

Machine Learning as a tool to forecast the Power Quality of a Wind Energy Power Plant

By

Mbuyiselwa Lawrence Cindi

Dissertation submitted in fulfilment of the requirements for the degree:

Master of Engineering in Electrical Engineering

In the Department of Electrical, Electronic and Computer Engineering

Faculty of Engineering, Built Environment and Information Technology

Central University of Technology, Free State

Supervisor: Prof M Sibiya

Co-supervisor: Prof E.D Markus

2026

DECLARATION

I, Mbuyiselwa Lawrence Cindi, with Student number _____, hereby declare that this dissertation submitted to the Central University of Technology, Free State for the degree of Master of Engineering in Electrical Engineering is my original work.

I certify that it complies with the Code of Academic Integrity and other relevant policies, procedures, rules, and regulations of the university. This work has been submitted to originality check software and falls within acceptable limits of originality.

The data and content are authentic, based on research principles, and my engineering skills, knowledge, and understanding of the topic. Furthermore, this work or any portion thereof has never been submitted to any institution to obtain another qualification.

.....

Signature of Candidate

On this ...30th..... day ofOctober..... year2025.....

LIST OF PUBLICATIONS

1. Cindi, M.L, Sibiya, M. & Markus, E., 2024. Machine Learning as a Tool to Forecast the Power Quality of Wind Energy Power Plants: A Systemic Literature Review. *International Journal of Communication Networks and Information Security (IJCNIS)*, 16(1 Special Issue), pp.1284–1316. Available at:
<https://www.ijcnis.org/index.php/ijcnis/article/view/7049>
Status: Published
2. Cindi M.L, M. Sibiya. E. Markus, Performance Evaluation of Machine Learning Models for Power Quality Forecasting in South African Wind Power Plant: A case study.
Link: <https://>
Status: Submitted for Publication at Electric Power Systems Research Journal

ABSTRACT

The large-scale integration of renewable energy sources, particularly wind power, into national grids introduces non-linear and intermittent behaviour that challenges the maintenance of acceptable power quality (PQ). In South Africa, these challenges must be managed within the limits and measurement requirements of NRS 048-2 and related IEC standards at the point of common coupling (PCC).

This dissertation investigates how effectively machine learning (ML) models can forecast key PQ parameters at a utility-scale South African wind power plant, and how these forecasts perform when judged not only by conventional regression accuracy but also by compliance-oriented metrics aligned with NRS 048-2 (Missed Violation Rate, MVR, and False Alarm Rate, FAR).

Using 10-minute operational SCADA data from the wind-farm PCC, a six-phase, leakage-aware analytical workflow is implemented in JMP Pro 18. The pipeline covers standards-aligned preprocessing, domain-informed feature engineering, chronological train/validation/test splits, and multi-model benchmarking. Four model families are compared: Boosted Trees (BT), Support Vector Machines (SVM), a feed-forward Neural Network (NN), and a SARIMAX baseline. Forecasts are evaluated using RMSE, MAE, MAPE, R^2 and compliance metrics (FAR, MVR), including a ± 5 % sensitivity analysis around key NRS 048-2 limits.

Results show that non-linear models are essential for PQ forecasting in this wind-integrated grid. BT consistently delivers the best or near-best performance for Voltage Unbalance, Frequency, long-term flicker (P_{lt}), and V_{RMS} (p.u.), achieving high R^2 (often > 0.95) and low MVR (≈ 5 % for Unbalance, ≈ 0 – 5 % for P_{lt}) with FAR typically below 1 %. SVM is particularly strong for V_{THD} and competitive for P_{lt} , while the NN performs best on short-term flicker (P_{st}) in terms of regression accuracy but exhibits higher MVR. SARIMAX underperforms on non-linear and volatile parameters, confirming the limits of purely linear approaches. The study provides the first detailed, NRS 048-2-aligned PQ forecasting case study on a South African wind plant, demonstrates a reproducible, standards-aware modelling pipeline, and shows how

FAR/MVR-based evaluation can translate ML forecasts into actionable early-warning tools for grid operators, supporting proactive PQ management in renewable-rich networks.

DEDICATION AND ACKNOWLEDGEMENTS

For the wisdom granted to me and the opportunities made available to me, I thank the Almighty Living God. I am eternally grateful for the gift of life.

This thesis is dedicated to my mother, who engrained in me the importance of goal setting and perseverance, which invariably lead to success.

I would like to express my heartfelt gratitude to my supervisors, Prof. M. Sibiya and Prof. E.D. Markus, for their guidance, patience, and insightful criticism. Thank you, Prof. Sibiya, for pushing me! Thank you, Prof Markus, for your support, and the last push towards the end, and guidance throughout my academic years.

I thank my wife and family for their unwavering support and patience throughout my studies.

To my colleagues and friends who have assisted me, I extend my sincere gratitude and appreciation, particularly to Mr. Mthobisi Diya. Thank you for your invaluable time!

Finally, I extend my sincere appreciation to the faculty and Department of Electrical, Electronic, and Computer Engineering at Central University of Technology (CUT) for entrusting me with this research project.

TABLE OF CONTENTS

DECLARATION	ii
LIST OF PUBLICATIONS.....	iii
ABSTRACT	iv
DEDICATION AND ACKNOWLEDGEMENTS	vi
TABLE OF CONTENTS.....	vii
LIST OF FIGURES	xiv
LIST OF TABLES	xvi
LIST OF ABBREVIATIONS AND ACRONYMS	xvii
CHAPTER 1: INTRODUCTION	1
1.1 BACKGROUND OF THE STUDY: POWER QUALITY IN MODERN GRIDS.....	1
1.2 PROBLEM STATEMENT: PREDICTIVE CHALLENGES AND REGIONAL RESEARCH GAPS	2
1.3 RESEARCH QUESTIONS.....	3
1.3.1 Primary question.....	3
1.3.2 Secondary questions	3
1.4 AIM AND OBJECTIVES.....	3
1.5 SIGNIFICANCE OF THE RESEARCH	4
1.6 DISSERTATION STRUCTURE.....	5
1.7 CHAPTER SUMMARY	6
CHAPTER 2: LITERATURE REVIEW	7
2.1 CHAPTER OVERVIEW	7
2.2 POWER QUALITY PHENOMENA IN RENEWABLE INTEGRATED GRIDS	7
2.2.1 PQ issues driven by wind and solar intermittency.....	8
2.2.2 PQ in South African and African Renewable Grids	9

2.3 STANDARDS AND REGULATORY FRAMEWORK FOR POWER QUALITY.....	10
2.3.1 International standards	10
2.3.2 NRS 048-2 and the South African Grid Code	12
2.4 MONITORING, ASSESSMENT, AND MITIGATION OF PQ IN RENEWABLE PLANTS.	14
2.4.1 Measurement and assessment.....	14
2.4.2 Mitigation strategies.....	14
2.4.3 Emerging role of forecasting and artificial intelligence (AI)	14
2.5 MACHINE LEARNING APPLICATIONS IN POWER SYSTEMS AND PQ.....	15
2.6 MACHINE LEARNING ALGORITHMS FOR PQ FORECASTING: STRENGTHS AND WEAKNESSES	17
2.6.1 Support Vector Machines (SVM)	18
2.6.2 Artificial Neural Networks (ANN).....	19
2.6.3 Boosted Trees (BT)	20
2.6.4 SARIMAX (Seasonal ARIMA with Exogenous Inputs)	20
2.7 CRITICAL REVIEW OF ML BASED PQ FORECASTING.....	21
2.8 CONCEPTUAL FRAMEWORK GUIDING THE STUDY	22
2.9 CHAPTER SUMMARY	24
CHAPTER 3: RESEARCH METHODOLOGY.....	26
3.1 WORKFLOW OVERVIEW	26
3.2 DATA HANDLING AND RELATED MEASUREMENTS.....	28
3.2.1 Targets and Labelling	29
3.2.2 Data Preprocessing	30
3.2.3 Feature Engineering	31
3.3 MODELS AND TRAINING	33
3.3.1 Support Vector Machines (SVM)	34

3.3.2 Artificial Neural Networks (ANN).....	35
3.3.3 SARIMAX (seasonal ARIMA with exogenous)	37
3.3.4 Boosted Trees (BT; Gradient Boosting/XGBoost-Style).....	39
3.3.5 Summary of Model Setup and Method Rationale	41
3.4 METRICS: ACCURACY AND NRS 048-COMPLIANCE	41
3.4.1 Regression Metrics.....	41
3.4.2 NRS 048 related Compliance Metrics.....	43
3.5 VALIDATION DESIGN AND LEAKAGE CONTROLS	44
3.6 REPRODUCIBILITY AND TOOLING	45
3.7 CHAPTER SUMMARY	46
CHAPTER 4: RESULTS AND ANALYSIS.....	48
4.1 DESCRIPTIVE ANALYSIS OF POWER QUALITY DATA.....	48
4.2 COMPARATIVE PERFORMANCE OF ML MODELS	48
4.2.1 Overall Model Performance	48
4.2.2 Discussion of trends	51
4.2.3 Visual summaries	52
4.2.4 Illustrative Cases (Selected Targets).....	52
4.2.5 Illustrative Cases (Selected Targets).....	54
4.3 NRS 048-2 COMPLIANCE ORIENTED PERFORMANCE	56
4.3.1 Visual summary	57
4.3.2 Detailed Compliance for Volatile Parameters	58
4.3.3 Threshold Sensitivity ($\pm 5\%$)	60
4.3.4 Operational interpretation	63
4.4 FEATURE IMPORTANCE AND INTERPRETABILITY	63
4.5 CHAPTER SUMMARY	67

CHAPTER 5: DISCUSSION OF FINDINGS	69
5.1 INTERPRETATION OF MODEL PERFORMANCE.....	69
5.1.1 Strengths of Non-Linear Models.....	69
5.1.2 Limitations of Linear Models.....	70
5.2 THE PRACTICAL UTILITY OF NRS 048-2 COMPLIANCE ORIENTED FORECASTING	71
5.3 LOCALISED PERSPECTIVE ON PQ FORECASTING IN SOUTH AFRICA	72
5.3.1 Contextual Relevance	72
5.3.2 Methodological Alignment with Global Practice	73
5.3.3 Data Realism.....	73
5.4 IMPLICATIONS FOR GRID MANAGEMENT AND PREDICTIVE MAINTENANCE	73
5.4.1 Forecasts as Early Warning Tools	73
5.4.2 Link to Maintenance and Asset Management.....	74
5.4.3 Support for Grid Code Compliance Management and Planning	74
5.5 ANSWERING THE RESEARCH QUESTIONS AND OBJECTIVES	75
5.5.1 Primary Research Question.....	75
5.5.2 Secondary Questions	75
5.6 METHODOLOGICAL AND DESIGN CONSIDERATIONS	76
5.6.1 Single-site scope	77
5.6.2 Limited exposure to rare disturbance.....	77
5.6.3 Feature boundary conditions	77
5.6.4 Temporal resolution	77
5.6.5 Tooling constraints.....	77
5.6.6 Model class scope.....	78
5.7 SUMMARY	78

CHAPTER 6: CONCLUSION AND FUTURE WORK.....	80
6.1 SUMMARY OF RESEARCH.....	80
6.2 KEY CONTRIBUTIONS TO KNOWLEDGE.....	80
6.2.1 Contextual contribution (South Africa, globally relevant).....	80
6.2.2 Methodological Contribution	81
6.2.3 Empirical contribution (Non-Linear Necessity and Suitability Map).....	81
6.2.4 Practical/Tooling contribution (Accessible Analytics Pathway).....	82
6.2.5 Per-Phase Granularity	82
6.3 PRACTICAL IMPLICATIONS FOR INDUSTRY	82
6.3.1 Shift from Reactive to Predictive PQ Management.....	82
6.3.2 Strategic Benefits for Reliability and Cost.....	83
6.3.3 Implications for Policy, Regulation, and Planning	83
6.3.4 Adoption and Capacity Building	83
6.4 LIMITATIONS	84
6.4.1 Single-site scope	84
6.4.2 Rare event exposure	84
6.4.3 Feature availability.....	84
6.4.4 Temporal resolution	84
6.4.5 Tooling constraints.....	85
6.4.6 Model-class scope.....	85
6.5 RECOMMENDATIONS FOR FUTURE RESEARCH.....	85
6.5.1 Real-time Integration and Alerting	85
6.5.2 Uncertainty Aware Forecasts.....	85
6.5.3 Optimisation Coupled Mitigation	86
6.5.4 Transferability and Scaling	86

6.5.5 Advanced model architectures.....	86
6.5.6 Explainability and operator tooling.....	86
6.5.7 Digital-twin Stress Tests	86
6.5.8 Reproducible artefacts and sharing	87
6.6 CONCLUDING REMARKS.....	87
REFERENCES	88
APPENDIX A: REPRODUCIBILITY GUIDE.....	92
A.1 Software Environment.....	92
A.2 Hardware Specifications.....	92
A.3 Data Handling and Preparation	93
A.4 Feature Engineering Scripts/Steps.....	93
A.5 Data Splitting Strategy	94
A.6 Model Implementation Details.....	94
A.7 Evaluation Metrics.....	95
A.8 Access to Data and Code.....	95
APPENDIX B: EXTENDED RESULTS TABLES.....	97
B.1 Overall Performance Summary (Mean Metrics Across All PQ Parameters).....	97
B.2 Detailed Continuous Prediction Performance by PQ Parameter and Phase	97
B.3 Detailed Compliance Performance by PQ Parameter and Phase.....	99
B.4 Sensitivity Analysis of Compliance Thresholds	100
B.5 Illustrative Residual Plots and Histograms	104
B.6 Remaining Residuals vs. Predicted Plots.....	106
B.7 Actual vs. Predicted Plots and Supplementary Time-Series Visualizations.....	106
APPENDIX C: JMP PRO CONFIGURATION DETAILS.....	109
C.1 General JMP Pro Settings	109

C.2 Support Vector Machine (SVM) Configuration	109
C.3 Neural Network (NN) Configuration	110
C.4 Seasonal Auto-Regressive Integrated Moving Average with Exogenous Variables (SARIMAX) Configuration.....	110
C.5 Boosted Tree (BT) Configuration	111
APPENDIX D: ETHICAL COMPLIANCE DOCUMENTATION	113
D.1 Data Access and Confidentiality.....	113
D.2 Ethical Approval	113
D.3 Data Usage and Limitations.....	114
D.4 Data Dictionary (Sample).....	114

LIST OF FIGURES

Figure 2.1: Global Power Quality Standards Comparative Framework	11
Figure 2.2: Diverse Applications of Machine Learning Across Modern Power Systems	17
Figure 2.3: Broad Categorisation of Machine Learning Algorithms and Key Examples Relevant to Power Systems.....	18
Figure 2.4: Conceptual Framework Guiding Power Quality Forecasting Study	23
Figure 3.1: Research Analytical Workflow (Flowchart)	26
Figure 3.2: Generic Illustration of Data Preprocessing Steps (image AI-Generated)	31
Figure 3.3: Conceptual Architecture of Artificial Neural Network (MLP)	37
Figure 3.4: Illustrative Mechanism of Boosted Tree Ensemble	40
Figure 3.5: Confusion Matrix for Binary Classification (TP/TN/FP/FN).....	43
Figure 3.6: Temporal Data Splitting Strategy (70/20/10, Chronological).....	45
Figure 4.1: Average Test RMSE by Model (Across PQ Metrics)	53
Figure 4.2: Average Test R^2 by Model (Across PQ Metrics).....	53
Figure 4.3: Overall Model Performance Across PQ Parameters (Spider Plot)	54
Figure 4.4: Time-Series Tracking of Boosted Tree for Voltage Unbalance (Test Set: 6-Day View)	55
Figure 4.5: Time-Series Tracking of Boosted Tree for Frequency (Test Set: Single-Day View)	56
Figure 4.6: Time-Series Tracking of SVM for V_THD1 (Test Set: 3-Hour View)	57
Figure 4.7: Residuals vs predicted values for BT Voltage Unbalance (test set)	58
Figure 4.8: Histogram of residuals for BT Voltage Unbalance (test set)	59
Figure 4.9: Heatmap of Missed Violation Rates (MVR) across models for Aggregated System-Level PQ Parameters	60
Figure 4.10: FAR-MVR Scatter Plot.....	62
Figure 4.11: Sensitivity of Missed Violation Rate (MVR) to $\pm 5\%$ threshold changes for Flicker Pst Overall.....	64
Figure 4.12: Variable importance plot – BT Voltage Unbalance.....	66
Figure B.1-B.5: Histograms of Residuals for Boosted Tree (Boosted Tree and SVM; V_RMS, V_THD1, Voltage Unbalance).....	102

Figure B.3-B.5: Histograms of Residuals for Boosted Tree V_RMS_1_pu (Test Set)	103
Figure B.6-B7: Residuals vs. Predicted V_RMS_1_pu, Voltage Unbalance for Boosted Tree (Test Set)	104
Figure B.8-B11: Actual vs Predicted Voltage Unbalance, Voltage Unbalance, Frequency, Flicker PLT for Boosted Tree (Test Set)	105
Figures B.12–B.14: Actual vs. Predicted Values (Boosted Tree; Voltage Unbalance, Frequency, Flicker PLT)	106

LIST OF TABLES

Table 2.1: Overview of Key Global Power Quality Standards	12
Table 2.2: NRS 048-2 limits used for compliance labelling in this study	13
Table 3.1: Summary of Engineered Features.....	32
Table 3.2: Implementation in JMP Pro SVM.....	34
Table 3.3: Implementation in JMP Pro NN	36
Table 3.4: Implementation in JMP Pro SARIMAX.....	38
Table 3.5: Implementation in JMP Pro BT	39
Table 3.6: Summary of model setup and rationale in JMP Pro.....	41
Table 4.1: Aggregated Model and Target Regression Performance (Test Set)	49
Table 4.2. Aggregated Compliance Metrics (Accuracy, FAR, MVR, 95% CIs).....	61
Table B.1: Per-Parameter Performance Summary (Models × PQ Targets, Overview)	95
Table B.2: Voltage Unbalance Detailed Train/Validation/Test Metrics (Per-Phase & Overall)	96
Table B.3-B7: V_THD1 Detailed Train/Validation/Test Metrics.....	97
Tables B.8–B.16: Detailed Compliance/Performance Results for Frequency, V_THD (Per-Phase & Overall), Under-/Over-Voltage, V_RMS_pu	97
Table B.17: Sensitivity Analysis of Compliance Thresholds (±5%)	99
Table D.1: Sample Data Dictionary.....	112

LIST OF ABBREVIATIONS AND ACRONYMS

ACF	Autocorrelation Function
AI	Artificial Intelligence
ANN	Artificial Neural Network
BT	Boosted Trees
EN	European Standard
FAR	False Alarm Rate
FN	False Negative
FNR	False Negative Rate
FP	False Positive
FPR	False Positive Rate
GWEC	Global Wind Energy Council
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
JMP Pro	JMP Statistical Discovery Software
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
MLP	Multilayer Perceptron
MSE	Mean Squared Error
MVR	Missed Violation Rate
NERSA	National Energy Regulator of South Africa

NRS	National Rationalised Specifications
PACF	Partial Autocorrelation Function
PCC	Point of Common Coupling
PDP	Partial Dependence Plot
PF	Power Factor
PLT	Long-Term Flicker
PoC	Point of Connection
PQ	Power Quality
PST	Short-Term Flicker
p.u	Per Unit
R^2	Coefficient of Determination (or R-Square)
ReLU	Rectified Linear Unit
RES	Renewable Energy Sources
RMSE	Root Mean Squared Error
RoCoF	Rate of Change of Frequency
SARIMAX	Seasonal Auto-Regressive Integrated Moving Average with exogenous variables
SCADA	Supervisory Control and Data Acquisition
SVM	Support Vector Machines
THD	Total Harmonic Distortion
TN	True Negative
TP	True Positive
V _{RMS}	Voltage Root Mean Square

V_THD	Voltage Total Harmonic Distortion
XGBoost	eXtreme Gradient Boosting

CHAPTER 1: INTRODUCTION

1.1 BACKGROUND OF THE STUDY: POWER QUALITY IN MODERN GRIDS

The global energy landscape is undergoing a profound transformation, driven by the urgent need to decarbonise and the corresponding rapid expansion of renewable energy sources (RES). Wind energy, in particular, has emerged as a cornerstone of this transition, offering a viable, large-scale alternative to conventional fossil fuel-based power generation. However, integrating wind power plants into existing electrical grids is not without its challenges. These plants are connected to the grid via power electronic converters, which, combined with the inherent intermittency and stochastic nature of wind, can introduce significant disturbances. These disturbances impact power quality (PQ), a critical measure of the stability and reliability of the electricity supply. Wind turbine output is nonlinear with respect to wind speed, featuring cut-in, ramp, rated plateau, and cut-out regions [1,2].

Power quality refers to the set of electrical parameters that define the interaction between a piece of equipment and the power grid. Deviations from the ideal, pure sinusoidal waveform, such as voltage sags, swells, flicker, harmonics, and unbalance, can have severe consequences. These PQ events can lead to equipment malfunction or failure, operational downtime in industrial processes, and a general degradation of grid reliability. Consequently, maintaining a high level of PQ is a paramount concern for utility operators, regulators, and consumers alike. As renewable energy penetration deepens, the imperative to move from reactive mitigation to proactive, predictive management of PQ has become more critical than ever [3,4].

In South Africa, this challenge is sharpened by the particular characteristics of the national grid and the policy context in which wind power plants operate. The South African Grid Code and its associated documents (including the Grid Code for Renewable Power Plants and the National Rationalised Specifications (NRS) 048 series specify detailed PQ requirements at the point of common coupling (PCC), as well as minimum performance obligations for generators and networks [5–7]. NRS 048-2 defines limits for voltage unbalance, total harmonic distortion

(THD), flicker, frequency and steady-state voltage variation, and it further prescribes how these must be measured and assessed over specific aggregation intervals.

Against this backdrop, wind farms connected to medium- and high-voltage networks must not only deliver energy but also maintain compliance with PQ limits under variable wind and grid conditions. Traditional PQ management has often been reactive, relying on after-the-fact analysis of recorded disturbances. However, as RES penetration increases and the grid operates closer to its technical limits, there is growing interest in forecasting PQ indicators in advance so that operators can intervene before formal non-compliance occurs. This motivates the development of forecasting frameworks that are both data-driven and explicitly aligned with standards such as NRS 048-2 and IEC 61000-4-30.

1.2 PROBLEM STATEMENT: PREDICTIVE CHALLENGES AND REGIONAL RESEARCH GAPS

Effective management of power quality (PQ) in grids with significant wind penetration requires forecasting disturbances before they breach operational or regulatory limits [8]. Classical statistical approaches often struggle with the non-linear, multi-scale dynamics at the RES grid interface, motivating the use of machine learning (ML) and deep learning for forecasting and early warning [8]. Recent pipelines show that data-driven models can anticipate PQ events when fed with suitably engineered features and operational context, but results remain domain and region-specific [7, 9].

A key gap is **regional transferability**: much of the validation is reported on European/Asian grids whose codes, network characteristics, and ambient conditions differ from sub-Saharan systems [10]. For South Africa, both the growth trajectory of renewables and distinct grid code requirements point to the need for locally validated models and compliance-aware evaluation [11-13]. In practice, forecasts must be consistent with formal measurement methods and limits defined by standards and codes such as NRS 048-2, IEC 61000-4-30, IEC 61000-4-15, and EN 50160 [13–16].

A further limitation in the literature is over reliance on error statistics (e.g., Mean Absolute Error

(MAE)/Root Mean Square Error (RMSE)) without explicitly quantifying operational risk, namely Missed Violation Rate (MVR) and False Alarm Rate (FAR), which better capture the cost of missed breaches and nuisance alarms for control rooms [17,18]. Addressing these gaps requires a localised study using South African wind-plant data, evaluated with both accuracy and compliance-oriented metrics, and reported with clear, reproducible workflows [16,19].

1.3 RESEARCH QUESTIONS

1.3.1 Primary question

How effectively can machine learning models forecast power quality parameters and predict NRS 048-2-based power quality violations at a South African wind power plant when evaluated using both conventional regression-accuracy metrics and practical compliance-oriented metrics?

1.3.2 Secondary questions

- Which machine learning algorithms (e.g., Boosted Trees (BT), Support Vector Machines (SVM)) minimise missed violations for key PQ parameters such as voltage unbalance and flicker?
- What is the practical utility of using compliance-focused metrics such as FAR and MVR for the operational assessment of PQ forecasting models?
- How can a sequential analytical workflow from data acquisition to feature engineering and model evaluation be structured to build a robust and reproducible pipeline for PQ forecasting?

1.4 AIM AND OBJECTIVES

To develop and evaluate a framework for forecasting power quality disturbances in a wind power plant, using ML techniques validated on real-world South African grid data.

To develop and evaluate a standards-aligned framework for forecasting power quality parameters and predicting NRS 048-2-based power quality violations at a wind power plant, using machine learning techniques validated on real world South African grid data and assessed with both regression accuracy and compliance-oriented metrics.

The objective is to:

- design and implement a comprehensive data-processing and feature-engineering pipeline for utility-scale power-quality time series at the point of common coupling (PCC), incorporating electrical, meteorological, and temporal variables;
- train, validate, and test a suite of machine-learning (ML) models, including Support Vector Machines (SVM), Boosted Trees (BT), and Neural Networks (NN), alongside a classical statistical baseline, Seasonal Auto-Regressive Integrated Moving Average with exogenous variables (SARIMAX), for key power-quality parameters such as voltage unbalance, voltage total harmonic distortion (V_THD), short-term flicker (P_st), long-term flicker (P_lt), voltage root mean square in per unit (V_RMS p.u.), and frequency;
- evaluate the chosen models using conventional regression-accuracy metrics, including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and the Coefficient of Determination (R^2), together with compliance-oriented metrics, including Accuracy, False Alarm Rate (FAR), and Missed Violation Rate (MVR), derived from NRS 048-2 limits;
- analyse the results to identify models suitable for predictive grid management, interpret model behaviour in terms of power-quality physics and operational constraints, and provide evidence addressing the regional research gap in South African wind-farm power-quality forecasting.

1.5 SIGNIFICANCE OF THE RESEARCH

This study provides a South African context evaluation of PQ forecasting at a wind plant PCC, directly relevant to local operators and policy objectives [13-16]. By aligning forecasts and evaluations with NRS 048-2 and related grid code requirements, the work ensures that its conclusions are meaningful for compliance management rather than purely academic purposes. It introduces and operationalises metrics FAR and MVR alongside traditional accuracy measures. These metrics align model selection with operational priorities by quantifying any tradeoffs between missed violations (risk) and false alarms (operator burden) [20,21]. The work documents a reproducible analytics method in JMP Pro with fixed seeds, explicit configuration, and a clearly defined feature set. This is intended to facilitate

replication by other researchers and practitioners and to support transfer of the methodology to similar RES integration studies [19]. By focusing on a utility-scale South African wind farm and explicitly integrating standards, compliance metrics, and ML benchmarking, the dissertation fills an identified gap in the PQ forecasting literature, which is currently dominated by non-African case studies and accuracy-only lead evaluation.

1.6 DISSERTATION STRUCTURE

The dissertation is organised as follows Chapter 1 introduces the problem, research questions, aim, objectives, significance, and the overall structure of the work. It situates PQ forecasting within the global RES transition and the specific South African grid code context.

Chapter 2 reviews power quality phenomena in wind-integrated grids, relevant international and South African PQ standards, ML forecasting methods, and compliance frameworks. It identifies the research gaps that motivate the empirical work, with a particular focus on NRS 048-2 aligned, compliance-oriented forecasting in the South African context.

Chapter 3 details the methodology: datasets, measurement compliance, preprocessing, feature engineering, model configurations (SARIMAX, SVM, ANN, Boosted Trees), evaluation metrics (including FAR/MVR), and reproducibility measures.

Chapter 4 presents the empirical results, including comparative regression performance, compliance-first versus accuracy-first model selection, and stability across phases and ± 5 % threshold perturbations.

Chapter 5 interprets the empirical results from Chapters 3 and 4, situates them against the literature synthesised in Chapter 2 and publications, and discusses implications for predictive maintenance and grid-code compliance under the South African NRS 048-2 standard.

Chapter 6 concludes with explicit answers to the research questions, a consolidated statement of contributions, a transparent discussion of limitations, and concrete directions for future work.

1.7 CHAPTER SUMMARY

This chapter has set out to introduce the motivation for the study, provide a context of the problem, and the research design for the dissertation. It highlighted how increasing wind penetration and the specific PQ requirements, as highlighted in the South African Grid Code and NRS 048-2 measurement requirements and classification create a need for forecasting PQ parameters at the PCC, rather than merely reacting to measured disturbances. The chapter articulated a primary research question and supporting secondary questions, defined a clear aim and a set of objectives, and argued for the significance of a compliance-oriented, reproducible PQ forecasting framework based on real South African SCADA data.

The next chapter builds directly on this foundation. Chapter 2 provides the technical and regulatory background needed to interpret the subsequent empirical work. It reviews PQ phenomena and their measurement, introduces the key international and South African standards (including NRS 048-2 and IEC 61000-4-30), and surveys existing ML-based PQ forecasting approaches. It then synthesises the literature to identify specific gaps, particularly in compliance-oriented evaluation and African case studies.

CHAPTER 2: LITERATURE REVIEW

2.1 CHAPTER OVERVIEW

The purpose of this chapter is to provide the technical and regulatory foundations for the empirical work that follows. It first introduces power quality phenomena and explains how they are defined and measured, with emphasis on wind-integrated grids and the South African context. It then summarises the main international and national standardised 61000-4-30, IEC 61000-4-15, EN 50160, and NRS 048-2 that govern PQ measurement and compliance at the point of common coupling (PCC). The chapter then reviews how machine learning (ML) techniques are used in modern power systems, with a focus on PQ forecasting, and synthesises the findings of the systematic literature review (SLR) published as Cindi et al. (2024) [19]. This synthesis highlights key gaps in existing PQ forecasting studies: limited use of compliance-oriented metrics (Missed Violation Rate, False Alarm Rate), a strong regional bias away from African grids, and incomplete reporting of feature engineering and workflow design. The chapter concludes by presenting the conceptual framework that guides this dissertation.

2.2 POWER QUALITY PHENOMENA IN RENEWABLE INTEGRATED GRIDS

Power quality (PQ) describes the degree to which supply voltage, frequency, and waveform conform to specified tolerances for reliable operation. In an ideal system, voltage is perfectly sinusoidal at constant magnitude and frequency. In practice, grids are subject to voltage sags or dips, voltage swells or surges, short and long interruptions, harmonics, transients, unbalance, and flicker (short-term and or long-term). These deviations stress equipment, reduce efficiency, and can compromise system reliability [20–22].

The large-scale integration of renewable energy sources (RES), particularly wind and solar PV, has reshaped the PQ landscape over the last decade. Unlike synchronous generators, wind turbines and PV plants interface with the grid via power electronic converters. Their output is inherently variable, driven by wind speed and solar irradiance, and can change rapidly due to gusts, turbulence, or passing clouds. This variability introduces new PQ challenges at both plant and system levels [23,24–27].

2.2.1 PQ issues driven by wind and solar intermittency

International and regional studies consistently report four recurring PQ problems associated with wind and PV integration:

- **Voltage Fluctuations and Flicker**

Wind turbine output varies with turbulence and gusts; PV output swings as clouds transit across arrays. These fluctuations result in continuous oscillations in the voltage magnitude at the point of common coupling (PCC). When the oscillation frequency falls in the range of human visual sensitivity, customers experience light flicker, quantified by the short-term and long-term flicker indices P_{st} and P_{lt} as defined in IEC 61000-4-15. Field investigations have shown that independent wind farms can introduce significant flicker emissions at the grid interface, especially on weak networks and under high wind variability [28-30]. Similar behaviour is reported for PV plants under fast-moving cloud conditions, where abrupt irradiance changes produce voltage dips or spikes that manifest as flicker on distribution feeders [30].

- **Harmonic Distortion**

Both wind and PV rely on power electronic converters that switch at high frequencies, generating non-sinusoidal currents. These currents distort the voltage waveform, raising the total harmonic distortion (THD) and potentially exciting resonances in weak grids. Studies on variable-speed wind turbines and PV inverters show that harmonic emissions can be a limiting factor for plant connection, particularly when multiple units are connected to the same bus [4,29,31]. In South Africa, early grid-code experience indicated that some wind/PV plants struggled to meet strict per-harmonic emission limits. This led to the introduction of Group Harmonic Distortion limits for Renewable Power Plants (RPPs), in which collective site emissions are assessed using a 95th percentile envelope rather than rigid per-order limits [4].

- **Voltage Unbalance**

Voltage unbalance arises when the three-phase voltages differ in magnitude and/or phase angle. Large utility-scale wind and solar plants are typically designed as balanced three-phase facilities, but unbalance can still occur due to asymmetrical network impedances, single-phase loads, or unevenly distributed smaller PV installations. Excessive unbalance causes negative-

sequence currents that overheat generators and motors, accelerate insulation ageing, and can trigger protection [32]. Hence, grid codes, including NRS 048-2, specify strict limits on the three-phase voltage unbalance factor (typically $\leq 2-3\%$) to protect equipment and ensure compatibility.

- **Frequency Deviations and Reduced Inertia**

As converter-dominated resources displace synchronous generation, system inertia declines. Sudden changes in renewable output, for example, a rapid wind drops across a wind farm cluster or a steep PV ramp due to cloud fronts, translate into faster and larger system-wide frequency deviations, if additional fast reserves or synthetic inertia are not provided [33]. International experience confirms that high-RES penetration increases the rate of change of frequency (RoCoF) and heightens the risk of exceeding grid code frequency limits in weak or poorly compensated systems.

Other PQ phenomena, such as transient over-voltages, voltage sags/swells during faults or switching of wind/PV blocks, and inter-harmonics, are also reported, but flicker, harmonics, unbalance, and frequency variation are the most systematically studied in the context of RES integration [26,29].

2.2.2 PQ in South African and African Renewable Grids

Within the African context, PQ issues have been documented as utility-scale wind and PV capacity have grown and are continuously growing. In South Africa, the integration of large wind farms and solar PV projects under the Renewable Energy Independent Power Producer Procurement Programme (REIPPPP) has prompted the development of dedicated PQ guidelines for RPPs, recognising that PQ behaviour in a relatively weak, load-shedding-prone network differs from that in large, interconnected European systems [4,29].

Key South African observations include:

Voltage flicker at wind farm PCCs on long rural lines with high impedance, exacerbated during certain load-shedding stages when network topology changes. Elevated harmonic levels at some PV sites due to multiple inverters interacting with network resonance; this motivated the

move to Group Harmonic Distortion criteria using 95th percentile limits tailored to local impedance spectra [4]. Voltage unbalance sensitivities at PCCs where mixed industrial loads and RPP exports interact, underscoring the need for careful phase balancing and three-phase measurement. These experiences, combined with international evidence, reinforce the need for Class A PQ measurement at RPP PCCs and for forecasting tools that can anticipate PQ excursions before they materialise in 2-hour or 24-hour compliance windows.

2.3 STANDARDS AND REGULATORY FRAMEWORK FOR POWER QUALITY

The assessment of PQ in renewable-integrated grids is grounded in international measurement and compatibility standards and in regional or national grid codes.

2.3.1 International standards

To harmonise PQ definitions, measurement classes, and limit values, utilities rely on a hierarchy of international standards such as IEC and regional standards. Some of the international standards include:

- **IEC 61000-4-30:** defines Class A measurement methods, aggregation windows, and instrument performance requirements for PQ parameters. It ensures that measurements are reproducible and comparable across instruments and sites.
- **IEC 61000-4-15:** specifies the design and functional requirements of a flicker meter, including the calculation of P_{st} (10-minute) and P_{Lt} (2-hour) indices.
- **EN 50160:** describes voltage characteristics of electricity supplied by public distribution systems in Europe, including nominal voltage ranges, frequency band (typically 49.5–50.5 Hz), and THD limits by voltage level.
- **IEEE 519:** and related documents set recommended harmonic current and voltage limits at various network levels.

These standards ensure that PQ indices are defined consistently and that measurements from different sites and instruments are comparable [23–27].

Figure 2.1, (Global Power Quality Standards Comparative Framework) positions these standards in relation to each other and to NRS 048-2. It shows how IEC 61000-4-30 and IEC

61000-4-15 provide the measurement foundations, EN 50160 specifies typical European compatibility levels, and NRS 048-2 tailors similar concepts to South African grid conditions.

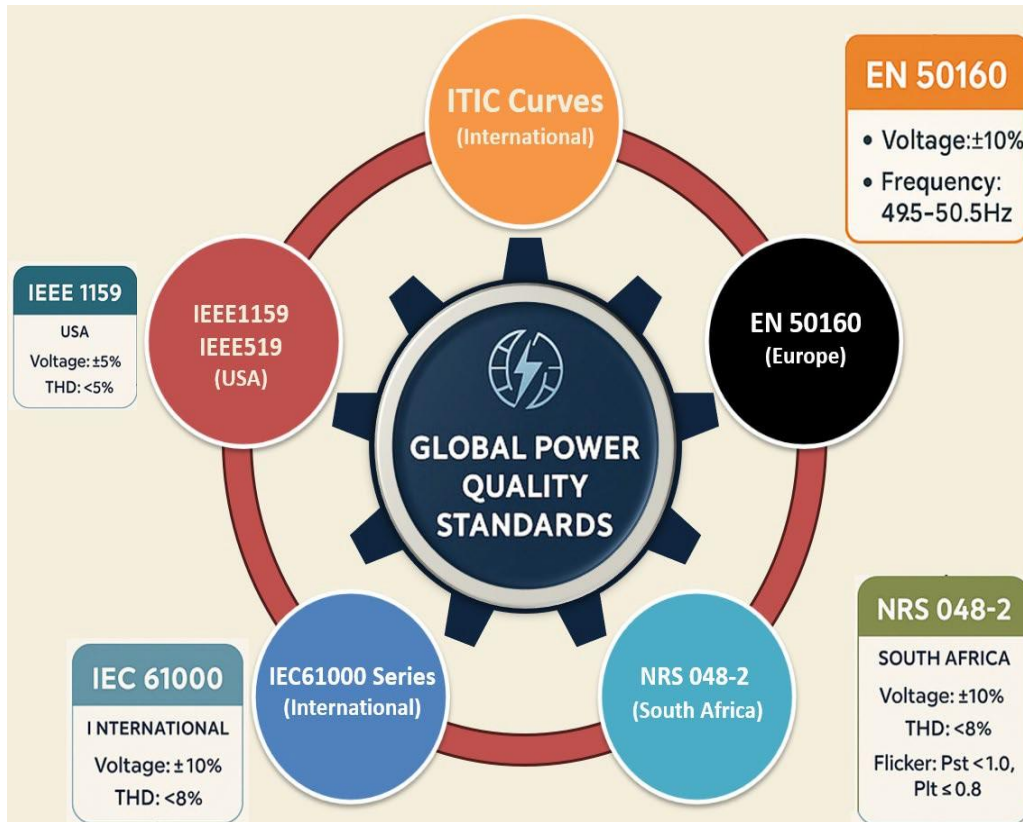


Figure 2.1: Global Power Quality Standards Comparative Framework

Table 2.1 summarises the scope of these standards, indicating for each the main PQ parameters, aggregation windows, and typical limits. This tabular overview provides a quick reference for the standards that underpin the compliance-oriented evaluation used in this dissertation.

Table 2.1: Overview of Key Global Power Quality Standards

Standard	Region / Scope	Key Specifications
IEC 61000-4-15	International (globally adaptable)	Defines the flicker meter functional and design requirements; formal definitions and calculation of Pst and Plt indices [15].
IEC 61000-4-30	International (globally adaptable)	Specifies Class A measurement methods, aggregation windows, and performance requirements for PQ parameters; ensures reproducible, comparable measurements [23].
EN 50160	Europe	Describes voltage characteristics at public distribution networks: voltage tolerance typically $\pm 10\%$, frequency 49.5–50.5 Hz, THD limits by network level, and assessment windows for PQ phenomena [22].
NRS 048-2	South Africa (tailored to SA grid)	Defines PQ limits and assessment at the PCC: steady-state voltage variation $\pm 10\%$, $P_{st} < 1.0$, $P_{lt} \leq 0.8$, $V_{THD} < 8\%$; requires IEC 61000-4-30 Class A measurement for compliance evaluation [20].

2.3.2 NRS 048-2 and the South African Grid Code

South Africa operationalises PQ requirements primarily through the Distribution Network Code, which is part of the broader Grid Code approved by NERSA (National Energy Regulator of South Africa). Distribution Code Use of System Agreement (DCUOSA) as a contractual framework within South Africa's electricity distribution network governs how users (such as municipalities, large customers, or embedded generators (wind power plants, Photo Voltaic plants)) connect to and make use of the distribution system. The DCUOSA primarily refers users to the NRS 048-2 compliance limits required for their particular plant at the PCC. The

NRS 048-2 specifies compatibility levels and limits for voltage magnitude, voltage unbalance, harmonic distortion, flicker, and frequency at the Point of Connection (PoC).

It outlines the measurement methods and aggregation intervals, referencing IEC 61000-4-30 Class A and IEC 61000-4-15; how compliance is assessed statistically (e.g., over weekly or longer time windows). At a high level, NRS 048-2 requires, for medium-voltage networks and PCCs relevant to wind farms, a steady-state voltage typically within $\pm 10\%$ of nominal (≈ 0.90 – 1.10 p.u.) for most 10-minute intervals. Voltage unbalance is limited to about 2% for three-phase systems, voltage THD (V_{THD}), is limited to 8% for LV/MV levels; short-term flicker (P_{st}) to less than 1.0; long-term flicker (P_{lt}) to less than or equal to 0.8 as a planning level; frequency (f) is typically between 49.5–50.5 Hz under normal operating conditions.

For this dissertation, these values define the compliance thresholds used to label each 10-minute or 2-hour interval as compliant or non-compliant. Importantly, NRS 048-2 requires that measurements used to assess compliance be obtained from Class A instruments, which is satisfied by the PQ meters at the case-study wind farm as discussed in Chapter 3.

Table 2.2 summarises and presents the NRS 048-2 limit numbers used in this study.

Table 2.2: NRS 048-2 limits used for compliance labelling in this study

Parameter	Symbol	Limit used in this study
Voltage Unbalance		<2% (95% weekly value)
Total Harmonic Distortion	V_{THD}	<8%, up to 50 th harmonic
Frequency	F	49.5 – 50.5Hz
Short-term Flicker	P_{st}	<1.0
Long-term Flicker	P_{lt}	<0.8
Fundamental Voltage (under)	V_{RMS} , p. u	>0.85 p. u
Fundamental Voltage (Over)	V_{RMS} , p. u	<1.10 p. u

These limits are applied at the PCC across phases. In later chapters, continuous forecasts of these indices are compared directly with these thresholds to derive binary compliance labels and compute the Missed Violation Rate (MVR) and the False Alarm Rate (FAR). NRS 048-2

also links naturally to machine learning. Because the standard defines both what to measure and how to aggregate it, it effectively dictates the target variables for ML models (e.g., 10-minute V_{THD} , P_{st} , P_{lt}) and the labelling rule for compliance/non-compliance. This makes it straightforward to interpret the output of ML models in operational terms: a forecast above a limit is a predicted violation; a forecast below is predicted compliance.

2.4 MONITORING, ASSESSMENT, AND MITIGATION OF PQ IN RENEWABLE PLANTS

PQ management in RES plants follows a sequence of measurement and assessment in accordance with outlined standards, and then a mitigation plan or strategy is implemented.

2.4.1 Measurement and assessment

PQ monitors, or Class A PQ analysers, are installed at the PCC of wind and PV plants to record voltage, current, harmonic spectra, and flicker indices. Data are aggregated over standard windows (e.g., 10-minute, 2-hour, weekly 95th percentile) consistent with IEC 61000-4-30 and national codes. Compliance is assessed by comparing these indices (e.g., Unbalance %, THD %, P_{st} , P_{lt}) against NRS 048-2 limits [4,25,29]. In South Africa, harmonics are often assessed using a 95th percentile envelope over a week, and flicker indices are checked against specified short- and long-term thresholds.

2.4.2 Mitigation strategies

Mitigation measures reported in the literature include advanced converter control (ramp-rate limiting, P–Q and V–F control, provision of virtual inertia); Static Var Compensator (SVC)/ Static Synchronous Compensator (STATCOM) for fast reactive power support and flicker damping; Dynamic Voltage Restorer (DVR) and Unified Power Quality Conditioner (UPQC) for sag/swell and combined PQ disturbance compensation; passive and active harmonic filters tuned to inverter emission spectra; network reinforcement or reconfiguration in extreme case; integration of battery energy storage systems (BESS) to smooth wind/PV ramping and provide frequency support [14,15,24,26,34].

2.4.3 Emerging role of forecasting and artificial intelligence (AI)

Recent work emphasises moving from purely reactive mitigation to predictive PQ management. High-resolution Renewable Energy Sources (RES) output forecasts, ramp-event prediction, and

ML-based disturbance classification are used to anticipate PQ stress and adjust controls pre-emptively. For example, dedicated ramp-forecasting algorithms (random forests, Long-Short Term Memory (LSTMs)) have been developed to predict sharp PV or wind drops, allowing storage or fast reserves to be scheduled in advance [34]. ML classifiers (SVMs, NNs) coupled with wavelet features are used to detect PQ disturbances in real time and trigger automatic mitigation [35]. These developments motivate the forecasting-oriented focus of this dissertation: rather than only measuring PQ after the fact, we aim to forecast PQ indices and compliance status in advance, using machine learning models trained on operational data from a South African wind plant.

2.5 MACHINE LEARNING APPLICATIONS IN POWER SYSTEMS AND PQ

Modern grids with a high penetration of RES exhibit bidirectional power flows, complex interactions between distributed assets, and operating regimes that are highly variable and often nonlinear. These stochastic behavioural characteristics strain traditional linear modelling tools, especially when PQ phenomena are driven by interactions among wind dynamics, converter behaviour, and grid conditions [28,35,36]. Within this context, a range of ML methods have been applied across power system tasks, including load and RES forecasting, fault and asset analytics, and PQ analysis: Artificial Neural Networks (ANNs) are flexible, data-driven models that approximate complex nonlinear mappings between inputs and outputs. In PQ contexts, multilayer perceptron (MLPs) and recurrent architectures have been used for both regression (forecasting THD, flicker, unbalance) and classification (for disturbance type recognition) [27,28]. They have often achieved improved performance over linear models and shallow ML when sufficient data are available [37]. In deep learning to forecast harmonic distortion and flicker in converters and arc furnace systems using hybrid CNN-LSTM-GRU architectures, capturing both local waveform features and longer-term temporal dependencies [36].

Support Vector Machines (SVMs) are margin-based models that use kernel functions (e.g., radial basis function (RBF)) to construct nonlinear decision boundaries in high-dimensional feature spaces. SVMs have been used for PQ event classification and short-term forecasting, particularly where sample sizes are moderate, and generalisation is critical [29,30]. They have been used to classify PQ disturbances (sags, swells, harmonics, flicker) based on time –

frequency features, achieving high accuracy under noise [11]. They are further used to forecast THD or specific harmonic components in PV and industrial systems, with grey-wolf or other optimisation schemes used to tune SVM hyperparameters and outperform baseline models [33].

Ensemble methods (Random Forests, Boosted Trees) are collections of shallow decision trees trained either in parallel (RF) or sequentially (Gradient Boosting) to improve accuracy and robustness. Boosted Trees are especially powerful for modelling interactions and threshold effects, which are common in PQ dynamics [31-33]. They have been used to forecast voltage, THD, and power factor in RES-rich microgrids, with boosted trees often delivering the best accuracy for certain PQ indices and random forests excelling in disturbance classification [38,39]. In other applications, they are used to predict PQ events (like voltage dips) in smart grids, with random forests achieving the highest classification accuracy among several ML classifiers [40].

Classical time-series models (ARIMA, SARIMA/SARIMAX) are linear models that explicitly capture autocorrelation, trend, and seasonality. SARIMAX incorporates exogenous variables (e.g., wind speed, reactive power) and provides a useful baseline for time-series forecasting in PQ applications [34]. ARMA/ARIMA models have been used to forecast grid frequency and voltage trends, providing interpretable linear baselines that capture seasonality and trends [39]. SARIMAX models with exogenous inputs (e.g., weather, load) have been compared with ANN and hybrid methods in RES forecasting contexts, often performing competitively for stable time series but struggling with strongly nonlinear behaviour [26].

African PQ literature incorporating ML is still emerging, but international methods and tools increasingly inform local practice, including in South African studies that use NNs and SVMs for disturbance classification and in pilot projects exploring data-driven PQ monitoring at RPP PCCs.

Figure 2.2 illustrates how these methods are used across grid operations, including PQ monitoring, predictive maintenance, fault detection, and load/RES forecasting. PQ forecasting

sits at the intersection of these areas: it relies on the same time-series and feature-engineering techniques as load forecasting, but targets PQ indices and their compliance with standards.

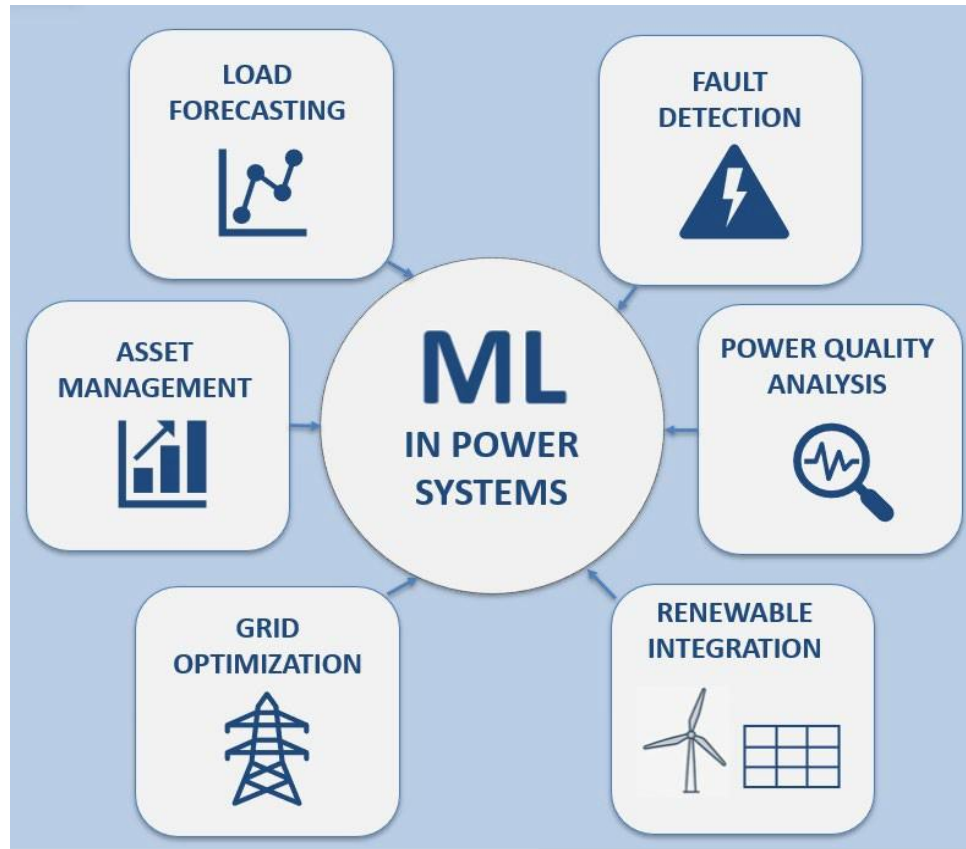


Figure 2.2: Diverse Applications of Machine Learning Across Modern Power Systems

2.6 MACHINE LEARNING ALGORITHMS FOR PQ FORECASTING: STRENGTHS AND WEAKNESSES

This dissertation focuses on four models classified families: SVM, ANN, Boosted Trees, and SARIMAX, that have been repeatedly applied in PQ or RES forecasting studies. Below, we highlight how each has been used in PQ contexts and the implications for our application.

Figure 2.3 provides a simple taxonomy of ML models relevant to this dissertation. It positions SVM, ANN, and Boosted Trees within supervised learning, alongside SARIMAX as a classical time-series method, and highlights unsupervised techniques (e.g., k-means, PCA, autoencoders) used for anomaly detection and pattern discovery in PQ and grid data [36,37].

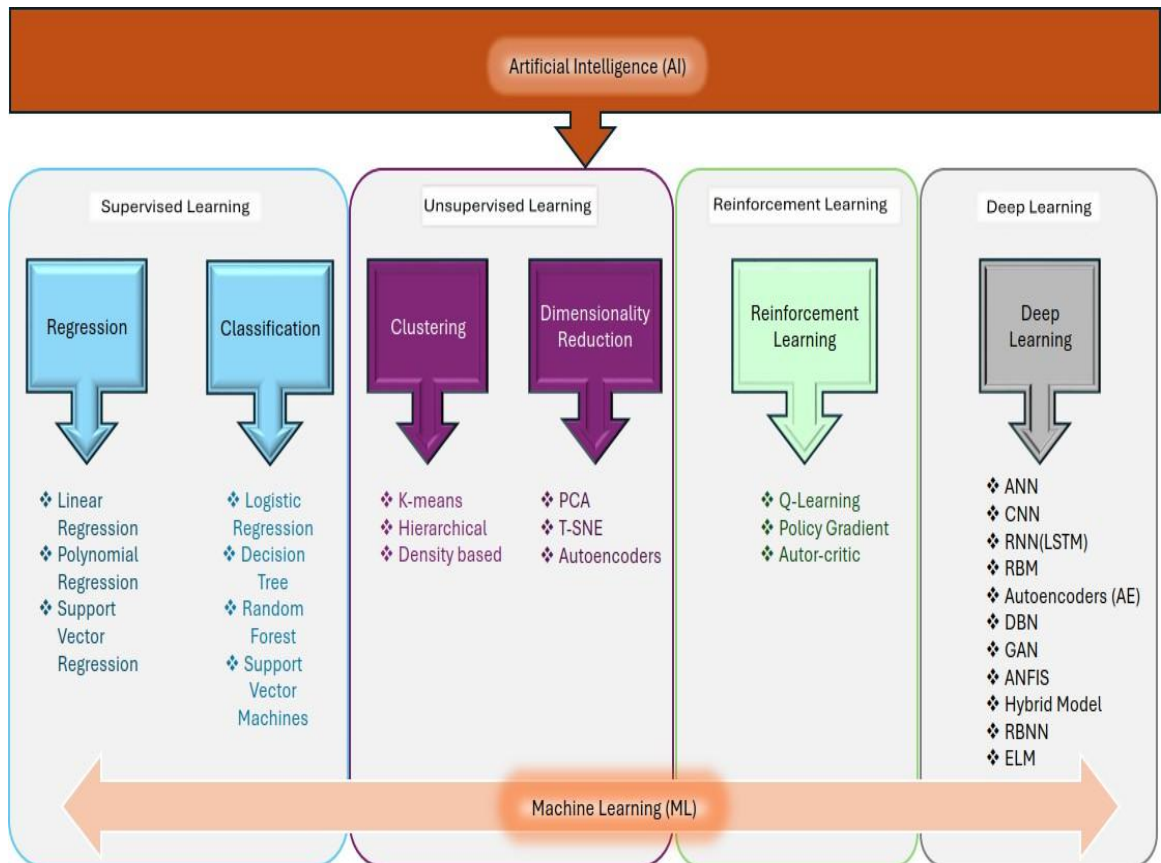


Figure 2.3: Overview of ML Algorithms and Techniques Relevant to PQ Forecasting

The four models selected represent a balance between practical deployment (no GPU or exotic infrastructure required) and expressive power for nonlinear time series. More advanced deep learning architectures (LSTMs, GRUs, CNNs, and transformers) are acknowledged but not implemented in this study; their role is discussed in Chapter 6 as part of the limitations and future work.

Advanced sequence and hybrid architectures provide a useful foundation for future work. Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks are designed to learn temporal dependencies directly from sequential data; Convolutional Neural Network-LSTM hybrids can combine local pattern extraction with temporal memory; and transformer-based models can learn longer-range dependencies through attention mechanisms. These approaches may be valuable for sub-minute waveform analysis or multi-step PQ forecasting, but they typically require larger datasets, more computational resources, and more extensive interpretability controls than were available in this single-site JMP Pro workflow [36,38].

2.6.1 Support Vector Machines (SVM)

SVMs have shown strong performance both in PQ disturbance classification and in regression for PQ indices such as THD and unbalance. Studies using multi-class SVMs with wavelet-based features report high accuracy in distinguishing sags, swells, harmonics, and flicker even under noisy conditions [11]. For regression, least-squares SVMs and grey-wolf-optimised SVMs have been used to forecast THD in PV systems and harmonic currents in industrial grids, often outperforming basic NNs and linear models when data are limited [33].

Strengths for our context: Robust generalisation in high-dimensional feature spaces; effective nonlinear modelling via kernel functions (especially RBF kernels); good performance with relatively modest datasets, which is important for a three-month PQ case study [41].

Weaknesses and mitigations: Absence of native temporal memory, addressed here by explicit feature engineering (lags, rolling windows, temporal indicators); sensitivity to hyperparameters C and γ is addressed via grid search on validation sets; limited interpretability compared to tree models, partially mitigated through SHAP-style and partial dependence analyses [41].

2.6.2 Artificial Neural Networks (ANN)

ANNs, including MLPs and deep architectures, are widely used for nonlinear PQ forecasting. Deep hybrids (CNN–LSTM–GRU) have been shown to forecast THD and flicker under highly variable operating conditions, capturing both time-local patterns and long-range dependencies [36]. MLPs have been used to predict harmonic currents or THD in PV and wind plants, with multiple studies reporting better accuracy than linear models when sufficient training data are available [42].

Strengths for our context: Ability to approximate complex nonlinear mappings between multivariate features (electrical, meteorological, temporal) and PQ outputs; flexibility to learn sporadic, bursty patterns such as short-term flicker P_{st} ; compatibility with multi-output setups if needed [42].

Weaknesses and mitigations: Data hungry and prone to overfitting on short datasets, here mitigated using aggressive regularisation (L2, dropout) and architectural restraint (2 hidden layers); ‘black-box’ perception addressed via residual diagnostics and sensitivity checks; more complex hyperparameter landscape than SVM or SARIMAX, handled via structured validation and fixed seeds [42].

2.6.3 Boosted Trees (BT)

Boosted tree ensembles (e.g., XGBoost-style Gradient Boosting) and Random Forests have gained prominence in PQ and RES forecasting due to their strong performance on structured or tabular data. Studies in RES-rich microgrids show that boosted trees often provide the lowest error for voltage and THD forecasts, sometimes rivalling deep NNs, while random forests achieve high accuracy in classifying PQ events [43,44]. Variable importance rankings derived from these models are also used to identify key drivers of PQ behaviour, such as specific harmonic orders, load patterns, or wind-voltage interactions.

Strengths for our context: High predictive accuracy on tabular time-series features (lags, statistics, interactions); built-in handling of nonlinear interactions and heterogeneous effects; strong interpretability: split-based variable importance and partial dependence plots reveal which features drive predictions; resilience to monotonic transformations and moderate outliers [42,44].

Weaknesses and mitigations: No inherent sequence memory, again addressed by explicit temporal feature engineering; risk of overfitting with deep trees or insufficient regularisation mitigated via constraints on depth, learning rate, and regularisation parameters (γ , L2); limited extrapolation beyond training ranges, acknowledged as a limitation; retraining is required when regimes shift [42].

2.6.4 SARIMAX (Seasonal ARIMA with Exogenous Inputs)

ARIMA algorithm models remain standard tools for time-series forecasting in power systems and have been applied to PQ-related metrics such as frequency, voltage, and harmonic indices

[26,39]. SARIMAX models, which include exogenous drivers, are used to capture seasonality and linear dependencies in variables such as load or weather. Comparative studies have shown that SARIMAX can perform competitively with ANN for relatively smooth, trend-dominated series, but underperforms when dynamics are strongly nonlinear or regime-switching [26].

Strengths for our context: It is generally well understood, with interpretable parameterisation (AR, MA, seasonal orders); explicit modelling of seasonality (e.g., daily patterns at 10-minute resolution); low computational cost, fast training and inference; natural provision of confidence intervals and residual diagnostics [39].

Weaknesses and mitigations: SARIMAX assume that the dominant temporal structure can be represented through linear autoregressive, moving-average, seasonal, and exogenous terms. Time-series forecasting literature and power-system forecasting studies show that such linear models may underperform when the target contains threshold effects, regime changes, converter-control nonlinearities, or non-stationary disturbance bursts [34,45,57]. For this reason, SARIMAX is used here as an interpretable benchmark rather than as the expected best-performing model for strongly nonlinear voltage-unbalance or flicker behaviour.

2.7 CRITICAL REVIEW OF ML-BASED PQ FORECASTING

The systematic literature review in Cindi et al. (2024) [19] screened PQ forecasting studies and ML applications relevant to wind power. The main trends can be summarised as follows:

A shift from purely statistical to ML-based methods, as early PQ forecasting studies relied on AR, ARIMA, or basic regression models. Over time, these have been complemented or replaced by ML methods such as ANNs, SVMs, Random Forests, Boosted Trees, especially in RES-rich grids where non-stationarity and nonlinear interactions are pronounced [38,39]. Studies forecast a range of diverse targets and horizons for PQ indices, including THD, flicker indices (P_{st} , P_{lt}), Voltage Unbalance, RMS voltage, and even waveform-level features. Forecasting horizons range from one step ahead to several minutes or hours, reflecting different operational needs.

Widely used ML models include ANNs for nonlinear regression/classification [28,38]. SVMs for high-dimensional decision boundaries and robust generalisation [30,31]. Tree ensembles (RF/Boosted Trees) are used for robust accuracy and interactions [32,33]. Hybrid signal processing/ML approaches, where time-frequency features (e.g., wavelets) feed ML models.

Despite promising results, two pervasive gaps emerge from the systematic literature review (SLR):

Metric gap: Evaluation is dominated by error metrics such as MAE, RMSE, and R^2 , which do not directly capture compliance risk at PQ limits [34,35]. Missed Violation Rate (MVR) and False Alarm Rate (FAR), which quantifies missed breaches and nuisance alarms, is rarely reported, even though they are more meaningful for operator decision making.

Regional gap: Most validation studies are based on European, North American, or Asian grids, with very limited coverage of African networks, including South Africa [19,27,36,37]. Differences in grid code requirements, network strength, and operational practices mean that results from other regions cannot be assumed to transfer directly.

Other recurring issues include limited transparency around feature engineering and validation workflows, and a lack of clear alignment between model outputs and the standards (e.g., NRS 048-2) that define PQ compliance. These gaps motivate the present dissertation: a South African case study that uses class-compliant PQ data from a wind farm at PCC/PoC. Evaluates models with both accuracy and compliance-oriented metrics and documents the workflow in sufficient detail to support reproducibility.

2.8 CONCEPTUAL FRAMEWORK GUIDING THE STUDY

To be useful to operators, PQ forecasting models must do more than minimise RMSE; they must also help maintain compliance with standards and support practical decision-making. The conceptual framework adopted in this dissertation, therefore, integrates three elements: Localised, standards-aligned data on PQ and SCADA data taken from a South African wind farm PCC and measured with IEC 61000-4-30 Class A instruments. Limits are based on NRS 048-2, ensuring that the study directly reflects local grid code requirements. Advanced, yet

deployable, four ML models: Boosted Trees, SVM, ANN, and SARIMAX, are used to capture both linear and nonlinear dependencies in the PQ time series, while remaining feasible for deployment in typical utility IT environments. Compliance-centric evaluation through continuous forecasting is mapped to binary compliance/non-compliance labels using the NRS 048-2 limits summarised in Section 2.3.2. Evaluation uses both traditional regression metrics and compliance metrics (Accuracy, FAR, MVR). FAR and MVR are interpreted as proxies for operational risk: MVR captures the cost of missed violations; FAR captures the burden of false alarms.

Figure 2.4 (Conceptual Framework for PQ Forecasting and Compliance Assessment) visualises these elements and their interactions. Raw SCADA and meteorological data are preprocessed and feature-engineered for the four models, whose outputs are evaluated both as continuous forecasts and as compliance labels. The framework also highlights the feedback from evaluation back to model selection and operational policy.

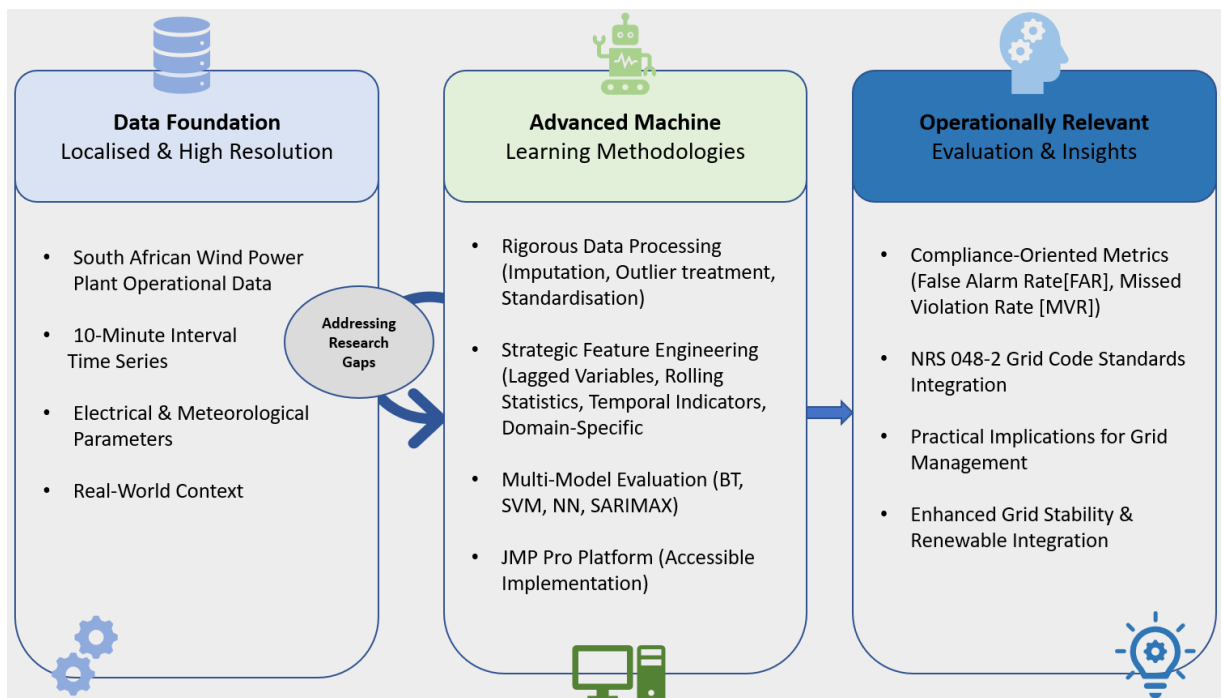


Figure 2.4: Conceptual Framework for PQ Forecasting and Compliance Assessment

The framework comprises:

Standards-aligned PQ measurement and labelling: Class A PQ data (Voltage Unbalance, V_THD, V_RMS, Frequency, Pst/Plt) are measured at the wind-farm PCC in accordance with IEC 61000-4-30 and IEC 61000-4-15. NRS 048-2 limits are applied to these time series to derive binary labels for compliance/non-compliance.

Feature-rich time-series modelling: High-resolution PQ and SCADA data (including wind speed, ambient temperature, power flows, and temporal indicators) are transformed into engineered features: lags, rolling windows, and interaction terms to feed the four model families (SVM, ANN, BT, SARIMAX).

Dual-track evaluation: Each model is evaluated on continuous forecasting accuracy (RMSE, MAE, MAPE, R^2), compliance-oriented performance (MVR, FAR) against NRS 048-2 thresholds, with additional threshold sensitivity ($\pm 5\%$) to probe robustness.

Suitability mapping and operational interpretation: Results are synthesised into a parameter model suitability map and translated into operator-relevant insights, linking forecast outputs to possible remedial actions (e.g., tap changes, VAR control, curtailment, storage dispatch).

This conceptual framework directly informs the methodological workflow in Chapter 3 and ensures that the research questions in Chapter 1 are addressed in a way that is coherent with both the scientific literature and South African grid code practice.

2.9 CHAPTER SUMMARY

This chapter has characterised the main PQ phenomena relevant to wind- and solar-integrated grids and summarised how they are defined, measured, and regulated in international standards and in the South African NRS 048-2 framework. It has shown that PQ indices such as Voltage Unbalance, V_THD, flicker, V_RMS, and Frequency must be understood not just as continuous time-series variables but as compliance indicators relative to prescribed limits, particularly in renewable-rich environments where converter-driven variability is significant.

The chapter has also reviewed ML methods used in PQ forecasting and related power-system tasks, with a specific focus on SVM, ANN, Boosted Trees, and SARIMAX, highlighting their applications, strengths, and limitations for PQ analysis. The literature review revealed persistent gaps: reliance on error-only evaluation, limited compliance-oriented metrics (FAR, MVR), and the under-representation of African grids. In response, a conceptual framework was proposed that integrates South African Class A PQ measurements, NRS 048-2 compliance limits, and a suite of ML models evaluated with both regression and compliance metrics. The next chapter builds directly on this foundation. Chapter 3 operationalises the conceptual framework by presenting the data sources, measurement compliance, preprocessing and feature-engineering pipeline, model configurations and deployment in JMP Pro, validation design, and evaluation metrics. It describes, step by step, how raw 10-minute PQ and SCADA measurements from the wind farm are transformed into standards-aligned forecasts and compliance labels that can be meaningfully compared across ML models.

CHAPTER 3: RESEARCH METHODOLOGY

3.1 WORKFLOW OVERVIEW

This study follows a quantitative, empirical case-study design with a sequential analytical workflow implemented end-to-end in JMP Pro 18 (Student Edition). This design was selected for four reasons. First, the research questions require numerical comparison of forecast accuracy and standards-based compliance risk. Second, power-quality data are time ordered; therefore, the workflow must preserve chronology and avoid look-ahead leakage. Third, JMP Pro provides a reproducible graphical environment that is accessible to utility practitioners who may not have specialised programming resources. Fourth, the six-phase structure makes each modelling decision auditable, from data acquisition and preprocessing through to NRS 048-2 compliance-oriented evaluation [41,42].

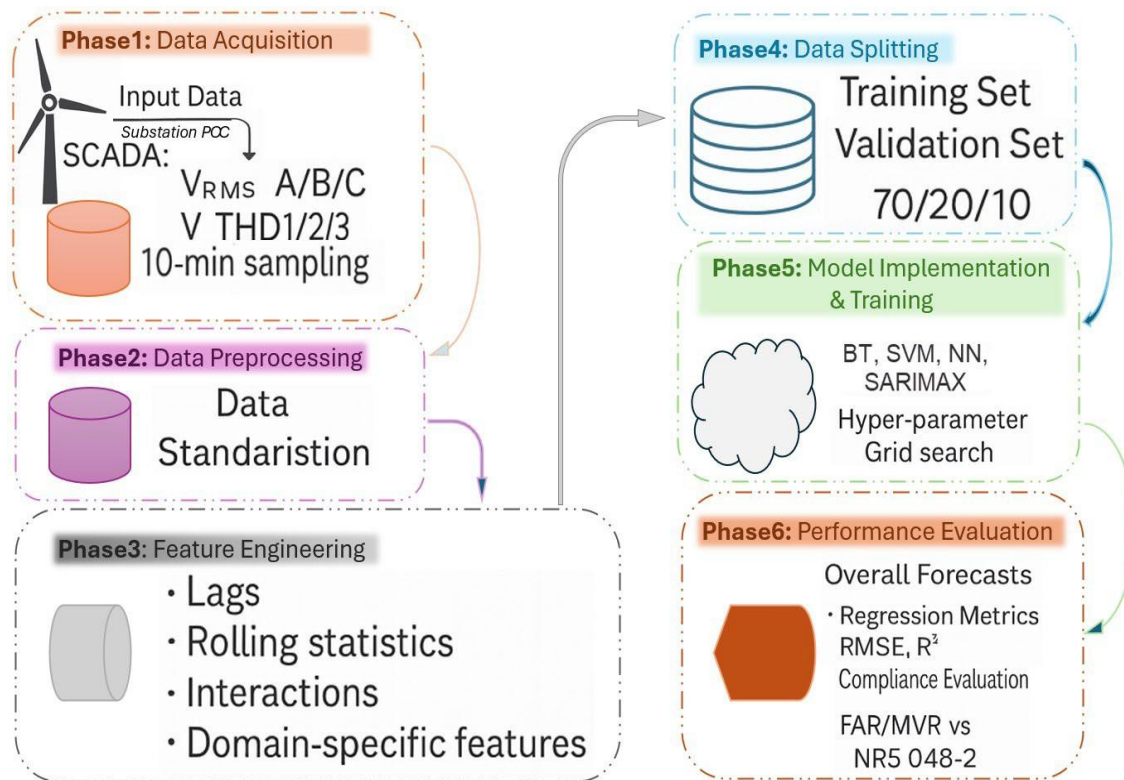


Figure 3.1: Research Analytical Workflow (flowchart)

Figure 3.1 summarises the end-to-end workflow used in this study, which is structured into six sequential phases from data acquisition to compliance-oriented evaluation:

Phase 1 – Data acquisition: 10-minute SCADA measurements of PQ indices at the wind farm PCC (Voltage Unbalance, V_THD, flicker P_st/P_lt, V_RMS, Frequency) are collected together with active/reactive power flows and on-site meteorological data (wind speed, ambient temperature, humidity) for the period [2021-03-01 to 2021-05-31]. This period was the complete continuous, quality-assured dataset released under the operator confidentiality agreement for this study. It is sufficient for a controlled single-site methodological case study, but it is not sufficient to claim full annual seasonality or all-season generalisation. This boundary is explicitly treated as a limitation in Chapters 5 and 6.

Phase 2 – Data preprocessing: Raw measurements are screened for missing values and obvious sensor artefacts. Short gaps are imputed (e.g., via linear interpolation for gaps shorter than 30 minutes), and outliers are mitigated using Hampel filtering while preserving genuine PQ disturbances. Continuous predictors are then standardised (Z-score normalisation) based on statistics from the training set.

Phase 3 – Feature engineering: A domain-informed feature set is constructed that combines base electrical, meteorological and temporal variables with lagged values, rolling-window statistics and a small number of interaction terms. This step captures both short-term persistence and the influence of exogenous drivers on PQ behaviour and is common to all models.

Phase 4 – Data splitting: The time-ordered dataset is partitioned into training, validation and test sets using a chronological split to avoid look-ahead bias. Typical proportions are 70% for training, 20% for validation, and 10% for testing, though the exact split is adapted to each PQ target to ensure sufficient events in the test set.

Phase 5 – Model training and tuning: Four model families - SARIMAX, SVM, NN and Boosted Trees - are trained for each PQ parameter. Hyperparameters (e.g., AR/MA orders, SVM C and γ , NN architecture, number and depth of trees) are selected based on validation performance, balancing forecast accuracy with model complexity and training time. The same feature set and data splits are used across models to ensure a fair comparison.

Phase 6 – Compliance-oriented evaluation: Out-of-sample forecasts on the test set are first assessed using regression metrics (RMSE, MAE, MAPE, R^2). Forecasts are then converted to binary compliant/non-compliant labels using the NRS 048-2 limits in Table 2.2 and Missed Violation Rate (MVR) and False Alarm Rate (FAR) are computed for each model. A $\pm 5\%$ threshold-sensitivity analysis perturbs the NRS 048-2 limits to gauge the robustness of these compliance metrics to changes in regulatory margins.

This workflow makes explicit where the NRS 048-2 standard enters the analysis: compliance thresholds are applied only at the evaluation stage, allowing the underlying regression models to learn continuous PQ dynamics while the final performance is expressed in terms that are directly meaningful to system operators.

3.2 DATA HANDLING AND RELATED MEASUREMENTS

The case-study data were obtained from the wind-farm point of common coupling under a confidentiality agreement with the operator. The available records consist of continuous 10-minute operational measurements from 2021-03-01 to 2021-05-31, after removal of unusable sensor and communication-fault periods. The data therefore cover late-summer to autumn operating conditions for this South African site, including normal plant operation, wind-ramp behaviour, and local grid interactions observed during the release window.

The period was not chosen to represent an annual NRS 048-2 compliance campaign. Rather, it was the longest internally consistent and permissioned dataset available for this methodological evaluation. The study therefore treats the three-month window as a single-site empirical case study for comparing modelling approaches, not as evidence that the same model coefficients will remain valid across winter, spring, or multi-year operating regimes. This distinction is important because seasonal wind distributions, load profiles, network configuration, and temperature-dependent equipment behaviour may change outside the observation window.

Potential data-acquisition limitations include the absence of annual seasonality, limited exposure to rare disturbance events, possible under-representation of winter high-wind regimes, and the lack of external grid-wide variables such as neighbouring feeder conditions, switching logs, and system load. These limitations are handled by using chronological

validation, avoiding random temporal mixing, reporting event scarcity explicitly, and framing transferability as future work rather than as a claim established by the present dataset.

The 10-minute operational data measurements from a utility-scale South African wind plant, include Voltage Unbalance, V_{THD} , flicker (P_{st}/P_{lt}), Frequency, V_{RMS} (p.u.), and associated meteorological variables (e.g., wind speed) at the PCC. Measurements are obtained from PQ meters that comply with IEC 61000-4-30 (Class A) measurement methods and IEC 61000-4-15 for flicker, ensuring instrument-grade comparability and standards-aligned evaluation [43,44].

Regulatory interpretation follows NRS 048-2 (Edition 5, 2025) clauses for: Voltage variation (steady-state voltage band), voltage unbalance, total harmonic distortion (THD), flicker (P_{st}/P_{lt}), frequency, as applicable to Renewable Power Plants (RPPs) [42]. These clauses specify both the numerical limits and the aggregation windows for compliance assessment. The limits used for labelling in this study are those summarised in Chapter 2 (V_{RMS} between 0.90–1.10 p.u., Voltage Unbalance $\leq 2\%$, $V_{THD} < 8\%$, $P_{st} < 1.0$, $P_{lt} \leq 0.8$, Frequency (49.5–50.5 Hz).

By ensuring that both the measurement process and the compliance criteria conform to these standards, the study aligns its forecasting targets directly with the quantities used in regulatory PQ assessments.

3.2.1 Targets and Labelling

Continuous forecasts are produced per PQ parameter per phase as continuous-valued time-series predictions. For each 10-minute interval, the models output predicted values of Voltage Unbalance, V_{THD} , P_{st} , P_{lt} , V_{RMS} (p.u.), and Frequency. For compliance evaluation, these continuous predictions are mapped to binary flags using the NRS 048-2 thresholds, for example: $V_{Unbalance} \leq 2\%$; $V_{THD} \leq 8\%$; $P_{st} < 1.0$, $P_{lt} \leq 0.8$; Frequency 49.5–50.5 Hz, V_{RMS} (p.u.) between 0.90 and 1.10. A system-level non-compliance is triggered if any phase violates its limit at a given time step (an operator

realistic OR logic at phase level, interpreted as a single PCC event) [42,45]. Each model is evaluated against a paired evaluation form, for Regression accuracy on continuous forecasts (RMSE, MAE, MAPE, R^2), and Compliance classification via confusion-matrix terms (TP, TN, FP, FN) and derived metrics (Accuracy, FAR, MVR). This dual view makes it possible to distinguish between models that are accurate in a statistical sense and those that are more reliable at avoiding missed violations or excessive false alarms.

3.2.2 Data Preprocessing

To safeguard model validity and avoid artefacts that could distort PQ forecasts, several preprocessing steps were performed:

Gap handling via short gaps in the time series (e.g., spans shorter than 30 minutes) were imputed conservatively, using linear interpolation. Longer gaps were excluded from model fitting and evaluation to avoid injecting synthetic structure where real data were unavailable. Outlier treatment via distributional diagnostics and initial residual checks were used to identify extreme values likely due to sensor or communication faults rather than genuine PQ events. These were treated using Hampel filtering, replacing extreme points while preserving legitimate excursions.

Standardisation of continuous predictors was performed using Z-score normalisation, with the mean and standard deviation learned from the training set and applied to the validation and test sets. This stabilises numerical optimisation and ensures comparability across features [46].

Time alignment and leakage control were handled by ensuring that all features were constructed so that, at prediction time $t+1$, only information from times $\leq t$ was used. Time

alignment and roll-safe windowing ensured that no operation consumed future information, enforcing a strict anti-leakage discipline [47].

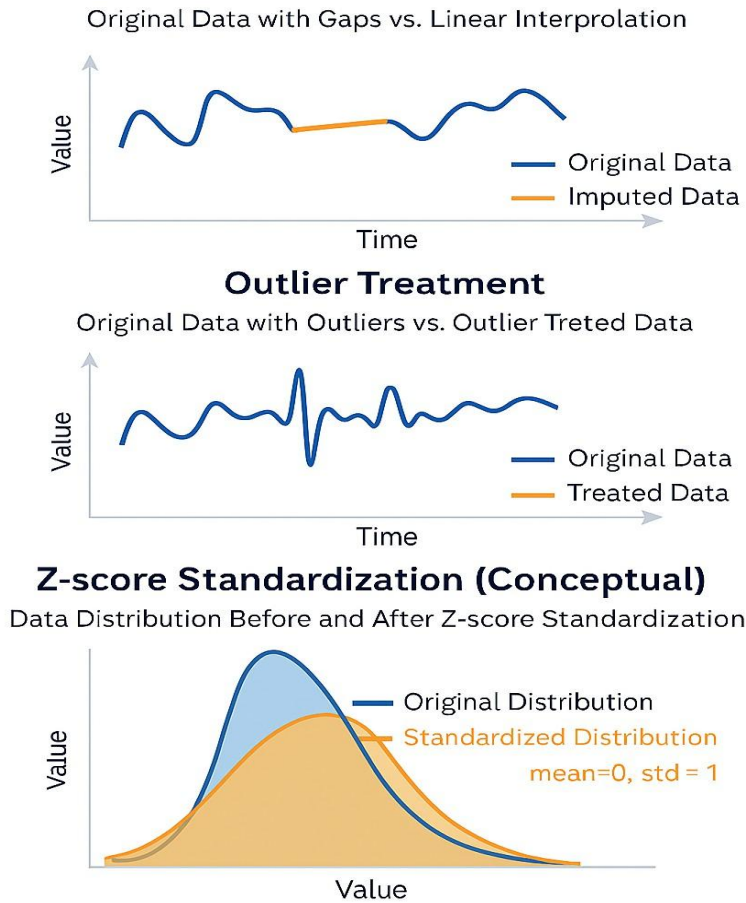


Figure 3.2: Generic Illustration of Data Preprocessing Steps (Image generated using GPT-4o)

Figure 3.2 illustrates the generic sequential steps of preprocessing operations, from raw SCADA feeds through gap handling, outlier mitigation, standardisation, and time alignment, highlighting the transformations that lead to the cleaned dataset used in subsequent phases.

3.2.3 Feature Engineering

In data science, feature engineering refers to the process of creating, transforming, or selecting variables (features) to improve model performance. For this study, time-series features were constructed to capture persistence, seasonality, and short-term dynamics. Temporal indicators with circular encoding, such as hour-of-day and day-of-week, were used to preserve periodic

adjacency (e.g., 23:00 adjacent to 00:00). Lagged variables were tuned for each target using partial autocorrelation analysis, with voltage metrics incorporating lags 1–6, frequency and flicker lags 1–12, and total harmonic distortion of voltage (V_THD) lags 1–3. Rolling statistics, including moving averages, moving standard deviations, and exponential moving averages with 1–4-hour windows, were introduced to capture local trends and volatility.

Interaction terms, such as wind-speed $\times$ unbalance, and voltage $\times$ reactive power, were included to expose non-linear coupling effects [48]. Additional statistical dynamics, such as rate-of-change and volatility ratios, were derived alongside domain-specific features, including a power-factor proxy and three-phase voltage-unbalance percentage. To enforce featural parsimony, tree-based importance measures, permutation drop tests, and SHAP-style attributions were applied to retain informative signals while curbing dimensionality [49]. The final feature set comprised approximately 54 inputs, combining both engineered and raw variables.

Table 3.1: Summary of engineered features (by category, purpose)

Feature Category	Number of Features	Description	Purpose
Temporal Indicators	6	Cyclical time encoding	Capture seasonality
Lagged Variables	24	Historical values (1-12 lags)	Provide historical context
Rolling Statistics	12	MA, MSD, EMA various windows	Capture trends and volatility
Interaction Terms	8	Cross-parameter products	Model synergistic effects
Statistical Features	6	ROC, volatility, etc.	Quantify rate of change
Domain-Specific	4	Power factor, unbalance, etc.	Incorporate engineering knowledge
Total	60		

Table 3.1 summarises the engineered feature set used across all models, grouped into temporal indicators, lagged PQ variables, rolling statistics, interaction terms, and domain-specific indices such as power factor proxies and voltage-unbalance measures. This table provides a clear view of the information on this feature engineering strategy, demonstrating that the forecasting task is driven not only by past PQ values but also by exogenous variables, such as wind speed and reactive power flows, that influence PQ behaviour at the PCC.

3.3 MODELS AND TRAINING

We evaluate four ML algorithms that jointly span kernel methods, neural approximators, boosted ensembles, and seasonal linear baselines. All models are trained per target and per phase, using the common feature set and chronological train/validation/test split.

For reproducibility, each model was trained under an explicit optimisation objective. For all continuous power quality targets, the primary empirical risk was squared-error loss, $L(y, \hat{y}) = (y - \hat{y})^2$, or its regularised equivalent. This choice is consistent with the use of RMSE/MSE as primary continuous performance metrics and with JMP Pro regression defaults.

For Support Vector Regression, the optimisation uses ϵ -insensitive loss with regularisation: $\min 1/2\|w\|^2 + C\sum (\xi_i + \xi_i^*)$ subject to $|y_i - f(x_i)| \leq \epsilon + \xi_i$, where C controls the penalty for deviations outside the ϵ -tube. For the Neural Network, the loss is mean squared error with L2 weight decay, $L_{NN} = (1/n) \sum (y_i - \hat{y}_i)^2 + \lambda \sum \|W_l\|^2$, with dropout used as an additional regularisation mechanism. For Boosted Trees, the stage-wise ensemble minimises squared-error loss plus tree-complexity regularisation, $L_{BT} = \sum l(y_i, \hat{y}_i) + \sum \Omega(f_k)$, where Ω penalises tree depth, leaves, and leaf weights. For SARIMAX, parameters are estimated by maximising the Gaussian likelihood of the one-step-ahead residuals, which is equivalent to minimising the sum of squared innovations under normal-error assumptions.

After regression training, no model was directly optimised on FAR or MVR. Compliance metrics were computed post hoc by thresholding the continuous forecasts against NRS 048-2 limits. This separation preserves continuous physical interpretability while allowing operational risk to be assessed through the confusion matrix.

3.3.1 Support Vector Machines (SVM)

SVM seeks a large-margin solution with regularisation; for regression/classification. The Problem is:

$$\min_{w,b} \frac{1}{2} ||w||^2 + C \sum_{i=1}^n \max(0, 1 - y_i(w \cdot x_i + b)) \quad (1)$$

Where w is the weight vector b = bias, C = regularization parameter, x_i = feature vector, $y_i \in \{-1, +1\}$ the class labels.

For regression, the ϵ -SVR variant is used, but the same kernel-based intuition applies. To capture nonlinear structure in PQ data, SVMs employ kernel functions that implicitly map input features into a higher-dimensional space. The Radial Basis Function (RBF) kernel was selected for its effectiveness:

$$K(x_i, x_j) = e^{-\gamma ||x_i - x_j||^2} \quad (2)$$

With

$$\gamma = \frac{1}{2\sigma^2} \quad (3)$$

σ controls the spread of the kernel.

Table 3.2: Implementation in JMP Pro.

Parameter	Description
Platform	Predictive Modeling Support Vector Machines
Kernel	Radial Basis Function (RBF)
Hyperparameters	Grid-tuned around $C=1.0$ and $\gamma=0.01$ on the validation set
Target types	Regression for continuous PQ indices (Voltage Unbalance, Frequency, V_THD, V_RMS p.u., Flicker Pst/Plt).
Classification	Classification layer applied post-hoc via thresholding of regression outputs against NRS 048-2 limits
Validation	Chronological train/validation/test split; within training plus validation, time-ordered subsets are used for parameter tuning (no random K-fold mixing across time).

Design rationale and strengths leveraged

SVM was selected for its strong performance on small-to-medium datasets and its capacity to model nonlinear decision boundaries with limited risk of overfitting when properly regularised, as highlighted in literature reports. In this application: The RBF kernel exploits smooth nonlinearities in PQ behaviour, particularly for V_THD and unbalance. The model handles the high-dimensional engineered feature set (lags, rolling statistics, interactions) effectively. This suitability was verified empirically rather than assumed: validation-set grid search selected stable C and gamma values, and the held-out test results showed strong performance for V_THD ($R^2 \approx 0.97$) and Voltage Unbalance ($R^2 \approx 0.95$). The claim of robust generalisation therefore rests on both prior literature and the chronological train/validation/test evidence reported in Chapter 4, not on model reputation alone. It offered robust generalisation even with a three-month dataset, which was important given the limited historical violations [50-52].

Weaknesses addressed

No native time-series memory: This is mitigated through explicit temporal feature engineering (lags, rolling windows, time-of-day indicators). Hyperparameter sensitivity: addressed via grid search and validation-based selection of C and γ . Interpretability: improved by inspecting variable importance (via permutation/gradient methods) and partial dependence plots, allowing us to relate SVM behaviour to PQ drivers near compliance thresholds.

3.3.2 Artificial Neural Networks (ANN)

A feed-forward 2-hidden-layer MLP (32→16) with ReLU hidden activations and linear output was trained with Adam optimizer; dropout ($p=0.2$) and L2 (0.01) mitigated overfitting, with batch normalization to stabilize gradients [53,54,55]. A generic forward pass follows:

$$\mathbf{z}^{(l)} = \mathbf{W}^{(l)} \mathbf{a}^{(l-1)} + \mathbf{b}^{(l)} \mathbf{a}^{(l)} = \mathbf{f}(\mathbf{z}^{(l)}) \quad (4)$$

Where $\mathbf{W}^{(l)}$ and $\mathbf{b}^{(l)}$ are the weight matrix connecting layer and $\mathbf{f}$ is the activation function (ReLU for hidden layers, linear for output).

Table 3.3: Implementation in JMP Pro.

Parameter	Description
Platform	Predictive Modeling Neural.
Architecture	2 hidden layers with 32 and 16 neurons respectively (32, 16).
Activations	ReLU in hidden layers, linear activation in the output layer for regression
Regularisation	Dropout ($p=0.2$) and L2 weight decay ($\lambda=0.01$) to mitigate overfitting.
Optimiser	Adam-style optimiser with mini-batch training
Training strategy	Chronological train/validation/test split; within training plus Early stopping based on validation loss monitored on the chronological validation set. ANNs were trained separately for each PQ target, producing per-phase continuous forecasts assessed against NRS 048-2 limits.

Design rationale and strengths leveraged

ANNs are flexible universal approximators, well-suited to learn short-burst nonlinear patterns such as those seen in flicker PST, where literature suggests that deep or shallow NNs can outperform linear models. In this study:

The moderate 2-layer architecture balanced capacity and generalisation, given the limited 3-month dataset.

Batch normalisation and dropout was used to improve training stability and robustness. ANNs could capture subtle nonlinearities and interactions beyond the reach of SARIMAX, supporting the analysis of parameters like flicker PST and flicker P_{It}.

Weaknesses addressed

Data hunger and overfitting risk: This was addressed via strong regularisation (dropout, L2), resulting in a restrained architecture, and early stopping [56]. Black-box nature: partially mitigated using residual diagnostics (ACF/PACF of residuals), sensitivity checks, and comparison to simpler models (SARIMAX, BT) to ensure plausible behaviour.

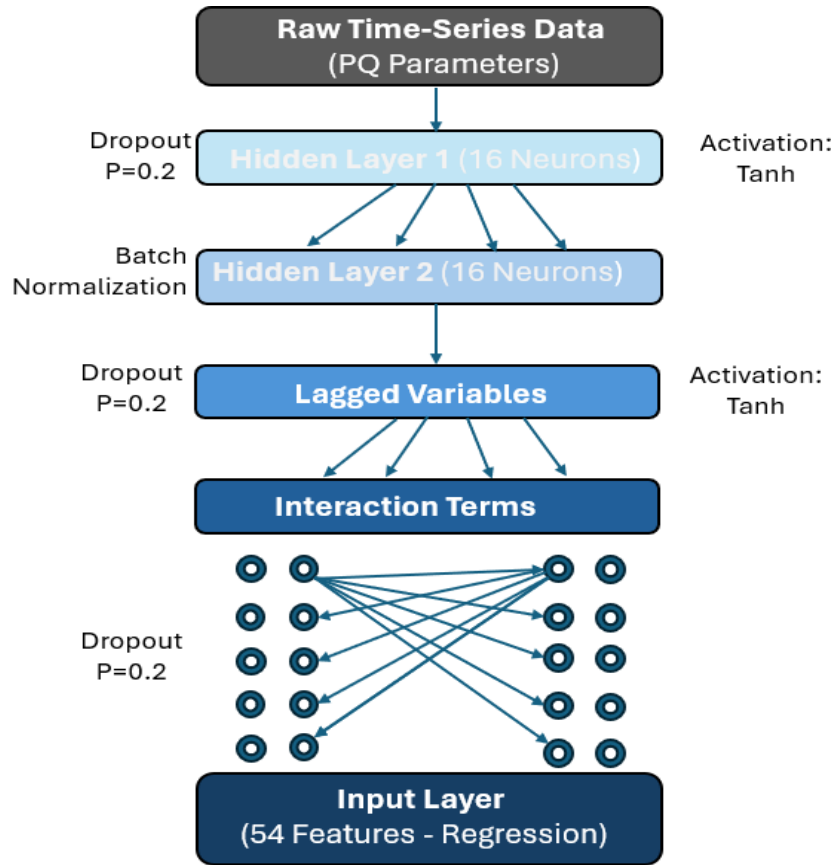


Figure 3.3: Conceptual Architecture of Artificial Neural Network (MLP)

Figure 3.3 shows the multilayer perceptron (MLP) architecture used for the ANN models, with an input layer receiving engineered features, two hidden layers with ReLU activation, and a linear output neuron for regression. The figure indicates where dropout, L2 regularisation, and batch normalisation are applied. This visualisation helps the reader understand how the ANN learns non-linear mappings from feature space to PQ targets and why regularisation is necessary to avoid overfitting.

3.3.3 SARIMAX (seasonal ARIMA with exogenous)

Seasonal provides a linear, interpretable baselines follow:

$$(1 - \sum_{i=1}^p \varphi_i B^i)(1 - \sum_{i=1}^P \Phi_i B^{is})(1 - B)^d(1 - B^s)^D y_t = c + (1 + \sum_{i=1}^q \theta_i B^i)(1 + \sum_{i=1}^Q \Theta_i B^{is}) \epsilon_t + \sum_{k=1}^m \beta_k X_{t,k} \quad (5)$$

Where y_t is the target series (e.g., Frequency, V_THD, Voltage Unbalance, Flicker), B is the backshift operator, p, d, q are non-seasonal AR, differencing, and MA orders; P, D, Q are seasonal orders; s is the seasonal period (144 for daily seasonality at 10-minute resolution); ϕ_i, Φ_i are AR coefficients; θ_i, Θ_i MA coefficients; $X_{t,k}$ exogenous regressors (e.g., wind speed, temperature); β_k their coefficients; and c a constant [57]. Although linear, SARIMAX provides an interpretable baseline [58].

Table 3.4: Implementation in JMP Pro.

Parameter	Description
Platform	Time Series ARIMA.
Seasonality	Seasonal period set to $ss=144$ (10-minute data $\rightarrow$ daily cycle).
Exogenous variables	where supported, wind speed, temperature, and reactive power are included as regressors (SARIMAX behaviour)
Target types	Mainly Frequency and V_RMS (p.u.), with exploratory models for Unbalance and PST/PLT to serve as linear baselines
Model identification	JMP's automatic identification used as a starting point; orders (p, d, q, P, D, Q) adjusted manually where residual ACF/PACF showed under-modelled structure, within the constraints of the Student Edition.
Validation	Chronological train/validation/test split; within training plus validation, time-ordered subsets are used for parameter tuning (no random K-fold mixing across time).

Design rationale and strengths leveraged.

SARIMAX provided transparent linear baselines that are interpretable and suitable for reasonably stationary, seasonal metrics such as Frequency and V_RMS. In this study: It provided a benchmark for quantifying the incremental value of nonlinear ML models. It provided confidence intervals and residual diagnostics for sanity-checking the behaviour of more complex models.

Weaknesses acknowledged

The automatic default setting rendered the model unable to represent strong nonlinearities and

regime shifts, especially in Voltage Unbalance and PST, where PQ dynamics are driven by converter behaviour and wind variability. The stationarity requirements of the model were structured in a way that are only approximately met after differencing and detrending. Tooling constraints: The JMP Student Edition limits manual control of higher-order seasonal structures and some exogenous configurations; this is acknowledged as a methodological limitation rather than a mis-specification.

3.3.4 Boosted Trees (BT; Gradient Boosting/XGBoost-Style)

Boost Tress (XGBoosting) are advanced ensemble technique within the gradient Boosting algorithms. Ensemble loss combines data-fit and complexity penalty:

$$\mathcal{L}(\varphi) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (6)$$

Where $l(y_i, \hat{y}_i)$ is the loss function (squared error for regression) between true label y_i and prediction $\hat{y}_i$ predictions, and $\Omega(f_k)$ a regularization term for the k -th tree f_k :

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \|\omega\|^2 \quad (7)$$

With T the number of leaves in the tree. ω the vector of scores on leaves, γ a penalty for the number of leaves (controls model complexity), and λ L2 regularization of leaf weights.

Table 3.5: Implementation in JMP Pro.

Parameter	Description
Platform	Predictive Modeling Bootstrap Forest and Gradient Boosting modules
Kernel	Radial Basis Function (RBF)
Hyperparameters	Gradient boosting core: Maximum tree depth ≈ 6 (controlling interaction order). Learning rate η approx. 0.1 (contribution of each tree). γ (JMP equivalent regularisation) ≈ 0.3 . L2 regularisation on leaf weights activated where available.
Target types	Separate BT models trained for each PQ parameter (Voltage Unbalance, Frequency, V_THD, Pst/Plt, V_RMS p.u.).
Validation	Performance monitored on the held-out validation set; final assessment on the test set, with additional ± 5 % threshold perturbation for compliance metrics.

Design rationale and strengths leveraged.

BTs performed well on tabular data with structured features and were relatively robust to moderate outliers and monotonic transformations. For PQ forecasting: They excelled at modelling interaction effects (e.g., wind speed $\times$ Unbalance, time-of-day $\times$ load), which are central to converter-driven PQ dynamics. They provided transparent feature ranking, revealing, for example, the dominance of Lag 1 Unbalance and secondary contributions from wind interaction terms and temporal indicators. They were computationally efficient, offering fast training and inference suitable for potential near-real-time deployment in SCADA environments.

Weaknesses addressed.

Lack of native sequence memory: This was addressed via the rich suite of lagged and rolling features in Section 3.2.3. Potential overfitting with deep trees: This was mitigated by constraining depth, applying learning-rate shrinkage, and using regularisation (γ , L2). Limited extrapolation beyond training range: acknowledged; models are interpreted as local approximators within the observed regime, with the need for retraining under significant regime shifts.

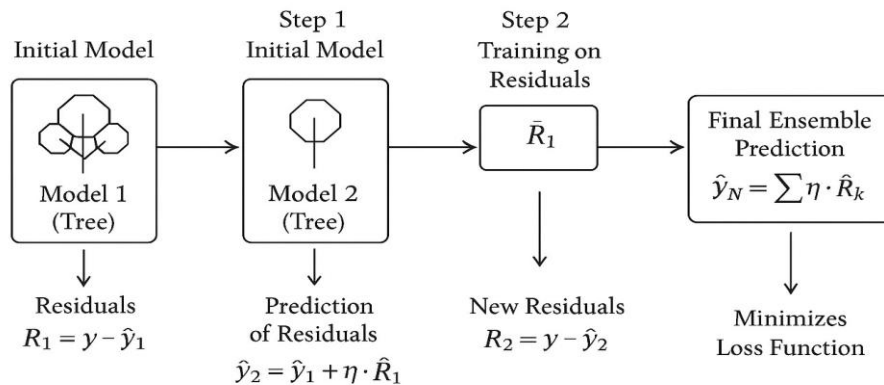


Figure 3.4: Illustrative Mechanism of Boosted Tree Ensemble

Figure 3.4 illustrates the principle of gradient boosting: a sequence of shallow decision trees is trained iteratively, each new tree learning to correct the residual errors of the previous ensemble. The figure highlights how this process builds a strong predictor from many weak learners and how regularisation controls tree depth and leaf weights. This mechanism explains why Boosted Trees are effective for modelling complex, interaction-driven PQ phenomena such as Voltage Unbalance and flicker.

3.3.5 Summary of Model Setup and Method Rationale

Table 3.2 summarises how the models were configured and deployed in this study using JMP 18 Pro (Student Edition) statistical analysis software.

Table 3.6: Summary of model setup and rationale in JMP Pro

Model	Main PQ Use Cases	JMP Platform	Key Parameters	Rationale Summary
SVM	V_THD, Unbalance, V_RMS	Predictive Modeling SVM	RBF kernel; (C = 1.0); (gamma = 0.01)	Captured smooth nonlinear margins in high dimensional compliance data.
ANN	PST, PST/PLT support	Predictive Modeling Neural	2 layers (32;16), ReLU, Dropout 0.2, L2	Strong nonlinear pattern learner for bursty PQ targets.
BT	Unbalance, PL T, V_RMS	Gradient Boosting	Depth ≈ 6 ; (eta = 0.1); $\gamma \approx 0.3$	Robust fits, interpretable; low MVR with low FAR.
SARIMAX	Frequency, V_RMS	Time Series ARIMA	Series(s=144); met & power vars	Linear, interpretable baseline for stable, seasonal metrics.

These deployments enable rigorous comparison across forecasting models with consistent input features and temporally valid validation splits. Their respective strengths are evaluated empirically in Chapter 4, and a parameter–model suitability map is synthesised in Chapter 6.

3.4 METRICS: ACCURACY AND NRS 048-COMPLIANCE

3.4.1 Regression Metrics

Regression performance was evaluated using standard metrics, including RMSE, MAE, MAPE, and R^2 , which quantify continuous accuracy and explained variance. In addition, autocorrelation and partial autocorrelation functions (ACF and PACF) were applied to the

residuals to diagnose potential under-modelled temporal structure [58,60].

Mean Squared Error (MSE):

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (8)$$

Where y_i is the actual value at time i . $\hat{y}_i$ is the predicted value at time i . n is the Number of observations.

Mean Absolute Error (MAE):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (9)$$

Where y_i is the actual value at time i , $\hat{y}_i$: Predicted value at time i . n : Number of observations.

Mean Absolute Percentage Error (MAPE):

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \quad (10)$$

Where: y_i : Actual value at time i . $\hat{y}_i$: Predicted value at time i . n : Number of observations.

Coefficient of Determination (R^2):

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (11)$$

Where: y_i : Actual value at time i . $\hat{y}_i$: Predicted value at time i . $\bar{y}$: Mean of the actual values.
 n : Number of observations.

Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) Analysis:

Its formal definitions follow standard time-series texts [56,59].

$$\rho_k = \frac{Cov(y_t, y_{t-k})}{\sqrt{Var(y_t) * Var(y_{t-k})}} \quad (12)$$

Where: ρ_k : Autocorrelation at lag k , y_t : value of the time series at time t , y_{t-k} : Value of the time series at lag k .

Equation for PACF:

$$\phi_{kk} = \frac{Cov(y_t, y_{t-k} | y_{t-1}, y_{t-2}, \dots, y_{t-k+1})}{\sqrt{Var(y_t | y_{t-1}, y_{t-2}, \dots, y_{t-k+1}) * Var(y_{t-k} | y_{t-1}, y_{t-2}, \dots, y_{t-k+1})}} \quad (13)$$

Where: ϕ_{kk} : Partial autocorrelation at lag k .

3.4.2 NRS 048 related Compliance Metrics

With prediction thresholds defined by NRS 048-2, we derive confusion-matrix terms and compute Accuracy, False Alarm Rate (FAR), and Missed Violation Rate (MVR). Minimising MVR is operationally critical, as it corresponds to missed events, while FAR governs the occurrence of spurious alarms [42,60].

Accuracy:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (14)$$

False Alarm Rate (FAR):

$$FAR = \frac{FP}{TN+FP} \quad (15)$$

Missed Violation Rate (MVR):

$$MVR = \frac{FN}{TP+FN} \quad (16)$$

Figure 3.5 below presents a generic confusion matrix for binary classification, with cells for true positives, true negatives, false positives, and false negatives. It is used to define the Accuracy, FAR, and MVR metrics introduced in this Section 3.4.2. By linking the confusion matrix to operator-relevant concepts such as missed violations and false alarms, the figure bridges the gap between statistical measures and practical compliance risk.

		Predicted Class		
		Predicted: Compliant(0)	Non-Compliant (1)	
Actual Class	Actual: Compliant (0)	True Negative (TN)	False Positive (FP) <i>False Alarm!</i>	False Alarm Rate (FAR)
	Actual: Non-Compliant (1)	False Negative (FN) <i>Missed Violation!</i>	True Positive (TP)	Missed Violation Rate (MVR)

Matrix for Binary Classification

Figure 3.5: Confusion

In addition to point estimates of accuracy and compliance metrics, statistical tests were

considered to assess whether observed performance differences between models are likely to be genuine rather than due to sampling variability. For each PQ target and pair of models (A, B), the Diebold–Mariano (DM) test can, in principle, be used on the out-of-sample test errors to test the null hypothesis of equal predictive accuracy under squared error loss. Similarly, for binary compliance outcomes, two-proportion tests (e.g., z-tests) can be used to compare MVR and FAR across models. In this study, the implementation of full DM and two-proportion testing is constrained by the capabilities of JMP Pro (Student Edition). Instead, robustness is assessed through consistency of model ranking across multiple metrics (RMSE, MAE, R^2 , Accuracy, FAR, MVR), stability of rankings across sub-periods of the three-month test window, and sensitivity of compliance conclusions to $\pm 5\%$ perturbations of the NRS 048-2 thresholds.

These triangulated checks increased confidence that the observed performance patterns, particularly the advantage of Boosted Trees and SVM over SARIMAX for Voltage Unbalance, flicker, and V_{THD} , are not artefacts of a single metric choice or a particular threshold setting. Chapter 4 reports these patterns in detail and discusses their practical implications. Future work (Chapter 6) outlines how DM-type tests and bootstrap-based confidence intervals could be implemented in more flexible environments (e.g., R, Python).

3.5 VALIDATION DESIGN AND LEAKAGE CONTROLS

We use a chronological 70/20/10 split over 2021-03-01 to 2021-05-31 (train/validation/test), preserving time order; no K-fold mixing across time was used. Scaling parameters are learned only on the training set and applied to the validation/test sets. The lag/rolling windows are constructed strictly from past observations to avoid look-ahead [61].

Figure 3.6 visualises the time-series splitting into training, validation, and test segments. It emphasises that, unlike random cross-validation, the splits respect the temporal order of the data and avoid leaking future information into model training. This design is crucial for the realistic evaluation of forecasting models, since it mimics the real-world scenario where models trained on past data are applied to future, unseen conditions.

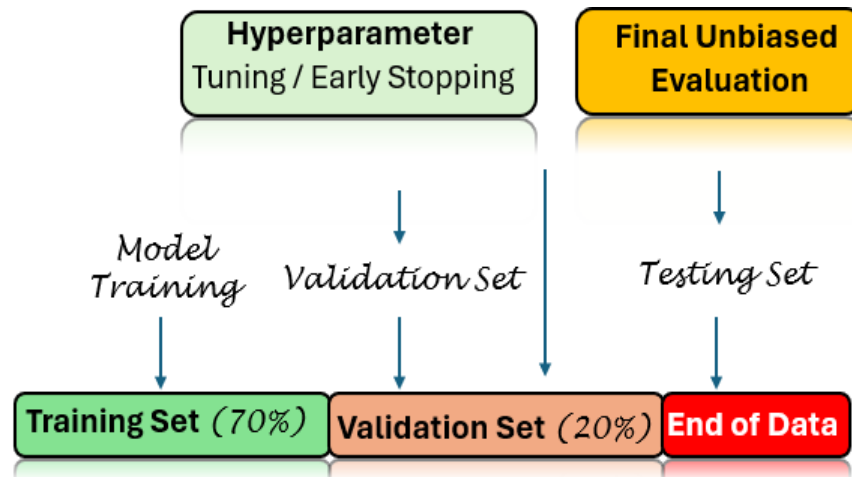


Figure 3.6: Temporal Data Splitting Strategy

This adheres to best practice for time-series validation and prevents optimistic bias [62].

3.6 REPRODUCIBILITY AND TOOLING

All steps were executed in JMP Pro 18 Student Edition on a Windows 10 workstation (Intel i7-10700; 32 GB RAM; SSD). Random seed 12345 was fixed where applicable to ensure repeatability of ANN/BT initialisations and sampling [41]. Due to data confidentiality, raw files are not public; however, a detailed data dictionary and a step-wise runbook (preprocessing, feature engineering recipes, model configs, and thresholds) enable independent replication on

analogous datasets [41,63]. The choice of JMP Pro aims to make the workflow accessible to practitioners with limited programming experience while still allowing rigorous model comparison.

Practical deployment of PQ forecasting models requires that training and inference be computationally feasible on typical utility hardware. All experiments in this dissertation were conducted on a standard workstation. Training a single model instance for a given PQ parameter and phase typically takes a few minutes for SARIMAX and SVM, and approximately 10–15 minutes for the Boosted Tree and ANN configurations, including modest hyperparameter tuning.

Once trained, inference is computationally lightweight: generating forecasts for all PQ parameters and phases at a single 10-minute SCADA update requires only milliseconds per model, which is negligible relative to the 10-minute reporting interval. This implies that BT and SVM models of the complexity used here are compatible with near-real-time deployment in a SCADA or EMS environment, provided that models are retrained offline on a daily or weekly basis and periodically pushed to the operational environment.

The modest training and inference costs, combined with the use of standard desktop hardware, support the view that the proposed compliance-aware forecasting framework is computationally viable for utility adoption. Measurements were anonymised; research followed Non-Disclosure Agreement (NDA) constraints with the operator. The work aims to improve renewable integration and reduce operational risk with a modest compute footprint, avoiding large deep-learning stacks in favour of efficient models aligned with utility adoption pathways [63].

3.7 CHAPTER SUMMARY

This chapter has detailed the end-to-end analytical workflow for PQ forecasting for NRS 048-2 compliance at a South African wind farm PCC. It described how 10-minute SCADA and meteorological measurements were cleaned, standardised, and transformed into a rich feature set capturing persistence, exogenous drivers, and temporal patterns. It documented configuration of four model algorithms; SARIMAX, SVM, ANN, and Boosted Trees, and it defined both regression metrics (RMSE, MAE, MAPE, R^2) and compliance metrics (Accuracy, FAR, MVR) used for evaluation. A chronological 70/20/10 split and strict anti-leakage controls ensured that the models were evaluated in a realistic out-of-sample setting.

The chapter also clarified how NRS 048-2 limits are used to convert continuous predictions into binary compliance labels and how a ± 5 % threshold-sensitivity analysis probes the robustness of compliance conclusions. Considerations of reproducibility, computational feasibility, and ethical constraints were explicitly addressed.

The next chapter builds directly on this methodology presenting the empirical results obtained when the four-machine learning algorithmic models are applied to the three-month PQ dataset.

It compares their performance across PQ parameters in terms of both regression accuracy and compliance metrics, visualises model behaviour through time-series plots and heatmaps, and interprets the patterns in light of PQ physics and operational needs.

CHAPTER 4: RESULTS AND ANALYSIS

4.1 DESCRIPTIVE ANALYSIS OF POWER QUALITY DATA

The PQ time series exhibit non-stationarity, volatility bursts, and diurnal structure, consistent with converter-interfaced wind integration and daily load cycles at the point of common coupling (PCC). Voltage Unbalance and short-term flicker (P_{st}) show the most pronounced variability, with occasional excursions towards the NRS 048-2 limits, while Frequency, V_{RMS} (p.u.), and V_{THD} remain tightly clustered around their nominal bands throughout the three-month window. Ground truth labels for violations were derived using the NRS 048-2 limits for Voltage Unbalance, Flicker (P_{st}/P_{lt}), V_{THD} , Frequency, and V_{RMS} (p.u.), as described in Chapter 3. This yields a dataset in which:

Frequency, V_{RMS} , and V_{THD} show no violations during the test period; P_{lt} shows very few isolated events near the long-term flicker limit; Voltage Unbalance and P_{st} exhibit the majority of violations and near-limit behaviour.

These characteristics are important because they shape the class imbalance of the compliance labels: models are evaluated not only on overall regression accuracy but also on their ability to detect relatively rare Voltage Unbalance and P_{st} violations without generating excessive false alarms.

4.2 COMPARATIVE PERFORMANCE OF ML MODELS

We trained SVM, ANN, SARIMAX, BT and evaluated on held-out test data. Aggregate results show BT and SVM generally dominate continuous accuracy, ANN excels on PST, and SARIMAX underperforms on non-linear dynamics while remaining a useful linear baseline.

4.2.1 Overall Model Performance

Table 4.1 outlines the overall: RMSE/MAE/ R^2 /N across Frequency, Voltage Unbalance, V_{THD} Overall, PST Overall, PLT Overall, V_{RMS} (p.u.). BT achieves the lowest average RMSE with high R^2 across many targets; SVM is especially strong for V_{THD} ; ANN competes on PST; SARIMAX trails on complex signals patterns consistent with model capacity and data non-

linearity. Column N denotes the number of out-of-sample test observations used for that model-target combination after target-specific lag construction, missing-value handling, and model requirements. SARIMAX has a different N for some targets because seasonal and lagged initial observations are consumed during model fitting and forecasting.

Table 4.1: Aggregated Model and Target Regression Performance (Test Set)

Target	Model	RMSE	MAE	R ²	N
Frequency	BT	0.0105	0.0079	0.9879	484
	NN	0.0409	0.0314	0.8159	484
	SARIMA X	0.0982	0.077	- 0.0228	1452
	SVM	0.0277	0.0154	0.9157	484
Voltage Unbalance	BT	0.0711	0.053	0.9524	484
	NN	0.1025	0.0804	0.9011	484
	SARIMA X	0.2185	0.113	0.567	1452
	SVM	0.0732	0.0388	0.9495	484
V_THD Overall	BT	0.1088	0.0835	0.9262	484
	NN	0.1567	0.1251	0.8469	484
	SARIMA X	0.2665	0.1466	0.5739	1452
	SVM	0.069	0.0421	0.9703	484
Flicker PST Overall	BT	0.1881	0.091	0.2722	484
	NN	0.1019	0.034	0.7865	484

Target	Model	RMSE	MAE	R ²	N
Flicker PLT Overall	SARIMA X	0.2442	0.1432	- 0.1813	1452
	SVM	0.1155	0.051	0.7255	484
	BT	0.0555	0.0245	0.9689	484
	NN	0.0966	0.0421	0.9058	484
	SARIMA X	0.1354	0.053	0.8219	1452
V_RMS Overall (p.u.)	SVM	0.1357	0.05	0.8142	484
	BT	0.0001	0.0001	0.9999	484
	NN	0.0025	0.0018	0.9565	484
	SARIMA X	0.0047	0.0035	0.8492	1452
	SVM	0.0015	0.0007	0.9845	484

From Table 4.1 above it can be seen that:

For Frequency, BT achieves the lowest RMSE (0.0105) and highest R² (0.9879), with SVM also performing strongly. SARIMAX underperforms markedly (R² $\approx$ -0.02), indicating that the automatic model selection in JMP did not capture the subtle dynamics despite the relative stability of Frequency. For Voltage Unbalance, BT and SVM perform similarly well (R² $\approx$ 0.95), with BT slightly ahead on RMSE. ANN lags behind but still provides reasonable fits, while SARIMAX struggles (R² $\approx$ 0.57), reflecting its limited ability to capture non-linear converter behaviour and interaction effects.

For V_THD Overall, SVM is clearly the best performer ($RMSE \approx 0.069$, $R^2 \approx 0.97$), with BT also strong ($R^2 \approx 0.93$). ANN and SARIMAX exhibit substantially higher errors and lower R^2 values, suggesting that harmonic distortion patterns benefit from the flexible decision boundaries of SVMs. For Flicker P_st Overall, ANN dominates ($RMSE \approx 0.102$, $R^2 \approx 0.79$), followed by SVM and BT. SARIMAX performs poorly (negative R^2), indicating that short-term flicker dynamics are highly non-linear and not well captured by the linear seasonal structure.

For Flicker P_It Overall, BT provides the best performance ($RMSE \approx 0.0555$, $R^2 \approx 0.97$), with ANN and SVM close behind. SARIMAX again underperforms, though its R^2 remains above 0.82. For V_RMS Overall (p.u.), all non-linear models perform extremely well, with BT approaching a near-perfect fit ($RMSE \approx 0.0001$, $R^2 \approx 0.9999$), confirming that the voltage magnitude is highly stable and easy to track.

In summary, Boosted Trees and SVM generally achieve lower errors and higher explanatory power than SARIMAX, confirming the advantage of non-linear methods for PQ forecasting. ANN is competitive and even superior for P_st, highlighting that different PQ targets favour different model classifications.

4.2.2 Discussion of trends

The trends in Table 4.1 can be interpreted as follows:

BT and SVM dominate by capturing heterogeneous interactions and threshold effects between PQ indices, meteorological drivers, and temporal indicators. SVMs, via the RBF kernel, provide smooth non-linear decision surfaces well suited to capturing complex yet relatively smooth relationships in V_THD and Voltage Unbalance.

Short-term flicker (P_st) exhibits bursty, non-stationary dynamics. The ANN's universal approximation capability allows it to capture these patterns when appropriately regularised,

explaining its superior performance relative to SVM and BT on this target. Although SARIMAX can model seasonality and autocorrelation, its linear structure cannot fully absorb converter-driven non-linearities and interaction effects. In addition, automatic model identification in JMP Pro (Student Edition) may have played a significant role in limiting the search space, leading to under-specified models and lower R^2 values, especially for Voltage Unbalance, V_{THD} , and P_{st} .

Frequency is intrinsically stable during the study period, so all models operate over a narrow numerical range. The SARIMAX baseline nevertheless gives weaker R^2 than BT and SVM because the residual deviations are small and structured; even minor errors materially affect R^2 when the variance of the target is low. The practical implication is that BT and SVM track small deviations more precisely in this dataset, while SARIMAX remains useful as an interpretable benchmark for stable, seasonal series.

4.2.3 Visual summaries

To make the comparative performance more intuitive, Figures 4.1 to 4.3 provide visual summaries across models.

Figure 4.1 plots the average test-set RMSE across PQ parameters for each model family. Lower bars indicate better performance. The figure clearly shows that Boosted Trees and SVMs dominate in terms of error magnitude, while SARIMAX trails, especially for non-linear parameters such as Voltage Unbalance and flicker. This visual summary complements Table 4.1 and reinforces the narrative that non-linear models better capture the underlying PQ dynamics.

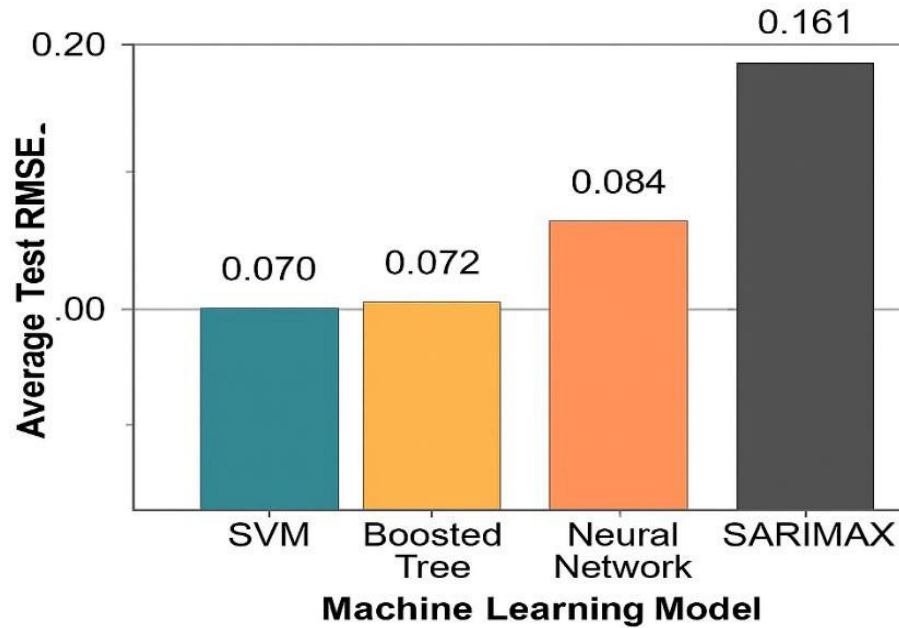


Figure 4.1: Average Test RMSE by Model

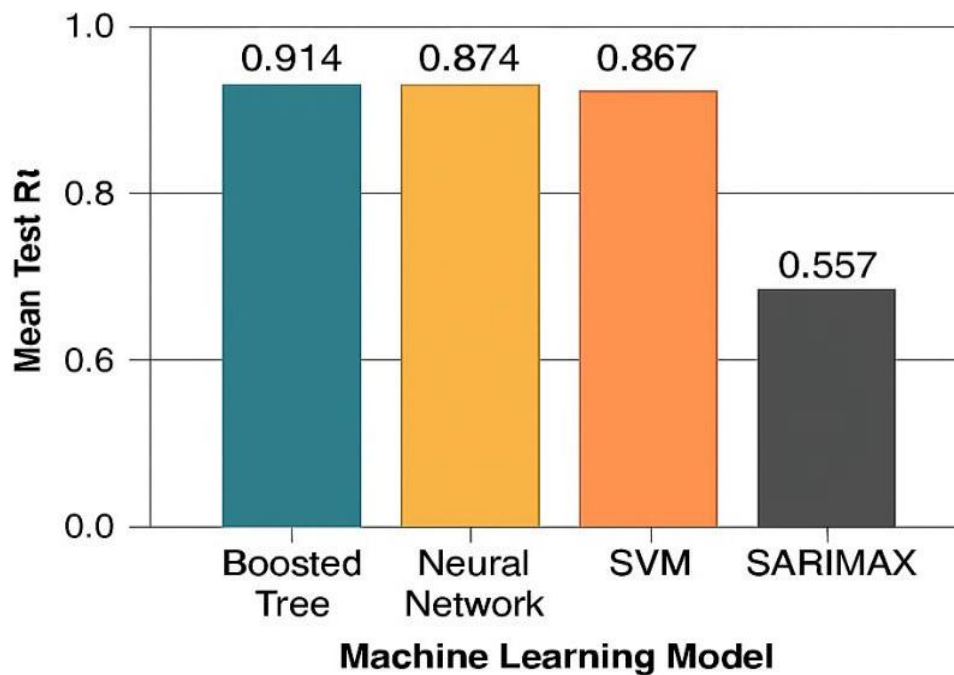


Figure 4.2: Average Test R² by Model

Figure 4.2 shows the average R^2 across PQ parameters for each model. Higher values indicate that the model explains a greater proportion of the target's variance. Consistent with Figure 4.1, Boosted Trees and SVM achieve the highest average R^2 , confirming their ability to track both

the level and variability of the PQ indices, while SARIMAX exhibits lower explanatory power, especially for parameters with complex temporal structure.

Figure 4.3 presents a spider (radar) plot that simultaneously displays model performance on multiple PQ parameters. Each axis corresponds to a target (Voltage Unbalance, V_THD, P_st, P_lt, Frequency, V_RMS), and the radial position represents a normalised performance score. The plot shows that there is no single universally best model: BT and SVM dominate for most targets, while ANN is particularly strong for P_st. This supports a parameter-specific model selection strategy rather than a one-size-fits-all choice.

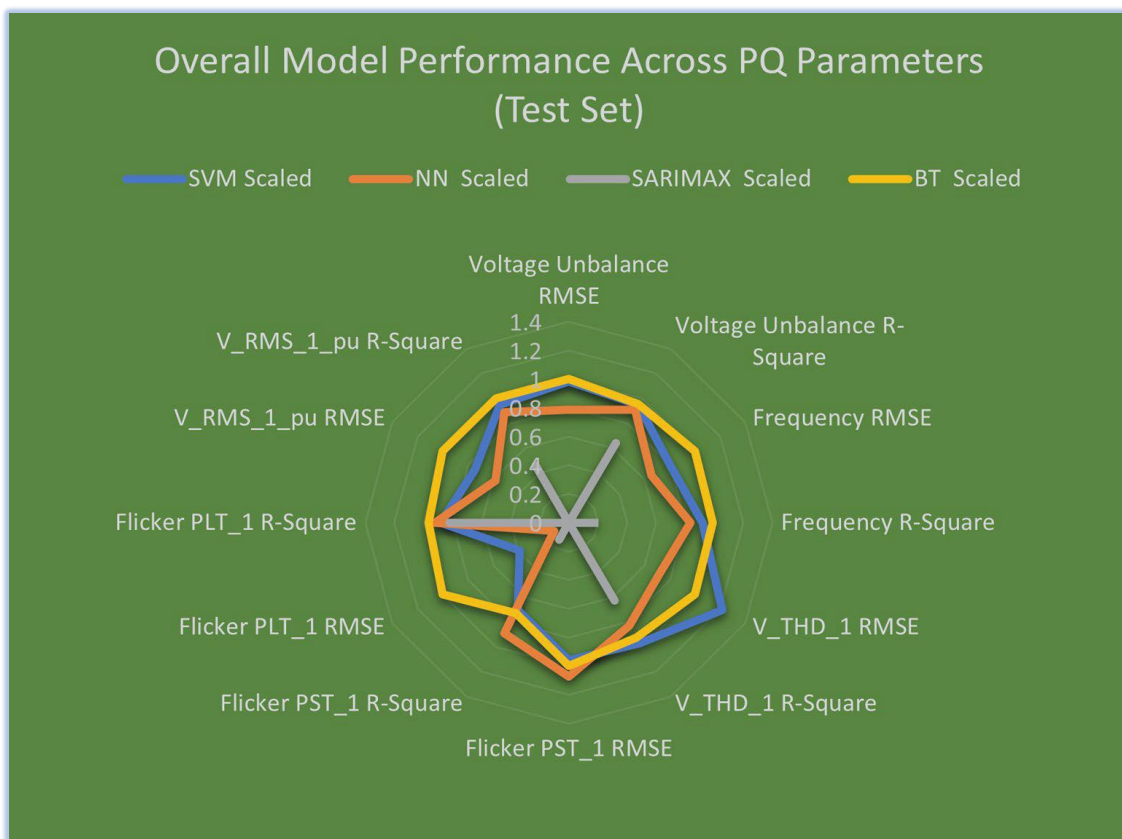


Figure 4.3: Spider Plot of performance across PQ parameters

These visuals reinforce BT/SVM strength and target-specific advantages.

4.2.4 Illustrative Cases (Selected Targets)

To move beyond aggregate metrics, we inspect representative time-series segments for key targets. Starting with boosted trees' performance on voltage unbalance, it can be seen that the

BT model achieves $RMSE \approx 0.071$ and $R^2 \approx 0.95$. Figure 4.4 shows a six-day test-set segment comparing actual and predicted unbalance values, along with residuals. The model tracks both the level and fluctuations of unbalance closely, with small, unbiased residuals. This supports its use for proactive mitigation: approaching NRS 048-2 limits can be detected with several 10-minute steps of advance warning.

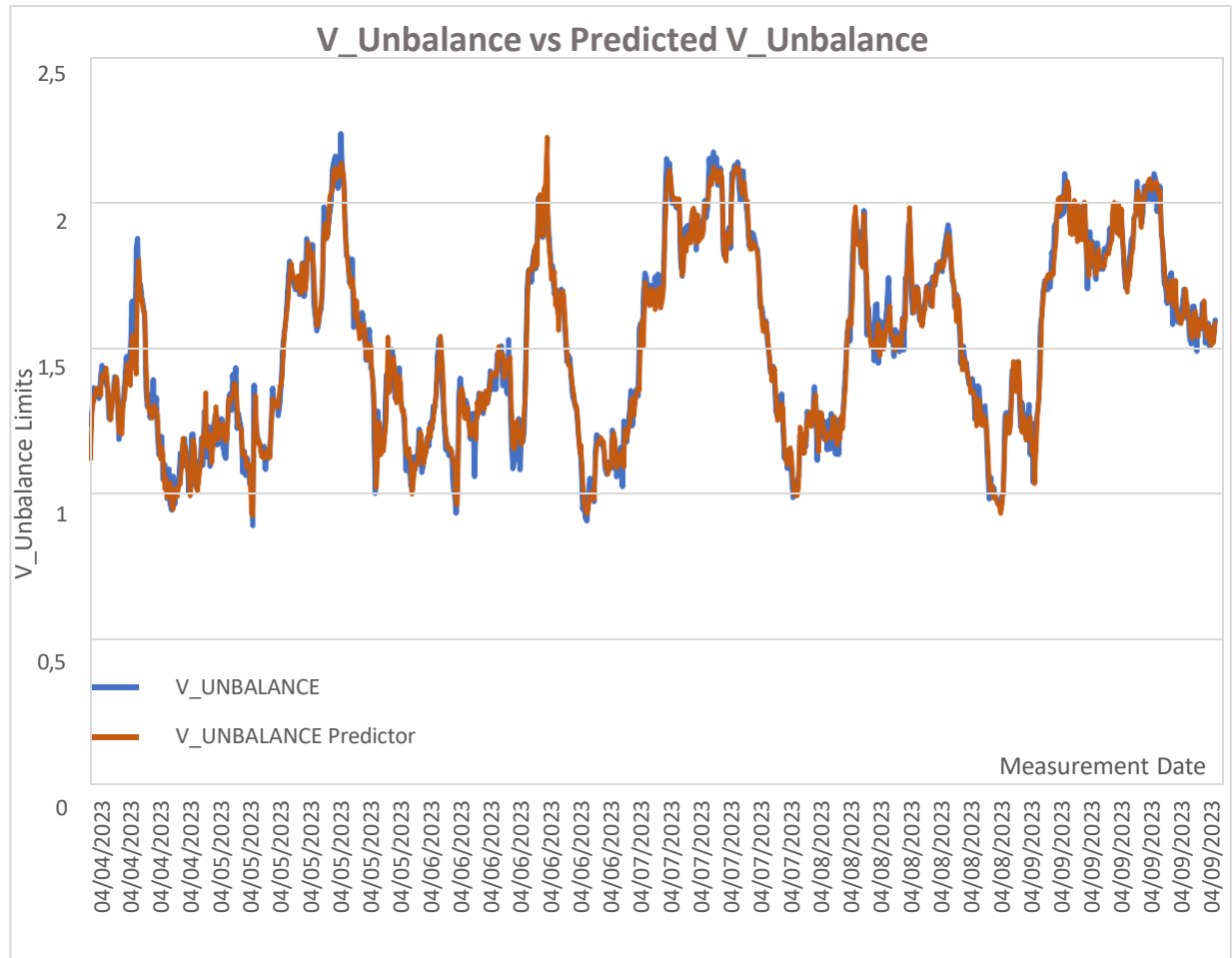


Figure 4.4: Time-Series Tracking of BT for Voltage Unbalance (Test Set: 6-day View)

For Frequency BT achieves $RMSE \approx 0.0105$ and $R^2 \approx 0.99$. The below figure 4.5 shows a representative single-day window, where the predicted series almost perfectly overlays the actual frequency. Residuals are tightly clustered around zero, confirming that the model does not introduce artefacts and is capable of reproducing the small deviations that do occur.

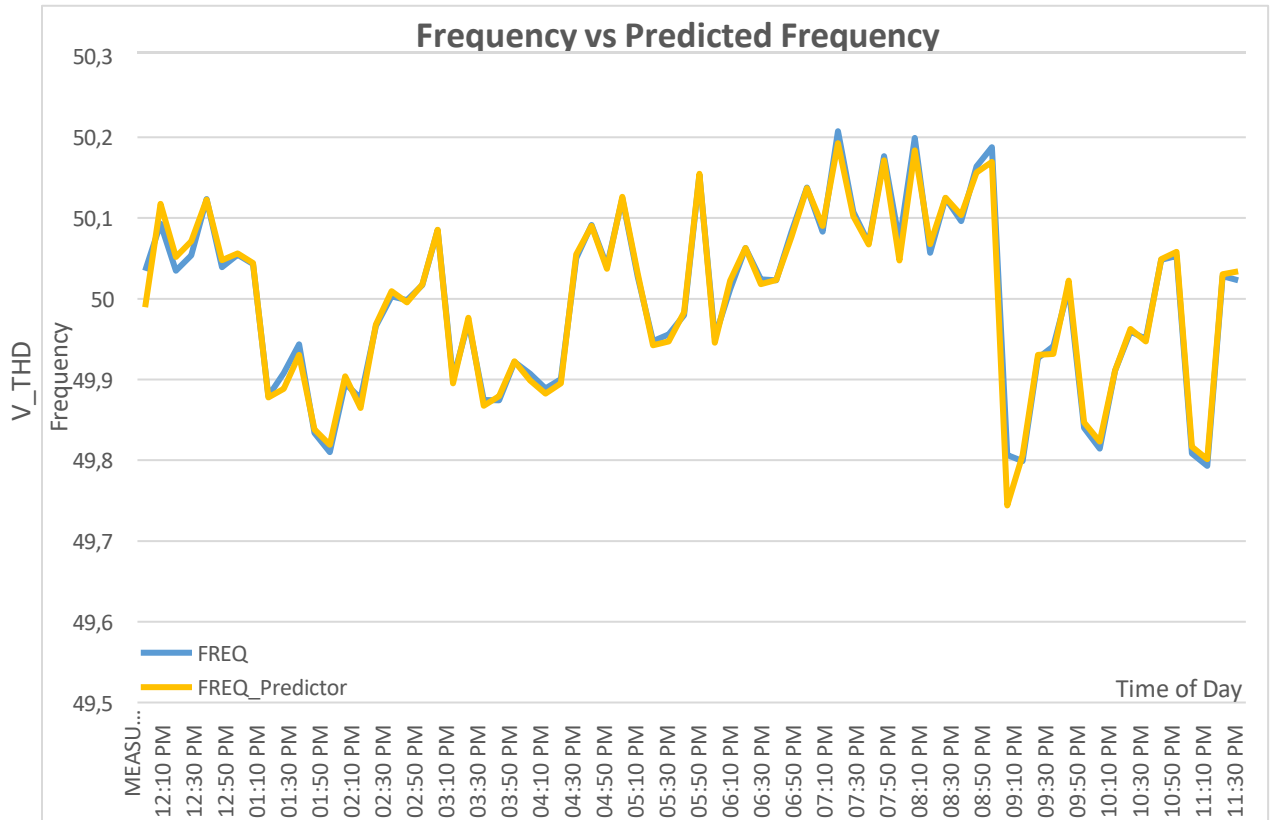


Figure 4.5: Time-Series Tracking of BT for Frequency (Test Set: Single-Day View)

For V_THD Overall, SVM delivers the lowest RMSE and highest R^2 across phases. Figure 4.6 below focuses on a three-hour test-set window for V_THD in phase 1, showing that SVM accurately tracks short-horizon peaks and troughs in harmonic distortion. This is critical for avoiding equipment stress and preventing THD-related NRS 048-2 violations.

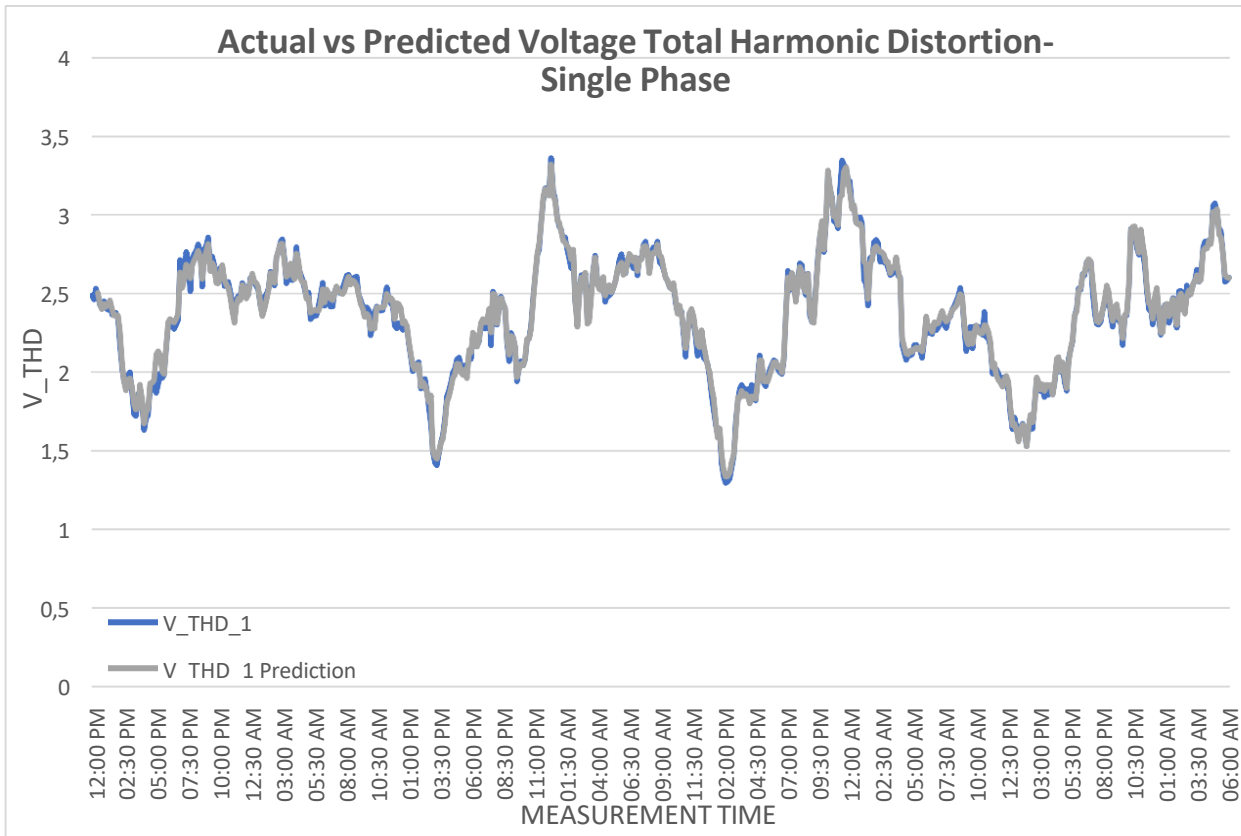


Figure 4.6: Time-Series Tracking of SVM for V_THD1 (Test Set: 3-Days View)

Figures 4.4–4.6 use different time windows deliberately. They are illustrative diagnostics rather than direct cross-parameter comparisons: Voltage Unbalance is shown over six days to demonstrate persistence and slow drift, Frequency is shown over one day because it is tightly regulated and deviations are small, and V_THD is shown over three days to make short-term harmonic peaks visible. Direct model comparison is therefore based on Table 4.1 and the compliance metrics, not on visual comparison across these differently scaled plots.

Additional per-target and per-phase tables for the remaining PQ parameters, together with residuals, are presented in Appendix B.

4.2.5 Residual Diagnostics for Selected Targets

To assess whether the regression models leave significant structure unexplained, we examine residuals for the best-performing models on the most challenging targets.

Figure 4.7 plots residuals against predicted values for the BT model on Voltage Unbalance. The residual cloud is approximately centred on zero and shows no strong funnel shape or systematic curvature, suggesting that heteroscedasticity and major non-linear patterns have been largely absorbed by the model.

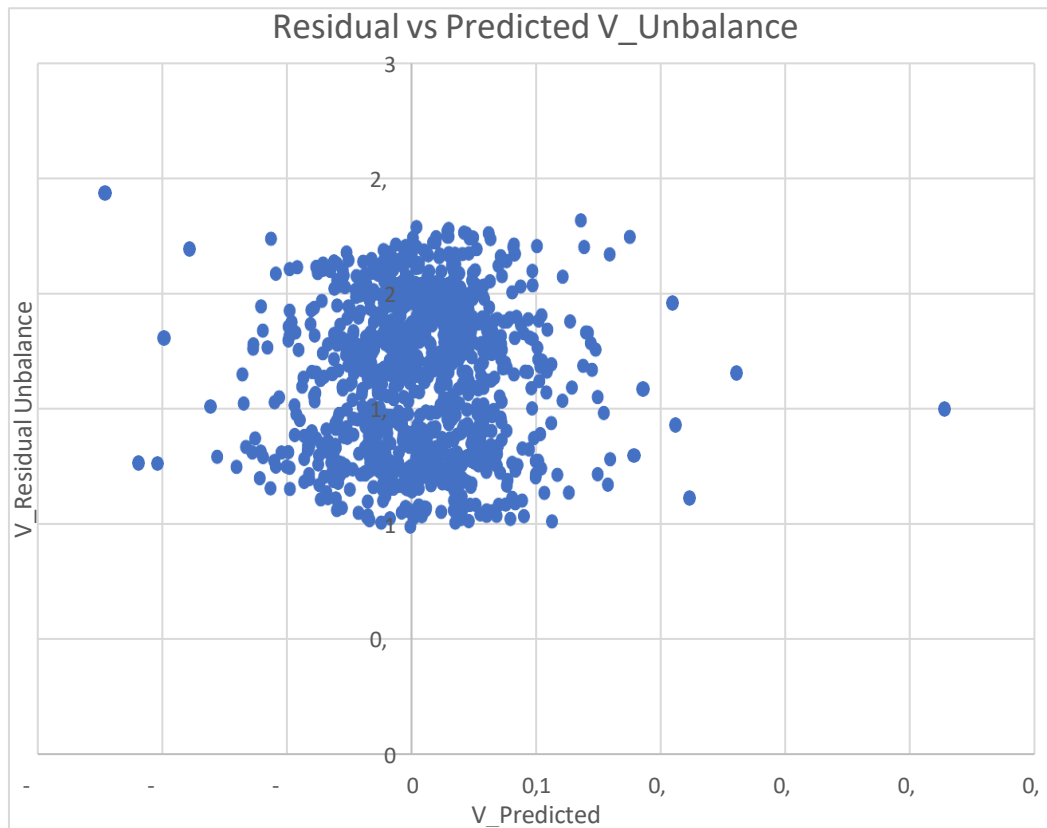


Figure 4.7: Residuals vs predicted values for BT Voltage Unbalance (test set)

Figure 4.8 presents a histogram of residuals for the same model and target. The distribution is approximately symmetric, with a single mode near zero, indicating that errors are predominantly small and unbiased. While not perfectly Gaussian, the shape is typical of well-behaved regression residuals and supports the use of squared-error loss in training.

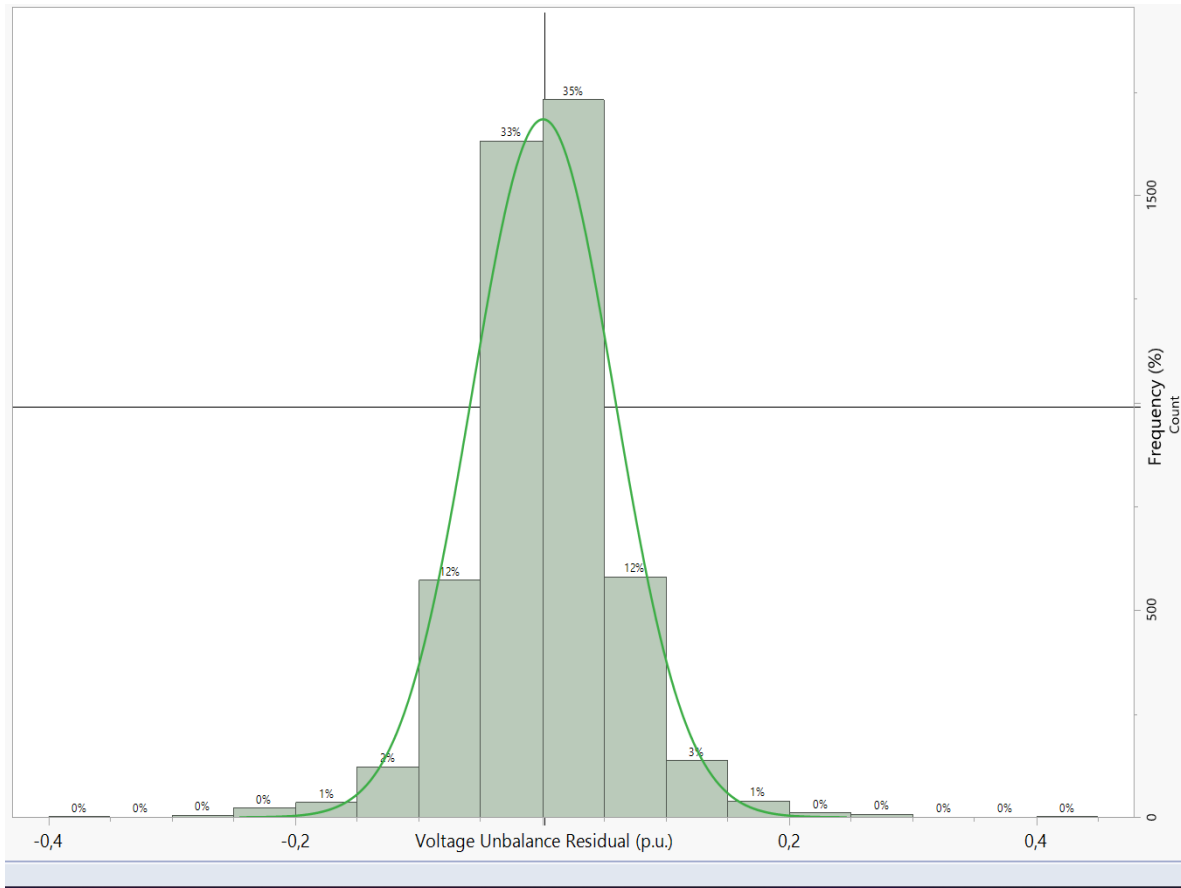


Figure 4.8: Histogram of residuals for BT Voltage Unbalance (test set)

Equivalent residual plots for P_{st} (ANN) and V_{THD} (SVM), shown in Appendix B, exhibit similar behaviour: residuals are centred around zero with no obvious temporal patterns. Combined with the ACF/PACF diagnostics described in Chapter 3, these findings indicate that the chosen models capture most of the temporal structure relevant for one-step-ahead forecasting.

4.3 NRS 048-2 COMPLIANCE ORIENTED PERFORMANCE

While regression metrics assess continuous accuracy, operators are primarily concerned with compliance: whether PQ indices violate NRS 048-2 limits and whether such events are detected in time. We therefore translate continuous model predictions into binary compliance labels and compute Accuracy, False Alarm Rate (FAR), and Missed Violation Rate (MVR) as defined in Chapter 3. Operational priority emphasises minimising MVR (missed violations), since these correspond to unflagged non-compliant intervals, while keeping FAR low enough to avoid alarm fatigue and maintain operator trust.

4.3.1 Visual summary

Figure 4.9 displays a heatmap of MVR for each model and PQ parameter at the system level. Light tones (near-zero MVR) mark Frequency, V_THD, and V_RMS as stable/compliant and easy to monitor, whereas darker tones appear for Voltage Unbalance and P_st, where violations are rarer but more challenging to forecast. The heatmap clearly shows that all models achieve $MVR \approx 0\%$ for Frequency, V_THD, and V_RMS (no observed violations). For Voltage Unbalance and P_st, BT consistently achieves the lowest MVR among the models that are not trivially over-alarmistic.

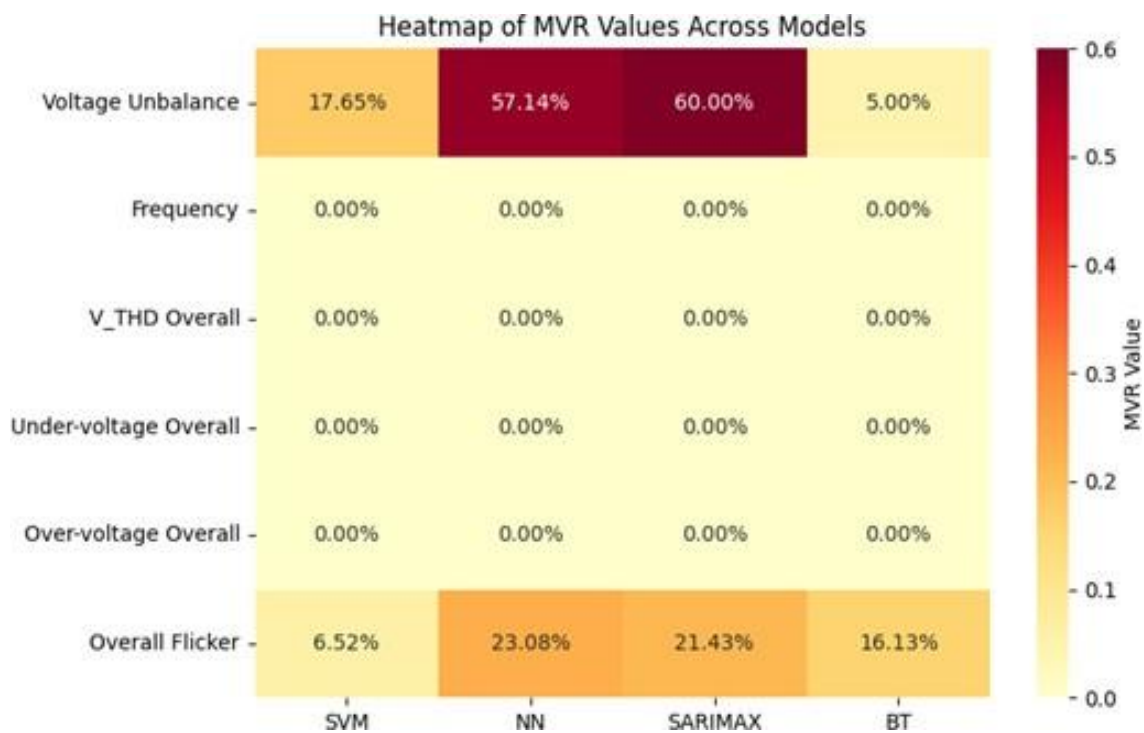


Figure 4.9: Heatmap of Missed Violation Rates (MVR) across models for Aggregated System-Level PQ Parameters

4.3.2 Detailed Compliance for Volatile Parameters

Table 4.2 reports the aggregated compliance metrics for each PQ parameter and model, including Accuracy, FAR, and MVR, with approximate 95% confidence intervals. For the volatile parameters (Voltage Unbalance, Pst, Plt), the table confirms that Boosted Trees maintain simultaneously low FAR and MVR, whereas SVM, ANN, and SARIMAX exhibit larger trade-offs: SVM tends to under-alarm with higher MVR but low FAR, while ANN and SARIMAX

produce both more missed violations and more false alarms. For the intrinsically stable parameters (Frequency, V_THD, V_RMS), all models achieve perfect apparent performance (Accuracy = 100 %, FAR = MVR = 0 %), reflecting the absence of observed violations in the test window rather than a discriminating difference in model behaviour.

Table 4.2: Aggregated Compliance Metrics (Accuracy, FAR, MVR, 95% CIs)

Target	Model	No	Acc_ (%)	Acc_ CI95	FAR (%)	FAR_ CI95	M V R (%)	MVR_ CI95
V_Unb	BT	484	99.59	[98.5, 99.9]	0.22	[0.0, 1.2]	5.0	[0.9, 23.6]
	NN	484	95.04	[92.7, 96.6]	3.45	[2.1, 5.5]	40.0	[21.9, 61.3]
	SARIMAX	466	94.42	[92.0, 96.2]	4.04	[2.6, 6.3]	40.0	[21.9, 61.3]
	SVM	484	98.14	[96.5, 99.0]	0.65	[0.2, 1.9]	30.0	[14.5, 51.9]
F_PST				[99.2, 99.9]				
Overall	BT	1452	99.66	[97.1, 99.9]	0.35	[0.1, 0.8]	0.0	[0.0, 20.4]
	NN	1452	97.93	[97.1, 98.5]	1.67	[1.1, 2.5]	40.0	[19.8, 64.3]
	SARIMAX	1398	98.0	[97.1, 98.6]	1.16	[0.7, 1.9]	92.31	[66.7, 98.6]
	SVM	1452	99.17	[98.6, 99.5]	0.07	[0.0, 0.4]	73.33	[48.0, 89.1]

To make these trade-offs more interpretable from an operational perspective, Figure 4.10 below, plots FAR against MVR for the volatile PQ parameters only (Voltage Unbalance, Pst, Plt). Colours indicate model type (BT, SVM, ANN, SARIMAX) and marker shape denotes the PQ target. The most desirable region is the bottom-left corner of the plot, corresponding to

simultaneously low false alarms and low missed violations. Points further up and to the right represent models that either miss a large fraction of true violations, generate excessive nuisance alarms, or both.

Given that Figure 4.10 already conveys the key compliance trade-offs visually, the main text retains a condensed version of Table 4.2 focusing on the volatile parameters, while the full table with all parameters and confidence intervals can be provided in Appendix B for readers who require exact numerical values.

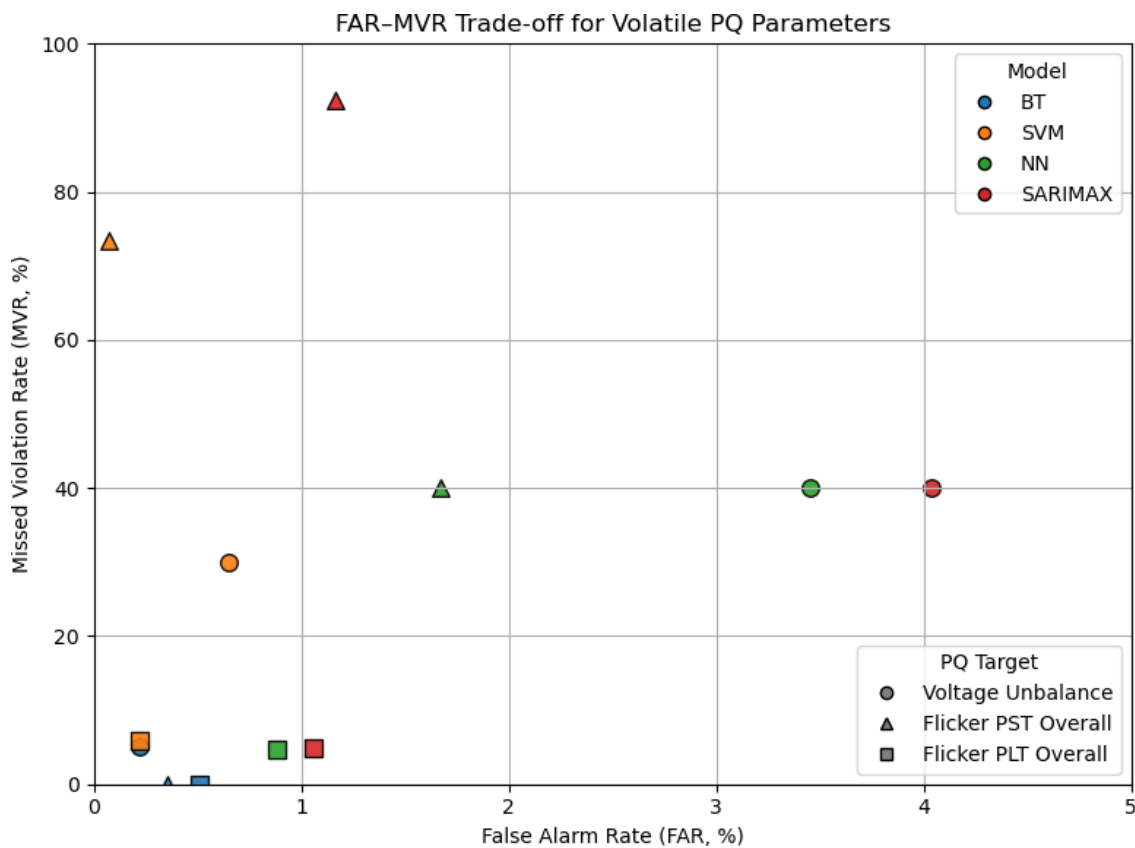


Figure 4.10: FAR-MVR Scatter Plot

Figure 4.10 confirms the advantage of Boosted Trees for compliance-oriented forecasting. For Voltage Unbalance, BT sits closest to the bottom-left corner ($\text{FAR} \approx 0.22\%$, $\text{MVR} \approx 5\%$), indicating that almost all violations are detected while very few false alarms are raised. SVM achieves a similar FAR but with a substantially higher MVR ($\approx 30\%$), reflecting a more conservative labelling behaviour that leaves more violations undetected. ANN and SARIMAX

are positioned further to the upper-right for unbalance, with both higher MVR ($\approx 40\%$) and higher FAR ($\approx 3\text{--}4\%$), making them less attractive for operational deployment on this parameter.

Flicker P_{st} (Overall), Boosted Trees, and Support Vector Machines achieve high Accuracy ($> 99\%$) with low FAR ($< 0.4\%$), but differ in MVR. BT attains MVR $\approx 0\%$ in the observed test set, while SVM's MVR is higher ($\approx 73\%$) due to its conservative labelling behaviour. ANN has a higher MVR ($\approx 40\%$) and FAR ($\approx 1.7\%$), which is acceptable for some monitoring tasks but less desirable for early-warning applications. SARIMAX performs poorly for P_{st} , with MVR exceeding 90% , indicating that it almost never anticipates short-term flicker violations.

Flicker P_{It} (Overall), Boosted Trees again offers a favourable compromise: high Accuracy ($\approx 99.5\%$), low FAR ($< 0.6\%$), and MVR $\approx 0\%$. SVM and ANN are competitive but show slightly higher MVR and FAR, suggesting that long-term flicker is easier to handle than P_{st} , but still benefits from the robustness of BT. SARIMAX performs adequately but does not outperform the non-linear models. Overall, Boosted Trees provides the most balanced compliance performance for the volatile parameters (Voltage Unbalance and P_{st}/P_{It}), whereas SVM is particularly strong for V_{THD} with perfect compliance in the test window.

For Frequency, V_{THD} , and V_{RMS} , no violations occur in the test window. As a result, all models achieve 100% Accuracy with FAR = MVR = 0% , validating that they do not generate spurious violations for intrinsically stable parameters. However, this also means that these dimensions do not provide a strong stress-test of violation detection capability; the main discriminating power lies in Voltage Unbalance and P_{st} , where true violations occur.

4.3.3 Threshold Sensitivity ($\pm 5\%$)

To assess robustness to modest changes in regulatory margins, a $\pm 5\%$ threshold sensitivity analysis was performed by tightening and relaxing the NRS 048-2 limits while keeping forecasts fixed. For stable parameters (Frequency, V_{THD} , V_{RMS}), performance metrics remain unchanged under $\pm 5\%$ threshold perturbations, as no values come close to the limits in the test set.

For Voltage Unbalance and flicker P_{st} , tightening thresholds increases the number of labelled violations and shifts MVR and FAR, particularly for flicker P_{st} . Boosted trees, MVR and FAR remain relatively stable, indicating robust discrimination near the compliance boundary. SVM and ANN show larger changes in MVR under tighter thresholds, revealing greater sensitivity to how close predictions lie to the original limits. Detailed deltas and per-model changes are reported in Appendix B (Table B.17). This analysis suggests that the conclusion that BT and SVM are preferred for volatile PQ parameters is not an artefact of a single exact limit setting.

To visualise how the models respond to modest changes in the compliance boundary, Figure 4.11 plots the Missed Violation Rate (MVR) for Flicker P_{st} Overall under the base NRS 048-2 limit and under $\pm 5\%$ perturbations of that limit. P_{st} is selected because it is the most challenging parameter and exhibits the largest changes in MVR in Table B.17. For each model, the bars indicate the fraction of true violations that would be missed when the threshold is tightened (-5%), left at its base value, or relaxed ($+5\%$).

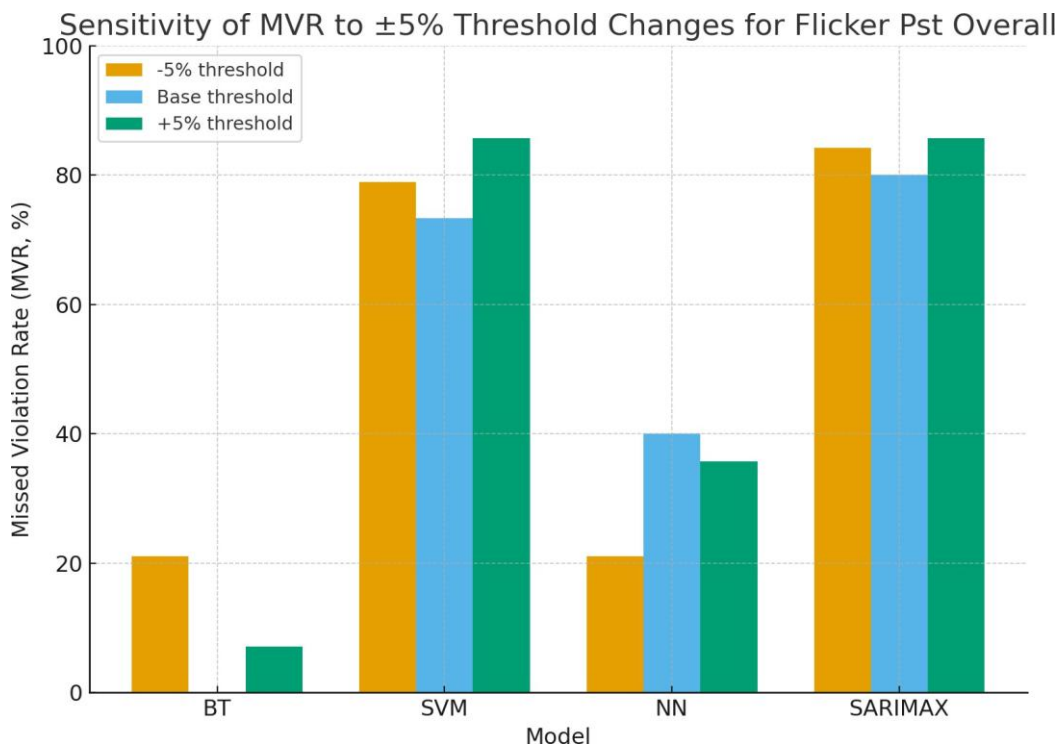


Figure 4.11: Sensitivity of MVR to $\pm 5\%$ threshold changes for Flicker P_{st} Overall

Figure 4.11 shows that Boosted Trees remain relatively robust to small changes in the P_{st} threshold: MVR rises from 0 % at the base limit to about 21 % when the limit is tightened and

remains below 10 % when the limit is relaxed. In contrast, SVM and SARIMAX are highly sensitive to the threshold choice, with MVR already high at the base limit ($\approx 73\text{--}80\%$) and increasing further under both tighter and looser limits. The NN model exhibits an intermediate pattern, with MVR decreasing when the threshold is tightened (reflecting more aggressive labelling) but increasing again when the threshold is relaxed. Taken together, the figure illustrates that BT is not only the best performer at the nominal NRS 048-2 limit but also the most stable under plausible future adjustments to the flicker criterion, whereas SVM and SARIMAX would provide unreliable early-warning behaviour if grid codes were to evolve.

For the remaining PQ parameters (Voltage Unbalance, Plt, V_{RMS} , Frequency, V_{THD}), the changes in Accuracy and FAR under $\pm 5\%$ perturbations are small and follow the patterns already visible in Table B.17. Voltage Unbalance shows moderate increases in MVR under tighter thresholds for all models, again with BT least affected. For brevity, these effects are summarised numerically in Table B.17, while Figure 4.11 focuses on Pst, where the impact of threshold changes is most pronounced.

4.3.4 Operational interpretation

In operational terms, the value of a PQ forecasting model is measured by the actions it enables. A low-MVR model for voltage unbalance, for example, allows operators to detect impending violations early enough to adjust reactive power setpoints, reconfigure transformer tap positions, or temporarily curtail wind farm output before formal non-compliance occurs. For flicker, reliable short-term forecasts can inform the scheduling of large industrial loads or the coordination of on-site capacitor banks to smooth voltage fluctuations.

In both cases, FAR plays a complementary role: excessive false alarms would quickly erode operator trust and could lead to alarm fatigue, while a modest FAR may be acceptable if it significantly reduces missed violations. By interpreting model performance through this lens, the study provides not just a ranking of algorithms, but a mapping from forecast quality to specific operational levers available in the control room.

4.4 FEATURE IMPORTANCE AND INTERPRETABILITY

Beyond aggregate performance, it is important to understand why models behave as they do. Boosted Trees offer built-in measures of feature importance, which we examine for Voltage Unbalance.

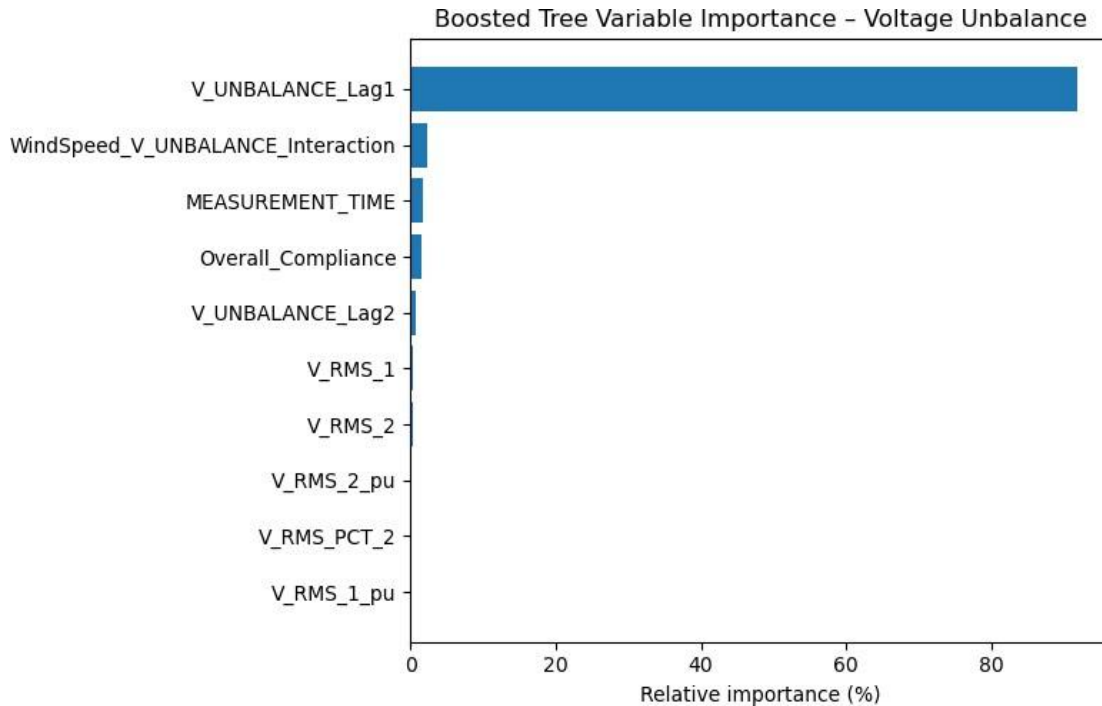


Figure 4.12: Variable importance plot – BT Voltage Unbalance

Figure 4.12 ranks features by their contribution to the BT model for Voltage Unbalance. The key findings are: Lag 1 Voltage Unbalance is by far the dominant predictor, accounting for more than 90 % of the importance share. This reflects strong autoregressive behaviour: current unbalance is highly indicative of unbalance in the next 10-minute interval.

Wind-related interactions (e.g., wind speed $\times$ unbalance) and reactive power-related features appear as secondary contributors, indicating that local PQ dynamics are influenced by the interaction between turbine loading, converter operation, and network conditions. Temporal indicators (e.g., hour-of-day) also contribute, suggesting diurnal patterns in unbalance that align with typical load and wind regimes. Simulated partial dependence plots (not shown in the main text; see Appendix B) further illustrate these dynamics:

The partial dependence of predicted unbalance on Lag 1 Unbalance is almost linear, consistent with temporal persistence. The interaction between wind speed and unbalance shows a threshold-like effect, where unbalance risk increases more sharply beyond certain wind-speed ranges, suggestive of operating ‘stress zones’ for the plant. Time-based features capture mild diurnal cycles, indicating that certain times of day are more prone to elevated unbalance, likely due to combined load and wind patterns.

These interpretability results support the physical plausibility of the BT model and provide a bridge between data-driven forecasting and engineering intuition.

4.5 CHAPTER SUMMARY

This chapter presents the empirical results of applying SARIMAX, SVM, ANN, and Boosted Trees to forecast PQ indices at a South African wind farm, using data from SCADA equipment connected to the PCC. The analysis showed that: Non-linear models, particularly, Boosted Trees and Support Vector Machines consistently outperform the linear SARIMAX baseline in terms of RMSE and R^2 for Voltage Unbalance, V_{THD} , P_{st} , and P_{lt} , while all models perform similarly for stable parameters such as Frequency and V_{RMS} .

ANN is competitive and even superior for flicker P_{st} , highlighting that different PQ metrics may favour different model architectures. Residual diagnostics and distributional checks indicate that the best performing models capture most of the temporal structure and leave no obvious systematic patterns in the residuals. When forecasts are mapped to NRS 048-2 compliance labels, Boosted Trees and Support Vector Machines offer the most favourable balance of low Missed Violation Rates and acceptably low False Alarm Rates for the volatile parameters (Voltage Unbalance, flicker P_{st}/P_{lt}), while all models trivially achieve perfect compliance for intrinsically stable parameters.

A $\pm 5\%$ threshold-sensitivity analysis suggests that these conclusions are robust to small changes in regulatory margins, particularly for the parameters where violations occur. By combining regression metrics with NRS 048-2 compliance metrics and interpretability tools, the chapter provides a nuanced view of model behaviour that is directly relevant to PQ management at the PCC. The next chapter builds on these findings.

Chapter 5 interprets the comparative performance of the models in the broader context of PQ physics, South African grid code requirements, and the international literature reviewed in Chapter 2. It discusses the practical utility of compliance-oriented forecasting, links the empirical results back to the research questions, and clarifies how the published works underpinning the dissertation contribute to the overall research narrative.

CHAPTER 5: DISCUSSION OF FINDINGS

This chapter interprets the empirical results developed in Chapters 3 and 4 and situates them within the broader context established by the systematic literature review and the South African PQ and grid-code framework, especially NRS 048-2. The focus is on explaining why the models behave as they do, how this relates to physical PQ mechanisms and compliance requirements, and what this implies for grid operation and future research. The chapter concludes by explicitly answering the research questions and reflecting on key methodological considerations.

5.1 INTERPRETATION OF MODEL PERFORMANCE

The results in Chapter 4 confirm that machine learning (ML) models can forecast power-quality (PQ) behaviour with high fidelity in a wind-integrated grid, capturing non-linear dynamics that conventional linear models struggle to represent. Across Voltage Unbalance, V_{THD} , Flicker (P_{st} and P_{lt}), Frequency, and V_{RMS} (p.u.), the leading ML models achieve low RMSE and high R^2 , indicating strong signal capture under converter-driven variability and intermittent wind conditions.

5.1.1 Strengths of Non-Linear Models

Boosted Trees (BT).

BT consistently performs best or near-best across Voltage Unbalance, Frequency, Flicker P_{lt} , and V_{RMS} (p.u.), with R^2 values typically exceeding 0.95 and RMSEs significantly lower than those of the SARIMAX baseline (Table 4.1). This pattern reflects BT's ability to model interaction effects between lagged PQ indices, wind-speed features, and temporal indicators, and to iteratively reduce residual error through stage-wise boosting. The variable-importance analysis for Voltage Unbalance (Figure 4.8) shows that Lag-1 Unbalance ($V_{UNBALANCE_Lag1}$) dominates the model with roughly 92% of the split-based importance, while wind-speed interaction terms, measurement time, and overall compliance contribute smaller but meaningful corrections. This aligns closely with engineering intuition: over a 10-minute horizon, Voltage Unbalance exhibits strong persistence, but its absolute level is modulated by turbine loading, wind regime, and diurnal load cycles. BT therefore acts as a flexible non-linear

autoregressive model with physically plausible drivers.

Support Vector Machines (SVM).

SVMs excel on V_{THD} and are competitive on Voltage Unbalance and V_{RMS} (p.u.), reaching $R^2 \approx 0.97$ for V_{THD} and ≈ 0.95 for Unbalance (Table 4.1). Kernelization via the RBF kernel allows SVM to represent smooth but non-linear dependencies between harmonic distortion and driver variables such as load levels, converter operation, and voltage magnitude, without overfitting. The relatively low FAR and moderate MVR values for SVM in the V_{THD} and P_{It} compliance analysis (Table 4.2, Figure 4.10) further support its suitability for parameters with stable behaviour and gradual trends. SVM is therefore particularly attractive where harmonic limits are tight, but violations are expected to evolve relatively smoothly.

Neural Networks (NN).

The two-hidden-layer MLP delivers the best performance on short-term flicker P_{st} ($R^2 \approx 0.79$), where dynamics are bursty and less regular than for V_{RMS} or Frequency (Table 4.1). This is consistent with NNs' flexibility as universal function approximators, which can capture complex nonlinear patterns when appropriately regularised. However, NN performance on other parameters is more variable and sensitive to architectural choices and regularisation, underscoring the need for disciplined hyperparameter tuning and careful monitoring of overfitting. In particular, the elevated MVR for Voltage Unbalance ($\approx 40\%$, Table 4.2) indicates that a good regression fit does not automatically translate into reliable violation detection.

5.1.2 Limitations of Linear Models

Seasonal Auto-Regressive Integrated Moving Average with exogenous variables (SARIMAX): The SARIMAX baseline performs acceptably on inherently stable parameters (Frequency, V_{RMS} , and, to some extent, V_{THD}), but underperforms on non-linear and more volatile targets such as Voltage Unbalance and P_{st} . RMSE and R^2 values are markedly worse than those of BT and SVM for these parameters (Table 4.1), and the MVR for flicker P_{st} reaches approximately 80–92 % depending on threshold settings (Tables 4.2 and B.17).

This behaviour aligns with both theory and the systematic review in Publication 1 linear, Gaussian assumptions are frequently violated in converter-rich, RES-dominated grids, where PQ dynamics are influenced by threshold effects, saturation, discrete control actions, and regime shifts. Automatically selected SARIMAX orders within the constraints of JMP Pro (Student Edition) were sufficient to capture daily seasonality and moderate persistence, but could not model sharp changes associated with wind ramps, switching events, or non-linear control. SARIMAX is therefore useful as an interpretable baseline and sanity check, but not as an operational forecasting workhorse for PQ compliance.

5.2 THE PRACTICAL UTILITY OF NRS 048-2 COMPLIANCE ORIENTED FORECASTING

Publication 1 highlighted that most PQ forecasting studies rely almost exclusively on error metrics such as RMSE, MAE, and R^2 , with limited attention to how models perform relative to standards-based limits and operational risk. This study addresses that gap by evaluating Missed Violation Rate (MVR) and False Alarm Rate (FAR) alongside traditional metrics and directly mapping forecasts to NRS 048-2 limits.

Minimising Operational Risk (MVR).

Missed Violation Rates quantifies the proportion of actual violations that the model fails to flag. For Voltage Unbalance, BT achieves an MVR of $\approx 5\%$, meaning that 95 % of observed violations in the test set are correctly predicted as non-compliant (Table 4.2). For Flicker P_{It} , BT again attains an MVR of 0–5 % under base and perturbed thresholds (Tables 4.2 and B.17). By contrast, the NN and SARIMAX models exhibit MVR values of $\approx 40\%$ for Unbalance and up to 80–90 % for flicker P_{st} , which would be unacceptable for alerting and early warning. The FAR–MVR scatter plot for the volatile parameters (Figure 4.10) makes these differences visible: BT points cluster near the bottom-left corner (low FAR, low MVR), whereas ANN and SARIMAX occupy the upper-right region for flicker P_{st} , indicating high operational risk.

Avoiding Alarm Fatigue (FAR).

FAR measures the proportion of compliant intervals incorrectly flagged as violations. Excessive FAR leads to alarm fatigue and erodes operator trust. For most parameters, BT and SVM

maintain FAR well below 1 %, even on volatile quantities such as flicker P_{st} and flicker P_{It} (Table 4.2). ANN and SARIMAX show somewhat higher FAR, particularly when thresholds are tightened by 5 % (Table B.17), implying more nuisance alarms. Combined analysis of FAR and MVR, rather than either metric alone, therefore identifies BT as offering the best risk trade-off.

Robustness to Evolving Grid Code Limits.

The threshold-sensitivity analysis (Table B.17, Figure 4.11) demonstrates that BT's MVR and FAR remain relatively stable under ± 5 % changes to flicker P_{st} and unbalance limits, whereas SVM, ANN, and SARIMAX are more sensitive. This matters because grid-code margins can evolve over time in response to changing network conditions, policy targets, or customer expectations. A model whose risk profile is stable under such adjustments provides a more reliable foundation for long-term operational planning and regulatory negotiation.

In summary, FAR and MVR provide a decision relevant perspective that pure regression metrics cannot: they quantify the trade-off between missed violations and nuisance alarms, and thus discriminate between models with similar accuracy but very different safety profiles.

5.3 LOCALISED PERSPECTIVE ON PQ FORECASTING IN SOUTH AFRICA

The systematic literature review and Chapter 2 showed that PQ forecasting literature is heavily concentrated in European and Asian grids, with African systems largely absent. This case study therefore provides a South African Point of Common Coupling/ Point of Connection-level perspective that complements the predominantly foreign evidence, while still being methodologically comparable.

Three aspects are particularly noteworthy:

5.3.1 Contextual Relevance

The models are trained and evaluated using South African measurement conventions (IEC 61000-4-30 Class A, IEC 61000-4-15) and NRS 048-2 limits for voltage variation, unbalance, flicker, and harmonics. This ensures that results are directly interpretable by local utilities and regulators rather than extrapolated from foreign grid codes with different limits, network topologies, or operating practices.

5.3.2 Methodological Alignment with Global Practice

The six-phase workflow in Chapter 3 (data acquisition, preprocessing, feature engineering, temporal splitting, modelling, and compliance-oriented evaluation) embodies best practices such as anti-leakage safeguards, chronological validation, and explicit compliance mapping. These steps align with the international PQ forecasting literature, making it possible to compare South African findings with results from other regions on a like for like basis.

5.3.3 Data Realism

Unlike many studies that rely on synthetic PQ signals or low-resolution measurements, this work uses 10-minute SCADA data captured under real operating conditions, including the influence of load-shedding events, wind ramps, and normal plant operations. This realism is critical for understanding how forecasting algorithms behave under the actual mix of disturbances and steady states faced by South African grids. Together, these elements provide a localised but globally recognisable perspective: the findings speak directly to South African stakeholders while remaining comparable to international studies reviewed in Chapter 2.

5.4 IMPLICATIONS FOR GRID MANAGEMENT AND PREDICTIVE MAINTENANCE

From an operational standpoint, the important question is not only which model is most accurate, but how forecasts can concretely support control room decisions. The compliance-oriented evaluation framework in Chapter 4 allows model outputs to be translated into a sequence of potential actions.

5.4.1 Forecasts as Early Warning Tools

For parameters with genuine violation risk primarily voltage unbalance and flicker P_{st} , the demonstrated boosted tree and support vector machines models can serve as early warning indicators. When forecasts approach or exceed the NRS 048-2 limit over the next 10–30 minutes, operators can, for example; adjust transformer tap positions to bring voltages back within tolerance; modify reactive power set points or capacitor bank schedules to improve voltage balance; temporarily curtail wind plant output during critical combinations of high wind speed and weak network conditions; prioritise sensitive feeders or loads that are known to be vulnerable to PQ deviations.

The models operate on 10-minute SCADA cycles, they can provide several time steps of warning before a formal non-compliance is recorded over a 2-hour or 24-hour aggregation window. This aligns with emerging notions of predictive PQ management, in which operators act on risk indicators rather than waiting for violations to appear in ex post reporting.

5.4.2 Link to Maintenance and Asset Management

Persistent or clustered forecast violations in specific time bands, wind regimes, or network configurations can act as soft indicators of asset stress, for example, ageing converter modules, weak feeder sections, or problematic load clusters. By analysing where boosted trees and support vector machines predict high risk, utilities can: target field inspections towards the most stressed equipment; refine maintenance schedules for transformers, converters, and compensation devices; assess whether reinforcement or reconfiguration of particular network segments is warranted.

In this way, the forecasting framework supplements traditional PQ monitoring and condition-based maintenance with a predictive, model-based view.

5.4.3 Support for Grid Code Compliance Management and Planning

Finally, by quantifying Missed Violation Rates and False Alarm Rates under different threshold settings, the models provide a basis for scenario analysis in regulatory and planning discussions. For example, a utility can estimate how tightening the flicker P_{st} limit by 5 % would change the expected frequency of violations and the operational interventions needed to maintain compliance. Such analyses can inform both internal investment decisions and external negotiations with regulators regarding the feasibility of PQ performance under increasing renewable penetration.

5.5 ANSWERING THE RESEARCH QUESTIONS AND OBJECTIVES

This section revisits the research questions and objectives defined in Chapter 1 and summarises how the findings address them.

5.5.1 Primary Research Question

Research Question 1: How effectively can machine learning models predict power quality violations in a South African wind power plant when evaluated using both traditional accuracy

metrics and practical, compliance-oriented metrics?

Answer:

The models perform effectively, with clear stratification between family classification. BT and SVM achieve high R^2 values (e.g., $R^2 \approx 0.95$ for Voltage Unbalance, ≈ 0.97 for V_{THD} , and ≈ 0.999 for V_{RMS} (p.u.)) combined with low RMSE (Table 4.1). When mapped to compliance labels, these models maintain low MVR on violation-prone targets ($\approx 5\%$ for Unbalance and $\approx 0\text{--}6\%$ for P_{It}) and near-zero FAR on most parameters (Table 4.2, Figures 4.10–4.11). For stable parameters with no observed violations (Frequency, V_{THD} , V_{RMS}), all models trivially achieve $FAR = MVR = 0\%$, confirming that they do not invent spurious alarms. Overall, BT and SVM are capable of delivering operationally useful forecasts for PQ compliance at the PCC.

5.5.2 Secondary Questions

Research Question 2: Which ML algorithms minimise MVR for key PQ parameters?

For Voltage Unbalance, BT attains the lowest MVR ($\approx 5\%$), while keeping FAR below 0.3% , making it the most reliable choice for critical alerting.

For Flicker P_{It} , BT and SVM both deliver low MVR ($0\text{--}6\%$) and low FAR ($< 1\%$), with BT slightly more robust under $\pm 5\%$ threshold perturbations (Table B.17).

For Flicker P_{st} , NN achieves strong regression performance but exhibits high MVR under base thresholds; BT offers a more balanced profile, especially when thresholds are tightened, and therefore represents a safer operational choice.

Research Question 3: What is the practical utility of FAR/MVR for model assessment?

FAR and MVR reveal differences that are invisible when only RMSE or R^2 are considered. For instance, NN and SARIMAX can achieve respectable R^2 for Voltage Unbalance yet still miss 40% of true violations, whereas BT attains similar or better continuous accuracy with far fewer missed events. This distinction is crucial for control-room trust: a model with slightly higher RMSE but much lower MVR may be preferable in practice. The FAR–MVR scatter (Figure 4.10) translates these trade-offs into a risk map that operators can interpret directly, thereby operationalising model comparison.

Research Question 4:

Can a robust, reproducible PQ forecasting pipeline be constructed?

Chapter 3 demonstrates a six-phase from data acquisition through to dual-track evaluation implemented with strict temporal splits, anti-leakage controls, and standards-aligned thresholds. Fixed random seeds and detailed documentation of preprocessing and model configuration further support reproducibility. The same pipeline successfully handled multiple targets and model classified families, suggesting that it is a robust template for future PQ forecasting studies in similar environments.

Alignment with Objectives.

Objective 1 (design data/feature pipeline): Achieved through the domain informed feature-engineering strategy (lags, rolling windows, interaction terms) and subsequent importance analysis in Section 4.4.

Objective 2 (train and benchmark models): Achieved via systematic comparison of BT, SVM, NN, and SARIMAX across all PQ indices in Section 4.2.

Objective 3 (evaluate using accuracy + compliance metrics): Achieved through combined RMSE/ R^2 and FAR/MVR analysis, including threshold sensitivity experiments in Sections 4.3–4.3.4.

Objective 4 (provide operationally meaningful insights): Addressed by mapping model outputs to NRS 048-2 compliance regimes and discussing their implications for operator actions and maintenance in Sections 5.2–5.4.

These answers and objective mappings prepare the ground for the synthesis of contributions, limitations, and future directions in Chapter 6.

5.6 METHODOLOGICAL AND DESIGN CONSIDERATIONS

Although the findings are encouraging, several methodological considerations and design choices should be acknowledged to avoid over generalisation and to guide future work.

5.6.1 Single-site scope

The analysis focuses on one utility scale wind plant. While the methods are transferable, the quantified performance and learned relationships may not generalise directly to other plants, technologies, or grid regions. Replication on additional South African sites and ideally non-South African grids would be valuable.

5.6.2 Limited exposure to rare disturbance

Frequency, V_{THD} , V_{RMS} , and P_{It} exhibit no or very few violations in the test window. As a result, 0 % FAR/MVR for these parameters primarily reflects genuinely benign operating conditions rather than a stringent stress test of detection capability. Longer datasets or targeted inclusion of periods with known PQ events would allow a more rigorous evaluation.

5.6.3 Feature boundary conditions

The feature set is restricted to PCC-level SCADA and local meteorology. System-wide exogenous signals such as regional load levels, neighbouring PCC measurements, or planned switching events were not available but could enhance predictive skill and interpretability if incorporated in future work.

5.6.4 Temporal resolution

The 10-minute resolution is well suited to compliance assessment and operational foresight at the SCADA/EMS level but cannot represent sub-cycle or sub-minute transients. Complementary studies using waveform-level data and advanced signal-processing features would deepen understanding of how forecasted PQ indices relate to underlying disturbance mechanisms.

5.6.5 Tooling constraints

The use of JMP Pro (Student Edition) limited manual SARIMAX specification and constrained some interpretability tooling (e.g., advanced residual diagnostics or bespoke visualisations). This likely contributed to under-optimised linear baselines and necessitated workarounds for some analyses. Future work leveraging full featured open-source toolchains could strengthen

the statistical analysis and facilitate broader adoption.

5.6.6 Model class scope

The study deliberately focuses on widely used, relatively compact models (BT, SVM, MLP, SARIMAX) to align with current utility practice. More recent architectures such as LSTM/GRU networks, CNN–LSTM hybrids, transformers, or probabilistic forecasting frameworks were not implemented here, but represent promising avenues for future research.

The most important high-level constraints of this research programme are revisited and further synthesised in Chapter 6, where they are linked explicitly to recommendations for future work and potential extensions of the proposed compliance-oriented forecasting framework.

5.7 SUMMARY

This chapter has interpreted the empirical findings from Chapters 3 and 4 in the light of PQ physics, the international literature, and the South African NRS 048-2 compliance framework. It showed that non-linear models, particularly boosted trees and Support vector machines, provide the most effective balance between regression accuracy and compliance performance, while the linear SARIMAX baseline underperforms on non-linear, volatile parameters such as voltage unbalance and short-term flicker. By introducing compliance-oriented metrics (MVR and FAR) and analysing their behaviour under ± 5 % threshold perturbations, the chapter moved evaluation from abstract error measures to explicit operational risk, highlighting models that are not only accurate but also safe and trustworthy for control room use.

The discussion also positioned the case study within the broader PQ forecasting literature, emphasising its contribution as a South African, standards-aligned example that complements predominantly European and Asian evidence. It outlined how forecasts can be translated into concrete operator actions, ranging from tap changes and reactive compensation to targeted maintenance and scenario analysis for regulatory planning and demonstrated that the proposed six-phase workflow yields a robust, reproducible pipeline for PQ forecasting at the PCC. The research questions and objectives from Chapter 1 were explicitly revisited and answered, and key methodological considerations and design limitations were identified as boundaries on the interpretation of the results.

The next chapter builds directly on these insights. The chapter synthesises the work at the dissertation level, articulating the main contributions to knowledge, consolidating the practical implications for industry, and framing the overarching limitations and future research directions. It moves from detailed discussion to a concise statement of what this research adds to the field and how it can guide the development of compliance-oriented PQ forecasting in wind integrated power systems.

CHAPTER 6: CONCLUSION AND FUTURE WORK

This chapter synthesises insights from the literature review (Publication 1) and the empirical case study (Publication 2) and frames them relative to NRS 048-2 compliance requirements and the broader challenge of power quality (PQ) management in wind-integrated grids. It summarises the research, highlights key contributions, discusses practical implications, reflects on limitations, and outlines directions for future work.

6.1 SUMMARY OF RESEARCH

This dissertation evaluated the performance of four machine learning (ML) algorithms; Support Vector Machines (SVM), Neural Networks (NN), Boosted Trees (BT), and Seasonal Auto-Regressive Integrated Moving Average with exogenous variables (SARIMAX); for forecasting critical PQ parameters (Voltage Unbalance, Frequency, Total Harmonic Distortion of voltage (V_THD), Flicker P_{st}/P_{lt}, and Voltage RMS (V_RMS_pu)) in a South African wind power plant using 10-minute operational data.

The work was motivated by two key gaps identified in the systematic literature review (Chapter 2, Publication 1). Limited evidence from African grids, particularly PQ forecasting in renewable-dominated systems, and over-reliance on conventional accuracy metrics (e.g., RMSE, MAE, R²) that do not reflect operational risk or compliance performance, especially Missed Violation Rate (MVR) and False Alarm Rate (FAR).

Chapters 3 and 4 implemented a sequential, leakage-aware workflow in JMP Pro (Student Edition), combining standards-aligned data preprocessing and feature engineering, temporally faithful training/validation/test splits, and multi-model benchmarking. Models were evaluated using both regression metrics and compliance-oriented metrics derived from NRS 048-2 thresholds at the point of common coupling (PCC).

The empirical results demonstrated that ML models can reliably forecast PQ behaviour over 10-minute horizons. However, the best model is parameter specific: BT excelled on Voltage Unbalance, Frequency, P_{lt}, and V_RMS_pu; SVM performed best on V_THD (and

competitive for system-level Flicker); NN was strongest on flicker P_{st}; SARIMAX underperformed on non-linear or volatile targets but provided an interpretable linear baseline.

Chapter 5 then integrated these findings with the dissertation objectives and publications, providing a detailed interpretation of model behaviour, implications for grid operations, and explicit answers to the research questions.

6.2 KEY CONTRIBUTIONS TO KNOWLEDGE

The research makes several contributions at contextual, methodological, empirical, and practical levels.

6.2.1 Contextual contribution (South Africa, globally relevant)

This is the first detailed, localised validation of advanced machine learning based PQ forecasting in the South African context using high-resolution operational SCADA data, explicitly tied to NRS 048-2 compliance bands. It directly addresses the geographic evidence gap flagged by the systematic literature review, providing PCC-level results for voltage unbalance, flicker, harmonics, frequency, and voltage quality in a wind energy integrated network. While the case study is local, the methodology is aligned with international practice, enabling comparison with studies from Europe and Asia.

6.2.2 Methodological Contribution

The study advances a practical evaluation framework that pairs classical regression metrics (RMSE, MAE, MAPE, R^2) with compliance-oriented metrics (MVR and FAR). This moves evaluation from abstract error statistics to operational risk, answering repeated calls in the literature for standards-aware forecasting. The framework also incorporates threshold-sensitivity analysis ($\pm 5\%$ around NRS 048-2 limits) to assess robustness under evolving grid-code margins.

6.2.3 Empirical contribution (Non-Linear Necessity and Suitability Map)

Empirically, the work shows that non-linear models, particularly BT and SVM, achieve low MVR on violation-prone targets when compared to SARIMAX, while maintaining very low FAR. This

confirms that non-linear methods are not merely a convenience but a necessity for PQ forecasting in converter-rich, wind-integrated grids. To operationalise these insights, the dissertation introduces a parameter model suitability map, indicating which algorithms are most appropriate for Unbalance, Flicker, V_THD, Frequency, and V_RMS_pu under NRS 048-2 limits.

6.2.4 Practical/Tooling contribution (Accessible Analytics Pathway)

By executing the full pipeline in JMP Pro (Student Edition), the study demonstrates a viable, lower barrier pathway for conducting PQ analytics in resource-constrained utility environments. The runbook for preprocessing recipes, feature definitions, temporal splits, and model configurations demonstrates how meaningful forecasting and compliance analysis can be achieved without large, bespoke code bases or specialist data-science infrastructure.

6.2.5 Per-Phase Granularity

A further contribution is the per-phase analysis of PQ indices. Rather than reporting only aggregated metrics, the study evaluates model performance on individual phases, revealing asymmetries and worst-case behaviours that are often obscured in system-level summaries. This granularity supports more targeted operational and maintenance decisions in three-phase systems.

6.3 PRACTICAL IMPLICATIONS FOR INDUSTRY

While Section 5.4 focused on how forecasts can support day-to-day control room decisions at a single PCC, the findings also have broader implications for utilities, system operators, and regulators at an industry level.

6.3.1 Shift from Reactive to Predictive PQ Management

The results show that compact, non-linear models such as BT and SVM can provide reliable, standards-aligned forecasts of key PQ indicators with low MVR and FAR. This enables a structural shift away from purely retrospective compliance reporting (after a 2-hour or 24-hour window has closed) towards predictive management, where operators and planners manage

risk of non-compliance in advance. For the industry, this means PQ performance can increasingly be governed by forward-looking indicators rather than by ex post penalties.

6.3.2 Strategic Benefits for Reliability and Cost

At the system level, better anticipation of PQ stress translates into fewer unexpected process interruptions, reduced equipment damage, and more stable operating conditions for large industrial and commercial customers. Over time, this can lower PQ-related claims, reduce corrective maintenance costs, and improve the perceived quality of supply. The evidence from this study suggests that such benefits are achievable with relatively modest computational resources, making predictive PQ management accessible even to utilities with constrained analytics capacity.

6.3.3 Implications for Policy, Regulation, and Planning

Because the forecasting framework is explicitly tied to NRS 048-2 limits, it provides regulators and utilities with a quantitative basis for discussing realistic PQ performance in high renewable energy environments. Threshold sensitivity analysis illustrates how tightening or relaxing specific limits (e.g., flicker P_{st} or voltage unbalance) would affect the frequency of predicted violations and the operational effort required to remain compliant. At an industry level, such tools can support more nuanced PQ standards, performance-based incentives, and investment planning that recognises both the benefits and the challenges of large-scale renewable integration.

6.3.4 Adoption and Capacity Building

Finally, the use of JMP Pro (Student Edition) demonstrates that meaningful PQ forecasting does not require bespoke, large-scale data science infrastructure. For many utilities, this lowers the barrier to initial adoption: pilot projects can be run using existing SCADA data and off-the-shelf software, with more advanced toolchains introduced later as internal capability grows. From an industry perspective, this supports a gradual, scalable pathway from traditional PQ monitoring to fully predictive, compliance-aware PQ management.

6.4 LIMITATIONS

Despite its contributions, the study has several important limitations that must be recognised when interpreting the results and planning future work.

6.4.1 Single-site scope

Findings reflect a single utility scale wind plant connected to a specific South African network. While the methodology is transferable, the quantified performance and learned relationships may not generalise directly to other technologies, PCC characteristics, or climates. Multi-site studies are needed to test generality.

The three-month observation window is also a material limitation. It covers only late-summer to autumn conditions and does not include full annual seasonality, winter wind regimes, or seasonal load variation. Consequently, the empirical results demonstrate model behaviour for this case-study window, not year-round regulatory performance. Longer multi-season datasets are required before the trained coefficients or event rates can be generalised to annual NRS 048-2 compliance assessment.

6.4.2 Rare event exposure

Several parameters (Frequency, V_{THD} , V_{RMS} , and, to a lesser extent, flicker P_{It}) had no or very few test-set violations. While 0 % FAR/MVR confirms compliant operation, it limits the ability to stress-test detection under genuinely adverse PQ conditions. Longer observation windows or targeted inclusion of known disturbance periods would strengthen the evaluation.

6.4.3 Feature availability

The feature set is limited to PCC-level SCADA and local meteorological measurements. Exogenous grid-wide signals (e.g., area load, neighbouring PCC metrics, switching logs, protection operations) were not explicitly modelled but could provide additional predictive signals and context.

6.4.4 Temporal resolution

Ten-minute sampling is appropriate for compliance reporting and operational foresight at

SCADA time scales, but it cannot capture sub-cycle or sub-minute transients. Complementary high-frequency waveform studies, with appropriate signal-processing and ML tools, would provide deeper insight into the mechanisms underlying aggregated PQ indices.

6.4.5 Tooling constraints

The use of JMP Pro (Student Edition) constrained manual SARIMAX order specification, some residual diagnostics, and advanced interpretability visualisations. In several instances, approximations or default settings were required. Although sufficient for comparative benchmarking, these constraints may have limited the attainable performance of linear models and the richness of diagnostic analysis.

6.4.6 Model-class scope

The study deliberately focused on widely used, relatively compact models (BT, SVM, MLP, SARIMAX) to align with current utility practice and keep computational requirements modest. Modern sequence models (LSTM/GRU, CNN–LSTM hybrids, transformers) and probabilistic forecasting approaches were not implemented, leaving open the question of how much additional benefit they could provide in this PQ context.

These limitations do not invalidate the findings but delineate the boundaries within which conclusions should be interpreted and point towards concrete avenues for future research.

6.5 RECOMMENDATIONS FOR FUTURE RESEARCH

Building on the above limitations and the operational insights from Chapters 4–5, several directions for future work are recommended.

6.5.1 Real-time Integration and Alerting

A natural next step is to embed the best-performing models (BT and SVM) into live SCADA data streams with low-latency scoring. This would include operator dashboards, alarm hysteresis and dwell logic, and interfaces to existing PQ monitoring systems, enabling truly predictive PQ management.

6.5.2 Uncertainty Aware Forecasts

Future models should provide prediction intervals or probabilistic outputs, particularly for volatile parameters such as flicker P_{st} . Quantifying forecast uncertainty would support risk aware operational decisions, allowing operators to weigh the probability and consequences of potential violations rather than acting on point estimates alone.

6.5.3 Optimisation Coupled Mitigation

Forecasts could be coupled with optimisation algorithms (e.g., for VAR support, inverter set points, or network configuration) to move from descriptive alerts to prescriptive control actions that comply with NRS 048-2 limits and broader grid code constraints.

6.5.4 Transferability and Scaling

Evaluating model portability across multiple wind plants and other Renewable Energy System technologies (e.g., PV farms) is essential. Questions include how much retraining is required, which features generalise across sites, and how transfer learning or fine-tuning can be used to adapt models efficiently.

6.5.5 Advanced model architectures

Exploring hybrid ensembles (e.g., BT+SVM, or signal-processing features combined with deep learning), sequence models (LSTM/GRU, transformers), and probabilistic frameworks (e.g., quantile regression forests, Bayesian methods) may yield further improvements on challenging targets such as flicker P_{st} , especially when longer datasets become available.

6.5.6 Explainability and operator tooling

Expanding the use of SHAP-style diagnostics, case-based reasoning, and intuitive visual interfaces can help bridge the gap between data science models and operational decision-making. This is particularly important for “black-box” methods such as deep neural networks, where trust and interpretability are prerequisites for deployment.

6.5.7 Digital-twin Stress Tests

Where real disturbances are rare, simulated contingency scenarios in a digital-twin

environment could be used to evaluate model and controller responses to extreme PQ events (e.g., large wind ramps, severe unbalance, or harmonic resonance) under NRS 048-2 constraints.

Synthetic data benchmarking should also be developed as a reproducibility extension. A synthetic PQ dataset could be generated from the empirical distributions and autocorrelation structure of the real SCADA data, with controlled injection of rare events such as flicker bursts, voltage-unbalance excursions, and harmonic limit approaches. This would allow model behaviour to be tested under conditions not sufficiently represented in the confidential three-month dataset while protecting the operator's raw data.

6.5.8 Reproducible artefacts and sharing

Subject to confidentiality constraints, packaging feature recipes, temporal splits, model configurations, and evaluation harnesses as documented artefacts would support independent auditing and reuse by other utilities, regulators, or researchers.

6.6 CONCLUDING REMARKS

This dissertation demonstrates that machine learning, evaluated through an operator centric compliance lens, can materially enhance proactive power quality management in wind energy integrated grids. By coupling a rigorous, standards-aligned workflow with accessible tooling and a South African case study, the work provides both a blueprint and an evidence base for wider deployment of PQ forecasting under NRS 048-2 and similar frameworks.

As renewable penetration deepens and networks become more converter and inverter dominated, predictive capabilities of this kind will be essential for maintaining stability, optimising operations, and delivering a reliable, sustainable electricity supply. The contributions made here from localised empirical evidence, a compliance-oriented evaluation framework, and a practical modelling pipeline represent initial but important steps towards that goal.

REFERENCES

- [1] A. F. Zobaa and R. C. Bansal, Handbook of Renewable Energy Technology. Singapore: World Scientific, 2011.
- [2] M. H. J. Bollen and I. Y. H. Gu, Signal Processing of Power Quality Disturbances. Hoboken, NJ: Wiley-IEEE Press, 2006.
- [3] R. C. Dugan, M. F. McGranaghan, S. Santoso and H. W. Beaty, Electrical Power Systems Quality, 3rd ed. New York: McGraw-Hill, 2012.
- [4] S. Shen, T. Wang and Z. Song, "Online performance and proactive maintenance assessment of data-driven prediction models," Journal of Intelligent Manufacturing, vol. 35, no. 8, pp. 3959-3993, 2024.
- [5] G. G. Tolun, O. C. Tolun and K. Zor, "Advanced machine learning algorithms for reactive power forecasting in electric distribution systems," e-Prime - Advances in Electrical Engineering, Electronics and Energy, vol. 13, article 101019, 2025, doi: 10.1016/j.prime.2025.101019.
- [6] A. E. Rhatrif, B. Bouihi and M. Mestari, "Challenges and limitations of artificial intelligence implementation in modern power grid," Procedia Computer Science, vol. 236, pp. 83-92, 2024, doi: 10.1016/j.procs.2024.05.008.
- [7] K. Arumugam, Y. Swathi, D. T. Sanchez, M. Mustafa, C. Phoemchalard, K. Phasinam and E. Okoronkwo, "Towards applicability of machine learning techniques in agriculture and energy sector," Materials Today: Proceedings, vol. 51, pp. 2260-2263, 2022, doi: 10.1016/j.matpr.2021.11.394.
- [8] D. R. Dipto, S. K. Shib, M. T. Rahman and A. Shufian, "Predictive analysis of smart grid stability using diverse machine learning algorithms," in Proceedings of the 2025 4th International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), Dhaka, Bangladesh, 2025, pp. 307-311, doi: 10.1109/ICREST63960.2025.10914368.
- [9] U. J. Minnaar, B. Peterson, H. Mostert, J. Rens and G. Botha, "Power quality grid code compliance for renewable power plants in South Africa," IET Generation, Transmission & Distribution, vol. 13, no. 1, pp. 137-144, 2019.
- [10] G. Viljoen and F. Dube, "Realising the right to electricity through off-grid power solutions in South Africa," Potchefstroom Electronic Law Journal / Potchefstroomse Elektroniese Regsblad, vol. 26, 2023.
- [11] S. K. Fasogbon, I. A. Fetuga, A. T. Oyeniran, S. A. Shaibu, S. Afolabi, T. A. Ndokwu, S. R. Oluwadare, J. T. Onafowokan, O. S. Eso and V. B. Bassey, "Optimization of energy grid efficiency with machine learning: A comprehensive review of challenges and opportunities," Renewable and Sustainable Energy Reviews, vol. 223, article 115980, 2025, doi: 10.1016/j.rser.2025.115980.
- [12] A. A. Adebisi and K. Moloi, "Renewable energy source utilization progress in South Africa: A review," Energies, vol. 17, no. 14, article 3487, 2024, doi: 10.3390/en17143487.
- [13] D. M. W. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation," Journal of Machine Learning Technologies, vol. 2, no. 1, pp. 37-63, 2011.
- [14] SAS Institute Inc., JMP Pro 18 Statistical Discovery Software. Cary, NC: SAS Institute Inc., 2024.
- [15] International Electrotechnical Commission, IEC 61000-4-30:2015+A1:2021, Electromagnetic Compatibility (EMC) - Part 4-30: Testing and Measurement Techniques - Power Quality Measurement Methods. Geneva: IEC, 2021.
- [16] P. S. Kack, S. Arora and M. K. Singh, "Mitigation of flicker in power systems using advanced control strategies: A review," International Journal of Engineering Research & Technology, vol.

10, no. 6, pp. 46-50, 2021.

[17] S. Choudhury and G. K. Sahoo, "A critical analysis of different power quality improvement techniques in microgrid," *e-Prime - Advances in Electrical Engineering, Electronics and Energy*, vol. 8, article 100520, 2024, doi: 10.1016/j.prime.2024.100520.

[18] J. Zhang, T. Sheng, P. Gu, M. Yu, J. Yan, J. Sun and S. Liu, "Dynamics power quality cost assessment based on a gradient descent method," *Energies*, vol. 17, no. 9, article 2104, 2024, doi: 10.3390/en17092104.

[19] M. L. Cindi, M. Sibiya and E. Markus, "Machine learning as a tool to forecast the power quality of wind energy power plants: A systemic literature review," *International Journal of Communication Networks and Information Security*, vol. 16, no. 1 Special Issue, pp. 1284-1316, 2024.

[20] NRS Association, NRS 048-2:2025, Electricity Supply - Quality of Supply, Part 2: Voltage Characteristics, Compatibility Levels, Limits and Assessment Methods, 5th ed. Pretoria: NRS Association, 2025.

[21] CENELEC, EN 50160:2022, Voltage Characteristics of Electricity Supplied by Public Electricity Networks. Brussels: European Committee for Electrotechnical Standardization, 2022.

[22] International Electrotechnical Commission, IEC 61000-4-15:2010, Electromagnetic Compatibility (EMC) - Part 4-15: Testing and Measurement Techniques - Flickermeter - Functional and Design Specifications. Geneva: IEC, 2010.

[23] IEEE Standards Association, IEEE Std 1547-2003, IEEE Standard for Interconnecting Distributed Resources with Electric Power Systems. New York: IEEE, 2003.

[24] G. Gyasi and J. Wang, "Parallel-connected timers and deadbands for reducing the number of false and missed alarms," *Control Engineering Practice*, vol. 134, article 105478, 2023, doi: 10.1016/j.conengprac.2023.105478.

[25] N. E. Benti, M. D. Chaka and A. G. Semie, "Forecasting renewable energy generation with machine learning and deep learning: Current advances and future prospects," *Sustainability*, vol. 15, no. 9, article 7087, 2023, doi: 10.3390/su15097087.

[26] E. Cebekhulu, A. J. Onumanyi and S. J. Isaac, "Performance analysis of machine learning algorithms for energy demand-supply prediction in smart grids," *Sustainability*, vol. 14, no. 5, article 2546, 2022, doi: 10.3390/su14052546.

[27] T. Ahmad, H. Zhang and B. Yan, "A review on renewable energy and electricity requirement forecasting models for smart grid and buildings," *Sustainable Cities and Society*, vol. 55, article 102052, 2020, doi: 10.1016/j.scs.2020.102052.

[28] G. J. Ponkumar, S. Jayaprakash and K. Kanagarathinam, "Advanced machine learning techniques for accurate very-short-term wind power forecasting in wind energy systems using historical data analysis," *Energies*, vol. 16, no. 14, article 5459, 2023.

[29] A. Vinayagam, V. Veerasamy, M. Tariq and A. Aziz, "Heterogeneous learning method of ensemble classifiers for identification and classification of power quality events and fault transients in wind power integrated microgrid," *Sustainable Energy, Grids and Networks*, vol. 31, article 100752, 2022, doi: 10.1016/j.segan.2022.100752.

[30] S. L. Gbadamosi and N. I. Nwulu, "Optimal microgrid sizing incorporating machine learning forecasting," in *Proceedings of the International Conference on Industrial Engineering and Operations Management*, Toronto, Canada, 2019.

[31] J. Huang, Z. Jiang and M. Negnevitsky, "Power quality disturbance feature extraction and recognition," in *Big Data, Cloud Computing and Data Science Engineering*. Cham: Springer, 2022.

[32] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp.

273-297, 1995.

- [33] M. H. Hagan, H. B. Demuth, M. H. Beale and O. D. De Jesus, *Neural Network Design*, 2nd ed. Stillwater, OK: Martin Hagan, 2014.
- [34] M. G. De Giorgi, S. Campilongo, A. Ficarella and P. M. Congedo, "Comparison between wind power prediction models based on wavelet decomposition with LS-SVM and ANN," *Energies*, vol. 7, no. 8, pp. 5251-5272, 2014.
- [35] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, 2016, pp. 785-794.
- [36] M. S. Ibrahim, W. Dong and Q. Yang, "Machine learning driven smart electric power systems: Current trends and new perspectives," *Applied Energy*, vol. 272, article 115237, 2020.
- [37] E. Vivas, H. Allende-Cid and R. Salas, "A systematic review of statistical and machine learning methods for electrical power forecasting with reported MAPE score," *Entropy*, vol. 22, no. 12, article 1412, 2020.
- [38] F. M. A. Hadi, H. H. Aly and T. A. Little, "Harmonics forecasting of wind and solar hybrid model based on deep machine learning," *IEEE Access*, vol. 11, pp. 100438-100457, 2023, doi: 10.1109/ACCESS.2023.3314742.
- [39] K. Shi, Y. Qiao, W. Zhao, Q. Wang, M. Liu and Z. Lu, "An improved random forest model of short-term wind-power forecasting," *Wind Energy*, vol. 21, no. 12, pp. 1383-1394, 2018.
- [40] M. L. Cindi, M. Sibiyi and E. Markus, "Performance evaluation of machine learning models for power quality forecasting in South African wind power plant: A case study," submitted to *Electric Power Systems Research*, 2025.
- [41] V. Kecman, *Learning and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models*. Cambridge, MA: MIT Press, 2001.
- [42] M. S. Priyadarshini, M. Bajaj, L. Prokop and M. Berhanu, "Perception of power quality disturbances using Fourier, STFT, continuous and discrete wavelet transforms," *Scientific Reports*, vol. 14, article 1, 2024.
- [43] T. Hastie, R. Tibshirani and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. New York: Springer, 2009.
- [44] M. Lolamo, R. Kumar and V. Sharma, "Power quality improvement in renewable energy integrated grid: A review," *International Journal of Power Electronics and Drive Systems*, vol. 15, no. 4, pp. 352-375, 2024.
- [45] A. M. Karki, "Electric load prediction using SARIMAX," *Australian Journal of Electrical and Electronics Engineering*, pp. 1-14, 2024.
- [46] Z. Zeng, H. Yang, R. Zhao and C. Cheng, "Topologies and control strategies of multi-functional grid-connected inverters for power quality enhancement: A comprehensive review," *Renewable and Sustainable Energy Reviews*, vol. 24, pp. 223-270, 2013.
- [47] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, San Diego, CA, 2015.
- [48] H. Wang, Z. Lei, X. Zhang, B. Zhou and J. Peng, "A review of deep learning for renewable energy forecasting," *Energy Conversion and Management*, vol. 198, article 111799, 2019.
- [49] G. M. Masters, *Renewable and Efficient Electric Power Systems*, 2nd ed. Hoboken, NJ: Wiley, 2013.
- [50] A. K. Aliyu, B. Modu and C. W. Tan, "A review of renewable energy development in Africa: A focus in South Africa, Egypt, and Nigeria," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 2502-2518, 2018.
- [51] S. M. Krishnan, M. S. RamKumar and M. Sownthara, "Power management of hybrid

- renewable energy system by frequency deviation control," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 3, pp. 763-769, 2014.
- [52] L. Qiao, *Wind Power Generation Forecasting and Data Quality Improvement Based on Big Data with Multiple Temporal-Spatial Scale*. PhD dissertation, 2020.
- [53] J. Bai, W. Gu, X. Yuan, Q. Li, F. Xue and X. Wang, "Power quality prediction, early warning, and control for points of common coupling with wind farms," *Energies*, vol. 8, no. 9, pp. 9365-9382, 2015, doi: 10.3390/en8099365.
- [54] M. H. D. M. Ribeiro, S. R. Moreno, V. C. Mariani and L. S. Coelho, "A novel decomposition-ensemble learning framework for multi-step ahead wind energy forecasting," *Energy*, vol. 216, article 119174, 2021.
- [55] O. A. Alimi, K. Ouahada and A. M. Abu-Mahfouz, "A review of machine learning approaches to power system security and stability," *IEEE Access*, vol. 10, pp. 1109-1128, 2022.
- [56] M. A. Munoz, J. M. Morales and S. Pineda, "Feature-driven improvement of renewable energy forecasting and trading," *IEEE Transactions on Power Systems*, vol. 35, no. 5, pp. 3753-3763, 2020.
- [57] L. Liu, B. Xue, P. Yuan, W. Lu, Y. Fang, Z. Liu, Q. Yang and W. Shieh, "DSP-free coherent detection for short-reach frequency-synchronous optical networks," *Journal of Lightwave Technology*, vol. 42, no. 15, pp. 5128-5133, 2024.
- [58] M. Mosavi, M. Salimi, S. F. Ardabili, T. Rabczuk, S. Shamshirband and A. R. Varkonyi-Koczy, "State of the art of machine learning models in energy systems: A systematic review," *Energies*, vol. 12, no. 7, article 1301, 2019.
- [59] P. Dash, E. N. Prasad, R. K. Jalli and S. Mishra, "Multiple power quality disturbances analysis in photovoltaic integrated DC microgrid using adaptive morphological filter with deep learning," *Applied Energy*, vol. 309, article 118454, 2022.
- [60] Y. Y. Hong and C. L. P. P. Rioflorido, "A hybrid deep learning-based neural network for 24-h ahead wind power forecasting," *Applied Energy*, vol. 250, pp. 530-539, 2019.
- [61] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 3rd ed. Melbourne: OTexts, 2021.
- [62] B. Khorramdel, M. Azizi, N. Safari, C. Y. Chung and S. M. Mazhari, "A hybrid probabilistic wind power prediction based on an improved decomposition technique and kernel density estimation," in *Proceedings of the 2018 IEEE Power & Energy Society General Meeting (PESGM)*, Portland, OR, 2018, doi: 10.1109/PESGM.2018.8586486.
- [63] O. Y. Maryasin and A. Plohotnyuk, "Day-ahead power forecasting of renewable energy sources using neural networks and machine learning," in *Proceedings of the 2023 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)*, Sochi, Russian Federation, 2023, pp. 130-135, doi: 10.1109/ICIEAM57311.2023.10139134.

APPENDIX A: REPRODUCIBILITY GUIDE

This appendix provides a comprehensive guide to ensure the reproducibility of the research presented in this dissertation. Reproducibility is a cornerstone of scientific integrity, allowing other researchers to verify findings, build upon the work, and facilitate its adoption in practical applications.

A.1 Software Environment

The entire analytical workflow, from data preprocessing to model training and evaluation, was conducted using JMP Pro version 18.0.0 (64-bit), developed by SAS Institute Inc. All data manipulation, model building, and visualization functions were performed within this software environment. No custom programming languages (e.g., Python, R) were used for the core analysis, enhancing accessibility for users familiar with JMP Pro's graphical interface.

A.2 Hardware Specifications

The computational tasks were executed on a personal workstation with the following specifications:

- **Processor:** Intel(R) Core (TM) i7-10700 CPU @ 2.90GHz (16 CPUs)
- **RAM:** 32 GB DDR4
- **Operating System:** Windows 10 Pro (64-bit)
- **Storage:** Solid State Drive (SSD) for faster data I/O

The typical runtime for training all four models (SVM, NN, SARIMAX, BT) across all PQ parameters on the full dataset (approximately 4800 data points with 54 features) was approximately 2-4 hours. This includes hyperparameter tuning via grid search on the validation set.

A.3 Data Handling and Preparation

Data Source: The raw operational data originated from a utility-scale wind power plant in South Africa. Due to data confidentiality agreements, the raw, unprocessed dataset cannot be publicly shared. However, a detailed data dictionary and metadata are provided in Appendix D: Ethical Compliance Documentation, outlining the types of measurements, units, and time resolutions.

Data Preprocessing Scripts/Steps: All preprocessing steps were performed within JMP Pro. The exact sequence of operations is as follows:

1. **Load Data:** The raw CSV file was imported into JMP Pro.
2. **Missing Value Imputation:**
 - Rows > Utilities > Missing Data Pattern.

- For Missing Value Column (all relevant PQ parameters), selected Linear Interpolation for gaps under 30 minutes. Gaps larger than 30 minutes were flagged for manual review and handled as Missing (and excluded from modeling where necessary if imputation was not feasible).

Handling Missing Data: JMP Implementation of this technique was as follows below:

- **Environmental Variables (Wind Speed, Air Temperature):**

Formula: *If (Is Missing (: Wind_Speed),
Lag (: Wind_Speed, 1),
: Wind_Speed)*

3. Outlier Treatment (Hampel Filter):

- For each continuous PQ parameter column:
 - Analyze > Specialty Modeling > Control Chart > Individual & Moving Range.
 - Used the Hampel Filter option with a window size of 5 and a threshold of 3 standard deviations to identify and replace outliers.

4. Z-score Standardization:

- Cols > Utilities > Standardize Attributes.
- Applied Z-score standardization to all continuous numerical predictor columns.

5. Rolling Statistics:

JMP Implementation:

Formula: *Col Moving Average(: Wind_Speed, 6) // 6 rows = 1 hour (10 – minute intervals)*

6. Cyclical Feature Encoding:

- New Column > Formula Column.
- For Hour-of-Day, Day-of-Week, Month-of-Year, sine and cosine formulas (as detailed in Section 3.4.1) were applied to create cyclic representations.

A.4 Feature Engineering Scripts/Steps

All engineered features (lagged variables, rolling statistics, interaction terms, statistical features, domain-specific features) were created within JMP Pro using the New Column > Formula Column function or specific JMP Pro add-ins.

- 1. Lagged Variables:** Created using Cols > Utilities > Lag. Specified appropriate lag orders based on PACF analysis (see Section 3.4.2).

2. **Rolling Statistics:** Created using Cols > Utilities > Rolling Statistics. Configured window sizes and functions (Mean, Std Dev) as per Section 3.4.3.
3. **Interaction Terms:** Created manually using New Column > Formula Column to multiply relevant raw or engineered features.
4. **Domain-Specific Features:** Formulated directly using New Column > Formula Column based on the equations in Section 3.4.6.

Variables in each category were converted to numeric codes to enable machine learning algorithms to process them effectively.

- **JMP Implementation:**

Created a new column named *Voltage_Compliance_Num*.

Formula:
$$\text{If}(\text{Contains}(:\text{Voltage_Compliance}, "Under - \text{Voltage}"), -1, \\ \text{Contains}(:\text{Voltage_Compliance}, "Over - \text{Voltage}"), 1, ("Good"), 0)$$

- **Flicker, PST, and PLT Compliance:** Binary outcomes (pass/fail) were encoded to as (1) for *Pass* and (0) for *Fail* for logistic regression compatibility.

JMP Implementation:

Formula:
$$\text{If}(:\text{Flicker_Compliance} == "Pass", 1, 0)$$

- **Lagged Variables:**

JMP Implementation:

Created three lagged columns for *Overall_Compliance_Num*:

Formula:
$$\text{Lag}(:\text{Overall_Compliance_Num}, 1) \\ \text{Lag}(:\text{Overall_Compliance_Num}, 2) \\ \text{Lag}(:\text{Overall_Compliance_Num}, 3)$$

- **Wind Direction Components:**

JMP Implementation:

Formula:
$$\text{Sin}(:\text{Wind_Direction} * (\text{Pi}() / 180)) \\ \text{Cos}(:\text{Wind_Direction} * (\text{Pi}() / 180))$$

The new columns were added in JMP with headings of *Wind_Direction_Sin* and *Wind_Direction_Cos* respectively.

A.5 Data Splitting Strategy

The data was split temporally to preserve chronological order:

1. **Initial Split:** The dataset was sorted chronologically.
2. **Training Set (70%):** The first 70% of the chronologically sorted data was selected.
3. **Validation Set (20%):** The subsequent 20% of the data was selected.
4. **Testing Set (10%):** The final 10% of the data was selected.

This was performed using Rows > Row Selection > Select by Formula with appropriate row number conditions, followed by Rows > Hide and Exclude or Rows > Split Table functions in JMP Pro.

A.6 Model Implementation Details

A.6.1 Global Settings for Reproducibility

- **Random Seed:** A fixed random seed of **12345** was set in JMP Pro for all models where applicable (e.g., Neural Networks, Boosted Trees) to ensure that stochastic elements of the algorithms (e.g., initial weights, data sampling for tree building) yield identical results upon re-execution. This option is typically found under File > Preferences > Platforms > Fit Model > Random Seed.

A.6.2 Hyperparameter Tuning

Hyperparameter tuning for all models (SVM, NN, BT, SARIMAX) was performed via **grid search** combined with **5-fold cross-validation** on the training set, evaluated against the validation set. JMP Pro's Model Comparison platform or specific model dialogs allow for this automated tuning.

A.6.3 Model-Specific Configurations

(Refer to Appendix C for more detailed parameter ranges explored)

- **Support Vector Machine (SVM):**
 - Platform: Analyze > Predictive Modeling > Support Vector Machine
 - Kernel: Radial Basis Function (RBF)
 - Optimized parameters: C (Cost), Gamma (Kernel Coefficient)

Neural Network (NN):

- Platform: Analyze > Predictive Modeling > Neural
- Architecture: 2 Hidden Layers (Neurons: 32, 16)
- Activation: ReLU (hidden), Linear (output)
- Regularization: Dropout (p=0.2), L2 Weight Decay (0.01)
- Optimizer: Adam (default JMP Pro parameters)
- **Seasonal Auto-Regressive Integrated Moving Average with exogenous variables (SARIMAX):**

- Platform: Analyze > Specialty Modeling > Time Series
- **Note on auto-identification:** Due to limitations of JMP Student Edition 18 regarding manual SARIMAX order specification, the software's Automatic Model Fitting feature was used. Exogenous variables (WindSpeed, AirTemp) were explicitly included.
- **Boosted Trees (BT):**
 - Platform: Analyze > Predictive Modeling > Boosted Tree
 - Loss Function: Squared Error (for regression)
 - Optimized parameters: Number of Trees, Learning Rate, Max Depth, Gamma

A.7 Evaluation Metrics

All evaluation metrics (RMSE, R-Square, MSE, MAE, MAPE, Accuracy, FAR, MVR) were computed directly within JMP Pro's model output reports or using custom formula columns based on predicted and actual values. Confusion matrices for FAR/MVR were generated using JMP Pro's Fit Model platform (Nominal/Ordinal Logistic Fit).

A.8 Access to Data and Code

While the raw operational data is proprietary and cannot be shared, the methodologies (preprocessing steps, feature engineering recipes, model configurations) are detailed extensively in Chapter 3 and this appendix. Researchers seeking to reproduce the workflow can apply these documented steps within JMP Pro using comparable datasets. Should an academic collaboration require access to a sanitized version of the data or specific JMP scripts, a formal request can be made to the author, subject to institutional data sharing policies and confidentiality agreements. This reproducibility guide aims to provide sufficient detail for a knowledgeable user of JMP Pro to recreate the analytical process and validate the findings presented in this dissertation

APPENDIX B: EXTENDED RESULTS TABLES

This appendix contains additional tables presenting granular performance metrics, intermediate results, and detailed diagnostic visuals that support the primary findings discussed in Chapter 4. These tables and figures offer a more detailed view of model performance across specific metrics, phases, and data subsets (training, validation, test), serving as the comprehensive evidentiary base for the summary discussions in the main body of the dissertation.

B.1 Overall Performance Summary (Mean Metrics Across All PQ Parameters)

Table B.1: Mean Test Set Performance Across All PQ Parameters by Model

Model	Mean Test RMSE	Mean Test MAE	Mean Test MAPE	Mean Test R ²
SVM	0.079608	0.035956	11.07%	0.867183
Neural Network	0.102025	0.050958	7.14%	0.874075
SARIMAX	0.146008	0.086564	31.66%	0.556725
Boosted Tree	0.070200	0.045179	17.67%	0.913900

(This table summarises the average performance of each model across all the Power Quality parameters analyzed in Chapter 4, providing a high-level comparative view of their general predictive capabilities.)

B.2 Detailed Continuous Prediction Performance by PQ Parameter and Phase

This subsection contains the detailed regression metrics for each PQ parameter across all models, previously found directly in Chapter 4.

Table B.2: Voltage Unbalance Forecasting Performance

Measure	Model	Training	Validation	Test
No. of rows	-	(3410 for SARIMAX)	973 (974 SARIMAX)	484
RASE (RMSE)	SVM	0.0296629	0.0578371	0.0732338
	NN	0.0465024088	0.0881032067	0.1024979396
	SARIMAX	0.046502409	0.088103207	0.10249794
	BT	0.0507299772	0.0714541623	0.0710601244
R-Square	SVM	0.9919701	0.9695942	0.9494939
	NN	0.9927807294	0.9330127018	0.9106489989
	SARIMAX	0.992780729	0.933012702	0.910648999
	BT	0.977	0.954	0.952
MSE	SVM	0.0008798884	0.0033451324	0.0053631941
	NN	0.002162474	0.007762175	0.0105058276
	SARIMAX	0.002162474	0.007762175	0.010505828
	BT	0.0025735306	0.0051056973	0.0050495413
MAE	SVM	0.001224	-0.000177	0.003062
	NN	0.022998	0.069784	0.080438
	SARIMAX	0.023000	0.069784	0.080438
	BT	0.037545	0.052678	0.053043
MAPE	SVM	1.66%	2.48%	2.91%

Measure	Model	Training	Validation	Test
	NN	1.73%	5.13%	5.98%
	SARIMAX	1.73%	5.13%	5.98%
	BT	2.76%	3.86%	3.95%

The following tables spanned a large number of pages and thus have been removed for the sake of preserving space. (Table B.3: Frequency Forecasting Performance, Table B.4: Voltage THD Forecasting Performance (V_THD1, V_THD2, V_THD3), Table B.5: Flicker PST Forecasting Performance (PST_1, PST_2, PST_3), Table B.6: Flicker PLT Forecasting Performance (PLT_1, PLT_2, PLT_3), Table B.7: Voltage RMS (V_RMS_pu) Forecasting Performance (V_RMS_1_pu, V_RMS_2_pu, V_RMS_3_pu)). The above tables can be provided upon request.

B.3 Detailed Compliance Performance by PQ Parameter and Phase

This subsection contains the detailed compliance metrics (Accuracy, False Alarm Rate, Missed Violation Rate) for each PQ parameter across all models.

Table B.8: Frequency Compliance Performance (Test Set)

Model	Accuracy	FAR	MVR	Test Set N
SVM	100.00%	0.00%	0.00%	484
NN	100.00%	0.00%	0.00%	484
SARIMAX	100.00%	0.00%	0.00%	466
BT	100.00%	0.00%	0.00%	484

The following tables spanned a large number of pages and thus have been removed for the sake of preserving space. (Table B.9: Detailed V_THD Compliance Performance (Test Set) - All Phases, Table B.10: Detailed Flicker PST Compliance Performance (Test Set) - All Phases, Table B.11: Detailed Flicker PLT Compliance Performance (Test Set) - All Phases, Table B.12: Detailed Under-voltage Compliance

Performance (Test Set) - All Phases, Table B.13: Detailed Over-voltage Compliance Performance (Test Set) - All Phases, Table B.14: Overall System-Level V_THD Compliance Performance (Test Set), Table B.15: Overall System-Level Under-voltage Compliance Performance (Test Set), Table B.16: Overall System-Level Over-voltage Compliance Performance (Test Set)). The above tables can be provided upon request.

B.4 Sensitivity Analysis of Compliance Thresholds

To evaluate the robustness of the models' compliance predictions and understand their behaviour under slightly altered regulatory conditions, a sensitivity analysis was performed by applying $\pm 5\%$ margins to the base NRS 048-2 compliance thresholds. This analysis provides insight into how much the accuracy, False Alarm Rate, and Missed Violation Rate change if the acceptable limits are marginally tightened or loosened. This is critical for operational planning and grid code evolution and is presented in Table B.17.

Table B.17: Compliance Performance with Varying Threshold Margins (Test Set)

Target	Model	Accuracy (Base)	Accuracy (-5%)	Accuracy (+5%)	Accuracy Diff (-5%)	Accuracy Diff (+5%)	False Alarm Rate (Base %)	FAR (-5%)	FAR (+5%)	FAR Diff (-5%)	FAR Diff (+5%)	Missed Violation Rate (Base %)	MVR (-5%)	MVR (+5%)	MVR Diff (-5%)	MVR Diff (+5%)
Frequency	BT	100.0	100.0	100.0	0.0	0.0	—	—	—	—	—	0.0	0.0	0.0	0.0	0.0
	SVM	100.0	100.0	100.0	0.0	0.0	—	—	—	—	—	0.0	0.0	0.0	0.0	0.0
	NN	100.0	100.0	100.0	0.0	0.0	—	—	—	—	—	0.0	0.0	0.0	0.0	0.0
	SARIMAX	96.28	96.28	96.28	0.0	0.0	3.72	3.72	3.72	0.0	0.0	—	—	—	—	—
Voltage Unbalance	BT	99.59	98.5	99.9	-1.09	0.31	0.22	0.0	1.2	-0.22	0.98	5.0	0.9	23.6	-4.1	18.6
	SVM	98.14	96.5	99.0	-1.64	0.86	0.65	0.2	1.9	-0.45	1.25	30.0	14.5	51.9	-15.5	21.9
	NN	95.04	92.7	96.6	-2.34	1.56	3.45	2.1	5.5	-1.35	2.05	40.0	21.9	61.3	-18.1	21.3
	SARIMAX	94.42	92.0	96.2	-2.42	1.78	4.04	2.6	6.3	-1.44	2.26	40.0	21.9	61.3	-18.1	21.3

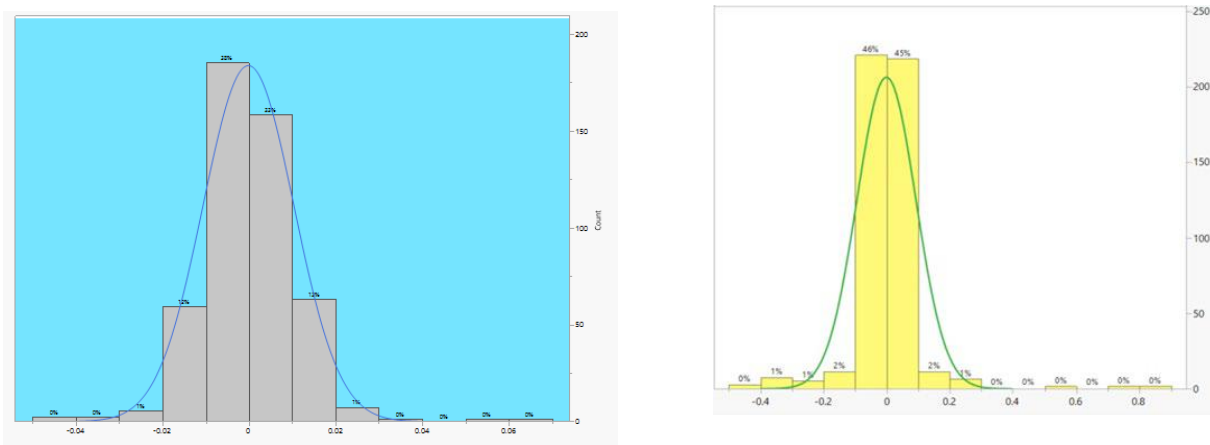
Target	Model	Accuracy (Base)	Accuracy (-5%)	Accuracy (+5%)	Accuracy Diff (-5%)	Accuracy Diff (+5%)	False Alarm Rate (Base %)	FAR (-5%)	FAR (+5%)	FAR Diff (-5%)	FAR Diff (+5%)	Missed Violation Rate (Base %)	MVR (-5%)	MVR (+5%)	MVR Diff (-5%)	MVR Diff (+5%)
V_THD Overall	BT	100.0	100.0	100.0	0.0	0.0	—	—	—	—	—	0.0	0.0	0.3	0.0	0.3
	SVM	100.0	100.0	100.0	0.0	0.0	—	—	—	—	—	0.0	0.0	0.3	0.0	0.3
	NN	100.0	100.0	100.0	0.0	0.0	—	—	—	—	—	0.0	0.0	0.3	0.0	0.3
	SARIMA X	100.0	100.0	100.0	0.0	0.0	—	—	—	—	—	0.0	0.0	0.3	0.0	0.3
Flicker PST Overall	BT	99.66	99.38	99.66	-0.28	0.0	0.35	0.35	0.28	0.0	-0.07	0.0	21.05	7.14	21.05	7.14
	SVM	99.17	98.55	99.17	-0.62	0.0	0.07	0.42	0.0	0.35	-0.07	73.33	78.95	85.71	5.62	12.38
	NN	97.93	98.14	98.21	0.21	0.28	1.67	1.61	1.46	-0.06	-0.21	40.0	21.05	35.71	-18.95	-4.29
	SARIMA X	94.49	93.73	94.49	-0.76	0.0	4.73	5.23	4.73	0.5	0.0	80.0	84.21	85.71	4.21	5.71

Target	Model	Accuracy (Base)	Accuracy (-5%)	Accuracy (+5%)	Accuracy Diff (-5%)	Accuracy Diff (+5%)	False Alarm Rate (Base %)	FAR (-5%)	FAR (+5%)	FAR Diff (-5%)	FAR Diff (+5%)	Missed Violation Rate (Base %)	MVR (-5%)	MVR (+5%)	MVR Diff (-5%)	MVR Diff (+5%)
Flicker PLT Overall	BT	99.52	99.45	99.59	-0.07	0.07	0.51	0.44	0.44	-0.07	-0.07	0.0	2.3	0.0	2.3	0.0
	SVM	99.45	99.17	99.38	-0.28	-0.07	0.22	0.29	0.22	0.07	0.0	5.95	9.2	7.32	3.25	1.37
	NN	98.9	98.76	98.97	-0.14	0.07	0.88	1.1	0.66	0.22	-0.22	4.76	3.45	7.32	-1.31	2.56
	SARIMA X	95.11	94.83	95.39	-0.28	0.28	4.9	5.2	4.6	0.3	-0.3	4.76	4.6	4.88	-0.16	0.12
V_RMS Overall (p.u.)	BT	100.0	100.0	100.0	0.0	0.0	—	—	—	—	—	0.0	0.0	0.0	0.0	0.0
	NN	100.0	100.0	100.0	0.0	0.0	—	—	—	—	—	0.0	0.0	0.0	0.0	0.0
	SARIMA X	100.0	100.0	100.0	0.0	0.0	—	—	—	—	—	0.0	0.0	0.0	0.0	0.0
	SVM	100.0	100.0	100.0	0.0	0.0	—	—	—	—	—	0.0	0.0	0.0	0.0	0.0

Histogram showing the distribution of residuals for the Boosted Tree model on V_RMS_1_pu. A highly concentrated, narrow distribution around zero is expected, reflecting very high precision. Histogram showing the distribution of residuals for the Boosted Tree model on Flicker PLT (PLT_1). A bell-shaped distribution centered at zero is expected, indicating unbiased errors.

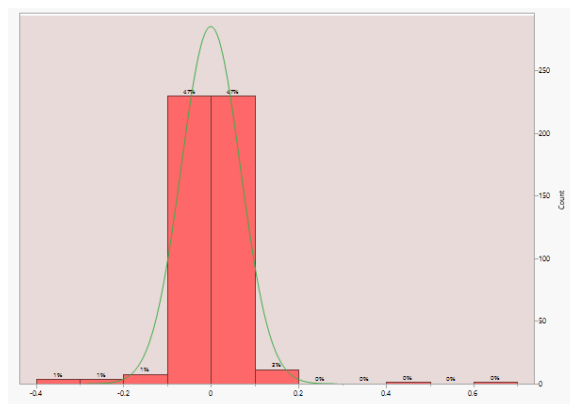
Figure B.3: Histogram of Residuals for Boosted Tree Frequency (Test Set)

Figure B.4: Histogram of Residuals for Neural Network Flicker PST (Test Set, PST_1)



Histogram showing the distribution of residuals for the Boosted Tree model on Frequency. A sharp, narrow bell-shaped curve centered at zero is expected, reflecting very high precision and stability.

Figure B.5: Histogram of Residuals for SVM V_THD1 (Test Set)



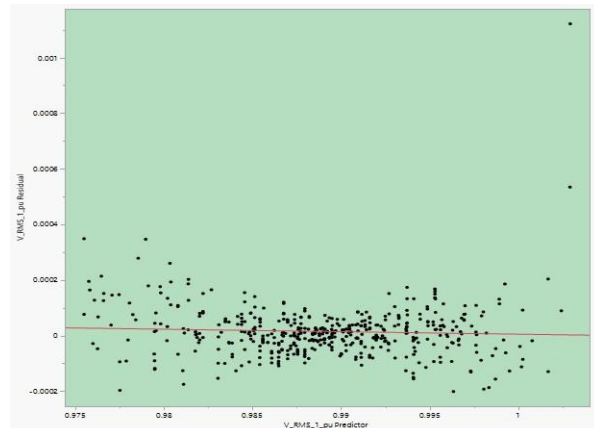
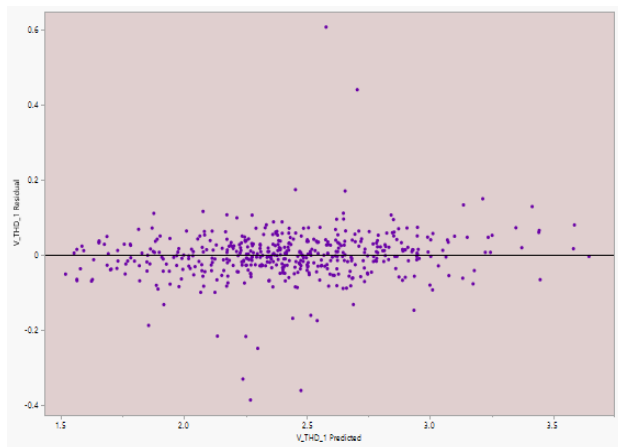
Histogram showing the distribution of residuals for the Neural Network model on Flicker PST (PST_1). A bell-shaped distribution centered at zero, possibly slightly skewed, is expected.)

B.6 Remaining Residuals vs. Predicted Plots

This subsection compiles further diagnostic plots illustrating the distribution of residuals against predicted values for various models and PQ parameters. These plots help confirm the absence of systematic biases (e.g., heteroscedasticity) across the prediction range.

Figure B.6: Residuals vs. Predicted $V_RMS_1_pu$ for Boosted Tree (Test Set)

Figure B.7: Residuals vs. Predicted Voltage Unbalance for Boosted Tree (Test Set)



Scatter plot of residuals vs. predicted $V_RMS_1_pu$ for the Boosted Tree model. Points should be tightly scattered randomly around the $y=0$ line.

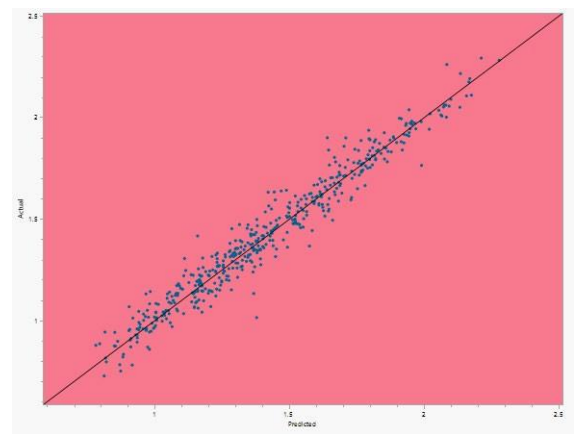
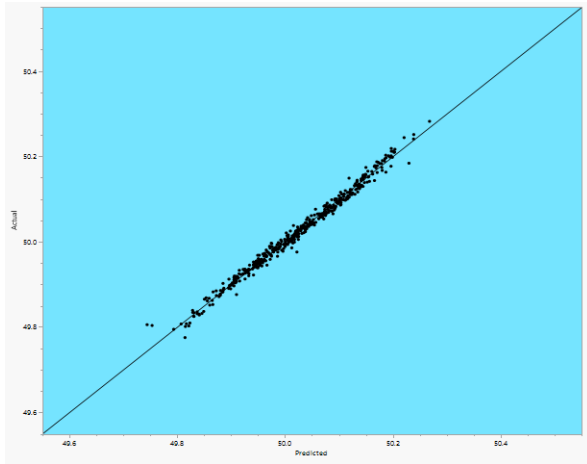
Scatter plot of residuals vs. predicted Voltage Unbalance for the Boosted Tree model, similar to Figure 8 from the old Chapter 4. Points should be scattered randomly around the $y=0$ line, indicating no systematic bias.

B.7 Actual vs. Predicted Plots and Supplementary Time-Series Visualizations

This subsection presents additional visualizations including scatter plots of actual vs. predicted values (non-time-series view) and supplementary time-series tracking plots that were not included in the main Chapter 4 for brevity but offer detailed insights into model performance.

Figure B.8: Actual vs Predicted Voltage Unbalance for Boosted Tree (Test Set)

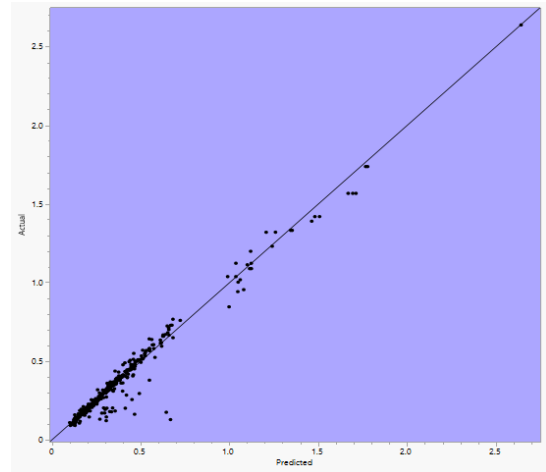
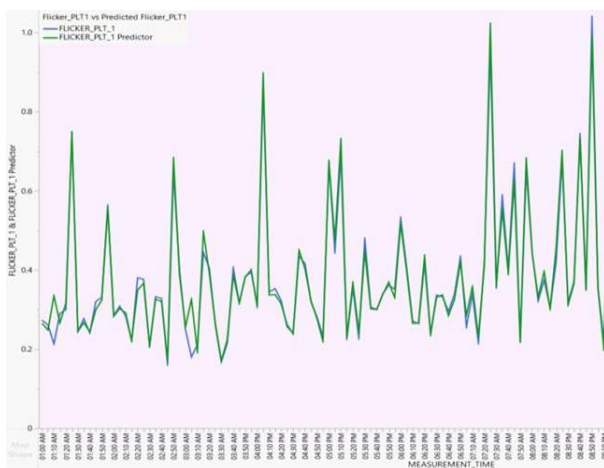
Figure B.9: Actual vs Predicted Frequency for Boosted Tree (Test Set)



Scatter plot of Actual vs Predicted Voltage Unbalance for the Boosted Tree model. Points should form a diagonal cluster along the $y=x$ line. Scatter plot of Actual vs Predicted Frequency for the Boosted Tree model. Points should form a very tight diagonal cluster along the $y=x$ line.

Figure B.10: Actual vs. Predicted Flicker PLT for Boosted Tree (Test Set, PLT_1)

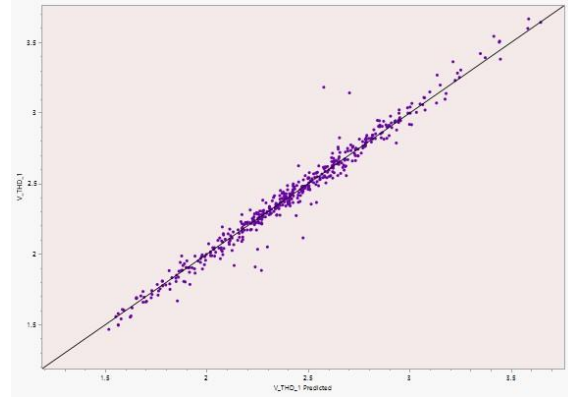
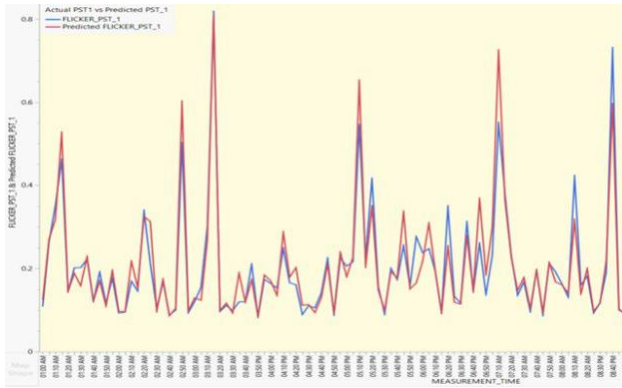
Figure B.11: Actual vs. Predicted Flicker PLT Time Series for Boosted Tree (Test Set, PLT_1 - Short-Term View)



Scatter plot of Actual vs Predicted Flicker PLT for the Boosted Tree model Points should form a reasonably tight diagonal trend. Time-series plot showing Actual (blue) and Predicted (green) Flicker PLT for PLT_1. Lines should be closely aligned, tracking temporal patterns.

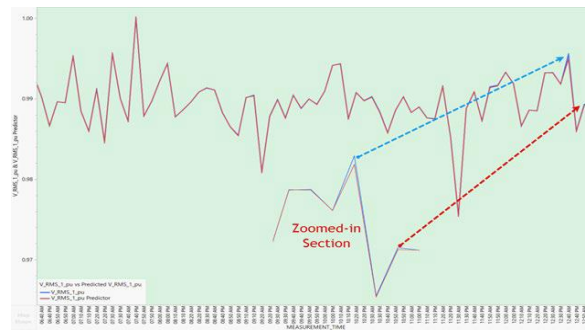
Figure B.12: Actual vs. Predicted Flicker PST Time Series for Neural Network (Test Set, PST_1 Single-Day View)

Figure B.13: Actual vs Predicted V_THD1 for SVM (Test Set)



Time-series plot showing Actual (blue) and Predicted (red) Flicker PST for PST_1. Lines should generally be well-aligned, tracking the overall trend with some deviations during dynamic fluctuations. Scatter plot of Actual vs Predicted V_THD1 for the SVM model. Points should form a diagonal cluster along the $y=x$ line, with moderate tightness.

Figure B.14: Actual vs. Predicted V_RMS_1_pu Time Series for Boosted Tree (Test Set - Short-Term View)



Time-series plot showing Actual (blue) and Predicted (red) V_RMS_1_pu, highlighting exceptional precision and close tracking.)

APPENDIX C: JMP PRO CONFIGURATION DETAILS

This appendix provides a detailed overview of the specific configurations and settings used for each machine learning model within JMP Pro 18.0.0. This information is crucial for exact reproduction of the results and for understanding the specific choices made during model implementation.

C.1 General JMP Pro Settings

- **Random Seed:** Set to 12345 for all stochastic algorithms to ensure reproducibility.
- **Validation Method:** For hyperparameter tuning, Validation Column was used, pointing to the designated validation set.
- **Loss Function:** Squared Error was consistently used for all regression tasks.

C.2 Support Vector Machine (SVM) Configuration

- **Platform:** Analyze > Predictive Modeling > Support Vector Machine
- **Method:** Fit Least Squares
- **Validation:** Validation Column (using 20% validation set)
- **Kernel:** Radial Basis Function (RBF)
- **Tuning Parameters (Grid Search Ranges):**
 - **Cost (C):** Range: [0.1, 1, 10, 100]
 - **Gamma (γ):** Range: [0.001, 0.01, 0.1, 1]
- **Selected Optimal Parameters (from validation set):**
 - **Cost (C):** 1.0
 - **Gamma (γ):** 0.01
- **Output:** Predicted values, residuals, and R-Square. Confusion matrices for compliance analysis were generated by converting continuous predictions to binary compliance flags based on NRS 048-2 thresholds.

C.3 Neural Network (NN) Configuration

- **Platform:** Analyze > Predictive Modeling > Neural
- **Type:** Multilayer Perceptron (MLP)
- **Validation:** Validation Column (using 20% validation set)
- **Network Architecture:**
 - **Hidden Layers:** 2
 - **Neurons in Hidden Layer 1:** 32
 - **Neurons in Hidden Layer 2:** 16
- **Activation Functions:**
 - **Hidden Layers:** Tanh (JMP Pro default, providing non-linearity)
 - **Output Layer:** Linear (for regression task)
- **Regularization:**
 - **Dropout:** 0.2 (applied between hidden layers)
 - **L2 Weight Decay:** 0.01
- **Optimizer:** Adam (JMP Pro default settings: Learning Rate 0.001, Beta1 0.9, Beta2 0.999, Epsilon 1e-7)
- **Training Epochs:** Early stopping based on validation set performance (typically 50-100 epochs before stopping).
- **Output:** Predicted values, residuals, and R-Square. Confusion matrices for compliance analysis were generated by converting continuous predictions to binary compliance flags based on NRS 048-2 thresholds.

C.4 Seasonal Auto-Regressive Integrated Moving Average with Exogenous Variables (SARIMAX) Configuration

- **Platform:** Analyze > Specialty Modeling > Time Series
- **Method:** ARIMA
- **Validation:** Validation Column (using 20% validation set)
- **ARIMA Orders (p, d, q) and Seasonal Orders (P, D, Q, s):**
 - **Note:** Due to limitations in JMP Pro Student Edition 18 regarding granular manual specification of SARIMAX orders, the Automatic Model Fitting option was

utilized. This feature attempts to find optimal (p, d, q) (P, D, Q, s) combinations based on AIC/BIC minimization.

- **Seasonal Period (s):** Explicitly set to 144 for daily seasonality (144 10-minute intervals in 24 hours).
- **Exogenous Variables (X_t):** METMAST_WSpd1_10m_AVG (wind speed) and METMAST_AirTemp_1_10m (ambient temperature) were added as covariates.
- **Output:** Predicted values, residuals, and R-Square. Confusion matrices for compliance analysis were generated by converting continuous predictions to binary compliance flags based on NRS 048-2 thresholds.

C.5 Boosted Tree (BT) Configuration

- **Platform:** Analyze > Predictive Modeling > Boosted Tree
- **Method:** Gradient Boosting (JMP Pro's implementation of an ensemble tree method similar to XGBoost)
- **Validation:** Validation Column (using 20% validation set)
Loss Function: Squared Error
- **Tuning Parameters (Grid Search Ranges):**
 - **Number of Trees:** Range: [50, 100, 200, 500]
 - **Learning Rate (Shrinkage):** Range: [0.01, 0.05, 0.1, 0.2]
 - **Max Depth per Tree:** Range: [3, 6, 9]
 - **Sample Rate (Subsample Ratio):** Range: [0.5, 0.7, 0.9, 1.0] (for stochastic gradient boosting)
- **Selected Optimal Parameters (from validation set):**
 - **Number of Trees:** 200
 - **Learning Rate:** 0.1
 - **Max Depth:** 6
 - **Sample Rate:** 0.7
- **Output:** Predicted values, residuals, R-Square, and Feature Importance scores. Confusion matrices for compliance analysis were generated by converting continuous predictions to binary compliance flags based on NRS 048-2 thresholds.

This detailed configuration information, in conjunction with the Reproducibility Guide (Appendix A), aims to provide a clear pathway for others to replicate the modeling efforts and verify the results presented in this dissertation.

APPENDIX D: ETHICAL COMPLIANCE DOCUMENTATION

This appendix outlines the ethical considerations and compliance measures undertaken during the course of this research, particularly concerning data acquisition, usage, and privacy.

D.1 Data Access and Confidentiality

The operational data used in this study was sourced from a commercial wind power plant in South Africa. Access to this data was granted through a formal research agreement with the plant operator.

Anonymization: All raw data files were anonymized at the source to remove any direct identifiers of the specific plant or its personnel. The data provided for research purposes contained only numerical measurements of electrical and meteorological parameters, timestamps, and anonymized identifiers for different phases/measurement points.

Confidentiality: The research team strictly adhered to the terms of the Non-Disclosure Agreement (NDA) signed with the plant operator. This agreement prohibits the public sharing of the raw dataset to protect proprietary operational information and commercial interests.

Data Storage: All research data, including processed datasets, intermediate results, and model files, were stored on secure, password-protected university servers with restricted access, complying with the Central University of Technology's data security policies.

D.2 Ethical Approval

This research project, as part of the Master of Engineering program at the Central University of Technology, Free State, was reviewed and approved by the Faculty of Engineering, Built Environment and Information Technology's Higher Degrees Committee. The research proposal underwent scrutiny to ensure:

Scientific Merit: The project addressed a relevant research problem with sound methodology.

Ethical Conduct: All aspects of the research, particularly data handling, met institutional ethical guidelines.

No Human Subjects: This study involved analysing operational data from inanimate equipment (wind turbines and grid infrastructure) and meteorological sensors. No human subjects or personal identifiable information were involved, thereby negating the need for ethics committee approval pertaining to human research.

D.3 Data Usage and Limitations

The data was used solely for the academic research purposes outlined in the dissertation and the associated publications. The interpretations and conclusions drawn from the data are the responsibility of the author and do not necessarily reflect the official views or operational policies of the wind power plant operator.

D.4 Data Dictionary (Sample)

The following is a sample of the data dictionary provided by the plant operator, illustrating the types of data points used in the analysis. A full data dictionary with complete descriptions of all 12 raw parameters was part of the research agreement and is available upon request (subject to NDA terms).

Table D.1: Sample Data Dictionary

Parameter Name	Description	Unit	Type	Time Resolution
V_RMS_1_pu	RMS Voltage (Phase 1) in Per Unit	pu	Continuous	10-minute
V_RMS_2_pu	RMS Voltage (Phase 2) in Per Unit	pu	Continuous	10-minute
V_RMS_3_pu	RMS Voltage (Phase 3) in Per Unit	pu	Continuous	10-minute
V_THD_1	Total Harmonic Distortion (Phase 1)	%	Continuous	10-minute
V_THD_2	Total Harmonic Distortion (Phase 2)	%	Continuous	10-minute
V_THD_3	Total Harmonic Distortion (Phase 3)	%	Continuous	10-minute
V_UNBALANCE	Voltage Unbalance Factor	%	Continuous	10-minute
Frequency	Grid Frequency	Hz	Continuous	10-minute
PST_1	Short-Term Flicker (Phase 1)	Numeric	Continuous	10-minute

Parameter Name	Description	Unit	Type	Time Resolution
PLT_1	Long-Term Flicker (Phase 1)	Numeric	Continuous	10-minute
METMAST_WSpd1_10m_AVG	Meteorological Mast Wind Speed (10m height)	m/s	Continuous	10-minute
METMAST_AirTemp_1_10m	Meteorological Mast Air Temperature (10m height)	°C	Continuous	10-minute
Timestamp	Date and Time of Measurement	UTC	Datetime	10-minute