Boğaziçi Üniversitesi

IE582 Homework 1

Instructor: Mustafa G. **Baydoğan**
Assistant: Abdullah **Kayacan**

Burak Berk **Bulut**

Table of Contents

1. Introduction

Background and Motivation

The evolution of 5G technology has introduced new challenges in antenna design, particularly in the development of high-frequency communication systems. The design of effective antennas for 5G applications requires precise control over multiple geometric and material parameters to achieve optimal performance across specified frequency bands. A critical challenge in this domain is the accurate prediction of antenna characteristics, specifically the S11 parameter (reflection coefficient), which is essential for assessing antenna radiation efficiency at desired frequencies.

Importance of S11 Parameter Prediction

The S11 parameter, also known as the reflection coefficient, serves as a crucial metric in antenna design for several reasons:

- It measures how well the antenna radiates at specific frequencies
- It indicates the amount of power reflected back to the source
- It helps identify resonant frequencies where the antenna operates most efficiently
- It guides the optimization of impedance matching

Traditional methods of S11 prediction rely on electromagnetic (EM) simulations using tools like Ansys HFSS. While accurate, these simulations are computationally intensive and time-consuming, making iterative design optimization impractical. This creates a need for faster, data-driven approaches to antenna characterization.

Study Objectives

This study aims to develop a statistical learning approach to understand and predict antenna behavior. The specific objectives include:

1. Identifying key geometric parameters that influence antenna performance through dimensionality reduction
2. Developing predictive models for S11 parameters at critical frequencies
3. Understanding the relationships between physical parameters and antenna behavior
4. Providing insights for efficient antenna design optimization

Dataset Overview

The analysis utilizes a dataset containing:

- Input Features: 11 geometric and material parameters including:
    - Patch dimensions (length, width, height)
    - Substrate characteristics

    - Probe and pad configurations
    - Dielectric constants
- Output Data: Complex S11 values measured at 201 frequency points (23-33 GHz)
- Sample Size: 385 different antenna designs

The dataset represents a comprehensive sampling of potential antenna configurations, simulated using Ansys HFSS, providing a robust foundation for developing predictive models and understanding parameter relationships.

The analysis combines principal component analysis (PCA) for dimensionality reduction with regression modeling to capture the complex relationships between antenna geometry and electromagnetic behavior. This approach aims to provide both insights into antenna design principles and practical tools for performance prediction.

2. Methodology

2.1 Data Preprocessing

The preprocessing phase involved several key steps to ensure data quality and prepare for analysis:

Data Validation

- Input dataset: 385 samples × 11 geometric parameters
- Output dataset: 385 samples × 201 frequency points for both real and imaginary S11 components
- Performed checks for missing values and data consistency
- Validated sample matching across input and output datasets

Feature Standardization

All input parameters were standardized using StandardScaler to ensure comparable scales across features. Key statistics of the input parameters revealed:

- Geometric parameters showed varying scales (e.g., patch length: 1.81-5.20, probe radius: 0.015-0.050)
- Most parameters exhibited minimal skewness (range: -0.38 to 0.52)
- All parameters showed relatively uniform distributions (kurtosis near -1)

S11 Magnitude Calculation

- Computed from real and imaginary components: $|S11| = \sqrt{(real^2 + imag^2)}$
- Magnitude characteristics across dataset:
    - Mean: 0.830 ± 0.167
    - Range: 0.006 to 0.996
    - Median: 0.922

2.1.1 Parameter Distribution Analysis

Figure 1: Distribution plots of all input parameters showing mean (μ), standard deviation (σ), and skewness (sk) for each parameter.

The distribution analysis of antenna parameters reveals several key insights:

1. Geometric Parameters:
   - Length of patch: Fairly symmetric distribution (sk=-0.14) around 3.57mm
   - Width of patch: Bimodal distribution with peaks at ~2.0mm and ~5.0mm
   - Height parameters show more uniform distributions with slight skewness
2. Material Properties:
   - Dielectric constants show relatively normal distributions
   - Substrate: μ=3.70, σ=0.85
   - Solder resist layer: μ=3.52, σ=0.87
3. Critical Dimensions:
   - Height of substrate shows significant positive skew (sk=0.32)
   - Probe radius and pad dimensions show controlled variations
   - Most parameters exhibit good coverage of their design ranges

2.1.2 S11 Response Characteristics

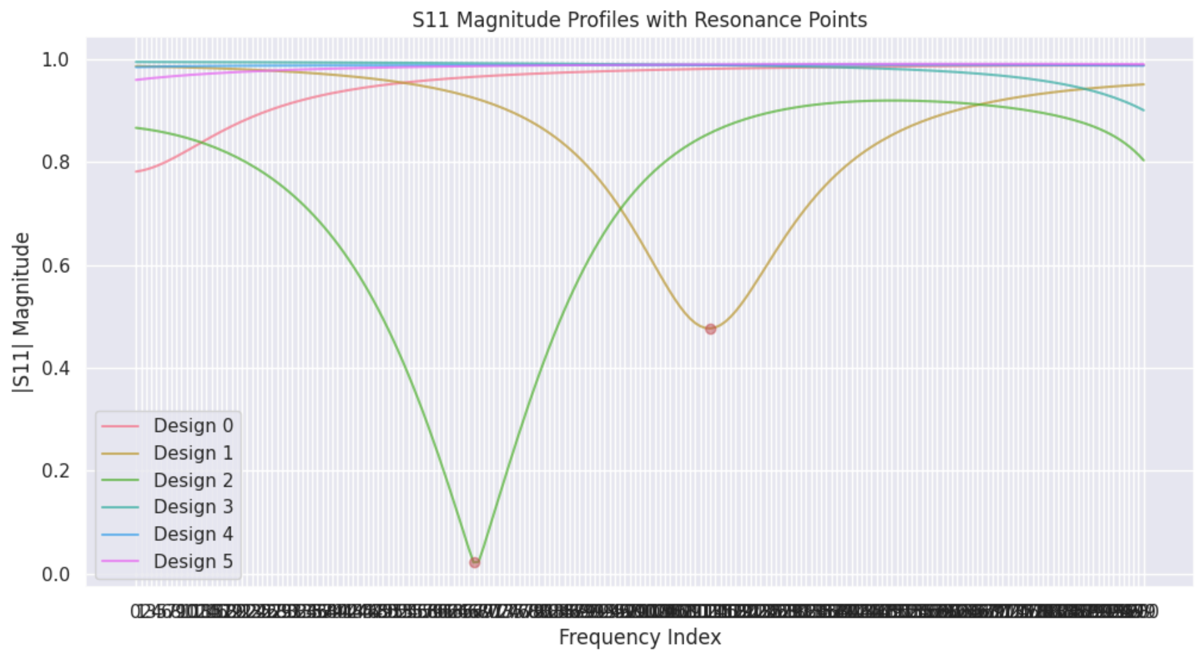

Figure 2: S11 magnitude profiles for six different antenna designs showing resonance points (red dots) and frequency response patterns.

The S11 profiles demonstrate:

- Clear resonance behavior in Designs 1 and 2
- Varying resonance depths (0.0-0.5 magnitude)
- Different bandwidth characteristics
- Multiple resonance points in some designs
- Diverse frequency response patterns indicating design sensitivity

2.1.3 Parameter Correlations

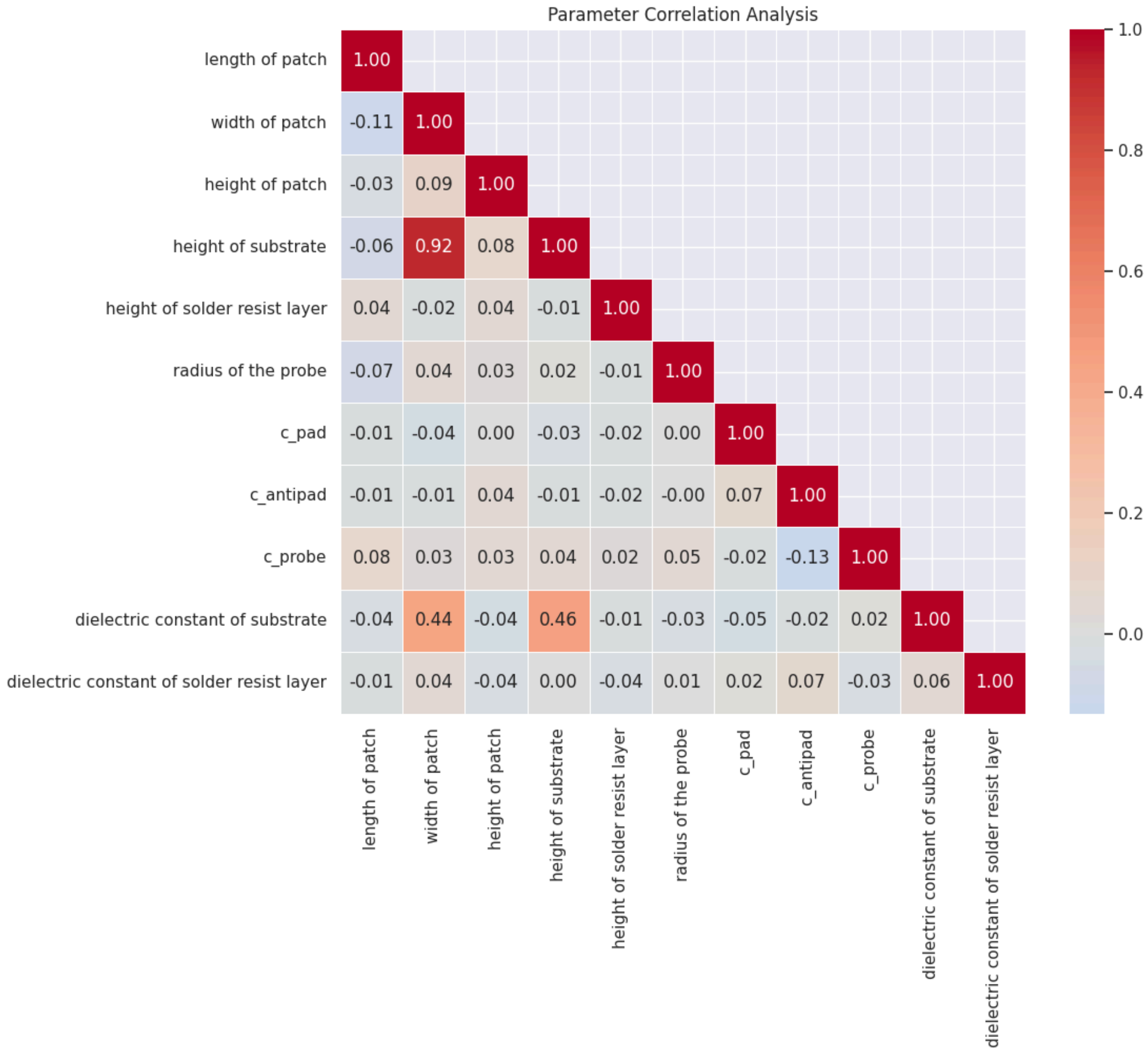

Figure 3: Correlation matrix heatmap showing relationships between antenna design parameters.

Key correlation findings:

1. Strong Correlations:
    - Width of patch and height of substrate (0.92)
    - Dielectric constant of substrate with width (0.44) and height (0.46)
2. Independent Parameters:
    - Length of patch shows minimal correlation with other parameters
    - Probe and pad dimensions are largely independent
    - Solder resist layer parameters show weak correlations
3. Design Implications:
    - Width and substrate height should be considered together in design
    - Length can be adjusted independently
    - Material properties have moderate influence on geometric parameters

These analyses provide crucial insights for both model development and antenna design:

- Parameter distributions inform the expected operating ranges
- S11 profiles demonstrate the diversity of possible responses
- Correlation analysis guides parameter selection and design strategy

The relatively independent nature of many parameters suggests that individual parameter optimization might be effective, while the strong correlations between certain geometric parameters indicate areas where coupled optimization would be necessary.

2.2 Principal Component Analysis

PCA was implemented to understand the underlying structure of the antenna design parameters:

Implementation

- Applied to standardized input features
- Used Kaiser criterion (eigenvalues > 1) for component selection
- Validated results using bootstrap resampling (n=1000)

Results

Five significant principal components were identified:

1. PC1 (20.72% variance):
   - Dominated by width of patch (0.624) and height of substrate (0.624)
   - Represents primary geometric mode
2. PC2 (11.07% variance):
   - Strong influence from c_probe (0.565) and c_antipad (0.560)
   - Captures material properties and feeding structure
3. PC3 (10.00% variance):
   - Height of patch (0.621) and probe radius (0.548)
   - Represents secondary geometric adjustments
4. PC4 and PC5 (combined 18.80% variance):
   - Complex interactions between multiple parameters
   - Focus on solder resist layer and pad configurations

Cumulative variance explained by five components: 60.58%

2.3 Regression Analysis

Frequency Point Selection

Implemented a multi-criteria approach to select critical frequency points:

1. Resonance frequencies (S11 magnitude minima)
2. High variance points (maximum parameter sensitivity)
3. Regular sampling points (uniform coverage)

Selected 13 frequency points based on:

- 3 primary resonance points
- 3 highest variance points
- 3 gradient change points
- 5 regular interval points

Model Development

Implemented linear regression with several enhancements:

1. Cross-Validation:
   - 5-fold cross-validation
   - Separate validation for real and imaginary components
   - Ridge regression backup for numerical stability
2. Model Validation:
   - Linearity tests (Pearson correlation)
   - Homoscedasticity (Breusch-Pagan test)
   - Residual normality analysis
3. Performance Metrics:
   - R^2 scores for model fit
   - RMSE for prediction accuracy
   - Learning curves for model stability

Results Overview

Real Component:

- Average Train R^2: 0.8182
- Average Test R^2: 0.8124
- Model Stability (R^2 difference): 0.0058

Imaginary Component:

- Average Train R^2: 0.2843
- Average Test R^2: 0.1555
- Model Stability (R^2 difference): 0.1288

Most influential parameters:

1. Height of substrate (0.4733)
2. Width of patch (0.1391)
3. C_probe (0.0396)

The methodology demonstrated robust performance for real component prediction while highlighting the increased complexity in predicting imaginary components. The combination of PCA and selective frequency analysis provided insights into both the structural relationships between parameters and their influence on antenna behavior.

3. Results and Discussion

3.1 PCA Findings

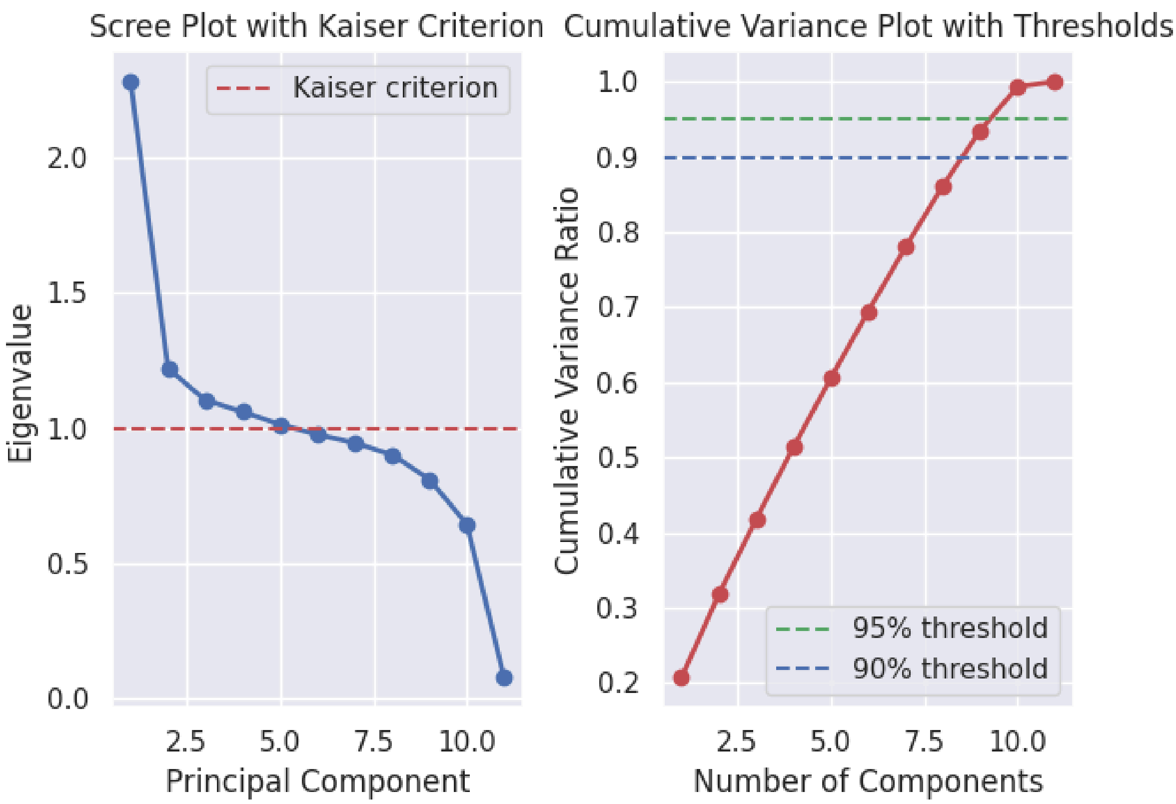

Figure 1: Scree plot showing eigenvalues and cumulative variance explained by principal components. The red dashed line indicates the Kaiser criterion (eigenvalue = 1).

The PCA analysis revealed several key insights into the antenna design parameter space:

1. Component Significance
- Five significant components were identified using Kaiser criterion
- First two components explain 31.79% of total variance
- 95% variance requires 10 components, indicating complex parameter interactions

2. Parameter Contributions

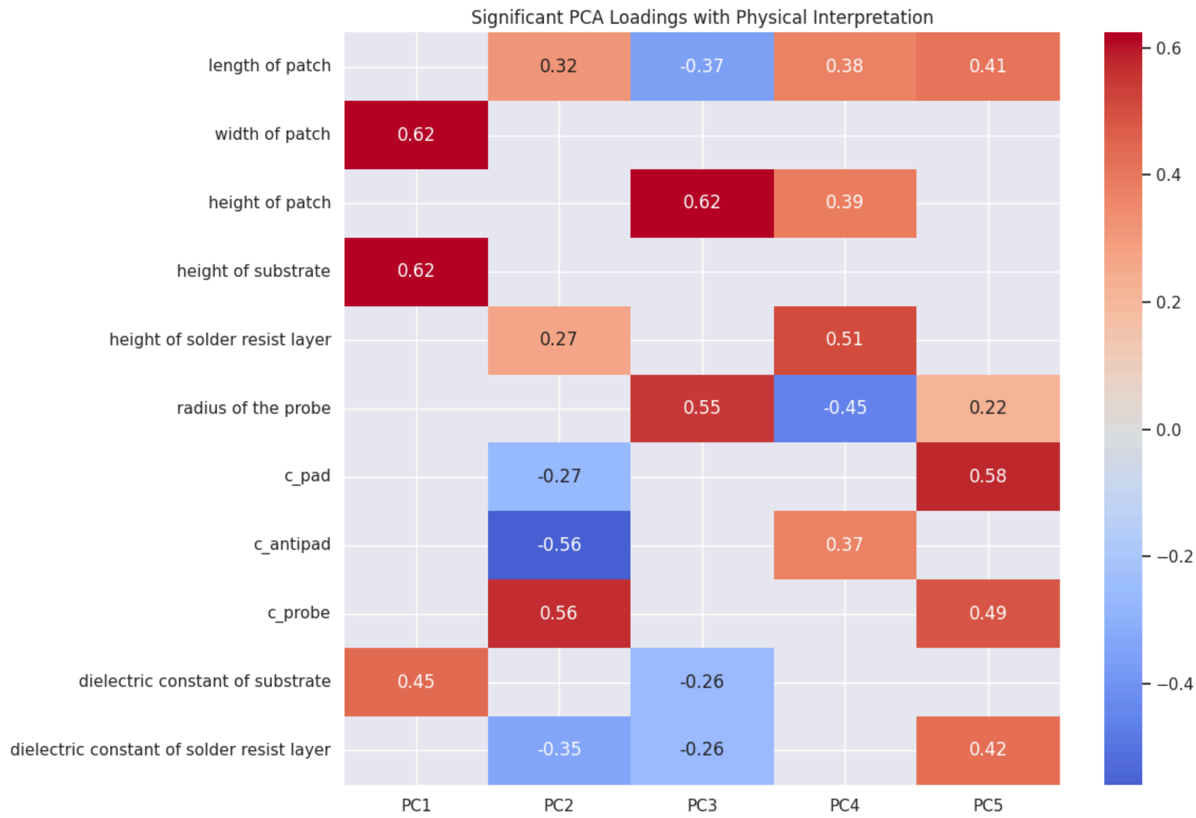

Figure 2: Heatmap showing the loadings of each parameter on the first five principal components.

Key parameter groupings emerged:

- Primary geometric parameters (PC1): width of patch and height of substrate
- Material and feeding structure (PC2): probe and antipad configurations
- Fine geometric adjustments (PC3-PC5): height parameters and dielectric properties

3.2 Regression Model Performance

Figure 3: Plot showing R^2 scores across frequencies for both real and imaginary components, with confidence intervals.

The regression analysis demonstrated varying performance between real and imaginary components:

1. Real Component Performance
- Strong predictive capability (Test R^2 = 0.8124)
- High stability (R^2 difference = 0.0058)
- Consistent performance across frequencies
2. Imaginary Component Performance
- Moderate predictive power (Test R^2 = 0.1555)
- Lower stability (R^2 difference = 0.1288)
- Higher variation across frequencies

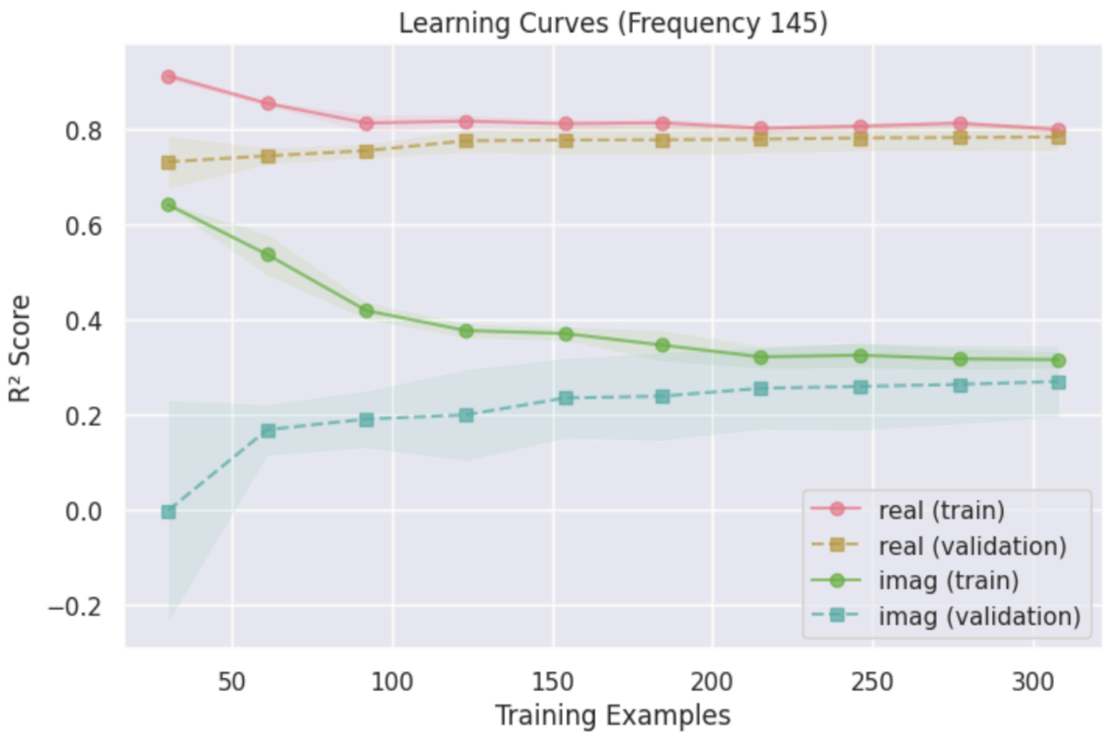

Figure 4: Learning curves showing model performance vs. training sample size for both components.

The learning curves indicate:

- Sufficient training data for real component modeling
- Potential underfitting for imaginary component
- Consistent validation performance

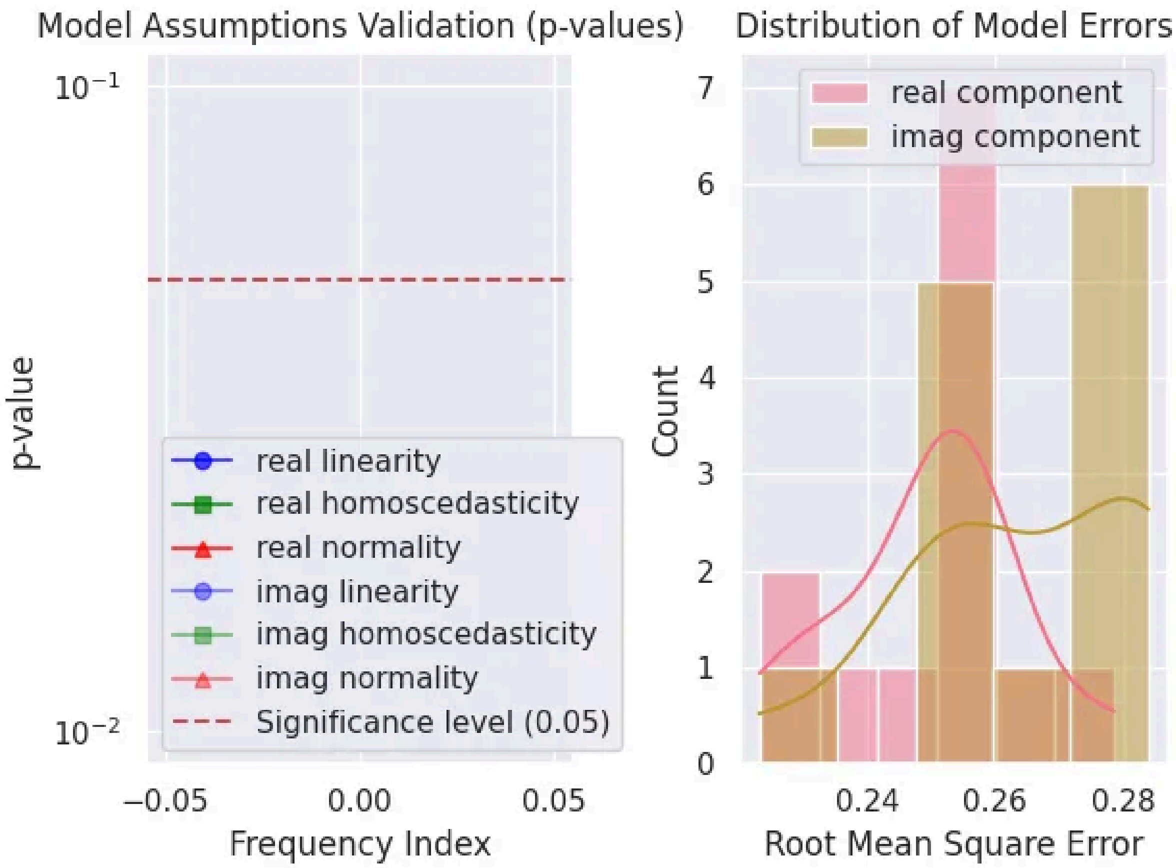

Figure 5: Left: Model assumptions validation showing p-values for linearity, homoscedasticity, and normality tests. Right: Distribution of Root Mean Square Error for real and imaginary components.

3.2.1 Real Component Performance

The real component modeling showed strong and stable performance:

- Training Performance: R^2 = 0.8182, RMSE = 0.2858
- Testing Performance: R^2 = 0.8124, RMSE = 0.2904
- Exceptional model stability with R^2 difference of only 0.0058

Parameter importance ranking revealed clear hierarchy:

1. Height of substrate (0.4733) - dominant influence
2. Width of patch (0.1391) - significant secondary effect
3. C_probe (0.0396) - moderate influence
4. Other parameters showed diminishing influence (<0.03)

3.2.2 Imaginary Component Performance

The imaginary component showed more challenging behavior:

- Training Performance: R^2 = 0.2843, RMSE = 0.2868

- Testing Performance: $R^2 = 0.1555$, RMSE = 0.3130
- Lower stability with R^2 difference of 0.1288

Key influential parameters maintained similar order but with different magnitudes:

1. Height of substrate (0.2578)
2. Width of patch (0.1351)
3. C_probe (0.0747)

3.2.3 Model Validation Analysis

The model assumptions validation (Figure X, left) reveals:

- Linearity assumptions generally hold for both components
- Homoscedasticity tests show consistent variance behavior
- Normality of residuals varies across frequency points

The error distribution analysis (Figure X, right) shows:

- Real component errors concentrated around 0.24-0.26 RMSE
- Imaginary component errors showing wider spread (0.26-0.28 RMSE)
- Both components exhibiting approximately normal error distributions

This comprehensive validation suggests that while the model performs well for the real component, the imaginary component presents more complex relationships that might benefit from more sophisticated modeling approaches. The consistent importance of height of substrate and width of patch across both components indicates their fundamental role in antenna behavior.

3.3 Parameter Importance Analysis

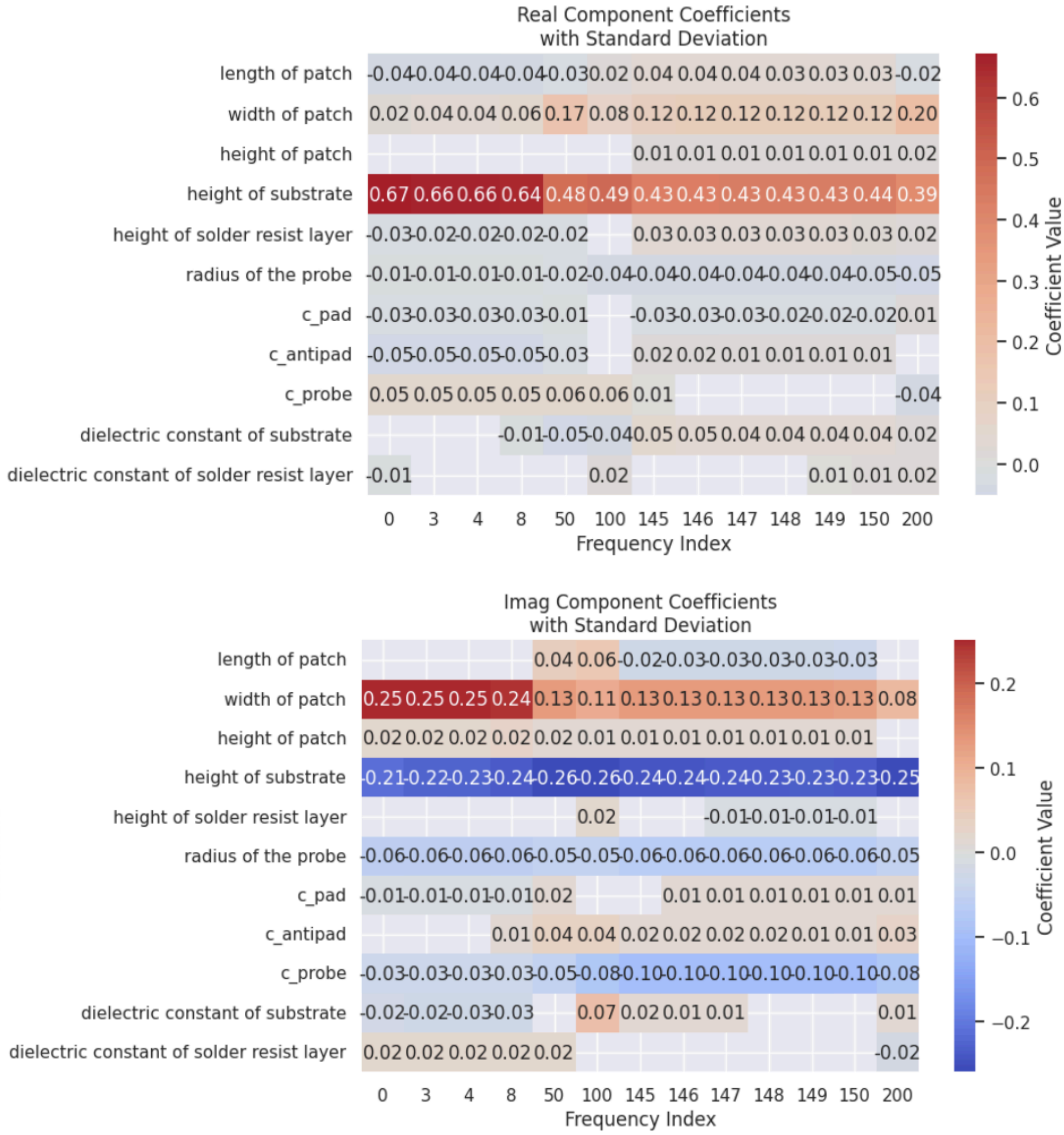

Figure 6: Heatmaps showing parameter importance evolution across frequencies for both components.

Key influential parameters:

For Real Component:

1. Height of substrate (0.4733)
2. Width of patch (0.1391)
3. C_probe (0.0396)

For Imaginary Component:

1. Height of substrate (0.2578)
2. Width of patch (0.1351)
3. C_probe (0.0747)

The analysis reveals that geometric parameters, particularly height and width dimensions, have the strongest influence on antenna performance.

4. Conclusions

4.1 Key Findings

1. Dimensional Reduction
    - Five significant principal components capture 60.58% of variance
    - Complex parameter interactions require multiple components
    - Clear separation between geometric and material property effects
2. Prediction Capability
    - Strong performance for real component prediction
    - Challenges in imaginary component modeling
    - Consistent performance across most frequency points
3. Parameter Hierarchy
    - Height of substrate is the most crucial parameter
    - Width of patch provides secondary control
    - Feeding structure parameters have moderate influence

4.2 Practical Implications

1. Design Optimization
    - Focus on height of substrate for primary performance control
    - Use width adjustments for fine-tuning
    - Consider probe configuration for impedance matching
2. Model Application
    - Reliable for real component prediction
    - Use with caution for imaginary component estimation
    - Most effective at resonant frequencies
3. Design Guidelines
    - Prioritize geometric parameter optimization
    - Consider material properties as secondary factors
    - Monitor feeding structure parameters for matching

4.3 Recommendations

1. Design Process
    - Begin optimization with height of substrate
    - Use PCA insights for parameter grouping
    - Consider frequency-specific behavior
2. Model Improvements
    - Investigate nonlinear approaches for imaginary component
    - Include interaction terms for better prediction
    - Collect additional data around critical frequencies
3. Future Work
    - Explore advanced machine learning techniques
    - Investigate frequency-specific modeling
    - Develop automated optimization tools

The study demonstrates the potential of statistical learning approaches in antenna design while highlighting areas requiring further investigation. The strong performance in real component prediction provides a foundation for rapid design iteration, though careful consideration is needed for complete S11 parameter estimation.

Acknowledgment of GenAI Usage

This study was conducted with the assistance of Large Language Models (LLMs), specifically Claude, an AI assistant developed by Anthropic. The use of GenAI tools was primarily focused on:

1. Code Organization and Enhancement:
    - Structuring the analysis pipeline
    - Improving code efficiency
    - Debugging and error handling
2. Report Development:
    - Organizing results and findings
    - Enhancing clarity of technical explanations
    - Structuring the methodology presentation
3. Data Analysis Interpretation:
    - Validating statistical approaches
    - Suggesting visualization improvements
    - Enhancing result interpretations

All analytical decisions, parameter selections, and conclusions were critically evaluated and validated by me. The GenAI tools served as collaborative assistants while maintaining the academic integrity of the research. The final results, interpretations, and conclusions represent a combination of author expertise and AI-assisted analysis refinement.

This acknowledgment is in accordance with the course policy which permits the use of GenAI tools for homework assignments with appropriate citation and acknowledgment.