to each other: neural networks scale better to large datasets and high-dimensional spaces, whereas Gaussian processes excel on small datasets and provide a principled uncertainty estimate.

Chapter 3 — the methodological core of the thesis — has translated these theoretical foundations into a concrete, automated pipeline. The FEM formulation used by HFSS has been recalled. The design specifications of the patch (target 6 GHz, FR-4 substrate, 50 Ω inset feed) have been fixed and analytically sized. The HFSS model has been constructed step by step, from the project setup to the lumped-port excitation and the adaptive frequency sweep. Most importantly, a four-step Python workflow (generate_lhs.py → run_LHS_simulations.py → extract_results.py → build_dataset.py) has been developed to drive HFSS in batch mode, extract the three key performance metrics (fr, $|S_{11}|$min, BW) and assemble a clean dataset_results.csv file ready for machine-learning consumption. A practical incident — the premature termination of HFSS after 74 of the 100 planned simulations — has been honestly reported and exploited as a real-world illustration of the robustness of the workflow. The final dataset comprises 70 designs (56 for training, 14 for validation). A composite cost function combining frequency accuracy, impedance matching and bandwidth reward has been formulated and will drive the optimization stage.

Main contributions of the thesis

- A unified, didactic presentation of microstrip patch antenna theory aimed at non-specialists in electromagnetics.
- A side-by-side comparison of the two most popular surrogate modelling techniques (MLP and Gaussian process) in the specific context of antenna design.
- A fully automated, four-stage Python pipeline that wraps Ansys HFSS in a reproducible batch workflow, from LHS sampling to the production of a clean ML-ready dataset.
- A composite cost function — equation (3.6) — that encodes the multi-objective nature of the optimisation problem in a single scalar.
- An honest engineering-style discussion of the practical issues encountered (HFSS memory leak, mesh-convergence sensitivity, choice between 100/74/70 samples) that will be useful to any future student tackling a similar topic.

Limitations of the present study

- The dataset is limited to 70 samples in six input dimensions; while sufficient for a small MLP, it remains modest by modern machine-learning standards.
- Only the inset-fed single patch on FR-4 has been considered. Other substrates (Rogers, foam, multilayer), other feed techniques and array configurations would require their own simulation campaigns.
- The HFSS workflow currently runs only under Windows because of the dependence on the Ansys IronPython API; portability to a Linux-based HPC cluster would require non-trivial re-engineering.

Perspectives for further work

- Extension of the dataset by completing the remaining 26 simulations (and possibly enlarging the LHS plan to 200–500 samples) in order to train deeper neural networks and to obtain a quantitative estimate of the model uncertainty.
- Integration of the trained surrogate with a metaheuristic optimizer (genetic algorithm, particle swarm, Bayesian optimization) to identify the design that minimizes the cost function (3.6) and to fabricate / measure the optimal patch.
- Generalization of the methodology to a 1×4 or 2×2 patch array, taking explicitly into account the mutual coupling between elements through additional inputs of the surrogate.
- Investigation of multi-band and reconfigurable antennas using more elaborate parameterizations (slots, parasitic elements, varactor loads) and more recent surrogate architectures (graph neural networks, transformer-based models).
- Development of an interactive web interface that exposes the trained surrogate to non-expert engineers, providing instantaneous antenna design feedback.

Beyond the specific case of the 6 GHz patch, the methodology proposed in this thesis is generic: it can be transposed to any antenna structure for which an HFSS model and a clear set of design variables can be defined. The combination of automated full-wave simulation, intelligent sampling and machine-learning surrogates constitutes today one of the most promising directions of microwave engineering, and we hope that the present work will contribute to its diffusion within the scientific community.

References

[1] C. A. Balanis, Antenna Theory: Analysis and Design, 4th ed. Hoboken, NJ, USA: Wiley, 2016.

[2] R. Garg, P. Bhartia, I. Bahl, and A. Ittipiboon, Microstrip Antenna Design Handbook. Boston, MA, USA: Artech House, 2001.

[3] J. R. James and P. S. Hall, Handbook of Microstrip Antennas. London, U.K.: Peter Peregrinus, 1989.

[4] D. M. Pozar and D. H. Schaubert, Microstrip Antennas: The Analysis and Design of Microstrip Antennas and Arrays. New York, NY, USA: IEEE Press, 1995.

[5] K. F. Lee and K. M. Luk, Microstrip Patch Antennas, 2nd ed. London, U.K.: Imperial College Press, 2011.

[6] W. L. Stutzman and G. A. Thiele, Antenna Theory and Design, 3rd ed. Hoboken, NJ, USA: Wiley, 2012.

[7] D. M. Pozar, "Microstrip Antennas," Proc. IEEE, vol. 80, no. 1, pp. 79–91, Jan. 1992.

[8] N. V. Queipo, R. T. Haftka, W. Shyy, T. Goel, R. Vaidyanathan, and P. K. Tucker, "Surrogate-based analysis and optimization," Progress in Aerospace Sciences, vol. 41, no. 1, pp. 1–28, Jan. 2005.

[9] Q. J. Zhang and K. C. Gupta, Neural Networks for RF and Microwave Design. Norwood, MA, USA: Artech House, 2000.

[10] J. W. Bandler, R. M. Biernacki, S. H. Chen, P. A. Grobelny, and R. H. Hemmers, "Space mapping: The state of the art," IEEE Trans. Microw. Theory Tech., vol. 52, no. 1, pp. 337–361, Jan. 2004.

[11] Ansys HFSS — High Frequency Structure Simulator, User's Guide, Release 2023 R2. Canonsburg, PA, USA: Ansys Inc., 2023.

[12] J.-M. Jin, The Finite Element Method in Electromagnetics, 3rd ed. Hoboken, NJ, USA: Wiley-IEEE Press, 2014.

[13] S. Koziel and L. Leifsson, Surrogate-Based Modeling and Optimization: Applications in Engineering. Berlin, Germany: Springer, 2013.

[14] S. Haykin, Neural Networks and Learning Machines, 3rd ed. Upper Saddle River, NJ, USA: Pearson, 2009.

[15] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," Neural Networks, vol. 2, no. 5, pp. 359–366, 1989.

[16] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.

[17] C. E. Rasmussen and C. K. I. Williams, Gaussian Processes for Machine Learning. Cambridge, MA, USA: MIT Press, 2006.

[18] J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn, "Design and analysis of computer experiments," Statistical Science, vol. 4, no. 4, pp. 409–423, Nov. 1989.

[19] D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient global optimization of expensive black-box functions," J. Global Optimization, vol. 13, no. 4, pp. 455–492, Dec. 1998.

[20] M. A. Bouhlel, J. T. Hwang, N. Bartoli, R. Lafage, J. Morlier, and J. R. R. A. Martins, "A Python surrogate modeling framework with derivatives," Adv. Eng. Software, vol. 135, p. 102662, Sep. 2019.

[21] J.-M. Jin, The Finite Element Method in Electromagnetics, 3rd ed. Hoboken, NJ, USA: Wiley-IEEE Press, 2014.

[22] Ansys Inc., "Ansys HFSS — High Frequency Structure Simulator," User Guide, Release 2023 R2, Canonsburg, PA, USA, 2023.

[23] C. A. Balanis, Antenna Theory: Analysis and Design, 4th ed. Hoboken, NJ, USA: Wiley, 2016, pp. 783–847.

[24] M. D. McKay, R. J. Beckman, and W. J. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," Technometrics, vol. 21, no. 2, pp. 239–245, May 1979.

[25] Q.-J. Zhang and K. C. Gupta, Neural Networks for RF and Microwave Design. Norwood, MA, USA: Artech House, 2000.

[26] S. Koziel and S. Ogurtsov, Antenna Design by Simulation-Driven Optimization. Cham, Switzerland: Springer, 2014.