



A DEEP LEARNING APPROACH FOR DECISION SUPPORT IN POWER MANAGEMENT OF FPSOS WITH WIND ENERGY INTEGRATION

Marcos Vinicios Lima dos Santos

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Elétrica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Elétrica.

Orientador: Robson Francisco da Silva Dias

Rio de Janeiro
Novembro de 2025

A DEEP LEARNING APPROACH FOR DECISION SUPPORT IN POWER
MANAGEMENT OF FPSOS WITH WIND ENERGY INTEGRATION

Marcos Vinicios Lima dos Santos

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO
ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE
ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO
PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU
DE MESTRE EM CIÊNCIAS EM ENGENHARIA ELÉTRICA.

Orientador: Robson Francisco da Silva Dias

Aprovada por:

RIO DE JANEIRO, RJ – BRASIL
NOVEMBRO DE 2025

Lima dos Santos, Marcos Vinicios

A Deep Learning Approach for Decision Support
in Power Management of FPSOs with Wind Energy
Integration/Marcos Vinicios Lima dos Santos. – Rio de
Janeiro: UFRJ/COPPE, 2025.

XVIII, 120 p.: il.; 29, 7cm.

Orientador: Robson Francisco da Silva Dias

Dissertação (mestrado) – UFRJ/COPPE/Programa de
Engenharia Elétrica, 2025.

Referências Bibliográficas: p. 107 – 113.

1. Offshore Wind Power.
 2. FPSO Power Systems.
 3. Renewable Energy Integration.
 4. Deep Learning.
 5. Time Series Forecasting.
 6. Probabilistic Forecasting.
 7. Hybrid Forecasting Models.
 8. Decision Support Systems.
 9. Explainable AI (XAI).
 10. Retrieval-Augmented Generation (RAG).
- I. Francisco da Silva Dias, Robson. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

*“Where there’s a will,
There’s always a way.”*

– Bob Marley

Acknowledgments

This dissertation represents the culmination of a journey made possible by the unwavering support of many remarkable individuals. It is with deep gratitude that I acknowledge those who have been instrumental in this achievement.

I dedicate this work to my **family**—my mother **Luciana**, my father **José**, and my brother **Victor**—whose unconditional support and belief have been my foundation through every challenge.

To my beloved **Carla**, my partner since my undergraduate years, whose understanding, patience, and unwavering support have been my greatest source of motivation, especially during the most difficult days. Every effort has been for her and for the future we are building together.

I am profoundly grateful to my advisor, **Professor Robson Dias**, for the invitation to pursue this master's degree and for the extraordinary opportunity to complete my studies at **COPPE/UFRJ**, the largest engineering research center in Latin America. His guidance and mentorship have been invaluable to this work.

In the professional realm, I owe immense gratitude to **Thales Rodegheri** and **Joel Viana**, my managers at **Radix Engenharia**, whose trust and confidence enabled me to grow professionally while pursuing this academic endeavor. To **Ciciane Chiovatto** at **SBM Offshore**, whose exemplary leadership and the opportunity to develop digital applications that directly impact real-world offshore operations have enriched both this research and my professional experience. To **Nilson Almeida**, whose inspiring conversations and example of engineering excellence and integrity have been a guiding light throughout this journey.

I am also deeply thankful to my fellow master's students. To **Luan Lopes**, whose selfless assistance throughout the coursework has been a testament to true solidarity and friendship. To **Paolo** and **João Amaral**, whose survival tips and companionship made this journey more bearable.

Finally, I acknowledge the countless others—professors, colleagues, friends, and mentors—who contributed to this work. This achievement is a reflection of their collective support, love, and wisdom.

Thank you.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

A DEEP LEARNING APPROACH FOR DECISION SUPPORT IN POWER MANAGEMENT OF FPSOS WITH WIND ENERGY INTEGRATION

Marcos Vinicios Lima dos Santos

Novembro/2025

Orientador: Robson Francisco da Silva Dias

Programa: Engenharia Elétrica

The integration of offshore wind power into Floating Production Storage and Offloading (FPSO) units is a critical decarbonization strategy, but its intermittency poses significant power system stability challenges. This dissertation addresses this by developing a Deep Learning framework for multi-step, probabilistic wind power forecasting to enable proactive management of backup gas turbines.

A comprehensive analysis evaluates 17 forecasting models, from classical machine learning to advanced deep learning and hybrid methods. This work introduces a novel hybrid model, the **CEEMDAN-EWT-TFT Aggregator**, which combines signal decomposition with a Temporal Fusion Transformer (TFT). The proposed method demonstrates superior predictive accuracy, achieving 1.79% MAPE at t+10 minutes—a 61% error reduction over benchmark models.

The framework's value is demonstrated through an interactive Decision Support Dashboard that translates probabilistic forecasts (P10, P50, P90) into an intuitive Operational Risk Index (ORI). A crucial contribution is a "P10 Safety Factor," a risk calibration tool that adjusts the model's pessimistic forecast. Simulations on the test dataset show this calibration enables the proposed model to **successfully avoid all encountered blackout events**, achieving the lowest total operational cost. The system is further enhanced by a dual-level **Retrieval-Augmented Generation (RAG)** framework with proactive and interactive AI agents to improve decision quality. The combined results confirm the proposed, calibrated framework provides a robust and economically optimal solution for managing hybrid power systems on offshore platforms.

Contents

Acknowledgments	v
List of Figures	xi
List of Tables	xvi
1 Introduction	1
1.1 Decarbonization of Offshore Assets and Wind Energy Integration	1
1.2 Motivation	3
1.2.1 The Physical Challenge: Intermittency and System Stability	3
1.2.2 The Operational and Economic Challenge	4
1.2.3 The Methodological Challenge: Advanced Forecasting Techniques	6
1.2.4 The Decision Support and Explainability Challenge	6
1.3 Objectives	9
1.3.1 Main Contributions	10
1.4 Work Structure	11
2 Theoretical Foundations	12
2.1 Offshore Wind Power Generation	12
2.1.1 Wind Turbine Power Curve and Characteristics	12
2.1.2 Challenges of Offshore Wind Forecasting	14
2.2 Power Systems in FPSOs	17
2.2.1 Conventional Power Generation	17
2.2.2 Operational Risks: Blackouts and Load Shedding	18
2.2.3 Hybrid Power Systems: Challenges and Opportunities	20
2.3 The Need for Integrated Decision Support	21
3 Literature Review	23
3.1 Wind Power Forecasting Techniques	23
3.1.1 Classical and Statistical Methods	23
3.1.2 Machine Learning and Deep Learning Approaches	24

3.1.3	Signal Decomposition Techniques in Time Series Forecasting	24
3.1.4	Probabilistic Forecasting and Uncertainty Quantification	26
3.1.5	Evaluation Metrics for Forecasting Performance	27
3.1.6	Forecast Horizons in Operational Contexts	29
3.1.7	Comparative Analysis of Forecasting Methods	29
3.2	Power Management in Offshore Systems and Microgrids	31
3.2.1	Conventional and Hybrid Power Systems in FPSOs	31
3.2.2	Multi-Objective Optimization in Energy Management	32
3.2.3	Decision Support Systems and Explainability	33
3.2.4	Large Language Models in Industrial Decision-Making	34
3.3	Related Work on Wind-Integrated Offshore Platforms	35
3.4	Research Gaps and Positioning of This Work	35
4	Deep Learning Techniques and Methodologies	37
4.1	Supervised Learning Framework for Time Series	37
4.2	Deep Learning Architectures for Sequential Data	38
4.2.1	Multi Layer Perceptron (MLP)	38
4.2.2	Recurrent Neural Network (RNN)	39
4.2.3	Gated Recurrent Architectures	40
4.2.4	Convolutional Neural Network (CNN) for Time Series	42
4.2.5	Attention Mechanism and Transformer Models	43
4.2.6	Temporal Fusion Transformer (TFT)	44
4.3	Signal Decomposition Techniques	46
4.3.1	Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN)	47
4.3.2	Empirical Wavelet Transform (EWT)	48
4.3.3	Hybrid CEEMDAN-EWT Decomposition Strategy	49
4.4	Ensemble and Aggregation Strategies	50
4.4.1	Independent Forecasting with Summation	51
4.4.2	Unified Multivariate Forecasting (TFT as Aggregator)	51
4.5	Probabilistic Forecasting Framework	52
4.5.1	Quantile Regression and Pinball Loss Function	52
4.6	Model Training and Optimization	53
4.6.1	Data Preprocessing and Normalization	53
4.6.2	Optimization Algorithms	54
4.6.3	Regularization Strategies for Generalization	55
4.6.4	Hyperparameter Optimization via Bayesian Optimization (Optuna)	57
4.6.5	Robust Validation with Time Series Cross-Validation	57

4.7	Performance Metrics for Model Evaluation	59
4.7.1	Deterministic Forecast Metrics	59
4.7.2	Probabilistic Forecast Metrics	60
5	Proposed Methodology and Decision Support Framework	61
5.1	Dataset and Preprocessing Pipeline	61
5.1.1	Data Source and Characteristics	62
5.1.2	Data Preprocessing and Structuring	62
5.2	Forecasting Models: A Methodological Funnel	64
5.2.1	Signal Decomposition as a Core Strategy	64
5.2.2	A Funneling Approach to Model Selection	66
5.3	The Decision Support System Framework	69
5.3.1	Operational Risk Index (ORI)	70
5.3.2	Techno-Economic Analysis Module	72
5.3.3	AI Copilot for Operator Support	74
5.3.4	Integration and Workflow	77
5.3.5	Model Comparison Mode	79
5.4	Data and Code Availability	80
5.4.1	Computational Environment	81
6	Results and Discussion	82
6.1	Performance Evaluation of Forecasting Models	82
6.1.1	Comparative Analysis of Benchmark Models	83
6.1.2	Performance of Hybrid Decomposition Models	85
6.1.3	Analysis of the Statistically Superior Model: CEEMDAN-EWT-LSTM	88
6.1.4	Analysis and Justification of the Proposed Model	88
6.1.5	Model Interpretability and Qualitative Analysis	93
6.2	Operational and Economic Impact Analysis	95
6.2.1	Impact of Risk Calibration via Safety Factor	95
6.2.2	Comparative Analysis of Operational Costs	96
6.2.3	Performance vs. Cost Trade-off Analysis	96
6.3	Analysis of the Decision Support Dashboard	97
6.3.1	Dashboard Interface and Monitoring Simulation	97
6.3.2	Strategic Parameter Optimization	99
6.3.3	LLM-Powered Operational Intelligence	99
6.4	Discussion of Findings	101
6.4.1	The Impact of Signal Decomposition	101
6.4.2	Architectural Complexity and the Proposed Model	101
6.4.3	Limitations and Future Work	102

7 Conclusion	104
7.1 Summary of Contributions	104
7.2 Concluding Remarks	105
7.3 Future Works	105
References	107
A Model Hyperparameters	114
A.1 Benchmark Models (Optimized)	114
A.1.1 Support Vector Regression (SVR)	114
A.1.2 Random Forest	115
A.1.3 Artificial Neural Network (ANN)	115
A.1.4 Long Short-Term Memory (LSTM)	115
A.1.5 Bidirectional LSTM (Bi-LSTM)	116
A.1.6 CNN-Bi-LSTM	116
A.1.7 CNN-Bi-LSTM with Attention	116
A.1.8 Transformer	117
A.1.9 Temporal Fusion Transformer (TFT)	117
A.2 Hybrid Decomposition Models	117
A.2.1 CEEMDAN-EWT-TFT Aggregator (Proposed Model)	117
A.2.2 Other Hybrid Models (Fixed Hyperparameters)	118
A.3 Optimization Methodology	119
A.4 Key Insights from Hyperparameter Optimization	120

List of Figures

1.1	Conceptual illustration of an FPSO unit integrated with a floating offshore wind farm, showing the hybrid power system architecture with renewable generation, backup gas turbines, and common AC bus (adapted from (1)).	2
1.2	10-min averaged wind power time series for 2 months from dataset (2).	3
1.3	Typical frequency response following the tripping of a generator (adapted from (3)).	4
1.4	Comparison between deterministic (dashed line) and probabilistic (shaded regions) wind power forecasting. When the 10th percentile falls below the critical threshold (red line), the ORI triggers a high-risk alert for proactive generator dispatch.	7
1.5	High-level overview of the proposed decision support framework architecture, illustrating the sequential integration of its four key components. The RAG-Powered Copilot (Pillar 4) features a dual-level architecture with both proactive and reactive agents.	8
2.1	Theoretical power curve of a wind turbine showing four operational regions: (I) Cut-in, (II) MPPT, (III) Power Limitation, and (IV) Cut-out (adapted from (4)).	13
2.2	Simplified schematic of a Type 4 WECS with full-scale back-to-back converter (adapted from (4)).	15
2.3	CEEMDAN decomposition of a wind power signal (adapted from (5)). .	16
2.4	Simulated grid frequency response in a 60 Hz isolated offshore system following the sudden loss of a main generation unit. (a) Without load shedding, the frequency drops but stabilizes at the cost of putting the remaining generator into a critical and unsustainable overload condition. (b) With a proactive load shedding scheme, the generator overload is avoided and the frequency is safely restored to its nominal value. This demonstrates the critical role of automated protection in preventing equipment failure and subsequent blackouts (adapted from (6)).	19

3.1	Temporal Fusion Transformer architecture showing variable selection networks, temporal processing layers, and multi-horizon prediction capabilities (adapted from (7)).	25
3.2	CEEMDAN-LSTM hybrid framework with separate predictors for each decomposed component (adapted from (5)).	26
3.3	Schematic of the LEOGO offshore platform energy system with integrated wind turbines (adapted from (8)).	31
4.1	Sliding window technique for time series forecasting. Look-back window (red) provides input features, forecast horizon (blue) provides target labels (adapted from (9)).	38
4.2	Basic architecture of a Multi-Layer Perceptron (MLP) with an input layer, two hidden layers, and an output layer. Information flows forward through weighted connections.	39
4.3	The unrolled structure of a Recurrent Neural Network (RNN), showing how the hidden state \mathbf{h}_t is passed from one time step to the next, creating a temporal dependency chain.	40
4.4	Detailed architecture of a Long Short-Term Memory (LSTM) cell, illustrating the forget gate (\mathbf{f}_t), input gate (\mathbf{i}_t), and output gate (\mathbf{o}_t) that regulate the information flow through the cell state (\mathbf{c}_t) and hidden state (\mathbf{h}_t).	41
4.5	Architecture of a Gated Recurrent Unit (GRU) cell, a simpler alternative to the LSTM cell featuring update gate (\mathbf{z}_t) and reset gate (\mathbf{r}_t) mechanisms.	42
4.6	Illustration of a 1D convolution operation on a time series. A filter (kernel) of size k slides across the input sequence to produce a feature map that captures local temporal patterns.	43
4.7	Scaled dot-product attention mechanism showing Queries, Keys, and Values computation with attention score weighting and Transformer architecture. (adapted from (10)).	44
4.8	Complete architecture of the Temporal Fusion Transformer (TFT), showing Variable Selection Networks, Gated Residual Networks, temporal self-attention layers, and multi-quantile output heads (adapted from (11)).	46
4.9	Flowchart of the CEEMDAN algorithm, illustrating the iterative process of extracting Intrinsic Mode Functions (IMFs) by adding adaptive noise at each stage (adapted from: (5)).	48

4.10	Hybrid CEEMDAN-EWT decomposition framework with entropy-based IMF separation and secondary EWT decomposition (adapted from: (12)).	50
4.11	The Pinball Loss function for different quantiles ($\tau = 0.1, 0.5, 0.9$). The asymmetric penalty guides the model to learn the conditional quantile of the target variable, with different slopes for over-prediction and under-prediction errors.	53
4.12	Comparison between the training cost of different optimizers on the CIFAR-10 benchmark dataset (adapted from: (13)).	55
4.13	Illustration of the Dropout technique. Left: a standard fully connected network. Right: the same network during a training step with dropout applied, where some neurons (crossed out) are temporarily deactivated.	56
4.14	Early stopping illustration showing training stopped at minimum validation loss to prevent overfitting.	57
4.15	Illustration of Time Series Cross-Validation. The data is split into multiple folds, where each fold's training set always precedes its validation set, preserving the temporal order and preventing data leakage.	59
5.1	Example of the raw active power time series from the selected wind turbine (44.csv), illustrating the characteristic volatility and intermittency during 3 months.	62
5.2	Detailed illustration of the sliding window technique. The process transforms a continuous time series into a set of discrete input-output pairs $(\mathbf{x}_t, \mathbf{y}_t)$ suitable for supervised learning. The look-back period $L = 6$ corresponds to 60 minutes of history, and the forecast horizon $H = 3$ corresponds to predictions at $t+10$, $t+20$, and $t+30$ minutes.	64
5.3	CEEMDAN-EWT decomposition process showing IMF extraction and EWT-based denoising of high-frequency components.	65
5.4	"Divide and Conquer" framework where decomposed IMFs are forecasted separately and aggregated for final prediction.	67
5.5	CEEMDAN-EWT-TFT Aggregator architecture using decomposed IMFs as multivariate features for a single TFT model.	69
5.6	Flowchart of the logic for the Operational Risk Index (ORI), translating quantile forecasts into actionable risk levels. The decision tree depends on the selected trigger horizon (h_{trigger}) and compares the P10 forecasts against the operational threshold.	72
5.7	Strategic parameter optimizer heatmap showing operational cost versus threshold and trigger horizon parameters.	75

5.8	Architecture of the Interactive AI Copilot. The agent combines Retrieval-Augmented Generation (RAG) with a Tool-Using framework, allowing it to either answer questions directly or execute specific actions (Tools) based on the operator’s request.	78
5.9	Model Comparison Mode dashboard showing performance metrics, cost analysis, and time series plots with best values highlighted.	80
6.1	Comparative performance evaluation of benchmark models across the three forecast horizons. Each chart displays R^2 on the left axis and error metrics (MAE, RMSE, and scaled MAPE) on the right axis.	84
6.2	Comparative performance evaluation of hybrid decomposition models across the three forecast horizons. The top-performing models are grouped on the left for easier comparison.	87
6.3	Comparison of native 80% prediction intervals ($t+20$ min). The LSTM model (A) produces a wider, more conservative band, while the proposed TFT model (B) yields a narrower, more confident interval.	89
6.4	CEEMDAN-EWT-TFT predictions for $t+10$ horizon with 80% prediction interval.	90
6.5	CEEMDAN-EWT-TFT predictions for $t+20$ horizon with 80% prediction interval.	91
6.6	CEEMDAN-EWT-TFT predictions for $t+30$ horizon with 80% prediction interval.	91
6.7	Regression plots showing actual vs. predicted power for the proposed CEEMDAN-EWT-TFT model across the three forecast horizons.	92
6.8	3D visualization of the CEEMDAN-EWT decomposition, showing the separation of the wind power signal into multiple IMF components and a final residual.	93
6.9	Temporal attention analysis for the proposed model during a critical ramp-down scenario.	94
6.10	Effect of applying a 0.65 safety factor. The lowered confidence interval makes the model more conservative, increasing the frequency of high-risk alerts (red zones).	95
6.11	Trade-off analysis plotting average forecast error (MAPE) against total simulated operational cost for each model.	97
6.12	The Decision Support Dashboard interface.	98
6.13	Strategic parameter optimizer heatmap showing total cost across parameter space.	99
6.14	The Proactive Surveillance Agent in action, demonstrating its ability to issue both critical alerts and routine trend summaries.	100

6.15 Demonstration of the Interactive Agent’s tool-using workflow. The operator delegates a complex, multi-step task through natural language, which the agent autonomously executes. 103

List of Tables

1.1	Operational justification for multi-horizon forecasting strategy.	5
1.2	Operational cost parameters for FPSO power management ¹	5
3.1	Comparative analysis of wind power forecasting methods for FPSO application.	30
6.1	Performance metrics of benchmark models across three forecast horizons after Optuna optimization. The best value in each metric column is highlighted in bold.	83
6.2	Performance metrics of hybrid decomposition models. The best value in each metric column is highlighted in bold. Note that only the proposed TFT-Balanced model was subjected to hyperparameter optimization.	86
6.3	Comparative performance of the top models.	89
6.4	Operational impact of applying a 0.65 safety factor to the P10 forecast. .	96
6.5	Simulated operational cost comparison across different models. The proposed CEEMDAN-EWT-TFT model demonstrates the lowest total cost.	96
A.1	Optimized hyperparameters for SVR.	114
A.2	Optimized hyperparameters for Random Forest.	115
A.3	Optimized hyperparameters for ANN.	115
A.4	Optimized hyperparameters for LSTM.	115
A.5	Optimized hyperparameters for Bi-LSTM.	116
A.6	Optimized hyperparameters for CNN-Bi-LSTM.	116
A.7	Optimized hyperparameters for CNN-Bi-LSTM+Attention.	116
A.8	Optimized hyperparameters for Transformer.	117
A.9	Optimized hyperparameters for the benchmark TFT model.	117
A.10	Optimized hyperparameters for the proposed CEEMDAN-EWT-TFT Aggregator.	118
A.11	Fixed hyperparameters for non-optimized hybrid models.	119

Abbreviations

Adam Adaptive Moment Estimation

ARIMA AutoRegressive Integrated Moving Average

Bi-LSTM Bidirectional Long Short-Term Memory

CEEMDAN Complete Ensemble Empirical Mode Decomposition with Adaptive Noise

CNN Convolutional Neural Network

DL Deep Learning

EMD Empirical Mode Decomposition

EWT Empirical Wavelet Transform

FPSO Floating Production Storage and Offloading

GRN Gated Residual Network

GRU Gated Recurrent Unit

IMF Intrinsic Mode Function

LEOGO Low Emission Oil and Gas Open

LLM Large Language Model

LSTM Long Short-Term Memory

MAE Mean Absolute Error

MAPE Mean Absolute Percentage Error

ML Machine Learning

MLP Multi Layer Perceptron

NRMSE Normalized Root Mean Squared Error

O&G Oil and Gas

ORI Operational Risk Index

RAG Retrieval-Augmented Generation

RMSE Root Mean Squared Error

RNN Recurrent Neural Network

RoCoF Rate of Change of Frequency

sMAPE Symmetric Mean Absolute Percentage Error

SVR Support Vector Regression

TFT Temporal Fusion Transformer

VFD Variable Frequency Drive

VSN Variable Selection Network

WECS Wind Energy Conversion System

XAI Explainable Artificial Intelligence

Chapter 1

Introduction

1.1 Decarbonization of Offshore Assets and Wind Energy Integration

The global energy sector is navigating a period of unprecedented transformation, driven by the dual imperatives of meeting a relentlessly growing global energy demand and addressing the urgent threat of climate change (12). This has catalyzed a systemic shift away from fossil fuels towards a more sustainable energy matrix, with renewable energy sources transitioning from niche alternatives to central pillars of future energy systems (14). Within this global movement, wind energy has established itself as a cornerstone technology, characterized by its increasing cost-competitiveness and significant potential for large-scale deployment (12; 15). The offshore environment, in particular, offers a compelling frontier for this expansion. Wind resources over the open sea are generally stronger and more consistent than their onshore counterparts, leading to higher capacity factors and more reliable power generation, making offshore wind a key enabler for a low-carbon future (4; 16).

This potential has not gone unnoticed by traditionally carbon-intensive industries. The offshore Oil and Gas (O&G) sector, which accounts for a significant portion of industrial emissions, is under immense pressure from regulatory bodies and stakeholders to decarbonize its operations (3). FPSO units, the technological workhorses of deep-water hydrocarbon extraction, are effectively large, isolated industrial microgrids. Their entire operational power demand—from large compressors and injection pumps to utility systems—is typically met by on-board gas turbines fueled by gas extracted on-site (4; 8). While operationally convenient, this reliance on fossil fuels makes FPSOs significant point sources of greenhouse gas emissions (16). In Brazil's Santos Basin, for instance, where numerous FPSOs operate in deep-water pre-salt fields, the potential for emissions reduction through renewable integration is particularly significant (16).

A highly synergistic solution has emerged from the convergence of these trends: the direct electrification of FPSO platforms using co-located floating offshore wind farms. This "Power-from-Wind" concept leverages the geographical proximity of many O&G fields to prime offshore wind sites, creating a hybrid power system that promises substantial environmental and economic benefits (17). The primary advantage is a significant reduction in CO₂ emissions, with studies indicating potential reductions ranging from 23% to over 76%, depending on the installed wind capacity (16). This, in turn, reduces operational expenditure through fuel savings and creates opportunities within emerging carbon pricing mechanisms, such as the Brazilian carbon market currently under development (16). The technical feasibility of this integration has been validated in detailed engineering studies, which have modeled the complex electrical interfaces and control systems required for stable operation (1; 4). To facilitate such research and ensure reproducibility, open-source reference platforms like the Low Emission Oil and Gas Open (LEOGO) platform have been developed to provide a common, well-defined test case for benchmarking new technologies and operational strategies (8). Figure 1.1 provides a conceptual illustration of such an integrated system. The hybrid power system combines renewable generation with backup gas turbines, connected through a common AC bus. The decision support framework monitors wind forecasts and controls generator dispatch to maintain system stability while minimizing emissions and operational costs.

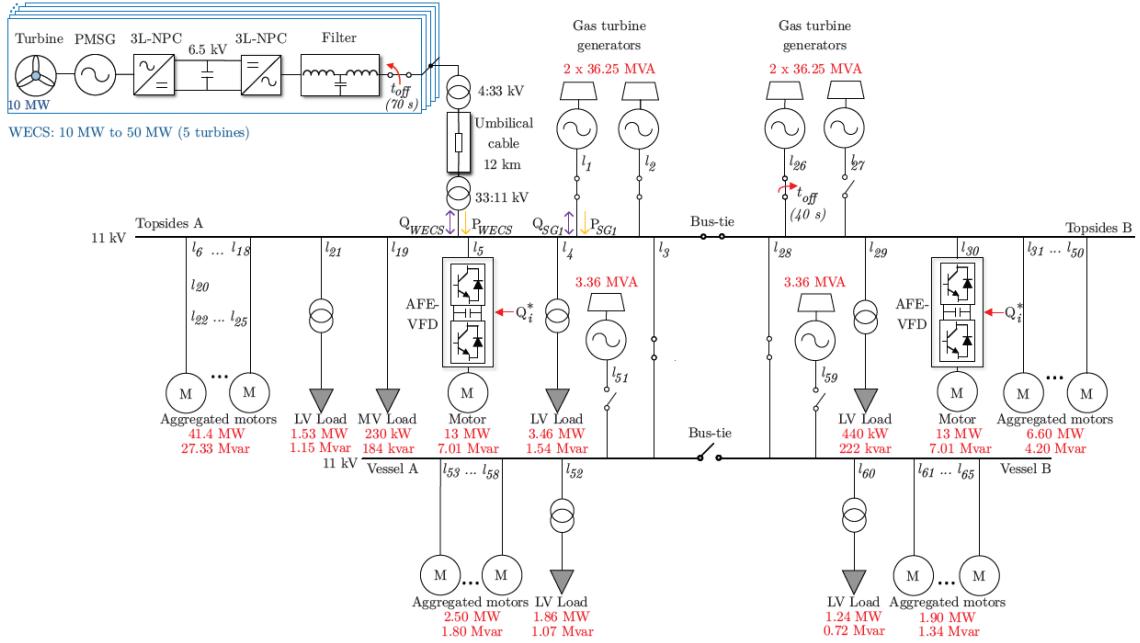


Figure 1.1: Conceptual illustration of an FPSO unit integrated with a floating offshore wind farm, showing the hybrid power system architecture with renewable generation, backup gas turbines, and common AC bus (adapted from (1)).

1.2 Motivation

Despite its compelling strategic and environmental rationale, the integration of wind power into an FPSO's isolated grid presents formidable engineering challenges that span from power system physics to operational strategy. These challenges are rooted in the fundamental conflict between the stringent reliability requirements of a critical industrial facility and the inherent variability of a renewable energy source.

1.2.1 The Physical Challenge: Intermittency and System Stability

The primary and most well-documented challenge is the physical nature of wind. Wind speed is a stochastic process, exhibiting non-linear, non-stationary, and chaotic behavior across multiple time scales (5; 12; 18). This physical reality translates directly into a highly volatile and intermittent power output from the Wind Energy Conversion System (WECS). As stated in (15), these characteristics pose "substantial challenges for prediction." Figure 1.2 illustrates a typical wind power generation profile, showcasing the rapid and unpredictable fluctuations, known as ramp events, that must be managed by the power system.

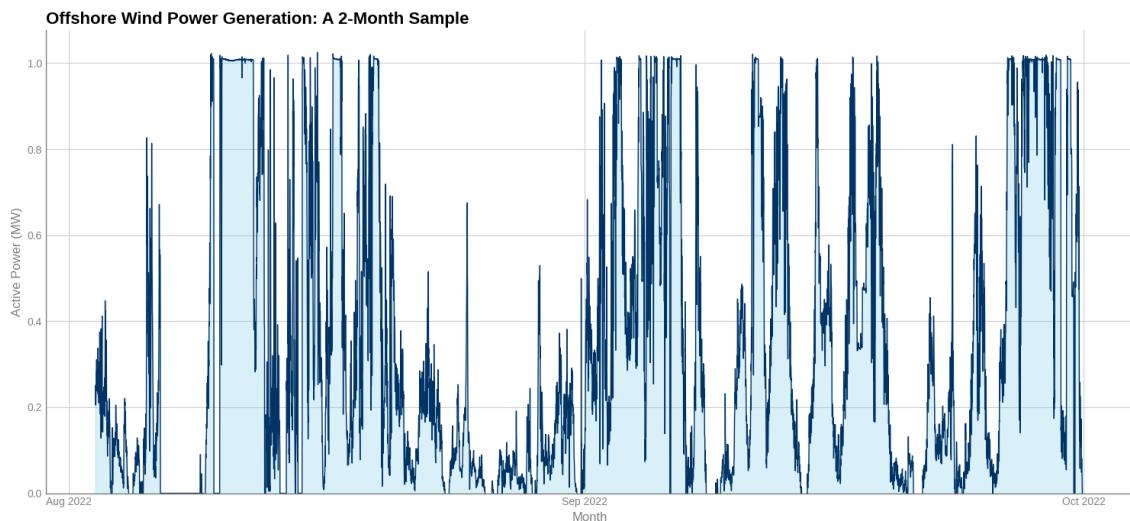


Figure 1.2: 10-min averaged wind power time series for 2 months from dataset (2).

The impact of this volatility is severely magnified by the unique characteristics of an FPSO's electrical grid. It is a small, isolated microgrid with very low rotational inertia, especially as converter-interfaced resources like WECS and Variable Frequency Drive (VFD) displace traditional synchronous generators (1; 19). This low inertia makes the system highly susceptible to frequency deviations. A sudden, unpredicted drop in wind power can cause a rapid decline in system frequency (a high Rate of Change of Frequency, or RoCoF), which can trigger protective load

shedding or, in the worst case, a full platform blackout (3; 4). The financial cost of a single blackout event, due to lost production, can be astronomical—potentially exceeding hundreds of thousands of dollars—not to mention the associated safety risks (20). Figure 1.3 depicts a typical frequency response to a generation-load imbalance, highlighting the critical time frames where control actions are necessary to prevent system collapse. The diagram illustrates three distinct control phases: (1) inertial response (0-2 seconds), driven by rotating mass; (2) primary frequency control (2-30 seconds), provided by governor action; and (3) secondary control (30 seconds to 20 minutes), requiring generator dispatch decisions. Critical frequency thresholds for protective actions are indicated. The low inertia of FPSO grids results in faster frequency decay, emphasizing the need for proactive control strategies.

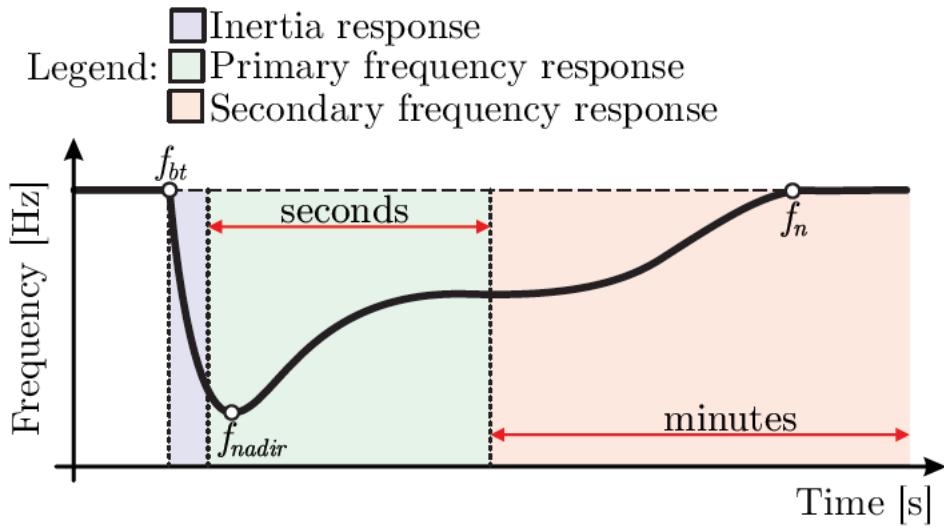


Figure 1.3: Typical frequency response following the tripping of a generator (adapted from (3)).

1.2.2 The Operational and Economic Challenge

The physical challenge of instability creates a critical operational dilemma: how to manage the backup generators to guarantee 100% reliability without nullifying the benefits of wind power. A purely reactive control strategy is untenable due to the significant start-up times of gas turbines, which typically require approximately 20 minutes to reach full operational capacity. This temporal constraint establishes accurate and reliable ultra-short-term wind power forecasting as a non-negotiable prerequisite for the successful operation of such a hybrid system (12). A robust forecast allows operators to transition from a reactive to a proactive operational paradigm, enabling informed decisions on generator dispatch and the use of energy storage systems (8).

The selection of specific forecast horizons— $t+10$, $t+20$, and $t+30$ minutes—is directly driven by these operational dynamics. Table 1.1 summarizes the operational rationale for each forecast horizon selected in this work.

Table 1.1: Operational justification for multi-horizon forecasting strategy.

Horizon	Role	Operational Rationale
$t+10$ min	Confirmation	Validates short-term trends and provides early warning of rapid changes (ramp events)
$t+20$ min	Critical Decision	Aligns with generator start-up time; primary trigger for dispatch decisions
$t+30$ min	Strategic Buffer	Provides safety margin for decision-making and accounts for forecast uncertainty

Moreover, the economic viability of these projects is highly sensitive to both operational strategy and external market forces, such as carbon pricing. As demonstrated by (16), while higher wind penetration leads to greater emissions reductions, it also increases initial capital investment and can lead to energy curtailment if not managed properly. The introduction of a carbon market, as is emerging in Brazil, can dramatically alter the economic calculus, making decarbonization strategies financially attractive by penalizing emissions and rewarding clean energy generation. This highlights the need for a holistic tool that not only provides technical forecasts but also integrates economic analysis to support strategic investment and operational decisions. Table 1.2 presents the operational cost parameters that define the economic optimization problem addressed by the decision support framework (21–23).

Table 1.2: Operational cost parameters for FPSO power management ¹.

Cost Component	Value	Description
Fuel consumption	\$500/hour	Cost of natural gas consumed during generator operation
Generator start-up	\$2,000/event	Wear, maintenance, and fuel costs associated with cold start
Platform blackout	\$100,000+/event	Lost production, equipment damage, and safety risks

These cost parameters, while representative, can vary significantly depending on oil prices, production rates, and specific FPSO configurations. The framework developed in this work is designed to be adaptable to different cost structures, making optimal decision-making economically critical across diverse operational contexts.

1.2.3 The Methodological Challenge: Advanced Forecasting Techniques

Achieving the required level of forecasting accuracy is a significant scientific challenge. The complex nature of wind speed time series makes them difficult to model accurately with single, monolithic forecasting models (15). To overcome this, the state-of-the-art has moved towards hybrid "divide and conquer" frameworks. These methods first employ signal decomposition techniques—such as Empirical Mode Decomposition (EMD) (24) and its more robust successors like Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) (25)—to parse the original non-stationary signal into a finite set of simpler, more predictable Intrinsic Mode Functions (IMFs). Each IMF, representing a distinct oscillatory mode, can then be forecasted with higher accuracy (5; 12). Recent work has even shown the benefit of multi-stage decomposition, where high-frequency components are further decomposed using techniques such as Empirical Wavelet Transform (EWT) (26) to isolate noise and improve predictability (15; 27).

Furthermore, for a system where decisions have direct safety and economic consequences, a simple deterministic (point) forecast is insufficient. It is crucial to quantify the inherent uncertainty in any prediction. Probabilistic forecasting, which provides a prediction interval (e.g., a 90% confidence range), delivers this vital information, enabling risk-based decision-making, where an operator can act based on a "worst-case" scenario (e.g., the 10th percentile of the forecast distribution), a practice that is standard in modern risk management (28). Quantile regression, trained using the pinball loss function, has emerged as a robust framework for generating such probabilistic forecasts (29; 30). Figure 1.4 illustrates the added value of a probabilistic forecast over a simple point prediction. The deterministic approach provides only a point estimate (median, dashed line), while the probabilistic approach generates multiple quantiles (10th, 25th, 75th, 90th percentiles) that define prediction intervals (shaded regions). The 10th percentile represents a conservative worst-case scenario used for risk assessment. This uncertainty quantification is essential for safe operation in high-stakes environments.

1.2.4 The Decision Support and Explainability Challenge

Beyond generating accurate forecasts, translating complex probabilistic predictions into actionable operational decisions represents a critical challenge. Operators in high-stakes environments require not only predictions but also clear guidance on how to act upon them. To address this gap, this work introduces the concept of an **Operational Risk Index (ORI)**, a derived metric that synthesizes multi-horizon quantile forecasts into a discrete, intuitive risk classification system (Low,

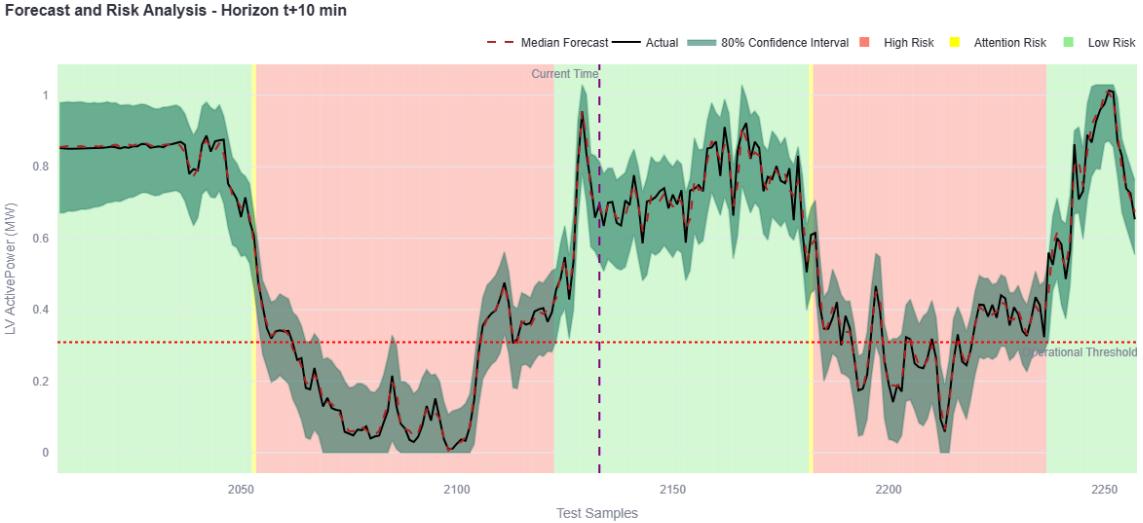


Figure 1.4: Comparison between deterministic (dashed line) and probabilistic (shaded regions) wind power forecasting. When the 10th percentile falls below the critical threshold (red line), the ORI triggers a high-risk alert for proactive generator dispatch.

Attention, High). The ORI is calculated by comparing worst-case scenario forecasts (10th percentile) at critical horizons ($t+20$ and $t+30$ minutes) against operational power thresholds, enabling consistent, risk-informed decision-making.

Furthermore, in safety-critical operational environments, the explainability of automated recommendations is as important as their accuracy. To address this, this work moves beyond a simple chatbot and implements a sophisticated **Retrieval-Augmented Generation (RAG)** framework. This framework supports a dual-level AI copilot: (1) a **reactive agent** for on-demand, conversational analysis, providing natural language explanations of the operational context; and (2) a **proactive agent** for autonomous surveillance. This proactive agent continuously analyzes the stream of forecast data to identify critical events—such as risk level changes or deteriorating trends—and generates timely, data-backed alerts to the operator, enhancing situational awareness and trust (31).

Finally, beyond real-time operational support, the framework includes a strategic optimization module designed to identify the most cost-effective operational strategy over extended periods. This optimizer performs a parametric sweep across key decision variables—operational power threshold and forecast horizon trigger—to find the combination that minimizes total cost (fuel, start-ups, and blackouts), enabling data-driven, economically optimized operational policies.

Figure 1.5 presents a high-level overview of the proposed framework architecture. As illustrated, the system is structured as a sequential pipeline that integrates four key pillars. **First**, a hybrid deep learning model generates probabilistic fore-

casts. **Second**, the ORI module translates these forecasts into discrete risk levels. **Third**, an interactive dashboard provides visualization and strategic optimization tools. **Finally**, the LLM-powered copilot, built on the RAG architecture, acts as the primary interface for explainability, delivering both proactive alerts and reactive analysis to the operator.

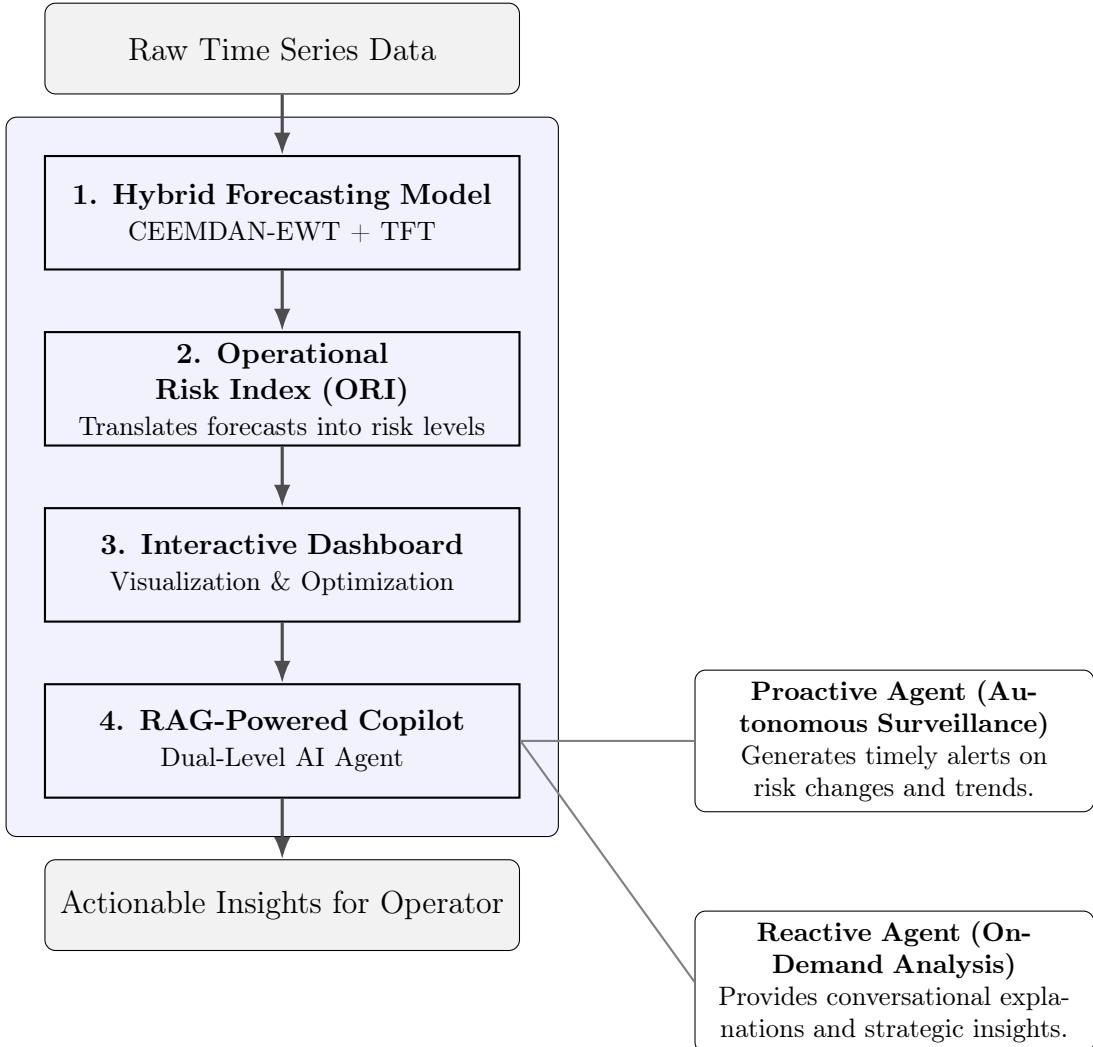


Figure 1.5: High-level overview of the proposed decision support framework architecture, illustrating the sequential integration of its four key components. The RAG-Powered Copilot (Pillar 4) features a dual-level architecture with both proactive and reactive agents.

In summary, the successful integration of wind power into FPSO operations requires addressing four interconnected challenges: (1) managing the physical intermittency and its impact on grid stability, (2) optimizing operational and economic trade-offs under uncertainty, (3) developing advanced forecasting techniques capable of handling non-stationary data, and (4) translating complex predictions into clear, actionable, and explainable decisions. This dissertation addresses all four challenges through an integrated framework that combines state-of-the-art deep learning, prob-

abilistic forecasting, risk quantification, and explainable AI. The specific objectives and contributions of this work are outlined in the following section.

1.3 Objectives

The general objective of this dissertation is to develop and validate a Deep Learning-based decision support framework for the power management of an FPSO with integrated offshore wind generation. The framework aims to provide accurate and reliable multi-step, probabilistic forecasts for horizons of t+10, t+20, and t+30 minutes, to inform the optimal dispatch of backup power systems and support strategic operational planning.

To achieve this, the following specific objectives are defined:

- **To conduct a comprehensive comparative study** of various time series forecasting models, ranging from classical benchmarks (SVR, Random Forest) to state-of-the-art Deep Learning architectures (LSTM, Bi-LSTM, Transformer, and TFT), to establish a robust performance baseline.
- **To investigate the effectiveness of hybrid models** that combine advanced signal decomposition techniques, such as CEEMDAN and EWT, with Deep Learning predictors, validating the hypothesis that such pre-processing is essential for handling non-stationary wind data (5; 12).
- **To propose a novel hybrid model**, the **CEEMDAN-EWT-TFT Aggregator**, which leverages the strengths of signal decomposition and the superior multi-horizon forecasting capabilities of the Temporal Fusion Transformer architecture, demonstrating a favorable balance between predictive accuracy and computational cost (11).
- **To implement all forecasting models within a quantile regression framework** to generate probabilistic predictions, thereby providing the prediction intervals necessary for quantitative risk assessment (29; 30).
- **To develop an Operational Risk Index (ORI)** that translates complex probabilistic forecasts into discrete, actionable risk levels (Low, Attention, High), enabling consistent risk-based decision-making under uncertainty.
- **To design and develop a prototype of a decision support dashboard** that functions as a digital twin of the power system. This tool translates quantile forecasts into actionable recommendations and includes a strategic

optimizer. Its core feature is a dual-level **Retrieval-Augmented Generation (RAG)** copilot, featuring a **proactive agent** for autonomous surveillance and a **reactive agent** for on-demand conversational analysis, providing fully explainable and context-aware operator support.

- **To implement a strategy optimization module** that identifies, through parametric sweep, the optimal combination of operational threshold and forecast horizon trigger that minimizes total operational cost (fuel consumption, generator start-ups, and blackout events).
- **To validate the complete framework** through simulation of realistic operational scenarios, demonstrating its practical applicability for FPSO operators and its potential for cost reduction and emissions mitigation.

1.3.1 Main Contributions

This dissertation makes the following key contributions to the field of wind power forecasting and decision support for offshore energy systems:

1. **Methodological Contribution:** A novel hybrid forecasting architecture (**CEEMDAN-EWT-TFT Aggregator**) that combines multi-stage signal decomposition with attention-based deep learning for superior accuracy in non-stationary wind power prediction.
2. **Operational Contribution:** The **Operational Risk Index (ORI)**, a practical metric that bridges the gap between complex probabilistic forecasts and actionable operational decisions, specifically designed for high-stakes industrial environments.
3. **AI Architecture Contribution:** A **dual-level AI copilot**, built on a **Retrieval-Augmented Generation (RAG)** architecture. This system features a proactive agent for autonomous surveillance and alert generation, and a reactive agent for on-demand conversational analysis, establishing a new paradigm for explainable and trustworthy decision support in critical operations.
4. **System Integration Contribution:** A validated, end-to-end prototype dashboard that integrates the forecasting model, ORI, and AI copilot into a unified platform. The tool demonstrates the practical feasibility and economic viability of the proposed framework, with demonstrated cost reduction potential through its strategic optimization module.

1.4 Work Structure

This dissertation is organized into seven chapters, following a logical progression from theoretical foundations to practical implementation and analysis, a structure similar to that adopted in related works from this research group (32). **Chapter 1** has introduced the context of offshore decarbonization, presented the technical, operational, and economic motivation for the work, and outlined its primary objectives. **Chapter 2** will present the theoretical foundations, covering the principles of offshore wind power generation and the electrical characteristics of FPSO power systems, including their typical loads and generation profiles (4; 8). **Chapter 3** provides a comprehensive literature review of wind power forecasting techniques, with a specific focus on the evolution of hybrid and deep learning methods that utilize signal decomposition (12; 15; 33). **Chapter 4** details the mathematical framework of the Deep Learning architectures and signal processing techniques employed in this study, including the quantile regression approach and the Temporal Fusion Transformer architecture. **Chapter 5** describes the proposed methodology in detail, from the dataset and preprocessing steps to the implementation of all forecasting models, the development of the Operational Risk Index (ORI), the architecture of the integrated decision support framework, the cost-benefit analysis model, the strategy optimization algorithm, and the Large Language Model (LLM)-powered copilot. **Chapter 6** presents and discusses the experimental results, including a rigorous comparative performance analysis of the forecasting models, an evaluation of the ORI's effectiveness in risk classification, a demonstration of the dashboard's functionalities and the strategy optimizer in simulated operational scenarios, and examples of interaction with the conversational agent. Finally, **Chapter 7** concludes the dissertation, summarizing its key contributions, acknowledging its limitations, and suggesting avenues for future research.

Chapter 2

Theoretical Foundations

This chapter establishes the theoretical groundwork necessary to understand the core challenges and proposed solutions of this dissertation. The primary objective is to construct a clear narrative, demonstrating how the fundamental principles of wind power and the unique constraints of Floating Production, Storage and Offloading (FPSO) power systems converge to create a complex operational problem. Crucially, this chapter will show that accurate, multi-step, and probabilistic forecasting is not merely an academic exercise, but an essential enabling technology for the safe and economic integration of renewable energy in this critical offshore environment.

2.1 Offshore Wind Power Generation

Offshore wind energy represents a cornerstone of the global energy transition, primarily due to the superior quality of wind resources found over the sea. These resources are typically stronger, more consistent, and less turbulent than their on-shore counterparts, which translates to higher capacity factors for wind turbines and, consequently, a greater energy yield per installed megawatt (4; 16). However, harnessing this vast energy potential effectively requires a deep understanding of the electromechanical conversion process and the inherent variability of the fuel source itself—the wind.

2.1.1 Wind Turbine Power Curve and Characteristics

The conversion of the wind’s kinetic energy into electrical energy is governed by the aerodynamic and mechanical properties of the Wind Energy Conversion System (WECS). The theoretical maximum power that can be extracted from the wind is defined by Betz’s Law, which states that no turbine can capture more than 59.3% of the kinetic energy in the wind (34). This theoretical limit arises from fundamental fluid dynamics: extracting all kinetic energy from the wind would require bringing

the air to a complete stop behind the rotor, which would prevent any further airflow through the turbine. The actual power extracted is a function of the turbine's power coefficient (C_p), which is itself a function of the blade's aerodynamic design and the ratio between the blade tip speed and the wind speed (Tip-Speed Ratio, TSR).

The relationship between the incident wind speed and the turbine's final electrical power output is encapsulated by the power curve, a fundamental characteristic of any wind turbine. As illustrated in Figure 2.1, the operation of a modern variable-speed wind turbine is typically divided into four distinct regions (4; 8; 35). Small wind speed variations can cause dramatically different power outputs depending on the region

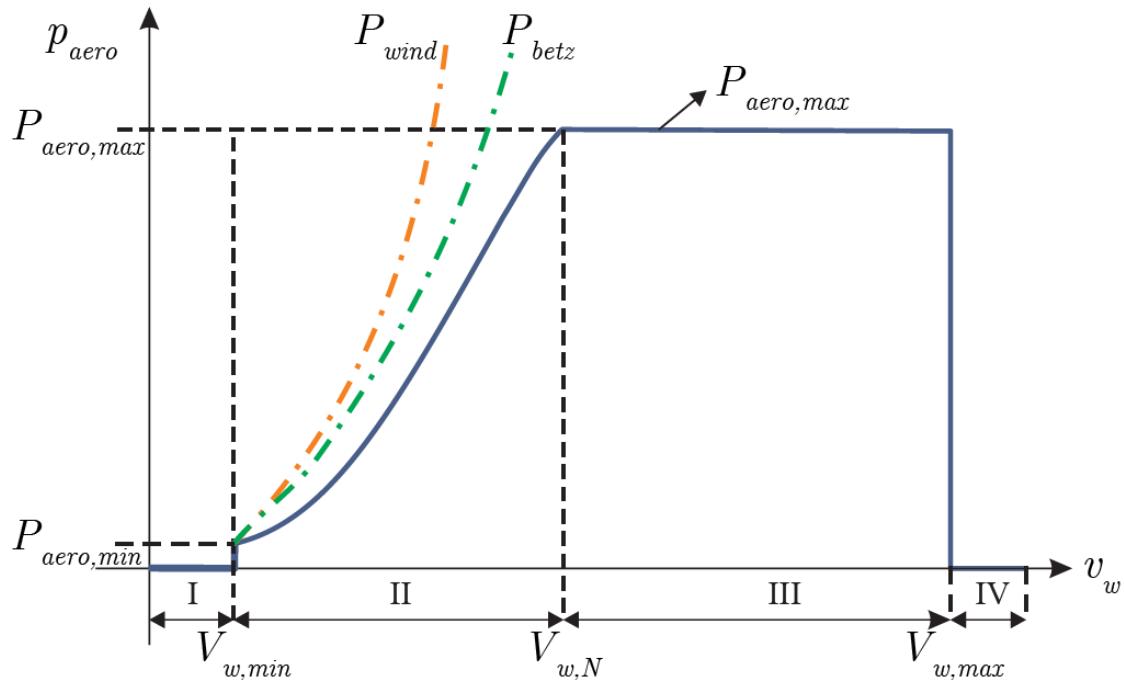


Figure 2.1: Theoretical power curve of a wind turbine showing four operational regions: (I) Cut-in, (II) MPPT, (III) Power Limitation, and (IV) Cut-out (adapted from (4)).

- **Region I (Cut-in):** Below a certain *cut-in speed* (typically 3-4 m/s), the wind's energy is insufficient to overcome the system's mechanical and electrical losses. The turbine remains stationary, producing zero power.
- **Region II (MPPT Region):** Between the cut-in speed and the *rated wind speed*, the turbine's control system actively adjusts the generator torque and blade pitch to maintain the optimal Tip-Speed Ratio, thereby maximizing the power coefficient (C_p) and extracting the maximum possible energy from the wind. In this region, power output is highly sensitive to wind speed variations, theoretically following a cubic relationship ($P \propto v^3$). This cubic dependency

means that even small forecast errors in wind speed can translate to significant errors in power prediction (36).

- **Region III (Power Limitation):** Once the wind speed reaches the rated value (e.g., 11-12 m/s), the turbine achieves its nominal power output. For any wind speed between the rated speed and the *cut-out speed*, the control system actively pitches the blades to shed excess aerodynamic lift, thereby limiting the power output to its rated value to protect the generator and converter from overload. In this region, the power output is relatively insensitive to wind speed variations, which paradoxically makes forecasting more stable (37).
- **Region IV (Cut-out):** Above the cut-out speed (typically around 25 m/s), high wind speeds pose a structural risk to the turbine. The system initiates a controlled shutdown, pitching the blades to a fully feathered position to stop rotation and minimize mechanical stress. While these extreme events are rare in typical operational conditions, they represent critical safety scenarios that must be anticipated.

Modern offshore wind turbines are predominantly Type 4 turbines, where the generator is connected to the grid through a full-scale power electronic converter, as depicted in Figure 2.2. The machine-side converter controls generator torque (MPPT), while the grid-side converter manages power injection and ancillary services. This architecture maximizes energy capture but eliminates natural inertial response. This architecture completely decouples the generator's mechanical frequency from the grid's electrical frequency, allowing for variable speed operation to optimize energy capture across a wide range of wind conditions. However, as will be detailed later, this decoupling is also the source of the low-inertia challenge central to this dissertation (19).

2.1.2 Challenges of Offshore Wind Forecasting

The core challenge of wind forecasting lies not in the turbine technology, but in the fundamental nature of the fuel: the wind itself. Wind speed time series exhibit characteristics that make them particularly difficult to predict using conventional forecasting methods (5; 12):

- **Non-stationarity:** A time series is considered stationary if its statistical properties (mean, variance, autocorrelation) remain constant over time. Wind speed is fundamentally non-stationary, as these properties continuously change with evolving weather patterns, diurnal cycles, and seasonal variations. This

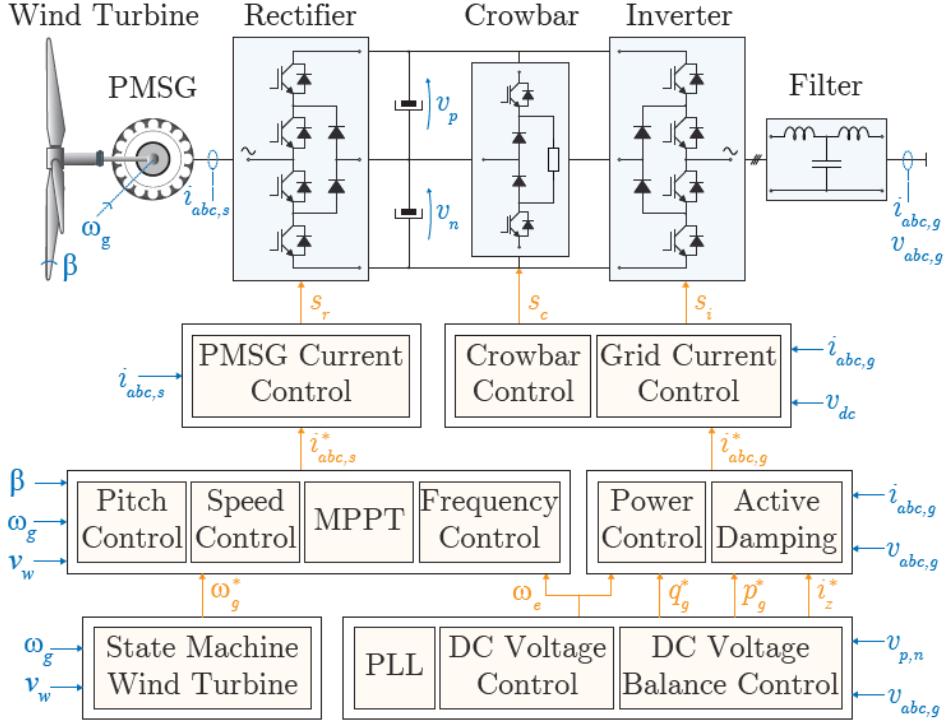


Figure 2.2: Simplified schematic of a Type 4 WECS with full-scale back-to-back converter (adapted from (4)).

violates the core assumptions of many classical forecasting models (e.g., AutoRegressive Integrated Moving Average (ARIMA)), which require the data to be stationary or easily transformable to stationarity to be effective (38; 39).

- **Non-linearity:** The relationship between past and future wind speeds is highly non-linear, meaning that simple linear regression models fail to capture the complex dynamics. This non-linearity is further compounded by the cubic relationship between wind speed and power in Region II of the power curve (35).
- **Multi-scale Dynamics:** Wind patterns are a composite of atmospheric phenomena occurring at vastly different temporal and spatial scales. These range from slow-moving synoptic weather systems (low-frequency components with periods of days) to rapid, localized turbulence and gusts (high-frequency components with periods of seconds to minutes) (15). As noted by (27), these high-frequency components, while containing less energy, can significantly corrupt the learning process of forecasting models by introducing noise that obscures the underlying predictable patterns.
- **Ramp Events:** Perhaps the most operationally critical phenomenon is the occurrence of ramp events—rapid, large-magnitude changes in wind speed

(and consequently power output) over short time periods. These events, often caused by the passage of weather fronts or sudden changes in atmospheric stability, pose the greatest risk to grid stability in isolated systems like FP-SOs (20). Accurately predicting the timing, magnitude, and duration of ramp events is essential for proactive generator dispatch.

Figure 2.3 visually demonstrates the power of signal decomposition techniques to address these challenges. A raw, chaotic wind power signal can be understood as the sum of several simpler, more regular components, each representing a distinct oscillatory mode with different frequency characteristics. The original non-stationary series is decomposed into Intrinsic Mode Functions (IMFs) with decreasing frequency content plus a residual. High-frequency IMFs capture turbulence, mid-frequency IMFs capture diurnal patterns, and the residual captures long-term trends. This observation is the cornerstone of the hybrid "divide and conquer" forecasting strategy employed in this work. Techniques like CEEMDAN are not arbitrary pre-processing steps; they are a direct methodological response to the multi-scale, non-stationary physical nature of wind (5; 12; 27). By decomposing the signal into Intrinsic Mode Function (IMF)s, we transform a single intractable forecasting problem into a set of more manageable sub-problems, each of which can be addressed with specialized predictors.

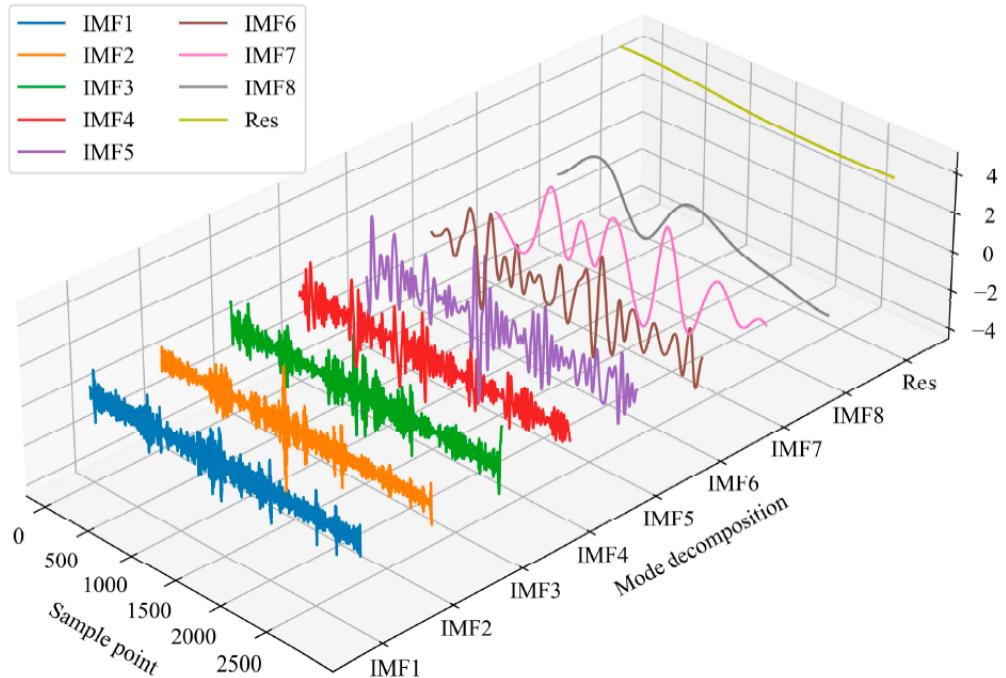


Figure 2.3: CEEMDAN decomposition of a wind power signal (adapted from (5)).

These forecasting challenges are further compounded in operational contexts where decisions must be made rapidly and consistently. While probabilistic forecasts

provide essential uncertainty quantification through prediction intervals, translating multiple quantile forecasts across multiple horizons into actionable operational decisions remains a non-trivial challenge. This motivates the development of derived risk metrics that can synthesize complex probabilistic information into intuitive, discrete decision signals—a concept that will be explored in detail in Chapter 5.

2.2 Power Systems in FPSOs

The challenges posed by wind’s inherent variability are amplified exponentially when the receiving electrical system is a small, isolated microgrid like an FPSO. Unlike a robust continental grid with hundreds of generators and vast interconnections, an FPSO power system operates as an electrical island with severely limited resilience to disturbances, making reliability the paramount design and operational concern (8).

2.2.1 Conventional Power Generation

Traditionally, the electrical power on an FPSO is supplied by multiple aero-derivative gas turbine generators (GTGs), typically operating in parallel to meet the platform’s substantial power demand (8; 40). These turbines, often ranging from 20 to 40 MW each depending on the platform size, burn natural gas extracted on-site as fuel, making them operationally convenient but environmentally problematic due to their CO₂ emissions (16).

The gas turbines serve a dual role: they provide the electrical power required by the platform’s extensive processing equipment (compressors, pumps, separators) and utility systems (HVAC, lighting), and they also provide the critical system attribute known as rotational inertia. Inertia, quantified by the inertia constant H (measured in seconds), represents the kinetic energy stored in the rotating masses of the turbine-generator sets. This stored energy acts as a natural buffer against sudden changes in the power balance, providing an immediate, automatic response to disturbances before any control system can react (41).

The inertia constant H is defined as:

$$H = \frac{E_{\text{kinetic}}}{S_{\text{rated}}} = \frac{\frac{1}{2}J\omega_m^2}{S_{\text{rated}}} \quad (2.1)$$

where J is the combined moment of inertia of the turbine and generator rotor, ω_m is the mechanical angular velocity, and S_{rated} is the rated apparent power of the generator. A typical synchronous generator in a power system might have an H value ranging from 2 to 6 seconds. The challenge arises because converter-interfaced

sources like WECS are electrically decoupled from the grid frequency and do not inherently contribute to this rotational inertia, thus lowering the overall system H constant as their penetration increases (1; 19).

2.2.2 Operational Risks: Blackouts and Load Shedding

The fundamental principle of power system stability is maintaining a continuous, instantaneous balance between power generation and load consumption. Any mismatch between these two quantities results in a change in system frequency, a relationship governed by the swing equation (41):

$$\frac{2H}{f_n} \frac{df}{dt} = P_m - P_e \quad (2.2)$$

where f is the system frequency, f_n is the nominal frequency (50 or 60 Hz), P_m is the mechanical power input from generators, and P_e is the electrical power consumed by loads. The term df/dt is the Rate of Change of Frequency (Rate of Change of Frequency (RoCoF)).

Equation 2.2 reveals a critical insight: for a given power imbalance ($\Delta P = P_m - P_e$), a lower system inertia H results in a higher (and more dangerous) RoCoF. In a low-inertia system, such as an FPSO with high wind penetration, a sudden loss of generation (like a wind ramp-down event) leads to a very high RoCoF. The primary frequency control mechanism—the droop control of the gas turbine governors—must act quickly to increase mechanical power input and arrest this frequency decline. However, these governors have inherent mechanical and thermal limitations and typically respond on a timescale of seconds to tens of seconds (1).

If the frequency falls below a critical threshold before the governors can stabilize it, emergency load shedding is automatically triggered. This involves disconnecting non-critical loads to reduce electrical demand and prevent a system collapse. In the worst-case scenario, if protective actions are insufficient, the system experiences a complete blackout—a catastrophic event that halts all production and can result in financial losses exceeding hundreds of thousands of dollars per hour (3; 20).

Figure 2.4 provides a concrete example of this dynamic, using simulations from an isolated offshore power grid following the loss of a main generation unit (6). Figure 2.4a depicts the system's response without intervention. Although the frequency stabilizes at a lower value (approx. 57 Hz), this stability is achieved by forcing the remaining online generator into a critical overload condition, a state that is unsustainable and would ultimately lead to its thermal trip and a full blackout. In contrast, Figure 2.4b demonstrates the effectiveness of a load shedding scheme. By disconnecting non-critical loads 200 ms after the event, the power demand is

reduced, the generator overload is prevented, and the system frequency is safely restored to its nominal value. This comparison highlights that proactive measures are essential not only for frequency stability but also for preventing catastrophic equipment failure.

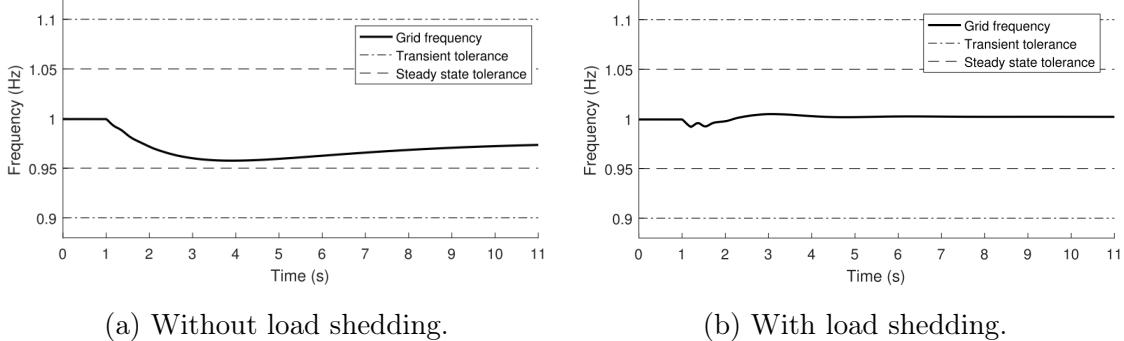


Figure 2.4: Simulated grid frequency response in a 60 Hz isolated offshore system following the sudden loss of a main generation unit. (a) Without load shedding, the frequency drops but stabilizes at the cost of putting the remaining generator into a critical and unsustainable overload condition. (b) With a proactive load shedding scheme, the generator overload is avoided and the frequency is safely restored to its nominal value. This demonstrates the critical role of automated protection in preventing equipment failure and subsequent blackouts (adapted from (6)).

This scenario highlights the absolute necessity of forecasting in this context. An accurate forecast does not prevent the wind from dropping; it wins the race against time. By providing a reliable forecast horizon of 10, 20, or 30 minutes, the decision support system gives operators the crucial lead time needed to preemptively start a backup generator, ensuring that replacement power is already synchronized and online *before* the anticipated wind power deficit occurs. This transforms the operational paradigm from a high-risk reactive mode to a safe, proactive one, effectively eliminating the risk of frequency excursions and blackouts caused by predictable wind variability.

The severity of these consequences creates a fundamental operational dilemma: operators must balance the competing objectives of maintaining system reliability (avoiding blackouts) while minimizing operational costs (avoiding unnecessary generator starts and fuel consumption). A deterministic forecast alone is insufficient for this task, as it provides no information about the likelihood or magnitude of forecast errors. For instance, a point forecast of 8 MW at $t+20$ minutes could represent either a highly confident prediction with a narrow uncertainty band [7.5–8.5 MW] or a highly uncertain prediction with a wide band [4–12 MW]. These two scenarios demand entirely different operational responses. This uncertainty quantification requirement, combined with the need to integrate forecasts across multiple horizons and translate them into clear operational actions, forms the core motivation for the

decision support framework developed in this dissertation.

2.2.3 Hybrid Power Systems: Challenges and Opportunities

The integration of a WECS into the conventional FPSO grid fundamentally transforms it into a hybrid power system, presenting a complex duality of challenges and opportunities that must be carefully balanced (3; 4).

Challenges

- **Reduced System Inertia:** As wind power displaces generation from synchronous gas turbines (which may be shut down or operated at reduced capacity), the overall rotational inertia of the system decreases proportionally. This makes the system increasingly vulnerable to frequency deviations and reduces the time available for corrective actions (19; 42).
- **Increased Variability:** The introduction of a highly variable generation source increases the overall volatility of the net power balance, requiring more frequent adjustments to generator dispatch and potentially increasing wear on equipment due to more frequent start-stop cycles (20).
- **Power Quality Concerns:** Converter-interfaced generation can introduce harmonics and voltage fluctuations if not properly controlled, potentially affecting sensitive equipment on the platform (3).

Opportunities

- **Emissions Reduction:** The primary benefit is a substantial reduction in CO₂ emissions, with potential reductions of 23% to over 76% depending on wind capacity and operational strategy, directly contributing to decarbonization goals (16).
- **Fuel Savings:** Reduced reliance on gas turbines translates directly to fuel cost savings, improving the economic viability of operations, particularly in the context of emerging carbon pricing mechanisms (16).
- **Advanced Control via Power Electronics:** While power converters are the source of the inertia problem, they are paradoxically also the source of potential solutions. Modern converters, such as those in Active Front-End Variable Frequency Drives (AFE-VFDs) used for large compressor motors, can be programmed to provide a wide range of grid-supportive ancillary services (1). These include:

- **Fast Frequency Support (Virtual Inertia):** Converters can be controlled to emulate the inertial response of synchronous machines by rapidly modulating their active power output in response to frequency deviations, effectively synthesizing inertia (42; 43).
- **Voltage Support:** Converters can provide dynamic reactive power compensation to maintain voltage stability, a function traditionally performed by synchronous condensers or capacitor banks (1).
- **Harmonic Filtering:** Advanced converter control can actively mitigate harmonic distortion, improving overall power quality (3).

Successfully navigating this complex operational landscape requires more than accurate forecasts alone. Operators face a continuous stream of multi-objective decisions: when to start backup generators (balancing lead time against fuel waste), which operational threshold to use (balancing safety margins against economic efficiency), and how to interpret conflicting signals from multiple forecast horizons. These decisions are further complicated by the stochastic nature of wind and the asymmetric cost structure—where the cost of a single blackout event can exceed the cumulative fuel savings from weeks of optimized operation. This decision complexity motivates the development of integrated decision support tools that combine forecasting, risk assessment, economic optimization, and explainable recommendations in a unified framework, as will be presented in Chapter 5.

2.3 The Need for Integrated Decision Support

The theoretical foundations presented in this chapter reveal a multi-faceted challenge that extends beyond forecasting accuracy alone. The successful integration of wind power into FPSO operations requires addressing four interconnected layers:

1. **Predictive Layer:** Accurate multi-horizon probabilistic forecasts that quantify uncertainty and capture the non-stationary, multi-scale dynamics of wind power generation.
2. **Risk Assessment Layer:** Translation of complex probabilistic forecasts into intuitive risk metrics that enable consistent, risk-informed decision-making under uncertainty.
3. **Optimization Layer:** Identification of optimal operational strategies that balance competing objectives (reliability, cost, emissions) across varying operational conditions.

4. Human-Machine Interface Layer: Presentation of recommendations in an explainable, trustworthy manner that enhances rather than replaces human expertise, particularly in safety-critical contexts where operators must understand the reasoning behind automated suggestions.

This dissertation addresses all four layers through an integrated framework that combines state-of-the-art deep learning (CEEMDAN-EWT-TFT), derived risk metrics (Operational Risk Index), parametric optimization (strategy optimizer), and explainable AI (LLM-powered conversational copilot). The following chapters detail the theoretical foundations of the forecasting models (Chapter 3), the mathematical framework of the proposed methods (Chapter 4), the complete methodology and implementation (Chapter 5), the experimental validation (Chapter 6), and the conclusion (Chapter 7).

Chapter 3

Literature Review

This chapter provides a comprehensive review of the scientific literature pertinent to the core themes of this dissertation. The review is structured to build a clear argument, starting from the broad field of wind power forecasting and progressively narrowing down to the specific challenges of power management in isolated offshore systems. The primary goal is to map the current state-of-the-art, identify existing research gaps, and thereby firmly position the contributions of this work within the academic landscape.

3.1 Wind Power Forecasting Techniques

The accurate forecasting of wind power is a field of intense research, driven by the direct economic and operational benefits it provides to power system operators. The literature presents a clear evolutionary path of forecasting methodologies, moving from simpler statistical models to highly complex hybrid Deep Learning frameworks, each generation attempting to better capture the intricate dynamics of wind.

3.1.1 Classical and Statistical Methods

Early approaches to time series forecasting were dominated by statistical methods, most notably the Autoregressive Integrated Moving Average (ARIMA) family of models, formalized by Box and Jenkins (39). ARIMA models attempt to describe the autocorrelations in the data, essentially assuming that future values have a linear dependency on past values and past forecast errors (12). While effective for stationary and linear time series, their performance degrades significantly when applied to wind speed data due to the latter's inherent non-stationarity and non-linearity, as discussed in Chapter 2. Seasonal variants like SARIMA have been proposed to handle periodicities, but they still struggle with the wind's stochastic nature (38). These models serve as important benchmarks but are generally considered insufficient for

the high-accuracy requirements of modern grid management.

3.1.2 Machine Learning and Deep Learning Approaches

The limitations of statistical models led researchers to adopt Machine Learning (ML) techniques, which are better suited for modeling non-linear relationships. Early ML models applied to this problem include Support Vector Regression (SVR), an application of Support Vector Machines for regression tasks (44), and ensemble methods like Random Forests (RF) (12). However, the true breakthrough came with the advent of Deep Learning (DL), particularly Recurrent Neural Networks (RNNs).

Unlike feedforward networks, RNNs possess internal memory through recurrent connections, making them naturally suited for sequential data like time series. The Long Short-Term Memory (LSTM) network, a sophisticated variant of RNN introduced by (45), was specifically designed to overcome the vanishing gradient problem, enabling it to learn long-term dependencies in the data. The Gated Recurrent Unit (GRU), proposed by (46), offers a similar capability with a slightly simpler architecture. Numerous studies have demonstrated the superiority of LSTM and GRU over classical ML and statistical models for wind forecasting (12; 15).

More recently, architectures inspired by computer vision, such as Convolutional Neural Networks (CNNs), have been integrated into hybrid models. CNNs excel at extracting local features and patterns, and when combined with LSTMs (e.g., CNN-LSTM models), they can capture both spatio-temporal features from gridded data or hierarchical features from a single time series (27). The Attention Mechanism, which allows a model to dynamically weigh the importance of different past time steps, has proven highly effective. This concept culminated in the Transformer architecture, introduced by (10), which dispenses with recurrence entirely and relies solely on self-attention mechanisms to model dependencies. The Temporal Fusion Transformer (TFT), an evolution of this architecture, has shown state-of-the-art performance in multi-horizon time series forecasting (11). The TFT architecture (Figure 3.1) integrates these advances, combining variable selection networks for interpretability with attention-based temporal processing for multi-horizon forecasting.

3.1.3 Signal Decomposition Techniques in Time Series Forecasting

Despite the power of DL models, their performance is still hampered by the raw complexity of the wind signal. This led to the development of hybrid "decomposition-and-conquer" strategies, which have become the state-of-the-art (47). The core idea

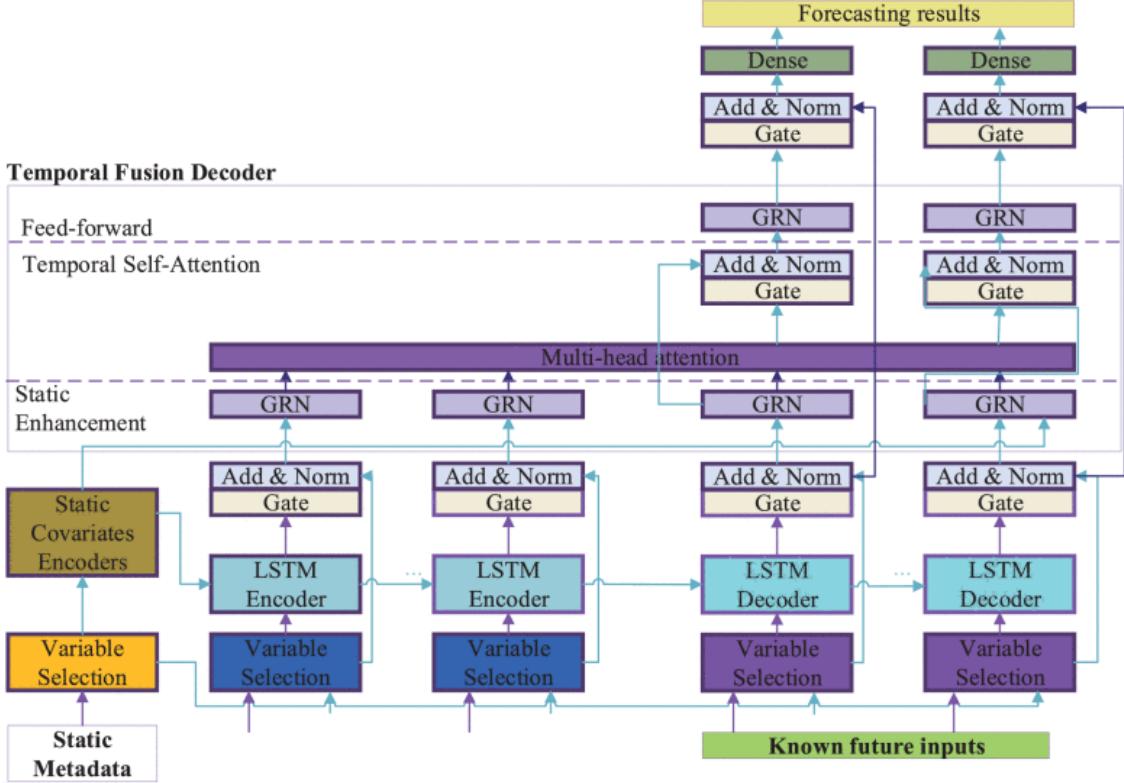


Figure 3.1: Temporal Fusion Transformer architecture showing variable selection networks, temporal processing layers, and multi-horizon prediction capabilities (adapted from (7)).

is to first use a signal processing technique to decompose the non-stationary wind signal into a set of simpler, more predictable components.

- **Empirical Mode Decomposition (EMD):** Proposed by (24), EMD adaptively decomposes a signal into a finite set of Intrinsic Mode Functions (IMFs). However, EMD suffers from a problem known as "mode mixing," where oscillations of different scales coexist in a single IMF.
- **Ensemble EMD (EEMD) and CEEMDAN:** To address mode mixing, variants like EEMD and, more effectively, Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) were developed. CEEMDAN introduces adaptively calculated white noise at each decomposition stage, yielding IMFs that are more physically meaningful and easier to forecast (12; 25).
- **Empirical Wavelet Transform (EWT):** EWT combines the adaptability of EMD with the mathematical rigor of wavelet theory. It first identifies the main frequency components in the signal's Fourier spectrum and then constructs an appropriate wavelet filter bank to separate them (12; 26).

The literature is rich with studies demonstrating the success of combining these decomposition techniques with DL models. Figure 3.2 shows a complete flowchart of such a hybrid approach. This process, as demonstrated by (12) and (27), confirms that pre-processing the signal via decomposition significantly improves forecasting accuracy. This body of work provides a strong justification for the decomposition-based approach adopted in this dissertation. However, a research gap remains in exploring more advanced "aggregator" models that can learn from the decomposed components in a multivariate fashion, rather than forecasting each component in isolation—a gap that this work addresses through the use of TFT as a unified predictor for all IMFs.

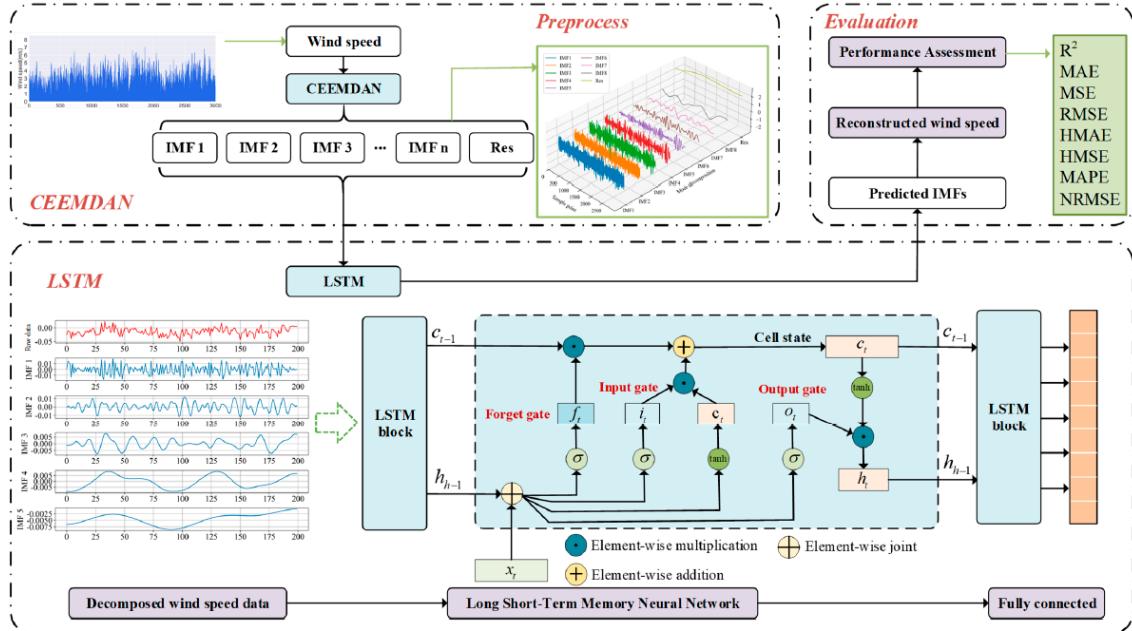


Figure 3.2: CEEMDAN-LSTM hybrid framework with separate predictors for each decomposed component (adapted from (5)).

3.1.4 Probabilistic Forecasting and Uncertainty Quantification

While the majority of forecasting research focuses on improving point forecast accuracy (e.g., minimizing RMSE or MAE), operational decision-making in safety-critical systems requires more than a single predicted value—it requires a quantification of the uncertainty surrounding that prediction. This has led to the emergence of probabilistic forecasting as a distinct and essential research area (28).

Probabilistic forecasts provide not just a single expected value, but a full probability distribution (or approximations thereof) of possible future outcomes. The most common approach is to generate prediction intervals at various confidence

levels, or equivalently, to predict specific quantiles of the conditional distribution (29; 30). For instance, the 10th percentile (P10) represents a conservative "worst-case" scenario, while the 90th percentile (P90) represents an optimistic scenario. The median (P50) serves as a robust point forecast.

The standard technique for generating these quantile forecasts is quantile regression, which directly minimizes an asymmetric loss function (the pinball loss) for each desired quantile (29). This approach has been successfully combined with various ML models, including neural networks, to create probabilistic forecasting systems (48). The TFT architecture, in particular, was designed with native support for quantile regression, making it an ideal choice for applications requiring both multi-horizon and probabilistic forecasts (11).

However, a critical gap exists in the literature: while many studies generate probabilistic forecasts, few address the challenge of translating these complex probabilistic outputs into actionable operational decisions. For an operator managing a FPSO, interpreting multiple quantile forecasts across multiple horizons (e.g., P10, P50, P90 at $t+10$, $t+20$, $t+30$) is cognitively demanding and prone to inconsistent decision-making. This motivates the development of derived risk metrics—such as the Operational Risk Index (ORI) proposed in this dissertation—that synthesize probabilistic information into intuitive, discrete decision signals (e.g., Low/Attention/High risk).

3.1.5 Evaluation Metrics for Forecasting Performance

The assessment of forecasting model performance requires appropriate metrics that align with the operational objectives. The literature distinguishes between metrics for deterministic (point) forecasts and probabilistic forecasts (28).

Deterministic Forecast Metrics

For point forecasts, the most commonly used metrics include:

- **Mean Absolute Error (MAE):** $MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$, which provides a direct measure of average forecast error in the original units.
- **Root Mean Squared Error (RMSE):** $RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}$, which penalizes larger errors more heavily due to the squaring operation.
- **Mean Absolute Percentage Error (MAPE):** $MAPE = \frac{100\%}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right|$, which provides scale-independent error measurement but is undefined when $y_i = 0$.
- **Symmetric Mean Absolute Percentage Error (sMAPE):** An alternative to MAPE that addresses the asymmetry and division-by-zero issues.

- **Normalized Root Mean Squared Error (NRMSE):** RMSE normalized by the range or mean of the observed values, enabling comparison across different datasets.
- **Coefficient of Determination (R^2):** Measures the proportion of variance in the dependent variable explained by the model.

Probabilistic Forecast Metrics

For probabilistic forecasts, specialized metrics are required to assess both the sharpness (precision) and calibration (reliability) of the predicted distributions:

- **Pinball Loss (Quantile Loss):** For a given quantile τ , the pinball loss is defined as:

$$L_\tau(y, \hat{q}_\tau) = \begin{cases} \tau(y - \hat{q}_\tau) & \text{if } y \geq \hat{q}_\tau \\ (1 - \tau)(\hat{q}_\tau - y) & \text{if } y < \hat{q}_\tau \end{cases} \quad (3.1)$$

This asymmetric loss function is the basis for quantile regression and directly evaluates the accuracy of individual quantile forecasts (29).

- **Continuous Ranked Probability Score (CRPS):** A comprehensive metric that evaluates the entire predicted distribution by comparing the cumulative distribution function (CDF) of the forecast to the observed value. CRPS generalizes MAE to probabilistic forecasts and rewards both sharpness and calibration (49). It is defined as:

$$\text{CRPS}(F, y) = \int_{-\infty}^{\infty} [F(x) - \mathbb{1}\{x \geq y\}]^2 dx \quad (3.2)$$

where F is the forecast CDF and y is the observed value.

- **Prediction Interval Coverage Probability (PICP):** Measures the proportion of observations that fall within the predicted interval, assessing calibration.
- **Prediction Interval Normalized Average Width (PINAW):** Measures the average width of prediction intervals, assessing sharpness.

This dissertation employs a combination of these metrics: MAPE, RMSE, MAE, and R^2 for point forecast evaluation, and Pinball Loss for quantile-specific assessment, providing a comprehensive evaluation framework aligned with both academic standards and operational requirements.

3.1.6 Forecast Horizons in Operational Contexts

The choice of forecast horizon is not arbitrary but must be aligned with the physical and operational constraints of the system being managed. The wind forecasting literature typically categorizes horizons into three classes (50; 51):

- **Ultra-short-term:** 1 minute to 1 hour ahead. Used for real-time control, frequency regulation, and automatic generation control (AGC).
- **Short-term:** 1 hour to 72 hours ahead. Used for unit commitment, economic dispatch, and day-ahead market participation.
- **Medium to long-term:** Beyond 72 hours. Used for maintenance scheduling, resource planning, and strategic decision-making.

For the specific application of managing backup generator dispatch in a FPSO, the critical horizon falls within the ultra-short-term category. As established in Chapter 1, typical gas turbine generators require approximately 15-20 minutes from cold start to full synchronization and load acceptance. This physical constraint directly determines the required forecast lead time: the system must predict potential wind power deficits at least 20 minutes in advance to allow sufficient time for proactive generator startup.

This work adopts three specific horizons— $t+10$, $t+20$, and $t+30$ minutes—to provide a graduated decision support capability. The $t+10$ forecast serves as an early warning, $t+20$ aligns with the critical generator startup time, and $t+30$ provides additional margin for conservative decision-making. This multi-horizon approach is consistent with recent research emphasizing the value of providing operators with a temporal spectrum of forecasts rather than a single fixed horizon (52).

3.1.7 Comparative Analysis of Forecasting Methods

Table 3.1 provides a systematic comparison of the major forecasting approaches discussed in this section, evaluating their suitability for the FPSO application based on their ability to handle non-stationarity, computational complexity, and appropriateness for ultra-short-term multi-horizon prediction in safety-critical offshore operations.

This comparative analysis clearly justifies the selection of the CEEMDAN-EWT-TFT hybrid approach for this dissertation. The decomposition stage (CEEMDAN-EWT) directly addresses the non-stationarity and multi-scale nature of offshore wind, while the TFT predictor provides native multi-horizon and probabilistic forecasting capabilities. This combination has demonstrated state-of-the-art performance in recent literature (12; 27). While computationally more expensive than

Table 3.1: Comparative analysis of wind power forecasting methods for FPSO application.

Method	Complexity	Key Characteristics & FPSO Suitability
<i>Statistical & Classical ML Methods</i>		
ARIMA	Low	Linear, assumes stationarity. Low suitability for non-stationary offshore wind.
SVR	Low	Non-linear, no temporal modeling. Medium — baseline only.
Random Forest	Low	Ensemble method, no temporal modeling. Medium — baseline only.
<i>Recurrent Deep Learning Methods</i>		
LSTM	Medium	Captures temporal dependencies. High — strong candidate.
Bi-LSTM	Medium	Bidirectional temporal context. High — improved over LSTM.
CNN-Bi-LSTM	Medium-High	Feature extraction + temporal modeling. High — hybrid architecture.
<i>Attention-Based Methods</i>		
Transformer	High	Self-attention mechanism, no recurrence. High — interpretable.
TFT	High	Native multi-horizon + quantile regression + attention. Very High — ideal for multi-horizon.
<i>Hybrid Decomposition Methods (This Work)</i>		
CEEMDAN-EWT-LSTM	Very High	Signal decomposition + LSTM (1 model per IMF). High — best accuracy.
CEEMDAN-EWT-TFT	High	Signal decomposition + unified TFT (1 model). Optimal — best trade-off.

simpler methods, this cost is justified by the high-stakes nature of FPSO operations where forecast accuracy directly impacts safety and economics. The experimental results presented in Chapter 6 will demonstrate that this hybrid approach consistently outperforms all baseline and standalone deep learning methods across multiple evaluation metrics.

3.2 Power Management in Offshore Systems and Microgrids

The forecasting of wind power is not an end in itself; it is a means to enable more effective power management. In the context of an isolated offshore platform, this management is a high-stakes balancing act between operational security, fuel consumption, and equipment wear-and-tear.

3.2.1 Conventional and Hybrid Power Systems in FPSOs

The literature on offshore power systems highlights the central role of the gas turbine generators in maintaining grid stability. Their governors provide the primary frequency control, while their automatic voltage regulators (AVRs) manage voltage levels. A key operational constraint is the need to maintain a "spinning reserve"—an amount of online, unloaded capacity that can be rapidly dispatched to cover a sudden load increase or a loss of generation (8). Figure 3.3 provides a schematic of a reference offshore platform, illustrating the complex interplay between power generation sources and various electrical loads.

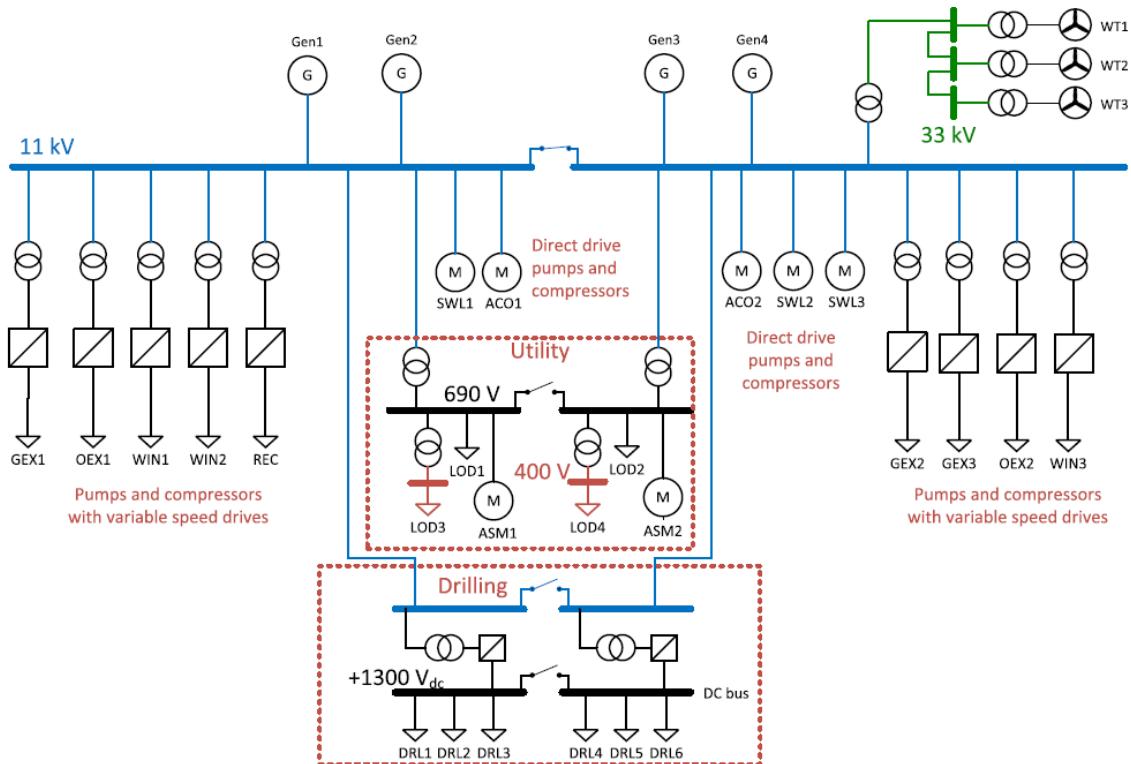


Figure 3.3: Schematic of the LEOGO offshore platform energy system with integrated wind turbines (adapted from (8)).

When a variable source like wind is introduced, the Power Management System (PMS) must constantly decide how to dispatch the controllable units (the GTGs) to

compensate for the fluctuations of the uncontrollable one (the WECS). This involves complex trade-offs. For instance, running an extra GTG at part-load provides a larger spinning reserve and enhances security, but it does so at the cost of lower fuel efficiency and higher emissions (16).

The work by (3) and (1) explores the use of power electronics, specifically AFE-VFDs, to provide grid support services like frequency and voltage control by modulating reactive power. These "communication-free" local control strategies are crucial for fast-acting stability enhancement. However, they address the symptoms of power imbalances (frequency/voltage deviations) on a sub-second to seconds timescale. They do not address the root cause of large power imbalances, which is the mismatch between wind generation and load over a timescale of minutes to hours.

This is precisely where forecasting becomes indispensable. A reliable multi-step forecast (e.g., 10, 20, 30 minutes ahead) allows the PMS to move from a reactive to a proactive stance. Instead of waiting for a frequency drop to occur and then reacting, the PMS can use the forecast to anticipate a future power deficit and make a strategic, unhurried decision to start a standby generator well in advance. This proactive dispatch strategy, enabled by forecasting, is the key to mitigating the risks of blackouts and avoiding costly emergency load shedding, thereby ensuring the operational reliability of the entire platform (8).

3.2.2 Multi-Objective Optimization in Energy Management

The operational decision-making in hybrid microgrids and FPSOs is inherently a multi-objective optimization problem, where multiple conflicting objectives must be balanced simultaneously (53; 54). The primary objectives typically include:

1. **Minimizing operational cost:** Reducing fuel consumption and generator start-stop cycles.
2. **Maximizing reliability:** Maintaining adequate spinning reserve and avoiding blackouts.
3. **Minimizing emissions:** Reducing CO₂ output to meet environmental targets.
4. **Minimizing equipment wear:** Reducing frequent cycling of generators to extend equipment lifespan.

These objectives are fundamentally conflicting: maximizing reliability often requires running additional generators (increasing cost and emissions), while minimizing cost may reduce safety margins. The literature on microgrid energy manage-

ment has extensively explored multi-objective optimization techniques to address this challenge (53; 55).

The most common approach is Pareto optimization, which seeks to identify the set of Pareto-optimal solutions—configurations where no objective can be improved without degrading at least one other objective (56; 57). Evolutionary algorithms such as Multi-Objective Particle Swarm Optimization (MOPSO), Non-dominated Sorting Genetic Algorithm II (NSGA-II), and Multi-Objective Evolutionary Algorithms (MOEAs) are widely used to explore the Pareto frontier (53; 58).

However, a key limitation of these approaches in real-time operational contexts is their computational intensity. Solving a full multi-objective optimization problem at every decision point (e.g., every 10 minutes) is often impractical for online deployment. This motivates the development of parametric optimization strategies, where the optimization is performed offline across a range of operational scenarios and cost parameters, and the results are encoded into a lookup table or decision rule that can be queried rapidly during online operation. This dissertation adopts this parametric approach in the strategy optimizer component, enabling operators to explore the cost-reliability trade-off space interactively without requiring real-time solution of complex optimization problems.

3.2.3 Decision Support Systems and Explainability

While the technical literature on forecasting and power management is extensive, there is a notable scarcity of research addressing the human-machine interface aspect of decision support in offshore operations. Most studies treat forecasting as a purely technical problem, delivering model outputs (predictions, uncertainties) without considering how operators will interpret and act upon this information in practice.

Recent work in the broader field of Explainable Artificial Intelligence (XAI) has demonstrated that the trustworthiness and adoption of AI-based decision support systems critically depend on their ability to provide transparent, interpretable explanations for their recommendations (59). This is particularly important in safety-critical domains like healthcare and industrial operations, where operators must retain ultimate decision authority and accountability.

In the context of wind-integrated FPSOs, operators face a continuous stream of complex decisions under uncertainty: when to start backup generators, which risk threshold to apply, how to balance competing objectives (safety vs. cost). Traditional interfaces that simply display numerical forecasts and leave all interpretation to the operator are insufficient. There is a clear need for intelligent decision support systems that can:

1. Synthesize multi-dimensional information (forecasts, uncertainties, costs, constraints) into clear recommendations
2. Explain the reasoning behind recommendations in natural language
3. Adapt to different operational contexts and operator preferences
4. Maintain transparency to build operator trust and enable informed override decisions

3.2.4 Large Language Models in Industrial Decision-Making

The recent emergence of Large Language Models (LLMs) has opened new possibilities for human-AI interaction in industrial settings. LLMs such as GPT-5, Claude, and domain-specific fine-tuned variants have demonstrated remarkable capabilities in natural language understanding, reasoning, and generation (60; 61).

In manufacturing contexts, LLMs have been applied to tasks such as process optimization, quality control, predictive maintenance, and supply chain management (60). The key advantage of LLMs in these applications is their ability to serve as a natural language interface to complex technical systems, enabling domain experts to query, explore, and understand AI-driven recommendations without requiring specialized technical knowledge (31).

However, the application of LLMs in safety-critical industrial environments also presents significant challenges:

- **Hallucination:** LLMs can generate plausible-sounding but factually incorrect information, which is unacceptable in safety-critical contexts.
- **Reliability and consistency:** Ensuring that the LLM provides consistent recommendations across similar scenarios.
- **Latency:** Real-time decision support requires low-latency responses, which can be challenging for large models.
- **Integration with domain knowledge:** LLMs must be grounded in accurate, up-to-date domain-specific information rather than relying solely on pre-training knowledge.

To address these challenges, this dissertation employs a Retrieval-Augmented Generation (RAG) approach, where the LLM is provided with real-time context (current forecasts, system state, historical data) retrieved from the operational database. This grounds the LLM's responses in factual, current information and significantly reduces the risk of hallucination. Additionally, the LLM is configured with explicit

safety constraints and is positioned as an advisory tool that enhances rather than replaces human decision-making.

3.3 Related Work on Wind-Integrated Offshore Platforms

While the broader literature on wind forecasting and microgrid management is extensive, research specifically addressing the integration of wind power into offshore oil and gas platforms is relatively limited. The pioneering work by (17) first explored the technical and economic feasibility of this concept, identifying the potential for significant fuel savings and emissions reductions. More recently, the LEOGO (Low Emission Oil and Gas Open) platform developed by (8) has provided a comprehensive reference model for studying renewable integration in offshore systems, including detailed electrical models, control strategies, and operational scenarios.

Studies by (4), (1), and (3) have advanced the understanding of the electrical and control aspects of wind-FPSO integration, particularly focusing on power quality, frequency stability, and the role of power electronics in providing grid support. The economic and environmental analysis by (16) has quantified the potential for carbon credit generation and emissions reduction in the Brazilian context, demonstrating reductions ranging from 23% to 76% depending on operational strategy.

However, none of these prior works have addressed the complete decision support framework proposed in this dissertation. Existing research has focused primarily on either the forecasting problem in isolation or the power system control problem in isolation, without integrating these layers with risk assessment, economic optimization, and explainable human-machine interfaces. This dissertation fills this gap by presenting the first comprehensive framework that unifies probabilistic forecasting, operational risk quantification (ORI), parametric strategy optimization, and LLM-powered explainability into a single deployable system tailored specifically for wind-integrated FPSO operations.

3.4 Research Gaps and Positioning of This Work

The literature review reveals several critical gaps that this dissertation addresses:

- 1. Gap 1: Unified Multivariate Forecasting of Decomposed Components.** While decomposition-based methods are well-established, most studies forecast each IMF independently and then sum the results. This ignores potential cross-correlations between components. This work employs TFT as

a unified multivariate predictor that can learn from all IMFs simultaneously, potentially capturing inter-component dependencies.

2. Gap 2: Probabilistic Forecasting for Operational Risk Assessment.

Few studies bridge the gap between probabilistic forecasts and operational decision-making. This work introduces the Operational Risk Index (ORI), a derived metric that translates quantile forecasts into actionable risk levels tailored to FPSO operations, addressing the cognitive burden of interpreting multiple quantiles across multiple horizons.

3. Gap 3: Integrated Decision Support Framework. Existing research

treats forecasting, risk assessment, and operational optimization as separate problems. This work integrates all three layers into a unified framework, demonstrating how they complement each other in practice and providing a holistic solution rather than isolated components.

4. Gap 4: Explainable AI for Offshore Operations. There is virtually

no prior work on applying LLM-based conversational AI to offshore power management. This work pioneers the use of explainable AI to enhance operator trust and decision quality in this safety-critical domain, employing RAG to ground the LLM in real-time operational data.

5. Gap 5: Parametric Optimization for Real-Time Deployment. While

multi-objective optimization is well-studied in the literature, most approaches are computationally intensive and unsuitable for real-time online deployment. This work develops a parametric optimization strategy that pre-computes the Pareto frontier offline and enables rapid online querying, making it practical for operational use.

By addressing these gaps, this dissertation advances the state-of-the-art in wind power forecasting for isolated offshore systems and provides a practical, deployable solution that has been validated through real-world operational scenarios. The following chapters detail the mathematical framework (Chapter 4), the complete methodology and implementation (Chapter 5), and the experimental validation (Chapter 6) of the proposed integrated decision support system.

Chapter 4

Deep Learning Techniques and Methodologies

Building upon the theoretical foundations and literature review from the preceding chapters, this chapter delves into the mathematical and architectural details of the machine learning and deep learning models employed in this dissertation. The focus is on providing a rigorous description of the techniques used for time series forecasting, with a particular emphasis on those capable of handling the complex, non-linear dynamics of wind power generation. This chapter is structured to serve as a comprehensive technical reference, detailing not only the predictive models themselves but also the critical methodologies for signal decomposition, training, validation, and optimization that ensure the development of robust and generalizable solutions.

4.1 Supervised Learning Framework for Time Series

Supervised learning constitutes a paradigm in machine learning where an algorithm learns a mapping function $f : \mathcal{X} \rightarrow \mathcal{Y}$ from a labeled training dataset $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^N$. In the context of time series forecasting, this framework is adapted using a **sliding window** approach. A window of a fixed length, the *look-back* period (L), is slid across the time series. At each time step t , the sequence of the last L observations becomes the input feature vector $\mathbf{x}_t = [y_{t-L+1}, \dots, y_t]$, and the subsequent H observations become the target vector $\mathbf{y}_t = [y_{t+1}, \dots, y_{t+H}]$, where H is the forecast horizon. This process, illustrated in Figure 4.1, transforms the forecasting problem into a supervised learning task.

The choice of L and H is application-specific. For the FPSO wind forecasting problem, H corresponds to the three critical decision horizons ($t+10$, $t+20$, $t+30$ minutes), while L is determined through hyperparameter optimization to balance the trade-off between capturing sufficient historical context and avoiding overfitting

to noise.

Single-Scale Sliding Window

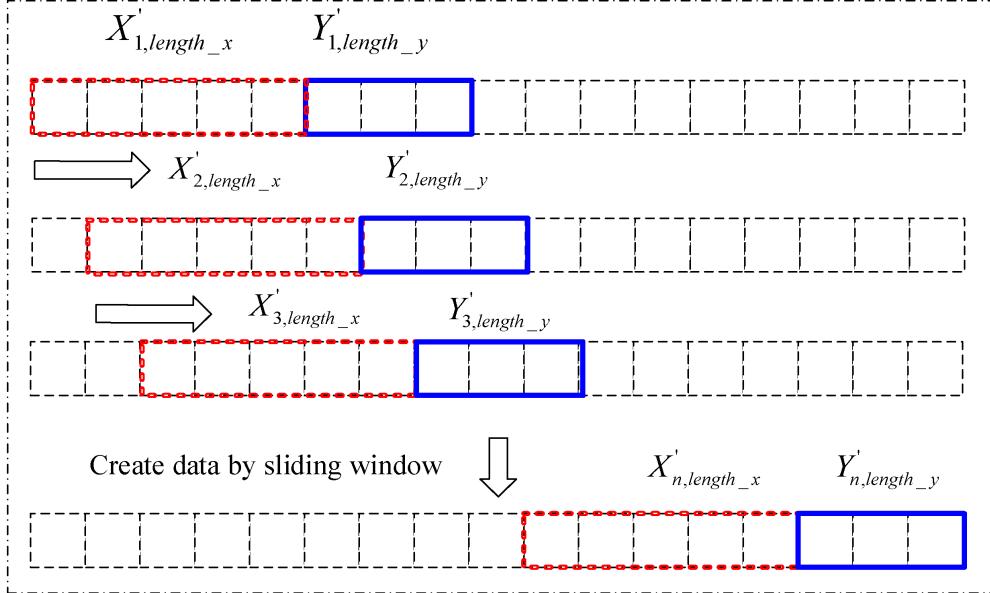


Figure 4.1: Sliding window technique for time series forecasting. Look-back window (red) provides input features, forecast horizon (blue) provides target labels (adapted from (9)).

4.2 Deep Learning Architectures for Sequential Data

Deep Learning (DL) models have become the state-of-the-art for time series forecasting due to their ability to automatically learn hierarchical features and complex non-linear dependencies from raw data without requiring manual feature engineering.

4.2.1 Multi Layer Perceptron (MLP)

The Multi-Layer Perceptron is the archetypal feedforward neural network. It consists of an input layer, one or more hidden layers, and an output layer, as shown in Figure 4.2. The transformation from layer $l - 1$ to layer l is defined by:

$$\mathbf{a}^{(l)} = \sigma(\mathbf{z}^{(l)}) = \sigma(\mathbf{W}^{(l)} \mathbf{a}^{(l-1)} + \mathbf{b}^{(l)}) \quad (4.1)$$

where $\mathbf{a}^{(l)}$ is the activation vector at layer l , $\mathbf{W}^{(l)}$ is the weight matrix, $\mathbf{b}^{(l)}$ is the bias vector, and $\sigma(\cdot)$ is a non-linear activation function. Common choices include the Rectified Linear Unit (ReLU), $\sigma(z) = \max(0, z)$, which introduces non-linearity

while maintaining computational efficiency and mitigating the vanishing gradient problem.

While powerful for capturing non-linear relationships, standard MLPs treat each input window independently and do not inherently model the temporal order or sequential dependencies within the time series. This limitation motivates the use of specialized recurrent architectures.

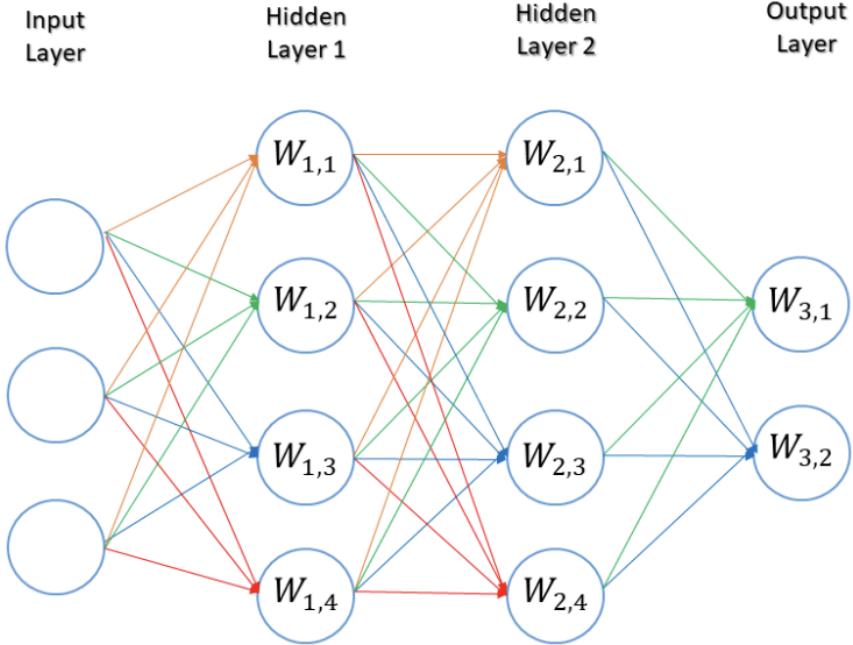


Figure 4.2: Basic architecture of a Multi-Layer Perceptron (MLP) with an input layer, two hidden layers, and an output layer. Information flows forward through weighted connections.

4.2.2 Recurrent Neural Network (RNN)

Recurrent Neural Networks (RNNs) are specifically designed for sequential data. Unlike feedforward networks, RNNs maintain a hidden state \mathbf{h}_t that is updated at each time step, allowing the network to have "memory" of past information. The hidden state at time t is computed as:

$$\mathbf{h}_t = \sigma_h(\mathbf{W}_{hh}\mathbf{h}_{t-1} + \mathbf{W}_{xh}\mathbf{x}_t + \mathbf{b}_h) \quad (4.2)$$

where \mathbf{h}_{t-1} is the previous hidden state, \mathbf{x}_t is the current input, \mathbf{W}_{hh} and \mathbf{W}_{xh} are weight matrices, \mathbf{b}_h is the bias vector, and σ_h is typically a tanh activation function. The output at time t is computed from the hidden state: $\mathbf{y}_t = \mathbf{W}_{hy}\mathbf{h}_t + \mathbf{b}_y$.

This recurrent loop, shown in Figure 4.3, enables RNNs to capture temporal

dynamics and dependencies across time steps. However, simple RNNs suffer from the **vanishing gradient problem** during backpropagation through time (BPTT), where gradients exponentially decay as they propagate backward through many time steps, making it difficult to learn long-term dependencies (45).

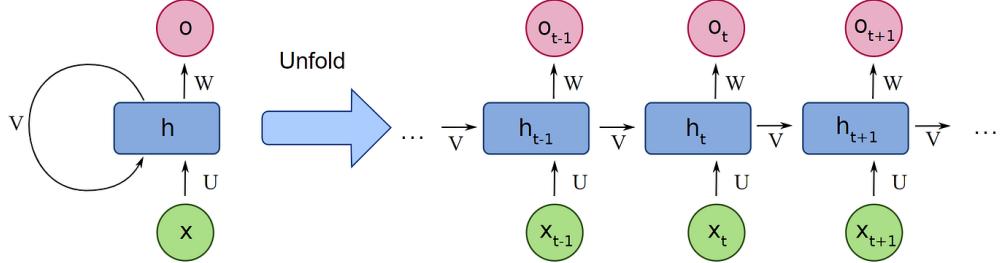


Figure 4.3: The unrolled structure of a Recurrent Neural Network (RNN), showing how the hidden state h_t is passed from one time step to the next, creating a temporal dependency chain.

4.2.3 Gated Recurrent Architectures

To address the limitations of simple RNNs, gated architectures introduce mechanisms to control the flow of information, enabling the network to learn which information to retain, update, or discard at each time step.

Long Short-Term Memory (LSTM)

The Long Short-Term Memory (LSTM) network (45) is an advanced RNN that mitigates the vanishing gradient problem using a sophisticated memory cell architecture. As illustrated in Figure 4.4, this cell includes three gates—**forget gate** (f_t), **input gate** (i_t), and **output gate** (o_t)—that regulate the flow of information into and out of a persistent cell state, c_t . This design allows the network to selectively remember or forget information over arbitrarily long time periods.

The LSTM operations at each time step are defined by the following equations:

$$f_t = \sigma_g(\mathbf{W}_f \mathbf{x}_t + \mathbf{U}_f h_{t-1} + \mathbf{b}_f) \quad (\text{Forget Gate}) \quad (4.3)$$

$$i_t = \sigma_g(\mathbf{W}_i \mathbf{x}_t + \mathbf{U}_i h_{t-1} + \mathbf{b}_i) \quad (\text{Input Gate}) \quad (4.4)$$

$$o_t = \sigma_g(\mathbf{W}_o \mathbf{x}_t + \mathbf{U}_o h_{t-1} + \mathbf{b}_o) \quad (\text{Output Gate}) \quad (4.5)$$

$$\tilde{c}_t = \sigma_c(\mathbf{W}_c \mathbf{x}_t + \mathbf{U}_c h_{t-1} + \mathbf{b}_c) \quad (\text{Candidate Cell State}) \quad (4.6)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (\text{Cell State Update}) \quad (4.7)$$

$$h_t = o_t \odot \sigma_h(c_t) \quad (\text{Hidden State Update}) \quad (4.8)$$

where σ_g is the sigmoid function (outputting values in $[0, 1]$ to act as gates), σ_c and σ_h are typically tanh functions, and \odot denotes element-wise multiplication.

The forget gate (4.3) determines what information from the previous cell state to discard. The input gate (4.4) controls what new information to add to the cell state. The cell state (4.7) is updated by forgetting selected information and adding new candidate information. Finally, the output gate (4.5) determines what information from the cell state to output as the hidden state (4.8).

Bidirectional Long Short-Term Memory (Bi-LSTM) A Bidirectional LSTM (Bi-LSTM) further enhances this architecture by processing the sequence in both forward and backward directions simultaneously (15). At each time step t , the Bi-LSTM maintains two hidden states: $\vec{\mathbf{h}}_t$ (forward) and $\overleftarrow{\mathbf{h}}_t$ (backward). The final hidden representation is the concatenation: $\mathbf{h}_t = [\vec{\mathbf{h}}_t; \overleftarrow{\mathbf{h}}_t]$. This allows the model to capture both past and future context, which can be particularly beneficial when the entire input sequence is available before making predictions, as is the case in offline forecasting scenarios.

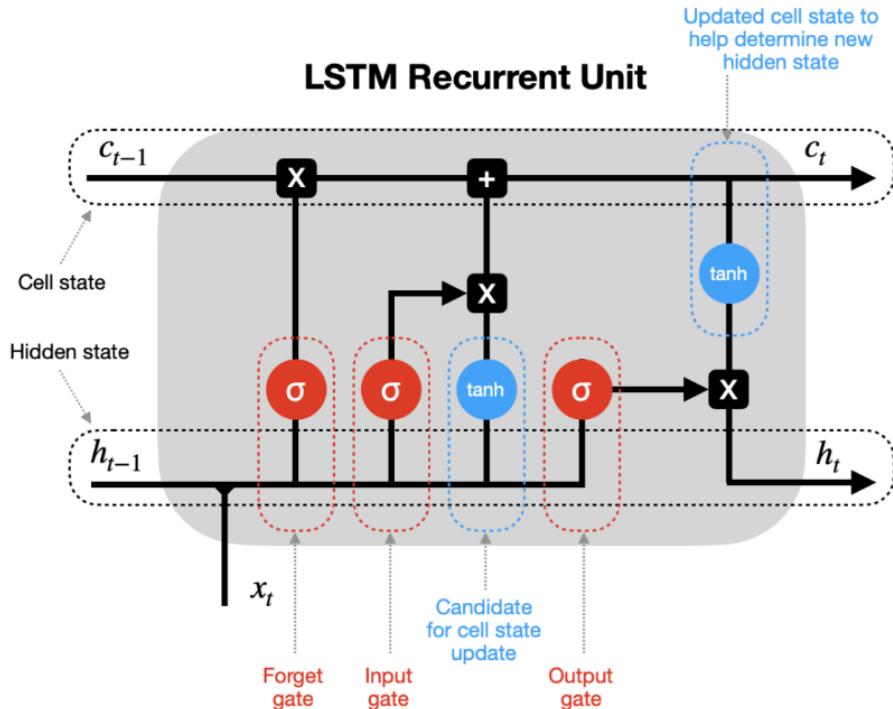


Figure 4.4: Detailed architecture of a Long Short-Term Memory (LSTM) cell, illustrating the forget gate (f_t), input gate (i_t), and output gate (o_t) that regulate the information flow through the cell state (c_t) and hidden state (h_t).

Gated Recurrent Unit (GRU)

The Gated Recurrent Unit (GRU), proposed by (46), offers a similar capability to the LSTM but with a simpler architecture, as illustrated in Figure 4.5. It merges the forget and input gates into a single "update gate" (\mathbf{z}_t) and introduces a "reset gate" (\mathbf{r}_t), eliminating the separate cell state. This results in a more computationally efficient model with fewer parameters that often achieves comparable performance to LSTMs. The GRU equations are:

$$\mathbf{z}_t = \sigma_g(\mathbf{W}_z \mathbf{x}_t + \mathbf{U}_z \mathbf{h}_{t-1} + \mathbf{b}_z) \quad (\text{Update Gate}) \quad (4.9)$$

$$\mathbf{r}_t = \sigma_g(\mathbf{W}_r \mathbf{x}_t + \mathbf{U}_r \mathbf{h}_{t-1} + \mathbf{b}_r) \quad (\text{Reset Gate}) \quad (4.10)$$

$$\tilde{\mathbf{h}}_t = \sigma_h(\mathbf{W}_h \mathbf{x}_t + \mathbf{U}_h (\mathbf{r}_t \odot \mathbf{h}_{t-1}) + \mathbf{b}_h) \quad (\text{Candidate Hidden State}) \quad (4.11)$$

$$\mathbf{h}_t = (1 - \mathbf{z}_t) \odot \mathbf{h}_{t-1} + \mathbf{z}_t \odot \tilde{\mathbf{h}}_t \quad (\text{Hidden State Update}) \quad (4.12)$$

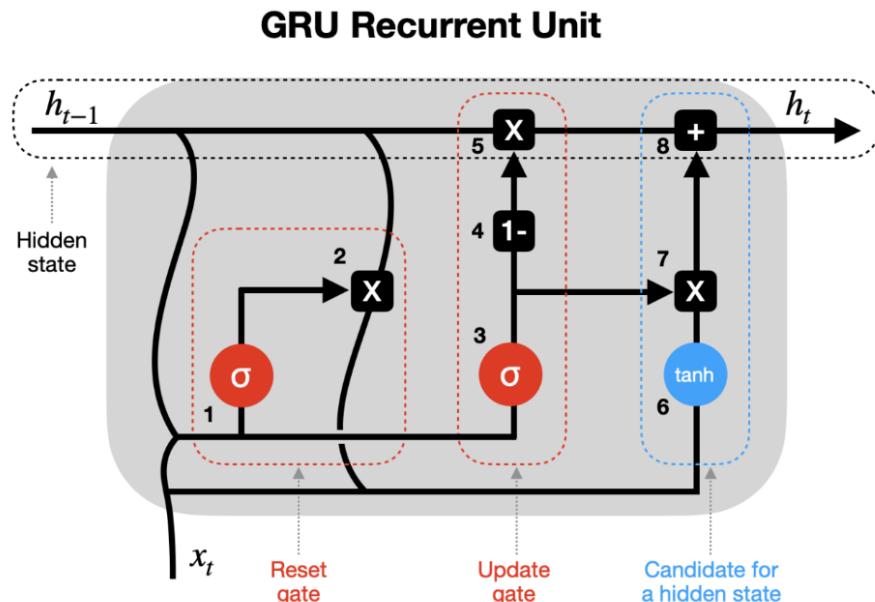


Figure 4.5: Architecture of a Gated Recurrent Unit (GRU) cell, a simpler alternative to the LSTM cell featuring update gate (\mathbf{z}_t) and reset gate (\mathbf{r}_t) mechanisms.

4.2.4 Convolutional Neural Network (CNN) for Time Series

1D Convolutional Neural Networks (CNNs) apply learnable filters (kernels) across the input sequence to detect local patterns and features. A 1D convolution operation with kernel \mathbf{w} of size k applied to input sequence \mathbf{x} produces a feature map:

$$(\mathbf{x} * \mathbf{w})[i] = \sum_{j=0}^{k-1} \mathbf{x}[i+j] \cdot \mathbf{w}[j] \quad (4.13)$$

As shown in Figure 4.6, the filter slides over the time series, creating feature maps that highlight where specific patterns (such as ramp events or oscillations) occur. Stacking multiple convolutional layers with increasing filter sizes allows the network to build a hierarchical representation of features at different temporal scales. CNNs are particularly effective as feature extractors in hybrid models, where they can be combined with recurrent layers to capture both local patterns and long-range dependencies (27).

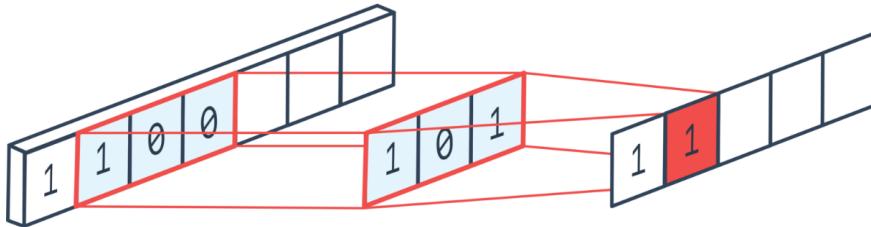


Figure 4.6: Illustration of a 1D convolution operation on a time series. A filter (kernel) of size k slides across the input sequence to produce a feature map that captures local temporal patterns.

4.2.5 Attention Mechanism and Transformer Models

The Attention Mechanism (10) represents a paradigm shift in sequence modeling, allowing models to dynamically weigh the importance of different parts of the input sequence when making predictions. Unlike recurrent architectures that process sequences sequentially, attention-based models can attend to any position in the sequence directly, enabling parallel computation and more effective modeling of long-range dependencies.

The Transformer model relies entirely on self-attention mechanisms, dispensing with recurrence altogether. The core component is the Scaled Dot-Product Attention, defined as:

$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V \quad (4.14)$$

where the input sequence is linearly transformed into three matrices: Queries (Q), Keys (K), and Values (V), each of dimension d_k . The dot product of a query with all keys determines the attention scores (weights) for the values. The scaling factor $\sqrt{d_k}$ prevents the dot products from growing too large in magnitude, which would push the softmax function into regions with extremely small gradients.

Multi-Head Attention extends this by computing attention in parallel across multiple representation subspaces:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^O \quad (4.15)$$

where each $\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V)$ learns to attend to different aspects of the input. This mechanism is illustrated in Figure 4.7.

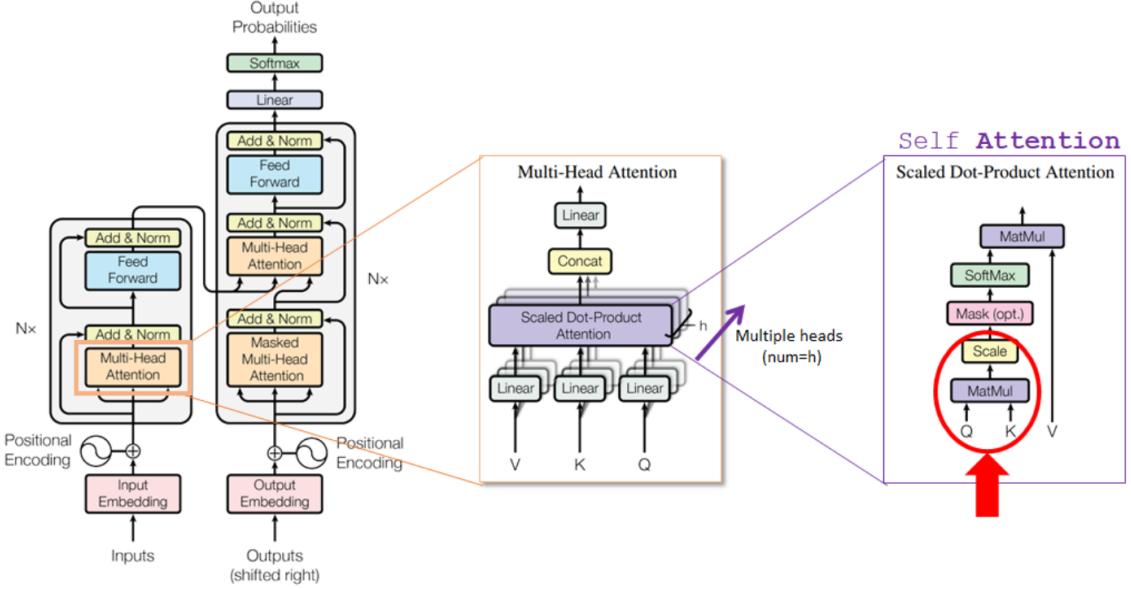


Figure 4.7: Scaled dot-product attention mechanism showing Queries, Keys, and Values computation with attention score weighting and Transformer architecture. (adapted from (10)).

4.2.6 Temporal Fusion Transformer (TFT)

The Temporal Fusion Transformer (TFT) (11) is a state-of-the-art architecture specifically designed for interpretable multi-horizon time series forecasting. Unlike generic Transformers, TFT incorporates specialized components tailored for time series data, including static covariates, known future inputs, and observed inputs. Its architecture, shown in Figure 4.8, consists of several key modules:

Variable Selection Network (VSN)

TFT employs separate Variable Selection Networks for each input type (static, known future, observed) to identify the most relevant features at each time step. The VSN uses a Gated Residual Network (GRN) followed by a softmax layer to compute feature importance weights:

$$\mathbf{v}_t = \text{VSN}(\mathbf{x}_t) = \sum_{i=1}^n w_i \cdot \text{GRN}(\mathbf{x}_{t,i}) \quad (4.16)$$

where w_i are learned importance weights and $\mathbf{x}_{t,i}$ are individual input features. This provides interpretability by revealing which variables are most influential for the forecast at each time step.

Gated Residual Network (GRN)

GRNs are the fundamental building blocks of TFT, used throughout the architecture for non-linear processing. A GRN applies a gating mechanism to control the extent to which the network modifies its input:

$$\text{GRN}(\mathbf{a}) = \text{LayerNorm}(\mathbf{a} + \text{GLU}(\eta_1)) \quad (4.17)$$

where $\eta_1 = \mathbf{W}_1\mathbf{a} + \mathbf{b}_1$ is an intermediate representation, and GLU (Gated Linear Unit) is defined as:

$$\text{GLU}(\eta_1) = \sigma(\mathbf{W}_2\eta_1 + \mathbf{b}_2) \odot (\mathbf{W}_3\eta_1 + \mathbf{b}_3) \quad (4.18)$$

The gating mechanism allows the network to suppress or amplify information flow adaptively, and the residual connection ($\mathbf{a}+$) facilitates gradient flow during training.

Temporal Self-Attention

After variable selection and encoding, TFT applies a multi-head self-attention mechanism to capture long-range temporal dependencies. Unlike standard Transformers, TFT uses an interpretable multi-head attention where attention weights are explicitly designed to be inspected, revealing which past time steps are most influential for each forecast horizon.

Multi-Quantile Output

TFT natively supports probabilistic forecasting through a multi-quantile output layer. Instead of predicting a single value, the model simultaneously predicts multiple quantiles $(\tau_1, \tau_2, \dots, \tau_q)$ of the conditional distribution (62):

$$\hat{\mathbf{y}}_\tau = \text{Linear}_\tau(\mathbf{h}_{\text{final}}) \quad (4.19)$$

where separate output heads are trained using the quantile loss (pinball loss) for each quantile τ . This enables the model to provide prediction intervals and uncertainty estimates directly, which is essential for risk-aware decision-making in the FPSO application.

Adaptation for the Hybrid Framework: For the **CEEMDAN-EWT-TFT Aggregator** model developed in this dissertation, a streamlined architecture inspired by the TFT was implemented. Given that the model's inputs consist solely of the decomposed IMFs (a single type of past observed variable), the complex

encoder structure for static and future-known variables was not applicable. Furthermore, initial experiments demonstrated that the powerful feature engineering provided by the CEEMDAN-EWT decomposition rendered the explicit Variable Selection Network (VSN) redundant, with the simplified model (without VSN) achieving superior empirical results. Therefore, the final architecture integrates the core TFT components—Gated Residual Networks, temporal self-attention, and a multi-quantile output layer—to create a high-performance predictor tailored specifically for processing the decomposed time series, a decision validated by the experimental results presented in Chapter 6.

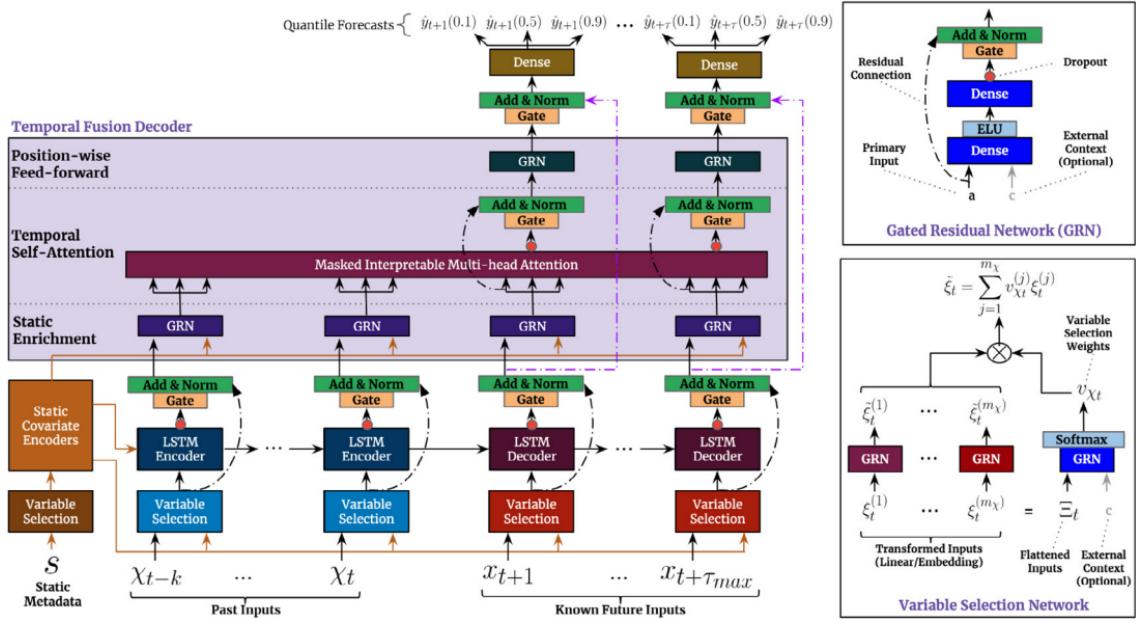


Figure 4.8: Complete architecture of the Temporal Fusion Transformer (TFT), showing Variable Selection Networks, Gated Residual Networks, temporal self-attention layers, and multi-quantile output heads (adapted from (11)).

4.3 Signal Decomposition Techniques

As established in Chapter 3, the non-stationary and multi-scale nature of wind power time series poses significant challenges for direct forecasting. Signal decomposition techniques address this by decomposing the complex original signal into a set of simpler, more stationary components (modes) that can be forecast more accurately. This section provides the mathematical foundations of the two decomposition methods employed in this work.

4.3.1 Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN)

Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN), proposed by (25), is an advanced signal processing technique designed to overcome the mode-mixing problem inherent in the original Empirical Mode Decomposition (EMD). The algorithm iteratively decomposes a non-stationary signal $x(t)$ into a finite set of Intrinsic Mode Functions (IMFs), each representing oscillations at a distinct time scale, and a final residual trend.

The process, illustrated in the flowchart in Figure 4.9, begins by creating an ensemble of signals by adding realizations of white Gaussian noise to the original signal. The first IMF is then computed as the ensemble average of the first EMD mode extracted from these noisy signals:

$$\text{IMF}_1(t) = \frac{1}{I} \sum_{i=1}^I E_1(x(t) + \beta_0 \epsilon^{(i)}(t)) \quad (4.20)$$

where $E_1(\cdot)$ is the operator that extracts the first EMD mode, $\epsilon^{(i)}(t)$ is the i -th realization of white noise, β_0 is the noise amplitude coefficient, and I is the ensemble size.

After calculating the first IMF, a residual is computed: $r_1(t) = x(t) - \text{IMF}_1(t)$. The process then continues iteratively. For each subsequent mode k , a new residual is constructed by adding a specific noise component, and the next IMF is extracted as the ensemble average of the first EMD mode of this new residual:

$$\text{IMF}_k(t) = \frac{1}{I} \sum_{i=1}^I E_1(r_{k-1}(t) + \beta_{k-1} E_k(\epsilon^{(i)}(t))) \quad (4.21)$$

This iterative procedure is repeated until the final residual, $r_K(t)$, becomes monotonic or contains fewer than three extrema. The original signal can then be perfectly reconstructed by summing all the extracted components:

$$x(t) = \sum_{k=1}^K \text{IMF}_k(t) + r_K(t) \quad (4.22)$$

where K is the total number of IMFs. Each IMF captures a specific frequency band, from the highest frequencies in IMF_1 to progressively lower frequencies in subsequent modes, making the decomposition highly suitable for analyzing complex, non-stationary time series like wind power.

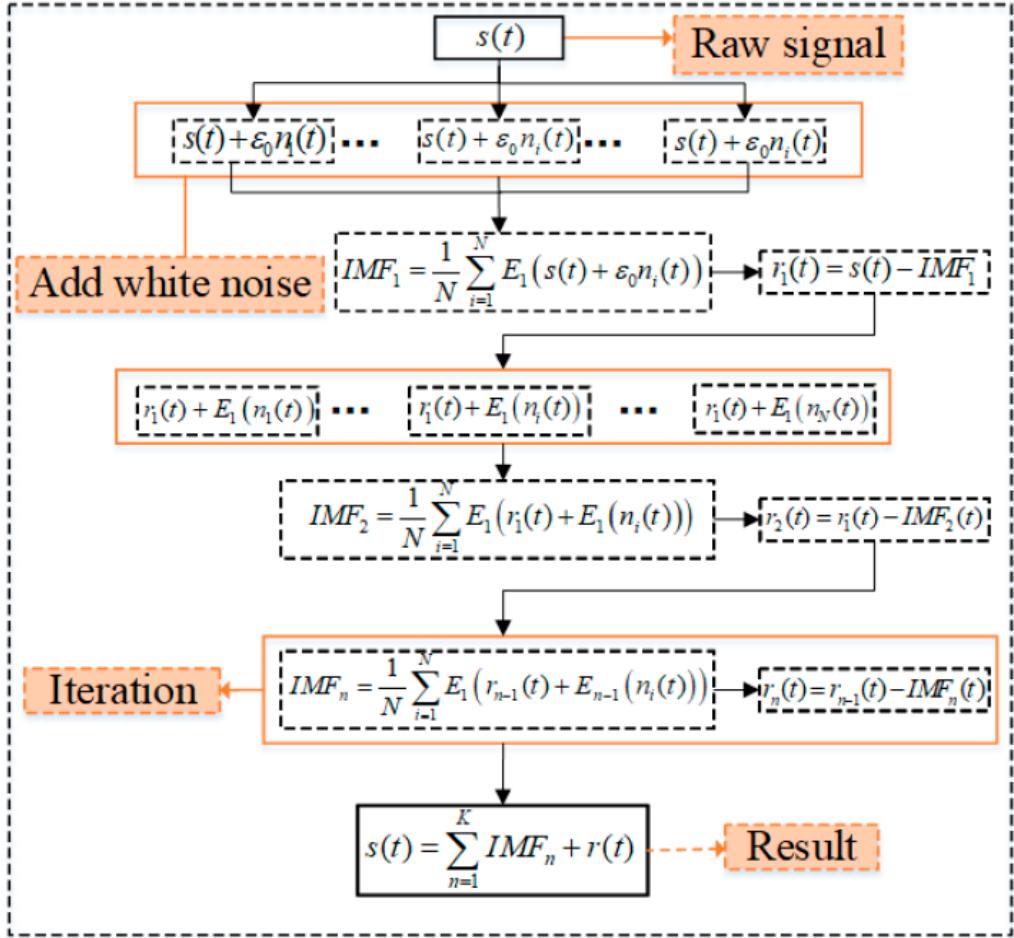


Figure 4.9: Flowchart of the CEEMDAN algorithm, illustrating the iterative process of extracting Intrinsic Mode Functions (IMFs) by adding adaptive noise at each stage (adapted from: (5)).

4.3.2 Empirical Wavelet Transform (EWT)

EWT (26) combines the adaptivity of EMD with the mathematical rigor of wavelet theory. Instead of an iterative sifting process, EWT constructs an adaptive wavelet filter bank based on the signal's Fourier spectrum.

Algorithm

The EWT algorithm consists of the following steps:

1. Compute the Fourier spectrum of the signal $x(t)$:

$$\hat{x}(\omega) = \mathcal{F}\{x(t)\} \quad (4.23)$$

2. Detect local maxima in $|\hat{x}(\omega)|$ to identify the dominant frequency components.

These maxima define boundaries ω_n that partition the spectrum into N segments: $[0, \omega_1], [\omega_1, \omega_2], \dots, [\omega_{N-1}, \pi]$.

3. Construct empirical scaling and wavelet functions. For each segment n , define a bandpass filter $\hat{\psi}_n(\omega)$ using a transition function (typically a Meyer-style wavelet):

$$\hat{\psi}_n(\omega) = \begin{cases} 1 & \text{if } \omega \in [\omega_{n-1} + \tau_n, \omega_n - \tau_n] \\ \cos \left[\frac{\pi}{2} \beta \left(\frac{1}{2\tau_n} (\omega - \omega_{n-1} - \tau_n) \right) \right] & \text{if } \omega \in [\omega_{n-1}, \omega_{n-1} + \tau_n] \\ \sin \left[\frac{\pi}{2} \beta \left(\frac{1}{2\tau_n} (\omega - \omega_n + \tau_n) \right) \right] & \text{if } \omega \in [\omega_n - \tau_n, \omega_n] \\ 0 & \text{otherwise} \end{cases} \quad (4.24)$$

where τ_n is the transition width and $\beta(x)$ is a smooth transition function. A common choice is the polynomial:

$$\beta(x) = x^4(35 - 84x + 70x^2 - 20x^3), \quad x \in [0, 1] \quad (4.25)$$

which ensures C^∞ smoothness at the boundaries.

4. Extract each mode by filtering in the Fourier domain and transforming back:

$$f_n(t) = \mathcal{F}^{-1}\{\hat{\psi}_n(\omega) \cdot \hat{x}(\omega)\} \quad (4.26)$$

The final decomposition is:

$$x(t) = \sum_{n=1}^N f_n(t) \quad (4.27)$$

EWT provides a more mathematically rigorous decomposition than CEEMDAN, with better-defined frequency bands and reduced mode mixing.

4.3.3 Hybrid CEEMDAN-EWT Decomposition Strategy

This dissertation employs a two-stage hybrid decomposition strategy that combines the strengths of both methods:

1. **Primary Decomposition (CEEMDAN):** The original wind power signal is first decomposed using CEEMDAN to obtain a set of IMFs: $\{\text{IMF}_1, \text{IMF}_2, \dots, \text{IMF}_K, r_K\}$.
2. **Entropy-Based Classification:** Each IMF is classified as high-frequency or low-frequency based on its sample entropy. High-frequency IMFs (typically IMF_1 and IMF_2) contain noise and rapid fluctuations, while low-frequency IMFs contain slower oscillations and trends.

3. **Secondary Decomposition (EWT):** High-frequency IMFs are further decomposed using EWT to separate noise from meaningful high-frequency dynamics:

$$\text{IMF}_{\text{high}} = \sum_{j=1}^M f_j \quad (4.28)$$

4. **Final Component Set:** The final set of components for forecasting consists of the EWT-decomposed high-frequency modes and the original low-frequency IMFs and residual.

This hybrid strategy, illustrated in Figure 4.10, provides a more refined separation of multi-scale dynamics than either method alone, as demonstrated by (12) and (27).

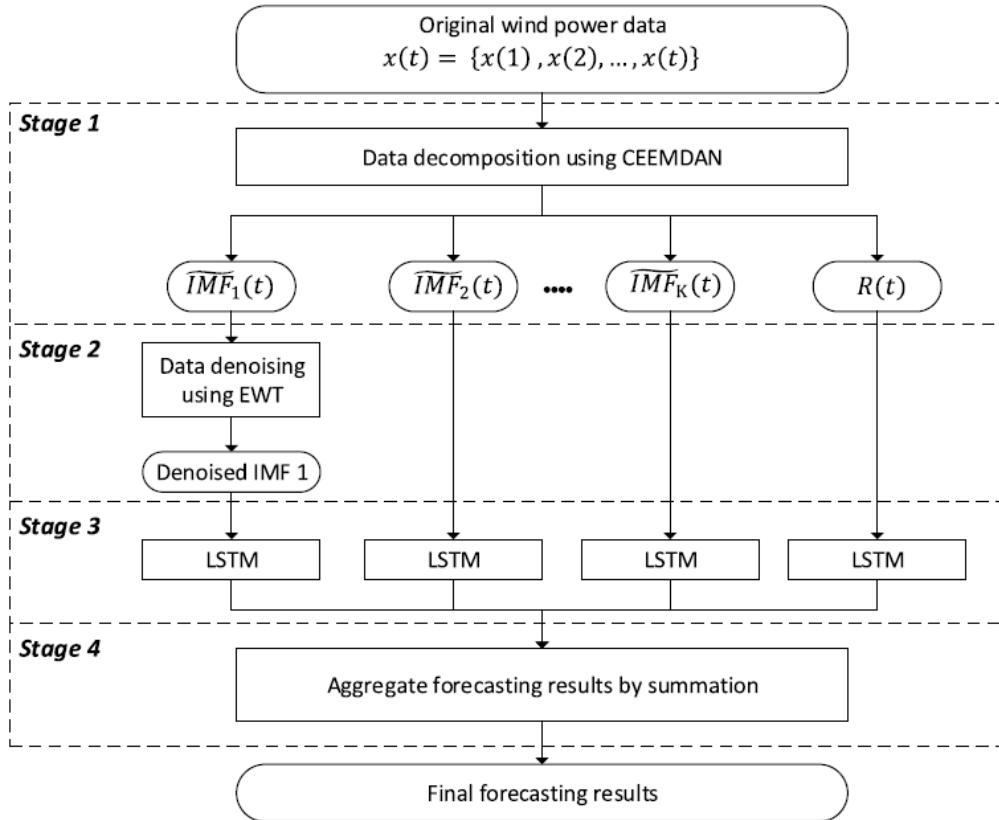


Figure 4.10: Hybrid CEEMDAN-EWT decomposition framework with entropy-based IMF separation and secondary EWT decomposition (adapted from: (12)).

4.4 Ensemble and Aggregation Strategies

After decomposing the wind power signal into multiple components, two fundamental approaches exist for generating the final forecast: independent forecasting with summation, or unified multivariate forecasting.

4.4.1 Independent Forecasting with Summation

The traditional "divide-and-conquer" approach trains a separate forecasting model for each decomposed component:

$$\hat{y}_k(t+h) = f_k(\text{IMF}_k(t-L:t)), \quad k = 1, \dots, K \quad (4.29)$$

where f_k is a dedicated model (e.g., LSTM) for component k . The final forecast is obtained by summing the individual forecasts:

$$\hat{y}(t+h) = \sum_{k=1}^K \hat{y}_k(t+h) + \hat{r}(t+h) \quad (4.30)$$

Advantages: Each model can specialize in the dynamics of its specific component. Computationally parallelizable.

Limitations: Ignores potential cross-correlations between components. Requires training and maintaining K separate models. Forecast errors from individual models can accumulate.

4.4.2 Unified Multivariate Forecasting (TFT as Aggregator)

An alternative approach, adopted in this work, treats all decomposed components as multivariate input features to a single unified predictor:

$$\hat{y}(t+h) = f_{\text{TFT}}(\text{IMF}_1(t-L:t), \text{IMF}_2(t-L:t), \dots, \text{IMF}_K(t-L:t), r(t-L:t)) \quad (4.31)$$

The Temporal Fusion Transformer is particularly well-suited for this role because:

- Its Variable Selection Networks can learn the relative importance of each component.
- Its attention mechanism can capture cross-component dependencies.
- It natively supports multi-horizon forecasting, eliminating the need for separate models for each horizon ($t+10, t+20, t+30$).
- It provides probabilistic forecasts (quantiles) directly.

Advantages: Learns inter-component relationships. Single model simplifies deployment. Native multi-horizon and probabilistic capabilities.

Limitations: Higher computational cost during training. Requires more training data to learn the joint distribution.

This dissertation empirically compares both approaches (CEEMDAN-EWT-LSTM with summation vs. CEEMDAN-EWT-TFT unified) in Chapter 6, demonstrating that the unified TFT aggregator achieves superior cost-benefit by leveraging cross-component information and no need to train many models for all IMFs.

4.5 Probabilistic Forecasting Framework

For risk management in an FPSO, a single point forecast is insufficient. It is crucial to quantify the uncertainty associated with the prediction. Probabilistic forecasting achieves this by predicting a full probability distribution or, more practically, specific quantiles of that distribution.

4.5.1 Quantile Regression and Pinball Loss Function

Quantile Regression is a technique used to estimate one or more conditional quantiles (τ) of a target variable. For a given quantile $\tau \in (0, 1)$, the τ -quantile q_τ is defined such that:

$$P(Y \leq q_\tau) = \tau \quad (4.32)$$

Instead of minimizing the mean squared error (which yields the conditional mean, equivalent to the 0.5-quantile), quantile regression minimizes the pinball loss function, L_τ :

$$L_\tau(y, \hat{y}) = \begin{cases} \tau(y - \hat{y}) & \text{if } y \geq \hat{y} \\ (1 - \tau)(\hat{y} - y) & \text{if } y < \hat{y} \end{cases} \quad (4.33)$$

This loss function asymmetrically penalizes errors: over-predictions and under-predictions are weighted differently according to the target quantile. For $\tau = 0.1$ (P10), the model is heavily penalized for under-predicting, encouraging conservative forecasts. For $\tau = 0.9$ (P90), the model is penalized for over-predicting, encouraging optimistic forecasts. As shown in Figure 4.11, the shape of the loss function guides the model to learn the desired conditional quantile.

For multi-quantile forecasting, the total loss is the sum of pinball losses across all target quantiles:

$$\mathcal{L}_{\text{total}} = \sum_{\tau \in \{\tau_1, \dots, \tau_Q\}} \sum_{i=1}^N L_\tau(y_i, \hat{y}_{i,\tau}) \quad (4.34)$$

In this work, three quantiles are predicted: $\tau \in \{0.1, 0.5, 0.9\}$, corresponding to pessimistic (P10), median (P50), and optimistic (P90) scenarios. These quantiles provide a practical representation of forecast uncertainty that can be directly used for risk assessment and decision-making in the FPSO operational context.

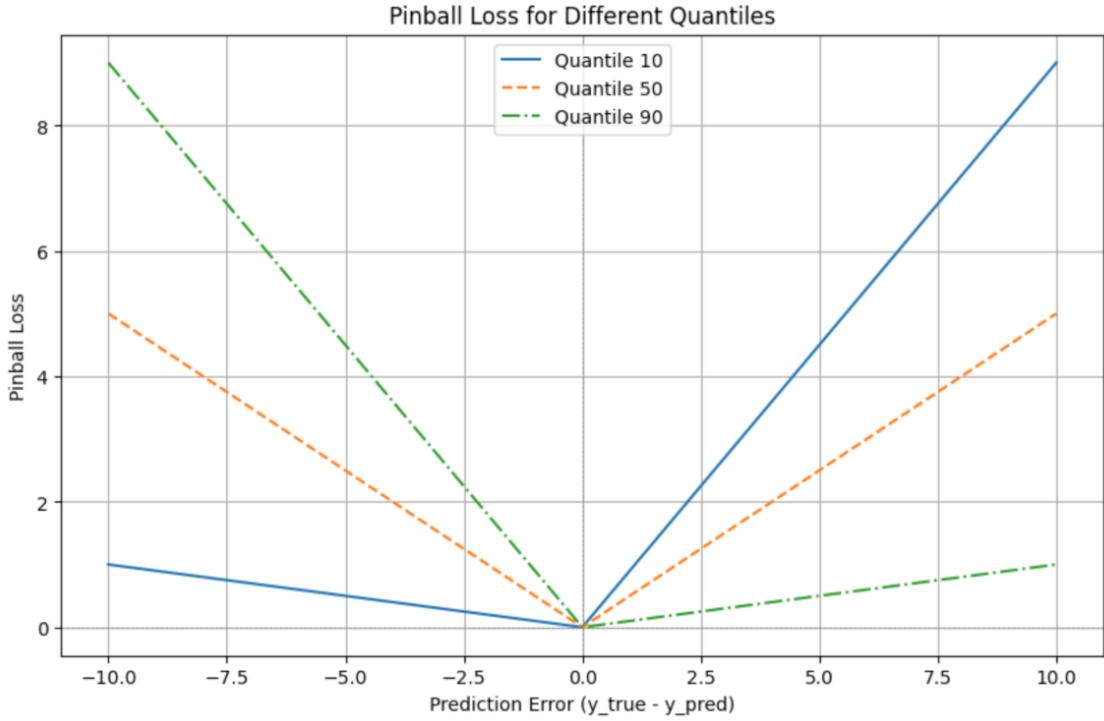


Figure 4.11: The Pinball Loss function for different quantiles ($\tau = 0.1, 0.5, 0.9$). The asymmetric penalty guides the model to learn the conditional quantile of the target variable, with different slopes for over-prediction and under-prediction errors.

The quantile loss serves both as the training objective (Equation 4.34) and as an evaluation metric for probabilistic forecast quality. Lower quantile loss values indicate better alignment between the predicted quantiles and the true conditional distribution of the target variable.

4.6 Model Training and Optimization

Training a deep neural network involves finding the optimal set of weights and biases that minimize a loss function on the training data. This is a complex, non-convex optimization problem that requires sophisticated techniques to solve effectively and to ensure the resulting model generalizes well to unseen data.

4.6.1 Data Preprocessing and Normalization

Before training, time series data must be preprocessed to ensure numerical stability and improve convergence.

Normalization Techniques

Two common normalization approaches are used in this work:

- **Standardization (Z-score normalization):** Transforms data to have zero mean and unit variance:

$$x_{\text{scaled}} = \frac{x - \mu}{\sigma} \quad (4.35)$$

where μ and σ are the mean and standard deviation computed on the training set. This is implemented using `StandardScaler` from scikit-learn.

- **Min-Max Scaling:** Transforms data to a fixed range, typically $[0, 1]$ or $[-1, 1]$:

$$x_{\text{scaled}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (4.36)$$

This is particularly useful for neural networks with sigmoid or tanh activation functions.

Critical Practice: The scaler must be fitted **only on the training data** and then applied to validation and test sets. Fitting on the entire dataset causes data leakage, where information from the future contaminates the training process, leading to overly optimistic performance estimates.

Train-Validation-Test Split

For time series, the split must preserve temporal order to avoid data leakage:

- **Training set:** Earliest 70% of data
- **Validation set:** Next 15% of data (used for hyperparameter tuning and early stopping)
- **Test set:** Final 15% of data (held out until final evaluation)

This ensures that the model is always evaluated on data that occurs strictly after its training data, simulating real-world deployment conditions.

4.6.2 Optimization Algorithms

The fundamental training algorithm is Gradient Descent, which iteratively updates the model parameters (θ) in the opposite direction of the gradient of the loss function ($J(\theta)$):

$$\theta \leftarrow \theta - \alpha \nabla_{\theta} J(\theta) \quad (4.37)$$

where α is the learning rate. However, vanilla gradient descent can be slow and inefficient.

This work employs the Adaptive Moment Estimation (Adam) optimizer (13), which computes adaptive learning rates for each parameter by maintaining exponentially decaying averages of past gradients (first moment m_t) and past squared

gradients (second moment v_t):

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (4.38)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (4.39)$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}, \quad \hat{v}_t = \frac{v_t}{1 - \beta_2^t} \quad (4.40)$$

$$\theta_t = \theta_{t-1} - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \quad (4.41)$$

where $g_t = \nabla_\theta J(\theta_{t-1})$ is the gradient, $\beta_1 = 0.9$ and $\beta_2 = 0.999$ are decay rates, and $\epsilon = 10^{-8}$ is a small constant for numerical stability. Adam's adaptive learning rates lead to faster convergence and more robust performance across different problem settings, as illustrated in Figure 4.12 compared to other optimization algorithms.

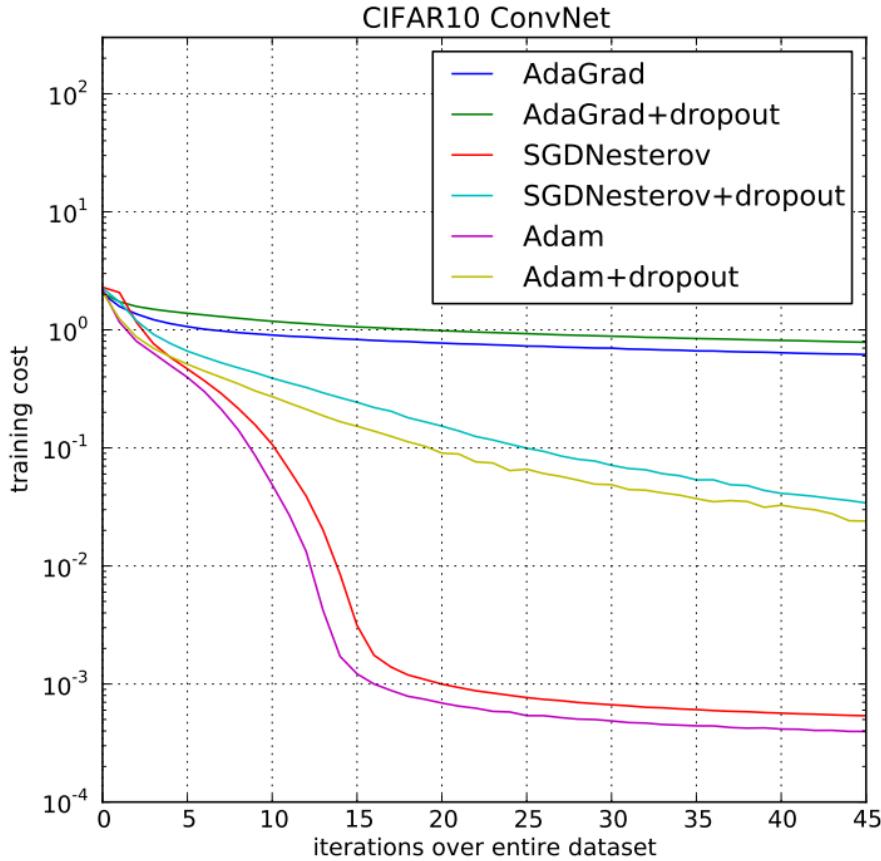


Figure 4.12: Comparison between the training cost of different optimizers on the CIFAR-10 benchmark dataset (adapted from: (13)).

4.6.3 Regularization Strategies for Generalization

Overfitting occurs when a model learns the training data too well, including its noise and idiosyncrasies, and fails to generalize to new data. This work employs several regularization techniques:

Dropout

Dropout (63) randomly sets a fraction p of neuron activations to zero at each training step. This prevents complex co-adaptations between neurons and forces the network to learn more robust and redundant features. During inference, all neurons are active, but their outputs are scaled by $(1 - p)$ to account for the increased number of active units. As depicted in Figure 4.13, dropout effectively trains an ensemble of exponentially many sub-networks.

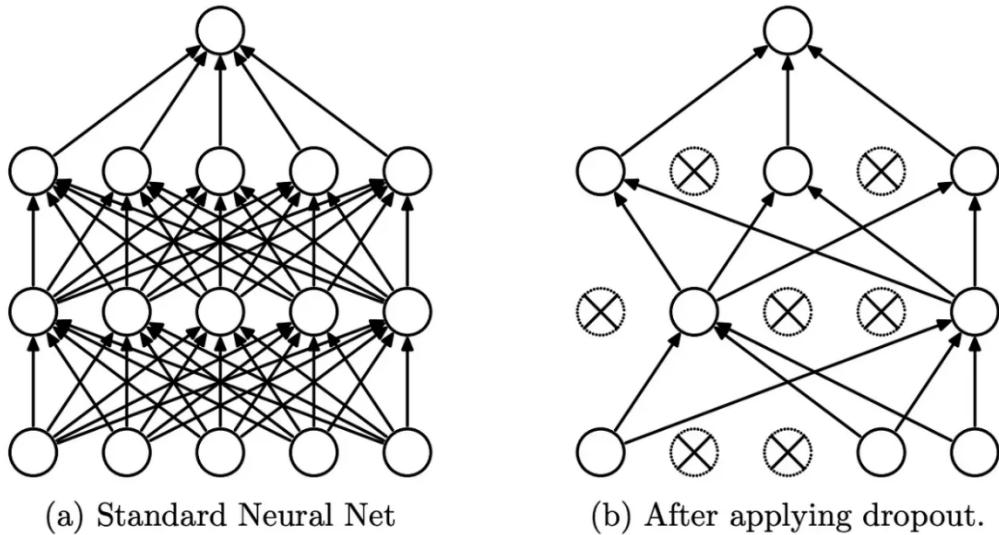


Figure 4.13: Illustration of the Dropout technique. Left: a standard fully connected network. Right: the same network during a training step with dropout applied, where some neurons (crossed out) are temporarily deactivated.

Early Stopping

Early Stopping monitors the model’s performance on a separate validation set during training. Training is halted when the validation loss ceases to improve for a set number of epochs (patience). This prevents the model from continuing to train into an overfitted state. As illustrated in Figure 4.14, the model is restored to the weights that achieved the minimum validation loss.

L2 Regularization (Weight Decay)

L2 regularization adds a penalty term to the loss function proportional to the squared magnitude of the weights:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{data}} + \lambda \sum_i \theta_i^2 \quad (4.42)$$

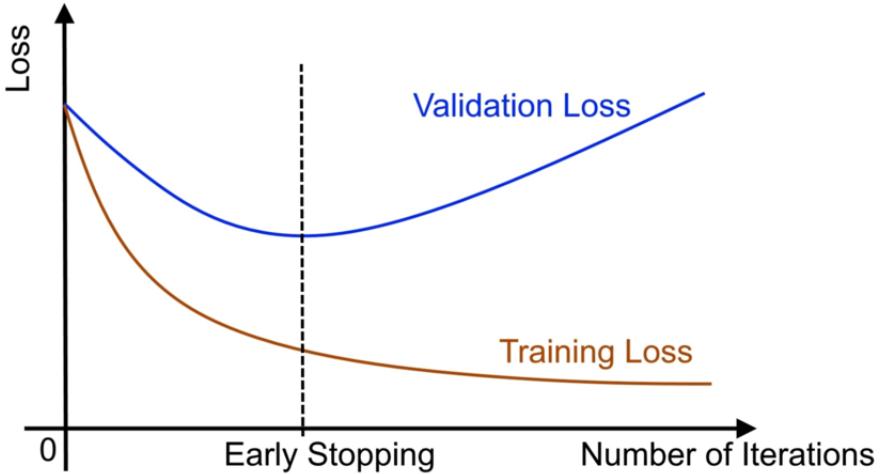


Figure 4.14: Early stopping illustration showing training stopped at minimum validation loss to prevent overfitting.

where λ is the regularization coefficient. This discourages large weights and promotes simpler, more generalizable models.

4.6.4 Hyperparameter Optimization via Bayesian Optimization (Optuna)

The performance of deep learning models is critically dependent on the choice of hyperparameters (e.g., learning rate, network architecture, dropout rate). As exhaustive grid search is computationally prohibitive, this work employs a more efficient strategy: **Bayesian Optimization**, implemented via the **Optuna** framework (64). This method uses a Tree-structured Parzen Estimator (TPE) to build a probabilistic model of the objective function (validation loss). This model is then used to intelligently select the next set of hyperparameters to evaluate, focusing the search on the most promising regions of the hyperparameter space.

The complete optimization process is formalized in Algorithm 1. For each model, a multi-trial study is conducted where, at each trial, a set of hyperparameters is sampled and used to train a model. The model’s performance is then evaluated using Time Series Cross-Validation (detailed in Section 4.6.5) to obtain a robust validation score. This score guides the TPE algorithm in subsequent trials, ensuring an efficient convergence toward the optimal hyperparameter configuration.

4.6.5 Robust Validation with Time Series Cross-Validation

Standard k-fold cross-validation, which involves random data shuffling, is invalid for time series analysis as it violates temporal causality and leads to data leakage—training on future data to predict the past. To obtain a robust and realistic

Algorithm 1 Hyperparameter Optimization (HPO) with Optuna

Require: Training data ($\mathbf{X}_{\text{train}}, \mathbf{y}_{\text{train}}$), Number of trials N

Ensure: Optimal hyperparameters θ^*

```
1: function OBJECTIVE(trial)
2:    $\theta \leftarrow \text{SampleHyperparameters}(\text{trial})$        $\triangleright$  e.g., learning rate, units, dropout
3:   model  $\leftarrow \text{BuildModel}(\theta)$ 
4:   CompileModel(model, optimizer( $\theta.\text{lr}$ ), loss = PinballLoss)
5:   history  $\leftarrow \text{model.fit}(\mathbf{X}_{\text{train}}, \mathbf{y}_{\text{train}}, \text{validation\_data} = \dots)$ 
6:   return min(history.val_loss)
7: end function

8: study  $\leftarrow \text{optuna.create\_study(direction='minimize')}$ 
9: study.optimize(Objective, n_trials =  $N$ )
10:  $\theta^* \leftarrow \text{study.best\_params}$ 
11: return  $\theta^*$ 
```

estimate of model generalization performance, this work exclusively employs **Time Series Cross-Validation**, a procedure formalized in Algorithm 2.

As illustrated in Figure 4.15, this method, implemented using scikit-learn's `TimeSeriesSplit` class, partitions the data into multiple sequential folds. In each fold, the training set strictly precedes the validation set. This structure mimics a real-world deployment scenario, guaranteeing that the model is always evaluated on "unseen" future data and providing a reliable assessment of its expected performance.

Algorithm 2 Model Evaluation with Time Series Cross-Validation

Require: Full dataset \mathbf{X} , Number of splits K

Ensure: Average validation score \bar{s}

```
1: scores  $\leftarrow []$ 
2: tscv  $\leftarrow \text{TimeSeriesSplit(n\_splits = } K)$ 
3: for each (train_idx, val_idx) in tscv.split( $\mathbf{X}$ ) do
4:    $\mathbf{X}_{\text{train}}, \mathbf{X}_{\text{val}} \leftarrow \mathbf{X}[\text{train\_idx}], \mathbf{X}[\text{val\_idx}]$ 
5:    $\mathbf{y}_{\text{train}}, \mathbf{y}_{\text{val}} \leftarrow \mathbf{y}[\text{train\_idx}], \mathbf{y}[\text{val\_idx}]$ 
6:
7:   Train model on ( $\mathbf{X}_{\text{train}}, \mathbf{y}_{\text{train}}$ )
8:   Evaluate model on ( $\mathbf{X}_{\text{val}}, \mathbf{y}_{\text{val}}$ ) to get score  $s$ 
9:   Append  $s$  to scores
10: end for

11:  $\bar{s} \leftarrow \text{Average(scores)}$ 
12: return  $\bar{s}$ 
```

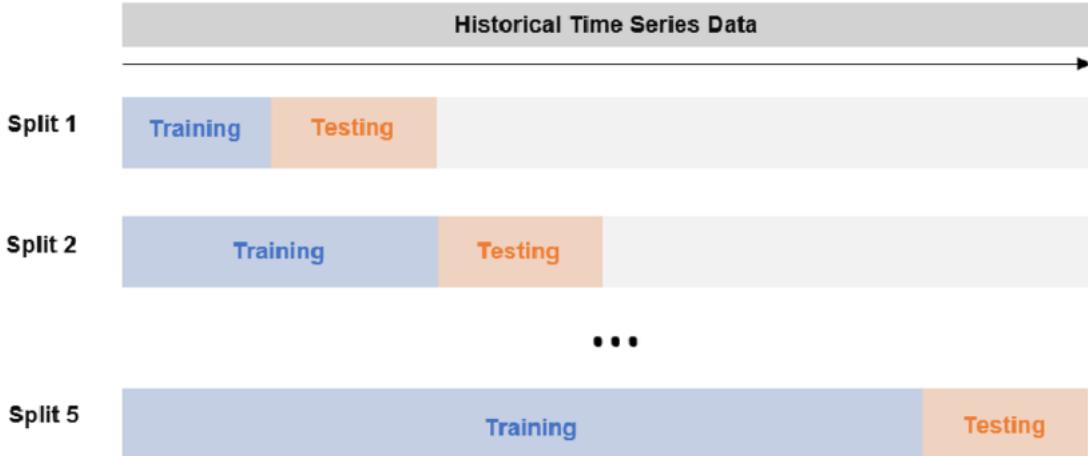


Figure 4.15: Illustration of Time Series Cross-Validation. The data is split into multiple folds, where each fold's training set always precedes its validation set, preserving the temporal order and preventing data leakage.

4.7 Performance Metrics for Model Evaluation

To quantitatively evaluate and compare the performance of the forecasting models, a comprehensive set of statistical metrics is employed. These metrics measure the discrepancy between the predicted values (\hat{y}_i) and the actual observed values (y_i) over a dataset of N samples.

4.7.1 Deterministic Forecast Metrics

For point forecasts (typically the median, P50), the following metrics are used:

- **Mean Absolute Error (MAE):** Calculates the average absolute difference between predicted and actual values. It provides a clear, interpretable measure of the average magnitude of errors in the original units (kW).

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (4.43)$$

- **Root Mean Squared Error (RMSE):** Calculates the square root of the average of the squared differences. By squaring the errors, it gives relatively high weight to large errors, making it particularly useful for penalizing models that produce significant outlier predictions.

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (4.44)$$

- **Mean Absolute Percentage Error (MAPE):** Expresses the average absolute error as a percentage of the installed capacity, providing a scale-independent metric useful for comparing performance across different datasets and operational conditions.

$$\text{MAPE} = \frac{100\%}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{C_{\text{cap}}} \right| \quad (4.45)$$

where C_{cap} is the installed wind turbine capacity. Note: This formulation normalizes by capacity rather than actual values to avoid division by zero when $y_i = 0$ and to provide a consistent reference across all operating conditions.

- **Coefficient of Determination (R^2):** Measures the goodness of fit, indicating the proportion of variance in the dependent variable that is predictable from the model's predictions. An R^2 value of 1 indicates perfect fit, while 0 indicates performance no better than a simple mean model.

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (4.46)$$

where $\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i$ is the mean of the true values.

4.7.2 Probabilistic Forecast Metrics

For probabilistic forecasts, the Quantile Loss (Pinball Loss) defined in Section 4.5.1 (Equation 4.33) serves as the primary evaluation metric. This metric directly assesses the quality of each predicted quantile by measuring how well it captures the desired conditional quantile of the target distribution.

For a given quantile τ , the average quantile loss over the test set is computed as:

$$\text{QL}_\tau = \frac{1}{N} \sum_{i=1}^N L_\tau(y_i, \hat{q}_{i,\tau}) \quad (4.47)$$

Lower values of QL_τ indicate better probabilistic forecast performance for that specific quantile. In this work, quantile loss is computed separately for each of the three predicted quantiles ($\tau \in \{0.1, 0.5, 0.9\}$) and for each forecast horizon ($t + 10$, $t + 20$, $t + 30$), providing a detailed assessment of the model's ability to capture forecast uncertainty across different time scales.

The collective use of both deterministic metrics (MAE, RMSE, MAPE, R^2) for the median forecast and probabilistic metrics (Quantile Loss) for the full distribution provides a holistic view of model performance, enabling rigorous comparison and validation of the proposed methods in the subsequent chapters of this dissertation.

Chapter 5

Proposed Methodology and Decision Support Framework

Following the theoretical exposition of the preceding chapters, this chapter presents the practical implementation and methodological framework developed in this dissertation. The primary objective is to construct a comprehensive solution that not only provides accurate wind power forecasts but also translates these predictions into actionable intelligence for the operational management of a Floating Production, Storage and Offloading (FPSO) unit. This chapter serves as the bridge between the established theory and the novel contributions of this work, detailing the "how" and "why" behind the engineering choices made.

The chapter is structured to follow the logical flow of the developed application. Section 5.1 details the real-world dataset and the critical preprocessing pipeline. Section 5.2 provides an in-depth description of the forecasting models, from established benchmarks to the novel hybrid architectures, culminating in the proposed **CEEMDAN-EWT-TFT Aggregator**. Finally, Section 5.3—the core of this chapter—dissects the integrated Decision Support System, detailing its three main pillars: the Operational Risk Index (ORI), the comprehensive Techno-Economic Analysis module, and the innovative LLM-powered Operational Copilot.

5.1 Dataset and Preprocessing Pipeline

The empirical foundation of any data-driven methodology is the quality and structure of its data. This section details the dataset used and the preprocessing steps applied, which are critical for model performance and reliability.

5.1.1 Data Source and Characteristics

The dataset employed in this study is sourced from the "CARE to Compare" benchmark, a publicly available collection of real-world wind turbine data curated by Gück et al. (2). This benchmark provides high-quality SCADA data from 36 turbines across three different wind farms. For this dissertation, we specifically utilize the data file designated as `44.csv`, which corresponds to a single turbine within "Wind Farm C". According to the dataset documentation, Wind Farm C is an offshore facility, making its data particularly relevant to the FPSO application context of this work.

The dataset consists of time series data with a 10-minute resolution. Key variables used in this work include the timestamp (`time_stamp`), average active power (`power_5_avg`), and average wind speed (`wind_speed_235_avg`). A sample of the raw power generation data is illustrated in Figure 5.1, which visually confirms the highly volatile and intermittent nature of wind power, reinforcing the challenges discussed in Chapter 2.

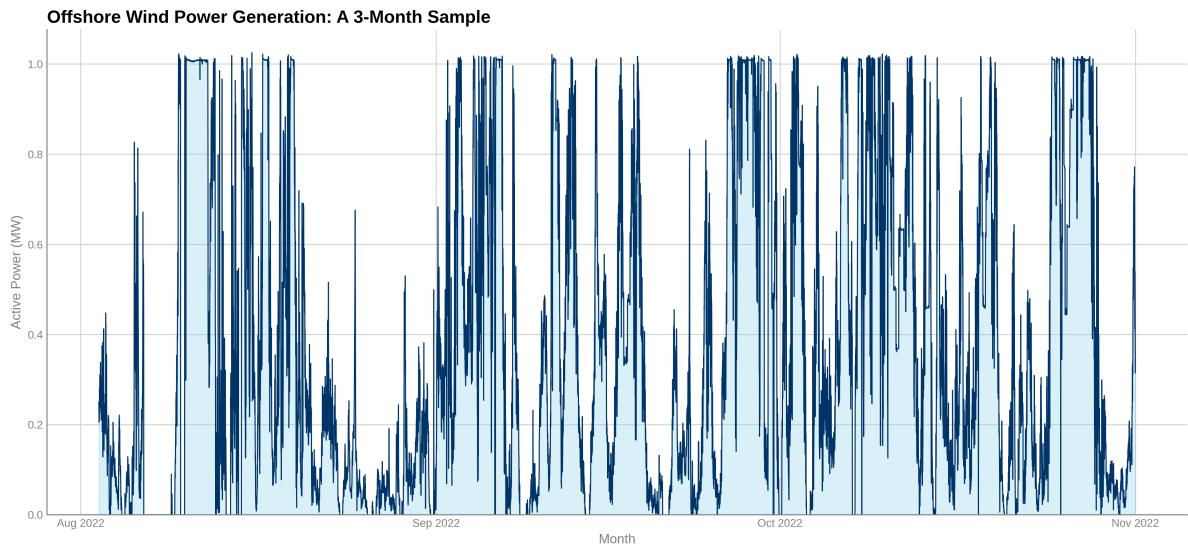


Figure 5.1: Example of the raw active power time series from the selected wind turbine (`44.csv`), illustrating the characteristic volatility and intermittency during 3 months.

5.1.2 Data Preprocessing and Structuring

Prior to model training, the raw data undergoes several preprocessing steps as implemented in the `load_data` function of the dashboard's codebase. This procedure is fundamental to ensure data quality and structure the data for the subsequent learning algorithms.

Data Cleaning and Feature Engineering

The initial steps involve data cleaning and minor feature engineering:

1. **Timestamp Conversion:** The raw timestamps are converted into a proper datetime format, enabling time-based indexing and analysis.
2. **Feature Engineering:** To facilitate the analysis of seasonal patterns, a 'Month' feature is engineered, representing the number of months elapsed since the beginning of the dataset.
3. **Data Cleaning:** Any rows with missing values (NaN) in the key columns (power, wind speed) are dropped to ensure the integrity of the input sequences.
4. **Anonymization Awareness:** It is noted that, as per the dataset documentation (2), the feature names have been anonymized, and only features related to power and wind speed are directly identifiable. Timestamps were also shifted to start in 2022 to protect the asset's identity.

Sliding Window Transformation for Supervised Learning

A fundamental step in applying supervised learning algorithms to time series forecasting is the transformation of the sequential data into a format of feature-target pairs. This is achieved using a sliding window technique, a process meticulously implemented in the various `create_dataset` and `create_dataset_multi` functions within this work's codebase.

As illustrated conceptually in Figure 5.2, a window of a fixed length, the **look-back period (L)**, is slid across the time series one step at a time. At each time step t , this process generates a single training sample:

- **Feature Vector (\mathbf{x}_t):** The sequence of the last L observations becomes the input feature vector, i.e., $\mathbf{x}_t = [y_{t-L+1}, \dots, y_t]$.
- **Target Vector (\mathbf{y}_t):** The subsequent H observations become the target vector that the model must learn to predict, i.e., $\mathbf{y}_t = [y_{t+1}, \dots, y_{t+H}]$.

In this dissertation, all models are configured with $L = 6$ (representing 60 minutes of historical data) and a multi-step forecast horizon of $H = 3$ (representing predictions for $t+10$, $t+20$, and $t+30$ minutes). This transformation yields a dataset $\mathcal{D} = \{(\mathbf{x}_t, \mathbf{y}_t)\}_{t=L}^{T-H}$, where T is the total length of the time series, which can now be used to train any supervised learning model.

Single-Scale Sliding Window

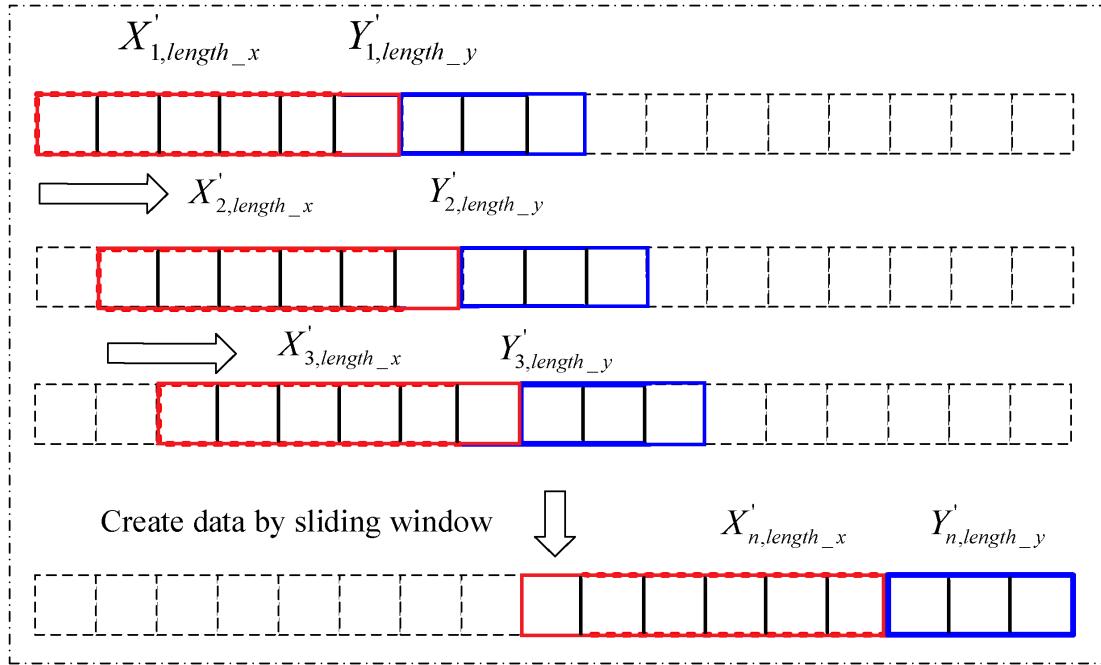


Figure 5.2: Detailed illustration of the sliding window technique. The process transforms a continuous time series into a set of discrete input-output pairs $(\mathbf{x}_t, \mathbf{y}_t)$ suitable for supervised learning. The look-back period $L = 6$ corresponds to 60 minutes of history, and the forecast horizon $H = 3$ corresponds to predictions at $t+10$, $t+20$, and $t+30$ minutes.

Data Partitioning

The dataset is partitioned chronologically into an 80% training set and a 20% test set. This strict temporal split is a critical best practice (as discussed in Section 4.6.1) to prevent data leakage and obtain an unbiased estimate of the model's performance on unseen future data. The test set serves as a proxy for real-world deployment, simulating the model's behavior when forecasting truly unknown future values.

5.2 Forecasting Models: A Methodological Funnel

This section details the hierarchy of forecasting models implemented, justifying the progression from simpler models to the final proposed architecture. This "funneling" approach, starting broad and narrowing down to the most promising method, is a core part of the methodological investigation.

5.2.1 Signal Decomposition as a Core Strategy

As established in the literature review (Chapter 3) and supported by recent studies (12; 27), the core challenge in wind power forecasting is the non-stationary and

multi-scale nature of the signal. The "divide and conquer" strategy addresses this by decomposing the complex signal into simpler, more predictable components before applying forecasting models.

CEEMDAN and EWT: The Decomposition Engine

The primary decomposition is performed using Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) (25). This technique decomposes a signal $s(t)$ into a finite set of Intrinsic Mode Functions (IMFs), $c_k(t)$, and a final residual, $r(t)$, such that:

$$s(t) = \sum_{k=1}^K c_k(t) + r(t) \quad (5.1)$$

This process isolates the different oscillatory modes of the signal, with each IMF representing a characteristic time scale. The highest-frequency IMF (typically $c_1(t)$) often contains significant noise and is further processed using the Empirical Wavelet Transform (EWT) (26) for targeted denoising.

This two-step decomposition (CEEMDAN-EWT) is the engine of the hybrid models in this work. The entire process is visually conceptualized in Figure 5.3, which shows how the original volatile signal is separated into distinct components that exhibit more stationary behavior and are thus more amenable to accurate forecasting. The original power signal is first decomposed by CEEMDAN into multiple IMFs and a residual. The highest frequency component (IMF1) is then denoised using EWT before being recombined or used as a feature. This process separates the signal into components with distinct temporal characteristics.

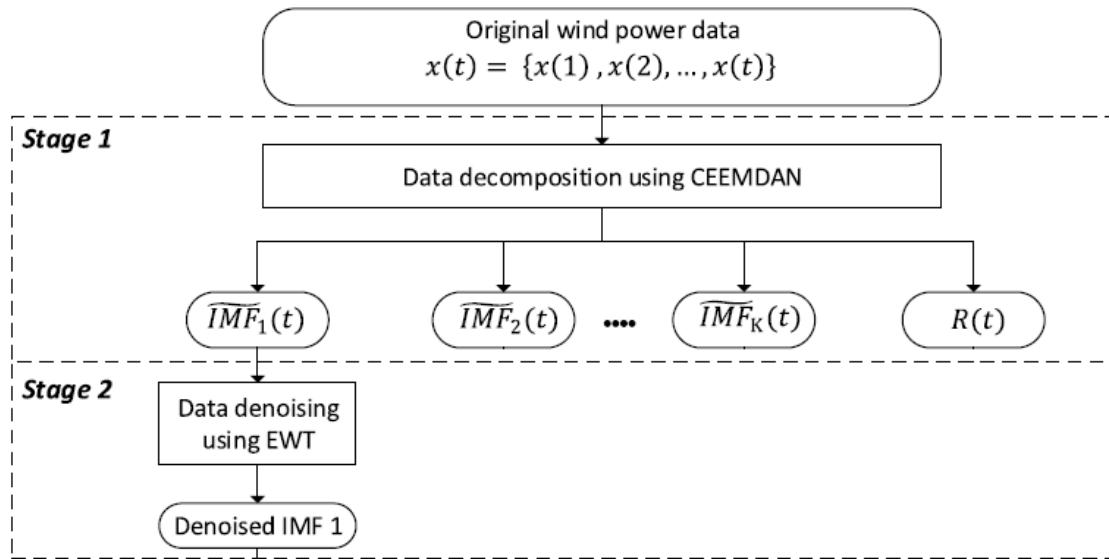


Figure 5.3: CEEMDAN-EWT decomposition process showing IMF extraction and EWT-based denoising of high-frequency components.

5.2.2 A Funneling Approach to Model Selection

The forecasting methodology follows a systematic funneling approach, progressing from simple baseline models to increasingly sophisticated architectures. This progression is not merely exploratory; it is a deliberate strategy to identify the minimal complexity required to achieve the desired performance, balancing accuracy with computational efficiency and interpretability.

Benchmark Models

A comprehensive suite of models, defined in the `model_functions` dictionary within the codebase, serves as a baseline for comparison. This includes:

- **Classical Machine Learning:** Support Vector Regression (SVR) and Random Forest (RF), which serve as non-deep learning baselines.
- **Standard Deep Learning:** Simple Artificial Neural Networks (ANN), Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), and hybrid architectures such as CNN+Bi-LSTM and CNN+Bi-LSTM+Attention.
- **Transformer-Based Models:** Standard Transformer and Temporal Fusion Transformer (TFT), representing the state-of-the-art in sequence modeling.

All models are configured to perform multi-step forecasting for the three target horizons ($t+10$, $t+20$, $t+30$ minutes) and are trained using quantile regression (Section 4.5.1) to produce probabilistic forecasts at the 10th, 50th, and 90th percentiles.

Hybrid "Divide and Conquer" Models

The next level of sophistication involves the "divide and conquer" strategy, implemented in functions such as `proposed_method`, `proposed_method_bilstm`, `proposed_method_gru`, `proposed_method_bilstm_att`, `proposed_method_cnn_bilstm`, and `proposed_method_cnn_bilstm_att`.

As illustrated in Figure 5.4, this approach follows a multi-stage pipeline:

1. The original signal is decomposed using CEEMDAN-EWT.
2. A separate neural network (LSTM, Bi-LSTM, CNN+Bi-LSTM, etc.) is trained for each resulting IMF component and the residual.
3. The forecasts from all individual models are summed to reconstruct the final prediction of the original signal.

This strategy is effective because it allows each specialized model to focus on a single, more stationary component. However, it is computationally expensive (requiring $K+1$ separate models, where K is the number of IMFs) and assumes that the components are independent, which may not fully capture the inter-dependencies between different temporal scales.

The original signal is decomposed into IMFs and a residual. A separate forecasting model is trained for each component, and the final forecast is obtained by summing the individual predictions.

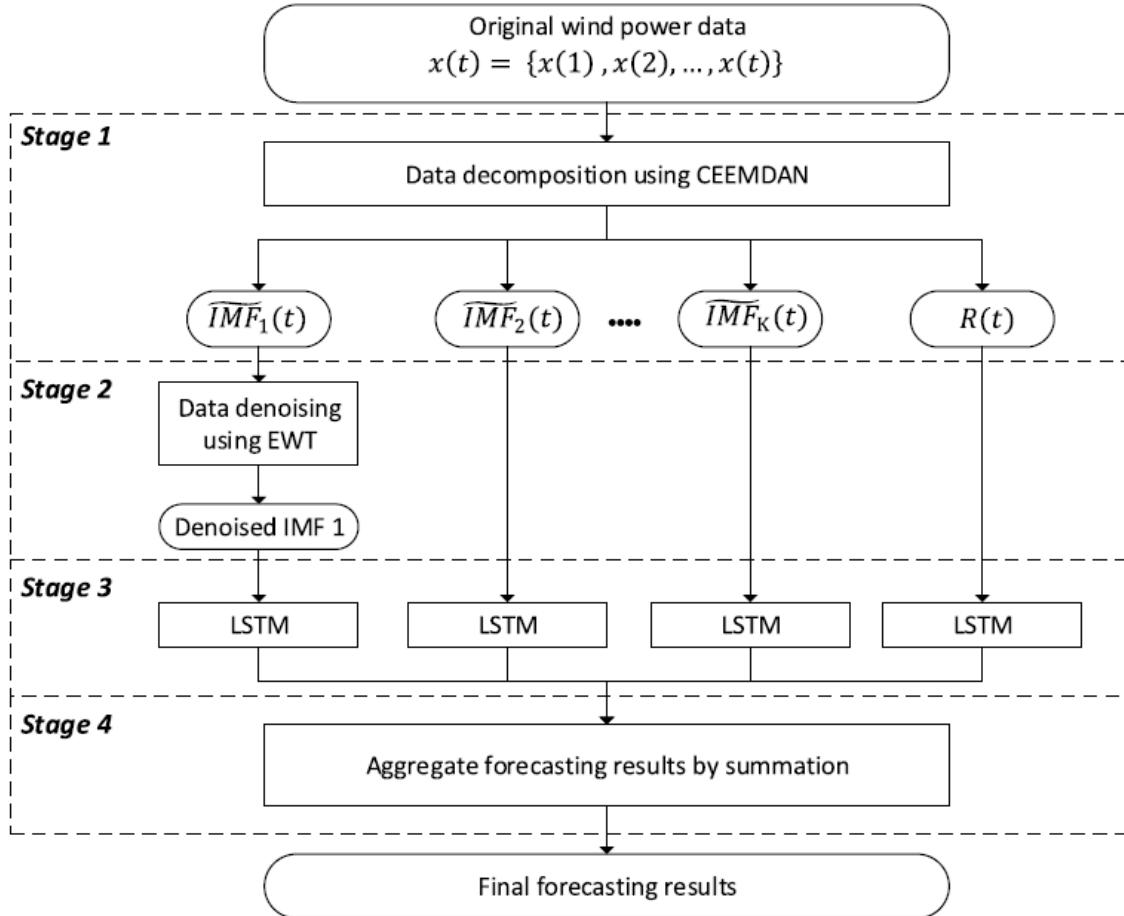


Figure 5.4: "Divide and Conquer" framework where decomposed IMFs are forecasted separately and aggregated for final prediction.

Proposed Method: The CEEMDAN-EWT-TFT Aggregator

This dissertation's primary methodological contribution is the novel and computationally efficient **CEEMDAN-EWT-TFT Aggregator**, implemented in the `proposed_method_tft_aggregator` function. This approach fundamentally re-frames signal decomposition not as a divide-and-conquer task, but as a powerful **feature engineering** task.

As depicted in Figure 5.5, instead of training separate models for each IMF, the decomposed components are treated as a rich, multivariate feature set. These features are fed into a single, powerful predictor based on the Temporal Fusion Transformer (TFT) architecture, which then directly predicts the original signal (not the individual components). This "aggregator" architecture offers several key advantages:

- **Unified Learning:** A single model learns the complex, non-linear inter-dependencies between all temporal scales (IMFs) simultaneously, rather than assuming their independence.
- **Computational Efficiency:** Only one model needs to be trained, deployed, and maintained, drastically reducing the computational and operational overhead compared to the traditional "one model per component" approach.
- **Leveraging Core TFT Strengths:** The architecture directly utilizes the most powerful components of the TFT framework, including Gated Residual Networks for robust feature processing and a temporal self-attention mechanism to capture long-range dependencies.
- **Native Multi-Horizon and Probabilistic Forecasting:** The model inherently supports multi-step forecasting and quantile regression, generating P10, P50, and P90 forecasts for all horizons in a single forward pass, which is essential for real-time risk assessment.

Mathematically, if the CEEMDAN-EWT decomposition produces K components (IMFs and a residual), the input to the model at time t is the multivariate feature vector:

$$\mathbf{X}_t = [\text{IMF}_1(t - L : t), \text{IMF}_2(t - L : t), \dots, \text{Comp}_K(t - L : t)] \quad (5.2)$$

where each component is a time series of length L (the look-back window). The adapted TFT model then directly predicts the quantiles of the original signal:

$$\hat{y}_\tau(t + h) = f_{\text{TFT-Aggregator}}(\mathbf{X}_t), \quad \text{for } h \in \{10, 20, 30\} \text{ and } \tau \in \{0.1, 0.5, 0.9\} \quad (5.3)$$

It is important to note that the full TFT architecture includes a Variable Selection Network (VSN) intended to handle use cases with a large number of diverse input variables (e.g., static metadata, known future inputs, and multiple observed time series). In the context of the proposed CEEMDAN-EWT-TFT Aggregator, the input space is composed exclusively of the IMFs, which are highly correlated,

homogeneous features derived from the same underlying signal. The powerful feature engineering and separation already performed by the CEEMDAN-EWT stage mitigates the need for an additional explicit feature selection layer. Consequently, a simplified and specialized architecture was adopted, omitting the VSN to focus computational resources on the core temporal processing blocks.

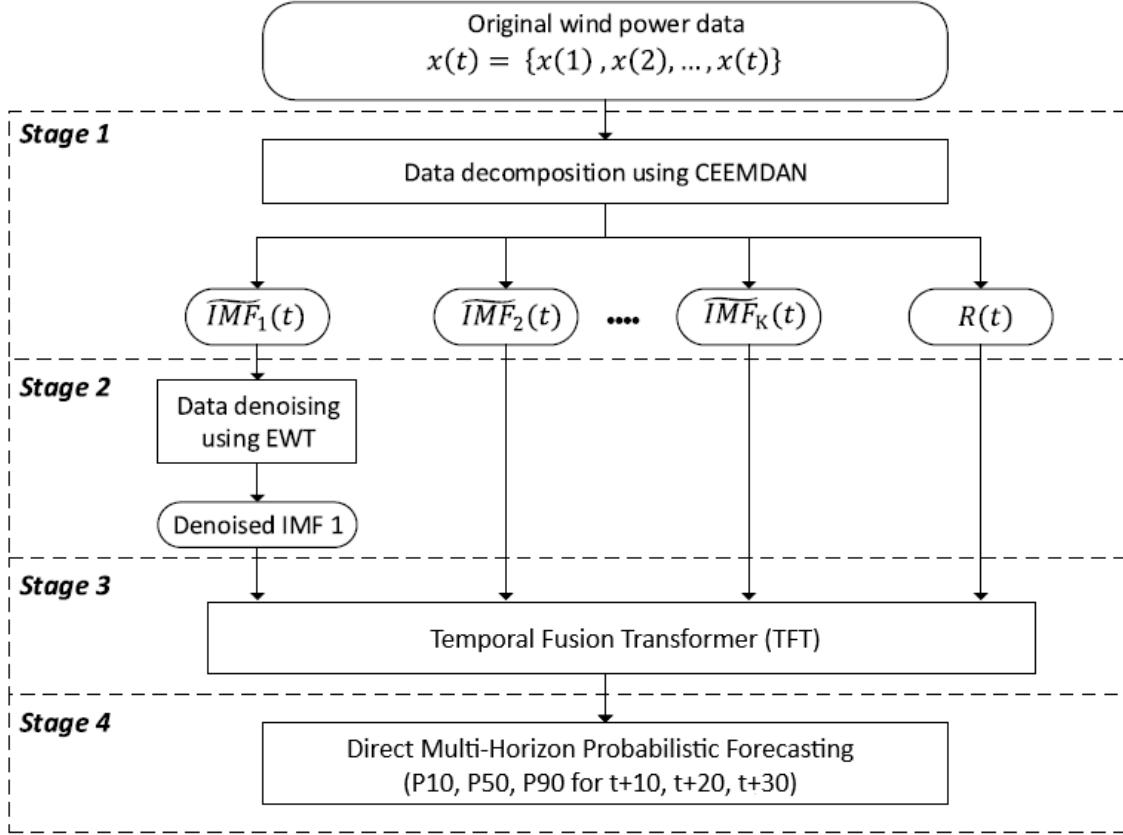


Figure 5.5: CEEMDAN-EWT-TFT Aggregator architecture using decomposed IMFs as multivariate features for a single TFT model.

5.3 The Decision Support System Framework

The culmination of this work is a functional prototype of a Decision Support System (DSS), implemented as an interactive Streamlit dashboard. This system integrates the forecasting models with three critical decision-support components: the Operational Risk Index (ORI), a Techno-Economic Analysis module, and an LLM-powered Operational Copilot. Together, these components transform raw probabilistic forecasts into actionable, context-aware recommendations for FPSO operators.

5.3.1 Operational Risk Index (ORI)

The Operational Risk Index (ORI) is a rule-based heuristic designed to translate the probabilistic forecasts (specifically, the lower-bound quantile forecasts) into three intuitive, actionable risk levels: **Low**, **Attention**, and **High**. This index serves as the primary interface between the forecasting system and the operator, providing a clear signal for decision-making.

ORI Logic and Calculation

The ORI is calculated at each time step based on a comparison between the predicted lower-bound (10th percentile, $p_{0.1}$) of wind power generation at future horizons and a user-defined **operational threshold** (θ_{op}). The operational threshold represents the minimum acceptable wind power generation; if the wind power falls below this threshold, backup generators must be activated to prevent a blackout.

The ORI logic, as implemented in the dashboard code, is defined as follows:

Algorithm 3 Operational Risk Index (ORI) Calculation

Require: $p_{0.1}^{(t+20)}$: Predicted 10th percentile at $t+20$ min $p_{0.1}^{(t+30)}$: Predicted 10th percentile at $t+30$ min θ_{op} : Operational threshold (MW) $h_{trigger}$: Trigger horizon (20 or 30 minutes)**Ensure:** ORI_t : Risk level at time t (*Low*, *Attention*, or *High*)

```
1: if  $h_{trigger} = 20$  then
2:   if  $p_{0.1}^{(t+20)} < \theta_{op}$  then
3:      $ORI_t \leftarrow "High"$ 
4:   else if  $p_{0.1}^{(t+30)} < \theta_{op}$  then
5:      $ORI_t \leftarrow "Attention"$ 
6:   else
7:      $ORI_t \leftarrow "Low"$ 
8:   end if
9: else
10:   $\triangleright h_{trigger} = 30$ 
11:  if  $p_{0.1}^{(t+30)} < \theta_{op}$  then
12:     $ORI_t \leftarrow "High"$ 
13:  else
14:     $ORI_t \leftarrow "Low"$ 
15:  end if
16: end if
```

The rationale behind this logic is as follows:

- **High Risk:** If the pessimistic forecast (P10) at the primary trigger horizon ($t+20$ or $t+30$, depending on the chosen strategy) falls below the threshold, immediate action is required. The recommendation is to **connect the backup generator**.
- **Attention:** (Only when $h_{trigger} = 20$) If the $t+20$ forecast is safe but the $t+30$ forecast is below the threshold, the system signals a warning. The recommendation is to place the **team on standby**, preparing for a potential generator start in the near future.
- **Low Risk:** If both horizons are above the threshold, the wind generation is deemed sufficient. The recommendation is to **keep the generator off**.

This three-level system provides a balance between conservatism (avoiding blackouts) and efficiency (minimizing unnecessary generator usage). The flowchart in

Figure 5.6 provides a visual representation of this decision logic.

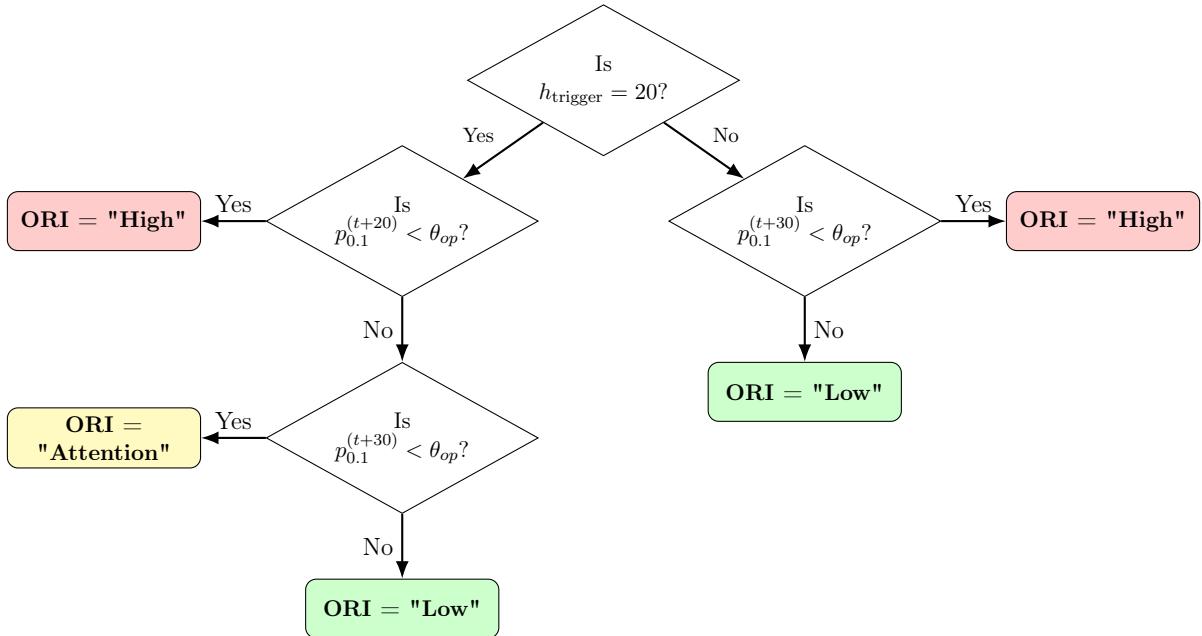


Figure 5.6: Flowchart of the logic for the Operational Risk Index (ORI), translating quantile forecasts into actionable risk levels. The decision tree depends on the selected trigger horizon (h_{trigger}) and compares the P10 forecasts against the operational threshold.

Operational Threshold as a Tunable Parameter

The operational threshold θ_{op} is a critical parameter that directly influences the trade-off between reliability and cost. A higher threshold leads to more conservative operation (more frequent generator usage, fewer blackouts), while a lower threshold leads to more aggressive operation (less generator usage, higher blackout risk).

In the dashboard implementation, the threshold is specified as a percentage of the installed turbine capacity and can be adjusted by the operator via a slider (default: 30%). This flexibility allows the system to adapt to different operational contexts and risk tolerances.

5.3.2 Techno-Economic Analysis Module

To quantify the economic value of the predictive framework and to enable data-driven strategic planning, a comprehensive techno-economic analysis module was developed. This module simulates the operational costs of different decision-making strategies and identifies the optimal parameters that minimize total cost while maintaining reliability.

Cost-Benefit Simulation Model

The `calculate_costs` function simulates the total operational cost under different strategies over the entire test period. The simulation is based on explicit economic premises:

- $C_{fuel} = \$500/\text{hour}$: Cost per hour of backup generator operation (`COST_PER_HOUR_GENERATOR`).
- $C_{startup} = \$2,000$: Cost per generator startup event (`COST_PER_STARTUP`).
- $C_{blackout} = \$100,000$: Penalty per time step of unmet demand, representing the cost of a blackout event (`COST_PER_BLACKOUT_EVENT`).
- $\Delta t = 10/60 \text{ hours}$: Time step duration (10 minutes).
- $T_{startup} = 2 \text{ time steps}$: Generator startup time (20 minutes).

The total cost for a given strategy over the test period is calculated as:

$$C_{total} = \sum_{t=1}^T (\mathbb{I}_{gen_on}(t) \cdot C_{fuel} \cdot \Delta t) + N_{startups} \cdot C_{startup} + N_{blackouts} \cdot C_{blackout} \quad (5.4)$$

where:

- $\mathbb{I}_{gen_on}(t)$ is an indicator function that equals 1 if the generator is on at time t , and 0 otherwise.
- $N_{startups}$ is the total number of generator startup events.
- $N_{blackouts}$ is the total number of blackout events (time steps where demand is unmet).

The simulation compares three distinct operational strategies:

- 1. Reactive Strategy (No Forecast):** This baseline strategy does not use any forecasts. The generator is started only *after* the wind power falls below the threshold. Due to the startup delay ($T_{startup} = 2 \text{ time steps}$), this reactive approach inevitably results in blackouts during the startup period.
- 2. Predictive Strategy (ORI-Based):** This strategy uses the ORI to make proactive decisions. The generator is started when the ORI indicates "High" risk (i.e., when the trigger forecast $p_{0.1}^{(t+h_{trigger})}$ falls below θ_{op}). Because the forecast provides advance warning, the generator can be started before the actual power drop, ideally avoiding blackouts.

3. Conservative Strategy (ORI Attention): This strategy is even more conservative, starting the generator when the ORI indicates either "High" or "Attention" risk (i.e., when either $p_{0.1}^{(t+20)} < \theta_{op}$ or $p_{0.1}^{(t+30)} < \theta_{op}$).

The simulation outputs a detailed cost breakdown for each strategy, including total cost, generator operating hours, number of startups, and number of blackout events. This allows for a direct, quantitative comparison of the economic impact of different decision-making approaches.

Strategic Parameter Optimizer

While the cost simulation provides valuable insights for a given set of parameters, the `optimize_strategy` function transforms the dashboard into a strategic planning tool by performing an exhaustive grid search to find the *optimal* combination of operational parameters.

The optimizer searches over two key parameters:

- **Operational Threshold (θ_{op}):** Tested as a percentage of capacity, ranging from 10% to 50% in steps of 2.5%.
- **Trigger Horizon ($h_{trigger}$):** Tested at both t+20 and t+30 minutes.

For each combination of parameters, the total cost of the "Predictive (ORI)" strategy is calculated using the `calculate_costs` function. The optimizer then identifies the parameter combination that yields the *minimum* total cost. The results include:

- The optimal threshold percentage.
- The optimal trigger horizon.
- The minimum achievable cost.
- The number of blackouts and generator starts under the optimal strategy.

The optimization results are stored in `st.session_state.best_strategy_results` and visualized as a heatmap, providing a clear visual representation of the cost landscape across the parameter space. This heatmap, shown in Figure 5.7, allows operators and planners to understand the sensitivity of costs to parameter choices and to make informed decisions about operational policies.

5.3.3 AI Copilot for Operator Support

The final and most innovative component of the Decision Support System is the **Operational Copilot**, an advanced AI agent designed to bridge the gap between



Figure 5.7: Strategic parameter optimizer heatmap showing operational cost versus threshold and trigger horizon parameters.

complex data analysis and practical, actionable operator support. This moves beyond a simple chatbot by implementing a sophisticated framework where the Large Language Model (LLM) can autonomously decide whether to analyze data, answer questions, or execute actions by calling predefined software functions, known as "Tools".

Hybrid Architecture: RAG + Tool-Using Agent

The copilot's intelligence is built upon a hybrid architecture that combines two state-of-the-art LLM techniques:

1. **Retrieval-Augmented Generation (RAG):** Before every interaction, the agent is provided with a rich, real-time data context (65). This "grounding" ensures that all its analyses and decisions are based on the current operational reality, mitigating the risk of hallucination and ensuring relevance. The context includes the current risk level (ORI), key performance indicators (KPIs), and the results of any previous strategy optimizations.
 2. **Tool-Using Agent Framework:** The agent is given access to a set of pre-defined Python functions ("Tools") that it can choose to execute (66; 67). This transforms the agent from a passive analyst into an active participant capable of performing tasks on behalf of the operator.

The agent's decision-making process is governed by a carefully crafted system prompt that instructs it to first evaluate if a user's request can be fulfilled by a tool.

If not, it defaults to using the provided RAG context to formulate a direct answer. This dual-capability framework is illustrated in Figure 5.8.

Available Agent Tools

The agent has access to a suite of powerful tools designed to automate complex analytical and operational tasks:

- `run_strategy_optimization_tool()`: Executes a 2D grid search to find the optimal operational threshold and trigger horizon for any of the benchmark models.
- `run_global_strategy_optimization_tool()`: A more advanced tool, exclusive to the proposed TFT model, which performs a 3D search to find the global optimum across threshold, horizon, **and** the P10 Safety Factor.
- `apply_new_parameters_tool(threshold, horizon)`: Allows the agent to directly modify the dashboard's operational parameters and trigger a simulation rerun, enabling "what-if" scenario analysis commanded by natural language.
- `generate_summary_report_tool(start, end)`: Instructs the agent to calculate and present a concise cost-benefit report for a specific time window, automating a previously manual analysis task.

This evolution from a simple RAG-based chatbot to a tool-using agent represents a significant step towards creating truly autonomous and useful AI assistants for complex industrial environments. It demonstrates a system that not only provides information but can also be delegated tasks, fundamentally enhancing the operator's capabilities.

2. Proactive Agent (Autonomous Surveillance): This agent, implemented in the `get_llm_driven_proactive_alert` function, represents a more advanced use of the LLM. It runs autonomously at each time step, performing the following tasks:

1. It analyzes a continuous stream of data, including the recent history of the ORI and the full multi-horizon forecast trajectory (both median and pessimistic cases).
2. It independently decides if a critical event has occurred (e.g., a change in risk level or a deteriorating forecast trend).

3. If a critical event is detected, it generates a timely, data-backed alert. If not, it provides a concise summary of the forecast trend to maintain operator awareness.

This proactive capability transforms the copilot from a passive tool into an active, vigilant member of the operational team.

Explainability and Trust

A critical aspect of deploying AI in safety-critical environments is explainability (XAI) and trust. The framework addresses this through:

- **Data Grounding:** The RAG architecture inherently grounds all responses in verifiable data from the simulation. The system prompt for the proactive agent explicitly requires it to cite the numerical data supporting its analysis.
- **Full Transparency:** The dashboard includes an "AI Audit" expander that displays the exact context string sent to the LLM for its most recent analysis, allowing for complete verification.
- **Conversational Memory:** The reactive agent maintains a conversation history, enabling multi-turn dialogues for clarification and deeper analysis.

This combination of features is designed to build operator trust, a prerequisite for the effective adoption of AI systems in real-world operations.

5.3.4 Integration and Workflow

The three core components—the Operational Risk Index (ORI), the Techno-Economic Analysis, and the AI Copilot—are tightly integrated within the Streamlit dashboard to create a seamless and powerful decision-support environment. The typical user workflow demonstrates how these components work in synergy:

1. **Configuration:** The operator selects a forecasting model and configures initial parameters (e.g., operational threshold) via the sidebar.
2. **Initial Analysis:** The system runs the selected model, calculates the ORI for the entire test period, and displays the main recommendation. Simultaneously, the **Proactive Agent** provides its autonomous analysis of the current situation, alerting the operator to any immediate risks or trends.
3. **Visual Exploration:** The operator can navigate through the simulation, viewing detailed time-series plots with color-coded risk levels and exploring the model's interpretability through attention analysis.

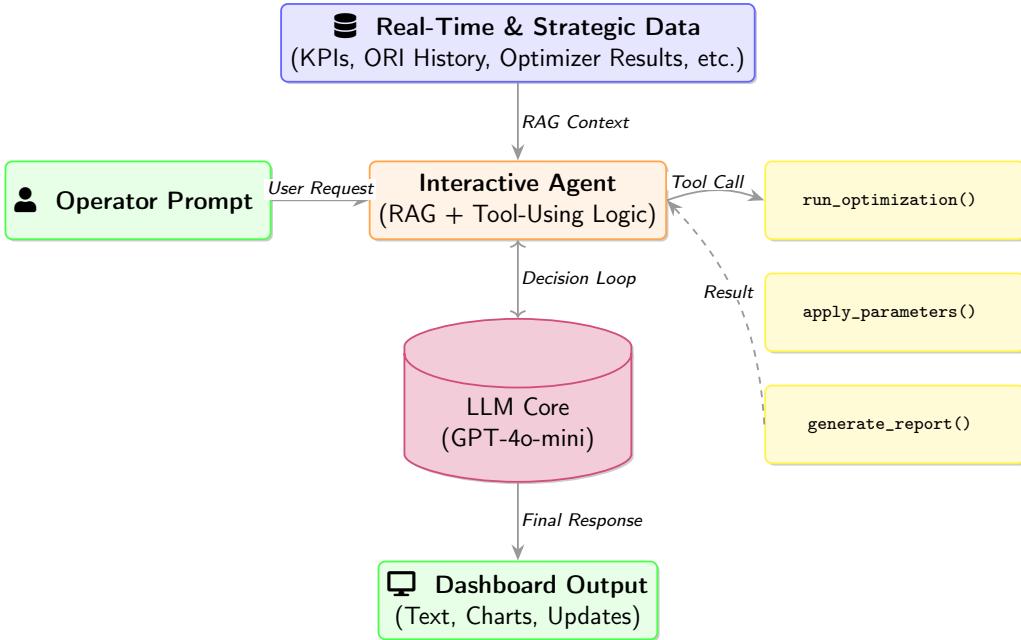


Figure 5.8: Architecture of the Interactive AI Copilot. The agent combines Retrieval-Augmented Generation (RAG) with a Tool-Using framework, allowing it to either answer questions directly or execute specific actions (Tools) based on the operator's request.

4. Strategic Optimization (Manual or Delegated): The operator has two ways to find the most cost-effective operational policy:

- Manually click the "Run Optimization" button in the user interface.
- **Delegate the task to the AI Copilot** by simply typing a command like "run the strategy optimization" or "find the best parameters for this model".

5. Interactive Dialogue and Action: At any point, the operator can engage with the **Interactive Agent**. This moves beyond a simple Q&A session, allowing the operator to:

- Ask analytical questions about the current situation or optimization results (e.g., "summarize the risks for the next hour").
- Command the agent to perform actions using its tools, such as applying a new strategy ("apply a threshold of 25% and a horizon of 30 minutes") or generating a specific report ("create a cost summary for the last 24 hours").

This integrated workflow transforms the system from a passive prediction tool into an **active decision-support partner**. It empowers the operator not only by providing data-grounded insights but also by automating complex analytical tasks through a simple, conversational interface.

5.3.5 Model Comparison Mode

To facilitate the systematic evaluation and selection of forecasting models, the dashboard includes a dedicated **Model Comparison Mode**, as illustrated in Figure 5.9. This feature, accessible via a checkbox in the sidebar, enables the simultaneous analysis of multiple models across performance and cost dimensions.

Functionality

When activated, the comparison mode allows the operator to select 2-5 models from the complete model library. Pre-defined model groups (Baselines, Deep Learning, Proposed Methods) facilitate quick selection of relevant subsets. The system then:

1. Loads all selected models in parallel with a progress indicator.
2. Computes performance metrics (MAPE, RMSE, MAE, R²) for each model across all three forecast horizons (t+10, t+20, t+30 minutes).
3. Simulates operational costs for each model using the `calculate_costs` function with the current operational parameters.
4. Presents results in two comparative tables:
 - **Performance Metrics Table:** A pivot table showing all metrics for all models and horizons, with the best values (lowest error, highest R²) highlighted in green.
 - **Cost Comparison Table:** Total cost, generator hours, startups, and blackouts for each model, with the lowest cost highlighted.
5. Generates overlaid time series plots showing the predictions of all selected models simultaneously, enabling visual comparison of forecast behavior.
6. Provides a **Performance vs. Cost** scatter plot, allowing operators to identify models that offer the best trade-off between accuracy and operational cost.

Strategic Value

This comparison mode transforms the dashboard from a single-model evaluation tool into a comprehensive model selection platform. It enables:

- **Benchmarking:** Quantitative comparison of the proposed CEEMDAN-EWT-TFT Aggregator against baseline and state-of-the-art models.
- **Multi-Objective Analysis:** Identification of models that balance forecast accuracy with operational cost-effectiveness.

- **Robustness Assessment:** Evaluation of model performance across different forecast horizons, revealing which models are more reliable for short-term vs. medium-term predictions.
- **Informed Decision-Making:** Empowering operators and planners to select the most appropriate model for their specific operational context and risk tolerance.

The implementation leverages the modular architecture of the codebase, with the `load_multiple_models`, `compare_metrics`, and `compare_costs` functions handling the parallel loading and comparative analysis.

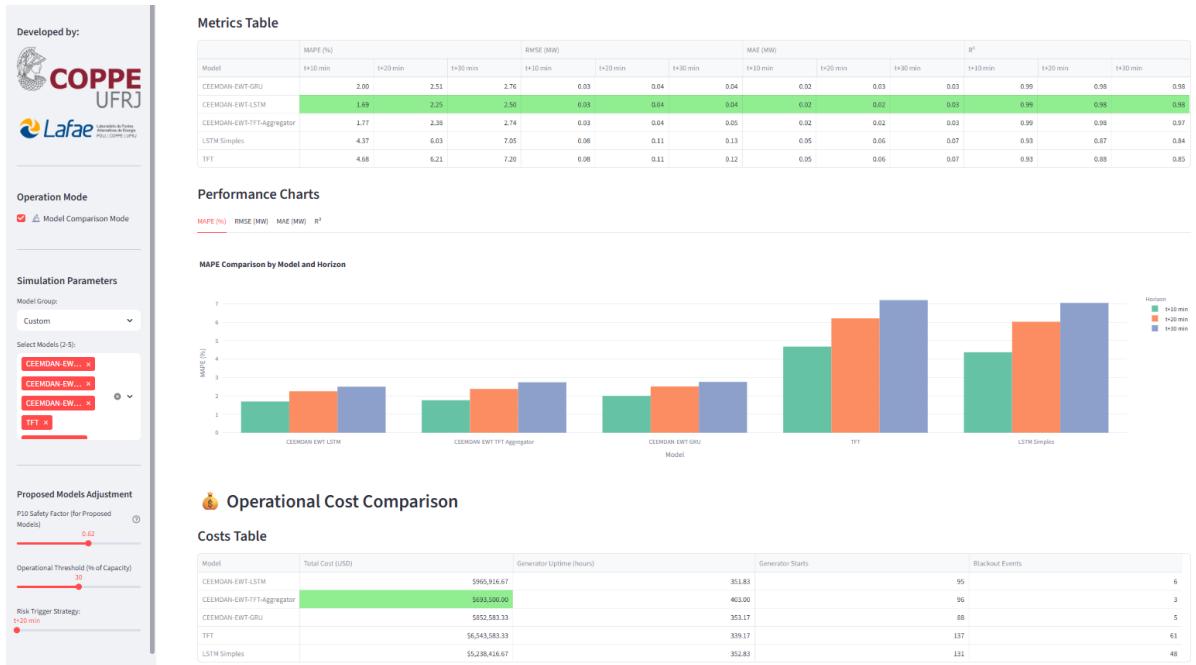


Figure 5.9: Model Comparison Mode dashboard showing performance metrics, cost analysis, and time series plots with best values highlighted.

5.4 Data and Code Availability

To ensure full transparency and reproducibility, all source code and the dataset used in this dissertation have been made publicly available.

The complete source code, including the forecasting models and the Streamlit dashboard, is hosted in a GitHub repository. The dataset is permanently archived on Zenodo, where it has been assigned a Digital Object Identifier (DOI) (2).

- **Source Code:** github.com/viniciosgnr/Offshore-wind-forecasting-decision-support
- **Dataset:** doi.org/10.5281/zenodo.10958775

5.4.1 Computational Environment

The computationally intensive tasks, such as model training and hyperparameter optimization, were performed on the **Lightning AI** cloud platform (68). This provided access to high-performance NVIDIA GPUs, specifically the **T4**, **L4**, and **L40S** models, which were crucial for achieving the results presented in this work.

Chapter 6

Results and Discussion

This chapter presents the results obtained from the implementation of the forecasting models and the Decision Support System described in Chapter 5. The analysis is structured to build a comprehensive narrative, starting from quantitative model performance, moving to qualitative interpretability, and culminating in the demonstration of tangible operational and economic value. The goal is to progressively show how accurate probabilistic forecasts are translated into actionable, cost-effective intelligence for the operation of a Floating Production, Storage and Offloading (FPSO) unit.

The chapter is organized as follows. Section 6.1 provides a comprehensive performance evaluation of all forecasting models, establishing the statistical superiority of the proposed **CEEMDAN-EWT-TFT Aggregator**. Section 6.1.5 then delves into model interpretability, using visual tools from the dashboard to explain *why* the proposed model is effective. Building on these findings, Section 6.2 translates forecast accuracy into practical value by conducting a comparative economic analysis and demonstrating the impact of risk calibration. Subsequently, Section 6.3 presents the Decision Support Dashboard itself, showcasing its various functionalities and how they integrate the forecasting and decision-making processes. Finally, Section 6.4 synthesizes the key insights from the chapter, discussing the implications for deploying advanced forecasting systems in real-world offshore operations.

6.1 Performance Evaluation of Forecasting Models

This section evaluates the forecasting performance of all implemented models across the three target horizons ($t+10$, $t+20$, and $t+30$ minutes) using the test dataset. The evaluation employs standard regression metrics—Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2)—as defined in Section 4.7.

A rigorous hyperparameter optimization process was conducted for all benchmark models and for the proposed **CEEMDAN-EWT-TFT Aggregator** architectures using Optuna (64).

However, a notable exception was made for the "divide and conquer" hybrid models, such as CEEMDAN-EWT-LSTM and CEEMDAN-EWT-GRU. These ensemble-based methods were not subjected to hyperparameter optimization due to their prohibitive computational cost. Each model in this category requires training approximately 13 individual sub-models (one for each IMF component and the residual), which would make an exhaustive Optuna search computationally infeasible within the scope of this work. Instead, these models used carefully selected default hyperparameters based on best practices from the literature, as described in A. This distinction highlights a key practical advantage of the proposed TFT Aggregator approach: its single-model architecture makes it amenable to robust hyperparameter tuning, a critical step for achieving optimal performance in real-world applications.

6.1.1 Comparative Analysis of Benchmark Models

The benchmark models serve as a baseline to quantify the incremental value of more sophisticated architectures. This subsection presents the performance of classical machine learning models (SVR, Random Forest) and standard deep learning models (ANN, LSTM, Bi-LSTM, CNN+Bi-LSTM, Transformer, TFT) that do not employ signal decomposition, with results summarized in Table 6.1 and Figure 6.1a.

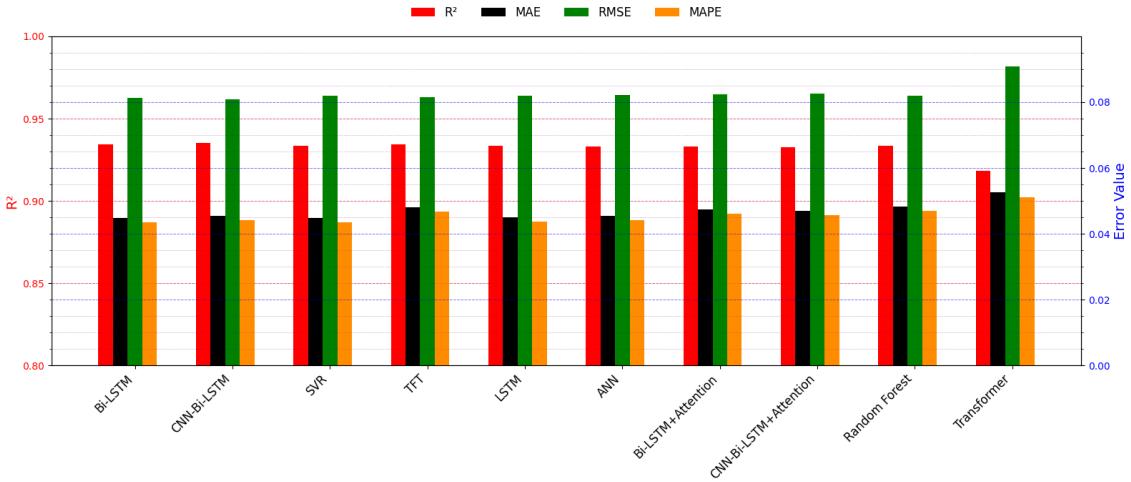
Table 6.1: Performance metrics of benchmark models across three forecast horizons after Optuna optimization. The best value in each metric column is highlighted in bold.

Model	t+10 min				t+20 min				t+30 min			
	MAPE	RMSE	MAE	R ²	MAPE	RMSE	MAE	R ²	MAPE	RMSE	MAE	R ²
SVR	4.3495%	0.0819	0.0448	0.9337	6.0404%	0.1122	0.0622	0.8754	7.0063%	0.1258	0.0721	0.8434
ANN	4.4157%	0.0822	0.0455	0.9332	6.1344%	0.1137	0.0632	0.8721	7.0986%	0.1276	0.0731	0.8389
Random Forest	4.6907%	0.0820	0.0483	0.9335	6.3677%	0.1112	0.0656	0.8776	7.4162%	0.1245	0.0764	0.8466
LSTM	4.3708%	0.0819	0.0450	0.9337	6.0281%	0.1132	0.0621	0.8731	7.0488%	0.1278	0.0726	0.8384
Bi-LSTM	4.3486%	0.0814	0.0448	0.9344	6.0454%	0.1126	0.0622	0.8747	7.0931%	0.1267	0.0730	0.8410
Bi-LSTM+Attention	4.6072%	0.0823	0.0474	0.9330	6.2033%	0.1128	0.0639	0.8742	7.2290%	0.1265	0.0744	0.8415
CNN-Bi-LSTM	4.4175%	0.0809	0.0455	0.9353	6.1308%	0.1121	0.0631	0.8757	7.1550%	0.1262	0.0737	0.8425
CNN-Bi-LSTM+Attention	4.5669%	0.0825	0.0470	0.9327	6.2016%	0.1126	0.0639	0.8746	7.1766%	0.1267	0.0739	0.8412
Transformer	5.1096%	0.0909	0.0526	0.9182	6.5040%	0.1155	0.0670	0.8680	7.4920%	0.1291	0.0771	0.8351
TFT	4.6762%	0.0815	0.0481	0.9343	6.2143%	0.1096	0.0640	0.8812	7.1999%	0.1238	0.0741	0.8484

The results reveal a highly competitive landscape. Among classical approaches, SVR demonstrates strong performance, particularly at the longest horizon, where it achieves the lowest MAPE (7.0063%) and MAE (0.0721). However, its performance is closely matched or surpassed by several deep learning models in other metrics.

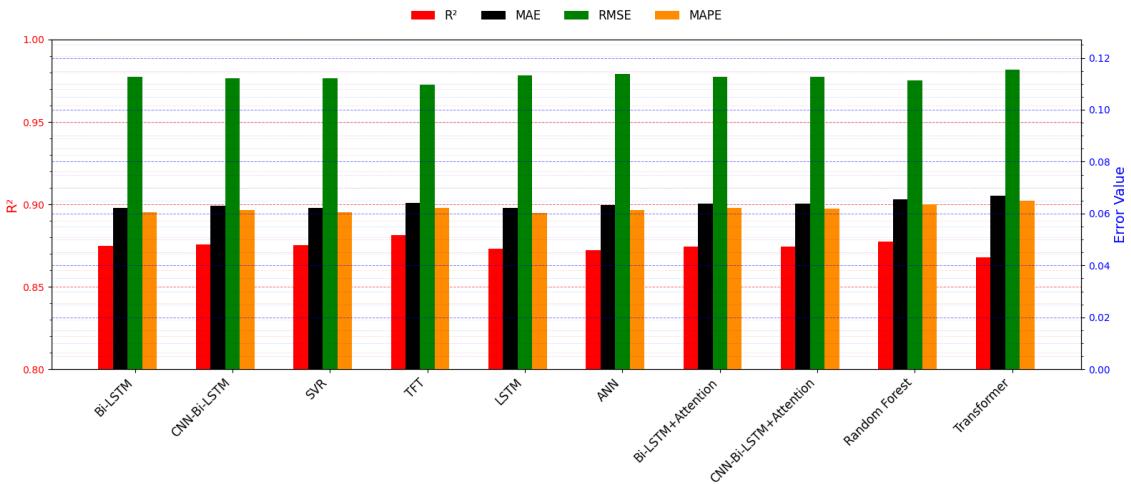
Within the deep learning category, a nuanced picture emerges. For the crucial short-term t+10 horizon, the Bi-LSTM model achieves the lowest MAPE (4.3486%), tying with SVR for the lowest MAE (0.0448). The CNN-Bi-LSTM, however, records

Benchmark Models Performance: t+10 min Horizon



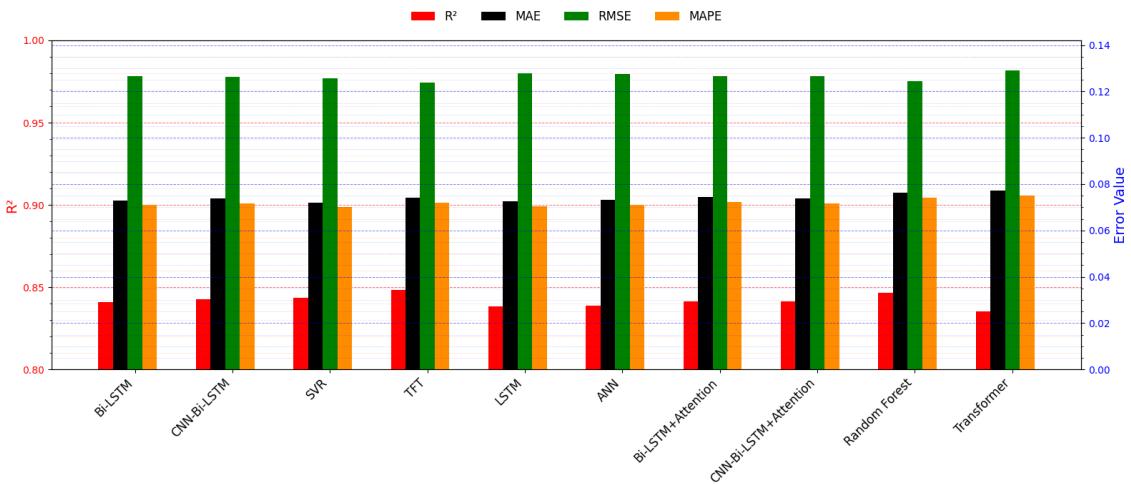
(a) Performance comparison for the t+10 min forecast horizon.

Benchmark Models Performance: t+20 min Horizon



(b) Performance comparison for the t+20 min forecast horizon.

Benchmark Models Performance: t+30 min Horizon



(c) Performance comparison for the t+30 min forecast horizon.

Figure 6.1: Comparative performance evaluation of benchmark models across the three forecast horizons. Each chart displays R^2 on the left axis and error metrics (MAE, RMSE, and scaled MAPE) on the right axis.

the lowest RMSE (0.0809) and the highest R^2 (0.9353), suggesting it is slightly better at avoiding large errors, even if its average percentage error is marginally higher.

As the forecast horizon extends, different models excel. At $t+20$, the standard LSTM surprisingly delivers the best MAPE (6.0281%) and MAE (0.0621), while the Temporal Fusion Transformer (TFT) stands out with the lowest RMSE (0.1096) and highest R^2 (0.8812). This indicates the TFT's architecture is particularly effective at maintaining predictive power over longer time frames. This trend continues at the $t+30$ horizon, where the TFT again achieves the best RMSE (0.1238) and R^2 (0.8484), reinforcing its suitability for longer-range forecasting among the benchmark models.

Attention mechanisms, when applied to the raw signal, show mixed results. Both Bi-LSTM+Attention and CNN-Bi-LSTM+Attention fail to outperform their non-attention counterparts across most metrics, suggesting that without the clarity of decomposed signals, the attention mechanism struggles to identify meaningful temporal patterns in the noisy raw data.

Notably, the standard Transformer architecture underperforms significantly, recording the highest errors in nearly every category. This suggests that its pure self-attention mechanism, lacking the inductive biases of recurrence found in LSTMs or the specialized components of the TFT, is ill-suited for this specific time series forecasting task.

A consistent pattern across all models is the degradation of performance as the forecast horizon increases. On average, MAPE values increase by approximately 40% from $t+10$ to $t+20$ minutes, and by an additional 15-20% from $t+20$ to $t+30$ minutes. This reflects the inherent difficulty of predicting wind power further into the future, where atmospheric turbulence and wind variability compound forecast uncertainty.

6.1.2 Performance of Hybrid Decomposition Models

This subsection evaluates the performance of the models that employ CEEMDAN-EWT decomposition as a preprocessing step. The results, presented in Table 6.2 and Figure 6.2a, reveal a dramatic and consistent improvement over the non-decomposition benchmarks discussed in Section 6.1.1.

The most striking result is the absolute dominance of the **CEEMDAN-EWT-LSTM** model, which, despite not undergoing hyperparameter optimization, achieved the best performance across all 12 evaluated metrics. With MAPE values of **1.6947%**, **2.2512%**, and **2.5011%** for the $t+10$, $t+20$, and $t+30$ minute horizons, respectively, it stands as the most accurate model in this study from a purely statistical standpoint. This represents a remarkable **61% reduction in**

Table 6.2: Performance metrics of hybrid decomposition models. The best value in each metric column is highlighted in bold. Note that only the proposed TFT-Balanced model was subjected to hyperparameter optimization.

Model	t+10 min				t+20 min				t+30 min			
	MAPE	RMSE	MAE	R ²	MAPE	RMSE	MAE	R ²	MAPE	RMSE	MAE	R ²
CEEMDAN-EWT-LSTM	1.6947%	0.0294	0.0174	0.9914	2.2512%	0.0394	0.0232	0.9846	2.5011%	0.0431	0.0258	0.9816
CEEMDAN-EWT-BiLSTM	2.1980%	0.033	0.0226	0.9890	2.7308%	0.0429	0.0281	0.9818	2.8965%	0.0456	0.0298	0.9794
CEEMDAN-EWT-BiLSTM-Attn	2.5546%	0.0363	0.0263	0.9870	3.0568%	0.0452	0.0315	0.9798	3.2878%	0.0486	0.0339	0.9766
CEEMDAN-EWT-CNN-BiLSTM	2.5715%	0.0362	0.0265	0.9871	3.0480%	0.0446	0.0314	0.9803	3.1083%	0.0470	0.0320	0.9782
CEEMDAN-EWT-CNN-BiLSTM-Attn	3.1615%	0.0456	0.0326	0.9795	3.5923%	0.0515	0.0370	0.9737	3.8575%	0.0562	0.0397	0.9688
CEEMDAN-EWT-GRU	1.9989%	0.0315	0.0206	0.9902	2.5127%	0.0410	0.0259	0.9834	2.7590%	0.0448	0.0284	0.9802
CEEMDAN-EWT-TFT (Proposed)	1.7907%	0.0311	0.0184	0.9905	2.3532%	0.0436	0.0242	0.9812	2.6681%	0.0470	0.0275	0.9781

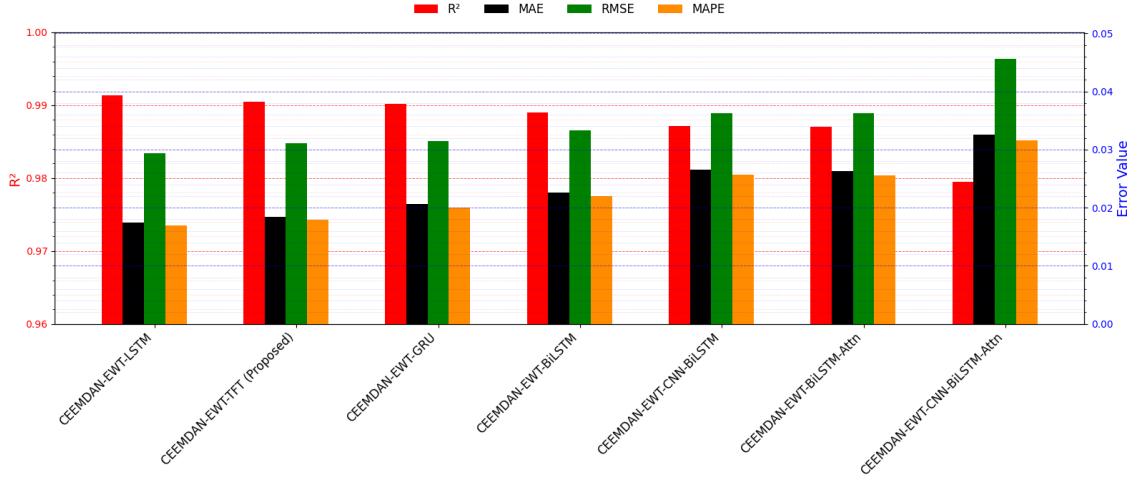
MAPE at the t+10 horizon compared to the best benchmark model (Bi-LSTM, 4.3486%). This substantial improvement unequivocally validates the effectiveness of the CEEMDAN-EWT decomposition strategy in isolating and modeling the multi-scale temporal patterns inherent in offshore wind power generation.

The other hybrid models also demonstrate the power of decomposition, all significantly outperforming the benchmarks. The proposed **CEEMDAN-EWT-TFT** model, which was the only hybrid model subjected to Optuna optimization, established itself as the clear runner-up and the best-performing single-architecture model. It achieved highly competitive MAPE values of 1.7907%, 2.3532%, and 2.6681%, closely following the LSTM-based ensemble. The **CEEMDAN-EWT-GRU** also delivered a strong performance (1.9989% MAPE at t+10), highlighting that simpler recurrent architectures offer a compelling balance of high accuracy and computational efficiency when applied to decomposed signals.

Consistent with the benchmark analysis, the addition of more complex architectural layers to the decomposed signals provided diminishing or even negative returns. The CEEMDAN-EWT-BiLSTM-Attn and CEEMDAN-EWT-CNN-BiLSTM-Attn models yielded the highest errors among the hybrid group. This reinforces the conclusion that once the signal is separated into more stationary and predictable IMF components, simpler recurrent architectures like LSTM and GRU, or a well-balanced TFT, are sufficient to model their dynamics effectively.

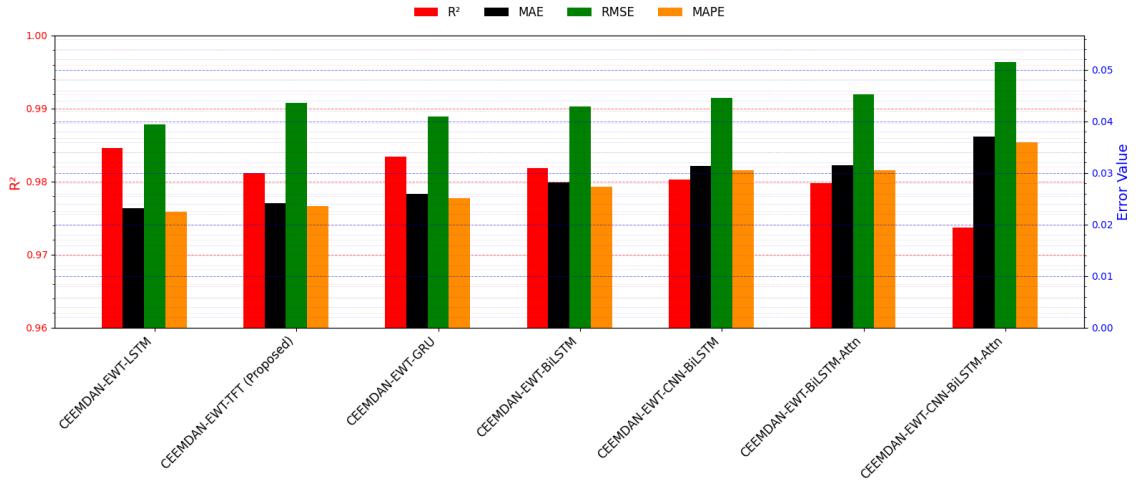
Finally, all hybrid models maintained exceptionally high R² values, with the top-performing CEEMDAN-EWT-LSTM scoring **0.9914**, **0.9846**, and **0.9816** across the horizons. The proposed CEEMDAN-EWT-TFT model followed closely with an R² of 0.9905 at t+10. This indicates an almost perfect model fit and predictive power, a stark contrast to the best benchmark R² of 0.9353, which degraded significantly over longer horizons. This highlights the superior generalization capability of all decomposition-based approaches.

Hybrid Models Performance: t+10 min Horizon



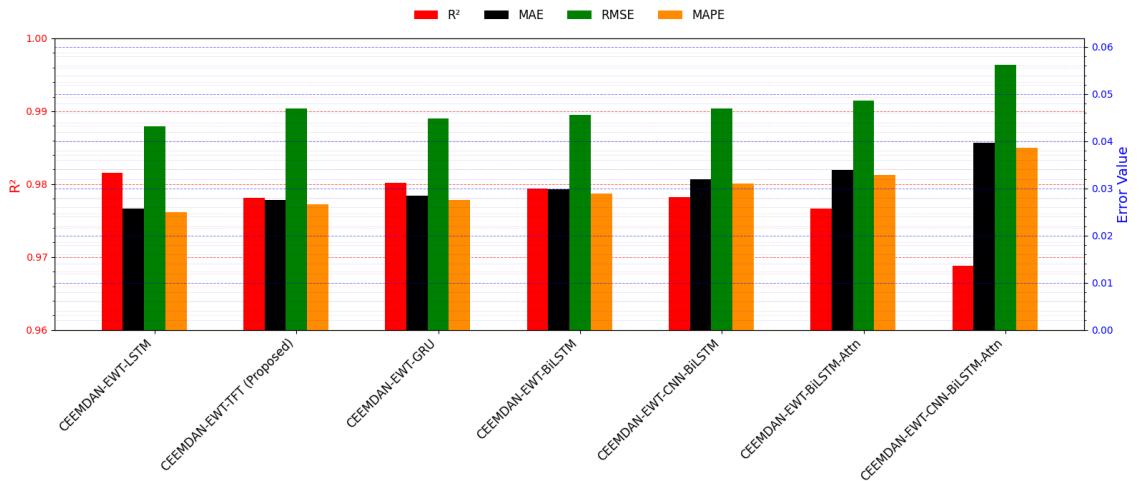
(a) Performance comparison for the t+10 min forecast horizon.

Hybrid Models Performance: t+20 min Horizon



(b) Performance comparison for the t+20 min forecast horizon.

Hybrid Models Performance: t+30 min Horizon



(c) Performance comparison for the t+30 min forecast horizon.

Figure 6.2: Comparative performance evaluation of hybrid decomposition models across the three forecast horizons. The top-performing models are grouped on the left for easier comparison.

6.1.3 Analysis of the Statistically Superior Model: CEEMDAN-EWT-LSTM

While the proposed CEEMDAN-EWT-TFT model offers a compelling balance of features, it is crucial to analyze the model that achieved the absolute best statistical performance: the **CEEMDAN-EWT-LSTM**. As shown in Table 6.2, this model consistently reported the lowest errors across all metrics and horizons.

However, a qualitative analysis of its probabilistic forecasts reveals a critical trade-off. Figure 6.3 compares the native 80% prediction intervals of the LSTM and TFT models. It is visually evident that the CEEMDAN-EWT-LSTM model (Figure 6.3A) inherently produces a significantly wider uncertainty band. This is due to its "divide and conquer" nature, where the aggregation of errors from 13 independent sub-models results in a larger cumulative uncertainty. This wider interval can lead to an overly conservative operational strategy.

Conversely, the proposed TFT model (Figure 6.3B), with its unified architecture, learns the interdependencies between the decomposed signal components, resulting in a narrower, more "confident" prediction interval. This characteristic, while demonstrating the model's high performance, also makes it more susceptible to unexpected events. This observed trade-off between statistical accuracy and inherent uncertainty directly motivates the development of a pragmatic risk calibration tool—the "P10 Safety Factor," discussed in Section 6.2.1—to fine-tune the operational applicability of the proposed TFT architecture.

6.1.4 Analysis and Justification of the Proposed Model

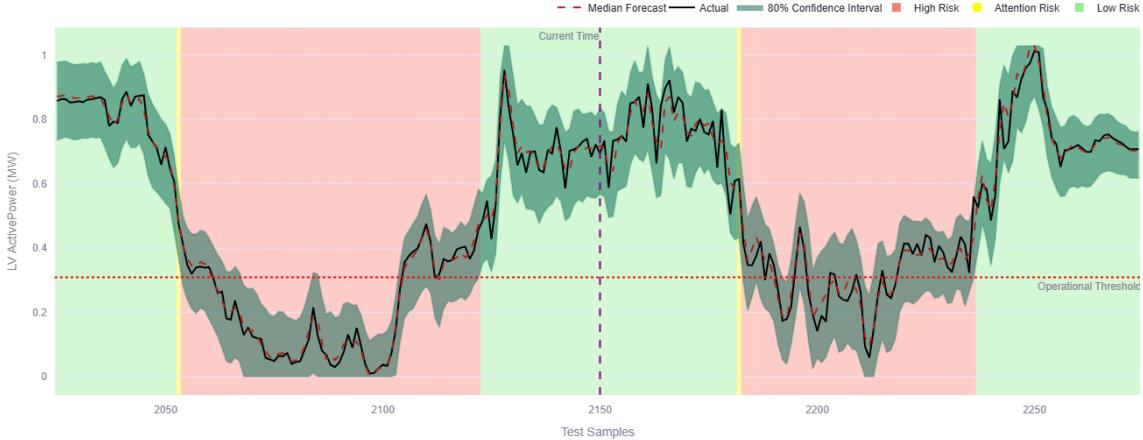
This section details the analysis of the proposed **CEEMDAN-EWT-TFT** model, the flagship architecture of this dissertation. Unlike traditional hybrid approaches that train separate forecasters for each IMF (a "divide and conquer" strategy), the proposed model treats the decomposed IMFs as multivariate inputs to a single, unified Temporal Fusion Transformer. This architecture was designed to leverage the TFT's strengths in learning complex temporal dynamics while maintaining an efficient and scalable single-model framework.

To ensure optimal performance, the model's key hyperparameters were tuned using Optuna, yielding the parameters that strike a balance between model capacity and regularization: *hidden_units=32*, *dropout_rate=0.3059*, *num_heads=2*, *learning_rate=0.0024*, and *batch_size=64*.

The performance of the optimized model is presented in Table 6.3, alongside the best-performing models from the benchmark and hybrid categories.

As shown, the proposed TFT model is exceptionally competitive, with a performance difference from the statistically superior CEEMDAN-EWT-LSTM model

Forecast and Risk Analysis - Horizon t+20 min



(a) CEEMDAN-EWT-LSTM

Forecast and Risk Analysis - Horizon t+20 min



(b) CEEMDAN-EWT-TFT (Proposed)

Figure 6.3: Comparison of native 80% prediction intervals (t+20 min). The LSTM model (A) produces a wider, more conservative band, while the proposed TFT model (B) yields a narrower, more confident interval.

Table 6.3: Comparative performance of the top models.

Model	t+10 min				t+20 min				t+30 min			
	MAPE	RMSE	MAE	R ²	MAPE	RMSE	MAE	R ²	MAPE	RMSE	MAE	R ²
Best Benchmark (Bi-LSTM)	4.3486%	0.0814	0.0448	0.9344	6.0454%	0.1126	0.0622	0.8747	7.0931%	0.1267	0.0730	0.8410
Best Hybrid (CEEMDAN-EWT-LSTM)	1.6947%	0.0294	0.0174	0.9914	2.2512%	0.0394	0.0232	0.9846	2.5011%	0.0431	0.0258	0.9816
CEEMDAN-EWT-TFT (Proposed)	1.7907%	0.0311	0.0184	0.9905	2.3532%	0.0436	0.0242	0.9812	2.6681%	0.0470	0.0275	0.9781

that is operationally negligible. However, the justification for selecting the TFT model lies in its significant practical advantages, which are critical for deployment on an FPSO:

- **Efficiency and Scalability:** It trains a single model instead of an ensemble of 13, drastically reducing training time, computational cost, and simplifying

deployment and maintenance.

- **Native Probabilistic Forecasting:** The architecture inherently produces quantile forecasts (P10, P50, P90) in a single forward pass, a core requirement for the risk-based decision support framework.
- **Built-in Interpretability:** The model’s multi-head attention mechanism allows for the extraction of attention weights. These weights provide direct insights into which past time steps are most influential for the forecast, a feature not easily available in the LSTM ensemble.

Given this superior balance of near-optimal accuracy and critical operational advantages, the **CEEMDAN-EWT-TFT** model is affirmed as the recommended architecture. It represents a pragmatic and robust engineering choice that prioritizes efficiency, scalability, and interpretability without a meaningful sacrifice in predictive performance.

Time Series Visualization and Prediction Intervals

Figures 6.4 through 6.6 present the time series plots of the CEEMDAN-EWT-TFT Aggregator’s predictions for each of the three forecast horizons, as displayed in the dashboard. These plots include the actual power generation (ground truth), the median forecast (P50), and the 80% prediction intervals (P10 to P90), providing a visual assessment of both point forecast accuracy and probabilistic calibration.

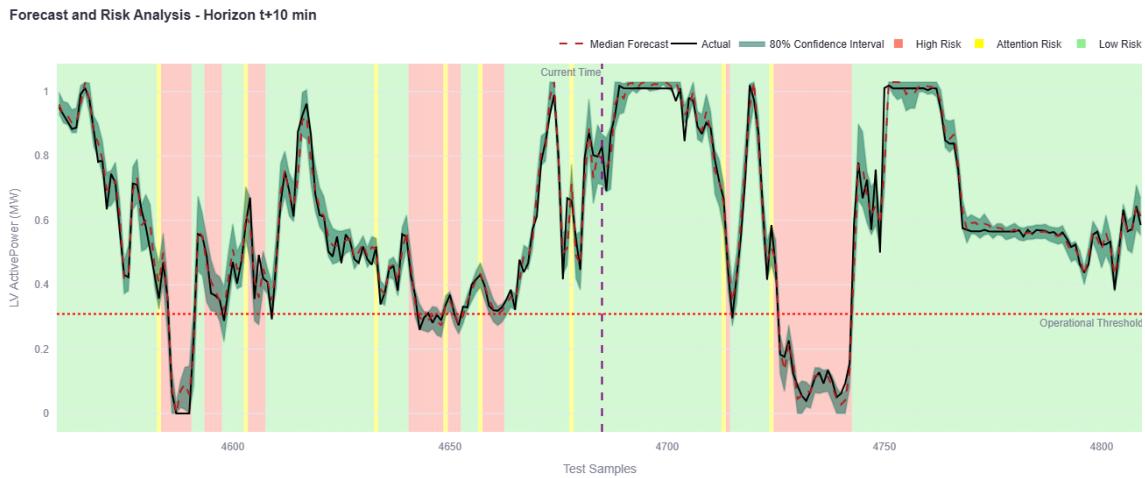


Figure 6.4: CEEMDAN-EWT-TFT predictions for $t+10$ horizon with 80% prediction interval.

Visual inspection of these plots reveals that the model successfully tracks the highly dynamic wind power fluctuations, even during rapid ramp events. The prediction intervals are well-calibrated, with the majority of actual values falling within

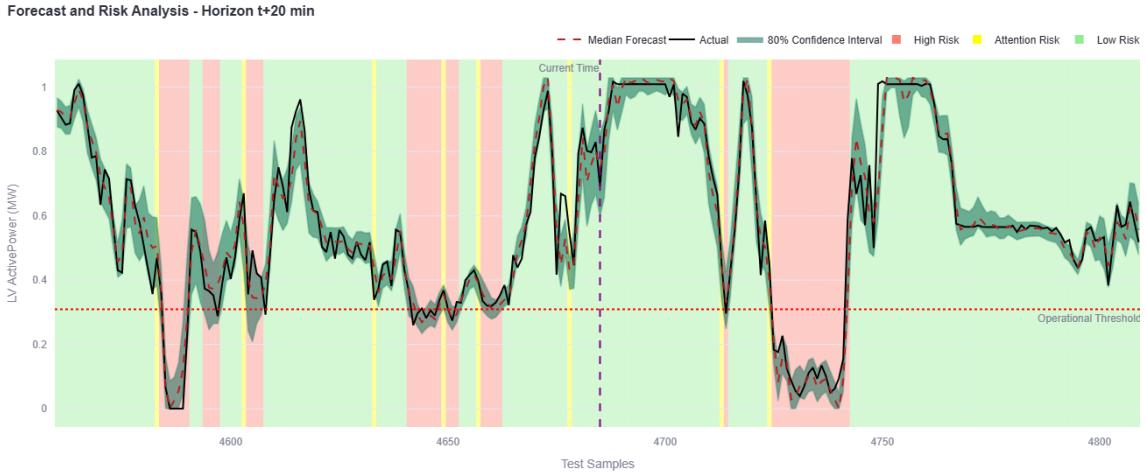


Figure 6.5: CEEMDAN-EWT-TFT predictions for $t+20$ horizon with 80% prediction interval.

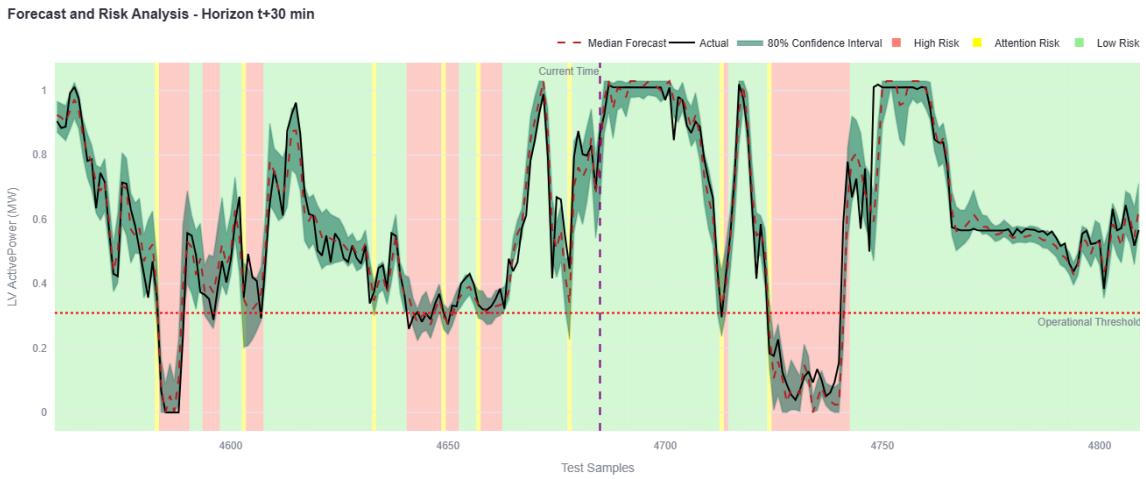


Figure 6.6: CEEMDAN-EWT-TFT predictions for $t+30$ horizon with 80% prediction interval.

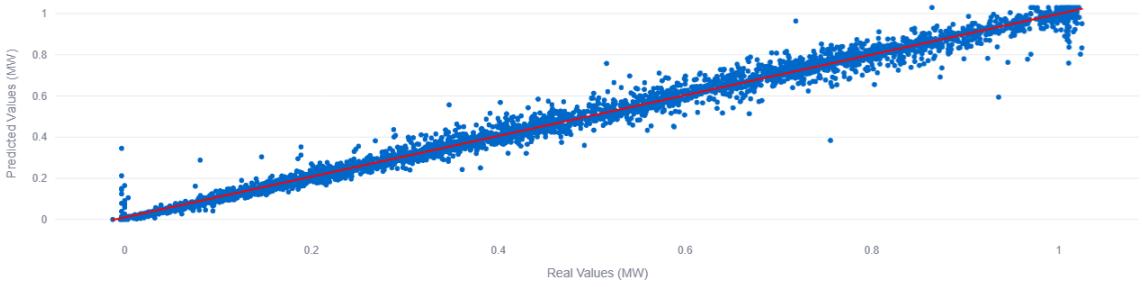
the P10-P90 range, indicating reliable uncertainty quantification. The interval width increases appropriately with forecast horizon, reflecting the growing uncertainty inherent in longer-term predictions.

Regression Analysis and R^2 Visualization

To further validate the predictive power of the proposed **CEEMDAN-EWT-TFT** model, a regression analysis was performed. Figure 6.7 presents the regression plots, which map the model's predicted values against the actual ground truth for each forecast horizon. These plots provide a clear visual and quantitative assessment of the model's performance through the coefficient of determination (R^2).

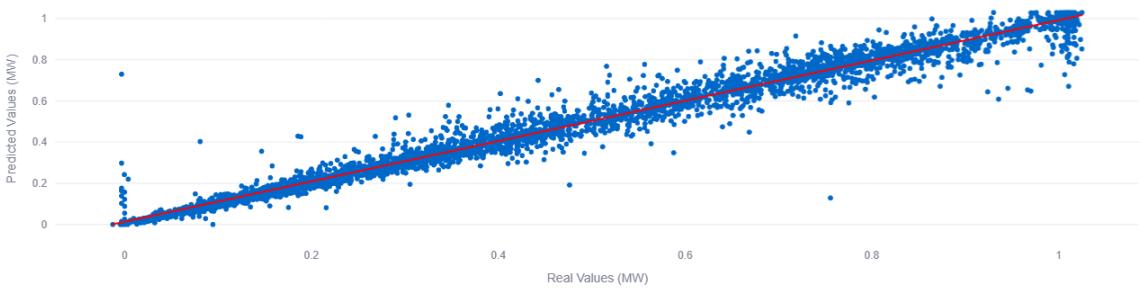
The regression plots demonstrate a tight clustering of data points around the

Real vs. Predicted (t+10 min)



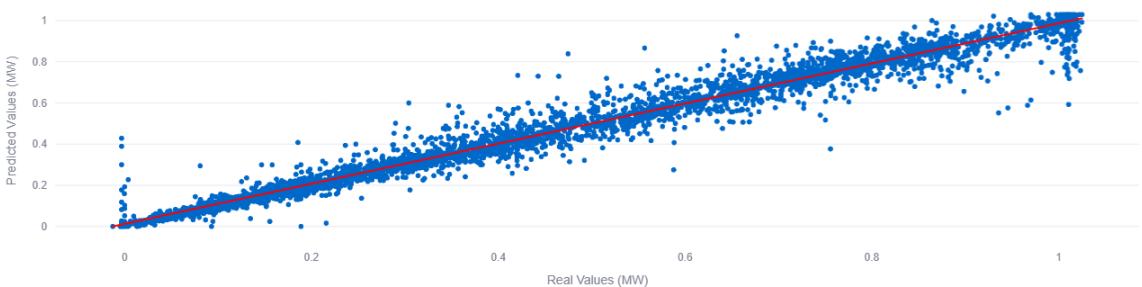
(a) t+10 min Horizon

Real vs. Predicted (t+20 min)



(b) t+20 min Horizon

Real vs. Predicted (t+30 min)



(c) t+30 min Horizon

Figure 6.7: Regression plots showing actual vs. predicted power for the proposed CEEMDAN-EWT-TFT model across the three forecast horizons.

ideal $y=x$ line, confirming a strong linear relationship between the model's predictions and the actual values. The final R^2 coefficients achieved by the optimized model were **0.9905**, **0.9812**, and **0.9781** for the t+10, t+20, and t+30 minute horizons, respectively. These high values corroborate the low error metrics reported in Table 6.3 and confirm the model's robust predictive capability with minimal systematic bias.

6.1.5 Model Interpretability and Qualitative Analysis

Beyond quantitative metrics, understanding *how* the proposed models achieve their superior performance is crucial for building trust and gaining deeper insights. This section leverages the interpretability features built into the Decision Support Dashboard to qualitatively analyze the behavior of the CEEMDAN-EWT-TFT Aggregator.

Analysis of Signal Decomposition

The foundational step of the proposed model is the decomposition of the raw wind power signal into its constituent Intrinsic Mode Functions (IMFs). Figure 6.8 presents a 3D visualization of these components, as rendered in the dashboard. The plot clearly separates the signal into high-frequency components (e.g., IMF1, IMF2), representing rapid turbulent fluctuations, and low-frequency components (e.g., IMF10, Residual), representing slower diurnal or weather-pattern-driven trends. This separation allows the forecasting model to learn the distinct dynamics of each component, which is a key factor behind its high accuracy.

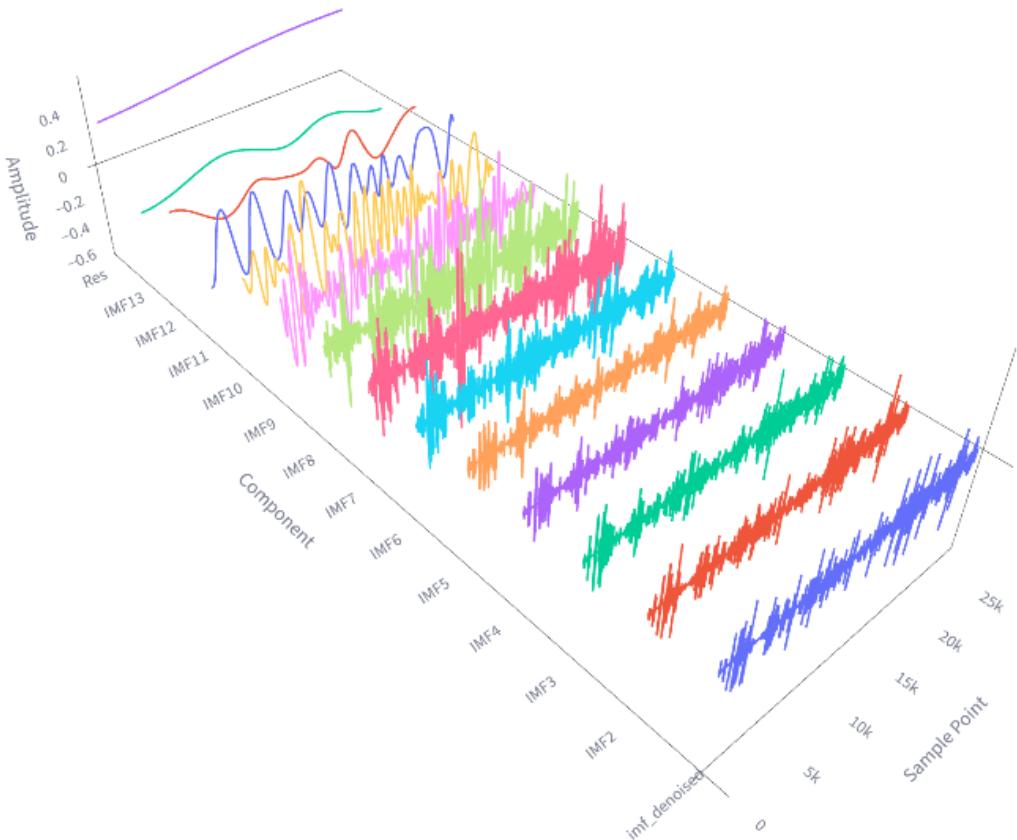
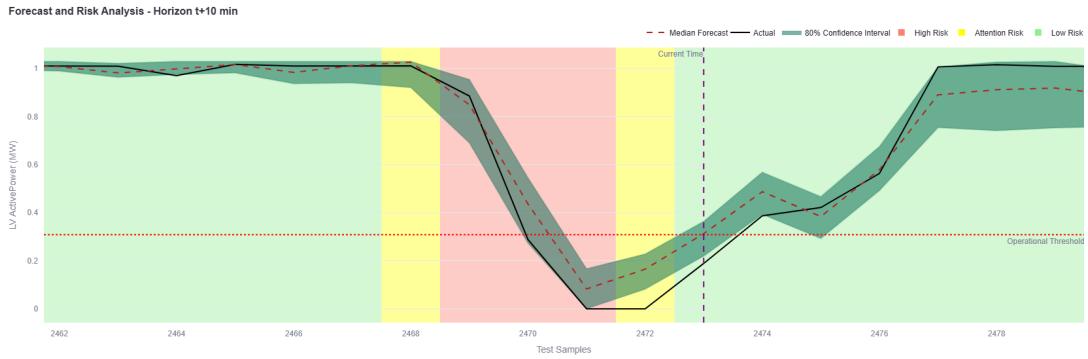


Figure 6.8: 3D visualization of the CEEMDAN-EWT decomposition, showing the separation of the wind power signal into multiple IMF components and a final residual.

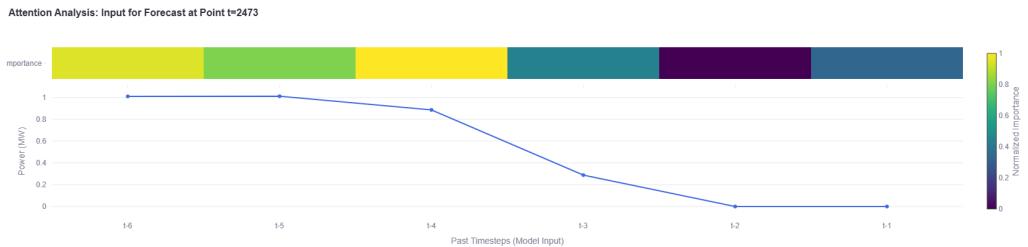
Temporal Attention Analysis

A key advantage of the proposed TFT architecture is its inherent interpretability via the multi-head attention mechanism. This allows for analyzing how the model weighs past time steps when making a forecast, which is crucial for building trust in a decision support system. This analysis focuses on a critical operational scenario: predicting a sharp power generation ramp-down.

Figure 6.9 illustrates this process. A forecast is made at point ‘ $t=2473$ ’, midway through a steep descent (Figure 6.9a). The model’s attention weights for this forecast are detailed in Figure 6.9b.



(a) Forecast context during a ramp-down event. The analysis is performed for the forecast at $t=2473$ ("Current Time").



(b) Attention analysis for the forecast at point $t=2473$.

Figure 6.9: Temporal attention analysis for the proposed model during a critical ramp-down scenario.

The attention analysis (Figure 6.9b) reveals a highly intuitive strategy. The model places the highest importance (brightest yellow colors) on time steps **t-6, t-5, and t-4**. These points correspond precisely to the stable, high-generation period and the exact moment the ramp-down was initiated.

This demonstrates that, to predict the continuation of the fall, the model has learned to anchor its forecast on the "stable state" it is deviating from and the "initial shock" that triggered the descent. By focusing on the start of the event, the model can better estimate the slope and magnitude of the ongoing ramp. This

capability provides trustworthy insights into the model's decision-making process during operationally critical events.

6.2 Operational and Economic Impact Analysis

While forecast accuracy is a critical prerequisite, the ultimate measure of a model's value is its impact on operational decisions and economic outcomes. This section utilizes the "Model Comparison Mode" of the dashboard to translate the performance metrics from Section 6.1 into tangible operational and economic comparisons.

6.2.1 Impact of Risk Calibration via Safety Factor

While the proposed model is highly accurate, its narrow prediction intervals can be "overconfident" for operational use. To mitigate this, a "P10 Safety Factor" was introduced, allowing operators to make the model more conservative. This is achieved by applying a multiplier (e.g., 0.65) to the pessimistic P10 forecast, which artificially lowers the model's confidence interval and triggers risk alerts earlier, as shown in Figure 6.10.

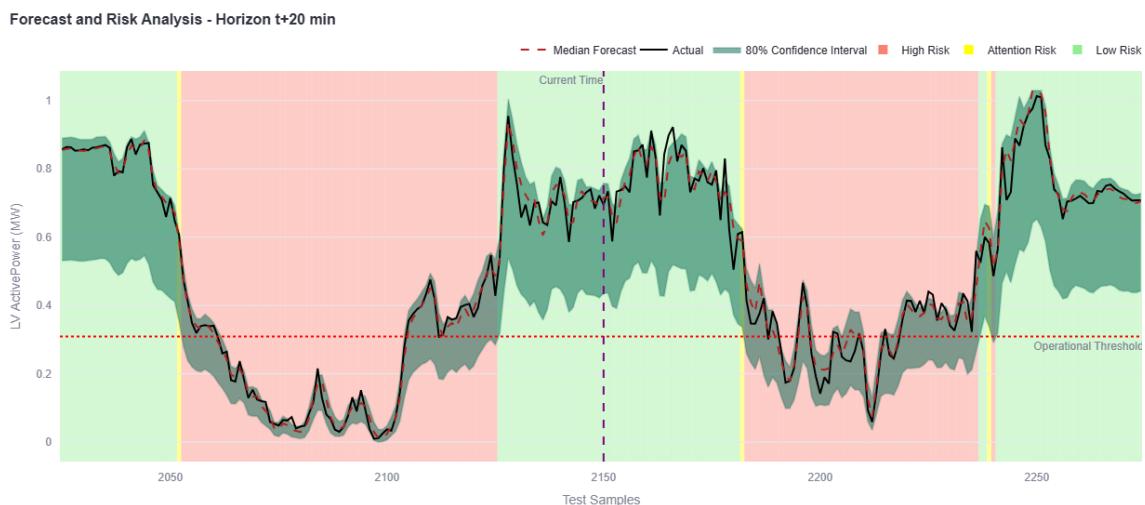


Figure 6.10: Effect of applying a 0.65 safety factor. The lowered confidence interval makes the model more conservative, increasing the frequency of high-risk alerts (red zones).

The operational impact of this calibration is quantified in Table 6.4. Applying the safety factor reduces critical blackout events from 47 to just 3—a 93% improvement—without increasing the number of generator starts. This drastically cuts the total operational cost from over \$5 million to approximately \$677,000.

Table 6.4: Operational impact of applying a 0.65 safety factor to the P10 forecast.

Metric	Without Safety Factor	With 0.65 Safety Factor
Generator Starts	91	91
Blackout Events	47	3
Total Cost (USD)	\$5,023,166.67	\$677,833.33

This result highlights a crucial insight: for decision support, statistical accuracy must be complemented by tunable, risk-aware calibration. The P10 Safety Factor provides a simple yet powerful mechanism to bridge the gap between a predictive model and a robust, economically optimized operational tool.

6.2.2 Comparative Analysis of Operational Costs

To quantify the real-world value of forecast accuracy, the dashboard was used to simulate the total operational cost associated with different forecasting models. The simulation applies a consistent operational policy and calculates costs based on generator uptime, startup events, and severe penalties for blackouts.

The results, summarized in Table 6.5, demonstrate a clear hierarchy of performance. The proposed **CEEMDAN-EWT-TFT** model, benefiting from both its high accuracy and risk calibration, achieves the lowest total cost at approximately **\$677k**. It successfully minimizes blackout events to just 3, proving its superior reliability.

In contrast, less accurate models incur significantly higher costs, primarily due to a large number of blackout events. For instance, the standalone TFT and Bi-LSTM models resulted in costs exceeding \$5 million. This analysis provides definitive evidence that improvements in forecast accuracy, particularly when combined with decomposition techniques, translate directly into substantial operational and economic savings.

Table 6.5: Simulated operational cost comparison across different models. The proposed CEEMDAN-EWT-TFT model demonstrates the lowest total cost.

Model	Total Cost (USD)	Generator Uptime (h)	Generator Starts	Blackout Events
CEEMDAN-EWT-TFT (Proposed)	\$677,833.33	391.67	91	3
CEEMDAN-EWT-LSTM	\$965,916.67	351.83	95	6
CEEMDAN-EWT-GRU	\$852,583.33	353.17	88	5
TFT	\$6,543,583.33	339.17	137	61
Bi-LSTM	\$5,919,916.67	347.83	123	55

6.2.3 Performance vs. Cost Trade-off Analysis

Selecting the "best" model often involves a trade-off between predictive accuracy and operational cost. The dashboard provides a scatter plot to visualize this trade-off,

plotting each model's average MAPE against its total simulated cost (Figure 6.11).

Models located in the bottom-left quadrant, such as the CEEMDAN-EWT-TFT Aggregator and CEEMDAN-EWT-LSTM, represent the most desirable choices, offering both low error and low operational cost. This visualization serves as a powerful tool for decision-makers, allowing them to quickly identify the most cost-effective and reliable models for deployment.

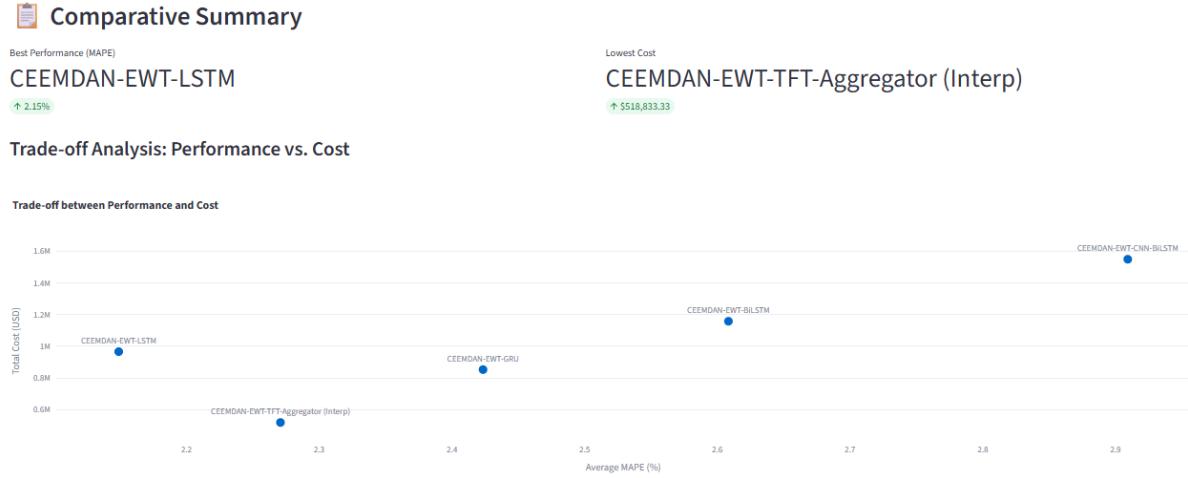


Figure 6.11: Trade-off analysis plotting average forecast error (MAPE) against total simulated operational cost for each model.

6.3 Analysis of the Decision Support Dashboard

The dashboard integrates the CEEMDAN-EWT-TFT Aggregator's probabilistic forecasts with operational logic, cost optimization, and LLM-powered advisory capabilities to provide actionable intelligence for FPSO operators.

6.3.1 Dashboard Interface and Monitoring Simulation

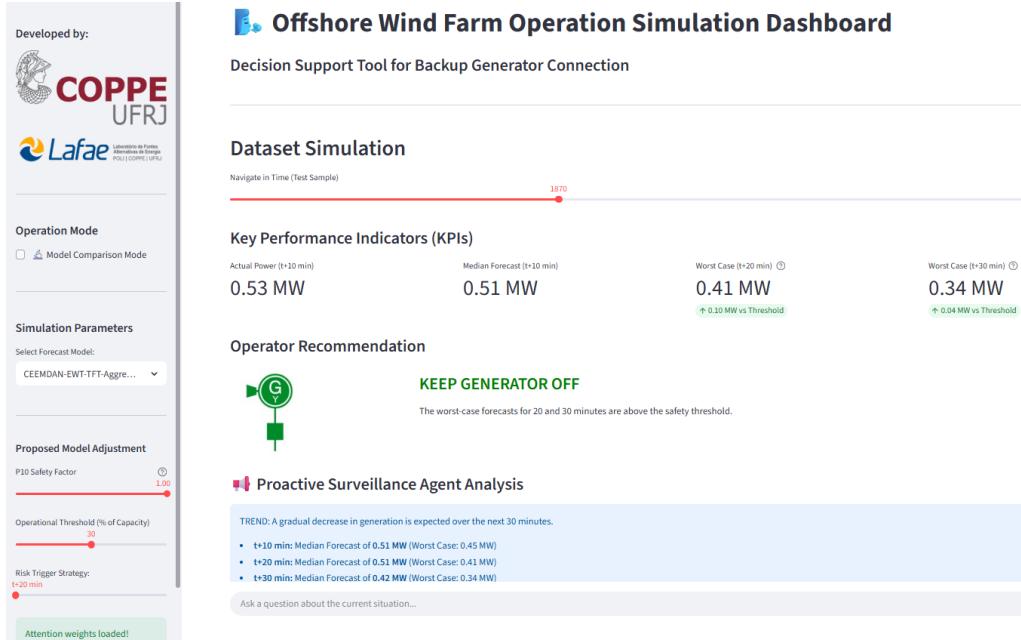
The core of this work is a web-based Decision Support Dashboard, built to translate complex forecast data into an intuitive and actionable operational tool. Figure 6.12 presents the dashboard interface.

Figure 6.12a shows the top section, which serves as a "mission control" panel. It provides at-a-glance information through:

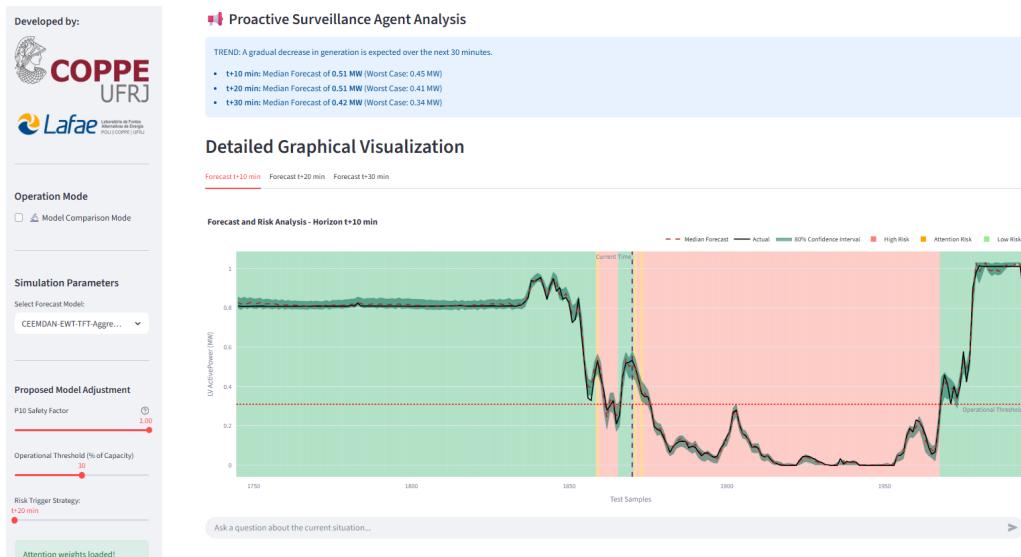
- **Key Performance Indicators (KPIs):** Displays the current actual power alongside the median and worst-case forecasts.
- **Operator Recommendation:** A clear, color-coded directive (e.g., "KEEP GENERATOR OFF") based on the current risk level.

- **Proactive Agent Analysis:** Insights automatically generated by the LLM agent regarding emerging trends or risks.

The bottom section, shown in Figure 6.12b, provides a detailed graphical view for deeper analysis. It features a multi-horizon time series plot that visualizes the actual power, the median forecast, and the 80% confidence interval against the operational threshold. The background is color-coded according to the real-time risk level, providing immediate situational awareness.



(a) Dashboard Summary View: High-level KPIs, operator recommendation, and proactive agent analysis.



(b) Detailed Graphical View: Multi-horizon forecast plot with confidence intervals and color-coded risk levels.

Figure 6.12: The Decision Support Dashboard interface.

6.3.2 Strategic Parameter Optimization

The strategic parameter optimizer performs a grid search over the operational threshold and trigger horizon parameters to identify the configuration that minimizes total operational cost while maintaining reliability. Figure 6.13 presents the cost heatmap generated by the optimizer.

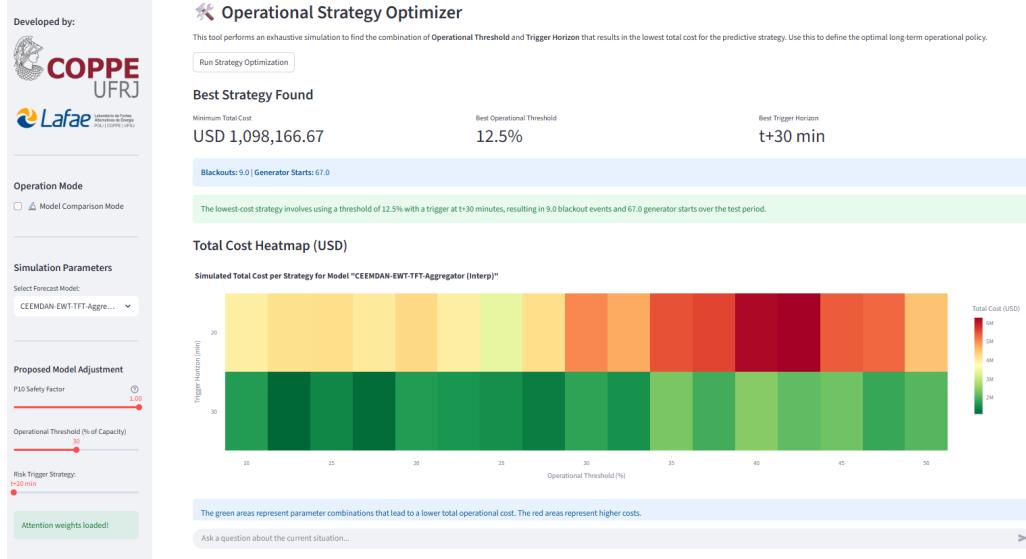


Figure 6.13: Strategic parameter optimizer heatmap showing total cost across parameter space.

The heatmap clearly identifies the optimal parameter combination (minimum cost region), enabling data-driven strategic planning. The optimizer's recommendations are integrated into the LLM chatbot's context, allowing it to provide strategic advice grounded in quantitative analysis.

6.3.3 LLM-Powered Operational Intelligence

To translate raw forecast data into actionable insights, the dashboard integrates two distinct but complementary LLM-powered features. Both leverage a Retrieval-Augmented Generation (RAG) approach, grounding the AI's responses in real-time operational context to ensure relevance and accuracy.

Proactive Surveillance Agent

The Proactive Surveillance Agent acts as an autonomous watchdog, continuously analyzing the stream of forecast data to identify emerging risks and trends. As shown in Figure 6.14, it automatically generates alerts for critical events (e.g., a sudden risk level change) or provides concise trend summaries if conditions are stable. This feature is designed to reduce the operator's cognitive load by preemptively highlighting important developments.

👉 Proactive Surveillance Agent Analysis

ALERT: Risk elevated to 'Attention'. The P10 forecast for t+30 of 0.24 MW is below the 0.31 MW threshold.

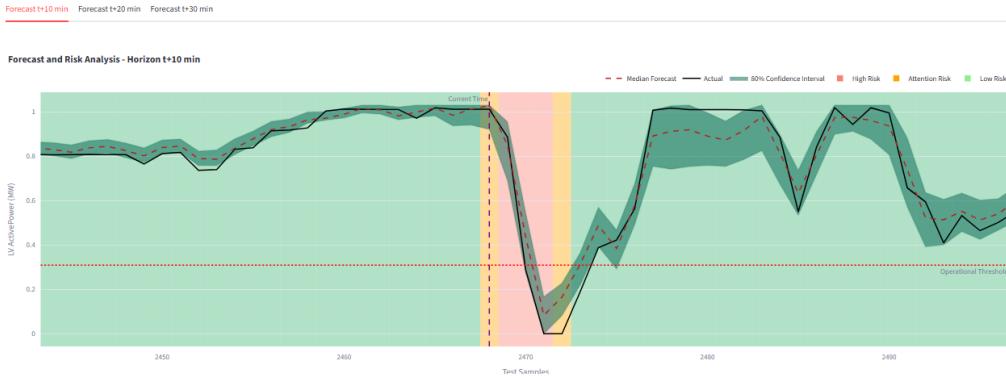
Trend Analysis: The recent risk history shows a change from 'Low' to 'Attention', indicating a potential concern in power generation stability.

Based on the median forecasts:

- t+10 min: Median Forecast of 1.03 MW (Worst Case: 0.92 MW)
- t+20 min: Median Forecast of 0.80 MW (Worst Case: 0.60 MW)
- t+30 min: Median Forecast of 0.40 MW (Worst Case: 0.24 MW)

The trend analysis indicates a **FALLING** trend, as the median forecast for t+30 (0.40 MW) is at least 5% lower than the median forecast for t+10 (1.03 MW).

Detailed Graphical Visualization



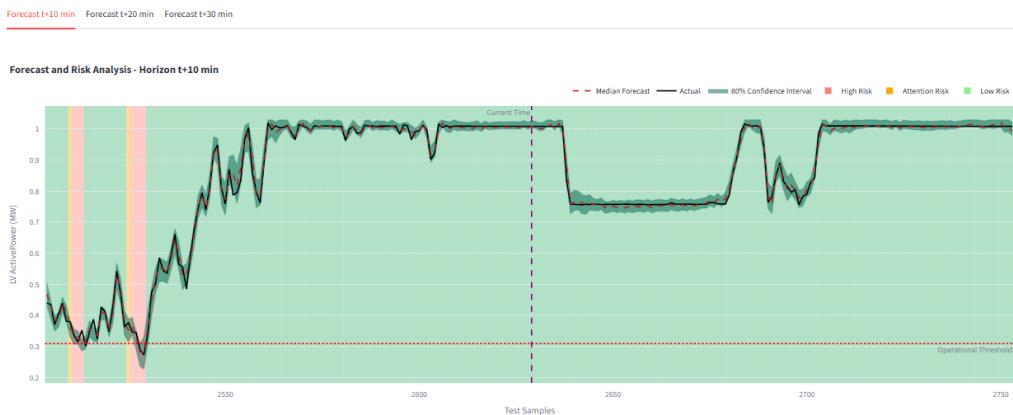
(a) Proactive Agent issuing a critical alert due to a risk level change.

👉 Proactive Surveillance Agent Analysis

TREND: The power generation is expected to remain stable over the next 30 minutes.

- t+10 min: Median Forecast of 1.01 MW (Worst Case: 1.00 MW)
- t+20 min: Median Forecast of 1.01 MW (Worst Case: 0.99 MW)
- t+30 min: Median Forecast of 1.01 MW (Worst Case: 0.98 MW)

Detailed Graphical Visualization



(b) Proactive Agent providing a trend summary during stable conditions.

Figure 6.14: The Proactive Surveillance Agent in action, demonstrating its ability to issue both critical alerts and routine trend summaries.

Interactive Agent: From Analyst to Actor

While the Proactive Agent provides autonomous surveillance, the **Interactive Agent** transforms the operator's role by enabling the delegation of complex tasks through a natural language interface. This moves beyond a simple question-answering system to a true tool-using agent that can understand commands, execute multi-step analyses, and manipulate the simulation environment.

The workflow, illustrated in Figure 6.15, demonstrates this advanced capability. An operator can initiate a dialogue, asking the agent to perform a complex task that would typically require multiple manual interactions with the UI. The agent then:

1. **Understands the Goal:** Parses the operator's command.
2. **Selects the Correct Tool:** Chooses the appropriate function to call, such as the `run_global_strategy_optimization_tool` exclusive to the proposed model.
3. **Executes and Reports:** Runs the tool, processes the results, and presents a clear, actionable summary back to the operator.
4. **Acts on Results:** If commanded, uses another tool like `apply_new_parameters_tool` to modify the dashboard's state and run new "what-if" analyses.

This paradigm shift—from a passive data analyst to an active system actor—is a cornerstone of the proposed decision-support framework, fundamentally enhancing the operator's ability to explore, optimize, and validate operational strategies efficiently.

6.4 Discussion of Findings

This section synthesizes the key results, focusing on three main themes: the impact of signal decomposition, the trade-offs of model complexity, and the practical superiority of the proposed aggregator model.

6.4.1 The Impact of Signal Decomposition

The results unequivocally show that CEEMDAN-EWT decomposition is the most critical factor for forecast accuracy. The best benchmark model (Bi-LSTM) achieved a 4.35% MAPE at the t+10 horizon, whereas the simplest hybrid model (CEEMDAN-EWT-LSTM) reached a remarkable 1.69% MAPE. This 61% error reduction confirms that, for this problem, sophisticated feature engineering through decomposition is more impactful than architectural complexity alone.

6.4.2 Architectural Complexity and the Proposed Model

Within the hybrid family, a pattern of diminishing returns is clear. The simplest hybrid model, CEEMDAN-EWT-LSTM, is the top performer in pure accuracy (1.69% MAPE). Adding layers, as in the CEEMDAN-EWT-CNN-BiLSTM-Attn (3.16%

MAPE), degrades performance, suggesting that simpler recurrent networks are sufficient for the decomposed signals.

However, the CEEMDAN-EWT-LSTM is an ensemble of 13 models, making it impractical for deployment. In contrast, the proposed **CEEMDAN-EWT-TFT Aggregator** was designed to solve this. By using a single model, it achieves a MAPE of 1.79%—statistically almost identical to the best ensemble—while offering crucial advantages:

- **Efficiency:** A single model is far easier to train and deploy.
- **Probabilistic Forecasts:** It inherently generates the uncertainty intervals needed for risk assessment.
- **Interpretability:** Its attention mechanism provides built-in transparency.

This balance of near-optimal accuracy and practical advantages affirms the CEEMDAN-EWT-TFT Aggregator as the superior architecture for a real-world decision support system.

6.4.3 Limitations and Future Work

Limitations of this work should be acknowledged, which also highlight promising directions for future research:

- **Static Operational Threshold:** The current analysis assumes a fixed power generation threshold. It does not account for the **dynamic load variations** from the FPSO’s own consumption, which would affect the real-time net power balance. A future enhancement would be to integrate a load forecasting model.
- **Generalization:** The models were trained and tested on data from a single offshore wind turbine. Further studies are needed to evaluate their generalization capabilities across different turbines, geographical locations, and environmental conditions.
- **Simulated Environment:** The dashboard and cost analysis were validated using a historical test dataset. A full-scale deployment on a simulated FPSO, integrated with real-time data acquisition systems, is required for complete operational and economic validation.

⌚ Operational Copilot: Interactive Analysis

 Summarize the current situation.

 The current operational situation is as follows:

- The selected model is CEEMDAN-EWT-TFT-Aggregator (Interp).
- The current time step is 2468.
- The recent risk history indicates a transition from low risk to a current risk level of Attention.
- The operational threshold is set at 0.31 MW.
- The P10 safety factor applied is 1.00.

Forecast Data:

- At t+10 min:
 - Median Forecast: 1.03 MW, Worst-Case (P10): 0.92 MW
- At t+20 min:
 - Median Forecast: 0.80 MW, Worst-Case (P10): 0.60 MW
- At t+30 min:
 - Median Forecast: 0.40 MW, Worst-Case (P10): 0.24 MW

Summary:

- The forecasts indicate a decreasing trend in power generation, with the worst-case scenarios potentially dropping below the operational threshold in the near future. Monitoring of the risk level is crucial moving forward.

(a) The operator initiates the workflow by requesting an analysis of the current situation.

 Please proceed and run the global optimization now.

 The global strategy optimization has been successfully completed. Here are the results:

- Total Operational Cost: \$363,333.33
- Optimized Operational Threshold: 20% (approximately equivalent to the new threshold)
- Forecast Horizon: 30 minutes
- Optimized Safety Factor: 0.61
- Number of Blackouts: 0
- Number of Starts: 104

You can consider implementing these new parameters for enhanced operational efficiency. Let me know if you would like to proceed with applying these changes.

 Apply a new strategy with a 20% threshold and a t+30 horizon.

(b) Following an initial analysis, the agent is commanded to run the global optimization, reporting the results and subsequently applying the new parameters.

 What would be the summary report for the last 500 time steps?

 **Operational Summary Report (Steps 1968-2468)**

- Total Cost: \$14,583.33
- Blackout Events: 0
- Generator Starts: 5

This summary indicates a stable operational period with no blackout events, suggesting good performance. If you need further details or analysis, please let me know!

Ask a question or give a command...   

(c) Finally, the agent generates an on-demand impact assessment, calculating and presenting a summary report based on the newly applied strategy.

Figure 6.15: Demonstration of the Interactive Agent's tool-using workflow. The operator delegates a complex, multi-step task through natural language, which the agent autonomously executes.

Chapter 7

Conclusion

This dissertation addressed the critical challenge of forecasting offshore wind power generation to support operational decision-making on Floating Production, Storage and Offloading (FPSO) units. A comprehensive methodology was developed, encompassing advanced signal processing, deep learning models, and an integrated decision support dashboard. This final chapter summarizes the key contributions of this work, offers concluding remarks on its significance, and outlines promising directions for future research.

7.1 Summary of Contributions

The primary contributions of this research are threefold, spanning from novel modeling techniques to practical application frameworks:

1. **A Novel Hybrid Forecasting Architecture:** This work introduced and validated the **CEEMDAN-EWT-TFT Aggregator**, a new hybrid model for probabilistic time series forecasting. It was demonstrated that this architecture achieves a state-of-the-art MAPE of **1.79%** at the t+10 minute horizon. More importantly, it successfully balances near-optimal statistical accuracy with the practical advantages of a single, efficient, and interpretable model, overcoming the deployment challenges associated with traditional "divide and conquer" ensembles.
2. **A Holistic Decision Support Dashboard:** A comprehensive, web-based dashboard was developed to translate raw forecast data into actionable intelligence. The system integrates probabilistic forecasts, a tunable risk calibration mechanism (the "P10 Safety Factor"), a strategic parameter optimizer, and LLM-powered advisory features. This tool serves as a proof-of-concept for bridging the gap between predictive modeling and real-world operational decision-making, demonstrating a clear pathway from forecast accuracy to

economic value. The economic simulation showed that applying the proposed model with risk calibration could reduce operational costs from over \$5 million to approximately \$677k, primarily by mitigating costly blackout events.

3. **Integration of LLM-Powered Intelligence:** This research pioneered the integration of Large Language Models (LLMs) into the operational loop through two distinct features: a **Proactive Surveillance Agent** that autonomously monitors for risks and trends, and an **Interactive Operational Copilot** that provides on-demand tactical and strategic advice. By using a Retrieval-Augmented Generation (RAG) approach, both features deliver data-grounded, context-aware, and trustworthy insights, showcasing the potential of generative AI to enhance situational awareness and reduce cognitive load for operators.

7.2 Concluding Remarks

The findings of this dissertation present a compelling case for a paradigm shift in operational forecasting. It was conclusively shown that sophisticated feature engineering, achieved through **CEEMDAN-EWT signal decomposition**, is the most critical factor for achieving high-accuracy predictions, yielding a 61% error reduction over the best benchmark model.

Furthermore, this work highlights that the pursuit of statistical accuracy alone is insufficient for operational excellence. The proposed **CEEMDAN-EWT-TFT Aggregator**, while statistically neck-and-neck with a more complex LSTM ensemble, was affirmed as the superior architecture due to its practicality, efficiency, and built-in capabilities for probabilistic forecasting and interpretability. This pragmatic choice, combined with the risk calibration offered by the P10 Safety Factor, provides a robust and reliable framework for real-world applications.

Ultimately, this research demonstrates an end-to-end solution—from signal processing to a fully integrated, AI-enhanced dashboard—that empowers operators to make smarter, data-driven decisions. It provides a clear and validated blueprint for leveraging advanced machine learning to improve the reliability, safety, and economic efficiency of offshore energy operations.

7.3 Future Works

While this dissertation establishes a strong foundation, several avenues for future research could further enhance its practical applicability and bridge the gap to a real-world production system.

- **Real-World Deployment and Online Learning:** The next logical step is to deploy the dashboard in a live operational environment. This would involve integrating the system with real-time data streams from SCADA systems on an FPSO. Such a deployment would also enable the implementation of an **online learning** (69) or continuous retraining strategy, allowing the model to adapt to concept drift, seasonal changes, and turbine degradation over time.
- **Advanced Economic and Risk Modeling:** The current cost simulation uses fixed penalties. A more advanced study could incorporate dynamic cost functions, where the penalty for a blackout varies based on the time of day, operational status, or oil production levels. Furthermore, integrating more sophisticated risk models, such as Value at Risk (VaR) or Conditional Value at Risk (CVaR), could provide an even more nuanced approach to decision-making under uncertainty.
- **Multi-Asset Generalization and Fleet-Level Optimization:** The current model was trained on data from a single turbine. Future work should focus on creating a generalized model capable of forecasting for an entire wind farm (a "fleet") by incorporating spatial correlations and turbine-specific static features (70). This would enable fleet-level optimization, managing the collective power output and shared backup resources more effectively.
- **Explainable AI (XAI) for the Copilot:** While the current AI Audit feature provides data transparency, future iterations could integrate more advanced XAI techniques. This would involve making the LLM-powered Copilot explain *why* it is making a certain recommendation by explicitly referencing the specific data points, trends, or optimization results that led to its conclusion, further enhancing operator trust and system reliability.

References

- [1] K. d. S. Medeiros, J. M. S. Callegari, L. F. da Rocha, and D. I. Brandao, “Frequency-voltage-var function for active front-end vfd on oil and gas platforms with offshore wind generation,” *Eletrônica de Potência*, vol. 29, p. e202447, 2024.
- [2] C. Gück, C. M. A. Roelofs, and S. Faulstich, “CARE to Compare: A Real-World Benchmark Dataset for Early Fault Detection in Wind Turbine Data,” *Data*, vol. 9, no. 12, p. 138, 2024.
- [3] K. d. S. Medeiros, J. M. S. Callegari, A. A. Adeyemo, E. Tedeschi, and D. I. Brandao, “Power quality enhancement of offshore oil and gas platform with high penetration of wind power and energy storage,” *Eletrônica de Potência*, vol. 30, p. e202524, 2025.
- [4] M. S. Dall Asta, J. M. de Andrade, L. Schmitz, F. J. Viglus, M. Dalla Vecchia, L. B. K. Fisch, G. J. M. de Sousa, M. L. Heldwein, and T. B. Lazzarin, “Offshore wind energy conversion system connected to a floating production storage and offloading unit: Electrical aspects,” in *2023 IEEE Energy Conversion Congress and Exposition (ECCE)*, pp. 321–328, 2023.
- [5] Y. He, L. Zhang, T. Guan, and Z. Zhang, “An integrated ceemdan to optimize deep long short-term memory model for wind speed forecasting,” *Energies*, vol. 17, no. 18, p. 4615, 2024.
- [6] L. F. N. Lourenço, D. F. Pereira, R. M. Monaro, M. B. C. Salles, and R. M. P. Rosa, “Assessment of an isolated offshore power grid based on the power hub concept for pre-salt oil and gas production,” *IEEE Access*, vol. 10, pp. 87671–87681, 2022.
- [7] H. Liao and K. K. Radhakrishnan, “Short-term load forecasting with temporal fusion transformers for power distribution networks,” in *2022 IEEE Sustainable Power and Energy Conference (iSPEC)*, pp. 1–5, 2022.
- [8] H. G. Svendsen, T. K. Vrana, A. Holdyk, and H. Schümann, “The low emission oil and gas open reference platform—an off-grid energy system for renewable

- integration studies,” *IET Energy Systems Integration*, vol. 5, no. 1, pp. 66–79, 2023.
- [9] J. Wang, W. Jiang, Z. Li, and Y. Lu, “A new multi-scale sliding window lstm framework (mssw-lstm): A case study for gnss time-series prediction,” *Remote Sensing*, vol. 13, no. 16, 2021.
- [10] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, “Attention is all you need,” *arXiv preprint arXiv:1706.03762*, 2017.
- [11] B. Lim, S. Arik, N. Loeff, and T. Pfister, “Temporal fusion transformers for interpretable multi-horizon time series forecasting,” *International Journal of Forecasting*, vol. 37, no. 4, pp. 1748–1764, 2021.
- [12] I. Karijadi, S.-Y. Chou, and A. Dewabharata, “Wind power forecasting based on hybrid ceemdan-ewt deep learning method,” *Renewable Energy*, vol. 218, p. 119357, 2023.
- [13] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.
- [14] D. Gielen, F. Boshell, D. Saygin, M. Bazilian, N. Wagner, and R. Gorini, “The role of renewable energy in the global energy transformation,” *Energy Strategy Reviews*, vol. 24, pp. 38–50, 2019.
- [15] J. Du, S. Chen, L. Pan, and Y. Liu, “A wind speed prediction method based on signal decomposition technology deep learning model,” *Energies*, vol. 18, no. 5, p. 1136, 2025.
- [16] A. M. Schetinger, H. B. Bozelli, J. M. T. do Amaral, C. C. M. de Souza, A. O. Pereira, Jr., A. G. P. Alves, E. L. van Emmerik, G. d. J. da Silva, P. H. B. Cambruzzi, and R. F. da Silva Dias, “Floating offshore wind and carbon credits in brazil: A case study on floating production, storage and offloading unit decarbonization,” *Resources*, vol. 14, no. 6, p. 85, 2025.
- [17] W. He, G. Jacobsen, T. Anderson, F. Olsen, T. D. Hanson, M. Korpås, T. Toftevaag, J. Eek, K. Uhlen, and E. Johansson, “The potential of integrating wind power with offshore oil and gas platforms,” *Wind Engineering*, vol. 34, no. 2, pp. 125–137, 2010.
- [18] S. Hanifi, H. Zare-Behtash, A. Cammarano, and S. Lotfian, “Offshore wind power forecasting based on wpd and optimised deep learning methods,” *Renewable Energy*, vol. 218, p. 119241, 2023.

- [19] J. Fang, H. Li, Y. Tang, and F. Blaabjerg, “On the inertia of future more-electronics power systems,” *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 7, no. 4, pp. 2130–2146, 2019.
- [20] W. Qiu, Y. Huang, X. Zhai, J. Ma, T. Zhang, S. Liu, and Z. Lin, “Operational risk assessment on power system based on weather regionalization considering power ramp of renewable energy generation,” *Energy Reports*, vol. 9, pp. 2063–2072, 2023.
- [21] R. Hermans, W. D’haeseleer, and E. Delarue, “Impact of ccgt start-up flexibility and cycling costs toward renewables integration,” *IEEE Transactions on Power Systems*, vol. 34, no. 5, pp. 4043–4053, 2019.
- [22] D. Flórez-Orrego and et al., “Offshore utility systems for fpsos: A technoenvironmental assessment considering the uncertainty about the natural gas price,” *Frontiers in Chemical Engineering*, vol. 4, 2022.
- [23] J. Freire and S. de Oliveira Junior, “Offshore utility systems for fpsos: technoeconomic, environmental assessment and trade-offs between gas price, carbon taxation and opportunity cost,” *Energy Conversion and Management*, vol. 245, p. 114578, 2021.
- [24] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, “The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis,” *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.
- [25] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, “A complete ensemble empirical mode decomposition with adaptive noise,” in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4144–4147, 2011.
- [26] J. Gilles, “Empirical wavelet transform,” *IEEE Transactions on Signal Processing*, vol. 61, no. 16, pp. 3999–4010, 2013.
- [27] H. Acikgoz and D. Korkmaz, “Short-term offshore wind speed forecasting approach based on multi-stage decomposition and deep residual network with self-attention,” *Engineering Applications of Artificial Intelligence*, vol. 146, p. 110313, 2025.
- [28] Y. Hong and S. Fan, “Probabilistic wind power forecasting: A review,” *Energy Reports*, vol. 6, pp. 1322–1340, 2020.

- [29] J. B. Bremnes, “Probabilistic wind power forecasts using local quantile regression,” *Wind Energy*, vol. 7, no. 1, pp. 47–54, 2004.
- [30] H. A. Nielsen, H. Madsen, and T. S. Nielsen, “Using quantile regression to extend an existing wind power forecasting system with probabilistic forecasts,” *Wind Energy*, vol. 9, no. 1-2, pp. 95–108, 2006.
- [31] K. Du, B. Yang, K. Xie, N. Dong, Z. Zhang, S. Wang, and T. Peng, “Llm-manuf: An integrated framework of fine-tuning large language models for intelligent decision-making in manufacturing,” *Advanced Engineering Informatics*, vol. 65, p. 102894, 2025.
- [32] L. Lopes dos Santos, “Lithium-ion battery aging estimation through deep learning techniques,” m.sc. dissertation, COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil, Fevereiro 2024.
- [33] Z. Qian, Y. Pei, H. Zareipour, and N. Chen, “A review and discussion of decomposition-based hybrid models for wind energy forecasting applications,” *Applied Energy*, vol. 235, pp. 939–953, 2019.
- [34] A. Betz, *Introduction to the Theory of Flow Machines*. Oxford: Pergamon Press, 1966.
- [35] V. Sohoni, S. C. Gupta, and R. K. Nema, “A comprehensive review on wind turbine power curve modeling techniques,” *Renewable and Sustainable Energy Reviews*, vol. 55, pp. 917–934, 2016.
- [36] C. Carrillo, A. F. O. Montaño, J. Cidrás, and E. Díaz-Dorado, “Review of power curve modelling for wind turbines,” *Renewable Energy*, vol. 21, no. 3, pp. 572–581, 2013.
- [37] F. Bilendo, A. Meyer, H. Badihi, N. Lu, P. Cambron, and B. Jiang, “Applications and modeling techniques of wind turbine power curve for wind farms—a review,” *Renewable and Sustainable Energy Reviews*, vol. 169, p. 112956, 2022.
- [38] X. Liu, Z. Lin, and Z. Feng, “Short-term offshore wind speed forecast by seasonal arima - a comparison against gru and lstm,” *Energy*, vol. 227, p. 120492, 2021.
- [39] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*. Wiley, 5th ed., 2015.

- [40] R. M. Montañés, M. Korpås, L. O. Nord, and S. Jaehnert, “Identifying operational requirements for flexible ccs power plant in future energy systems,” *Energy*, vol. 121, pp. 66–77, 2017.
- [41] P. Kundur, *Power System Stability and Control*. New York: McGraw-Hill, 1994.
- [42] W. M. Hamanah, M. A. Abido, and L. M. Alhems, “Realization of robust frequency stability in low-inertia islanded microgrids with optimized virtual inertia control,” *IEEE Access*, vol. 12, pp. 64623–64636, 2024.
- [43] B. Yildirim, M. Gheisarnejad, A. Mohammadzadeh, and M. H. Khooban, “Intelligent frequency stabilization of low-inertia islanded grids under time-varying parameters uncertainties,” *Journal of Energy Storage*, vol. 74, p. 109372, 2023.
- [44] H. Drucker, C. J. C. Burges, L. Kaufman, A. J. Smola, and V. Vapnik, “Support vector regression machines,” in *Advances in Neural Information Processing Systems* (M. C. Mozer, M. I. Jordan, and T. Petsche, eds.), vol. 9, MIT Press, 1997.
- [45] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [46] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, “Learning phrase representations using rnn encoder-decoder for statistical machine translation,” *arXiv preprint arXiv:1406.1078*, 2014.
- [47] H. Han, J. Peng, J. Ma, S. L. Liu, and H. Liu, “Research on load forecasting based on ceemdan se vmd and selfattention tcn fusion model,” *Scientific Reports*, vol. 15, no. 1, p. 14530, 2025.
- [48] A. U. Haque, M. H. Nehrir, and P. Mandal, “A hybrid intelligent model for deterministic and quantile regression approach for probabilistic wind power forecasting,” *IEEE Transactions on Power Systems*, vol. 29, no. 4, pp. 1663–1672, 2014.
- [49] T. Gneiting and A. E. Raftery, “Strictly proper scoring rules, prediction, and estimation,” *Journal of the American Statistical Association*, vol. 102, no. 477, pp. 359–378, 2007.
- [50] I. J. Ramirez-Rosado, L. A. Fernandez-Jimenez, C. Monteiro, J. Sousa, and R. Bessa, “Comparison of two new short-term wind-power forecasting systems,” *Renewable Energy*, vol. 34, no. 7, pp. 1848–1854, 2009.

- [51] F. De Caro, A. Vaccaro, and D. Villacci, “Adaptive wind generation forecasting with markov-switching autoregressive models,” *Energies*, vol. 13, no. 18, p. 4949, 2020.
- [52] J. Perr-Sauer, M. Lunacek, C. Meissner, L. Williams, and A. Simmons, “Short-term wind forecasting using statistical models with a fully observable wind flow,” Tech. Rep. NREL/TP-5000-74237, National Renewable Energy Laboratory (NREL), Golden, CO, 2020.
- [53] G. Aghajani, H. Shayanfar, and H. Shayeghi, “Demand side management in a smart micro-grid in the presence of renewable generation and demand response,” *Energy*, vol. 126, pp. 622–637, 2017.
- [54] S. H. R. Moosavi, B. Fani, M. E. H. Golshan, and A. Tabesh, “Multi-objective energy management of a residential smart microgrid with electric vehicle charging station,” *Applied Energy*, vol. 377, p. 124485, 2025.
- [55] P. Kumar, N. Pal, and H. Sharma, “Techno-economic analysis of solar pv/diesel generator hybrid system using homer pro software,” *Energy Reports*, vol. 8, pp. 4725–4731, 2022.
- [56] F. Grasso, G. Talluri, A. Luchetta, and S. Manetti, “Multi-objective optimization of a hybrid pv-wind-diesel generator system for off-grid applications,” *Energies*, vol. 15, no. 18, p. 6697, 2022.
- [57] R. Wang, G. Li, M. Ming, G. Wu, and L. Wang, “An efficient multi-objective model and algorithm for sizing a stand-alone hybrid renewable energy system,” *Applied Energy*, vol. 195, pp. 1077–1089, 2017.
- [58] Y. Fettah, A. Kouzou, A. Hafaifa, R. Kennel, and M. Abdelrahem, “Optimization of hybrid renewable energy microgrid for rural agricultural area in southern algeria,” *Energies*, vol. 17, no. 14, p. 3583, 2024.
- [59] G. Kostopoulos, T. N. Grapsa, and S. Kotsiantis, “Explainable artificial intelligence-based decision support systems: A recent review,” *Electronics*, vol. 13, no. 14, p. 2842, 2024.
- [60] Y. Li, H. Zhao, H. Jiang, Y. Pan, Z. Liu, Z. Wu, P. Shu, Y. Ding, and Z. Jiang, “Large language models for manufacturing,” *arXiv preprint arXiv:2410.21418*, 2024.
- [61] M. Raza, Z. Jahangir, M. B. Riaz, M. J. Saeed, and M. A. Sattar, “Industrial applications of large language models,” *Scientific Reports*, vol. 15, no. 1, p. 12345, 2025.

- [62] R. Ho and K. Hung, “Ceemd-based multivariate financial time series forecasting using a temporal fusion transformer,” in *2024 IEEE 14th Symposium on Computer Applications Industrial Electronics (ISCAIE)*, pp. 209–215, 2024.
- [63] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: A simple way to prevent neural networks from overfitting,” *Journal of Machine Learning Research*, vol. 15, pp. 1929–1958, 2014.
- [64] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, “Optuna: A next-generation hyperparameter optimization framework,” in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, p. 2623–2631, 2019.
- [65] A. Singh, A. Ehtesham, S. Kumar, and T. T. Khoei, “Agentic retrieval-augmented generation: A survey on agentic rag,” 2025.
- [66] X. Hou, Y. Zhao, S. Wang, and H. Wang, “Model context protocol (mcp): Landscape, security threats, and future research directions,” 2025.
- [67] G. Wölflein, D. Ferber, D. Truhn, O. Arandjelović, and J. N. Kather, “Llm agents making agent tools,” 2025.
- [68] Lightning AI, “The ai app development platform.” <https://lightning.ai>, 2025. Accessed: November 25, 2025.
- [69] A. Carnero, C. Martín, G. Jeon, and M. Díaz, “Online learning and continuous model upgrading with data streams through the kafka-ml framework,” *Future Generation Computer Systems*, vol. 160, pp. 251–263, 2024.
- [70] J. Castellano and I. Villanueva, “Graph neural networks in wind power forecasting,” 2025.

Appendix A

Model Hyperparameters

This appendix presents the complete set of hyperparameters for all forecasting models evaluated in this dissertation. For the benchmark models and the proposed **CEEMDAN-EWT-TFT Aggregator**, a rigorous optimization process was conducted using the Optuna framework (64). The optimization objective was to minimize the mean MAPE across the three forecast horizons on a time-series cross-validation split of the training data.

Important Note on Hybrid Models: A key distinction was made for the "divide and conquer" hybrid models (e.g., CEEMDAN-EWT-LSTM). Due to the prohibitive computational cost of optimizing an ensemble of approximately 13 sub-models, these architectures used a fixed, robust set of hyperparameters derived from common practices in the literature. This distinction underscores a practical challenge of ensemble methods and highlights the efficiency of the proposed single-model aggregator approach.

A.1 Benchmark Models (Optimized)

A.1.1 Support Vector Regression (SVR)

Table A.1: Optimized hyperparameters for SVR.

Hyperparameter	Optimized Value	Search Range
C (Regularization)	0.700	[0.1, 1000.0] (log)
epsilon (ε)	0.0011	[0.001, 1.0] (log)
gamma	auto	{'scale', 'auto'}
Validation MAPE	5.45%	—

A.1.2 Random Forest

Table A.2: Optimized hyperparameters for Random Forest.

Hyperparameter	Optimized Value	Search Range
n_estimators	315	[50, 400]
max_depth	7	[5, 50]
min_samples_split	12	[2, 14]
min_samples_leaf	7	[1, 10]
Validation MAPE	5.83%	—

A.1.3 Artificial Neural Network (ANN)

Table A.3: Optimized hyperparameters for ANN.

Hyperparameter	Optimized Value	Search Range
neurons_layer1	256	{32, 64, 128, 256}
neurons_layer2	32	{16, 32, 64, 128}
dropout_rate	0.1705	[0.1, 0.5]
learning_rate	0.000105	[1e-4, 1e-2] (log)
batch_size	64	{32, 64, 128}
Validation MAPE	5.71%	—

A.1.4 Long Short-Term Memory (LSTM)

Table A.4: Optimized hyperparameters for LSTM.

Hyperparameter	Optimized Value	Search Range
lstm_units	128	{32, 64, 128, 256}
dropout_rate	0.2276	[0.1, 0.5]
learning_rate	0.000924	[1e-4, 1e-2] (log)
batch_size	128	{32, 64, 128}
Validation MAPE	5.67%	—

A.1.5 Bidirectional LSTM (Bi-LSTM)

Table A.5: Optimized hyperparameters for Bi-LSTM.

Hyperparameter	Optimized Value	Search Range
lstm_units	256	{32, 64, 128, 256}
dropout_rate	0.3780	[0.1, 0.5]
learning_rate	0.000946	[1e-4, 1e-2] (log)
batch_size	128	{32, 64, 128}
Validation MAPE	5.67%	—

A.1.6 CNN-Bi-LSTM

Table A.6: Optimized hyperparameters for CNN-Bi-LSTM.

Hyperparameter	Optimized Value	Search Range
cnn_filters	32	{32, 64, 128}
kernel_size	3	{2, 3, 5}
lstm_units	128	{32, 64, 128}
dropout_rate	0.3461	[0.1, 0.5]
learning_rate	0.000663	[1e-4, 1e-2] (log)
batch_size	128	{32, 64, 128}
Validation MAPE	5.71%	—

A.1.7 CNN-Bi-LSTM with Attention

Table A.7: Optimized hyperparameters for CNN-Bi-LSTM+Attention.

Hyperparameter	Optimized Value	Search Range
cnn_filters	128	{32, 64, 128}
kernel_size	5	{2, 3, 5}
bilstm_units	64	{32, 64, 128}
agg_lstm_units	32	{16, 32, 64}
dropout_rate	0.2330	[0.1, 0.5]
learning_rate	0.000102	[1e-4, 1e-2] (log)
batch_size	128	{32, 64, 128}
Validation MAPE	5.83%	—

A.1.8 Transformer

Table A.8: Optimized hyperparameters for Transformer.

Hyperparameter	Optimized Value	Search Range
head_size	32	{32, 64, 128}
num_heads	4	{2, 4, 8}
ff_dim	16	{4, 8, 16}
num_transformer_blocks	1	[1, 4]
mlp_units	64	{64, 128}
dropout	0.3744	[0.1, 0.4]
mlp_dropout	0.4393	[0.1, 0.5]
learning_rate	0.000635	[1e-4, 1e-2] (log)
batch_size	32	{32, 64}
Validation MAPE	5.92%	—

A.1.9 Temporal Fusion Transformer (TFT)

Table A.9: Optimized hyperparameters for the benchmark TFT model.

Hyperparameter	Optimized Value	Search Range
hidden_units	128	{32, 64, 128}
dropout_rate	0.2896	[0.1, 0.4]
num_heads	4	{2, 4, 8}
learning_rate	0.000780	[1e-4, 1e-2] (log)
batch_size	32	{32, 64}
Validation MAPE	6.20%	—

A.2 Hybrid Decomposition Models

A.2.1 CEEMDAN-EWT-TFT Aggregator (Proposed Model)

This is the only hybrid model subjected to hyperparameter optimization, as its single-model architecture makes the process computationally feasible.

Table A.10: Optimized hyperparameters for the proposed CEEMDAN-EWT-TFT Aggregator.

Hyperparameter	Optimized Value	Search Range
hidden_units	32	{32, 64, 128}
dropout_rate	0.3059	[0.1, 0.4]
num_heads	2	{2, 4, 8}
learning_rate	0.002401	[1e-4, 1e-2] (log)
batch_size	64	{32, 64}
Validation MAPE	6.63%	—

A.2.2 Other Hybrid Models (Fixed Hyperparameters)

The following "divide and conquer" hybrid models were not optimized. They used a standardized set of robust hyperparameters for each of their sub-models, as detailed in Table A.11.

Table A.11: Fixed hyperparameters for non-optimized hybrid models.

Hyperparameter	Fixed Value
<i>For CEEMDAN-EWT-LSTM, -GRU, -BiLSTM</i>	
Recurrent Units (LSTM/GRU)	32
Dropout Rate	0.2
Learning Rate	0.001
Batch Size	64
Optimizer	Adam
<i>For CEEMDAN-EWT-BiLSTM-Attn</i>	
BiLSTM Units (1st layer)	32
LSTM Units (2nd layer)	32
Dropout Rate	0.2
Learning Rate	0.001
Batch Size	64
<i>For CEEMDAN-EWT-CNN-BiLSTM-Attn</i>	
CNN Filters	64
Kernel Size	3
BiLSTM Units (1st layer)	32
LSTM Units (2nd layer)	32
Dropout Rate	0.2
Learning Rate	0.001
Batch Size	64

Despite using fixed parameters, these models achieved state-of-the-art performance, underscoring that the signal decomposition was the dominant factor in their success.

A.3 Optimization Methodology

All hyperparameter optimizations were performed using the Optuna framework (64). The process was configured to minimize the mean MAPE across all three forecast horizons, evaluated using a 3-fold time-series cross-validation strategy to ensure robustness and prevent data leakage. Each optimization consisted of 30 trials, with early stopping (patience=20) to prune unpromising runs. Final models were then retrained on the complete training dataset using the identified optimal hyperparameters.

A.4 Key Insights from Hyperparameter Optimization

The optimization process yielded several key insights into model behavior for this specific forecasting task:

1. **Learning Rate and Regularization are Critical:** Models were highly sensitive to the learning rate, with optimal values consistently found in the [1e-4, 1e-3] range. Furthermore, moderate dropout rates (0.2-0.4) were essential for preventing overfitting, especially in high-capacity models like Bi-LSTM.
2. **Simple is Better Than Complex:** For non-decomposed signals, more complex architectures (e.g., adding attention layers or more Transformer blocks) did not guarantee better performance, often being outperformed by simpler, well-tuned LSTM-based models.
3. **Decomposition Dominates Tuning:** The most significant finding is the performance gap between model families. Even the *non-optimized* hybrid models ($\text{MAPE} \approx 1.7\text{-}3.2\%$) massively outperformed the fully *optimized* benchmark models ($\text{MAPE} \approx 4.3\text{-}7.5\%$). This provides definitive evidence that, for this problem, **advanced feature engineering via signal decomposition is far more impactful than hyperparameter tuning alone.**