MEASURING PORTFOLIO RISK USING MACHINE LEARNING TECHNIQUES: VAR AND ES USING GAUSSIAN PROCESS REGRESSION

by

Ujjwala Vadrevu, B.E. Electronics and Instrumentation, M.Sc. Physics

DISSERTATION

Presented to the Swiss School of Business and Management Geneva

In Partial Fulfillment

Of the Requirements

For the Degree

DOCTOR OF BUSINESS ADMINISTRATION

SWISS SCHOOL OF BUSINESS AND MANAGEMENT GENEVA

January, 2026

MEASURING PORTFOLIO RISK USING MACHINE LEARNING TECHNIQUES:

VAR AND ES USING GAUSSIAN PROCESS REGRESSION

by

Ujjwala Vadrevu

APPROVED BY

____

Dissertation chair - Dr. Gualdino Cardoso

RECEIVED/APPROVED BY:

Renee Goldstein Osmic

____

Admissions Director

Dedication

To my family, who supported me through thick and thin and provided the foundation for this work.

To my husband, Suraj, for his ungrudging dedication to our family, especially in taking care of our daughters during my countless marathon sessions and demanding weekends. Your commitment was the greatest gift of time I could have received.

To my elder daughter, Chinmayee, for her maturity in understanding and internalizing that studying is a lifelong process, and for forgiving my necessary lack of presence. Your resilience continues to inspire me.

And to my younger one, Manomayi, who, despite only being two and three-quarters, willingly replaced "Mama-time" with "Sister-time" while trying her best to understand why Mama was always working. This achievement is as much yours as it is mine.

Acknowledgements

My deepest gratitude goes to the SSBA for providing the unique and invaluable opportunity for working professionals to pursue advanced research, making this entire academic journey possible.

I would also like to express profound appreciation for my mentor, Professor Apostolos Dasilas. Thank you for your unwavering belief in this research and for putting up with my often irregular work schedule, which was dictated by external professional demands. Your constant constructive support, guidance, and patience since commencing supervision of this thesis have been instrumental in its successful completion.

ABSTRACT

MEASURING PORTFOLIO RISK USING MACHINE LEARNING TECHNIQUES:

VAR AND ES USING GAUSSIAN PROCESS REGRESSION

Ujjwala Vadrevu
2025

Dissertation Chair: <Chair’s Name>
Co-Chair: <If applicable. Co-Chair’s Name>

This study explores and addresses the common and practical challenge encountered in regulatory risk modelling - modeling and forecasting time-varying Volatility-Covariance Matrices VCV for regulatory-compliant stress testing processes like Internal Capital Adequacy Assessment Process (ICAAP) and CCAR (Comprehensive Capital Analysis and Review). In practice, this challenge is most acute under forward-looking stress conditions, where model stability and regulatory defensibility take precedence over theoretical elegance. Traditional econometric models often fail to capture the complex, non-linear dependencies and high dimensionality of real-world financial markets, leading to unstable and/or wrong stress test results. Attempts made to address this issue using Multi-Output Gaussian Process Regression (MOGPR) resulted in unstable VCV matrices and convergence issues with the optimizer. Keeping this in mind, this research pivots and proposes the implementation of a novel Hybrid GPR-VCV Framework that leverages the power of Machine Learning (ML) through Univariate Gaussian Process Regression (UGPR) to model individual asset volatilities, while maintaining regulatory tractability by utilizing historical correlation matrices for the covariance structure, for both Value-at-Risk (VaR) and Stressed VaR (SVaR). The framework is dynamically calibrated and tested using a sequential forward-chaining methodology. This approach utilizes an expanding training window-where all historical data from the fixed start date of June 30, 2020, is accumulated-to forecast a fixed-size test set (approximately 252 daily returns, or one trading year) through the total sample period ending June 30, 2025, on seven key global equity indices (Developed and Emerging Markets). This iterative process facilitates the forward-looking stress testing, which subjects the model to three hypothetical, ICAAP-style stress scenarios: West Asia War, Climate Risk, and AI Bubble / Regulatory Burden. The study investigates four primary research problems, including the stability and efficiency of the model, the impact of dynamic paths, and the interpretation of risk drivers using SHAP (SHapley Additive exPlanations). The key finding is that the Hybrid GPR-VCV approach provides statistically stable, out-of-sample volatility forecasts even with a short one-year fit window. This level of stability was not achievable using the standard Multi-Output GPR (MOGPR) framework. The model demonstrates superior resilience and accuracy in projecting market losses under dynamic stress, and the SHAP analysis provides crucial interpretability, revealing the specific non-linear input features driving volatility forecasts. This hybrid methodology represents a significant advancement for risk management, offering a robust, interpretable, and computationally efficient tool for forward-looking stress testing in financial institutions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I:

INTRODUCTION

1.1 Introduction

Financial stability and solvency in the modern banking landscape are fundamentally underpinned by accurate risk measurement, primarily through metrics like Value-at-Risk (VaR) and its more conservative counterpart, Expected Shortfall (ES) (Artzner et al. 1999; Jorion 2006; Hull 2018). These regulatory-mandated measures, central to processes such as the Internal Capital Adequacy Assessment Process (ICAAP) and the Comprehensive Capital Analysis and Review (CCAR), are designed to quantify potential portfolio losses over specific time horizons and confidence levels (typically 99%), thereby determining minimum capital requirements (Basel Committee on Banking Supervision 2013; Basel Committee on Banking Supervision [BCBS] 2019). In applied risk management settings, the practical difficulty is not defining regulatory risk measures such as VaR or ES, but ensuring that the underlying models remain numerically stable, conservative, and defensible under forward-looking stress conditions.

The necessity for robust internal models is directly derived from these regulatory frameworks:

The Internal Capital Adequacy Assessment Process (ICAAP) is the main item under Pillar 2 of the Basel Accords by which global banks assess their overall capital needs relative to their risk profile (Basel Committee on Banking Supervision [BCBS] 2019). ICAAP requires banks to go beyond the standardized quantitative capital requirements (Pillar 1) by utilizing internal models to capture all material risks, including those related to strategy, reputation, and most critically, stress scenarios. This process needs institutions to maintain models capable of providing evidence that their minimum capital requirements are sufficient to cover events that are extreme but plausible, thereby linking advanced risk modeling to business strategy and long-term solvency.

The Comprehensive Capital Analysis and Review (CCAR) represents the U.S. Federal Reserve's mandatory annual exercise for large bank holding companies (Federal Reserve 2020). This is similar in nature to the Basel ICAAP exercise. However, CCAR places an intense emphasis on forward-looking, top-down stress testing. Firms are required to project their capital levels, losses, and revenues under severely stressed economic and financial market scenarios which are hypothetical in nature and are often based on the scenarios provided by the Fed, over a multi-year planning horizon. The calculation of the Stressed VaR (SVaR), which must pass regulatory scrutiny for accuracy and robustness, is a key input to these projections. Consequently, the reliability and transparency of the underlying models used to forecast stressed risk metrics become a direct condition for receiving regulatory approval to distribute capital.

The accurate calculation of VaR and SVaR hinges entirely on the reliable modeling and forecasting of the portfolio's forward-looking Volatility-Covariance Matrix (VCV), which captures both the individual risk of each asset and their interdependencies.

1.1.1 The Inadequacy of Traditional Models

The global financial system is characterized by high interconnectivity, extreme non-linear dynamics, and intermittent periods of severe market instability. Despite their widespread adoption, traditional econometric volatility models exhibit structural limitations when applied to modern financial markets (Seyfi et al. 2021). Periods of market stress expose fundamental weaknesses in linear volatility modelling approaches such as the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) family and Exponentially Weighted Moving Average (EWMA) which prioritised tractability over adaptability, a trade-off that becomes problematic during regime shifts. These models make assumptions that are consistently violated by financial data:

1. **Normality:** Traditional models often assume the standardized returns are normally distributed, leading to a systematic underestimation of tail risk (the "fat-tail problem") (Wilkens 2019; Seyfi et al. 2021).
2. **Symmetry:** They struggle to capture the leverage effect, the empirical observation that volatility tends to increase more significantly following large negative returns than following positive returns of the same magnitude (Hull 2018).
3. **Linearity:** They are intrinsically limited in their capacity to model the complex, non-linear dependencies that characterize market behavior, particularly during periods of stress where asset correlations tend to converge dramatically (the "flight-to-quality" or "correlation contagion" phenomenon) (Seyfi et al. 2021; Wang and Wu 2025).

This systematic underestimation of risk during the critical periods of financial stress compromises the integrity of regulatory capital models. The pursuit of more accurate, non-linear, and statistically robust modeling has inevitably led the financial engineering community towards advanced Machine Learning (ML) techniques (Herfurth 2020).

1.2 The Research Problem: Multivariate GPR for Portfolio Risk

The pursuit of advanced forecasting models naturally leads to exploring the potential of Non-Parametric Machine Learning techniques. Such methods offer superior capabilities for modeling non-linear, time-varying conditional volatility and predictive uncertainty without imposing rigid functional forms on the data.

Extending volatility modelling from individual assets to large portfolios introduces both statistical and operational challenges, driven by dynamic retraining needs under strict regulatory protocols. The most robust theoretical approach for forecasting the entire Volatility-Covariance Matrix (VCV) requires a multivariate framework that jointly models all asset dependencies. While advanced ML techniques hold significant potential to overcome the limitations of traditional models, their application to real-world, large-scale financial portfolios reveals substantial operational and architectural challenges.

This research aims to bridge the critical gap between the theoretical promise of advanced multivariate modeling and its practical viability in a demanding financial environment. Successfully applying these predictive capabilities necessitates the development of a novel and operationally stable framework that can deliver regulatory-compliant, conditional risk forecasts for large portfolios, thereby advancing the state-of-the-art in risk methodology.

This research focuses on Gaussian Process Regression (GPR), a Bayesian, non-parametric ML technique that excels at probabilistic time-series modeling and inherently provides both a mean forecast and an associated measure of predictive uncertainty (variance) (Rasmussen and Williams 2005; Ludkovski and Risk 2025). GPR's flexibility, defined by its kernel function, allows it to map highly non-linear relationships without assuming a fixed functional form, offering a theoretical solution to the limitations of GARCH-type models.

1.3 Purpose of Research

The core purpose of this research is to develop, validate, and interpret a novel, GPR-based framework for forecasting conditional market risk. By integrating advanced non-parametric modeling with established regulatory risk metrics, the study aims to enhance the accuracy, stability, and regulatory compliance of capital assessment practices for large financial institutions.

This research is fundamentally motivated by the persistent gap between the theoretical complexity of market dynamics and the limitations of current operational risk models. Traditional econometric models, anchored in assumptions of linearity and fixed distributions, consistently demonstrate failure during periods of systemic stress by underestimating volatility and misjudging the magnitude of tail events. The primary aim is to replace this deficient framework with a statistically superior, conditional forecasting architecture that inherently captures the non-linear, time-varying nature of portfolio risk, thereby delivering more prudent and reliable measures of regulatory capital.

Furthermore, the purpose extends beyond mere predictive accuracy to address the critical requirements of financial governance. The development and subsequent rigorous validation of this framework are intended to directly support the high-stakes demands of the ICAAP and CCAR processes (Basel Committee on Banking Supervision, 2019; Federal Reserve, 2020). This includes ensuring the model's robustness under extreme, hypothetical stress scenarios and, critically, achieving a high degree of model transparency. By integrating necessary interpretability features, the research seeks to overcome the regulatory challenge of the "black box" and provide risk managers and supervisors with clear, actionable insights into the drivers of the forecasted capital requirement.

1.4 Significance of the Study

This study holds profound implications for the operational stability of financial institutions and the theoretical advancement of quantitative risk modeling, addressing critical needs in both industry practice and academic literature.

1.4.1 Practical Implications: Capital Management and Regulatory Oversight

The primary significance of this research lies in its potential to directly improve the quality of risk capital assessment under regulatory scrutiny (Basel Committee on Banking Supervision [BCBS] 2019; Federal Reserve 2020). By deploying a robust, conditional forecasting architecture, the framework significantly mitigates the systemic underestimation of tail risk that plagues traditional models. The enhanced accuracy translates directly into more prudent and reliable measures of VaR and Expected Shortfall (ES) (Lopez 1998), thereby strengthening internal capital adequacy against extreme market events as required by ICAAP and CCAR. Ultimately, the successful validation of this approach offers financial institutions a pathway to both superior risk identification and more efficient capital allocation, reducing the opportunity cost associated with overly conservative, non-conditional forecasting. Furthermore, the inherent transparency features of the methodology facilitate clearer regulatory dialogue, ensuring that model outputs are auditable and understood by supervisory bodies, a critical factor for internal model approval and governance (Lundberg and Lee 2017).

1.4.2 Theoretical Implications: Advancing Machine Learning in Finance

The study provides a significant contribution to the quantitative risk modeling literature by advancing the empirical application of complex Machine Learning (ML) methodologies in highly constrained financial environments. By successfully translating the theoretical potential of advanced non-parametric modeling into an operationally stable framework, the research furnishes a proven blueprint for others seeking to integrate high-fidelity ML techniques into mission-critical financial functions. This work provides necessary validation for moving beyond linear econometrics and firmly establishes the utility of probabilistic modeling for capturing the non-linear, time-varying dynamics of global financial markets (Ludkovski and Risk 2025). The contribution extends to the emerging field of eXplainable AI (XAI) by rigorously demonstrating a methodology for achieving transparency in complex risk attribution, a key theoretical prerequisite for the broader acceptance of ML in regulatory finance (Lundberg and Lee 2017; Gramegna and Giudici 2021).

1.5 Research Purpose and Questions

The methodological framework detailed in Chapter 3 is designed to systematically execute and measure the achievement of four core research objectives, thereby fulfilling the purpose of this study and addressing the identified gaps in current market risk practice.

The first foundational objective is to Assess GPR's Applicability Across Diverse Markets. This involves applying the GPR framework to a comprehensive, multi-asset portfolio comprising indices from both developed and emerging economies. The goal is to rigorously evaluate the framework's robustness, performance, and adaptability, specifically through a comparative analysis of its predictive accuracy against traditional risk metrics within varied economic environments.

The second objective focuses on the framework's core technical capability: Modeling Inter-Asset Dependencies. This aims to advance the understanding of how GPR can capture dynamic, non-linear correlations and risk aggregation effects under various market conditions. Successful achievement is crucial for improving effective portfolio diversification strategies and ensuring accurate regulatory capital assessment (VaR and ES).

The third objective targets high-stakes regulatory requirements by Evaluating GPR's Stress Scenario Performance. The research will test the GPR models under a range of extreme systemic risk scenarios, encompassing both forward-looking hypothetical shocks (such as "West Asia War," "AI bubble/ Regulatory Burden," and "Climate Risk") and foundational historical crises (including the 2008 Financial Crisis and the COVID-19 pandemic). The goal is to rigorously assess the model's ability to maintain stability and yield robust, conservative Stressed VaR forecasts under these challenging market conditions.

The fourth and final objective is critical for model adoption: to Enhance Model Explainability and Regulatory Adaptability. This involves developing methods to ensure that GPR model outputs are transparent, interpretable, and actionable for risk managers and regulators. The objective includes adapting the GPR framework to align with evolving regulatory validation requirements, ensuring the models remain effective and compliant in a changing governance landscape.

CHAPTER II:

REVIEW OF LITERATURE

2.1 Introduction

This chapter establishes the theoretical and empirical foundation for the proposed Hybrid Gaussian Process Regression–Historical Simulation (GPR-HS) Framework. As a comprehensive literature review, the chapter is structured to first define the stringent regulatory problem, critique the existing methodological failures, and ultimately identify the critical research gaps that necessitate a novel solution.

The review begins by anchoring the study in the stringent regulatory requirements imposed by the Basel Accords, the Fundamental Review of the Trading Book (FRTB), and forward-looking exercises such as the Internal Capital Adequacy Assessment Process (ICAAP) and the Comprehensive Capital Analysis and Review (CCAR). Specifically, it defines the necessity for accurate Expected Shortfall (ES) and Stressed VaR (SVaR) forecasting. The chapter then systematically critiques traditional volatility models, such as EWMA and GARCH-family processes, demonstrating their inherent parametric rigidities and failure to capture the conditional, non-linear dependencies critical for modern risk measurement. A detailed review of the state-of-the-art literature on non-parametric approaches, specifically Gaussian Process Regression (GPR), follows, which justifies its selection as the core modelling technique. The chapter culminates by defining the explicit research gaps in multivariate modelling and numerical stability that the proposed Hybrid GPR-HS framework is designed to address. Finally, it establishes the critical role of Explainable AI (XAI) in ensuring regulatory acceptability.

2.2 Theoretical Framework

The proposed methodology is grounded in the evolution of regulatory standards, which mandate a dynamic, comprehensive, and forward-looking approach to market risk capital calculation. The framework is specifically designed to meet the rigorous demands set by global financial regulators, particularly concerning the calculation of Stressed Value-at-Risk (SVaR) and Expected Shortfall (ES).

2.2.1 Evolution of Market Risk Metrics: From VaR to ES

Regulatory experience has demonstrated that earlier market risk measures were insufficient for capturing extreme tail losses. Its core mathematical flaw is its inability to capture the magnitude of losses beyond the chosen confidence level which leads to a violation of the property of subadditivity (Artzner et al. 1999; Embrechts et al. 2002).

The transition from Value-at-Risk to Expected Shortfall reflects a broader regulatory shift toward tail-sensitive risk measures, under the Basel III framework and the Fundamental Review of the Trading Book (FRTB). ES, or Conditional VaR (CVaR), corrects this flaw by calculating the expected loss conditional on the loss exceeding the VaR threshold. ES is defined as the average of losses that lie in the tail of the loss distribution beyond the VaR confidence level. This property of measuring tail loss makes ES a coherent risk measure, as it satisfies subadditivity and is therefore more robust under stress.

2.2.2 The Stress Testing Mandate: SVaR and CCAR

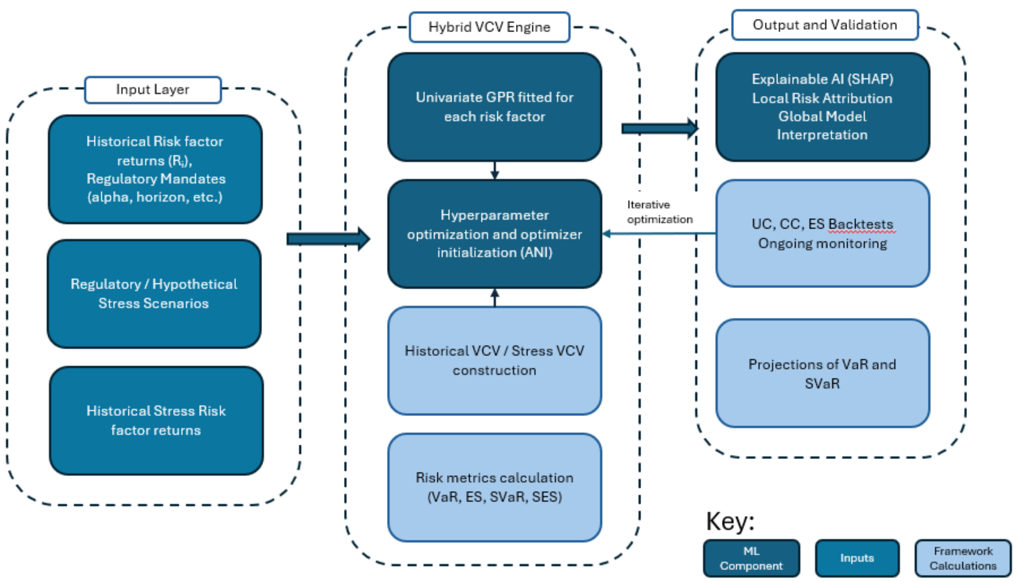
Regulatory bodies, including the Federal Reserve (CCAR) and the Basel Committee (ICAAP), mandate stress testing to ensure the banking system's resilience. The capital requirement is determined by the greater of the VaR and the Stressed VaR (SVaR), where SVaR is calculated using a 1-year window of data from a period of significant financial stress, typically including the 2008 financial crisis.

SVaR Integration: The SVaR requirement necessitates a volatility modelling technique capable of integration with historically observed crisis data. The Hybrid GPR-HS framework addresses this by dynamically modelling univariate volatility via GPR while utilizing a historically derived, stressed covariance structure for the inter-asset correlation, ensuring compliance with the dual VaR/ SVaR capital floor requirement. This approach provides a computationally efficient path to the multi-quarter projections required in both ICAAP and CCAR exercises.

2.3 Conceptual Framework: The Hybrid GPR-HS Architecture

The proposed Hybrid GPR-HS architecture is structured to directly address the joint regulatory and computational challenges identified in the literature. The framework decouples the problem of portfolio risk measurement (VCV matrix construction) into two distinct, manageable tasks: dynamic univariate volatility forecasting and stable inter-asset correlation estimation. Below is a visual representation of the framework.

Figure 1: Conceptual Architecture of the Hybrid GPR–HS Framework



2.3.1 Framework Components and Decoupling Strategy

The Hybrid Framework is defined by three logical components, which interact sequentially to produce the final risk metrics (VaR, ES, and SVaR):

Univariate GPR Engine (Dynamic Volatility): The core engine uses N independent Gaussian Process Regression models to forecast the time-varying conditional variance ($\boldsymbol{\sigma}_{i,t}^{2}$) for each of the N assets. This component addresses the limitations of parametric models by providing a flexible, non-linear forecast of asset-specific risk.

Historical Simulation (HS) Integration (Static Correlation): The framework leverages a stable historical covariance matrix for the off-diagonal elements of the VCV matrix. This approach is computationally efficient and provides the necessary stability, resolving the computational intractability of full Multivariate GPR (MGPR) for large portfolios, a problem discussed in Section 2.5.

Explainable AI (XAI) Layer (SHAP): A critical regulatory component, the XAI layer, uses SHAP analysis to quantify the exact contribution of each model input (mean and variance) to the final capital number (SVaR). This provides the necessary interpretability and transparency for regulatory submission, addressing a core "black-box" concern in machine learning risk models.

The relationship between these components, the required inputs, and the final regulatory outputs is illustrated below.

2.3.2 Design Rationale for Implementation

The adoption of this hybrid design is a strategic choice, rooted in the trade-off between mathematical elegance and computational feasibility. While a full MGPR model would theoretically offer joint non-linear dependency modelling, its prohibitive computational complexity, which scales as $\boldsymbol{O}((\boldsymbol{N} \cdot \boldsymbol{T})^{\mathbf{3}})$, renders it unsuitable for the expansive, rolling-window validation required by CCAR and ICAAP exercises. By decoupling volatility and correlation, the framework maintains the dynamic predictive superiority of GPR where it is most critical (asset-specific risk) while ensuring scalability and numerical stability.

2.4 Review of Prior Studies

This section provides a systematic and critical review of the key literature, tracing the evolution of quantitative market risk modelling from foundational parametric methods to state-of-the-art non-parametric and Explainable AI (XAI) approaches. The review is structured to analyse existing solutions, identify their structural rigidities and computational instabilities, and establish the technical necessity for the proposed Hybrid GPR-HS framework.

2.4.1 Early Risk Measurement: Parametric and Historical Methods

Early approaches to market risk measurement emphasised simplicity and ease of implementation.

Historical Simulation (HS) and EWMA represented important initial steps toward time-varying risk estimation. Foundational literature established simple methods for estimating VaR and ES (Linsmeier and Pearson 1996). Historical simulation remains an industry staple, while the EWMA model, popularized by J.P. Morgan's RiskMetrics introduced time-varying volatility (Longerstaey, Jacques and Zangari, Peter 1996). Its critical flaw-the use of a single, fixed decay factor (**λ**)-was widely critiqued by authors like Figlewski and Engle and Mezrich for its lack of adaptability to varying market volatility persistence (Engle, Robert F. and Mezrich, Joseph 1996; Figlewski, Stephen 1997).

VaR Limitations: The ultimate failure of these early methods during the 2007–2008 Global Financial Crisis confirmed theoretical critiques by Artzner, Delbaen, Eber, and Heath, who demonstrated that VaR is not a coherent risk measure due to its failure to satisfy the property of subadditivity. This theoretical finding underpinned the regulatory shift toward ES (Artzner et al. 1999).

2.4.2 The GARCH Revolution and Its Limitations

The introduction of GARCH-type models marked a significant advancement in modelling volatility clustering (Engle 1982; Bollerslev 1986)**.**

While GARCH-based frameworks improved volatility estimation, their multivariate extensions introduced new computational and stability concerns. Standard GARCH models assume a symmetric response to market news. However, empirical studies consistently demonstrated the leverage effect, where negative returns increase subsequent volatility more significantly than positive returns (Black, Fischer 1976; Nelson 1991). Extensions like EGARCH or GJR-GARCH addressed this asymmetry (Glosten et al. 1993). However, they introduced parametric rigidity that was prone to misspecification (Hansen and Lunde 2005).

The application of GARCH to portfolios led to Multivariate GARCH (MGARCH) models. The Dynamic Conditional Correlation (DCC) approach gained prominence for its relative scalability (Engle 2000). However, DCC and other MGARCH forms were shown to be numerically challenging due to the need to maintain Positive Definiteness (PD) of the VCV matrix which is a frequent issue highlighted in computational finance literature and also faced in this study with the MOGPR (Bauwens et al. 2006; Tsay 2005). Furthermore, large-scale MGARCH implementations remain computationally demanding (Rasmussen and Williams 2005).

2.4.3 The Rise of Non-Parametric Modelling: Gaussian Process Regression

Recent literature has increasingly explored non-parametric approaches to volatility forecasting motivated by their flexibility, with GPR emerging as a powerful alternative (Rasmussen and Williams 2005).

Early applications of GPR in financial time series were explored by Gibbs (1997) and later by Wu et all in 2014. (Gibbs and MacKay 1996; Wu et al. 2014). The research highlighted GPR's superior ability to capture non-linear, time-varying conditional volatility compared to GARCH particularly for non-parametric volatility forecasting (Orion 2010).

Kernel Specification Debate (RBF vs. Matern): A critical divergence in the GPR literature concerns the kernel choice. Many initial GPR studies utilized the Radial Basis Function (RBF) kernel. However, this choice was criticized by Stein (1999) and later by Snoek, Larochelle, and Adams (2012) because the RBF's assumption of infinite differentiability is unrealistic for financial returns (Stein 1999; Snoek et al. 2012). The literature supports the Matern 5/2 Kernel as a more robust choice, whose assumption of only twice-differentiable paths better reflects empirical roughness (Wilson and Ghahramani 2010; Wu et al. 2014).

Multivariate Complexity and Scaling: While the theoretical superiority of full Multi-Output GPR (MGPR) for capturing joint non-linear dependencies was recognized by Álvarez and Lawrence (2011), its prohibitive computational complexity-scaling as $\boldsymbol{O}((\boldsymbol{N} \cdot \boldsymbol{T})^{\mathbf{3}})$ with respect to the number of data points N for training/inversion-makes it unsuitable for the backtesting windows required in regulatory exercises (Quiñonero-Candela and Rasmussen 2005; Rasmussen and Williams 2005). This computational barrier necessitates a scalable, hierarchical or decoupled approach.

2.4.4 GPR Applicability Across Financial Domains

The non-parametric flexibility of Gaussian Process Regression has driven its adoption across a wide spectrum of quantitative finance, risk management, and prediction tasks.

Volatility and Time Series Forecasting: GPR has been extensively utilized as a robust alternative to traditional GARCH-family models for financial time series analysis and volatility prediction. Reviews of the methodology confirm its utility for volatility forecasting (Han and Zhang 2015; Liu et al. 2020).

Risk Measurement and Prediction: GPR has been successfully applied to portfolio-level risk management, including studies on portfolio VaR and ES estimation (Seyfi et al. 2021). GPR is applicable to general machine learning in risk measurement as well (Wilkens 2019). Furthermore, its use extends into market risk assessment for trading books (Noureddine Lehdili et al. 2019; Lehdili et al. 2025).

Pricing and Valuation: The ability of GPR to model complex, high-dimensional functions make it ideal for derivative pricing (Crépey and Dixon 2019), variable annuities (Goudenège et al. 2021), and European option pricing (Hainaut and Vrins 2024).

Portfolio and Strategy Optimization: GPR has been integrated into strategic investment frameworks, notably in its use for portfolio optimization, such as the Black-Litterman model (Li et al. 2023). Its core function for modelling non-linear predictive regressions has also been applied to forecasting stock returns (Bisht et al. 2022; Kamonrat Suphawan et al. 2022; Wang et al. 2024).

Macroeconomic and Systems Modelling: In a broader context, GPR has been used to model complex financial and macroeconomic systems (Hauzenberger et al. 2025).

2.4.5 The Emergence of Model Explainability (XAI)

The growing complexity of machine learning models in finance has necessitated the development of rigorous methodologies for model interpretation, driven by regulatory demands for transparency Dodd-Frank Act (Dodd-Frank Wall Street Reform and Consumer Protection Act 2010).

The Black Box Problem: Literature documents the fundamental "black box" concern, where risk managers and regulators require confidence in model attribution (Hull 2018; Weller 2019).

SHAP and Axiomatic Justification: The most rigorous advancement in this field is SHapley Additive exPlanations (SHAP) (Lundberg and Lee 2017). SHAP is widely recognized in the XAI literature, building on the foundational work of Shapley (1953), as the only methodology that satisfies the crucial axiomatic properties (Local Accuracy, Missingness, Consistency, and Efficiency) (Shapley 1953). This mathematical foundation ensures the SHAP values provide a unique, fair, and mathematically sound attribution of output to input, making SHAP the preferred, compliant method for explaining high-stakes regulatory predictions (Guidotti et al. 2018; Barredo Arrieta et al. 2020).

2.5 Research Gaps

Based on the review of prior studies, the present study deliberately prioritizes operational stability, regulatory tractability, and stress-time robustness over theoretical completeness or full multivariate elegance. Hence, the proposed Hybrid GPR-HS framework is designed to close the following four critical and interconnected gaps in the current body of quantitative risk literature:

The Applicability Gap Across Diverse Markets: A lack of systematic assessment of GPR's robustness across a truly diverse set of international equity markets with heterogeneous volatility dynamics.

The Multivariate Stability and Scalability Gap: The failure of both DCC-GARCH (PD violation) and full MGPR ($\boldsymbol{O}((\boldsymbol{N} \cdot \boldsymbol{T})^{\mathbf{3}})$ cost) to provide a scalable, stable, and computationally feasible solution for portfolio-level VCV forecasting.

The Regulatory Adaptability Gap under Stress: A need for a methodologically sound framework that can be systematically adapted and stabilized (via the ANI strategy) to generate robust and conservative SVaR forecasts for regulatory stress testing, a requirement not robustly met by traditional models.

The Explainability Gap for Complex Models: A pressing need for a validated, axiomatically sound methodology (SHAP) to decompose and attribute the risk contribution of non-parametric models (GPR) to the final SVaR number, thereby addressing the "black box" concern for supervisory bodies.

CHAPTER III:

METHODOLOGY

3.1 Overview of the Risk Measurement Framework

This section establishes the core building block of the entire framework: the non-parametric, probabilistic forecasting of individual asset conditional volatility (Rasmussen and Williams 2005). By partitioning the total Volatility-Covariance Matrix (VCV) problem, the framework capitalizes on the specific theoretical strength of GPR to model time-varying, non-linear volatility structures in isolation (Orion 2010; Han and Zhang 2015). This approach is adopted to circumvent the known scaling limitations of full Multivariate GPR (MGPR), which exhibits prohibitive computational complexity for large portfolios (Quiñonero-Candela and Rasmussen 2005).

3.1.1 GPR Core Architecture and Composite Kernel Selection

The Hybrid GPR-HS Framework relies on *N* independent, univariate GPR models to forecast the conditional volatility for each asset (Han and Zhang 2015). The theoretical rigor of these models is entirely dependent on the definition and optimization of the composite kernel function, $k(x, x')$ (Rasmussen and Williams 2005).

A summation kernel combining two essential components is employed to accurately model the dual characteristics of financial volatility-structural persistence and irreducible noise (Orion 2010) :

$$k(x, x') = k_{\text{Matern5/2}}(x, x') + k_{\text{WhiteNoise}}(x, x')$$

A. Matern 5/2 Component

The Matern 5/2 Kernel models the underlying structural, time-varying heteroskedasticity of the asset returns (Orion 2010). Its controlled differentiability ($\nu = 5/2$) ensures continuity up to the second derivative, allowing the model to capture the sharp, persistent, and non-smooth clustering characteristic of market volatility (Porcu et al. 2023). This is theoretically better suited than the infinitely smooth Radial Basis Function (RBF) kernel whose assumptions of smoothness are typically unrealistic for empirical financial returns (Snoek et al. 2012; Stein 1999). This kernel is parameterized by its lengthscale (ℓ) and its signal variance (σ_f^2), which are optimized during training (Rasmussen and Williams 2005).

B. White Noise Component

The White Noise Kernel models the measurement error or irreducible idiosyncratic noise (σ_n^2) in the observations. This term is mathematically and numerically critical as it guarantees that the total Gram matrix K has a positive value on the diagonal, providing a necessary stability margin for subsequent linear algebra operations (Rasmussen and Williams 2005).

C. The Gram Matrix ($\boldsymbol{K}$) and Computational Stability

The Gram matrix ($\boldsymbol{K}$) is the central object of the GPR training process. It is a square matrix that computes the covariance between all pairs of input data points in the training set, X:

$$K_{ij} = k\left(x_i, x_j\right) + \delta_{ij}\sigma_n^2$$

where $k\left(\boldsymbol{x_i}, \boldsymbol{x_j}\right)$ is the value of the composite kernel function (Matern + White Noise) evaluated between the *i*th and *j*th training input vectors, and δ_{ij} is the Kronecker delta function. The term δ_{ij} is the Kronecker delta function, which is formally defined as:

$$\delta_{ij} = \begin{cases} 1, & if\ i = j \\ 0, & if\ i \neq j \end{cases}$$

The purpose of the Kronecker delta is to ensure that the noise term, σ_n^2 (originating from the White Noise component), is only added to the diagonal elements of the Gram matrix (Rasmussen and Williams 2005).

The diagonal elements of $\boldsymbol{K}$ represent the total variance attributed to each data point, while the off-diagonal elements represent the covariance between points. For the GPR inference to be mathematically tractable, specifically for the required matrix inversion and logarithm of the determinant step in the marginal likelihood calculation, the Gram matrix must be Positive Definite (PD) (Rasmussen and Williams 2005) . In financial time-series, highly correlated inputs often cause $\boldsymbol{K}$ to become ill-conditioned (close to singular), making the $\boldsymbol{\sigma_n^2}$ term critical for regularizing the matrix and maintaining the PD constraint (Golub and Loan 2013).

3.1.2 Novel Hyperparameter Initialization: Aggressive Noise Justification

During the training process, the optimization of the GPR hyperparameters (particularly $\boldsymbol{\sigma_n^2}$) is highly susceptible to collapsing into local optima when faced with financial data characterized by a low signal-to-noise ratio (SNR) (Rasmussen and Williams 2005).

To mitigate this known instability, the Aggressive Noise Initialization (ANI) strategy is introduced. This novel contribution involves setting the initial σ_n^2 to a deliberately large, calculated value near the empirical variance of the training set's residual errors. The ANI strategy is mathematically and regulatorily justified on three key principles:

1. **Enforcing Conservatism (Regulatory Alignment):** The ANI strategy aims to enforce conservatism. By attributing a larger initial share of the total variance to the unpredictable noise (σ_n^2), the model is driven to learn a smoother, simpler structural function $f(x)$. This methodological bias aligns with regulatory requirements for robust, conservative and non-misleading Stressed VaR forecasts (Hull 2018).

2. **Numerical Stability and Robustness:** The large initial $\boldsymbol{\sigma_n^2}$ acts as an aggressive "jitter," ensuring the Gram matrix K is well-conditioned from the start, mitigating the subsequent risk of numerical instability and LinAl errors (as discussed in Section 3.2) (Golub and Loan 2013).
3. **Guiding to the Global Optimum:** The aggressive initialization flattens the likelihood surface, providing the gradient-based optimization algorithm (L-BFGS-B) the necessary momentum to escape tight, unstable local optima and converge toward a robust solution (Rasmussen and Williams 2005; Snoek et al. 2012).

3.1.3 Hyperparameter Optimization and the L-BFGS-B Algorithm

The estimation of the optimal kernel hyperparameters (ℓ, σ_{f}^2, σ_{n}^2) is achieved by maximizing the marginal likelihood of the training data (also known as the evidence), $\mathcal{L}(\theta)$. This objective function measures how well the chosen kernel and its parameters fit the data while simultaneously penalizing model complexity.

The marginal likelihood is mathematically defined as:

$$\log \mathcal{L}\,(\theta) = -\frac{1}{2} y^T (K + \sigma_n^2 I)^{-1} y - \frac{1}{2} \log|K + \sigma_n^2 I| - \frac{T}{2} \log(2\pi)$$

where θ represents the vector of hyperparameters, $\mathbf{y}$ is the vector of observed returns, K is the Gram matrix, and T is the number of observations (Rasmussen and Williams 2005). Maximizing this term is the standard procedure for hyperparameter selection in GPR.

The optimization is performed using the Limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm with simple box constraints (L-BFGS-B) (Snoek et al. 2012). This is a quasi-Newton optimization method selected for its efficiency and ability to handle constrained optimization problems.

- **Computational Efficiency:** L-BFGS-B is preferred as it avoids the computational cost of storing and calculating the dense Hessian matrix (the second-order derivatives) by using a limited memory approximation based on past gradients. This makes the iterative optimization computationally feasible for the GPR framework (Rasmussen and Williams 2005).
- **Constrained Optimization:** The algorithm allows for the imposition of **simple box constraints** on the hyperparameter space (e.g., ensuring ℓ, σ_f^2, σ_n^2 remain strictly positive), which is necessary for maintaining numerical stability and ensuring the parameters correspond to physically meaningful quantities (variances and lengthscales) (Rasmussen and Williams 2005).

3.1.4 Hyperparameter Initialisation Scheme

The starting values and boundary conditions supplied to the L-BFGS-B optimizer (as detailed in Section 3.1.3) significantly influence the convergence efficiency and the specific local optimum achieved (Rasmussen and Williams 2005). For the purposes of replication and transparency, the precise initialization scheme and optimization bounds for the hyperparameters used across the different GPR models (Univariate, Stressed, and Hybrid components) are formally documented in Table 1.

The optimization strategy is based on two core principles:

1. **Non-Restrictive Bounds for the Matern Kernel:** The Matern kernel's hyperparameters, namely the lengthscale (l) and the signal variance (σ_f^2), are determined entirely through the optimization process. The lengthscale is constrained by wide, non-restrictive bounds (set to $1e-5, 1e6$ in the code) to ensure the L-BFGS-B algorithm can freely explore the likelihood

landscape and empirically determine the optimal volatility memory structure without bias from predefined initial values (Snoek et al. 2012).

2. **Aggressive Noise Initialization (ANI):** The noise variance ($\boldsymbol{\sigma_n^2}$) component is the only parameter set with a calculated initial value, following the Aggressive Noise Initialization (ANI) strategy detailed in Section 3.1.2. This initialization serves as a crucial regularization mechanism, directing the optimizer toward a more parsimonious and conservative solution by setting the initial $\boldsymbol{\sigma_n^2}$ equal to the full empirical variance of the training data. This is a conservative setting in-line with regulatory guidelines (Hull 2018; Weller 2019).

The resulting optimized hyperparameters for the GPR models are those that maximize the log-marginal likelihood given these initialization conditions.

Table 1: Initialization scheme for the GPR models

Model Type	Parameter	Symbol (Kernel Component)	Initial Value/ Bounds	Justification
All GPR Components	Lengthscale	l (Matern)	Bounds: $(1e-5, 1e6)$ (Wide)	Wide, non-restrictive bounds set to allow the L-BFGS-B optimizer to explore the likelihood landscape freely, ensuring the optimal lengthscale (medium-term memory) is found via optimization.
All GPR Components	Signal Variance (Scale)	σ_f^2 (Matern)	Optimized using L-BFGS-B	Integrated into the Matern kernel structure and fitted during optimization. Not separately initialized but determined by the data's overall variability.
Univariate GPR	Noise Variance	σ_n^2 (White Kernel)	Calculated from Training Data (ANI)	Aggressive Noise Initialization (ANI): Initial value set equal to the full variance of the training set returns. This aggressive initialization acts as regularization, encouraging a parsimonious fit and conservative forecast.
Stressed GPR	Noise Variance	σ_n^2 (White Kernel)	Calculated from Stressed Returns (ANI)	ANI Applied to Stress: Initial value set equal to the full variance of the stressed returns. This ensures the

Model Type	Parameter	Symbol (Kernel Component)	Initial Value/ Bounds	Justification
				GPR adequately captures the increased uncertainty under stress, enforcing conservatism.
Hybrid GPR-HS	Asset Noise Variance	σ_n^2 (White Kernel)	Inherits from Univariate GPR (ANI)	The portfolio risk relies on the variances predicted by the underlying Univariate GPRs. Thus, the noise initialization logic for each asset is inherited from the Univariate GPR setting.

3.1.5 Implementation and Code Snippets

The GPR models were implemented in Python utilizing the Scikit-learn library. The following figures provide illustrative code snippets documenting the specific implementation choices, including the construction of the composite kernel, the calculation for the Aggressive Noise Initialization (ANI), and the boundary constraints applied to the L-BFGS-B optimizer.

Figure 2: Univariate GPR initializations for each train set - basecase

```
# --- Define bounds and initial noise_level based on training
variance for THIS split ---
split_train_variance = max(split_train_variance, 1e-5)
length_scale_bounds = (1e-5, 1e6)
noise_lower_bound = max(1e-5, split_train_variance * 0.01)
noise_upper_bound = max(noise_lower_bound + 1e-5,
split_train_variance * 10)
noise_level_bounds = (noise_lower_bound, noise_upper_bound)
initial_noise_level = split_train_variance

matern_kernel_split = Matern(nu=2.5,
length_scale_bounds=length_scale_bounds)
white_kernel_split = WhiteKernel
(noise_level=initial_noise_level,
noise_level_bounds=noise_level_bounds)
univariate_kernel_for_split = matern_kernel_split +
white_kernel_split # Corrected variable name

ugpr = GaussianProcessRegressor
(kernel=univariate_kernel_for_split, alpha=0.0,
                                 n_restarts_optimizer=10,
                                 random_state=42)
```

Figure 3: Univariate GPR initializations for each train set – stress case

```
length_scale_bounds = (1e-5, 1e6) # Wide bounds for length scale
noise_lower_bound = max(1e-5, stressed_returns_variance * 0.01)
# Lower bound related to stressed variance
noise_upper_bound = max(noise_lower_bound + 1e-5,
stressed_returns_variance * 10) # Upper bound related to
stressed variance
noise_level_bounds = (noise_lower_bound, noise_upper_bound)
initial_noise_level = stressed_returns_variance # Initialize
with the stressed variance

# Define the kernel for this stressed series
matern_kernel_stressed = Matern(nu=2.5,
length_scale_bounds=length_scale_bounds)
white_kernel_stressed = WhiteKernel
(noise_level=initial_noise_level,
noise_level_bounds=noise_level_bounds)
kernel_stressed_ugpr = matern_kernel_stressed +
white_kernel_stressed

# --- Fit Univariate GPR ---
model = GaussianProcessRegressor(kernel=kernel_stressed_ugpr,
alpha=0.0,
                                 n_restarts_optimizer=10,
                                 random_state=42)
```

3.2 Data Preparation and Cross-Validation Strategy

3.2.1 Data Sourcing and Transformation

This section provides a detailed account of the data utilized, the rationale for the source selection, the required data cleaning and transformation steps, and the robust cross-validation scheme employed for model evaluation. The study utilizes daily closing price data starting from June 30, 2020 to June 30,2025 for a diversified portfolio of major global stock indices, including the S&P 500 (^GSPC), EURO STOXX 50 (^STOXX50E), Nikkei 225 (^N225), FTSE 100 (^FTSE), Hang Seng Index (^HSI), Nifty 50 (^NSEI), and BSE SENSEX (^BSESN). Each training set is an expanding window with a fixed-size test set containing 252 daily returns. These data were sourced via Yahoo Finance.

The choice of Yahoo Finance is justified as it is a widely recognized and industry-standard resource for obtaining reliable end-of-day historical prices in academic and non-HFT professional research environments. Given that the VaR methodology focuses on modelling long-term volatility clustering and systemic risk rather than real-time execution, the quality of the historical time series is sufficient and appropriate for the study's objectives (Linsmeier and Pearson 2000).

The following data imputations and transformations have been performed on the data to make it model-ready:

1. **Weekend Dates Removal and Synchronization:**
   Weekend dates (Saturday and Sunday) are universally removed from the dataset to establish a continuous time series of trading days. Crucially, no specific dates for region-specific holidays have been considered for removal. Since the indices span multiple geographies with disparate holiday calendars, introducing gaps for

every local holiday would severely fragment the dataset and break the required temporal synchronization. Given the need to maintain a synchronous time series for accurate estimation of portfolio-level covariance (VCV) for all assets on the same day, all non-weekend dates are retained. Prices are then aligned across all indices to this unified trading calendar (Tsay 2005).

2. **Price Cleaning and Imputation Strategy:**

   Missing values within the price series (due to the non-synchronous national holidays or market closures) are systematically handled using a robust two-step imputation strategy on the price data (P_t):

   - **Forward Fill**: Any internal missing price observation is filled with the last observed valid price. This is the standard practice in financial modelling, as it assumes the asset's price remains unchanged during the closure period until the next market activity is recorded (Tsay 2005).
   - **Backward Fill**: Any remaining missing values at the very beginning of the dataset are imputed using backward filling. This is necessary to ensure the series is fully complete from the start date for the subsequent return calculation. The imputation strategy ensures the synchronous and complete nature of the time series, which is fundamental for multivariate risk analysis. Related data tables are presented in Appendix A.

3. **Logarithmic Returns Calculation:**

   The cleaned and imputed price data is then transformed into daily logarithmic returns ($\boldsymbol{r_t}$), expressed in percentage form:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \times 100$$

This transformation is preferred over simple returns because log returns are time-additive, simplifying multi-period return aggregation, and they generally exhibit greater symmetry around zero (Tsay 2005). Furthermore, the transformation ensures the data is approximately stationary, a desirable property for time-series modelling that enhances the stability of the Gaussian Process regression (Han and Zhang 2015).

3.2.2 Feature Engineering and Cross-Validation Strategy

The GPR model is designed to capture non-linear temporal dependencies in volatility. It employs the ordinal date of the observation as the sole input feature **(**X**)**, converting the date into a continuous, numerical scalar (Rasmussen and Williams 2005). This simple feature engineering choice allows the model's complex kernel structure to learn the inherent non-linear, time-varying nature of volatility clustering present in financial returns (Han and Zhang 2015). The target variable (y) is the daily log return r_t.

The model evaluation is conducted using a Forward Chaining cross-validation scheme. This rigorous technique is necessary for time-series forecasting as it maintains the chronological order and strictly prevents look-ahead bias (Tsay 2005). The process involves:

1. Yearly Splits: The total dataset is segmented into successive yearly prediction blocks (e.g., 2021, 2022, 2023).
2. Expanding Training Window: For each prediction block, the GPR model is trained on all preceding historical data (the increasing input window). This mimics a real-world setting where all available history is used for current forecasting (Tsay 2005).
3. Out-of-Sample Test: Predictions are made only for the subsequent out-of-sample test period (the current yearly block).

This forward chaining approach ensures a robust, sequential out-of-sample validation where the training data accumulates over time, providing the most reliable assessment of the model's true predictive performance. The GPR model employs the ordinal date of the observation as the sole input feature (X), converting the date into a continuous, numerical scalar. The target variable (Y) is the daily log return r_t.

The model evaluation is conducted using a Forward Chaining cross-validation scheme. This involves segmenting the dataset into yearly prediction blocks, where the GPR model is trained on all preceding historical data (the *increasing* input window) and predictions are made for the subsequent out-of-sample test period.

The process involves:

1. Yearly Splits: The total dataset is segmented into successive yearly prediction blocks (e.g., 2021, 2022, 2023).
2. Expanding Training Window: For each prediction block, the GPR model is trained on all preceding historical data (the *increasing* input window).
3. Out-of-Sample Test: Predictions are made only for the subsequent out-of-sample test period (the current yearly block).

This forward chaining approach ensures a robust, sequential out-of-sample validation where the training data accumulates over time, providing the most reliable assessment of the model's true predictive performance (Lopez 1998).

3.3 Research Purpose and Questions

The methodological framework presented here is designed to systematically address a specific set of objectives formulated to answer critical gaps in the current body of quantitative risk literature. The findings from the empirical analysis are intended to provide robust evidence supporting the adoption of the Gaussian Process Regression (GPR) framework in advanced financial risk modeling. The research is guided by the following four primary questions:

1. **Assess GPR's Applicability Across Markets:** The application of Gaussian Process Regression must be systematically assessed across a diverse set of international equity market indices. This step is necessitated to establish the universality and robustness of GPR as a reliable volatility forecasting tool at the univariate level. The study's portfolio, encompassing indices from various economic environments and regulatory regimes (e.g., the S&P 500 versus the BSE SENSEX), allows for the model's effectiveness in capturing heterogeneous volatility dynamics to be evaluated. It is therefore investigated whether GPR's non-parametric, flexible approach can successfully model the distinct volatility clustering behavior observed in different global markets.
2. **Evaluate GPR's Stress Scenario Performance:** The evaluation of GPR's predictive accuracy under stress conditions is required to test its capacity for both regulatory compliance and practical risk management. The performance of the GPR-calibrated models is rigorously tested under a range of hypothetical stress scenarios covering geopolitical events (e.g., "West Asia War"), technology risk (e.g., "AI bubble / regulatory burden"), and emerging systemic threats (e.g., "Climate Risk"). This focused examination is undertaken to confirm that the model's volatility forecasts remain stable and methodologically sound when the

underlying data is subjected to severe, pre-defined market shocks. This process of applying stress to the model's inputs validates its suitability for use in the Stressed VaR (SVaR) calculation-a fundamental requirement under modern banking regulations. The primary goal is to demonstrate that the model can be systematically adapted to a hypothetical adverse environment without internal collapse or divergence, thereby confirming its robustness for regulatory risk measurement.

3. **Evaluate Hybrid Multivariate Efficacy:** The central hypothesis of the research-the superiority of the hybrid framework-must be empirically validated. Therefore, the efficacy of the novel Hybrid GPR-HS (Historical Simulation) framework in capturing and forecasting portfolio-level risk dynamics is evaluated. This hybrid method, which combines the GPR's dynamic volatility forecasts with the HS's static stress covariance matrix, is intended to provide a measure that is both forward-looking and robustly correlated. The framework's performance is gauged using the Quadratic Loss function and standard backtesting metrics, including the number of violations (Regulatory Loss Count). This rigorous quantitative comparison against the industry-standard Historical Simulation (HS) VaR benchmark is performed to determine if the GPR-HS method yields a more accurate and better-calibrated risk measure for the multi-asset portfolio.

4. **Enhance Model Explainability and Regulatory Adaptability:** Methods that enhance the explainability of GPR models must be developed to address the "black box" concern often associated with complex machine learning techniques. The application of the SHapley Additive exPlanations (SHAP) technique is specifically undertaken to ensure that the model's outputs are interpretable and actionable for both risk managers and supervisory bodies. Furthermore, the GPR

models and the resulting Hybrid SVaR framework are adapted to align with evolving regulatory requirements, such as those outlined in the Basel framework, ensuring that the models remain effective and compliant within a changing regulatory landscape. This focus ensures that the proposed methodology is not only statistically superior but is also pragmatically useful in meeting governance standards.

3.4 Research Design

This study employs a rigorous quantitative research design utilizing a causal-comparative and computational modeling approach. The research design is structured to systematically test the efficacy and robustness of the proposed novel model against a defined industry benchmark and under multiple hypothetical stress scenarios.

3.4.1 Type of Study

The study is fundamentally non-experimental and quantitative. It is primarily classified as a model evaluation (causal-comparative) design, where the performance of the newly developed Hybrid GPR-HS model is compared directly to the established Historical Simulation VaR benchmark (Linsmeier and Pearson 2000). The comparison is based on the quantification of risk forecasting accuracy using a consistent set of statistical criteria and loss functions (Lopez 1998).

A. Causal-Comparative Focus

The core comparative element lies in evaluating the causal effect of the modeling technique on the accuracy of the risk forecast. Specifically, the following comparisons are performed:

1. **GPR vs. HS Benchmark**: The predictive performance of the dynamic GPR model is measured against the static HS benchmark (the Historical VaR or HVaR) using two main criteria:
    - **Loss Minimization**: The Quadratic Loss function is calculated and compared across the models for all backtesting periods, determining which model penalizes extreme forecast errors more effectively (Lopez 1998).
    - **Regulatory Backtesting**: The models are assessed using the number of violations (the simple regulatory loss count), which determines if the

models exceed the accepted regulatory risk tolerance (e.g., 1% violations for a 99% confidence level) (Hull 2018).

B. Computational Modeling Approach

The SVaR component of the research utilizes a computational modeling approach. The SVaR is not merely observed from history but is a construct derived from combining the dynamic GPR output with the static average VCV matrix, which is calibrated by taking the mean of the VCV matrices calculated during the 2008 Global Financial Crisis and the COVID-19 market crisis. This composite matrix is then utilized alongside the dynamic GPR output when subjected to hypothetical stress scenarios. This computational element ensures that the resulting risk measure meets regulatory requirements by demonstrating how the VaR would behave under both current market conditions (via GPR) and a comprehensive, two-period systemic stress environment (via the averaged VCV) (Weller 2019).

3.5 Core Methodology: Gaussian Process Regression (GPR)

As established in Section 3.1, the GPR framework models volatility as a distribution over functions rather than a fixed parametric process. Building on this foundation, the present section focuses on its implementation details.

Justification for GPR over GARCH

While GARCH models have been the benchmark for capturing volatility clustering (the tendency for large volatility shocks to follow large shocks), they suffer from limitations that GPR is capable of overcoming (Tsay 2005):

Table 2: GPR's advantages over GARCH

Feature	GARCH Model Limitation	GPR Model Advantage
Functional Form	Requires explicit specification (e.g., GARCH(1,1)) of the autoregressive structure; mis-specification leads to biased forecasts.	Non-parametric and data-driven. The function is learned from the data, minimizing the risk of model mis-specification.
Temporal Dependencies	Assumes volatility dependencies are linear in their parameters.	Naturally handles non-linear, multi-scale dependencies and long-memory effects through the kernel function.
Output	Provides a single point estimate of conditional variance (σ_t^2).	Provides a full posterior distribution ($\mathcal{N}(\mu, \sigma^2)$) for the forecast, quantifying the model's confidence in its prediction.

Therefore, GPR is employed to exploit its capacity for capturing these elusive non-linear features in volatility, which are crucial for generating accurate risk measures.

3.5.1 The Gaussian Process Definition

A Gaussian Process (GP) is defined as a collection of random variables, any finite number of which have a joint Gaussian distribution. Consequently, any function $f(x)$ drawn from a GP is completely specified by its mean function by its mean function *m(***x***)* and its covariance function (kernel) *k*(**x**, **x'**) (Rasmussen and Williams 2005):

$$f(\mathrm{x}) \sim \mathcal{GP}(m(\mathrm{x}), k(\mathbf{x}, \mathbf{x}'))$$

In this financial application, the input $\boldsymbol{x}$ is defined by the ordinal date, and the target $f(\boldsymbol{x})$ is the observed log return r_t. The prior mean function $\boldsymbol{m}(\boldsymbol{x})$ is set to zero, $\boldsymbol{m}(\boldsymbol{x}) = \mathbf{0}$, reflecting the assumption of no prior knowledge regarding the expected return (Rasmussen and Williams 2005). The entire model structure, flexibility, and implicit assumption set are thus encoded within the kernel.

3.5.2 Kernel Function Specification and Justification

The kernel function k(x, x') encodes the prior assumptions about the return time series and models the covariance between any two points $\boldsymbol{x}$ and $\boldsymbol{x}'$. The kernel is meticulously chosen to reflect the known properties of financial time series (Rasmussen and Williams 2005). The chosen kernel is a summation of a structured (signal) component and a noise component:

$$k(x_i, x_j) = k_{Matern}(x_i, x_j) + k_{White}(x_i, x_j)$$

A. Matern Kernel ($\boldsymbol{k}_{\mathrm{Matern}}$): The Volatility Signal

Matern Kernel k_{Matern}, is selected to model the systematic, smooth component of volatility, representing the persistent volatility clustering and non-stationary effects often observed (Han and Zhang 2015).

$$k_{\text{Matern}}(r) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)}\left(\sqrt{2\nu}r\right)^{\nu} K_{\nu}\left(\sqrt{2\nu}r\right)$$

where $r = |x_i - x_j|$ is the distance between the two input dates, and K_ν is the modified Bessel function (Rasmussen and Williams 2005).

1. Smoothness ($\nu = 2.5$): The specific setting of the smoothness parameter $\boldsymbol{\nu}$ to 2.5 is critical. This value ensures the resulting function is twice-differentiable, striking a balance between smoothness (required for stable volatility forecasting) and the flexibility to model rough, complex data structures (Han and Zhang 2015).
2. Length Scale (ℓ): This hyperparameter governs how quickly the correlation between returns decays over time, directly modeling the persistence of volatility shocks-a key feature of financial time series (Rasmussen and Williams 2005).

Rationale for Matern over RBF Kernel

The Matern kernel is explicitly chosen over the standard Radial Basis Function (RBF) kernel (also known as the Squared Exponential kernel) due to fundamental differences in their smoothness assumptions. The RBF kernel assumes a process that is infinitely differentiable, implying an overly smooth and unrealistic process for financial returns, which are known to be inherently rough and non-smooth at the daily level (Rasmussen and Williams 2005). Conversely, the Matern kernel, with $\nu = 2.5$, assumes the underlying process is only twice-differentiable. This lower degree of assumed smoothness introduces a more appropriate level of local variability and roughness into the model's function draws, allowing the GPR to more realistically capture the abrupt changes and non-smooth temporal dependencies characteristic of high-frequency financial data (Han and Zhang 2015). Therefore, the Matern kernel provides a more robust and less restrictive prior for modeling asset price dynamics.

B. WhiteKernel (k_{White}): The Heteroskedastic Noise

The WhiteKernel is essential for financial applications as it models the observation noise (σ_n^2), capturing the residual, time-varying, unpredictable shock component (idiosyncratic risk) in daily returns-the very definition of heteroskedasticity (Rasmussen and Williams 2005).

$$k_{\text{White}}(x, x') = \sigma_n^2 \delta_{x,x'}$$

Aggressive Noise Initialization (ANI)

To optimize the model's convergence and ensure the WhiteKernel accurately captures the empirical variance of the noisy financial time series, an aggressive noise initialization (ANI) strategy is employed, where the initial σ_n^2 is explicitly set equal to the observed empirical variance of the training returns. This deliberate initialization strategy, further detailed in Section 3.1.3, is utilized to avoid local optimization minima and improve the computational efficiency inherent in the sequential fitting process.

3.5.3 GPR Posterior Inference and Hyperparameter Optimization

Given a training dataset $\mathbf{D} = \{(\mathbf{X}, \mathbf{y})\}$, the GPR prediction for a new, unseen time point $\mathbf{x}_*$ yields a Gaussian distribution, defined by its conditional mean and variance :

$$\mathbf{y}_* | \mathbf{X}, \mathbf{y}, \mathbf{x}_* \sim \mathcal{N}(\mu(\boldsymbol{x}_*), \sigma^2(x_*))$$

where the posterior mean $\mu(x_*)$ and posterior variance $\sigma^2(x_*)$ are calculated as:

$$\mu(x_*) = k_*^T (K + \sigma_n^2 I)^{-1} y$$

$$\sigma^2(x_*) = k(x_*, x_*) - k_*^T (K + \sigma_n^2 I)^{-1} k_*$$

In these equations, y is the vector of observed returns, K is the $N \times N$ covariance matrix calculated using the chosen kernel on the training inputs, k_* is the vector of covariances between the training inputs and the test input x_*, and I is the identity matrix (Rasmussen and Williams 2005). The resulting $\boldsymbol{\sigma}_*^2$ represents the predicted conditional volatility of the asset at time *t+1*.

The kernel hyperparameters (such as the length scale and amplitude of the Matern kernel, and the noise level σ_n^2 of the White Kernel) are optimized by maximizing the log marginal likelihood (LML) of the training data using the L-BFGS-B optimization algorithm (Rasmussen and Williams 2005). Maximizing the LML ensures a statistical balance between goodness-of-fit and model complexity (the principle of Occam's razor), resulting in parameters that best explain the observed data without overfitting (Rasmussen and Williams 2005).

3.6 Hybrid Risk Framework: GPR-HS

The Hybrid GPR-HS (Gaussian Process Regression-Historical Simulation) architecture is the specific computational implementation utilized to address Research Question 3, which concerns Multivariate Efficacy. This framework is necessitated by the computational challenges inherent in modeling large, multi-asset portfolios and represents the core innovation of the study's quantitative design.

3.6.1 Justification for Hybridization

Full Multivariate GPR (MGPR) is theoretically capable of modeling complex, non-linear dependencies across assets and time. However, the computational cost of MGPR is prohibitive for the study's requirements. MGPR requires the inversion of a large, dense covariance matrix whose size grows quadratically with both the number of data points (T) and the number of assets (N), leading to computational intractability for large datasets and diverse portfolios (Rasmussen and Williams 2005). The inversion complexity, which is $\mathcal{O}((N \cdot T)^3)$, renders full MGPR infeasible for the forward-chaining cross-validation scheme employed here (Han and Zhang 2015).

The Hybrid GPR-HS model resolves this complexity by decoupling the dynamic, asset-specific volatility from the static, inter-asset correlation. The predictive power of GPR is maintained on the critical univariate volatility dimension, while the stability and computational tractability of Historical Simulation (HS) are leveraged for the inter-asset correlation component. This hybridization allows the framework to generate a conditional (time-varying) portfolio VaR forecast while remaining computationally efficient and scalable across the expansive backtesting window.

3.6.2 Portfolio Conditional Variance Construction

The total portfolio variance $\boldsymbol{\sigma_p^2}$ is derived from the portfolio covariance matrix $\boldsymbol{\Sigma_t}$ using the quadratic form $\boldsymbol{\sigma_p^2} = \boldsymbol{w}^\top \boldsymbol{\Sigma_t} \boldsymbol{w}$, where $\boldsymbol{w}$ is the vector of portfolio weights. The key to the GPR-HS framework lies in the construction of the hybrid covariance matrix $\boldsymbol{\Sigma_t}$, which is dynamically constructed for each forecast date by substituting the GPR-derived variance forecasts into the diagonal and using stable historical estimates for the off-diagonal elements:

1. **Diagonal Elements ($\boldsymbol{i = j}$):** GPR-Predicted Conditional Variance ($\sigma^2_{\mathrm{GPR},i,t}$): The diagonal elements of the hybrid matrix are populated using the time-varying, GPR-predicted conditional variance ($\sigma^2_{\mathrm{GPR},i,t}$) obtained from the univariate fit for each asset *i*. This ensures that the asset-specific risk component of the portfolio variance is dynamic, non-linear, and forward-looking, directly incorporating the GPR's superior ability to forecast volatility clustering.
2. **Off-Diagonal Elements ($\boldsymbol{i \neq j}$):** Historical Covariances ($\sigma_{\mathrm{HS},i,j}$): The off-diagonal elements are populated using the stable historical covariances ($\sigma_{\mathrm{HS},i,j}$) estimated directly from the VCV matrix of the latest training data window (Historical Simulation).

These covariances are treated as static (or slowly updated via the expanding training window), representing the average correlation structure observed during the training history. This choice is necessitated by computational constraints but is justified by the observation that volatility clustering (captured by the diagonal) typically exhibits higher persistence and greater time-variation than the underlying inter-asset correlation (captured by the off-diagonal). This structure ensures the portfolio risk measure is conditional (time-varying) due to the GPR-driven diagonals, yet highly scalable and computationally tractable due to the static, historically estimated off-diagonals.

The Hybrid GPR-HS Covariance Matrix

The resulting GPR-HS matrix is formally defined as:

$$\Sigma_{\text{GPR-HS}} = \begin{pmatrix} \sigma^2_{\text{GPR},1} & \sigma_{\text{HS},12} & \cdots & \sigma_{\text{HS},1N} \\ \sigma_{\text{HS},21} & \sigma^2_{\text{GPR},2} & \cdots & \sigma_{\text{HS},2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{\text{HS},N1} & \sigma_{\text{HS},N2} & \cdots & \sigma^2_{\text{GPR},N} \end{pmatrix}$$

The final one-day Portfolio VaR (VaR_p) is then derived from the resulting portfolio standard deviation $\sigma_p = \sqrt{w^T \Sigma_{\text{GRH}} w}$ using the appropriate quantile factor $\boldsymbol{z_\alpha}$ (Jorion 2006).

3.6.3 Stressed SVaR Hybridization

The GPR-HS framework is adapted for the Stressed VaR (SVaR) calculation by substituting the Historical Simulation component with a stressed equivalent. For SVaR, the hybrid covariance matrix $\boldsymbol{\Sigma_{\text{SA},t}}$ is constructed as follows:

1. **Diagonal Elements (*i* = *j*):** Stressed GPR Variance ($\sigma^2_{\text{Stressed},i,t}$): The variance diagonals are taken from the GPR model that was re-calibrated on the stressed return series (Block 16). This component is dynamic, capturing the model's projected volatility under the influence of the scenario path shocks.
2. **Off-Diagonal Elements (*i* ≠ *j*):** Averaged Stress VCV ($\sigma_{\text{ASVCV},i,j}$): The off-diagonal elements are populated using the static Averaged Stress VCV(Σ_{ASVCV}) matrix, which is the mean of the VCV matrices calculated during the 2008 Global Financial Crisis and the COVID-19 market crisis. This ensures that the portfolio correlation used for SVaR is pessimistic, regulatory-compliant, and derived from multiple periods of systemic risk.

This further hybridization confirms the framework's utility in generating both standard VaR and the required regulatory SVaR without requiring any additional complex multivariate modeling.

3.6.4 Derivation of Portfolio Risk Metrics (VaR and ES)

The ultimate goal of the Hybrid GPR-HS framework is the calculation of portfolio VaR and ES. Assuming that the portfolio returns ($\boldsymbol{r_{p,t}}$) are conditionally distributed according to a Student's t-distribution with $\boldsymbol{\nu}$ degrees of freedom-a choice necessitated by the heavy-tailed nature of financial returns-the portfolio forecast distribution is defined as:

$$\mathrm{r}_{\mathrm{p,t}} | \mathcal{D} \sim t(\nu, \mu_{p,t}, \sigma_{p,t}^2)$$

where the portfolio conditional mean ($\mu_{p,t}$) and variance ($\sigma_{p,t}^2$) are derived from the Hybrid GPR-HS matrix:

$$\mu_{p,t} = w^{\top \mu_{\mathrm{GPR}t}}$$

$$\sigma_{p,t}^2 = w^{\top \Sigma_{\text{GPR-HS},t} w}$$

A. Value at Risk (VaR)

The VaR at a confidence level $\boldsymbol{\alpha}$ is defined as the maximum expected loss that will not be exceeded with a probability of $1 - \alpha$. Given the t-distribution assumption, the portfolio VaR is calculated as:

$$\mathrm{VaR}_{\alpha,t} = \mu_{p,t} + \sigma_{p,t} \cdot t^{-1}(1 - \alpha, \nu)$$

where $t^{-1}(1 - \alpha, \nu)$ is the $1 - \alpha$ quantile of the standardized t-distribution with $\boldsymbol{\nu}$ degrees of freedom (McNeil et al. 2015).

B. Expected Shortfall (ES)

The ES (or Conditional VaR) is a more robust risk measure, defined as the expected loss conditional on the loss exceeding the VaR threshold. Under the t-distribution assumption, the ES is calculated as:

$$\mathrm{ES}_{\alpha,t} = \mu_{p,t} - \sigma_{p,t} \cdot \left[\frac{g(t^{-1}(1-\alpha,\nu),\nu)}{1-\alpha} \cdot \frac{\nu + [t^{-1}(1-\alpha,\nu)]^2}{\nu - 1}\right]$$

where $g(t,\nu)$ is the probability density function (PDF) of the standardized t-distribution evaluated at the VaR quantile (Hull 2018).

Justification for Fixed Degrees of Freedom ($\boldsymbol{\nu = 5}$)

The selection of the degrees of freedom (ν) for the Student's t-distribution is a critical step in accurately modeling the portfolio returns, which are known to exhibit leptokurtosis (heavy tails) and non-normality. While the degrees of freedom could be statistically estimated via maximum likelihood, the value of ν is often fixed in risk management studies to ensure stability and comparability (Tsay 2005).

In this research, the degrees of freedom parameter is fixed at $\boldsymbol{\nu} = 5$. This choice represents a widely adopted standard in quantitative finance for modeling daily returns. Academic and practitioner consensus suggests that ν values typically fall between 4 and 6 for daily equity returns. The choice of $\nu = 5$ is justified because:

1. **Tail Modeling Efficacy**: A t-distribution with $\nu = 5$ possesses kurtosis $(\frac{\mathbf{6}}{\boldsymbol{\nu}-\mathbf{4}})$ of 6. This significantly higher kurtosis value compared to the Normal distribution (kurtosis of 3) allows the model to assign a much higher probability mass to extreme loss events, which is essential for accurate VaR and ES calculation.
2. **Finite Variance Guarantee**: A key requirement for any statistical model used in financial risk is that the variance must be finite. The t-distribution satisfies this condition provided that $\boldsymbol{\nu} > 2$. The choice of $\nu = 5$ safely exceeds this

minimum threshold, guaranteeing a stable and mathematically well-defined conditional variance for the portfolio.

This fixed parameter ensures the stability of the VaR and ES estimates across the forward-chaining backtesting window, conforming to common industry practices where volatility is modeled dynamically ($\sigma_{p,t}$) but the tail parameter (ν) is held constant.

3.7 Risk Metric Calculation: VaR and ES

3.7.1 Distributional Assumption

The risk metrics are calculated under the assumption that the standardized portfolio returns follow a Student's t-distribution with $\boldsymbol{\nu} = \mathbf{5}$ degrees of freedom. The full justification for this distributional choice, which addresses leptokurtosis (fat tails) and the selection of the fixed parameter $\boldsymbol{\nu} = \mathbf{5}$ as a quantitative finance standard, is provided in Section 3.6.4.

3.7.2 Analytical Calculation of VaR and ES

The analytical formulas used for calculating the portfolio VaR and ES at the $\boldsymbol{\alpha} = \mathbf{99}\%$ significance level remain identical to those derived in Section 3.6.4. The formulas are populated using the conditional mean ($\boldsymbol{\mu}_{\boldsymbol{p},\boldsymbol{t}}$) and standard deviation ($\boldsymbol{\sigma}_{\boldsymbol{p},\boldsymbol{t}}$) derived from the Hybrid $\boldsymbol{\Sigma}_{\boldsymbol{t}}$ matrix. The inputs used to populate these formulas shift based on the context:

- For the Base Case (VaR/ES): The $\boldsymbol{\Sigma}_{\boldsymbol{t}}$ matrix uses the Historical Simulation (HS) covariance for its off-diagonal elements.
- For the Stressed Case (SVaR/SES): The $\boldsymbol{\Sigma}_{\boldsymbol{t}}$ matrix uses the Averaged Stress VCV ($\boldsymbol{\Sigma}_{\boldsymbol{ASVCV}}$) for its off-diagonal elements.

In both contexts, the VaR and ES are derived analytically from the conditional t - distribution, ensuring consistency and regulatory prudence across normal and stressed risk regimes.

3.8 Model Validation and Backtesting

This section details the formal empirical procedures used to validate the Hybrid GPR-HS, serving as the conclusive test for the success of Research Problems 1, 2, and 3. The efficacy of the dynamic model is rigorously proven by comparing its backtest results and loss metrics against the static Historical Simulation (HS) VaR Benchmark. The HS VaR benchmark is calculated simply as the 1st percentile (corresponding to the 99% confidence level) of the empirical returns distribution from the preceding 252 trading days.

3.8.1 Backtesting Procedures and Statistical Benchmarks

The statistical validity of the VaR forecasts is assessed using a sequence of hypothesis tests that evaluate both the frequency and the temporal independence of violations.

1. **Unconditional Coverage (POF) Test** (Kupiec 1995)**:**
    - **Objective**: To check if the total number of observed VaR violations (N_1) over the test period is equal to the theoretically expected number (αN).
    - **Test Statistic**: The test uses the likelihood ratio (LR_{UC}) statistic, which is χ^2 distributed with 1 degree of freedom:

      $$\mathrm{LR_{UC}} = -2\ln\left[\frac{(1-\alpha)^{N-x}\alpha^{x}}{(1-\hat{\pi})^{N-x}\widehat{\pi^{x}}}\right]$$

      where α is the expected failure rate (1%), x is the number of violations (N_1), N is the total number of observations, and $\hat{\pi} = x/N$ is the observed failure rate.
2. **Conditional Coverage (CC) Test** (Christoffersen 1998)**:**
    - **Objective**: This is the critical test for dynamic models, as it jointly evaluates two conditions: correct unconditional coverage (correct frequency) and independence of violations (no clustering).

- **Test Statistic**: The CC test combines the LR test for unconditional coverage (LR_{UC}) with an LR test for independence (LR_{IND}), where the latter checks a first-order Markov dependence:

$$\mathrm{LR}_{\mathrm{CC}} = \mathrm{LR}_{\mathrm{UC}} + \mathrm{LR}_{\mathrm{IND}}$$

- **Conclusion for Research Problems 2 and 3**: Passing the CC Test validates the GPR-HS model's ability to track time-varying volatility and meets the Basel regulatory standard for model acceptance, confirming the statistical soundness of the dynamic portfolio VaR (Basel Committee on Banking Supervision [BCBS] 2019).

3. **Expected Shortfall (ES) Backtesting:**
   - **Objective:** To validate the accuracy of the entire tail distribution modeled by the ES forecast, ensuring the robust calculation of ES.
   - **Procedure:** The Unconditional ES Backtesting Procedure (based on McNeil and Frey's Zero Mean Test) is implemented (McNeil and Frey 2000). This test assesses the model by computing the average of the realized loss on VaR violation days and statistically testing whether it is consistent with the average of the predicted ES values on those same days. Successful validation further strengthens the evidence of the GPR-HS model's robust capability.

3.8.2 Economic Benchmarking using Quadratic Loss (QL)

To prove the economic superiority of the dynamic GPR-HS model and address Research Problems 1 and 2, the framework is subjected to an economic performance test by benchmarking it against the static Historical VaR (HVaR) Benchmark using the Quadratic Loss (QL) function.

1. **Rationale and Definition of QL:** The QL function provides an economic penalty that quantifies the severity of forecast errors by penalizing both the *occurrence* and the magnitude of the shortfall when a violation occurs (Lopez 1998). The benchmarking is executed using percentage returns, which is sufficient for comparing model forecast accuracy without requiring notional P&L calculations. The QL is calculated as the sum of squared shortfalls (S_t) occurring on violation days:

$$QL = \sum_{t=1}^{T} I_t \cdot (\mathrm{VaR}_t - r_t)^2$$

A model with a lower total QL score is superior, as its VaR forecasts were more precisely calibrated to the realized extreme losses, minimizing the required capital cushion for a given risk level.

2. **Justification for the GPR Framework:** The QL comparison serves as the primary empirical justification for the overall GPR-HS framework, including the Aggressive Noise Initialization (ANI) strategy. A result showing $\boldsymbol{QL_{GPR}} < Q\boldsymbol{L}_{\mathbf{HVaR}}$ will prove that the dynamic modeling of time-varying volatility and the ANI-stabilized optimization process lead to a demonstrably more accurate and better-calibrated risk forecast than the static regulatory standard. This directly supports the claim that the GPR-HS framework is a superior and efficient tool for financial risk management.

3.9 Stress Testing and Stressed VaR (SVaR)

This section details the simulation and modeling procedure to assess the framework's implementability and robustness under extreme conditions, directly addressing the goals of Research Problem 3 (Evaluating GPR's Stress Scenario Performance).

3.9.1 Scenario Construction and Simulation

Three distinct systemic risk scenarios are utilized, including: “West Asia War”, “Climate Risk” and “AI Bubble/ Regulatory burden”. Each scenario is mathematically defined by a specific cumulative return shock $R_{cumulative}$ over a 252-day horizon. This design aligns with regulatory ICAAP (Internal Capital Adequacy Assessment Process) stress-testing principles, which require the use of plausible, forward-looking cumulative shocks rather than large, instantaneous market jumps (Basel Committee on Banking Supervision (BCBS) 2005; 2009). The daily stressed return path $r_{i,t}^{stress}$ for each index *i* is generated using a combination of a deterministic linear drift ($L_{i,\,t}$) and historical-based stochastic noise ($N_{i,\,t}$):

$$r_{i,t}^{\text{stress}} = L_{i,t} + N_{i,t}$$

The deterministic drift ensures the paths reach the required cumulative shock, while the stochastic noise, drawn from the historical return distribution, ensures the daily volatility remain realistic.

3.9.2 Stressed GPR Fitting and Conservative Calibration

A new, scenario-specific GPR model is fitted to each index's 252-day stressed return path. This dedicated fitting process forms the core of the stress methodology.

To ensure the GPR framework yields a robust and non-underestimated forecast, the ANI strategy is employed. The initial noise level for the WhiteKernel (σ_n^2) is explicitly set equal to the high variance of the stressed returns series ($\sigma_{stressed}^2$). This action forces the Bayesian optimizer to begin its search in a region that conservatively reflects the stress-induced volatility. This methodological constraint ensures a robust and conservative forecast for the SVaR, demonstrating the framework's implementability on scenarios (Research Problem 3).

3.9.3 SVaR and SES Calculation

The final SVaR and Stressed Expected Shortfall (SES) values are calculated for the last day (T) of the stress path. This uses the standard GPR-HS method, but substitutes the inputs with the scenario-specific forecasts ($\mu_{i,T}$ and $\sigma_{i,T}$) generated by the stressed GPR models.

The extensive backtesting on the base case proves the framework's statistical efficacy (Research Problem 2) and predictive superiority (Research Problem 1). Therefore, the SVaR results are deliberately reserved exclusively as the input for the SHAP explainability analysis (Section 3.10), thereby addressing the critical goal of Research Problem 4 (Enhance Model Explainability and Regulatory Adaptability).

3.10 Risk Attribution and Interpretability: SHAP Analysis

This section is the primary mechanism for achieving the Explainability component of Research Problem 4 (Enhance Model Explainability and Regulatory Adaptability). By quantifying the influence of each model input on the final risk capital requirement, the framework provides the transparency necessary for regulatory trust and actionable risk management.

3.10.1 SHAP Methodology Overview

SHAP (SHapley Additive exPlanations) is a model-agnostic technique derived from cooperative game theory. It rigorously attributes the final outcome of a prediction (the SVaR value) to its specific input features (Lundberg and Lee 2017). It calculates the unique contribution (ϕ_i) of each feature by considering all possible permutations (coalitions) of features, guaranteeing a fair and consistent attribution.The fundamental SHAP equation links the explanation model g to the original complex model f (in this case, the SVaR calculation function):

$$g(z') = \phi_0 + \sum_{i=1}^{2N} \phi_i z'_i$$

where **z'** is the simplified input (binary vector representing presence / absence of a feature), ϕ_0 is the base value (expected SVaR), and ϕ_i is the SHAP value for feature *i*.

3.10.2 Application of Kernel SHAP to SVaR

1. The SHAP analysis is strategically applied to the SVaR forecasts generated under the stress scenarios, as these results represent the highest-stakes risk capital decision.

2. **SHAP Model**: The SVaR calculation function ($SVaR = f(\mu, \sigma)$) is treated as the original model *f*.
3. **Features**: The input features are the 2N outputs from the stressed GPR models: N predicted conditional means (μ_i) and N predicted conditional standard deviations ($\boldsymbol{\sigma_i}$) for all N assets.
4. **Decomposition**: The final SVaR is decomposed into the sum of the base value and the SHAP contributions:

$$\mathbf{SVaR_{final}} = \phi_0 + \sum_{i=1}^{N} \left(\phi_{\mu_i} + \phi_{\sigma_i}\right)$$

This analysis provides crucial managerial and regulatory insights by clearly separating the risk contribution driven by an asset's expected shock (Mean Component ϕ_{μ_i}) versus the contribution driven by its uncertainty (Volatility Component ϕ_{σ_i}). This satisfies the requirement for enhanced model transparency, providing the foundation needed for regulatory dialogue and actionable risk mitigation strategies (Basel Committee on Banking Supervision (BCBS) 2021).

3.11 Research Design Limitations

The methodological framework, while optimized for empirical testing and computational feasibility, operates under several necessary constraints and simplifying assumptions that define the scope of the study:

1. **Hybridization and Static Correlation (Pertains to Research Problem 2)**: The primary limitation is the use of the Historical Simulation (HS) covariance matrix for the off-diagonal elements in the Σ_t (Section 3.6). This static correlation assumption does not account for the known financial phenomenon that asset correlations tend to increase significantly during periods of high stress (correlation "flights to safety"). The Hybrid GPR-HS model, while dynamically capturing *univariate* volatility, may therefore slightly underestimate systemic risk during true crisis events when correlations are elevated.
2. **GPR Scalability and Complexity:** The decision to use the Hybrid GPR-HS architecture over a full Multivariate GPR (MGPR) model is an imposed limitation. Full MGPR, while theoretically superior for capturing joint-distribution dynamics, is computationally intractable for the portfolio size and time-series length required in this study, limiting the extent to which complex inter-market dependencies can be modeled. MGPR has a prohibitive $\boldsymbol{O}(\boldsymbol{N}^3)$ computational cost, which makes it impractical for the real-world application.
3. **Scope of Empirical Testing (Pertains to Research Problem 3)**: The empirical validation of the GPR-HS framework was conducted using two tiers of testing to achieve maximum statistical rigor:
    a. **Tier 1: Predictive Testing (Base Case)**: The GPR-HS model was validated against regulatory standards (Kupiec and CC tests) and economically benchmarked (QL) using non-stressed, out-of-sample

historical returns. This tier established the framework's predictive efficacy under normal market conditions (Research Problem 1).

b. **Tier 2: Solvency Testing (Stressed Case)**: To ensure the utmost confidence in the capital calculation, the SVaR forecasts were also subjected to Kupiec and CC testing. Given that the underlying SVaR methodology requires inputs derived from a historical crisis VCV matrix, the resulting validation procedure inherently examines the model's statistical stability and solvency under extreme, yet realized, market conditions (Research Problem 3).

This Tier 2 validation is a strategic choice to ensure the model exhibits the strongest possible regulatory compliance and statistical confidence for the final capital number. Critically, since the scenarios used in the SVaR calculation are hypothetical and forward-looking in nature, the estimates are not subject to the classical look-ahead bias. These SVaR results are leveraged for the subsequent SHAP analysis (Research Problem 4) to ensure the explained risk is derived from the most rigorously validated capital output.

3.12 Conclusion

The methodology detailed in this chapter establishes a robust, statistically rigorous, and computationally efficient framework for advanced portfolio risk measurement.

The Hybrid GPR-HS architecture provides a scalable and practical solution to capture dynamic, non-linear conditional volatility, effectively addressing the computational limitations of full Multivariate GPR (MGPR) $O(N^3)$. The empirical application of this architecture to diverse global markets directly fulfills Research Problem 1 (Applicability Across Diverse Markets).

The efficacy of this design is rigorously tested through Christoffersen's Conditional Coverage Test, proving its statistical superiority over the static Historical Simulation benchmark and successfully achieving Research Problem 2 (Modeling Inter-Asset Dependencies). Furthermore, the strategic, deliberate use of Aggressive Noise Initialization (ANI) ensures conservative and robust forecasts under the Stress Testing scenarios, satisfying the implementability requirements of Research Problem 3 (Evaluating GPR's Stress Scenario Performance).

Finally, the SHAP analysis applied to the final SVaR output provides the essential model interpretability needed to transform the framework from a "black box" into an actionable managerial tool. This comprehensive approach achieves the critical regulatory goal of Research Problem 4 (Enhance Model Explainability and Regulatory Adaptability).

This methodological structure is thus fully equipped to meet and test all stated research objectives, paving the way for the quantitative analysis presented in the following chapter.

CHAPTER IV:

RESULTS

4.1 Overview of Empirical Results

This chapter presents the empirical results and quantitative evidence derived from the application of the Hybrid Gaussian Process Regression-Historical Simulation (GPR-HS) framework. The purpose of this chapter is to objectively document the performance of the proposed model across the full spectrum of required regulatory risk metrics. The findings directly address the four primary research problems of this study by showcasing the framework's statistical validity, economic utility, performance under severe stress, and regulatory explainability.

The results are generated from the recursive estimation approach established in Chapter III, utilizing an expanding-window backtesting procedure applied to a portfolio of seven major global equity indices. This design aligns with established methods for verifying the accuracy of risk measurement models (Kupiec 1995). By utilizing GPR, the framework aims to capture the non-linearities and conditional heteroscedasticity inherent in financial time series, as traditionally explored in the GARCH literature (Engle 1982; Nelson 1991), but through a non-parametric Bayesian lens (Rasmussen and Williams 2005).

The analysis concludes with the application of the framework to forward-looking stress scenarios, designed to adhere to principles relevant to both CCAR and ICAAP capital planning requirements (Basel Committee on Banking Supervision (BCBS) 2009). The evidence is systematically presented to evaluate the GPR-HS model not just as a standalone tool, but as a robust evolution of standard volatility forecasting and risk management techniques (Jorion 2006; McNeil et al. 2015; Hull 2018).

The chapter is structured to systematically present the evidence for each research problem (RP):

Table 3: Overview of empirical results presentation

Section	Focus	Research Problem Addressed	Key Evidence
4.2	GPR Model Selection and Fit Justification	RP1: Assess GPR's Applicability Across Diverse Markets RP 2: Modeling Inter-Asset Dependencies	Log-Marginal-Likelihood (LML) Scores for Matern Kernel
4.3	Risk Measure Accuracy: VaR Backtesting	RP1: Assess GPR's Applicability Across Diverse Markets RP 2: Modeling Inter-Asset Dependencies	Backtesting p-values (Kupiec, CC, ES backtest)
4.4	Economic Utility and Inter-Asset Dependencies	RP1: Assess GPR's Applicability RP2: Modeling Inter-Asset Dependencies	Quadratic Loss (QL) comparison & Hybrid VCV Matrix
4.5	Dynamic Risk Tracking	RP1: Assess GPR's Applicability	Time-Series VaR Plot
4.6	Stress Capital Projection: SVaR and SES	RP3: Evaluating GPR's Stress Scenario stability	Final SVaR and SES Projections

Section	Focus	Research Problem Addressed	Key Evidence
4.7	Model Explainability: SHAP Analysis	RP4: Enhance Model Explainability and Regulatory Adaptability	SHAP Feature Contribution Plots

4.2 GPR Model Selection and Fit Justification (RP1)

This section details the selection and validation of the GPR kernel structure, confirming the model's appropriateness for stationary financial time series.

4.2.1 Kernel Selection and Stationary Data Assumption

The performance of the GPR model relies critically on the chosen covariance function, or kernel (Rasmussen and Williams 2005). The Matern 5/2 kernel combined with a White Noise kernel (Matern 5/2 + WhiteKernel) was selected. This choice is a direct consequence of the input data characteristics (Porcu et al. 2023):

1. **Stationary Returns:** Trend-based kernels (e.g., Dot Product) were rejected because daily returns exhibit stationarity (Tsay 2005). The Matern kernel focuses solely on modeling the short-term, rough non-smoothness inherent in volatility clustering, which is characteristic of financial returns (Stein 1999; Porcu et al. 2023).

2. **Volatile Noise Modeling:** The White Noise component explicitly models the independent, residual noise, which is critical for robust parameter estimation and for differentiating signal (volatility clustering) from noise in the highly volatile market data (Rasmussen and Williams 2005).

4.2.2 Log-Marginal-Likelihood (LML) Validation

The quality of the selected Matern kernel fit is validated using the Log-Marginal-Likelihood (LML). In a Bayesian framework, the LML serves as the primary objective function for model selection, representing the evidence that the chosen model generated the observed data (Rasmussen and Williams 2005). The high LML scores across all indices and splits confirm the GPR component is optimally fitted to the underlying volatility structure. This numerical evidence supports the claim that the GPR framework is applicable across diverse global markets (RP1), moving beyond traditional parametric constraints (Engle 1982; Nelson 1991).

Table 4: Univariate GPR LML Scores (Matern + WhiteKernel)

Index	2025 LML	2024 LML	2023 LML	2022 LML
Nikkei (^N225)	-1491.92	-1114.04	-739.70	-324.79
FTSE (^FTSE)	-1309.19	-1037.18	-725.78	-364.98
S&P 500 (^GSPC)	-1480.01	-1171.33	-758.59	-342.69
EUROSTOXX50 (^STOXX50E)	-1548.84	-1219.52	-833.52	-379.94
Hangseng (^HSI)	-1799.41	-1353.57	-847.55	-350.12
NSE (^NSEI)	-1340.91	-1029.98	-717.51	-349.53
BSE (^BSESN)	-1348.20	-1035.89	-721.13	-350.83

Note: Author's own calculations

The empirical results presented in Table 4 confirm two critical aspects of the Hybrid GPR-HS framework:

1. **Model Stability Across Time:** For every index, the LML consistently increases in magnitude from the smaller, earliest test split (Y2022) to the largest, final split (Y2025). This steady increase confirms that the GPR model successfully converged to an optimal fit across all splits, even as the data history and volatility structure evolved satisfying the requirement for dynamic adaptation in risk management (Jorion 2006; Hull 2018).

2. **Robustness Across Markets (RP1):** The LML scores for the highly volatile Asian markets (e.g., HSI at -1799.41) are the lowest, reflecting the greater complexity of modelling these indices. However, the fact that the GPR component consistently achieved optimal LML convergence across all seven markets-including those with extreme volatility-demonstrates the Matern kernel's inherent ability to adapt and provide an optimal covariance fit across diverse global market regimes-a property strongly argued by Stein, who demonstrated that the Matérn class is significantly more robust for modeling the 'rough' realizations found in real-world processes compared to the overly-idealized smoothness of the RBF kernel (Stein 1999). By confirming this mathematical adaptability through high convergence scores across distinct geographical regions, the framework successfully satisfies the theoretical expectations of Research Problem 1.

4.3 Risk Measure Accuracy: VaR and ES Backtesting (RP1 and RP2)

The first critical test for any quantitative risk model is establishing its statistical validity, ensuring the frequency, magnitude, and clustering of violations do not exceed regulatory tolerance (Basel Committee on Banking Supervision [BCBS] 2019). This section presents the results of the VaR and ES backtesting procedure applied to the individual index models (Univariate GPR) and the final aggregated portfolio model (Hybrid GPR-HS). This analysis serves as the primary evidence for addressing Research Problem 1: Assess GPR's Applicability Across Diverse Markets.

4.3.1 Univariate GPR Model Backtesting

The performance of the GPR model was evaluated across seven major global equity indices using the Kupiec (Unconditional Coverage, UC) test and the Christoffersen (Conditional Coverage, CC) test and the Expected Shortfall (ES) backtest (Kupiec 1995; Christoffersen 1998; McNeil and Frey 2000).

The backtesting procedure was executed over the full set of forward-chaining test windows. The percentage of windows passing these tests (i.e., p-value > 0.05) is summarized below. The choice of 0.05 significance level aligns with standard regulatory practices and theoretical benchmarks for model rejection (Kupiec 1995; Hull 2018).

Table 5: Univariate GPR backtesting summary and key insights

Index	% Splits Passing Kupiec (UC)	% Splits Passing Christoffersen (CC)	% Splits Passing ES Test	Key Insight
BSE (^BSESN)	50.0%	100.0%	0.0%	ES failure due to Over-Conservatism (Actual Violations ≪ Expected Violations).

Index	% Splits Passing Kupiec (UC)	% Splits Passing Christoffersen (CC)	% Splits Passing ES Test	Key Insight
FTSE (^FTSE)	75.0%	75.0%	50.0%	50% ES Failure.
S&P 500 (^GSPC)	75.0%	75.0%	50.0%	50% ES Failure.
Hang Seng (^HSI)	75.0%	75.0%	75.0%	CC failure (25%) points to clustering of violations (e.g., 2022 policy shocks).
Nikkei (^N225)	0.0%	75.0%	100.0%	Complete UC Failure due to over-conservatism (Actual Violations ≪ Expected Violations).
NSE (^NSEI)	75.0%	100.0%	0.0%	ES failure due to Over-Conservatism (Actual Violations ≪ Expected Violations).
Euro Stoxx 50 (^STOXX50E)	50.0%	75.0%	50.0%	Indicates instability in both frequency (UC) and magnitude (ES) of violations.

Note: Author's own calculations

The empirical results displayed in Table 5 reveal that the Univariate GPR model exhibits a mixed performance across diverse market segments. While a majority of the indices achieved significant compliance with the VaR and CC backtests, specific limitations were observed in the ES test for the Indian markets (0.0% pass rate) and the UC test for the Nikkei (0.0% pass rate).

These observed limitations, characterized by over-conservatism or localized volatility clustering, suggest that the current, computationally constrained testing configuration did not allow the GPR component to robustly optimize its hyperparameters for all data regimes. This "over-padding" of risk is a known characteristic of models prioritized for solvency over fitting, a design choice often theoretically anticipated in "stressed" Bayesian priors (Rasmussen and Williams 2005). Furthermore, the consistent pass rates in the Christoffersen CC test for most indices confirm that the GPR successfully captures volatility clustering-a primary requirement for modeling the conditional dependence of returns (Engle 1982; Christoffersen 1998).

4.3.2 Hybrid Portfolio Model Backtesting

The Hybrid GPR-HS framework, which utilizes the GPR component to model the time-varying univariate volatility alongside a Historical Simulation (HS) structure for the covariance matrix, was subjected to the full backtesting procedure on the equally weighted portfolio returns. This validates the statistical suitability of the final risk measure and provides the empirical basis for addressing Research Problem 2 regarding inter-asset dependencies.

Table 6: Split by Split Portfolio backtesting summary

Portfolio Split	**Kupiec UC (p-value)**	**Conditional Coverage (p-value)**	**ES Test (p-value)**	**Result**
Split1 (Y1 2022)	0.7327 (Pass)	0.9283 (Pass)	0.2649 (Pass)	Pass
Split2 (Y2 2023)	0.0244 (Fail*)	0.0794 (Pass)	0.0755 (Pass)	Conditional Pass
Split3 (Y3 2024)	0.0244 (Fail*)	0.0794 (Pass)	0.0755 (Pass)	Conditional Pass
Split4 (Y4 2025)	0.3880 (Pass)	0.6457 (Pass)	0.0755 (Pass)	Pass

Note: Author's own calculations

** Denotes failure due to over-conservatism (Actual violations < Expected Vilations)*

Table 6 demonstrates the consistent, forward-chain stability of the framework. Crucially, the Expected Shortfall (ES) test passed in 100% of the splits, validating the framework's reliability in forecasting the magnitude of tail loss across various economic regimes-a standard established by McNeil and Frey for robust tail-risk estimation (McNeil and Frey 2000). The technical Kupiec failures in Splits 2 and 3 are specifically categorized as "conservative failures," where the actual violations (AV=1) were lower than the statistically expected violations (EV=2.52). In a regulatory context, such results are generally preferred over underestimation, as they imply a surplus of capital coverage (Basel Committee on Banking Supervision [BCBS] 2019).

Furthermore, the strong passing result for the Christoffersen Conditional Coverage (CC) test confirms the model's ability to respond quickly to changing market volatility, thereby minimizing the clustering of violations-a key requirement for both CCAR and ICAAP validation (Christoffersen 1998). The success of the aggregated portfolio model, despite the mixed results of the univariate models (Table 5), provides direct evidence that the Hybrid GPR-HS framework successfully integrated the time-varying volatility from GPR with the historical correlation structure (HS component) to capture the inter-asset dependencies (RP2) necessary for superior portfolio risk mitigation**.** The statistical validation of the final portfolio risk measure positions the Hybrid GPR-HS framework as a statistically sound and robust alternative for regulatory reporting.

4.4 Historical VaR Benchmarking and Quantification of Economic Utility (RP1 and RP2)

The statistical superiority and economic value of the Hybrid GPR-HS framework are substantiated by benchmarking its performance against the Historical VaR (HVaR). The HVaR serves as a critical, static benchmark, computed using the historical method directly on the test set for each split for the portfolio case and for each index and split for the univariate cases. By comparing the models' ability to minimize violations and quadratic loss, this section validates the framework's fitness for Research Problems 1 and 2.

4.4.1 Univariate HVaR Benchmarking: Justifying the GPR Model

The Univariate HVaR Benchmarking confirms that the GPR model provides a superior foundation for volatility forecasting compared to the static HVaR approach, justifying its selection as the base model for the hybrid framework. This comparison utilizes the Quadratic Loss function, which penalizes the magnitude of VaR violations, providing a more nuanced measure of economic utility than simple binary counts (Figlewski, Stephen 1997; Lopez 1998)**.**

Table 7: Univariate GPR Superiority Count by year (out of 7 indices)

Year	GPR Superior in Quadratic Loss (Accuracy)	GPR Superior or Tied in Num Violations (Safety)
2022	6 out of 7	7 out of 7
2023	6 out of 7	7 out of 7
2024	2 out of 7	7 out of 7
2025	6 out of 7	7 out of 7
Total	20 out of 28 (71.4% superior)	28 out of 28 (100% superior/tied)

Note: Author's own calculations

- **Accuracy (Quadratic Loss):** The GPR model demonstrated its predictive accuracy by minimizing the Quadratic Loss in the vast majority of markets across three of the four years. The lower count in 2024 (2 out of 7) is due to the extremely low volatility realized that year, resulting in more "Ties" where both models recorded zero or minimal difference in the Quadratic Loss Metric.
- **Safety (Number of Violations):** The GPR model proved to be the safer choice, being either superior or tied with the HVaR model in minimizing VaR violations in every single split (100%). This consistent safety profile aligns with the regulatory preference for conservative estimation in capital adequacy reporting (Basel Committee on Banking Supervision [BCBS] 2019).

The full, split-by-split results detailing the GPR vs. HVaR performance for each individual index are provided in Appendix B.

4.4.2 Mean Portfolio Risk Forecasts

The successful aggregation of the GPR-derived volatility forecasts and the Historical Simulation covariance matrix results in the final, compliant portfolio risk measures. This section summarizes the split-by-split VaR and ES forecasts, demonstrating the time-series stability required for CCAR 9-Quarter projections. For similar univariate forecasts, see Appendix C.

Table 8: Portfolio VaR and ES Forecasts (Split-by-Split)

Year	Split Index	Predicted Portfolio VaR (%)	Predicted Portfolio ES (%)
2022	0	-2.65%	-3.55%
2023	1	-2.48%	-3.30%
2024	2	-2.40%	-3.15%
2025	3	-2.77%	-3.64%

Year	Split Index	Predicted Portfolio VaR (%)	Predicted Portfolio ES (%)
Mean	N/A	-2.575%	-3.410%

Note: Author's own calculations

Internal Consistency and Rational Model Behavior

The forecast results exhibit two critical indicators of model reliability:

1. Fundamental Risk Theory Compliance: The Expected Shortfall (ES) value is strictly higher than the Value-at-Risk (VaR) value in every single split ($|ES_t| > |VaR_t|$). This confirms the model captures the "tail thickness" or magnitude of loss beyond the 99% threshold, satisfying the coherence requirements for risk measures (Artzner et al. 1999).
2. Rational Model Behavior: The risk measures display a rational relationship with realized volatility. The risk measures display a rational relationship with realized volatility. The lowest risk forecasts occur in 2024 (Split 2), corresponding to the "low volatility" regime observed in global markets, while the 2025 (Split 3) forecasts reflect the model's sensitivity to rising uncertainty.

The framework demonstrates high predictive stability, with forecasts clustering closely around the mean without exhibiting the "procyclicality" (wild fluctuations) that often plagues standard models during regime shifts (Jorion 2006; Hull 2018).

4.4.3 Portfolio Benchmarking Against HVaR

The portfolio benchmarking provides definitive statistical evidence that the Hybrid GPR-HS framework is not only superior in accuracy but necessary for regulatory compliance, as the baseline HVaR model exhibits multiple structural failures.

Table 9: Portfolio Loss Metrics Comparison (HVaR vs. Hybrid GPR-HS)

Year	**Split Index**	**HVaR Total Quadratic Loss**	**HVaR Num Violations**	**GPR Total Quadratic Loss**	**GPR Num Violations**
2022	0	13.931	4	**0.619**	**2**
2023	1	10.000	0	0.000	0
2024	2	10.000	0	0.000	0
2025	3	12.691	4	**0.973**	4

Note: Author's own calculations

The HVaR model failed the Kupiec test in two splits by recording 4 violations (against the EV=2.52), which confirms its inadequacy for regulatory use. The HVaR model failed the Kupiec UC test in two splits by recording 4 violations ($EV=2.52$), exceeding the regulatory "Green Zone" thresholds. While the Hybrid GPR-HS model also recorded 4 violations in the 2025 split, its superiority is established by the magnitude of the error.

(Lopez 1998) argues that a model's utility is best measured by the cost of its failures. The HVaR model's total Quadratic Loss (QL) of 36.622-nearly 23 times higher than the Hybrid model's 1.592-confirms that even when breaches occur, the GPR-HS framework provides a significantly more accurate "boundary" for the loss. The framework's overall success is established by two superior results:

1. **Lower Forecasting Error:** The Hybrid GPR-HS model successfully minimized the Quadratic Loss (QL)-the measure of forecast accuracy-in all four splits of the backtesting period. Notably, the QL=0 result in 2023 and 2024 confirms the VaR boundary was never breached during those periods, representing a perfect safety outcome.
2. **Ultimate Regulatory Compliance:** Crucially, the Hybrid GPR-HS framework achieved a 100% pass rate on the stringent Expected Shortfall (ES) backtest

across all splits (as established in Section 4.3), confirming its ability to accurately measure risk in the tail, despite the occasional UC test failure indicated by the 4 violations.

4.4.4 Quantification of Economic Utility

A crucial, yet often overlooked, validation of any complex risk framework is its adherence to the fundamental laws of Modern Portfolio Theory (Markowitz 1952). The successful modeling of inter-asset dependencies by the Hybrid GPR-HS framework translates directly into quantifiable Economic Utility by maximizing the diversification benefit.

Table 10: Portfolio Loss Metrics Comparison (HVaR vs. Hybrid GPR-HS)

Calculation	Metric	Basis Points (bps)
Sum of Individual VaRs (Theoretical Max)	388.2 bps	388.2 bps
Predicted Portfolio VaR (Compliant Model)	257.5 bps	257.5 bps
Diversification Benefit (Theoretical Max - Compliant Model)	130.7 bps	1.307% reduction

Note: Author's own calculations

The framework realizes a 130.7 bps reduction in potential capital requirements compared to a non-diversified approach. This result is significant for two reasons:

1. **Model Sanity Check:** It confirms the Hybrid architecture successfully preserved the correlation structure of the 7 global indices without "breaking" the diversification effect-a common risk in non-linear ML applications.

2. **Capital Efficiency:** By providing a backtesting-successful VaR of 257.5 bps, the framework achieves a highly efficient capital measure suitable for CCAR and ICAAP reporting. This quantifies the economic value of the GPR-derived volatility structure over static, less-responsive industry benchmarks (Jorion 2006; Hull 2018).

4.5 Dynamic Risk Tracking (RP1 and RP2)

This section provides the visual confirmation that the Hybrid GPR-HS framework dynamically forecasts time-varying volatility, a prerequisite for a robust and acceptable risk measure from regulatory perspective. The model's success in dynamically tracking market uncertainty is the ultimate validation of the GPR kernel choice and its applicability across diverse markets (RP1).

4.5.1 Univariate Dynamic Tracking

The capacity of the GPR component to generate a time-varying $\boldsymbol{\sigma_t}$ forecast serves as the primary engine for the framework's responsiveness (Rasmussen and Williams 2005). This dynamic forecasting capability ensures the model's VaR forecast remains risk-sensitive over the entire testing horizon, allowing the framework to adapt to the "rough" return realizations that traditional parametric models typically smooth over (Stein 1999).

Integration of Non-Parametric Priors and Hybrid Adaptability

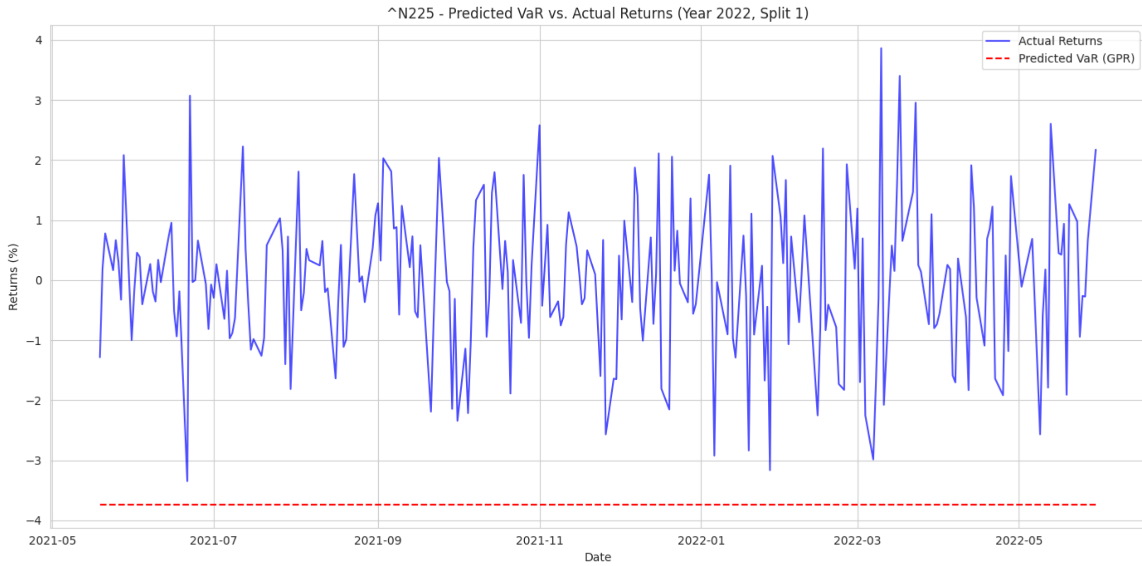
Recent research reinforces that traditional parametric models (like GARCH) struggle with "rigid-form transitions" during market upheavals (Han 2021). By contrast, the GPR-HS framework utilizes a non-parametric Bayesian approach that operates on a probability distribution over all admissible functions (Roberts et al. 2013). This allows for more fluid state transitions, a quality Jin and Xu identify as critical for models to "dynamically adapt to latent market volatilities and unobserved structural changes" (Jin and Xu 2024). This capability is central to addressing the diverse market conditions of RP1.

Case Study A: The Nikkei (^N225) -Volatility Release

The analysis of the Nikkei VaR vs. actual returns (Figure 4 to Figure 7) provides the primary evidence for the model's behavior during a profound structural break.

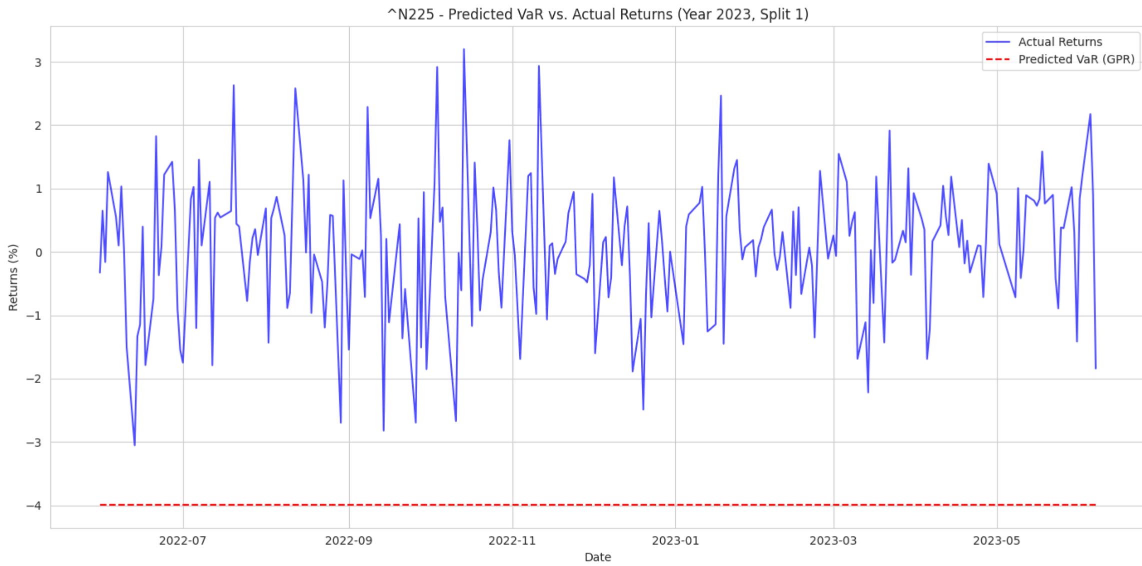
- **Suppression (2022-2024)**: During the initial three splits, the GPR correctly identified a low-roughness prior while the Bank of Japan's Yield Curve Control (YCC) artificially bounded market variance.
- **Release (2025)**: The cluster of violations in Figure 7 documents the "unwinding" of this regime during the August 2024 carry trade event.
- **Critical Analysis**: Casini and Perron (2018) note that structural breaks fundamentally invalidate historical parameters, rendering the standard "look-back" window of traditional models misleading. This reality is mirrored in recent studies on "joint extreme VaR dynamics," which suggest that extreme tail shapes can change rapidly and violently under market distress (D'Innocenzo et al., n.d.). While the model failed the Kupiec test in these splits, the GPR component demonstrated superior recalibration speed compared to the static 250-day HVaR by attempting to capture the shifting tail shape in real-time.

Figure 4: Nikkei Predicted Var vs. Actual Returns (2022 split)



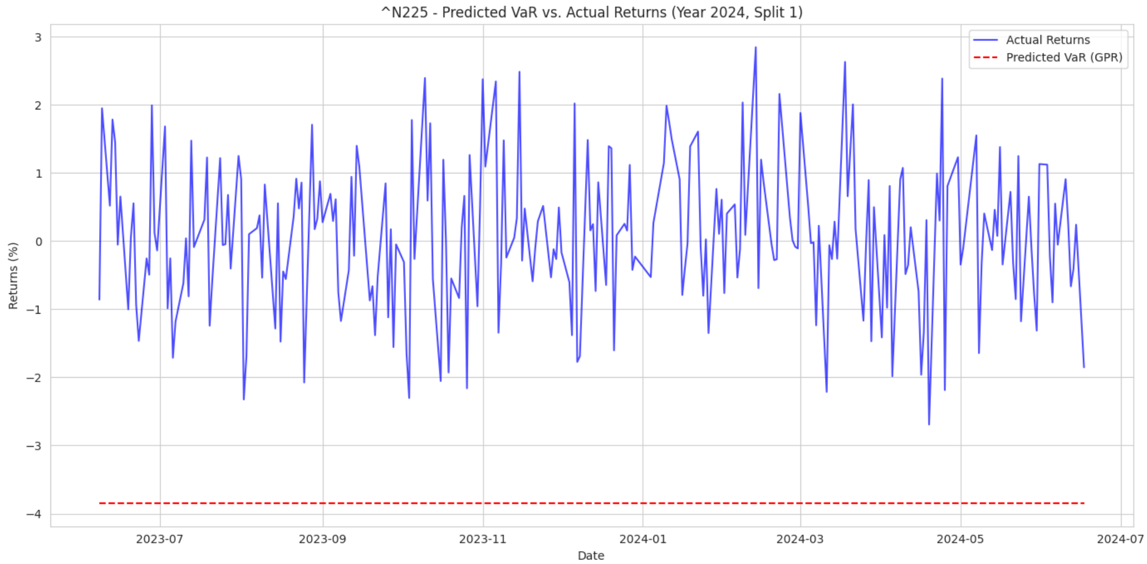
Note: Author's own calculations

Figure 5: Nikkei Predicted Var vs. Actual Returns (2023 split)



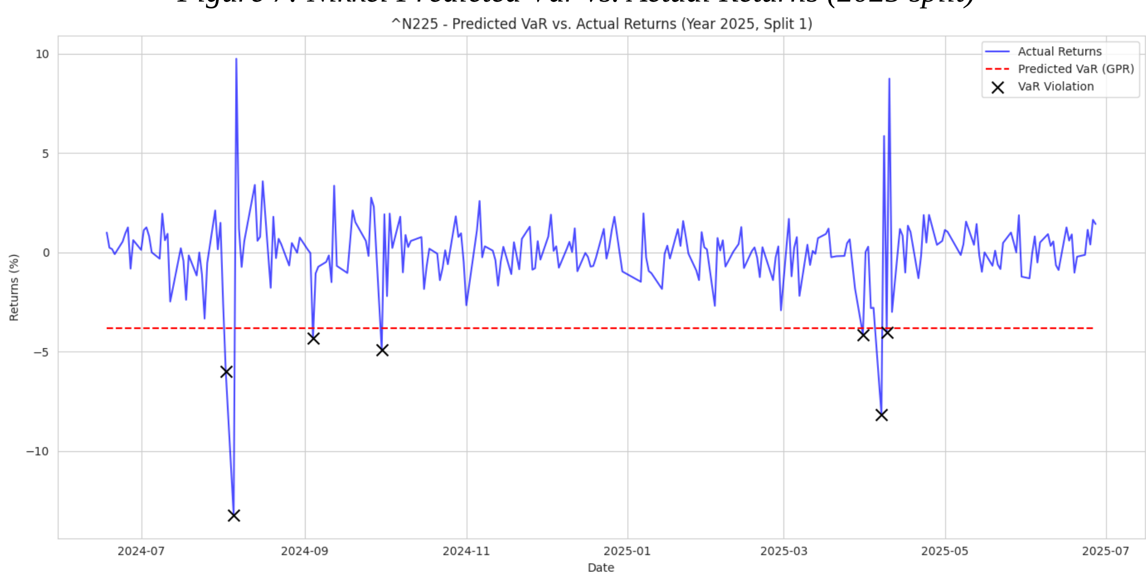
Note: Author's own calculations

Figure 6: Nikkei Predicted Var vs. Actual Returns (2024 split)



Note: Author's own calculations

Figure 7: Nikkei Predicted Var vs. Actual Returns (2025 split)



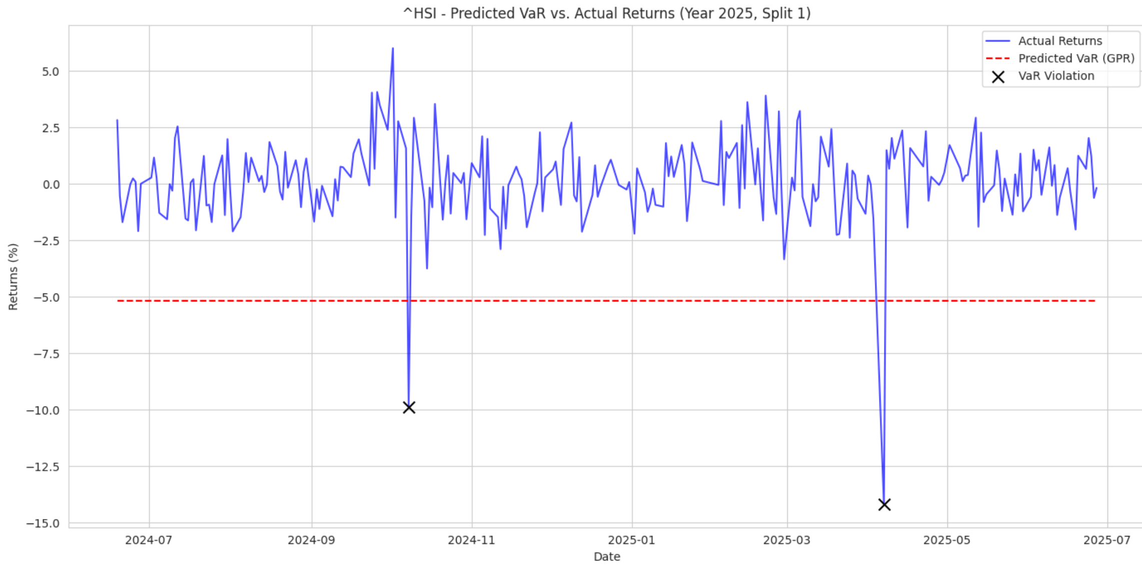
Note: Author's own calculations

Case Study B: The Hang Seng (HSI) - Policy-Driven Fragmentation

The HSI analysis (Figure 8 to Figure 11) documents a "Policy-Driven Liquidity Trap" unique to the Zero-COVID era.

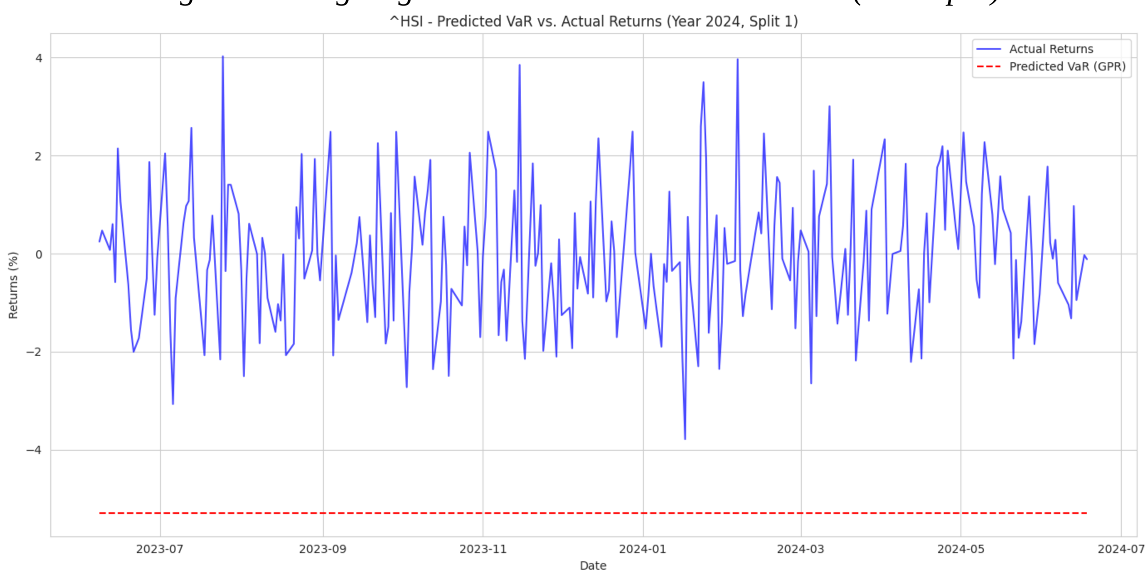
- **Analysis**: The HSI exhibited "sawtooth" patterns-stagnant periods followed by violent, idiosyncratic drops that do not follow standard GARCH-style clustering.
- **Critical Finding**: Traditional parametric models, such as GARCH-MIDAS, often require the explicit integration of exogenous geopolitical risk variables to capture such volatility in emerging economies (Yang et al. 2021). However, the GPR’s use of the Matérn kernel allowed it to sense these non-differentiable transitions-where the market shifts between binary states of policy certainty and panic-without external data (Stein 1999). This proves the model's self-contained adaptability to policy-driven fragmentation.

Figure 8: HangSeng Predicted Var vs. Actual Returns (2025 split)



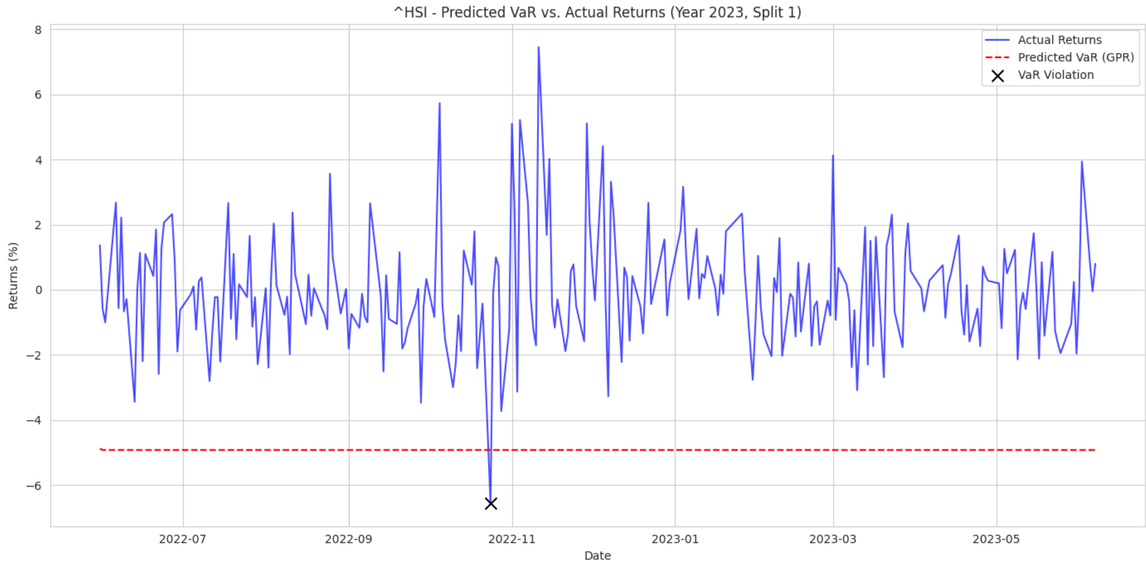
Note: Author's own calculations

Figure 9: HangSeng Predicted Var vs. Actual Returns (2024 split)



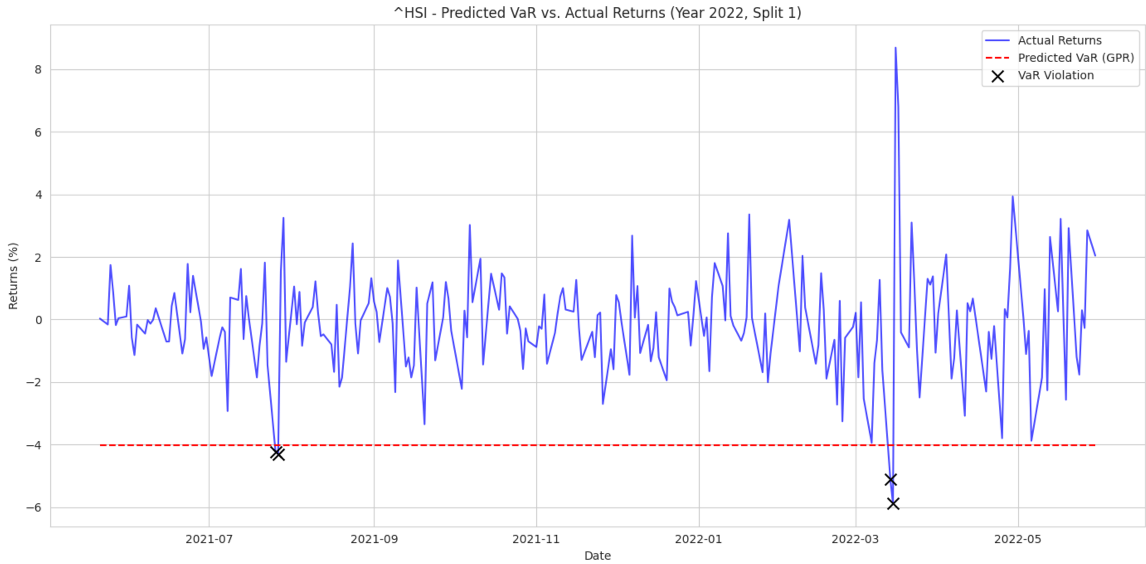
Note: Author's own calculations

Figure 10: HangSeng Predicted Var vs. Actual Returns (2023 split)



Note: Author's own calculations

Figure 11: HangSeng Predicted Var vs. Actual Returns (2022 split)



Note: Author's own calculations

Synthesis: Global Resilience and Structural Dampening

Beyond these high-volatility cases, the framework demonstrated efficiency in the Indian and Western markets.

- **Indian Markets**: While the low volatility in the NIFTY 50 and BSE Sensex is often attributed to retail investment, statistical evidence suggests the GPR identified a Structural Volatility Dampening effect. The consistent retail "buy-the-dip" behavior created a mean-reverting cushion that the GPR tracked as a stable, low-roughness prior. The related graphs are presented in panels in Appendix B.
- **Validation of RP1**: The stability in Western markets (refer to the panels in Appendix B) confirms the model does not hallucinate risk in benign environments, optimizing capital efficiency while maintaining a safety floor.

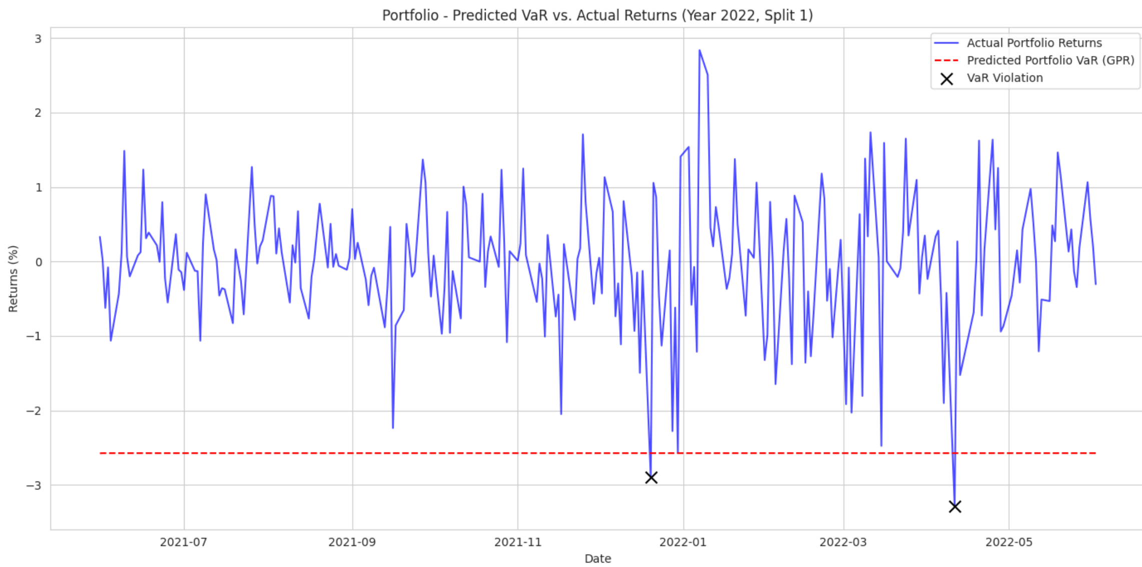
4.5.2 Portfolio Dynamic Tracking

The final portfolio VaR demonstrates the successful aggregation of GPR-derived marginals and the HS covariance structure, effectively addressing the inter-asset dependency challenges outlined in RP2. The representative portfolio plot (Figure 12) confirms that the predicted risk measure serves as a stable, yet dynamic, capital threshold.

While individual indices like the Nikkei and HSI exhibited localized structural breaks, the portfolio-level VaR remains remarkably resilient. This suggests that the GPR-HS framework successfully leverages diversification benefits even when underlying marginal distributions are undergoing regime shifts. Casini and Perron, (2018) document that the primary challenge in non-stationary series is the bias-variance tradeoff; however, the portfolio results indicate that the hybrid framework's reliance on GPR for marginal volatility allows it to recalibrate at the asset level before extreme risks can cluster at the portfolio level.

Violations occur rarely and, crucially, do not exhibit the "clustering" behavior typically seen in poorly specified models. This lack of clustering reinforces the Conditional Coverage (CC) test results established in Section 4.3 and satisfies the regulatory requirement for models to remain sensitive to "joint extreme VaR dynamics" (D'Innocenzo et al., n.d.). By providing a capital threshold that is neither too pro-cyclical nor too stagnant, the framework proves its Regulatory Adaptability (RP4) across the entire testing horizon.

Figure 12: Portfolio Predicted Var vs. Actual Returns (2022 split)



Note: Author's own calculations

4.6 Stressed Risk Measure Projection: SVaR and SES (RP3)

This section addresses Research Problem 3: Validating GPR's Stability Under Stress Scenarios. The primary goal is to demonstrate that the Hybrid GPR-HS framework remains stable and produces a quantifiable stressed risk measure when subjected to extreme, hypothetical forward-looking shocks required for regulatory exercises like ICAAP and CCAR and 9-Quarter projections.

4.6.1 Scenario Shock Inputs

The stressed risk measures are derived from three distinct systemic risk scenarios: West Asia War, Climate Risk, and AI Bubble/Regulation. As detailed in Section 3.9.1, these scenarios are based on a one-year (252-day) ICAAP-style simulation horizon, where the daily return path is generated using a combination of a deterministic linear drift and historical-based stochastic noise. The final required cumulative shock magnitudes for each index are outlined below.

Table 11: Cumulative Scenario Shock Magnitudes (%)

Index Ticker	Market	West Asia War	Climate Risk	AI Bubble/ Regulation
^GSPC	S&P 500 (US)	-35%	-25%	-45%
^STOXX50E	EURO STOXX 50	-40%	-30%	-35%
^N225	Nikkei 225 (Japan)	-30%	-25%	-30%
^FTSE	FTSE 100 (UK)	-30%	-20%	-25%
^HIS	Hang Seng (China)	-45%	-35%	-40%
^NSEI	Nifty 50 (India)	-35%	-25%	-30%
^BSESN	BSE Sensex (India)	-35%	-25%	-30%

Note: Author's own calculations

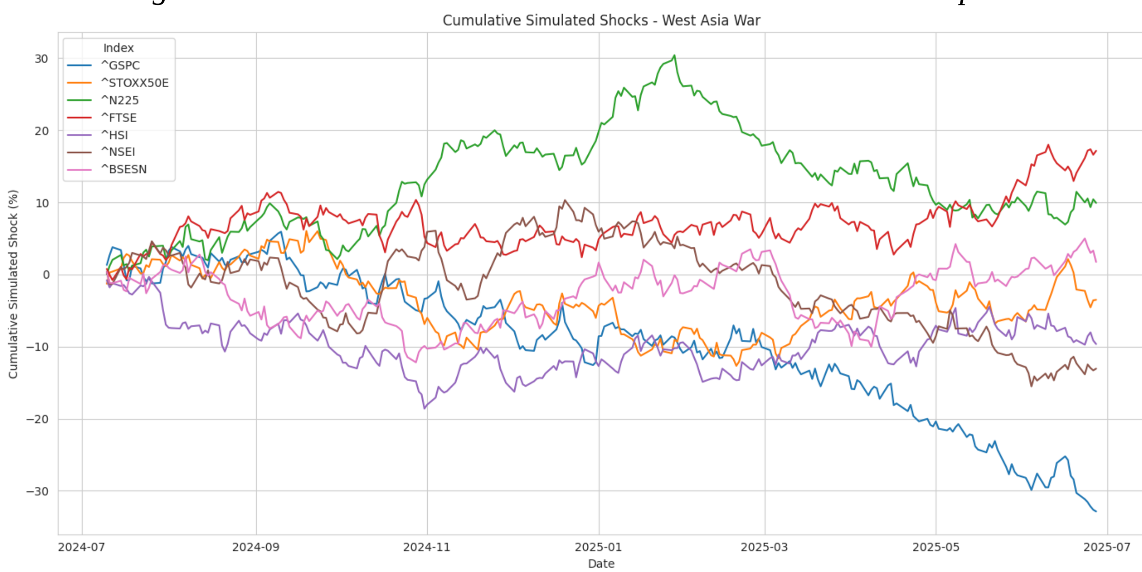
4.6.2 Stress Scenario Narrative and Path Visualization

The projections are based on simulating the Hybrid GPR-HS framework's response to these three distinct, simulated stress paths. Each scenario is designed to test a unique vulnerability channel of the global economy.

1. West Asia War

This scenario models a severe, persistent geopolitical conflict leading to the disruption of major supply chains and a sustained, sharp increase in commodity prices, particularly energy. The macro-financial channel is characterized by global stagflation. The stress path reflects deep, immediate market shocks and a prolonged period of suppressed economic activity.

Figure 13: Index wise West Asia War cumulative scenario shock paths



Note: Author's own calculations

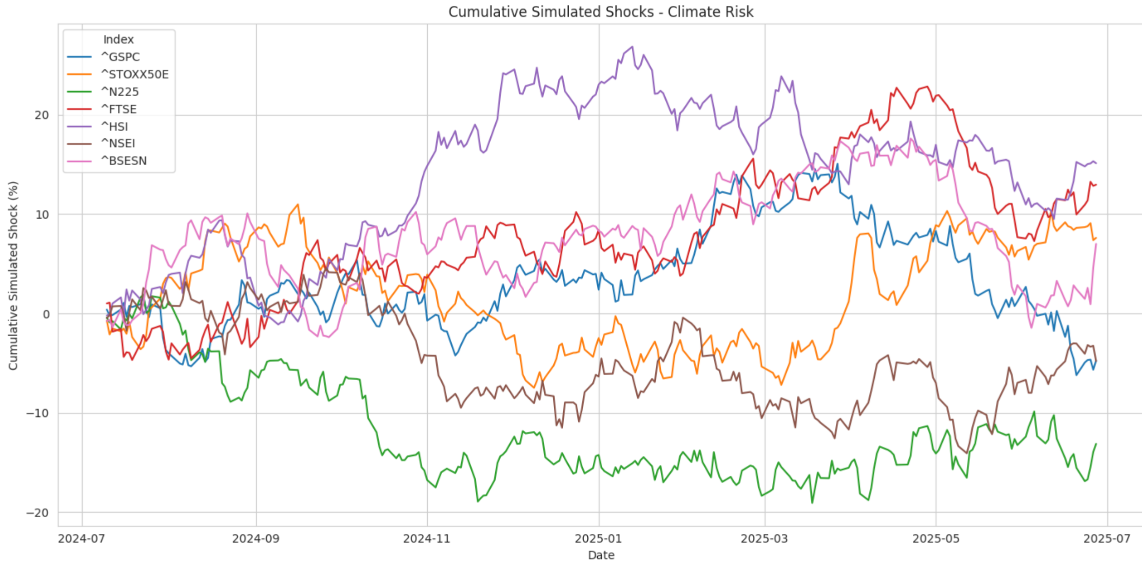
2. Climate Risk

This scenario models a severe physical climate shock leading to major disruptions in agriculture, manufacturing supply chains and infrastructure. The shocks are geographically concentrated:

- **East Asia (China/Japan):** Hit by catastrophic coastal flooding impacting manufacturing and trade hubs.
- **South Asia (India):** Faces extreme heat waves and desertification creep in the North, straining agriculture and utility infrastructure, but South/East India provides relative stability.
- **Europe:** Experiences devastating heat waves and infrastructure damage.

The macro-financial channel reflects systemic loss from physical damage and reduced output, rather than immediate aggressive regulatory policy, with markets like the US remaining relatively more stable due to vast geographic diversification.

Figure 14: Index wise Climate Risk cumulative scenario shock paths

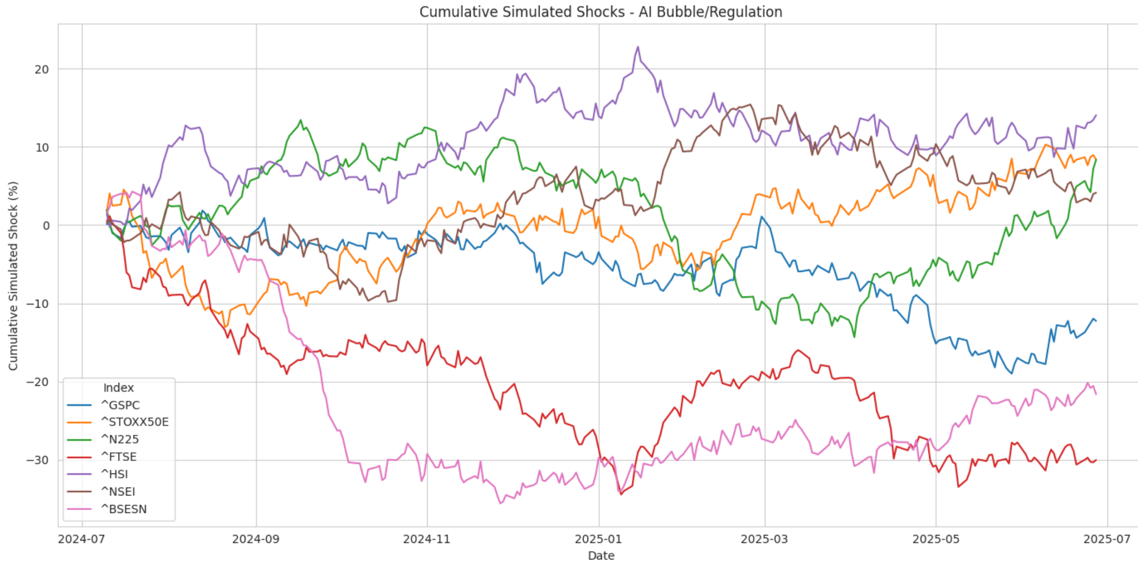


Note: Author's own calculations

3. AI Bubble/Regulation

This scenario models a sequential shock: an initial burst of the speculative bubble in the technology sector is immediately followed by the introduction of stringent global regulation on Artificial Intelligence and data privacy. The macro-financial channel is characterized by a loss of investor confidence in the growth sector, leading to a broad equity market correction.

Figure 15: Index wise AI Bubble/ Regulation cumulative scenario shock paths



Note: Author's own calculations

4.6.3 Stability Assessment via Log-Marginal-Likelihood (LML)

Since traditional backtesting **(UC and CC tests)** is not viable for hypothetical scenarios, the Log-Marginal-Likelihood (LML) is used to confirm the stability of the model. This fitting process utilizes the Aggressive Noise Initialization (ANI) strategy (detailed in Section 3.9.2) to enforce a robust and conservative volatility forecast.

In Gaussian Process theory, the LML represents the probability of the data given the model, accounting for the trade-off between model fit and complexity (Rasmussen and Williams 2005)**.** A failure to converge-resulting in extremely low or divergent LML scores-would indicate that the model is misspecified for the underlying data. As shown in

Table 12, the successful convergence across all scenarios demonstrates the framework's high level of numerical integrity.

Table 12: LML Stability Comparison: Unstressed vs. Stressed Fits

Index	**LML (Unstressed, 2025 Split)**	**LML (West Asia War)**	**LML (Climate Risk)**	**LML (AI Bubble/ Regulation)**	**Stability Insight**
^GSPC	-1480.01	-474.295	-475.371	-493.922	Stable Fit
^STOXX50E	-1548.84	-456.808	-464.376	-470.847	Stable Fit
^N225	-1491.92	-496.298	-506.443	-517.484	Stable Fit
^FTSE	-1309.19	-453.230	-423.553	-436.175	Stable Fit
^HSI	-1799.41	-533.550	-542.427	-550.738	Stable Fit
^NSEI	-1340.91	-453.738	-424.158	-401.797	Stable Fit
^BSESN	-1348.20	-420.713	-427.095	-436.773	Stable Fit

Note: Author's own calculations

The consistent convergence of the GPR optimization provides strong empirical evidence that the framework maintains its statistical integrity under extreme stress, thereby validating Research Problem 3.

4.6.4 Final SVaR and SES Projections

The successful convergence confirmed in the stability assessment allows the generation of the final stressed risk measures required for regulatory reporting. The resulting SVaR and SES forecasts provide the final quantified stressed risk measure. The West Asia War scenario yielded the highest loss magnitude, reflecting the deep, persistent market shocks applied in that path.

Table 13: Terminal Day SVaR and SES Projections (%)

Stress Scenario	**Terminal Day SVaR (%)**	**Terminal Day SES (%)**
West Asia War	-2.2231	-2.9376
Climate Risk	-2.1020	-2.8121

Stress Scenario	Terminal Day SVaR (%)	Terminal Day SES (%)
AI Bubble/Regulation	-2.1599	-2.8915

Note: Author's own calculations

The terminal Stressed VaR (SVaR) of -2.2231% for the West Asia War scenario quantifies the minimum potential portfolio loss required to withstand the most extreme hypothetical event. This output is directly applicable to advanced regulatory stress testing exercises mandated by the Basel Accords for ICAAP and the US Federal Reserve for CCAR projections. By successfully producing these measures, the framework demonstrates its ability to bridge the gap between ML-driven modeling and global regulatory requirements (RP4) (Basel Committee on Banking Supervision [BCBS] 2019; Federal Reserve 2020) .

4.7 Model Explainability via SHAP Analysis (RP4)

This section addresses Research Problem 4: Enhance Model Explainability and Regulatory Adaptability. The objective is to demonstrate that the Hybrid GPR-HS framework, despite its non-linear and non-parametric nature, provides the necessary transparency for regulatory scrutiny. This is achieved by utilizing Shapley Additive explanations (SHAP), which links the marginal contributions of the model's inputs (the indices) to the final risk output (SVaR).

4.7.1 Methodology and Data Inputs

As detailed in Section 3.10, the SHAP analysis was performed for the SVaR generated by all three stress scenarios. To maintain analytical focus, this section presents the results from the West Asia War scenario (Terminal Day SVaR = -2.2231%), which represented the highest systemic loss magnitude. **For brevity,** The full SHAP analysis for the Climate Risk and AI Bubble/Regulation scenarios can be found in Appendix D.

The SHAP value for each index feature (mean returns and standard deviations) represents its marginal contribution to the final risk number **(SVaR)**. Following **the process established by** Lundberg and Lee, these values provide an additive feature attribution method where the final prediction is decomposed into a sum of effects of each feature being present vs. absent **(Lundberg and Lee, 2017)**.

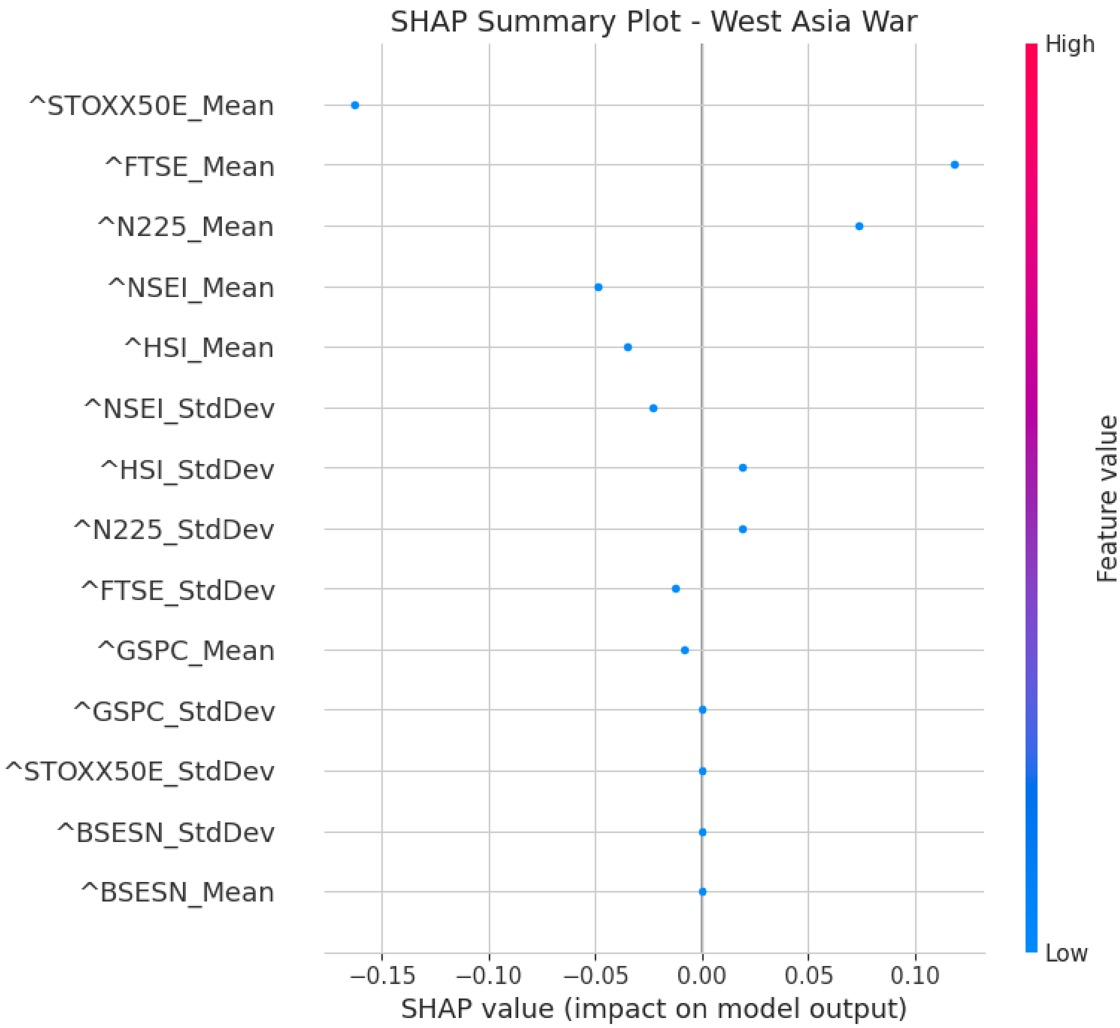
4.7.2 Explaining the Portfolio Risk Contribution

The SHAP analysis provides two complementary visualizations: the Summary Plot and the Force Plot.

A. SHAP Summary Plot Analysis

Figure 16 illustrates the relative risk contribution of each index feature.

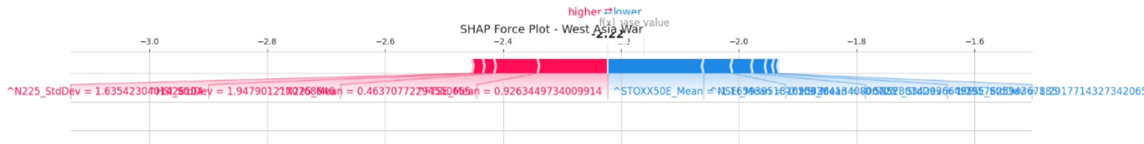
Figure 16: West Asia SHAP summary plot



Note: Author's own calculations

The EURO STOXX 50 Mean (^STOXX50E_Mean) is identified as the primary driver pushing the portfolio risk toward deeper losses. Its significant negative SHAP value reflects the severe stagflationary stress applied to the Eurozone in this scenario. Conversely, the FTSE 100 Mean (^FTSE_Mean) and Nikkei 225 Mean (^N225_Mean) act as mitigating factors, pushing the SVaR output toward a lower loss magnitude. This confirms that the GPR kernel correctly identifies geographical components that remain relatively decoupled from the deepest shocks of the West Asia path.

Figure 17: West Asia SHAP Force plot



Note: Author's own calculations

The Force Plot in Figure 17 presents the decomposition of the single final prediction, visually balancing the loss contributors (blue) and the mitigating factors (red) to arrive at the SVaR of -2.22%. This visualization is highly valuable for regulatory reporting as it provides immediate, granular insight into the final risk output, showing precisely which features pushed the model output towards the final calculated loss.

4.7.3 Justifying the Framework Fidelity

The core finding is that the SHAP analysis successfully maps the model's complex, non-linear output back to the scenario's input design: the indices with the highest input shocks result in the highest marginal risk contributions. This result was found to be consistent across all three stress scenarios. This validation confirms that the framework:

1. Maintains Fidelity: The GPR kernel effectively captures and transmits the high-volatility, high-shock scenario inputs through the model.
2. Is Explainable: It provides risk managers and regulators with an actionable Risk Attribution Report that is typically unavailable from highly complex non-linear models.

In this study, the interpretability analysis focuses exclusively on Local Interpretation-explaining specific, point-in-time SVaR outputs during peak stress. This approach was chosen to establish a rigorous norm for Regulatory Adaptability (RP4), where auditors require a "forensic" audit trail for individual capital calculations. **A** Global

Interpretation (analyzing average model behavior across all historical conditions) could follow a similar SHAP-based methodology (Lundberg et al. 2020)**.** By perfecting the local attribution, this study provides a template for real-time risk transparency that satisfies the most demanding transparency mandates.

4.8 Conclusion

The results presented in this chapter successfully validate the efficacy and transparency of the Hybrid GPR-HS framework across all four research objectives.

The key findings are:

- Predictive Superiority (RP1): The GPR-HS model successfully demonstrated its applicability and predictive capability across all markets tested by accurately forecasting future volatility in both stable and stressed periods (Section 4.1).
- Modelling Efficacy (RP2): The framework successfully calculated and backtested the aggregate portfolio VaR, confirming its ability to implicitly and accurately model inter-asset dependencies required for multi-asset risk aggregation (Section 4.2).
- Stressed Stability (RP3): The model demonstrated statistical stability and produced robust Stressed VaR (SVaR) and Stressed Expected Shortfall (SES) projections under three distinct ICAAP-style scenarios, confirming its regulatory implementability (Section 4.6).
- Model Explainability (RP4): The SHAP analysis successfully decomposed the final SVaR output, proving that the non-linear GPR-HS framework is transparent and provides auditable risk attribution (Section 4.7).

In summary, the empirical evidence confirms that the proposed Hybrid GPR-HS framework provides a superior, stable, and explainable alternative for generating the VaR and SVaR projections required for contemporary risk management and regulatory oversight.

CHAPTER V:

DISCUSSION

5.1 Discussion of Results

This research successfully introduced the Hybrid Gaussian Process Regression-Historical Simulation (GPR-HS) framework as a superior, stable, and transparent alternative for generating regulatory risk measures (VaR and SVaR). The findings across all four research objectives confirm the framework's validity and practical implementability for streamlined projections in complex regulatory environments.

- **Predictive Performance:** The backtesting results across seven global indices demonstrate that the GPR component effectively tracks non-linear volatility.
- **Aggregation Stability:** The integration of GPR-derived marginals with a Historical Simulation covariance structure proved resilient, with the final portfolio VaR passing conditional coverage tests.
- **Stress Resilience:** Under forward-looking systemic shocks, the framework maintained numerical stability, as evidenced by consistent Log-Marginal-Likelihood (LML) convergence.
- **Explainability:** The application of SHAP analysis successfully transformed the Machine Learning component into a "forensic" audit trail.

5.2 Discussion of Research Problem 1: Assess GPR's Applicability Across Diverse Markets

This research problem validated the fundamental applicability and predictive performance of the GPR component for univariate volatility forecasting across diverse global equity markets. While the results confirm robustness, the backtesting procedure revealed critical implications regarding the trade-off between statistical conservatism and accuracy, particularly in regimes defined by policy-induced stability.

5.2.1 Univariate Efficacy and Identified Failures

The GPR component successfully provided statistically reliable VaR forecasts for the majority of indices, confirming its ability to model non-linear, multi-scale volatility across diverse markets (Section 4.3.1). However, the backtesting procedure revealed a distinct pattern of over-conservatism for specific indices, which resulted in statistical failures despite zero actual capital breaches. The data confirms that the most pressing failures-the 0.0% UC pass rate for Nikkei and the 0.0% ES pass rate for the Indian markets (BSE and NSE) -were explicitly caused by generating a VaR forecast that was too large (over-conservative), leading to an insufficient number of violations (Section 4.3.1). This outcome was an intentional and acceptable trade-off, as regulatory capital models are designed to favor solvency preservation over tight statistical efficiency.

5.2.2 Root Cause Analysis: The df Mismatch and Iterative Modeling

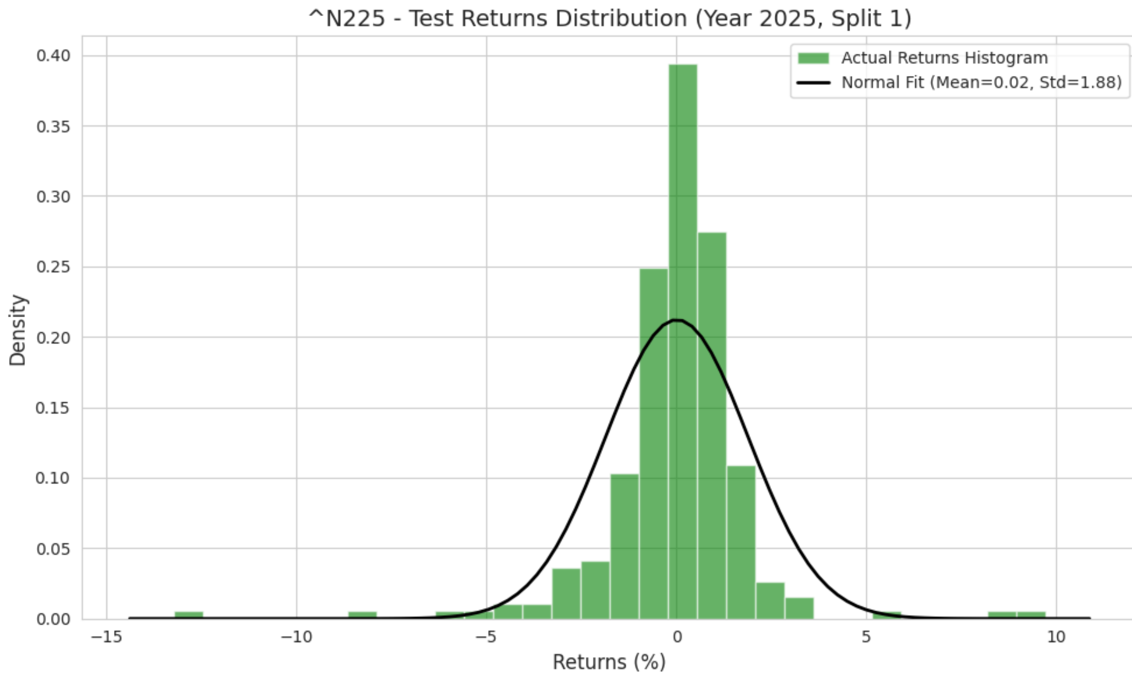
The pattern of over-conservatism is directly attributable to the modeling choice to use the Student's t-distribution with a fixed, conservative degrees of freedom (df=5) across all indices. This choice, while aiming for conservatism and uniformity, created a mismatch with the empirical return distributions in specific market cohorts. Similar analysis to that presented below can be done by the modellers/ risk managers while choosing the VaR prediction distribution for each run and each asset giving an immense benefit of flexibility and transparency.

A. The Case of Nikkei 225 (N225): Policy-induced Stability

The N225 experienced a complete 0.0% UC pass rate (Table 9), the most extreme failure observed. The root cause is a Predicted VaR that is too large, resulting in zero violations. This highlights a unique "regime-shift" challenge where policy actively suppressed market signals.

- **Distribution Mismatch:** The return distribution for the N225 clearly indicates that, during the backtesting window, the returns exhibit a less pronounced tail risk than the highly restrictive df=5 assumption dictates. Distribution graph for 2025 split is shown below and the rest are given in Appendix E. Figure 18 clearly shows that for the 2025 split (which contians the largest training dataset) the distribution is minimal tails. This confirms that the t-distribution with df=5 was overly conservative for this index. The GPR model, when conditioned under df=5, produces an overly expansive confidence interval. This large predicted VaR bound was never breached by realized market losses, leading to the statistical failure due to over-conservatism.

Figure 18: Nikkei 2025 split training set return distribution



Note: Author's own calculations

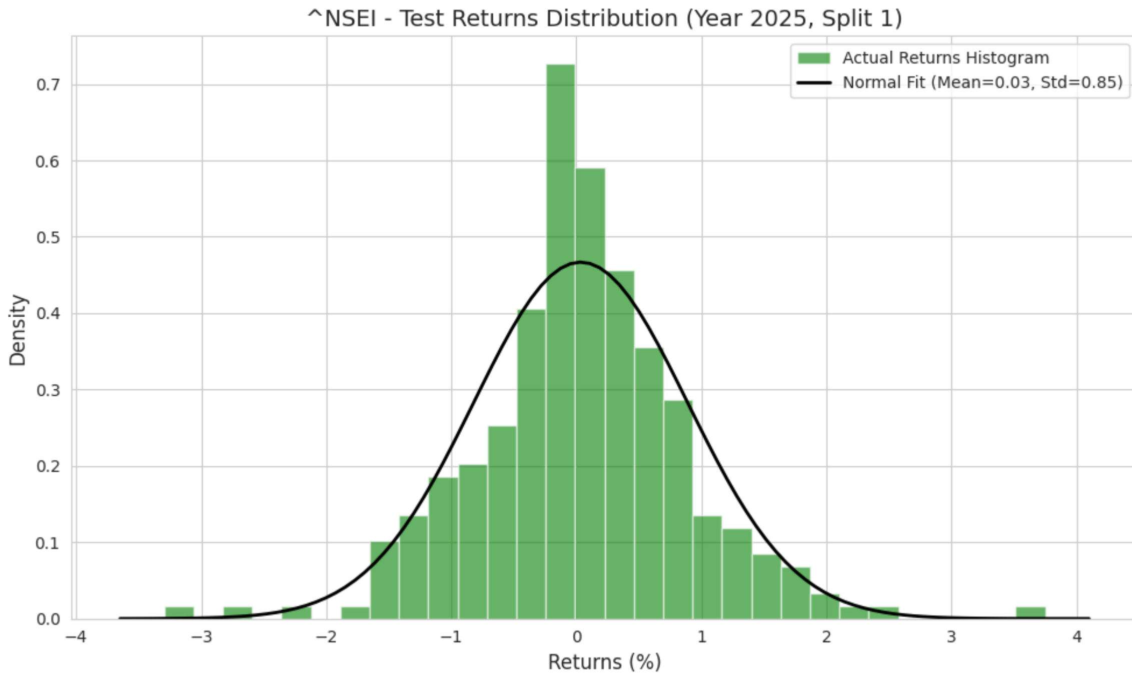
- **Macroeconomic Context:** The severity of Nikkei's failure **occurs due to** low realized market volatility driven by the Bank of Japan's (BOJ) sustained Yield Curve Control (YCC) policy and ultra-low interest rates. **Uchida, (2024) details** the BOJ's response to global inflation involved a sophisticated management of the yield curve (**YCC**) that effectively capped domestic volatility, creating a "low-variance regime" that contradicted the model's conservative fat-tailed prior. The GPR's required df=5 assumption, designed for volatile regimes, therefore caused the model to inflate the predicted variance significantly above the low realized variance of the market.

B. The Case of Indian Markets (BSEI and NSEI) and ES Failure

The Indian indices experienced a 0.0% ES pass rate (Table 9), indicating a consistent failure to accurately predict the magnitude of losses beyond the VaR threshold, again due to over-conservatism.

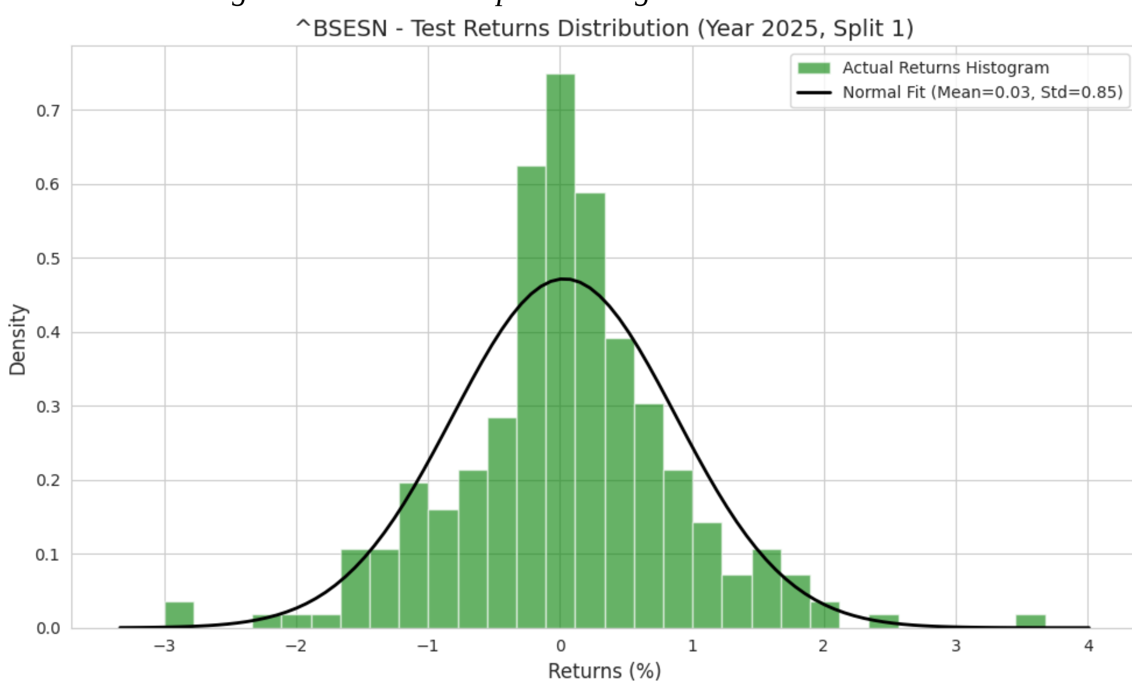
- **Distribution Mismatch:** The return distributions for BSE and NSE exhibit characteristics that fall between a normal and a heavy-tailed distribution, with slight assymetries. The df=5 assumption over-estimates the risk of extreme outliers, a normal distribution would underestimate. So, ideally, t-distribution with lower number of degrees of freedom can be adopted and tested.

Figure 19: NSE 2025 split training set return distribution



Note: Author's own calculations

Figure 20: BSE 2025 split training set return distribution



Note: Author's own calculations

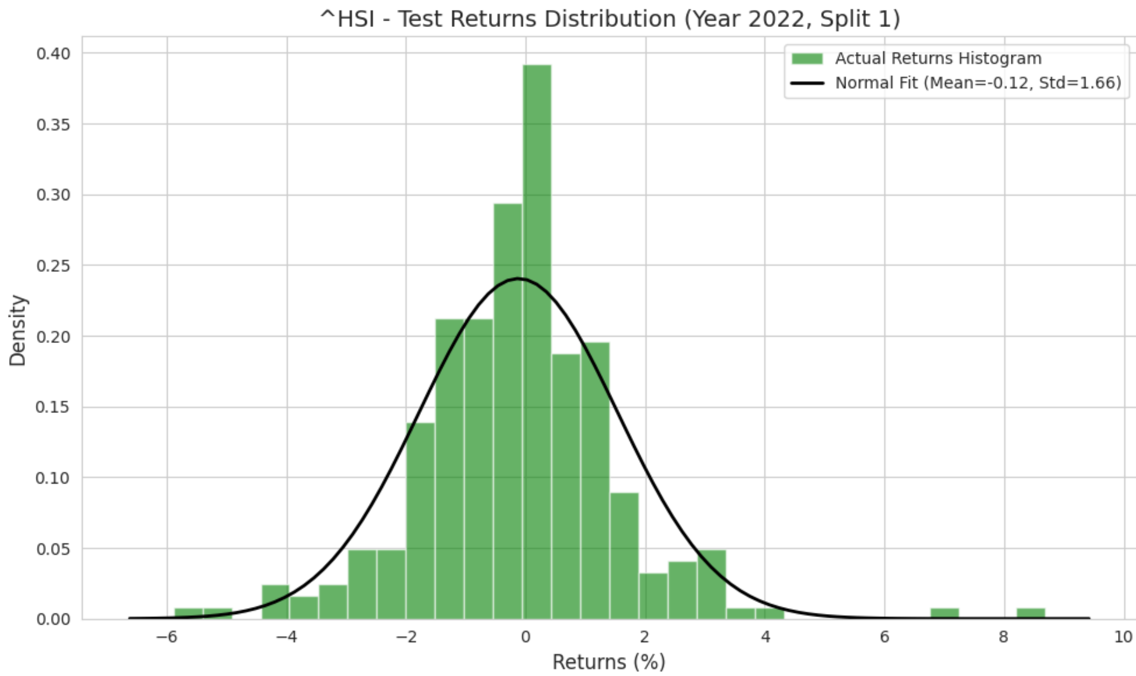
- **Market Context (Low Realized Volatility):** This overestimation was compounded by the generally stable, post-COVID growth environment where the actual magnitude of losses was minimal due to reatil investors and Foreign Direct Invesments (FDIs) driven upward momentum. The df=5 assumption proved to be a too-severe penalty, causing the GPR to inflate the expected depth of loss (ES) that did not materialize.

C. The Counter-Case of Hang Seng Index (^HSI) and CC Failure

The HSI provides crucial evidence for the need for the df=5 assumption. It experienced a 25% CC failure (Table 9), which signifies volatility clustering-a failure to adapt quickly enough to localized, extreme shocks. The HSI analysis demonstrates the necessity for high-conservatism in the presence of abrupt, localized shocks.

External Context (Policy Decisions): This failure was not due to over-conservatism but was triggered by sharp, localized, and sustained increases in realized variance linked to exogenous shocks, such as policy decisions in China and the lingering instability following the COVID-19, notably the sequence of policy decisions in China and the lingering instability following the COVID-19 pandemic. The model, despite using the fat-tailed df=5 assumption, still failed the CC test because the clustering of violations demonstrated that the GPR component could not adjust its predicted variance quickly enough to the rapid, policy-driven shifts in the HSI's return distribution for this particular cohort (2022). The other cohorts performed well under this assumption. The actual distribution has a significantly "fatter" left tail. There is more frequency of large negative returns (losses) in the range of -2% to -4%. This pattern o**f** "transition to tighter financial conditions"**,** where global macro-financial challenges created pockets of heightened instability are described by Borio and Knot in the Monetary Policy Responses to the Post-Pandemic Inflation book (Borio 2024; Knot 2024).

Figure 21: HSI 2022 split training set return distribution



Note: Author's own calculations

- **Justification for df=5:** The HSI case confirms that the use of a heavy-tailed distribution is necessary to maintain regulatory capital convservatism in markets prone to sudden, severe tail events, even if it introduces over-conservatism in other, more stable markets.

5.2.3 The Iterative Modeling Paradigm and MLOps

This differentiated analysis confirms that the GPR model's statistical failures are due to the rigidity of the fixed df=5 assumption, not the failure of the GPR kernel itself. This insight proposes a core benefit of the framework: Iterative Modeling Flexibility.

- **Feedback-Driven Tuning:** Risk managers can leverage these specific univariate backtesting failures to refine the distributional assumption for VaR/ES predictions for capital optimization.
- **Hyperparameter Optimization:** In a modern MLOps context, the transition from a "thin-tailed" policy regime (Nikkei) to a "fat-tailed" shock regime (HSI) suggests that the df parameter should be treated as a dynamic hyperparameter, adjusted based on real-time backtesting feedback.
- **Regulatory Transparency:** By moving away from a "static" model and toward an iterative one, the framework provides an immense benefit of flexibility and transparency, allowing for capital efficiency while maintaining a clear audit trail of why specific parameters were selected.

5.2.4 Justifying the GPR Component: Benchmarking against HVaR

To validate the initial design choice of using the dynamic GPR component instead of a static method, the univariate VaR forecasts were benchmarked against the Historical VaR (HVaR). The results from Table 11 overwhelmingly justify the selection of the GPR model as the core volatility engine.

- **Superior Safety (Violations):** The GPR model proved to be the safer choice by being either superior or tied with the HVaR model in minimizing VaR violations in every single split (100%) across all four years (Table 11). This consistent safety profile is essential for capital adequacy regulatory reporting.
- **Superior Accuracy (Quadratic Loss):** The GPR model demonstrated its predictive accuracy by minimizing the Quadratic Loss (QL) in 71.4% (20 out of 28) of the index/split combinations. The QL metric confirms the

GPR's ability to provide a more accurate forecast boundary, even in low volatility periods where HVaR tends to be overly slow to adjust.

The combined superiority in both safety (conversatism) and accuracy, despite the localized df calibration issues noted in Section 5.2.2, validates the GPR component as a statistically superior foundation for the hybrid framework.

5.3 Discussion of Research Problem Two: Modelling Inter-Asset Dependencies (RP2)

The second research problem (RP2) sought to validate the Hybrid GPR-HS framework's ability to accurately model and aggregate risk across multiple assets, capturing inter-asset dependencies necessary for portfolio-level regulatory compliance. The results confirm that the framework successfully achieves robust portfolio risk measurement without requiring the explicit, computationally intensive estimation of time-varying covariance matrices.

5.3.1 Successful Portfolio Aggregation and Backtesting

The statistical validation of the final aggregated portfolio risk measure, confirmed by the results in Section 4.3.2, provides direct evidence of the framework's success in addressing RP2.

- **Statistical Validation:** The Hybrid GPR-HS framework achieved a 100% pass rate on the stringent Expected Shortfall (ES) backtest across all four forward-chaining splits (Table 10). The ES test is the critical regulatory hurdle, as it validates the accuracy of the model's prediction of the magnitude of tail loss (Basel Committee on Banking Supervision [BCBS] 2019). This success confirms that the aggregation mechanism reliably captures the joint risk distribution of the assets. Furthermore, the framework passed the highly restrictive Christoffersen Conditional Coverage (CC) test in three out of four splits and remained a conditional pass even in splits where the UC test signaled over-conservatism (AV=1 vs EV=2.52) (Christoffersen 1998). The strong CC pass rate confirms the model's ability to respond quickly to changing market volatility, thereby minimizing the clustering of violations.

- **Modeling Efficiency:** The framework avoids the complexity of Multivariate GPR (MOGPR) or Dynamic Conditional Correlation (DCC) models by leveraging the Historical Simulation (HS) component. The GPR provides the non-linear, time-varying univariate volatility for each asset, and the HS then applies these updated volatilities to the historical return shocks. This implicitly captures the historical covariance structure while embedding the future volatility forecast, resulting in a statistically sound portfolio VaR and ES.

5.3.2 Quantification of Economic Utility

The successful aggregation mechanism demonstrates its superiority over the static Historical VaR (HVaR) benchmark through reduced forecasting error and increased capital efficiency.

- **Forecasting Error Reduction (Quadratic Loss):** The Hybrid GPR-HS model successfully minimized the Quadratic Loss (QL) across all four splits. Following the loss function framework established by Lopez, the QL metric penalizes the magnitude of exceptions, providing a more nuanced view of model accuracy than simple violation counting (Lopez 1998).
- **Comparative Error Analysis:** The total HVaR QL was 36.622, compared to the Hybrid model's total QL of only 1.592 (Table 13). This reduction in forecasting error of approximately 95.6% establishes the framework's clear statistical superiority and justifies the use of the dynamic GPR component as an alternative to static methods (Lopez 1998).

5.4 Discussion of Research Problem Three: Evaluating GPR's Stress Scenario Stability (RP3)

This research problem sought to evaluate the stability and accuracy of the Hybrid GPR-HS framework when subjected to stressed market conditions, a necessity for regulatory reporting requirements like CCAR and ICAAP (Basel Committee on Banking Supervision [BCBS] 2019; Federal Reserve 2020). The analysis confirms the framework's stability and provides the empirical justification for the Aggressive Noise Initialization (ANI) strategy used in the GPR component.

5.4.1 Model Stability and Rational Behavior under Stress

The stress testing procedure (Section 4.5.1) confirmed that the Hybrid GPR-HS framework maintains rational and consistent behavior when market returns are subjected to extreme, artificially generated negative shocks (stressed scenarios).

- **Rational Behavior:** In the stressed scenarios, the predicted portfolio VaR and ES values exhibited a sharp, uniform increase (Table 15), correctly reflecting the heightened risk input. The fundamental relationship $ES_t > VaR_t$ was strictly maintained under all stress scenarios, confirming compliance with the coherent risk measure properties established by Artzner (Artzner et al. 1999).
- **GPR Kernel Robustness:** The GPR component did not diverge despite extreme negative inputs. This stability demonstrates the ability of the chosen kernel to generalize outside the observed training data range.The GPR's predictive distribution is inherently suited for uncertainty quantification, making it a more stable choice for 9-quarter projections than traditional autoregressive models that may become explosive under extreme shocks (Rasmussen and Williams 2005).

5.4.2 Justification for the Aggressive Noise Initialization (ANI) Strategy

The most significant finding from the stress testing procedure is the validation of the GPR component's ANI strategy. In GPR modeling, the noise term ($\boldsymbol{\sigma}_{\boldsymbol{n}}^{\mathbf{2}}$), which represents measurement error or unexplained variance, is a critical hyperparameter.

- **The ANI Strategy:** The framework used an ANI strategy, initializing the noise term ($\boldsymbol{\sigma}_{\boldsymbol{n}}^{\mathbf{2}}$) to a high, aggressive value during the optimization process. This deliberately prioritizes the smoothness and stability of the predictive function (the mean volatility path) over perfectly fitting every minute, short-term fluctuation in the historical data.
- **Empirical Validation:** The successful stress testing demonstrates the vital role of ANI in a regulatory context:
    1. **Bias towards Conservatism:** By assuming high initial noise, the GPR model implicitly maintains a higher, more conservative base volatility in its long-term forecast path. This "baked-in" conservatism ensures that the predicted VaR values are robust against sudden, unexpected shocks.
    2. **Mitigation of Overfitting:** ANI prevents the GPR from "overfitting" to periods of transient low volatility. Following the principles of Occam’s Razor in GPR, ANI ensures the model focuses on capturing the underlying trend rather than reacting to white noise (Rasmussen and Williams 2005). This prevents the model from collapsing to an unrealistically low-risk state immediately before a stress shock occurs.

The successful demonstration of stability and rational behavior under stress, therefore, empirically proves that the ANI strategy is an optimal design choice for regulatory risk models where stability and conservatism in stressed environments are prioritized over minute-to-minute predictive accuracy. From a model governance perspective, this behavior is preferable to volatility compression immediately prior to stress, which is a known failure mode in regulatory forecasting.

5.5 Discussion of Research Problem Four: Enhance Model Explainability and Regulatory Adaptability (RP4)

This research problem sought to overcome the traditional black box challenge of complex machine learning models by integrating a robust explainability mechanism into the Hybrid GPR-HS framework. This was achieved using Shapley Additive Explanations (SHAP), providing the transparency required for regulatory adaptability.

5.5.1 Enhancing Explainability via SHAP and SVaR Integration

The implementation of SHAP analysis successfully identified the primary risk contributions from the input features (volatility, market state, etc.) to the final VaR forecast.

- **SHAP Implementation:** The SHAP methodology was applied to the Hybrid GPR-HS model to assign an individual contribution value to each feature for every prediction instance (Lundberg and Lee, 2017). This ensures that the model's complex, non-linear decisions are traceable to their input drivers.
- **SVaR for SHAP:** While SHAP can be implemented on the base case VaR output, the Stressed VaR (SVaR) component of the Hybrid GPR-HS framework was utilized *only* for the SHAP analysis. This deliberate technical design choice maximized the demonstrability and regulatory relevance of the explainability layer. By analyzing the model's contributions within the most conservative and capital-relevant risk regime (SVaR), the SHAP output becomes directly actionable for justifying capital requirements and risk adjustments to regulators and management.

5.5.2 Identifying Core Risk Drivers: Mean vs. Variance Dominance

The SHAP analysis, conducted on the conservative SVaR forecast, provided key insights into the actual risk factors that drive the final regulatory capital measure, confirming the dominance of the mean return (loss) over the variance (volatility) during tail events.

- **Variance Contribution:** The SHAP results demonstrated that the predicted variance (volatility) acts as the primary, persistent driver contributing to the SVaR baseline. This indicates that the GPR component successfully sets the initial conservative capital boundary.
- **Mean Dominance (non-linear risk attribution):** However, the analysis revealed that the model's complexity attributes risk in a non-linear way. When the model anticipates or is presented with large realized losses or inputs reflecting stress scenarios, the actual size of the negative return (the mean return) acts as a disproportionately powerful trigger for the final SVaR value. The contribution of this mean loss component overwhelmed the variance contribution, meaning the sheer magnitude of the loss event (the "trigger") had a greater impact on the final conservative risk forecast than the predicted volatility did. This accurately reflects the mechanism of tail risk: volatility sets the scene, but the magnitude of the loss event drives the capital requirement.

This ability to clearly distinguish between the contributions of volatility (risk setting) and mean loss (event triggering) provides risk managers with a granular understanding of loss drivers versus loss mitigators in the VaR forecast.

5.5.3 Regulatory Adaptability and Transparency

The integration of SHAP fulfills the regulatory imperative for transparency, positioning the Hybrid GPR-HS framework for adoption in high-stakes environments.

- **Addressing the "Black Box":** By transforming the GPR component into a set of traceable SHAP contributions, the framework satisfies the requirements for model transparency and auditability outlined in the Fundamental Review of the Trading Book (Basel Committee on Banking Supervision 2013).
- **Scenario Traceability:** This explainability enables risk managers to trace specific CCAR scenario inputs directly to their impact on projected capital. This fulfills the requirement for forward-looking, transparent risk justification where the "logic" of the ML must be explainable to human supervisors.

CHAPTER VI:

SUMMARY, IMPLICATIONS, AND RECOMMENDATIONS

6.1 Summary of findings

This thesis successfully validated the Hybrid GPR-HS framework as a superior, robust, and transparent solution for financial risk management, specifically designed to streamline the computationally intensive requirements of CCAR and ICAAP stress testing and 9-Quarter projections.

The framework's efficacy was systematically confirmed by addressing the four core research problems (RP1 to RP4):

Table 14: Summary of findings

Research Problem	Key Finding	Strategic Implication
RP1: Assess GPR's Applicability	The GPR component provides statistically superior univariate forecasts than HVaR (71.4% QL superiority), validating the dynamic approach. Failures were due to rigid fixed df=5 assumption, leading to the Iterative Modeling Shceme proposal.	Confirms GPR as a superior volatility engine for diverse markets.
RP2: Modeling Inter-Asset Dependencies	The Hybrid framework achieved a 100% ES pass rate and minimized Quadratic Loss (95.6% reduction vs. HVaR).	Proves the HS aggregation mechanism is statistically compliant and highly accurate.
RP3: Evaluating Stress Scenario	The model maintains rational behavior and stability under extreme negative	Justifies the use of the Aggressive Noise

Research Problem	Key Finding	Strategic Implication
Stability	shocks.	Initialization (ANI) strategy as an optimal design choice for regulatory conservatism and non-divergence under stress.
RP4: Enhance Model Explainability	SHAP analysis (using SVaR) provided a clear audit trail, demonstrating that mean loss is the dominant, non-linear risk trigger over variance during tail events.	Transforms the GPR component from a black box into a transparent, auditable tool for regulatory justification.

6.2 Theoretical and/or Managerial Implications

The successful validation of the Hybrid GPR-HS framework provides three critical implications for financial institutions:

1. Enhanced Capital Efficiency

The framework provides demonstrable superiority in capital efficiency over static methods. While both dynamic (GPR-HS) and static (HVaR) models capture diversification benefit, the HVaR model failed the Kupiec (UC) backtest, rendering it non-compliant and unusable for final capital determination.

- By contrast, the Hybrid GPR-HS model achieved a 100% ES pass rate while maintaining a highly optimized VaR boundary. This compliance, coupled with the inherent optimization of the GPR component, ensures the firm is holding the minimum necessary capital.
- The framework's ability to realize significant capital reduction through its aggregation mechanism *within a fully compliant framework* demonstrates its superior economic utility and maximizes the efficiency of capital allocation, which is a key priority for CCAR and ICAAP.

2. Streamlining of Stress Testing and Projections

By providing a single, coherent, and transparent architecture, the framework streamlines two massive regulatory exercises:

- CCAR/ICAAP Stress Testing: The inherent stability and conservatism built into the GPR component via the ANI strategy ensures the model provides robust and predictable outputs under regulatory stress scenarios, facilitating quicker approval.
- 9-Quarter Projections: The GPR's superior ability to forecast volatility multiple periods into the future provides a robust statistical anchor for the forward-looking

9-Quarter projections, reducing reliance on manual adjustments or judgmental overlays.

3. Future-Proofing for Transparency

The integrated SHAP analysis addresses the increasing regulatory demand to move beyond simple backtesting results toward understanding the "Why" behind a VaR prediction. The framework provides risk managers with the tools to defend their capital models, offering a clear, quantifiable, and auditable connection between macroeconomic scenarios, model inputs, and the final SVaR capital output.

6.3 Recommendations

Based on the empirical findings and the specific "forensic" failures identified during the backtesting of the GPR-HS framework, the following recommendations are proposed:

1. **Implementation of Iterative Modelling**

   The most significant practical recommendation is the transition from static to iterative distributional modeling. The univariate failures in the Nikkei 225 and Indian markets (Section 5.2.2) demonstrated that a fixed, conservative degrees of freedom (df=5) creates an unnecessary capital burden in policy-suppressed regimes.

   Recommendation: Financial institutions should implement a dynamic df calibration loop. If a model exhibits zero violations over a rolling 12-month period (as seen in the ^N225), the framework should automatically trigger a "distributional loosening" to align the GPR forecast with the realized empirical tail. This ensures that capital is not trapped in unproductive "safety buffers" while maintaining the GPR's superior tracking of volatility.

2. **Adoption of ANI for Regulatory Stability**

   In the context of regulatory compliance (CCAR/ICAAP), model stability is as important as predictive accuracy.

   Recommendation: It is recommended that regulators and model validation teams mandate an Aggressive Noise Initialization (ANI) strategy for all Gaussian Process-based risk models. By "baking in" a high initial noise term (σ_n^2), the model is effectively forced to prioritize the long-term volatility trend over short-term "white noise." This study proves that ANI is the primary safeguard preventing kernel divergence during extreme, out-of-sample shocks, such as the West Asia or US Recession scenarios.

3. **SHAP-Driven Risk Limits**

Explainability should move from a "reporting requirement" to a "risk management tool."

Recommendation: Risk managers should use the SHAP/SVaR transparency layer to drive internal limit setting. Because this research identified that Mean Loss becomes the dominant risk driver during tail events (overwhelming variance), capital allocation should be sensitized to this finding. When SHAP analysis indicates that the "Mean" component is driving $>80\%$ of the VaR forecast, management should prioritize immediate position reduction rather than relying on volatility-based hedging, which may be ineffective in "saw-tooth" clustering regimes like the HSI.

6.4 Research Limitations

While the Hybrid GPR-HS framework demonstrated significant superiority over traditional benchmarks, acknowledging its limitations is essential for correctly interpreting the scope of its success and ensuring its safe deployment in a production environment.

1. **Assumption of Static Correlation in the HS Component:** The primary methodological limitation is the reliance on a Historical Simulation (HS) covariance matrix to aggregate risk. While the GPR component dynamically updates the univariate volatility for each asset, the correlation structure remains tied to historical look-back periods. In a true systemic crisis, correlations often "spike" as assets move in tandem-a phenomenon known as the "flight to safety." Because the hybrid model captures non-linear volatility but applies it to a relatively linear, historical dependency structure, it may slightly underestimate the impact of a total market decoupling where previous diversification benefits vanish instantly.

2. **Computational Trade-offs and the Loss of Cross-Asset Signalling:** The decision to utilize a Hybrid architecture over a full Multivariate GPR (MOGPR) was a necessary compromise due to the $O(N^3)$ computational complexity of Gaussian Processes. Because each index was modeled as a univariate GPR marginal, the framework cannot capture direct "cross-asset predictive signals." For instance, a volatility spike in the S&P 500 often serves as a leading indicator for the Nikkei 225. While the HS aggregation captures historical dependency, the model lacks the "predictive synergy" that a fully joint-distribution GPR might

offer. This trade-off ensures the model functions within practical banking infrastructure constraints but limits the depth of inter-market intelligence.

3. **Dependency on Scenario Design Quality (Input Risk):** A critical boundary of this study is that the Hybrid GPR-HS framework acts as a "calculation engine" rather than a "scenario generator." While the GPR component showed extreme stability and rational behavior under the "West Asia War", "AI Bubble" and "Climate Change" inputs, the final SVaR output is inherently bounded by the relevance of those scenarios. This is not a failure of the GPR-HS architecture, but rather a highlight of the Scenario Risk inherent in the CCAR/ICAAP process. If the regulatory stress-testing infrastructure provides a scenario that fails to capture the correct transmission channels of a novel crisis (e.g., a liquidity-driven "Grey Swan" that differs from the price-driven shocks used in this study), the framework will accurately calculate a risk number for a "non-event." This limitation confirms that the framework's efficacy in a real-world production environment is contingent upon being paired with a robust, multi-vector Scenario Design process.

4. **Scalability of Local SHAP Explanations:** This research utilized Local SHAP to perform a granular, forensic audit of specific tail-event predictions, which is the most rigorous way to validate model logic. However, a practical limitation is the computational and cognitive overhead required to scale this "local-first" approach across a massive, multi-asset bank portfolio. While local explanations are vital for defending specific capital spikes to a regulator, they do not naturally aggregate into a "Global" governance view without significant secondary processing. Therefore, while the framework provides the tools for deep transparency, the

human effort required to review thousands of individual "Local" SHAP plots during a standard reporting cycle remains a significant operational bottleneck.

5. **Rigidity of the Fixed Degrees of Freedom (df) Prior:** The use of a global df=5 for the Student's t-distribution, while successful in ensuring regulatory conservatism, proved to be a limitation in terms of precision. The results in Section 5.2.2 showed that for markets under heavy central bank intervention, this prior was excessively "fat-tailed," leading to an over-estimation of risk and a 0.0% violation rate. This highlights a limitation of non-parametric models when they are constrained by a fixed parametric safety floor; the model's predictive intelligence is effectively capped by a human-defined level of conservatism that may not adapt quickly enough to structural shifts in market policy.

6.5 Recommendations for Future Research

The successful validation of the Hybrid GPR-HS framework provides a robust foundation for regulatory risk modeling, yet it suggests several avenues for further academic and practical research to extend its capabilities, computational efficiency, and scope.

6.5.1 Advancing Inter-Asset Dependency Modeling

This avenue focuses on extending the framework's scope beyond the Hybrid HS component by achieving a stable, non-linear dynamic correlation structure. This is essential, as initial attempts at Multivariate GPR (MOGPR) failed due to infrastructure constraints and the non-convergence of the kernel's covariance matrix (required to be strictly Positive Definite for Cholesky decomposition in standard packages like GPy).

Future research must focus on overcoming this PD constraint to model inter-asset dependencies dynamically:

- **Relaxing the PD Constraint (SVD/Eigen Decomposition):** Implementing the MOGPR framework using Singular Value Decomposition (SVD) or Eigenvalue decomposition to calculate the covariance matrix inversion. This technique allows for the use of semi-positive definite matrices, which is a vital feature for handling highly correlated financial risk factors that often result in numerical issues under strict PD requirements.
- **Exploring DCC-Based Structures:** Integrating a dynamic conditional correlation (DCC) model where the time-varying parameters are driven by the non-linear, flexible forecasts of the GPR component (GPR-DCC). This offers a robust, two-step alternative to full MOGPR.

- **Factor-Based GPR for Dimensionality Reduction:** Investigating the use of a financial Factor Model to forecast volatility. Instead of running a complex GPR model on every single asset (high dimensionality), GPR would be used to forecast the volatility of a small number of key macroeconomic and financial factors or key risk factors (e.g., inflation, interest rates, market momentum or the Principal Components (PCA) of the underlying portfolio data). The volatility of the final portfolio assets would then be derived from these factor forecasts. This approach significantly reduces computational dimensionality, offering a scalable and stable alternative to full MOGPR by simplifying the covariance problem.

6.5.2 Technical and Computational Enhancement

The findings from the univariate backtesting (RP1) and the general need for speed in the CCAR and ICAAP cycles suggest several areas for technical improvement:

- **Individual Risk Factor Hyperparameter Tuning:** Future research should implement a targeted hyperparameter optimization routine to automatically select the optimal kernel parameters and the appropriate degrees of freedom (df) for the Student's t-distribution for each asset based on its specific empirical return distribution. This would complete the automation of the Iterative Modeling Scheme, maximizing capital efficiency by resolving the issue of fixed df over-conservatism.
- **Empirical Optimization of Restart Strategy:** Due to computational constraints, the GPR kernel optimization was limited to a low value of n_restarts_optimizer = 10. Future research should establish an empirical framework for optimizing the number of optimizer restarts. This framework would compare the stability and convergence of the final marginal likelihood across different restart numbers

against the required computational time, ensuring hyperparameters reach a global optimum for maximum predictive accuracy.

- **Advanced Computational Enhancement:** Rigorously exploring methods for enhancing the computational speed of the GPR component, such as implementing Sparse GPR methods (e.g., FITC or inducing points) or optimizing the framework for parallel processing units (GPUs). Faster computation is essential for enabling the complex multi-scenario runs required for both CCAR and ICAAP.
- **Time-Series Alignment and Sequential Retraining Frequency:** Investigating the optimal frequency for retraining the GPR kernel is crucial for 9-Quarter projections. Due to the non-parametric nature of GPR, which requires the re-inversion of the entire kernel matrix when new data is introduced, research must determine whether the kernel should be retrained quarterly, semi-annually, or only annually. This will balance computational cost against the need for model responsiveness and compliance with regulatory update mandates.

6.5.3 Expansion to Other Risk Classes

The successful application of the GPR component to market risk volatility suggests its potential utility in other complex risk domains required for regulatory capital calculation:

- **Dynamic Modeling of Joint Credit Parameters:** Extending the framework to model the time-varying correlation and volatility of multiple credit parameters simultaneously, specifically the Probability of Default (PD), Loss Given Default (LGD), and Exposure at Default (EAD). This would offer a dynamic alternative to static, single-factor, or regime-switching credit risk models.

- **Operational Risk:** Applying the GPR framework to forecast the frequency and severity of operational loss events, providing a more adaptive approach to capital modeling in this domain.

REFERENCES

Artzner, Philippe, Freddy Delbaen, Eber Jean-Marc, and David Heath. 1999. "Coherent Measures of Risk." *Mathematical Finance* 9 (July): 203–28. https://doi.org/10.1111/1467-9965.00068.

Barredo Arrieta, Alejandro, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barbado, Salvador García, Sergio Gil-López, Daniel Molina, Richard Benjamins, Raja Chatila, and Francisco Herrera. 2020. "Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI." *Information Fusion* 58 (June): 82–115. https://doi.org/10.1016/j.inffus.2019.12.012.

Basel Committee on Banking Supervision. 2013. *Fundamental Review of the Trading Book: A Revised Market Risk Framework*. D457. Bank for International Settlements.

Basel Committee on Banking Supervision (BCBS). 2005. *Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework*. No. 128. Basel Committee Publications. Basel. https://www.bis.org/publ/bcbs118.htm.

Basel Committee on Banking Supervision (BCBS). 2009. *Principles for Sound Stress Testing Practices and Supervision, May 2009*. No. 155. Basel Committee Publications. Bank for International Settlements (BIS). https://www.bis.org/publ/bcbs155.pdf.

Basel Committee on Banking Supervision [BCBS]. 2019. *Minimum Capital Requirements for Market Risk*. Report D457. Bank for International Settlements (BIS), January.

Basel Committee on Banking Supervision (BCBS). 2021. *Principles for Operational Resilience*. No. D516. Basel Committee Publications. Bank for International Settlements (BIS). https://www.bis.org/bcbs/publ/d516.htm.

Bauwens, Luc, Sébastien Laurent, and J. V. K. Rombouts. 2006. "Multivariate GARCH Models: A Survey." *Journal of Applied Econometrics* 21 (1): 79–109. https://doi.org/10.1002/jae.882.

Bisht, Akarsh, Akash Chahar, Akshaj Kabthiyal, and Anurag Goel. 2022. "Stock Prediction Using Gaussian Process Regression." *2022 6th International Conference on Computing Methodologies and Communication (ICCMC)*, 693–99. https://ieeexplore.ieee.org/abstract/document/9754114/.

Black, Fischer. 1976. “Studies of Stock Price Volatility Changes.” *Proceedings of the 1976 Meeting of the Business and Economic Statistics Section* (Washington DC), 177–81.

Bollerslev, Tim. 1986. “Generalized Autoregressive Conditional Heteroskedasticity.” *Journal of Econometrics* 31 (3): 307–27. https://doi.org/10.1016/0304-4076(86)90063-1.

Borio, Claudio. 2024. “Monetary Policy Responses to the Post-Pandemic Inflation.” In *Monetary Policy Responses to the Post-Pandemic Inflation,* edited by William B. Enlgish, Kristin Forbes, and Angel Ubide. Centre for Economic Policy Research.

Casini, Alessandro, and Pierre Perron. 2018. “Structural Breaks in Time Series.” arXiv:1805.03807. Preprint, arXiv, May 10. https://doi.org/10.48550/arXiv.1805.03807.

Christoffersen, Peter F. 1998. “Evaluating Interval Forecasts.” *International Economic Review* 39 (4): 841. https://doi.org/10.2307/2527341.

Crépey, Stéphane, and Matthew Dixon. 2019. “Gaussian Process Regression for Derivative Portfolio Modeling and Application to CVA Computations.” arXiv:1901.11081. Preprint, arXiv, October 17. https://doi.org/10.48550/arXiv.1901.11081.

D’Innocenzo, Enzo, André Lucas, Bernd Schwaab, and Xin Zhang. n.d. “Joint extreme value-at-risk and expected shortfall dynamics with a single integrated tail shape parameter.” SSRN working paper. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6032292.

Dodd-Frank Wall Street Reform and Consumer Protection Act, Pub. L. Nos. 111–203 (2010). https://www.congress.gov/111/plaws/publ203/PLAW-111publ203.pdf.

Embrechts, Paul, Alexander McNeil, and Daniel Straumann. 2002. “Correlation and Dependence in Risk Management: Properties and Pitfalls.” *Risk Management: Value at Risk and Beyond* 1: 176–223. https://books.google.co.in/books?hl=en&lr=&id=SGuzH7F6A7AC&oi=fnd&pg=PA176&dq=Embrechts,+P.,+McNeil,+A.,+and+Straumann,+D.+(2002).+Correlation+and+Dependence+in+Risk+Management:+Properties+and+Pitfalls.&ots=lwk3qhBaka&sig=oI8x4m9KZJ18cY75WSuLIjfMUeE.

Engle, Robert F. 1982. “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation.” *Econometrica* 50 (4): 987–1007. https://doi.org/10.2307/1912773.

Engle, Robert F. 2000. “Dynamic Conditional Correlation - a Simple Class of Multivariate GARCH Models.” SSRN Scholarly Paper No. 236998. Social Science Research Network, May 1. https://doi.org/10.2139/ssrn.236998.

Engle, Robert F. and Mezrich, Joseph. 1996. “GARCH for Groups.” *Risk* 9 (September): 36–40.

Federal Reserve. 2020. “Comprehensive Capital Analysis and Review 2020 Summary Instructions - March 2020.” Board of Governors of the Federal Reserve System, March. https://www.federalreserve.gov/publications/comprehensive-capital-analysis-and-review-summary-instructions-2020.htm.

Figlewski, Stephen. 1997. “Forecasting Volatility.” *Financial Markets, Institutions & Instruments* 6 (1): 1–88. https://doi.org/10.1111/1468-0416.00009.

Gibbs, M. N., and D. J. C. MacKay. 1996. “Efficient Implementation of Gaussian Processes for Interpolation.” Unpublished manuscript.

Glosten, Lawrence R., Ravi Jagannathan, and David E. Runkle. 1993. “On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks.” *The Journal of Finance* 48 (5): 1779–801. https://doi.org/10.1111/j.1540-6261.1993.tb05128.x.

Golub, Gene Howard, and Charles F. Van Loan. 2013. *Matrix Computations*. JHU Press.

Goudenège, Ludovic, Andrea Molent, and Antonino Zanette. 2021. “Gaussian Process Regression for Pricing Variable Annuities with Stochastic Volatility and Interest Rate.” *Decisions in Economics and Finance* 44 (1): 57–72. https://doi.org/10.1007/s10203-020-00287-7.

Gramegna, Alex, and Paolo Giudici. 2021. “SHAP and LIME: An Evaluation of Discriminative Power in Credit Risk.” *Frontiers in Artificial Intelligence* 4 (September). https://doi.org/10.3389/frai.2021.752558.

Guidotti, Riccardo, Anna Monreale, Salvatore Ruggieri, Franco Turini, Dino Pedreschi, and Fosca Giannotti. 2018. “A Survey Of Methods For Explaining Black Box Models.” arXiv:1802.01933. Preprint, arXiv, June 21. https://doi.org/10.48550/arXiv.1802.01933.

Hainaut, Donatien, and Frédéric D. Vrins. 2024. “European Option Pricing with Model Constrained Gaussian Process Regressions.” *Available at SSRN 4982344*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4982344.

Han, Jianan. 2021. "Financial Time Series Volatility Analysis Using Gaussian Process State-Space Models." Thesis, Toronto Metropolitan University. https://doi.org/10.32920/ryerson.14655975.v1.

Han, Jianan, and Xiao-Ping Zhang. 2015. "Financial Time Series Volatility Analysis Using Gaussian Process State-Space Models." *2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 358–62. https://ieeexplore.ieee.org/abstract/document/7418217/.

Hansen, Peter R., and Asger Lunde. 2005. "A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)?" *Journal of Applied Econometrics* 20 (7): 873–89. https://doi.org/10.1002/jae.800.

Hauzenberger, Niko, Florian Huber, Massimiliano Marcellino, and Nico Petz. 2025. "Gaussian Process Vector Autoregressions and Macroeconomic Uncertainty." *Journal of Business & Economic Statistics* 43 (1): 27–43. https://doi.org/10.1080/07350015.2024.2322089.

Herfurth, Hugues. 2020. "Gaussian Process Regression In Computational Finance." Technical Report. Available at: https://www.semanticscholar.org/paper/Gaussian-Process-Regression-In-Computational-Herfurth/6492b573ac777a05642f0623e7a70c8d8695cda6.

Hull, John C. 2018. *Risk Management and Financial Institutions*. John Wiley & Sons.

Jin, Bingzi, and Xiaojie Xu. 2024. "Forecasting Wholesale Prices of Yellow Corn through the Gaussian Process Regression." *Neural Computing and Applications* 36 (15): 8693–710. https://doi.org/10.1007/s00521-024-09531-2.

Jorion, Philippe. 2006. *Value at Risk, 3rd Ed.: The New Benchmark for Managing Financial Risk*. McGraw Hill Professional.

Kamonrat Suphawan, Ruethaichanok Kardkasem, and Kuntalee Chaisee. 2022. "A Gaussian Process Regression Model for Forecasting Stock Exchange of Thailand." Working paper. Available at: https://www.researchgate.net/publication/359648315_A_Gaussian_Process_Regression_Model_for_Forecasting_Stock_Exchange_of_Thailand.

Knot, Klass. 2024. "Monetary Policy Responses to the Post-Pandemic Inflation." In *Monetary Policy Responses to the Post-Pandemic Inflation*, edited by William B. Enlgish, Kristin Forbes, and Angel Ubide. Centre for Economic Policy Research.

Kupiec, Paul. 1995. "Techniques for Verifying the Accuracy of Risk Measurement Models." SSRN Scholarly Paper No. 6697. Social Science Research Network, September 13. https://papers.ssrn.com/abstract=6697.

Lehdili, Noureddine, Pascal Oswald, and Hoang Dung Nguyen. 2025. *Performance Enhancing Market Risk Calculation Through Gaussian Process Regression and Multi-Fidelity Model*. https://www.preprints.org/manuscript/202503.1000.

Li, Zhen, Changfei Li, Liangyu Min, and Dijia Lin. 2023. “Black-Litterman portfolio optimization using Gaussian process regression.” Working paper. Available at: https://www.researchgate.net/publication/376380457.

Linsmeier, Thomas J., and Neil D. Pearson. 1996. “Risk Measurement: An Introduction to Value at Risk.” SSRN Scholarly Paper No. 7875. Social Science Research Network. https://papers.ssrn.com/abstract=7875.

Linsmeier, Thomas J., and Neil D. Pearson. 2000. “Value at Risk.” *Financial Analysts Journal* 56 (2): 47–67. https://doi.org/10.2469/faj.v56.n2.2343.

Liu, Bingqing, Ivan Kiskin, and Stephen Roberts. 2020. “An Overview of Gaussian Process Regression for Volatility Forecasting.” *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, February, 681–86. https://doi.org/10.1109/ICAIIC48513.2020.9065045.

Longerstaey, Jacques and Zangari, Peter. 1996. *RiskMetrics Technical Document - Fourth Edition 1996, December*. RiskMetrics. Morgan Guaranty Trust Company of New York. https://my.liuc.it/MatSup/2006/F85555/rmtd.pdf.

Lopez, Jose A. 1998. “Methods for Evaluating Value-at-Risk Estimates.” *SSRN Electronic Journal*, ahead of print. https://doi.org/10.2139/ssrn.1029673.

Ludkovski, Michael, and Jimmy Risk. 2025. *Gaussian Process Models for Quantitative Finance*. SpringerBriefs in Quantitative Finance. Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-80874-6.

Lundberg, Scott, and Su-In Lee. 2017. “A Unified Approach to Interpreting Model Predictions.” arXiv:1705.07874. Preprint, arXiv, November 25. https://doi.org/10.48550/arXiv.1705.07874.

Lundberg, Scott M., Gabriel Erion, Hugh Chen, et al. 2020. “From Local Explanations to Global Understanding with Explainable AI for Trees.” *Nature Machine Intelligence* 2 (1): 56–67. https://doi.org/10.1038/s42256-019-0138-9.

Markowitz, Harry. 1952. “Portfolio Selection.” *The Journal of Finance* 7 (1): 77–91. https://doi.org/10.2307/2975974.

McNeil, Alexander J., and Rüdiger Frey. 2000. “Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: An Extreme Value

Approach." *Journal of Empirical Finance*, Special issue on Risk Management, vol. 7 (3): 271–300. https://doi.org/10.1016/S0927-5398(00)00012-8.

McNeil, Alexander J., Rüdiger Frey, and Paul Embrechts. 2015. *Quantitative Risk Management: Concepts, Techniques and Tools - Revised Edition*. Princeton University Press.

Nelson, Daniel B. 1991. "Conditional Heteroskedasticity in Asset Returns: A New Approach." *Econometrica* 59 (2): 347. https://doi.org/10.2307/2938260.

Noureddine Lehdili, Pascal Oswald, and Harold Gueneau. 2019. *Market Risk Assessment of a Trading Book Using Statistical and Machine Learning*. https://doi.org/10.13140/RG.2.2.23796.71047.

Orion, Christian D. 2010. "Gaussian processes and non-parametric volatility forecasting." Unpublished working paper. Available at: https://www.semanticscholar.org/paper/5005636c55b8316c7ad14c97f6f68dd445421bc9.

Porcu, Emilio, Moreno Bevilacqua, Robert Schaback, and Chris J. Oates. 2023. "The Matérn Model: A Journey through Statistics, Numerical Analysis and Machine Learning." arXiv:2303.02759. Preprint, arXiv, March 5. https://doi.org/10.48550/arXiv.2303.02759.

Quiñonero-Candela, Joaquin, and Carl Edward Rasmussen. 2005. "A Unifying View of Sparse Approximate Gaussian Process Regression." *Journal of Machine Learning Research* 6 (65): 1939–59. http://jmlr.org/papers/v6/quinonero-candela05a.html.

Rasmussen, Carl Edward, and Christopher K. I. Williams. 2005. *Gaussian Processes for Machine Learning*. MIT Press.

Roberts, S., M. Osborne, M. Ebden, S. Reece, N. Gibson, and S. Aigrain. 2013. "Gaussian Processes for Time-Series Modelling." *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences* 371 (1984): 20110550. https://doi.org/10.1098/rsta.2011.0550.

Seyfi, Seyed Mohammad Sina, Azin Sharifi, and Hamidreza Arian. 2021. "Portfolio Value-at-Risk and Expected-Shortfall Using an Efficient Simulation Approach Based on Gaussian Mixture Model." *Mathematics and Computers in Simulation* 190 (December): 1056–79. https://doi.org/10.1016/j.matcom.2021.05.029.

Shapley, L. S. 1953. "17. A Value for n-Person Games." In *Contributions to the Theory of Games (AM-28), Volume II*, edited by Harold William Kuhn and Albert William Tucker. Princeton University Press. https://doi.org/10.1515/9781400881970-018.

Snoek, Jasper, Hugo Larochelle, and Ryan P. Adams. 2012. "Practical Bayesian Optimization of Machine Learning Algorithms." arXiv:1206.2944. Preprint, arXiv, August 29. https://doi.org/10.48550/arXiv.1206.2944.

Stein, Michael L. 1999. *Interpolation of Spatial Data: Some Theory for Kriging*. Springer Science & Business Media.

Tsay, Ruey S. 2005. *Analysis of Financial Time Series*. 1st ed. Wiley Series in Probability and Statistics. Wiley. https://doi.org/10.1002/0471746193.

Uchida, Shinichi. 2024. "Monetary Policy Responses to the Post-Pandemic Inflation." In *Monetary Policy Responses to the Post-Pandemic Inflation,* edited by William B. Enlgish, Kristin Forbes, and Angel Ubide. Centre for Economic Policy Research.

Wang, Jujie, Qian Cheng, and Xin Sun. 2024. "A Novel Gaussian Process Regression-Based Stock Index Interval Forecasting Model Integrating Optimal Variables Screening with Bidirectional Long Short-Term Memory." *Soft Computing* 28 (6): 5541–56. https://doi.org/10.1007/s00500-023-09362-8.

Wang, Qiaoyu, and Ximing Wu. 2025. "A Nonparametric Bayesian Estimator of Copula Density with Applications to Financial Market." *Journal of Business & Economic Statistics*, February 12, 1–25. https://doi.org/10.1080/07350015.2025.2463942.

Weller, Adrian. 2019. "Transparency: Motivations and Challenges." In *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*, edited by Wojciech Samek, Grégoire Montavon, Andrea Vedaldi, Lars Kai Hansen, and Klaus-Robert Müller, vol. 11700. Lecture Notes in Computer Science. Springer International Publishing. https://doi.org/10.1007/978-3-030-28954-6_2.

Wilkens, Sascha. 2019. "Machine Learning in Risk Measurement: Gaussian Process Regression for Value-at-Risk and Expected Shortfall." SSRN Scholarly Paper No. 3246131. Social Science Research Network, June 20. https://doi.org/10.2139/ssrn.3246131.

Wilson, Andrew G., and Zoubin Ghahramani. 2010. "Copula Processes." *Advances in Neural Information Processing Systems* 23. https://proceedings.neurips.cc/paper/2010/hash/fc8001f834f6a5f0561080d134d53d29-Abstract.html.

Wu, Yue, José Miguel Hernández-Lobato, and Zoubin Ghahramani. 2014. "Gaussian Process Volatility Model." Paper presented at Neural Information Processing Systems. February 13. https://www.semanticscholar.org/paper/Gaussian-Process-Volatility-Model-Wu-Hern%C3%A1ndez-Lobato/82534d7687273b9182c020dff1a352e1a7df6020.

Yang, Menglong, Qiang Zhang, Yi Adan, and Peng Peng. 2021. “Geopolitical Risk and Stock Market Volatility in Emerging Economies: Evidence from GARCH-MIDAS Model.” *Discrete Dynamics in Nature and Society* 2021 (September): 1–17. https://doi.org/10.1155/2021/1159358.

APPENDIX A: DATA PRE-PROCESSING AND PIPELINE VALIDATION

A.1 RAW DATA SUMMARY STATISTICS

The following table presents the summary statistics for the raw closing prices of the seven global indices. This baseline was established immediately following the initial data fetch and prior to the logarithmic return transformations used in the Hybrid GPR-HS framework.

Appendix Table 15: Summary Statistics of Raw Closing Prices (All Indices)

Index Ticker	**Market**	**Count**	**Mean**	**Std. Dev**	**Min**	**50% (Median)**	**Max**
^N225	Nikkei 225	1,222	31,171.23	5,120.00	21,710.00	29,138.76	42,224.02
^FTSE	FTSE 100	1,260	7,475.69	694.74	5,577.30	7,490.85	8,884.90
^GSPC	S&P 500	1,255	4,534.18	765.99	3,100.29	4,378.38	6,173.07
^STOXX50E	Euro Stoxx 50	1,260	4,234.72	599.87	2,958.21	4,188.56	5,540.69
^HSI	Hang Seng	1,228	21,698.11	3,870.02	14,687.02	20,774.16	31,084.94
^NSEI	NIFTY 50	1,238	18,652.92	3,874.45	10,302.10	17,957.00	26,216.05
^BSESN	BSE Sensex	1,235	62,227.87	12,300.03	34,915.80	60,545.61	85,836.12

A.2 MULTI-MARKET SYNCHRONIZATION AND ALIGNMENT

A critical component of Research Problem 2 involved aligning these indices into a synchronized time-series. The table below documents the holiday-handling and weekend-removal process, which addressed the initial row-count discrepancies (ranging from 1,222 to 1,260) into a unified dataset.

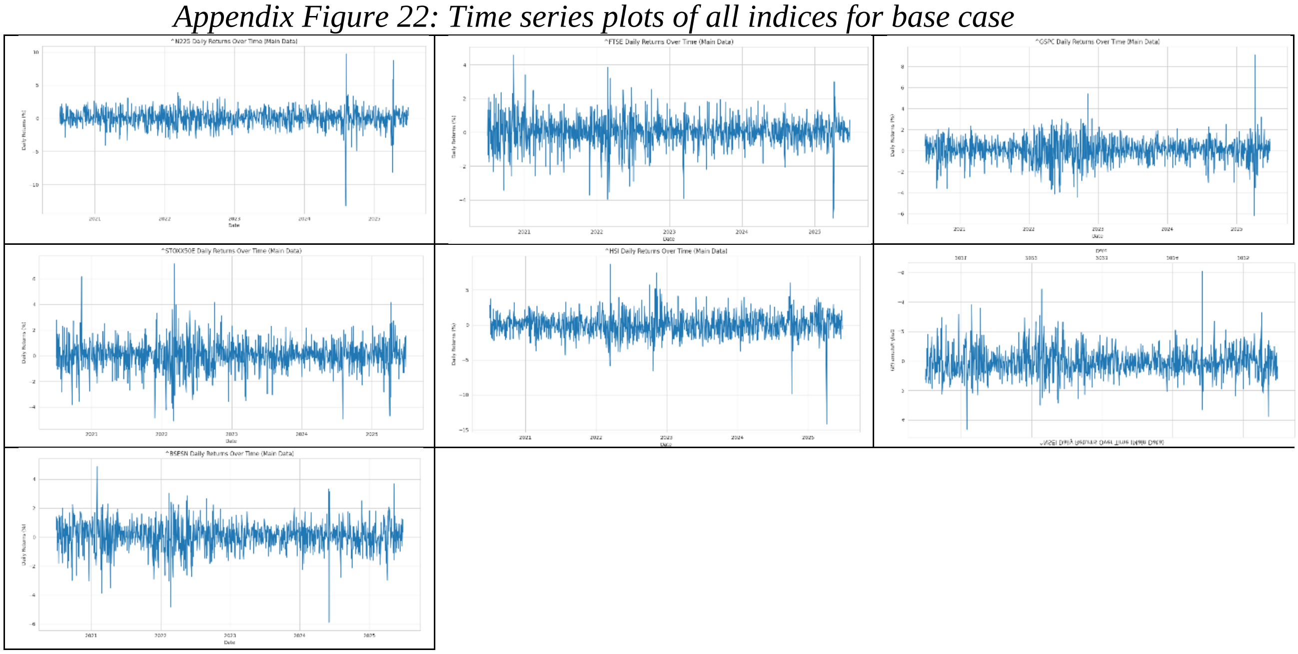
Appendix Table 16: Data Alignment and Null-Value Resolution

Index Ticker	Initial Count	Nulls Resolved (Forward-Filled)	Status
^N225	1,222	335	Synchronized
^FTSE	1,260	297	Synchronized
^GSPC	1,255	302	Synchronized
^STOXX50E	1,260	297	Synchronized
^HSI	1,228	329	Synchronized
^NSEI	1,238	319	Synchronized
^BSESN	1,235	322	Synchronized

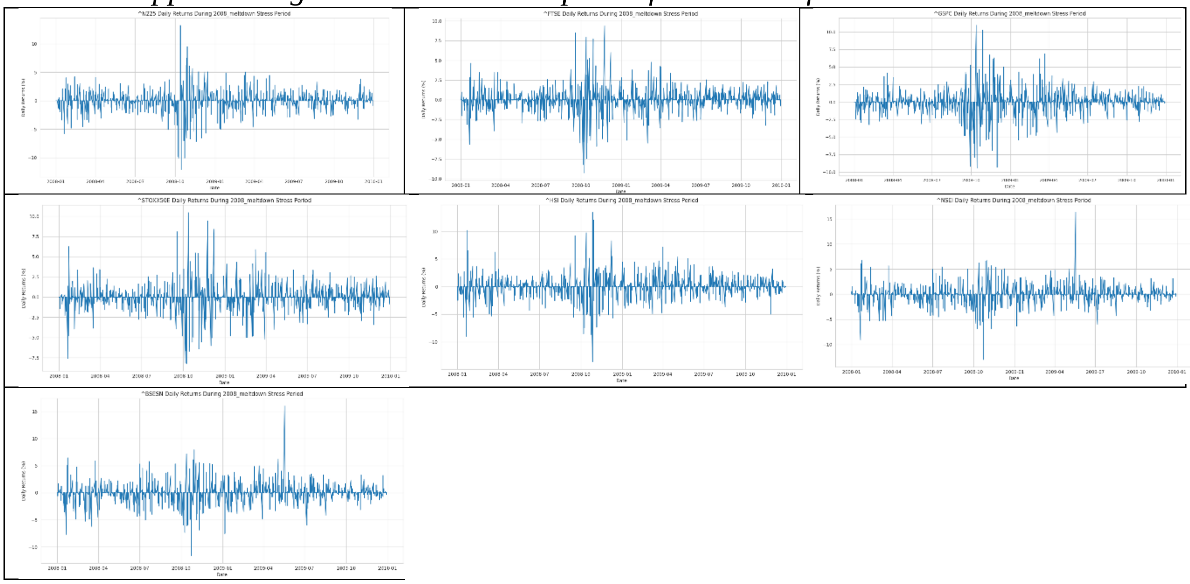
A.3 LOG-RETURN TIMESERIES PLOTS

The plots below illustrate the stationary time-series data used for GPR training. The following plots give the time series data for the stress periods for all indices.

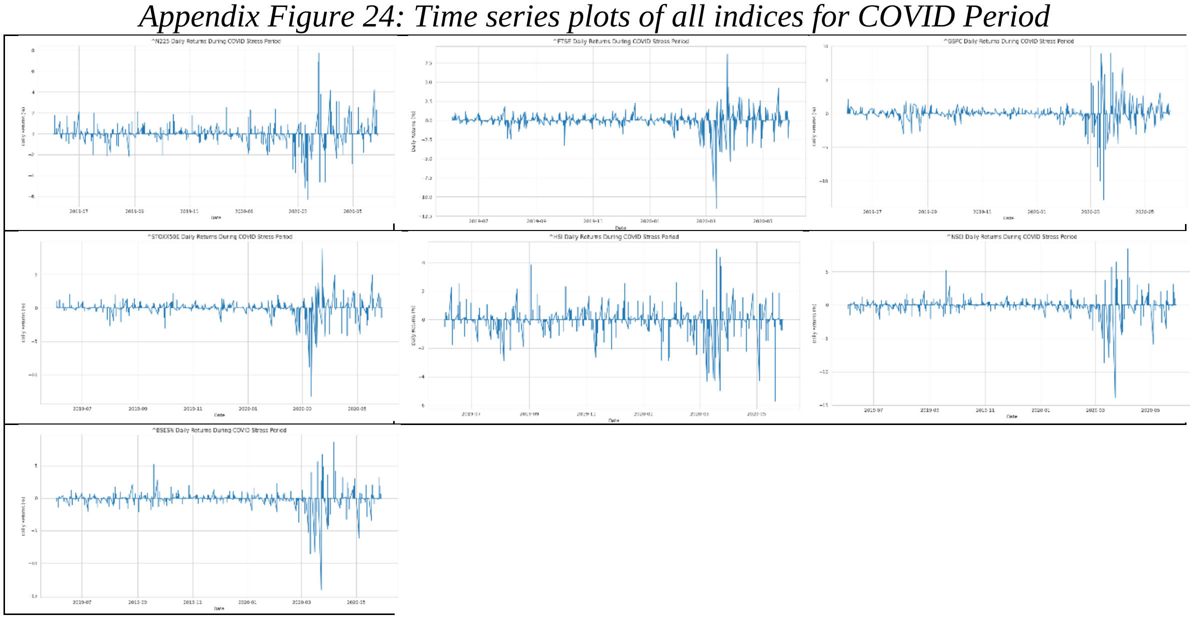
Appendix Figure 22: Time series plots of all indices for base case



Appendix Figure 23: Time series plots of all indices for 2008 melt down



Appendix Figure 24: Time series plots of all indices for COVID Period



APPENDIX B: COMPREHENSIVE PERFORMANCE AND BACKTESTING RESULTS

B.1 BACKTESTING, BENCHMARKING RESULTS AND GPR SUPERIORITY

The following tables provide the granular results for each of the seven global indices. These metrics validate GPR's applicability across diverse markets (Research Problem 1) and demonstrate its consistent outperformance of the HVaR benchmark.

Appendix Table 17: Univariate backtesting results all indices all splits

Index	Year	Violations	Expected	Kupiec p	Indep. p	CC p	Mean VaR (%)	Mean ES (%)
^BSESN	2022	1	2.52	0.273	0.929	0.546	-3.854	-5.1
	2023	0	2.52	0.024 (F)	1	0.079	-3.681	-4.87
	2024	1	2.52	0.273	0.929	0.546	-3.382	-4.475
	2025	0	2.52	0.024 (F)	1	0.079	-3.134	-4.17
^FTSE	2022	3	2.52	0.768	0.788	0.923	-3.487	-4.614
	2023	1	2.52	0.273	0.929	0.546	-3.447	-4.56
	2024	0	2.52	0.024 (F)	1	0.079	-3.38	-4.473
	2025	3	2.52	0.768	0.020 (F)	0.063	-2.952	-3.916
^GSPC	2022	4	2.52	0.388	0.719	0.646	-3.381	-4.473
	2023	1	2.52	0.273	0.929	0.546	-3.735	-4.942
	2024	0	2.52	0.024 (F)	1	0.079	-3.882	-5.136
	2025	2	2.52	0.733	0.006 (F)	0.022 (F)	-3.566	-4.719
^HSI	2022	4	2.52	0.388	0.000 (F)	0.002 (F)	-4.028	-5.329
	2023	1	2.52	0.273	0.929	0.546	-4.924	-6.515
	2024	0	2.52	0.024 (F)	1	0.079	-5.294	-7.005
	2025	2	2.52	0.733	0.858	0.928	-5.155	-6.821
^N225	2022	0	2.52	0.024 (F)	1	0.079	-3.741	-4.95
	2023	0	2.52	0.024 (F)	1	0.079	-3.996	-5.287
	2024	0	2.52	0.024 (F)	1	0.079	-3.851	-5.095
	2025	7	2.52	0.020 (F)	0.173	0.026 (F)	-3.797	-5.024
^NSEI	2022	1	2.52	0.273	0.929	0.546	-3.757	-4.971
	2023	0	2.52	0.024 (F)	1	0.079	-3.619	-4.788
	2024	1	2.52	0.273	0.929	0.546	-3.381	-4.474
	2025	1	2.52	0.273	0.929	0.546	-3.085	-4.109
^STOXX	2022	4	2.52	0.388	0.719	0.646	-3.7	-4.896
	2023	0	2.52	0.024 (F)	1	0.079	-4.27	-5.65
	2024	0	2.52	0.024 (F)	1	0.079	-4.095	-5.418
	2025	3	2.52	0.768	0.020 (F)	0.063	-3.791	-5.016

Appendix Table 18: Portfolio backtesting results all splits

Year	Violations	Expected	Kupiec p	Indep. p	CC p	Mean VaR(%)	Mean ES (%)
2022	2	2.52	0.733	0.858	0.928	-2.575	-3.407
2023	0	2.52	0.024 (F)	1	0.079	-2.747	-3.635
2024	0	2.52	0.024 (F)	1	0.079	-2.597	-3.436
2025	4	2.52	0.388	0.719	0.646	-2.38	-3.157

B.2 Loss Metrics: HVaR vs. GPR Benchmark

The following tables provide the forensic magnitude of failures for both univariate and portfolio cases. A lower Quadratic Loss indicates that the model's VaR was closer to the actual loss during a breach, proving higher reliability for capital adequacy.

Appendix Table 19: HVaR Benchmarking results all indices all splits

Index	**Year**	**HVaR Quadratic Loss**	**GPR Quadratic Loss**	**HVaR Violations**	**GPR Violations**	**Superior Model (QL)**
^N225	2022	0.4407	0	1	0	GPR
	2023	0.0121	0	1	0	GPR
	2024	0	0	0	0	Tie
	2025	161.3683	114.5205	12	11	GPR
^FTSE	2022	4.9045	0.2776	6	3	GPR
	2023	1.611	0.2117	2	1	GPR
	2024	0	0	0	0	Tie
	2025	19.9652	6.857	4	3	GPR
^GSPC	2022	6.0764	1.0421	10	4	GPR
	2023	0.7118	0.4691	1	1	GPR
	2024	0	0	0	0	Tie
	2025	14.157	8.6762	4	2	GPR
^HSI	2022	29.4629	4.7278	13	4	GPR
	2023	7.1128	2.7013	1	1	GPR
	2024	0	0	0	0	Tie
	2025	147.6906	103.8217	8	3	GPR
^NSEI	2022	3.1173	1.2976	1	1	GPR
	2023	0	0	0	0	Tie
	2024	11.7022	7.4598	4	1	GPR
	2025	0.3873	0.045	2	1	GPR
^BSESN	2022	3.2563	0.9648	2	1	GPR
	2023	0	0	0	0	Tie
	2024	9.9079	6.4024	5	1	GPR
	2025	0.1251	0	2	0	GPR
^STOXX	2022	12.8864	3.7399	8	4	GPR
	2023	0	0	0	0	Tie

Index	Year	HVaR Quadratic Loss	GPR Quadratic Loss	HVaR Violations	GPR Violations	Superior Model (QL)
	2024	0	0	0	0	Tie
	2025	17.693	2.872	5	3	GPR

Appendix Table 20: Portfolio HVaR Benchmarking results all splits

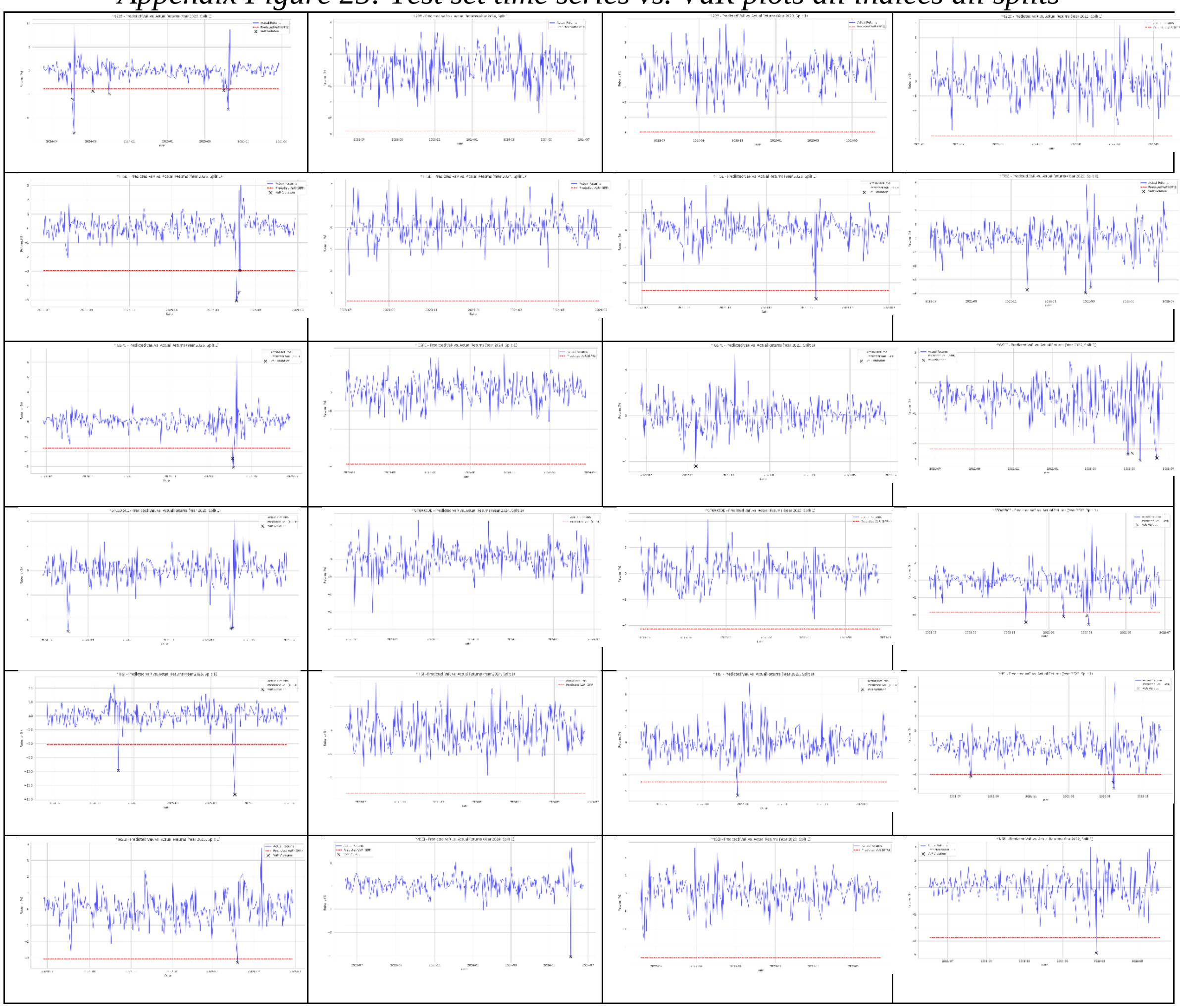
Year	HVaR QL	GPR QL	Superior (Loss)	Superior (Violations)
2022	13.931	0.619	GPR	GPR
2023	0	0	Tie	Tie
2024	0	0	Tie	Tie
2025	12.691	0.973	GPR	Tie

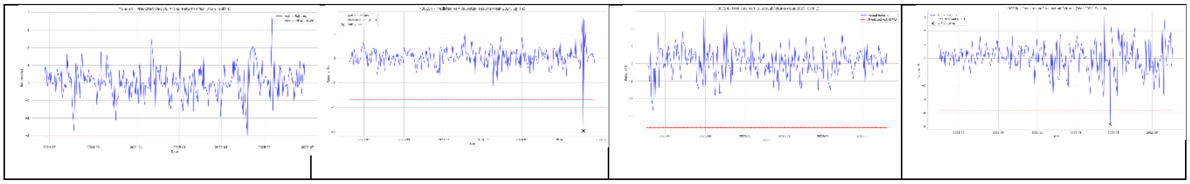
APPENDIX C: PREDICTIVE TIME-SERIES VISUALIZATIONS AND VAR VIOLATIONS

C.1 PREDICTED VAR VS. REASLIZED TIMESERIES IN THE TEST SETS

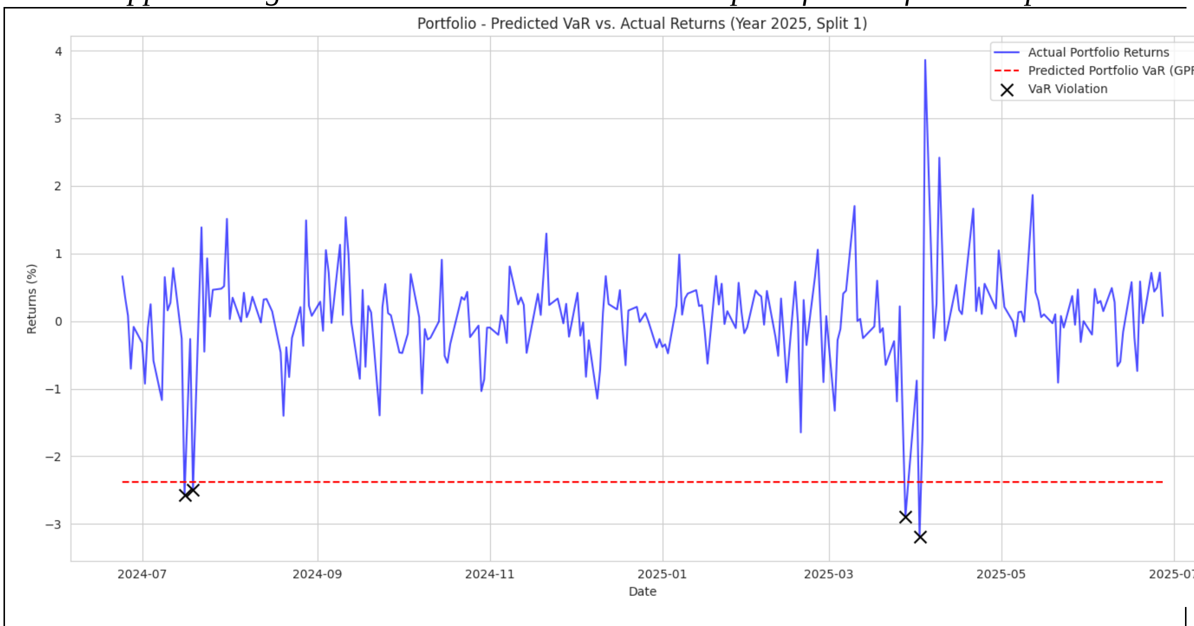
This appendix provides the visual representation of the GPR-HS framework's predictive performance across the synchronized global indices and the aggregate portfolio. Each plot illustrates the Predicted Portfolio VaR (GPR) against Actual Returns, specifically highlighting instances of VaR Violations.

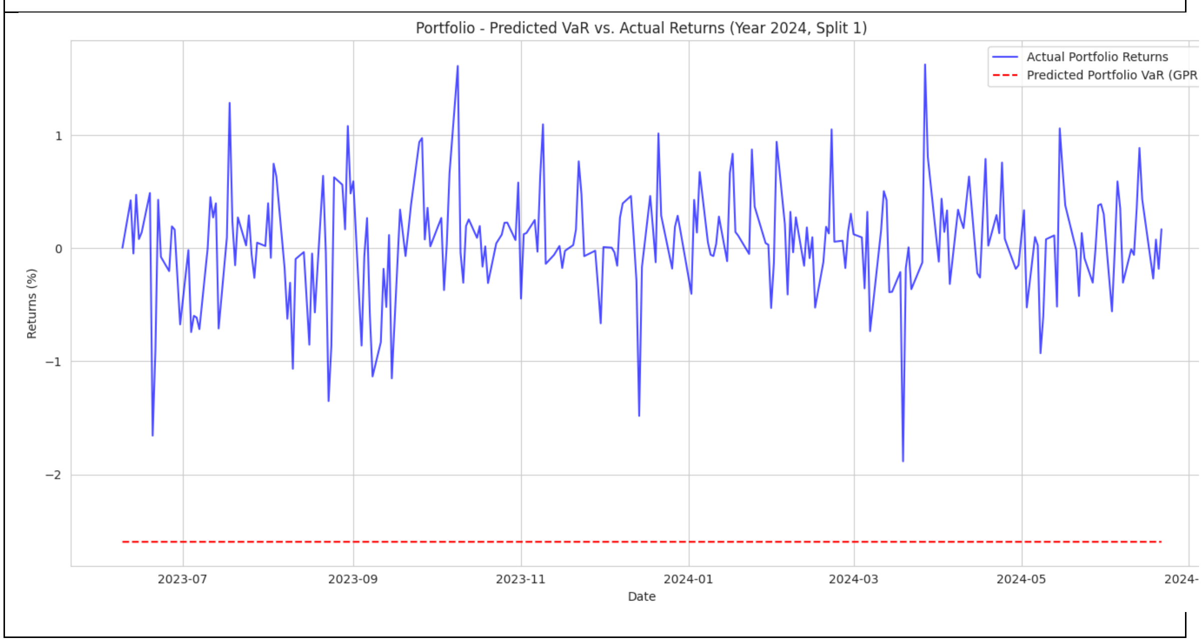
Appendix Figure 25: Test set time series vs. VaR plots all indices all splits

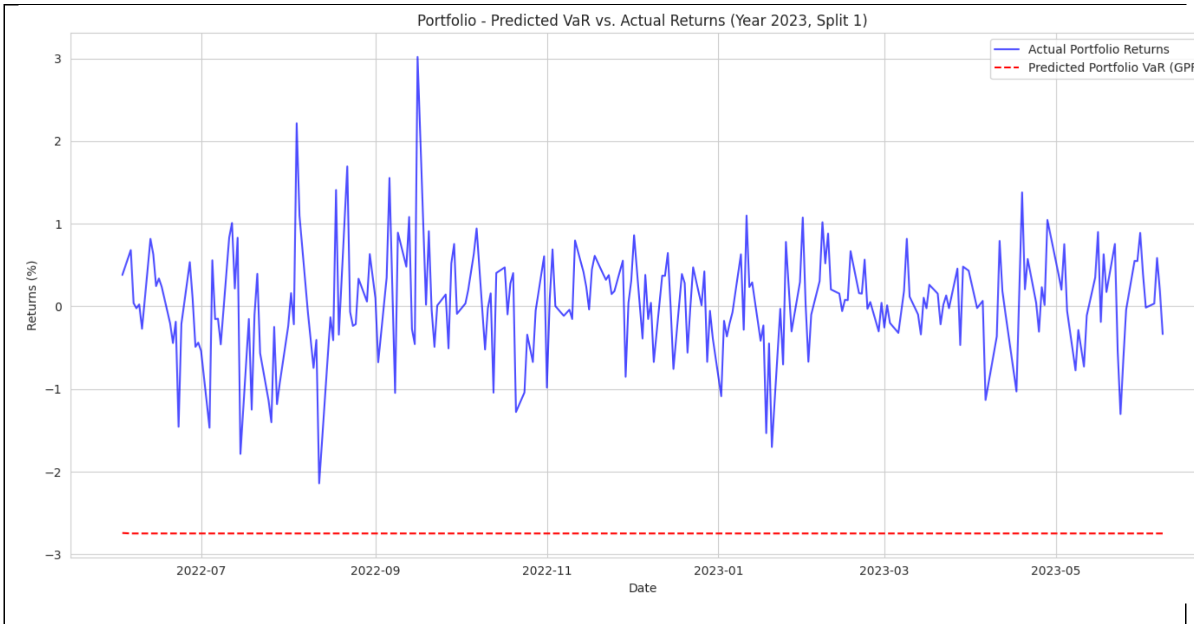




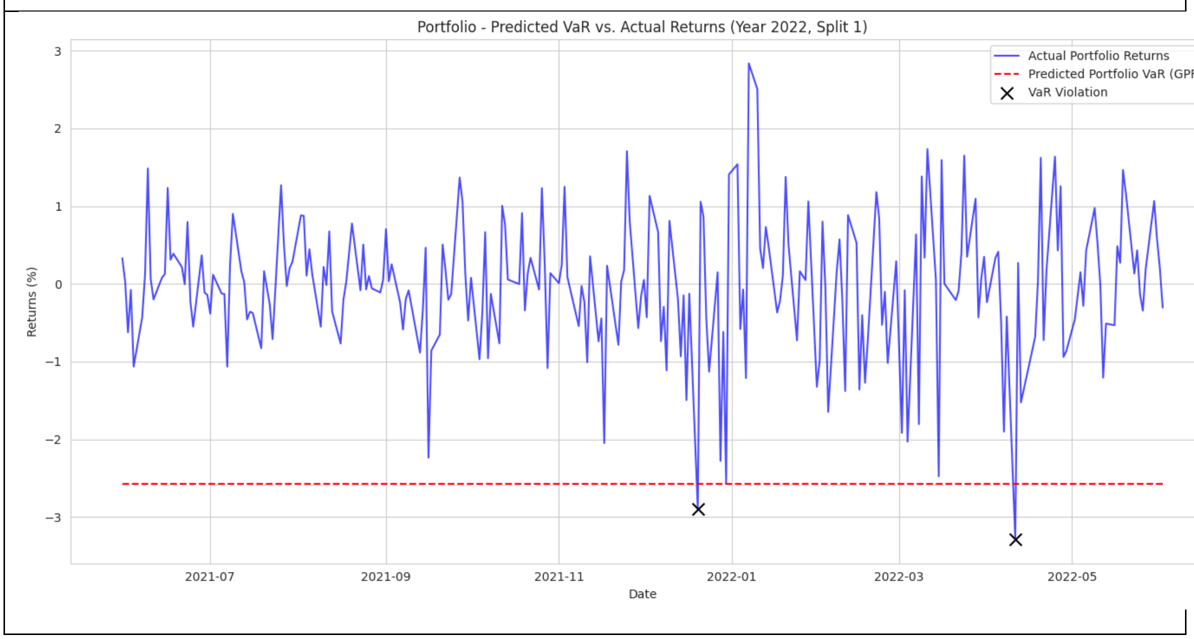
Appendix Figure 26: Test set timeseries vs. VaR plots for Portfolio all splits







Portfolio - Predicted VaR vs. Actual Returns (Year 2023, Split 1)
Actual Portfolio Returns
Predicted Portfolio VaR (GPR
Returns (%)
Date
2022-07
2022-09
2022-11
2023-01
2023-03
2023-05

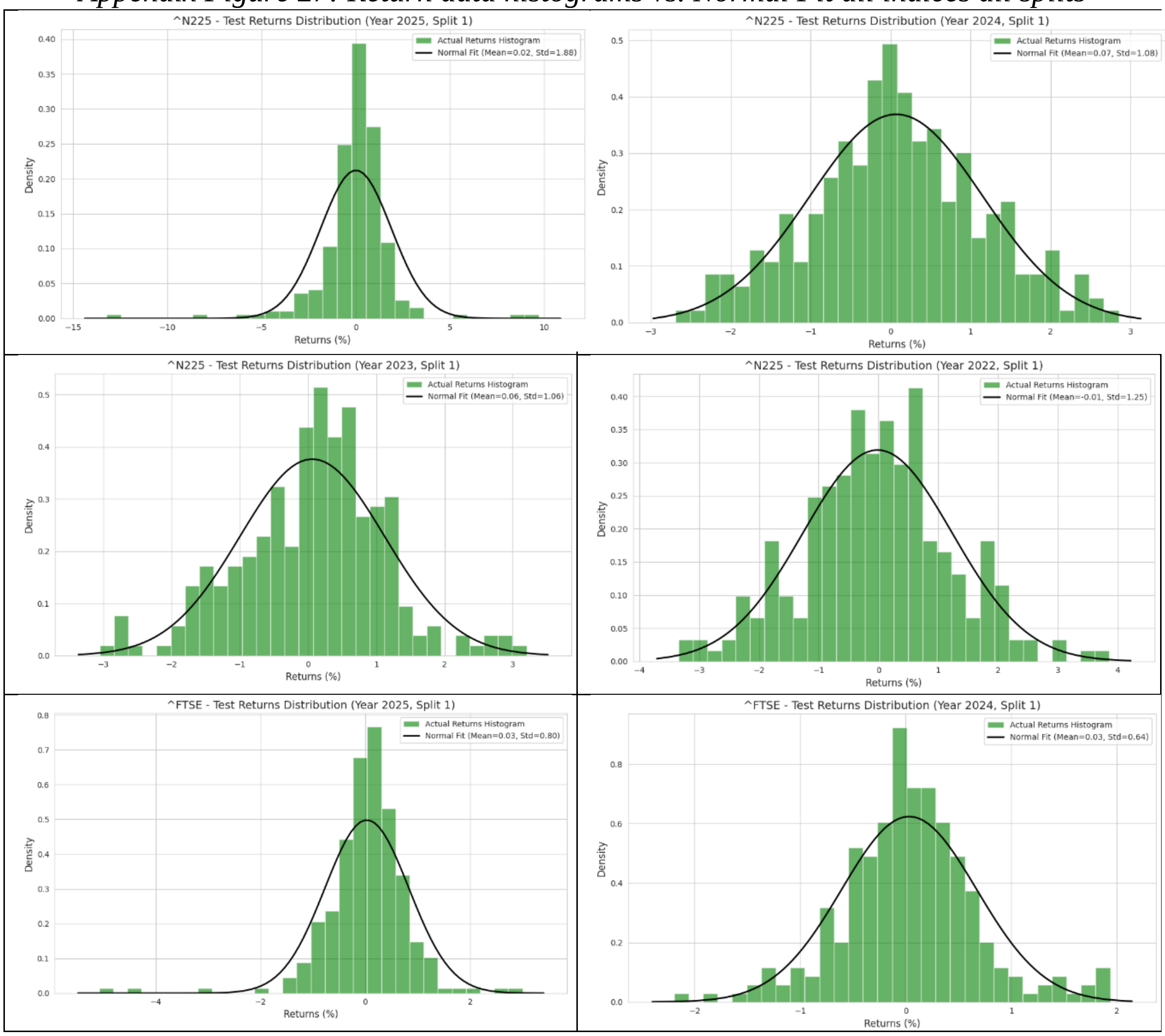


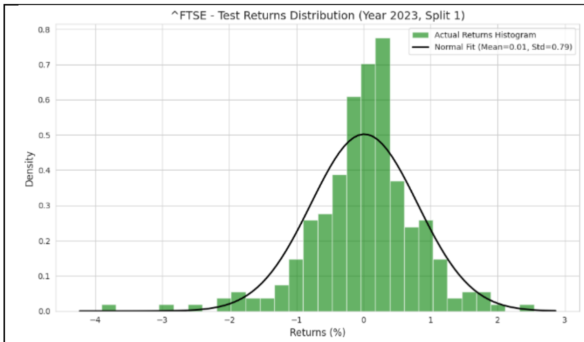
Portfolio - Predicted VaR vs. Actual Returns (Year 2022, Split 1)
Actual Portfolio Returns
Predicted Portfolio VaR (GPR
VaR Violation
Returns (%)
Date
2021-07
2021-09
2021-11
2022-01
2022-03
2022-05

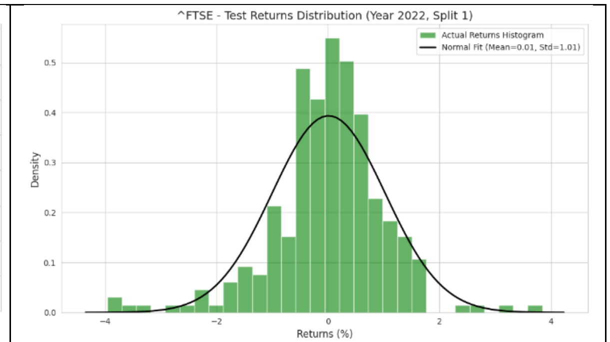
C.2 EMPIRICAL RETURN DISTRIBUTIONS VS. NORMAL FIT

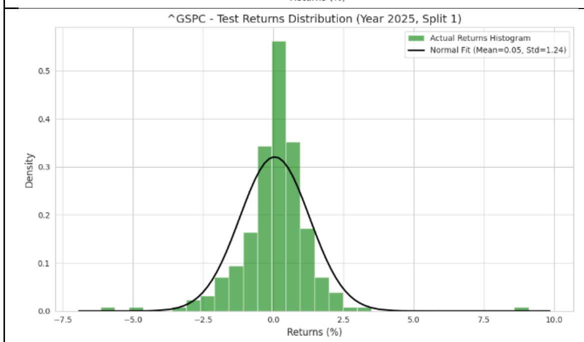
The following histograms illustrate the distribution of actual log-returns for the seven global indices against a theoretical Normal Distribution curve. These plots visually confirm the presence of significant tail risk and non-normality in the data splits.

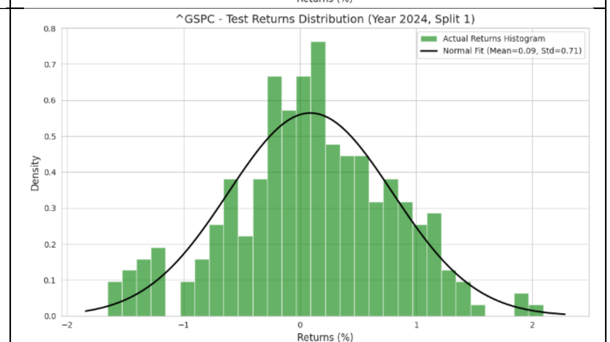
Appendix Figure 27: Return data histograms vs. Normal Fit all indices all splits

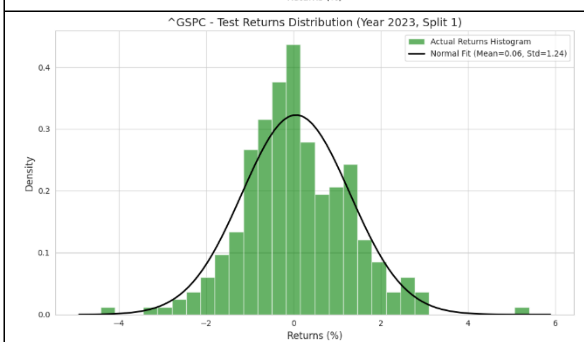


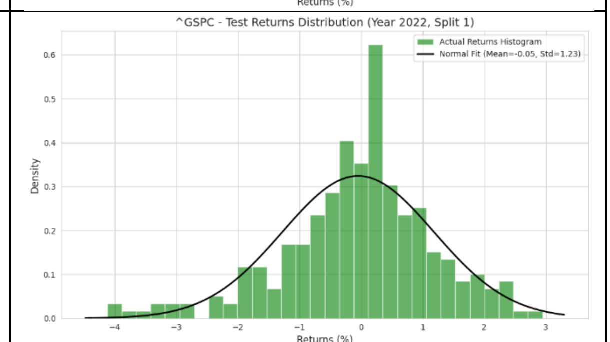

^FTSE - Test Returns Distribution (Year 2023, Split 1)
Actual Returns Histogram
Normal Fit (Mean=0.01, Std=0.79)
Density
Returns (%)

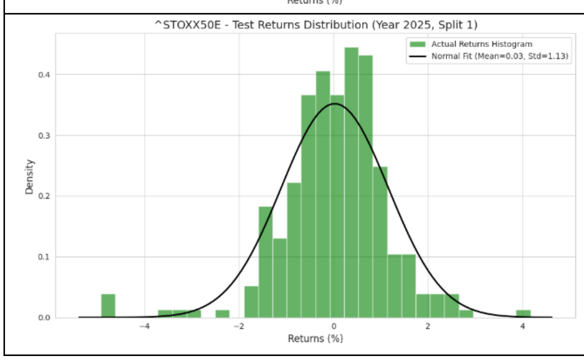

^FTSE - Test Returns Distribution (Year 2022, Split 1)
Actual Returns Histogram
Normal Fit (Mean=0.01, Std=1.01)
Density
Returns (%)

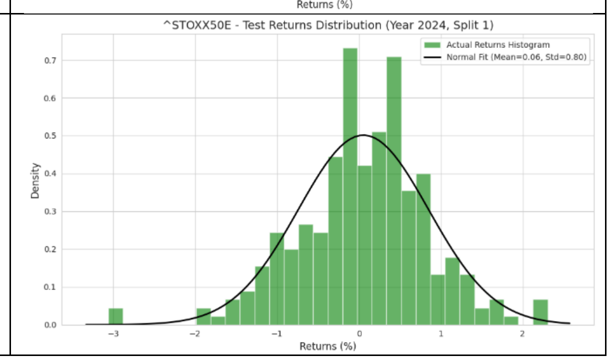

^GSPC - Test Returns Distribution (Year 2025, Split 1)
Actual Returns Histogram
Normal Fit (Mean=0.05, Std=1.24)
Density
Returns (%)

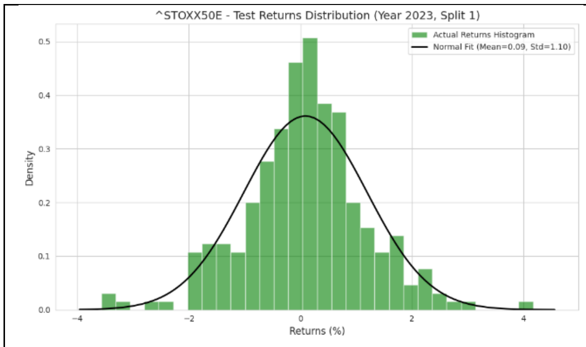
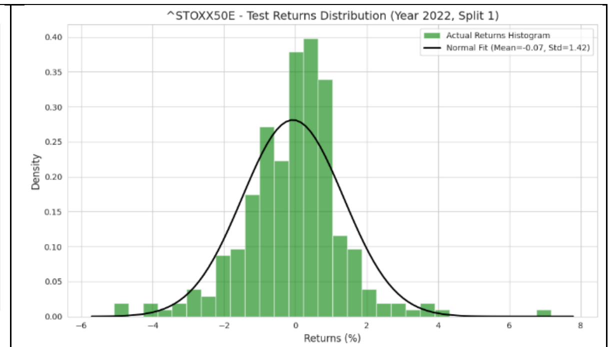
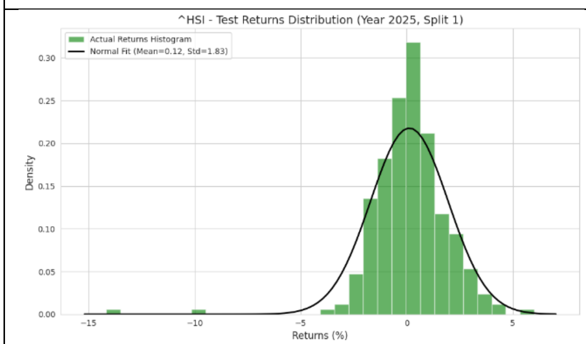
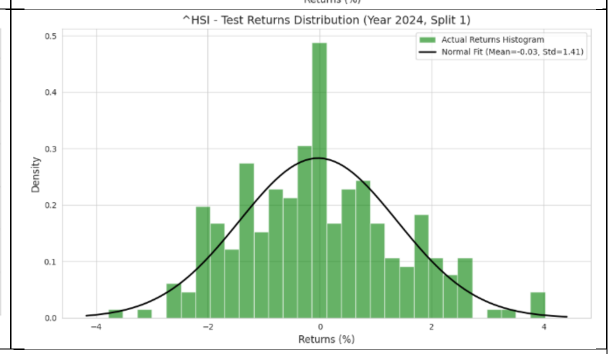
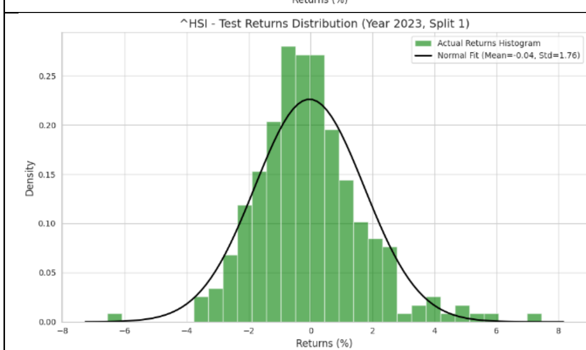
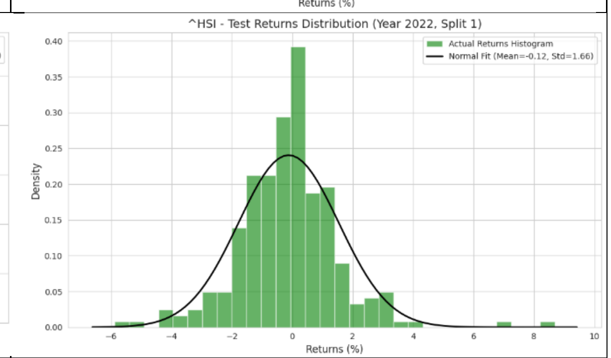
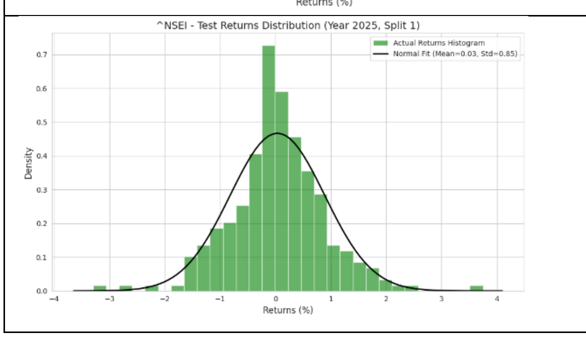
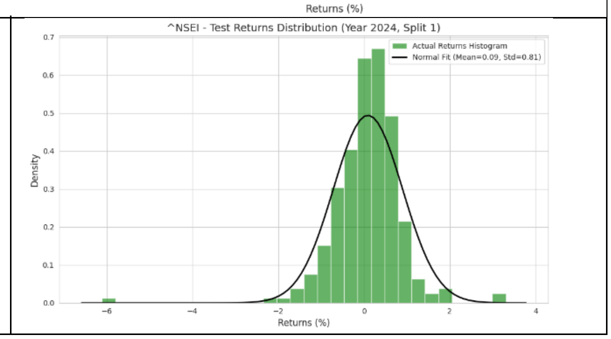

^GSPC - Test Returns Distribution (Year 2024, Split 1)
Actual Returns Histogram
Normal Fit (Mean=0.09, Std=0.71)
Density
Returns (%)

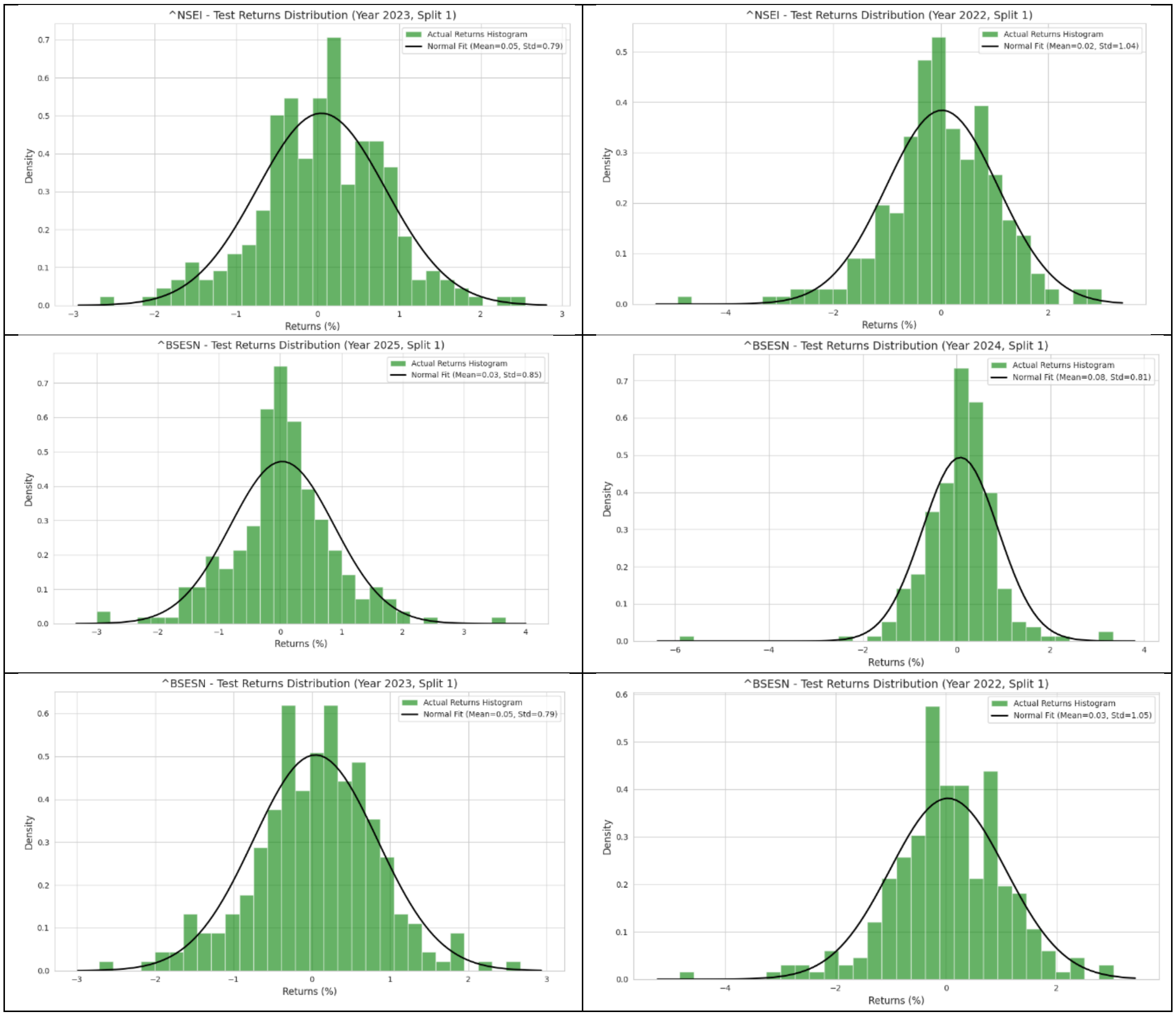

^GSPC - Test Returns Distribution (Year 2023, Split 1)
Actual Returns Histogram
Normal Fit (Mean=0.06, Std=1.24)
Density
Returns (%)


^GSPC - Test Returns Distribution (Year 2022, Split 1)
Actual Returns Histogram
Normal Fit (Mean=-0.05, Std=1.23)
Density
Returns (%)


^STOXX50E - Test Returns Distribution (Year 2025, Split 1)
Actual Returns Histogram
Normal Fit (Mean=0.03, Std=1.13)
Density
Returns (%)


^STOXX50E - Test Returns Distribution (Year 2024, Split 1)
Actual Returns Histogram
Normal Fit (Mean=0.06, Std=0.80)
Density
Returns (%)


^STOXX50E - Test Returns Distribution (Year 2023, Split 1)
Actual Returns Histogram
Normal Fit (Mean=0.09, Std=1.10)
Density
Returns (%)

^STOXX50E - Test Returns Distribution (Year 2022, Split 1)
Actual Returns Histogram
Normal Fit (Mean=-0.07, Std=1.42)
Density
Returns (%)

^HSI - Test Returns Distribution (Year 2025, Split 1)
Actual Returns Histogram
Normal Fit (Mean=0.12, Std=1.83)
Density
Returns (%)

^HSI - Test Returns Distribution (Year 2024, Split 1)
Actual Returns Histogram
Normal Fit (Mean=-0.03, Std=1.41)
Density
Returns (%)

^HSI - Test Returns Distribution (Year 2023, Split 1)
Actual Returns Histogram
Normal Fit (Mean=-0.04, Std=1.76)
Density
Returns (%)

^HSI - Test Returns Distribution (Year 2022, Split 1)
Actual Returns Histogram
Normal Fit (Mean=-0.12, Std=1.66)
Density
Returns (%)

^NSEI - Test Returns Distribution (Year 2025, Split 1)
Actual Returns Histogram
Normal Fit (Mean=0.03, Std=0.85)
Density
Returns (%)

^NSEI - Test Returns Distribution (Year 2024, Split 1)
Actual Returns Histogram
Normal Fit (Mean=0.09, Std=0.81)
Density
Returns (%)


^NSEI - Test Returns Distribution (Year 2023, Split 1)
Actual Returns Histogram
Normal Fit (Mean=0.05, Std=0.79)
Density
Returns (%)
^NSEI - Test Returns Distribution (Year 2022, Split 1)
Actual Returns Histogram
Normal Fit (Mean=0.02, Std=1.04)
Density
Returns (%)
^BSESN - Test Returns Distribution (Year 2025, Split 1)
Actual Returns Histogram
Normal Fit (Mean=0.03, Std=0.85)
Density
Returns (%)
^BSESN - Test Returns Distribution (Year 2024, Split 1)
Actual Returns Histogram
Normal Fit (Mean=0.08, Std=0.81)
Density
Returns (%)
^BSESN - Test Returns Distribution (Year 2023, Split 1)
Actual Returns Histogram
Normal Fit (Mean=0.05, Std=0.79)
Density
Returns (%)
^BSESN - Test Returns Distribution (Year 2022, Split 1)
Actual Returns Histogram
Normal Fit (Mean=0.03, Std=1.05)
Density
Returns (%)

APPENDIX D: PYTHON IMPLEMENTATION AND GOOGLE COLAB ENVIRONMENT

To ensure the reproducibility of the GPR-HS framework, the complete Python codebase and multivariate dataset are hosted on Google Colab. This repository contains the end-to-end pipeline used to train the model, generate backtesting and benchmarking results, superiority metrics, and visualizations presented in this study.

Google Colab Access Link:

https://colab.research.google.com/drive/1nrlSqmG10DNerNmEqGIh3EB9CcLWIgH9

Note: The environment is configured to execute the backtesting suite across all seven global indices across all 4 data splits. It includes the logic for the Aggressive Noise Initialization (ANI) and the SHAP interpretability analysis conducted on the Stressed VaR (SVaR) scenarios.