PORTFOLIO OPTIMIZATION BY MEAN-VARIANCE-KULLBACK-LEIBLER DIVERGENCE MEASURE USING THE AR-GJR-GARCH FILTRATION

Léon Mishindo Mbucici[1], John W. Muteba Mwamba[2], Jules Clement Mba[3]
School of Economics, University of Johannesburg, South Africa[1,2,3]
Email: julemba@gmail.com

Abstract:
The classical Markowitz Mean-Variance optimization framework remains foundational. Still, it is often criticized for its sensitivity to estimation error and assumption of normality, leading to poorly diversified and non-robust portfolios. We propose a novel hybrid framework that integrates the Kullback-Leibler divergence measure into the mean-variance objective, regularizing the solution towards an investor-defined target distribution. This paper presents an analytical approach to portfolio optimization by integrating the classical mean-variance framework with the Kullback-Leibler (KL) divergence measure. While the mean-variance method, pioneered by Markowitz, seeks to balance expected return and risk (as measured by variance), it often assumes perfect knowledge of asset return distributions. We utilize AR-GJR-GARCH models to estimate and forecast volatility on all asset returns. To address model uncertainty and distributional robustness, we investigate the KL divergence as a regularization term, penalizing deviations from a reference distribution. This fusion results in a robust optimization framework that accounts for uncertainty in the estimated parameters. We derive closed-form solutions under certain assumptions and explore the impact of the divergence parameter on the efficient frontier. The proposed method enhances stability and reliability in portfolio allocation, particularly in data-scarce or high-volatility environments. Empirical results across global markets show that our Mean-KL (M-KL) model achieves superior diversification and higher absolute returns, though with higher volatility, demonstrating a compelling trade-off for target-oriented investors.

Keywords: Portfolio optimization, Mean-variance, Kullback-Leibler, Divergence

INTRODUCTION

Portfolio optimization has long been a central theme in modern finance and investment theory. Since the seminal work of Markowitz (1952), the mean–variance optimization (MVO) framework has served as the cornerstone for constructing efficient portfolios that balance expected return and risk. The model assumes that investors are rational and risk-averse, seeking to maximize expected returns for a given level of variance, or equivalently, to minimize risk for a desired level of return. Despite its theoretical elegance and practical appeal, the classical mean–variance approach is built on restrictive assumptions—most notably, that asset returns are normally distributed, and risk is fully captured by variance. Empirical evidence, however, consistently reveals that financial returns display non-normal characteristics, such as asymmetry, heavy tails, and volatility clustering (see Cont, 2001). These features violate the foundational premises of the MVO model and often lead to portfolios that are suboptimal or unstable in real-world conditions.

To address these limitations, recent literature has explored the use of conditional volatility models, particularly the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) family, to model time-varying risk. Among these, the Glosten–Jagannathan–Runkle GARCH (GJR-GARCH) model (Glosten et al., 1993) extends the standard GARCH framework by incorporating asymmetric responses of volatility to positive and negative shocks, a phenomenon known as the leverage effect. When coupled with autoregressive (AR) components, the AR–GJR–GARCH model effectively captures both mean persistence and volatility asymmetry in financial return series. This filtration process allows for a more realistic and dynamic estimation of the conditional mean and variance, thus providing a richer input for portfolio optimization.

Parallel to these developments in econometric modeling, information-theoretic measures have gained increasing attention in financial optimization. The Kullback–Leibler (KL) divergence, in particular, provides a measure of the discrepancy between two probability distributions (Kullback & Leibler, 1951). In the context of portfolio theory, the KL divergence can be used to quantify model uncertainty, regularise optimization problems, and penalize deviations from reference distributions. Integrating KL divergence into mean–variance optimization introduces an additional layer of robustness, allowing investors to control the sensitivity of optimal portfolios to estimation errors and distributional changes.

However, despite these conceptual advancements, the integration of the KL divergence measure within a mean–variance framework filtered through AR–GJR–GARCH dynamics remains largely unexplored. Existing studies tend to examine these components in isolation, either focusing on volatility modeling through GARCH-type processes or incorporating information divergence in static optimization setups. This separation overlooks the potential synergies that arise when information-theoretic measures are used to adjust portfolio weights dynamically based on filtered conditional distributions of returns. Furthermore, the interaction between distributional robustness (via KL divergence) and conditional heteroskedasticity (via AR–GJR–GARCH) is yet to be systematically investigated.

Another notable gap of this study lies in the empirical validation and computational feasibility of such hybrid models. Implementing KL-based mean–variance optimization using AR–GJR–GARCH filtered returns entails high-dimensional estimation and optimization procedures. The challenge of parameter uncertainty, data frequency, and sample window selection further complicates practical deployment. Few studies have rigorously compared the out-of-sample performance, risk-adjusted returns, or stability of such models against traditional benchmarks.

Hence, there exists a significant research gap in developing and testing an integrated framework that combines the mean–variance paradigm, Kullback–Leibler divergence measure, and AR–GJR–GARCH filtration. Addressing this gap can enhance portfolio construction methods by accounting for time-varying volatility, asymmetric risk behavior, and model uncertainty, ultimately producing more robust and adaptive portfolio strategies suitable for complex financial markets.

In this paper, the Kullback-Leibler divergence measure is incorporated as a penalty factor into the objective function to obtain an optimal portfolio according to the analytical approach, giving an exact solution. The Kullback-Leibler divergence enriches portfolio optimization by providing a flexible, informative metric that accommodates complex return distributions and emphasizes risk assessment. The use of the Kullback-Leibler divergence allows the preference for one portfolio or another to be assessed relative to a targeted return distribution. This is therefore a comparative approach adopted in particular by Clark et al. (2022) and Lassance and Vrins (2023). Its advantages can significantly enhance portfolio decision-making and risk management in dynamic market environments. By integrating Kullback-Leibler divergence into the optimization process, investors can achieve more robust and adaptive portfolio strategies. The essential contributions of this work are: – The Kullback-Leibler divergence serves as a performance metric for evaluating how closely a portfolio's return distribution aligns with desired outcomes or benchmarks by measuring the divergence from a benchmark or target distribution – Incorporated into the objective function, Kullback-Leibler divergence plays the role of a regularization term. This will help to avoid overfitting by penalizing portfolios that deviate too much from a target distribution, thus promoting robustness in the investment strategy, this measure allows for greater diversification in portfolio construction – Choosing a reference portfolio simplifies the optimization problem with several constraints where the weights are fixed in the lower and upper bounds – We propose an analytical analysis providing the exact solution of mean-variance-Kullback-Leibler divergence model (MVKLD). Masri et al. (2010) were the first to introduce simultaneous constraints for the upper and lower bounds of the weights to be invested in each security.

Our study differs from the minimum extended Kullback-Leibler (MEKL) model due to Lassane and Vrins (2023) by considering discrete probability distributions; in addition, the target distribution is the result of simulations of possible distributions that can reflect investors' preferences.

However, it is not obvious to consider that the Markowitz model is entirely obsolete. This is why we incorporate the first two moments of return into our model so that, by minimizing the divergence between the target distribution and the optimal distribution, we can at the same time maximize the mean portfolio return and minimize its variance. Markowitz portfolio model still provides a fundamental framework for understanding the trade-off between risk and return. For example, Sinon and Mba (2024) added entropy to the MAD model to better diversify the weights of assets in the portfolios. That's why this paper aims to optimize both the mean and variance of the portfolio as well as the Kullback-Leibler divergence measure.

Problem Statement

The traditional Markowitz mean-variance optimization framework, while foundational in portfolio theory, faces limitations due to its reliance on normally distributed returns and historical data. These assumptions often fail to capture the complex behaviors of asset returns, such as skewness and kurtosis, leading to suboptimal diversification and heightened sensitivity to estimation errors. Additionally, the model's concentration on high-return, low-risk assets undermines diversification benefits. Despite advancements in fitting mean and volatility returns of assets and incorporating entropy and alternative risk measures, there remains a need for a robust approach that integrates investor preferences for target return distributions while balancing mean returns and variance.

Main Objective

This paper aims to develop and empirically validate a robust portfolio optimization model (MV-KL) that simultaneously optimizes for expected return, variance, and proximity to a flexible, investor-specific target return distribution via Kullback-Leibler divergence.

Specific Objectives:

1. Integrate Kullback-Leibler (KL) Divergence: Develop a portfolio optimization model that incorporates KL divergence to measure alignment with a target return distribution, addressing non-normality and enhancing diversification.
2. Hybrid Optimization: Combine mean-variance principles with KL divergence to create a multi-objective framework (MV-KL) that optimizes return, risk, and distributional proximity.
3. Empirical Validation: Compare the performance of the MV-KL model against traditional models (e.g., minimum variance, equally weighted) using real-world datasets, evaluating metrics like Sharpe ratio and diversification.
4. Flexibility in Target Distributions: Explore the impact of different target distributions (e.g., kernel density estimates) on portfolio outcomes to reflect investor preferences. Python software gives the better Target Distribution according to the processed data.

Research questions:

1. How does the inclusion of KL divergence in the objective function improve portfolio diversification and robustness compared to traditional mean-variance models?
2. What is the analytical solution for the MV-KL model, and how does it balance trade-offs between return maximization, risk minimization, and target distribution alignment?
3. How do portfolios optimized under the MV-KL framework perform empirically in terms of risk-adjusted returns (e.g., Sharpe ratio, Sortino ratio, VaR and CVaR) across different market regions (e.g., African vs. global markets)?
4. What role does the choice of target distribution play in the MV-KL model's effectiveness, and how can it be tailored to investor-specific preferences?
5. Under what conditions does the MV-KL model converge to traditional models (e.g., mean-variance when KL aversion is zero), and how does this inform practical implementation?
6. The paper is organized as follows. Section 2 presents the relevant literature review. A theoretical framework and a conceptual model based on the mean-variance-Kullback-Leibler (MV-KL) model are presented in Section 3. In Section 4, we proceed with an empirical study. Section 5 presents the results and the interpretation. Section 6 concludes the paper.

LITERATURE REVIEW

Risk measures in portfolio optimization

Portfolio optimization is the process of selecting the best asset distribution from a set of assets to fulfill an objective (Reilly & Brown, 2012). Markowitz's mean–variance portfolio optimizes a trade-off between expected return and risk measured by the variance. Martinez-Nieto et al. (2021) argue that although the mean-variance model has been accepted and approved by researchers, this model nevertheless presents two major limitations, namely: sensitivity to input errors and concentration. In fact, Britten-Jones (1999) observes that the investor obtains portfolios with high weights in overvalued assets and low weights in undervalued assets because the model does not take into consideration factors such as subjective investor knowledge or specific market performance as inputs to the system. According to Black & Litterman (1992), the asset weights are very sensitive to estimation errors, thus causing large changes in portfolio securities via slight changes in the values of the return and risk estimators.

Schmidt (2019) indicates that securities with high expected returns will be overweighted and thus lose the benefit of diversification, which optimization is supposed to provide; therefore, the mean-variance optimized portfolios are highly concentrated on a few assets with high return and low risk.

However, recent studies in the field of portfolio optimization have shown that the objective function based only on the first two moments (mean and variance) does not capture enough information about the portfolio; see Chunhachinda et al. (1997); Harvey et al. (2010); Usta and Kantar (2011) and Le Courtois and Xu (2023).

However, the variance and volatility are no longer the only criteria for measuring risk. A variety of alternative and more sophisticated measures have been proposed over the past seven decades, such as downside risk, semi-variance, value-at-risk, conditional value-at-risk, expected shortfall, drawdown (Palomar, 2025), and the Shannon information measure, and how to incorporate such measures in the function of the portfolio to be optimized.

The conditional value at risk (CVaR) is a well-known coherent risk measure introduced by Rockafellar and Uryasev in 2000. Lleo (2010) proved that the CVaR is superior to Value-at-Risk (VaR) because CVaR quantifies tail risk and has been shown to be subadditive.

Ahmadi-Javid and Fallh-Tafti (2019) suggest the entropic value-at-risk (EVaR) as a new coherent risk measure, which is an upper bound for both the value-at-risk (VaR) and conditional value-at-risk (CVaR). The result is that the EVaR approach outperforms the others as the sample size increases. The comparison of the portfolios obtained for a real case by the EVaR and CVaR approaches shows that the EVaR-based portfolios can have better best, mean, and worst return rates as well as Sharpe ratios.

AR-GJR-GARCH Model

AR-GJR-GARCH refers to an Autoregressive (AR) model combined with a GJR-GARCH model, a statistical technique used in econometrics to model financial time series data. The AR component captures temporal dependencies in the mean of the time series. In contrast, the GJR-GARCH (Glosten, Jagannathan, and Runkle GARCH) model specifically captures asymmetric volatility, meaning it accounts for the different impacts of positive and negative news or shocks on future volatility. The GARCH models and their extensions have been widely used in financial econometrics to model the conditional volatility dynamics of financial returns. The paper utilizes a conditional extreme value theory (EVT) based model that combines the GJR-GARCH model, which takes into account the asymmetric shocks in time-varying volatility observed in financial return series, and EVT, which focuses on modeling the tail distribution to estimate extreme currency tail risk. The autoregressive conditional heteroscedastic model (ARCH Model) was first introduced by Engle (1982), and the extended ARCH Model and introduced the Generalized conditional Heteroscedastic model (GARCH Model). The GARCH Model was applied in the asset return model by Nelson (1991). Studying the relation between the Expected Value and the Volatility of the Nominal Excess Return on Stock, Glosten et al. (1993) extended the GARCH Model by combining it with the GJR Model and formed the GJR-GARCH Model. GARCH and GJR (Glosten-Jagannathan-Runkle) models are time-series models used to analyze and forecast financial market volatility, with GARCH being a symmetric model that treats positive and negative shocks equally. In contrast, the GJR model is an asymmetric extension that assigns greater impact to negative shocks due to the "leverage effect," making it more realistic for financial data. The GJR model adds a threshold component, activated by negative past returns, to capture this asymmetric reaction. In our study, we will add the Autoregressive (AR) Model to the GJR-GARCH Model to form the AR-GJR-GARCH Model. This model will better fit both the volatility and the mean return of assets.

Kullback-Leibler as risk measure

Bera & Park (2008) minimize the KL divergence between portfolio weights and target weights, subject to moment constraints. Glasserman and Xu (2013) used the Kullback-Leibler divergence measure in distributionally robust portfolio optimization, which maximizes worst-case performance assuming that the portfolio-return distribution lies in an ambiguity set at some prescribed distance from a nominal distribution.

Lassance and Vrins (2023) improve the higher moments of mean-variance-efficient portfolios by designing the target so that its first two moments match those of the chosen efficient portfolio but has more desirable higher moments. The authors show theoretically that the optimal portfolio is in general different from the mean-variance portfolio, but remains mean-variance efficient when asset returns are Gaussian. The Kullback-Leibler divergence measure has been used between the target-return density that captures the investor’s preferences and the return density. Then the optimal portfolio has a return density that is as close as possible to the target-return density. According to Jiang et al. (2018), the Kullback-Leibler divergence measure is a standard measure of divergence and is used in many financial applications.

The choice of a different target portfolio is the main challenge of this problem. Lassance and Vrins (2019) proposed an approach where the authors target a reference portfolio and then fix the first two moments of the investor’s target as the same as the reference portfolio. The first reference portfolio would be the mean-variance efficient portfolio, which could be the minimum variance portfolio or the one that maximizes the Sharpe ratio. Chalabi and Wurtz (2014) investigated the impact of divergence measures in portfolio optimization, where they present the φ divergence measure and explore the main advantages of using this divergence. The Rényi divergence was introduced by Rényi (1951) as a measure of information and depends on a parameter α called the order of the function. The Rényi divergence extends the Kullback-Leibler divergence in the sense that the limit of the Rényi divergence when $\alpha = 1$ equals the Kullback-Leibler divergence. Chamakh

(2021) explores different divergence measures between the portfolio-return and target-return distributions; his study proved that it is the Wasserstein distance and Stein discrepancy which deliver portfolios robust to misspecification about the distribution of returns.

Gaps

While Lassance and Vrins (2023) focus on continuous target distributions (Shift-Log-Normal, Gaussian), our approach utilizes a non-parametric kernel density estimate for the target, avoiding parametric assumptions and more directly reflecting empirical investor preferences. Furthermore, we derive analytical solutions under the MVKL framework, providing computational and theoretical advantages. In addition, we highlighted AR-GJR-GARCH forecasting models for asset returns in order to optimize the portfolio on the uncertain component of returns, which is an original approach for our study.

METHODOLOGY

AR-GJR-GARCH Model

The modeling of mean and volatility returns will precede the optimization framework. This procedure consists of capturing enough information on these returns in order to create a more realistic portfolio.

AR Model

This model is an extension of the GjR-GARCH model obtained by taking into account the portion of the return made in the previous moment. In fact, Glosten, Jagannathan, and Runkle extended the standard GARCH (generalized Autoregressive Conditional Heteroskedasticity) model that specifically addresses the asymmetry observed in financial time series. This model is formulated in the two equations below.

a) The mean equation : $r_t = \mu + \varepsilon_t$ (1)
b) The variance equation : $\sigma_t^2 = w + \alpha\varepsilon_{t-1}^2 + \gamma I_{t-1}\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$ (2)

where:

σ_t^2 : The conditional variance at time t.
w : A constant term, similar to the baseline variance in a standard GARCH model.
$\alpha\varepsilon_{t-1}^2$: The standard GARCH component, representing the impact of the squared previous shock.
$I_{t-1}\varepsilon_{t-1}^2$: The threshold component that accounts for the leverage effect :
I_{t-1} : Is an indicator function that equals 1 if the previous return (ε_{t-1}) is negative, and 0 otherwise.
γ : The leverage parameter, which scales the impact of squared negative residuals.
$\beta\sigma_{t-1}^2$: The persistence component, showing how past volatility influences current volatility.

The Leverage Effect in GJR-GARCH

i) Positive Shock: If ε_{t-1} is positive, then $I_{t-1} = 0$ and the term $\gamma I_{t-1}\varepsilon_{t-1}^2$ becomes zero. The effect on variance is simply a ε_{t-1}^2 .
ii) Negative Shock: If ε_{t-1} is negative, then $I_{t-1} = 1$ and the term $\gamma I_{t-1}\varepsilon_{t-1}^2$ is active. The total impact on variance is (α+γ) ε_{t-1}^2.
iii) Asymmetry: If $\gamma > 0$, negative returns will cause greater next-period variance than positive returns of the same absolute size, a phenomenon known as the leverage effect.

σ_t^2 : is the conditional variance at time t
ε_t : is the residue calculated at time t

The AR-GJR-GARCH (1, 1) model, presents the mean equation as an autoregressive process. The Akaike Information Criterion led us to choose the first-order autoregressive model for this study.

$r_t = \phi r_{t-1} + \mu + \varepsilon_t$ is the mean equation in the AR-GJR-GARCH model.

In this model, we have the following parameters:

- Phi (ϕ) : expresses the temporal dependence between returns from one instant to the next. It measures the inertia or the memory of the process.
- Constant or the drift (μ) : is a constant, the autonomous portion of the return
- Omega (w) : is a constant, the autonomous portion of the variance
- Alpha (α) : is the coefficient for the lagged squared residuals (ε_{t-1}^2)
- Gamma (γ) : is the coefficient for the asymmetric term that captures the impact of negative shocks.
- Beta (β) : is the coefficient for the lagged conditional variance (σ_{t-1}^2)

The AR-GJR-GARCH allows for modeling the variance in the case when asymmetries are present in the time series and also to measure the memory of the process under study.

Divergence measures

In this work we introduce another way for portfolio optimization based on divergence measures. In this process, the investor needs the portfolio that is as close as possible to his target portfolio; this allows the investor to reduce a multitude of constraints in this target. In the literature, several divergence measures exist, we present below the most used in the context of portfolio optimization. Mahalanobis (1936) was the first to introduce the notion of distance between probability measures.

Given a probability distribution $Q \in \mathbb{R}^n$ with mean $\mu = (\mu_1, \mu_2, \dots, \mu_n)^T$ and covariance matrix Σ, the Mahalanobis distance of point $P = (x_1, x_2, \dots, x_n)^T$ from Q is defined by:

$$d_M(P, Q) = \sqrt{(P-\mu)^T \Sigma^{-1} (P-\mu)} \qquad (3)$$

In our context, we will consider a target-return distribution $V = (v_1, v_2, \dots, v_n)$ that captures the investor's preferences and the optimal portfolio has a return distribution

$W = (w_1, w_2, \dots, w_n)$ that is as close as possible to V and then the Kullback-Leibler divergence between W and V is defined as follows:

$$D_{KL}(w \,\|/ v) = \sum_{i=1}^{n} w_i log\left(\frac{w_i}{v_i}\right) = E_W\left[log\left(\frac{w}{v}\right)\right] \qquad (4)$$

where

$$\frac{W}{V} = \left(\frac{w_1}{v_1}, \frac{w_2}{v_2}, \dots, \frac{w_n}{v_n}\right) \qquad (5)$$

$D_{KL}(w \,\|/ v)$ is the amount of information lost when V is used to approximate W. The investor is therefore interested in obtaining the smallest possible loss from this information.

$D_{KL}(w \,\|/ v)$ is a standard divergence measure because $D_{KL}(w \,\|/ v) > 0 \; for \; all \; w \neq v$ and $D_{KL}(w \,\|/ v) = 0$ when **w = v**, Jiang et al; (2018) show its applications in the financial field.

Among the reasons for preferring the Kullback-Leibler measure, Lassance and Vrins (2023) observe that this measure accounts for higher moments and can capture investors' preferences beyond mean and variance. From this measure, many others have been constructed, such as the Reny (1951), the Jensen-Shannon and the jensen divergence measure. They are respectively defined as follows:

Determining the target distribution: Theoretical framework and hypotheses

The major problem of portfolio optimization reflecting a given target distribution lies in the choice of this target. In the case of continuous distributions, Lassane and Vrins (2023) argue that the distributions commonly used in finance to model asset returns, such as the skew-Student distribution or the hyperbolic distribution, may not be those that investors would wish to target. Indeed, investors would typically target distributions with positive skewness and thin tails, which is at odds with the characteristics of asset returns. These authors propose two classes of target distributions, namely: the shift log normal and the Gaussian normal distributions, to solve the minimum extended Kullback-Leibler (MEKL) divergence problem. This approach is relevant for continuous distributions. These distributions are often characterized by a number of unknown parameters which will be estimated. Moreover, the more parameters there are, the more the target distribution is appreciated; this is the case of the Skew T-student distribution, which has four parameters. Note that the number of parameters corresponds to the number of constraints imposed by the investor. For example, when targeting a Gaussian distribution, it implicitly sets the skewness and kurtosis preferences to 0 and 3, respectively.

The Skew T-student distribution is defined by:

$$f(x;\mu,\sigma,\lambda,q) = \frac{\Gamma\left(\frac{1}{2}+q\right)}{v\sigma(\pi q)^{\frac{1}{2}}\Gamma(q)\left[1+\frac{|x-\mu+m|^2}{q(v\sigma)^2(1+\lambda sgn(x-\mu+m))^2}\right]^{\frac{1}{2}+q}} \quad (6)$$

Where

$$m = \lambda v\sigma \frac{2q^{\frac{1}{2}}\Gamma\left(q-\frac{1}{2}\right)}{\pi^{\frac{1}{2}}\Gamma(q)} \quad (7)$$

gives a mean of μ

$$v = \frac{1}{q^{\frac{1}{2}}\sqrt{(1+3\lambda^2)\frac{1}{2q-2}-\frac{4\lambda^2}{\pi}\left(\frac{\Gamma\left(q-\frac{1}{2}\right)}{\Gamma(q)}\right)^2}} \quad (8)$$

gives a variance of σ^2

In practice, there are several formulations of Skew T-student distributions.

In the context where we only have observations from a sample, the probability distribution can be schematically represented by a histogram, in this case the characteristic parameters of this distribution are not known. The Kernel function are used in this case to construct these histograms. That's why, in this study, the target distribution will be a Kernel distribution because we want to avoid making assumptions about the distribution of data. The Kernel density estimator's formula is given by:

$$\hat{f}_{nh}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-x_i}{h}\right) \quad (9)$$

Where X_1, X_2, …, X_n are random sample for an unknown distribution, n is the sample size, K(.) is the Kernel smoothing function, and h is the bandwidth.

The Python software, we will use here, automatically generates a target distribution adapted to the processed data.

Portfolio optimization model based on the Kullback-Leibler measure

Problem statement

In this section, we will present models that incorporate the Kullback-Leibler divergence as the performance measure of the portfolio optimization problem. The aim of this incorporation is to improve results concerning the weight allocation. We will analyze the simple mean-Kullback-Leibler divergence and the mean-variance-Kullback-Leibler models for portfolio optimization.

We introduce a new way for portfolio selection based on divergence measures. The framework introduces a portfolio selection process where the investor seeks the portfolio that is as close as possible to his ideal portfolio. This method allows an investor to define various constraints in the portfolio selection and to optimize a single objective function.

Analytical Approximation and Numerical Solution of the MV-KL Problem

In this section we derive a controlled first–order approximation for the Mean

Variance-Kullback-Leibler (MV-KL) optimization problem, followed by a robust primal–dual Newton method for computing the exact solution. Throughout we denote by $r\epsilon\mathbb{R}^n$ the expected return vector, $\Sigma \epsilon \mathbb{R}^{n\times n}$ the positive semidefinite covariance matrix, $v \in \Delta^n$ a target distribution with strictly positive components, and ,$w \in \Delta^n$ the portfolio weights (where, $\Delta^n = \{w\epsilon\mathbb{R}^n\colon 1^T w = 1, w_i > 0\}$). We use ρ > 0 for risk aversion and κ > 0 for the KL–penalty parameter.

MV-KL Formulation and First-Order Conditions

We consider the scalarized MV-KL objective

$$\max_{w\epsilon\Delta^n}\left\{r^T w - \frac{\rho}{2} w^T \Sigma w - \kappa D_{KL}(w \parallel v)\right\} \quad (10)$$

where

$$D_{KL}(w \parallel v) = \sum_{i=1}^n w_i log\left(\frac{w_i}{v_i}\right) \quad (11)$$

The Lagrangian associeted with the equality constraint $\mathbf{1}^T w = 1$ is

$$\mathcal{L}(w,\lambda) = r^T w - \frac{\rho}{2} w^T \Sigma w - \kappa D_{KL}(w \parallel v) + \lambda(1 - \mathbf{1}^T w). \quad (12)$$

Assuming an interior solution $w_i > 0$, the first-order conditions are

$$r - \rho\Sigma w - \kappa(logw - logv + 1) - \lambda\mathbf{1}0, \quad \mathbf{1}^T w = 1 \quad (13)$$

where log$w$ is taken elementwise.

First-Order Taylor Linearization of the KL Term

Let $w^{(0)} \ \epsilon \ \Delta^n$ be a reference point for linearization. Using the Taylor expansion of each coordinate,

$$logw_i \approx logw_i^{(0)} + \frac{1}{w_i^{(0)}}\left(w_i - w_i^{(0)}\right) \quad (14)$$

we obtain the linear approximation

$$\log w \approx \log w^{(0)} + diag\left(1/w^{(0)}\right)\left(w - w^{(0)}\right) \quad (15)$$

Substituting into (13) yields the linearized FOC :

$$\left(\rho\Sigma + diag\left(1/w^{(0)}\right)\right) w = r - \kappa\left(\log w^{(0)} - \log v\right) - \lambda\mathbf{1} \quad (16)$$

Define

$$A \coloneqq \rho\Sigma + \kappa diag\left(1/w^{(0)}\ \right)$$

$$b \coloneqq r - \kappa\left(\log w^{(0)} - \log v\right). \quad (17)$$

Since A is positive (as Σ is PSD and $w^{(0)} > 0$), the libear system becomes

$$Aw = b - \lambda\mathbf{1}. \quad (18)$$

Imposing $\mathbf{1}^T w = 1$ gives the scalar equation

$$\lambda = \frac{\mathbf{1}^T A^{-1} b - 1}{\mathbf{1}^T A^{-1}\mathbf{1}} \quad (19)$$

Substituting (19) yields the closed-form approximate solution :

$$\boxed{w^{(approx)} = A^{-1}\left(b - \frac{\mathbf{1}^T A^{-1} b - 1}{\mathbf{1}^T A^{-1}\mathbf{1}}\mathbf{1}\right)} \quad (20)$$

This approximation is accurate when $w^{(0)}$ is close to the true optimum. It serves both as (i) an analytic insight into the structure of the MV-KL solution and (ii) an efficient warm start for numerical solvers.

Approximation Error Bound

For each coordinate, the Taylor remainder of log is

$$R_i = \log w_i - \left(\log w_i^{(0)} + \left(w_i - w_i^{(0)}\right)/w_i^{(0)}\right) = -\frac{\left(w_i - w_i^{(0)}\right)^2}{2\xi_i^2}, \quad (21)$$

for some ξ_i between w_i and $w_i^{(0)}$. Hence

$$|R_i| \leq \frac{\left(w_i - w_i^{(0)}\right)^2}{2\xi_i^2}, \qquad \underline{\xi} \coloneqq \min_i\left\{w_i, w_i^{(0)}\right\} \quad (22)$$

The resulting FOC error satisfies

$$\|FOC_{true} - FOC_{lin}\|_2 \leq \frac{\kappa}{2\xi^2}\left\|w - w^{(0)}\right\|_2^2 \quad (23)$$

so the approximation error is $0\left(\left\|w - w^{(0)}\right\|_2^2\right)$

Iterated Linearization

Iterating the approximation

$$w^{(k+1)} = w^{(approx)}\left(w^{(k)}\right) \quad (24)$$

yields a fixed-point refinement analogous to a Newton or Picard method. Starting from $w^{(0)} = v$ or $w^{(0)} = 1/n$

typically gives rapid convergence when the KL penalty κ is moderate or large.

Primal-Dual Newton Method (Exact Solution)

The exact KKL system is

$$g(w,\lambda) = \begin{bmatrix} r - \rho\Sigma w - \kappa(logw - logv + 1) - \lambda\mathbf{1} \\ \mathbf{1}^T w - 1 \end{bmatrix} = \mathbf{0} \tag{25}$$

The Jacobian is

$$J(w,\lambda) = \begin{bmatrix} -\rho\Sigma - \kappa diag(1/w) & -1 \\ \mathbf{1}^T & \mathbf{0} \end{bmatrix} \tag{26}$$

Let $(\Delta w, \Delta\lambda)$ solve

$$J(w,\lambda)\begin{bmatrix} \Delta w \\ \Delta\lambda \end{bmatrix} = -g(w,\lambda). \tag{27}$$

A damped Newton step is then performed:

$$w^+ = w + \alpha\Delta w, \qquad \lambda^+ = \lambda + \alpha\Delta\lambda \tag{28}$$

where $\alpha \epsilon (0\,,1]$ is chosen by backtracking to ensure $w^+ > 0$ and a sufficient decrease in $\|g\|$.

After the step, we enforce

$$w_i^+ \longleftarrow max\{w_i^+, \varepsilon\}, \quad w^+ \longleftarrow \frac{w^+}{\mathbf{1}^T w^+} \tag{29}$$

for a small floor $\varepsilon > 0$ (e.g. 10^{-12}).

A stopping criterion such as $\|g(w,\lambda)\|_\infty < 10^{-8}$ ensures numerical convergence.

Algorithmic Summary

Below is a pseudocode description of the primal-dual Newton method.

Algorithm : Primal–Dual Newton Method for MV–KL

Input: r, Σ, v, ρ, κ, tolerance tol

Initialize $w \longleftarrow w^{(0)} \epsilon \Delta^n$ (e.g. v or uniform), $\lambda \longleftarrow 0$;

for k = 0,1,2, . . . **do**

Compute gradient: $g_{w=r-\rho\Sigma w-\kappa(logw-logv+1)-\lambda\mathbf{1}}$;

Compute constraint residual: $g_{eq} = \mathbf{1}^T w - 1$;

if $\max\{|\,g_w\,|,\ |\,g_{eq}\,|\} <$ tol **then**

stop;

Form Hessian block: $H = -\rho\Sigma - \kappa diag(1/w)$;

Solve KKT system

$$\begin{bmatrix} H & -1 \\ \mathbf{1}^T & 0 \end{bmatrix}\begin{bmatrix} \Delta w \\ \Delta\lambda \end{bmatrix} = -\begin{bmatrix} g_w \\ g_{eq} \end{bmatrix}$$

Backtracking: choose $\alpha\epsilon\ (0,1]$ such that $w + \alpha\Delta w > 0$ and $\|g\|$ decreases;

Update $w \longleftarrow w + \alpha\Delta w$, $\lambda \longleftarrow \lambda + \alpha\Delta\lambda$;

Project $w_i \longleftarrow max\{w_i, \varepsilon\}$ and renormalize $w \longleftarrow w(\mathbf{1}^T w)$;

This method converges rapidly for moderate dimensions ($n \lesssim 500$) and provides an exact numerical solution to (10).

In practice, the investor's target depends on many parameters which cannot be limited to only the moments of the returns. To better assess the optimal portfolio weightings, we will set or simulate some targets reflecting the investor's preferences. For example, the case of weights equally distributed, the case of weights classified in the order of mean assets or classified in the inverse order of assets' variances, and so on.

Portfolio Performance Measures

Marhfor (2016) provides a review of the main measures of portfolio performance. This study shows that the conditional approach addresses one major shortcoming of the traditional approach based on the risk stability assumption. The author groups, on the one hand, the unconditional performance evaluation indicators (Sharpe ratio, Information ratio, Sortino ratio, Treynor ratio, Value at Risk measure, Jensen's Alpha,…, etc.) and, on the other hand, the conditional indicators (conditional Jensen's Alpha, conditional GARCH Volatility,…). In this section, some of these measures are presented and will be used to evaluate the performance of the portfolios obtained by applying the mean-Kullback-Leibler (M-KL) and the mean-variance-Kullback-Leibler (MV-KL) models compared to traditional optimization problems.

Portfolio performance measures

The Sharpe ratio (SR) was developed by Sharpe (1966), this ratio employs the standard deviation of the portfolio to measure risk rather than focusing solely on systematic risk. It's like having an extra magnifying glass to examine the investment, taking diversification into account. A higher Sharpe ratio (Sharpe, 1966) means the investment provides a better return for the amount of risk that is being taken, its formula is given as follows:

$$SR = \frac{E(R)}{\sqrt{V(R)}} \qquad (30)$$

where R is the return of the portfolio.

The Adjusted Sharpe ratio (ASR) was introduced by Pezeir (2004) to adjust for skewness and kurtosis by incorporating a penalty factor for negative skewness and excess kurtosis. It is defined as follows:

$$ASR = SR\left[1 + \frac{Sk(R)}{6}SR - \frac{(Kur(R)-3)}{24}SR^2\right] \quad (31)$$

The Sortino-Satchell ratio (SSR) is a performance measure based on the partial moments. It calculates the average active return divided by a lower partial moment of the portfolio return distribution. Its formula is as follows:

$$SSR = \frac{E(R)}{\sqrt{E[max(-R,0)^2]}} \qquad (32)$$

Where $E[max(-R,0)^2]$ is the lower partial moment of order 2.

Sortino ratio SR(T) with a threshold T is given by:

$$SR(T) = \frac{E(R)-T}{\sqrt{E_{R\le T}(T-R)^2}} \qquad (33)$$

If T=0 we onbtain the classical Sortino ratio (Sortino and Satchell, (2001).

The Value at Ris (VaR) is one well-known performance measure which represents the predicted maximum loss with a specified probability level over a given period of time. Given a confidence level $\alpha \in (0,1)$, the VaR of the portfolio at the confidence level α is given by:

$$VaR_\alpha(X) = inf\{\gamma \epsilon \mathbb{R}: P(X \ge \gamma) \le 1-\alpha\} \qquad (34)$$

The CVaR (conditional value at risk) is the expected loss that exceeds VaR and is an alternative to VaR that is more sensitive to the Sharpe of the loss distribution in the tail of the distribution. The CVaR for a given confidence level α is defined as follows:

$$CVaR_\alpha = \frac{1}{1-\alpha}\int_{f(x,h)\ge VaR_\alpha(x)}^{\infty} f(x,h)p(h)\,dh \qquad (35)$$

The traditional portfolio optimization problem

The standard portfolio optimization model known as the Markowitz' mean-variance portfolio optimization problem can be stated as follows.

Minimize the variance: Minimum variance

In this model, the weights of assets are obtained by minimizing only the variance-covariance matrix of the return of the portfolio. It is formulated as follows:

Minimize V(R)

s. t. $\sum_{j=1}^{n} w_j = 1$ and $w_j \ge 0 \ \forall j$

Minimize variance with a target expected return

Minimize V(R)

s. t. $\sum_{j=1}^{n} w_j r_j = m$

$\sum_{j=1}^{n} w_j = 1$ and $w_j \ge 0 \ \forall j$

Maximize the expected return considering a tolerance limit for the portfolio variance.

Maximize E(R)

s. t. $V(R) \le Va$

$\sum_{j=1}^{n} w_j = 1$ and $w_j \ge 0 \ \forall j$

Maximize expected return with risk aversion

$$\text{Maximize} E(R) - \frac{\rho}{2} V(R)$$

$$\text{s. t.} \sum_{j=1}^{n} w_j = 1 \text{ and } w_j \geq 0 \ \forall j$$

In the literature, we had the Equally Weighted model (EWM). This model considers the portfolio weights to be equal and does not involve any optimization. Moreover, Usta and Kantar (2011) argue that various studies in the literature show that the EWM works well for the out-of-sample cases.

$$w_j = \frac{1}{n} \text{ for } j = 1, 2, \dots, n$$

RESULT AND DISCUSSION

Empirical results and discussion

In this section, we give the descriptions of empirical datasets used in this study and present the results of the empirical study

Data description

The data is obtained from two websites as follows:

- Datasets from IRESS: FTSE-JSE ALSI, FTSE-JSE TOP 40, FTSE-JSE INDUSTRIALS, FTSE-JSE FINANCIALS, FTSE-JSE PRECIOUS METALS & MINING, HANG SENG INDEX- China, NIKKEI 225-Japan, CAC 40-France, FTSE 100-UK and DJI-US
- Datasets from investing.com; Nigeria Stock Exchange (NSE)-Nigeria, FTSE NSE Kenya 25-Kenya, EGX 30 – Egypt, KOSPI-Korea, DAX-Germany and S&P500-US.

These empirical datasets are considered in the empirical evaluation of the MKL and MVKL. The period of the datasets includes daily stock prices of assets from January 2010 to April 2025. The price at time t, such that P_t will be transformed in log returns R_t given by the following formula:

$$R_t = log\left(\frac{P_t}{P_{t-1}}\right) \quad (36)$$

It should be noted that these two datasets are adjusted and merged to harmonize the stock price observation dates in order to calculate the log returns. The following tables give the key statistics describing all of these data. We have divided the stock markets into two groups according to two geographical zones, the first comprising the markets operating in Africa and the second those located outside Africa.

On the African continent, we have selected the most important markets in terms of capitalization. There are the South African, Egyptian, Nigerian, and Kenyan markets. Regarding the South African market, the analysis was carried out in two stages, the first consisted of considering all the indices, both global such as FTSE-JSE ALSI and the indices specific to certain sectors of economic activity. The second stage retained only the specific indices, thus excluding the ALSI index.

Outside of Africa, three major regions were included in this study: Western Europe (FTSE100, CAC40, and DAX), The United States of America (NASDAQ100, S&P500, and DJI) and Asia-Pacific (HSIX-HANGSENG, KOSPI, and NIKKEI225).

This analysis is further developed by including all places in the same region.

Table 1. The key descriptive statistics

	mean	std	min	25%	50%	75%	max	Skewness	Kurtosis
CAC40	0,00015	0,01808	-0,41011	-0,00539	0,00064	0,00621	0,39860	-0,43060	228,620

	mean	std	min	25%	50%	75%	max	Skewness	Kurtosis
DAX	0,00031	0,02399	-0,69931	-0,00505	0,00080	0,00650	0,68681	-0,36863	441,416
DJI	0,00032	0,03347	-0,97893	-0,00348	0,00040	0,00497	0,98978	0,17068	568,212
EGX30 Hstorica Data	0,00041	0,07615	-2,29871	-0,00592	0,00072	0,00759	2,30273	0,09990	839,481
FTSE 100	0,00010	0,01108	-0,14174	-0,00437	0,00032	0,00498	0,14298	-0,37419	30,112
FTSE NSE Kenya 25	0,00033	0,03510	-0,69859	-0,00529	0,00033	0,00600	0,54469	-1,22348	112,201
FTSE_JSE ALSI	0,00027	0,03083	-0,75098	-0,00503	0,00054	0,00612	0,75194	-0,22299	357,596
FTSE-JSE FINANCIALS	0,00022	0,03934	-0,97655	-0,00592	0,00056	0,00682	0,96963	-0,24678	384,596
FTSE_JSE INDUSTRIALS	0,00007	0,06070	-2,30223	-0,00622	0,00021	0,00649	2,30877	0,07198	1074,905
FTSE_JSE OIL_GAS_Coal	0,00055	0,14847	-3,20659	-0,01242	0,00000	0,01264	3,19247	0,09493	288,457
FTSE-JSE Prcious Metals_M	0,00016	0,02865	-0,66393	-0,01062	0,00017	0,01051	0,65111	-0,09073	205,239
FTSE_JSE TOP 40	0,00028	0,05772	-2,29677	-0,00558	0,00059	0,00651	2,30637	0,14873	1295,017
HSIX_HANGSENG	0,00002	0,03646	-0,71459	-0,00649	0,00000	0,00659	0,69513	-1,55431	267,772
KOSPI	0,00010	0,01891	-0,63198	-0,00479	0,00041	0,00548	0,63000	-0,20207	634,993
NASDAQ100	0,00056	0,06803	-2,29730	-0,00440	0,00080	0,00701	2,30407	0,01388	747,945
NIKKEI 225	0,00029	0,02681	-0,76516	-0,00588	0,00008	0,00698	0,74304	-0,65255	413,412
NSE All Share_Nigeria	0,00041	0,27398	-3,24239	-0,00379	0,00014	0,00466	2,51881	-0,55347	74,798
S&P 500	0,00038	0,09537	-1,65401	-0,00384	0,00064	0,00570	1,66643	0,00173	138,464

Note. This table presents key descriptive statistics such as the mean, standard deviation, minimum, maximum, 25th, 50th, and 75th percentile, skewness and kurtosis of each of the 18 asset returns (CAC40, …, S&P 500).

To complete the description of the behavior of returns for these different assets, we present Figures No. 1 and No. 2 below, which highlight the major characteristics and outliers.

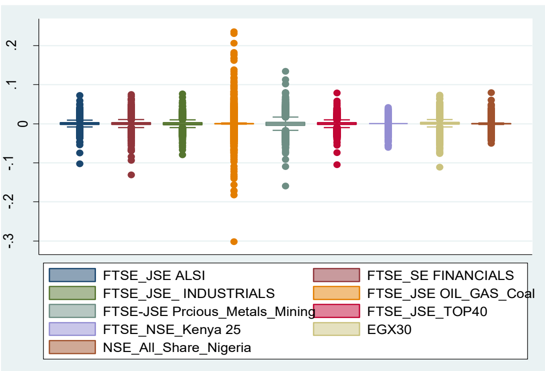

Figure 1. Box plot of the African's assets returns

Note. This figure shows how returns vary across assets observed in African stock markets. It clearly shows outliers that deviate significantly from other data in the same distribution.

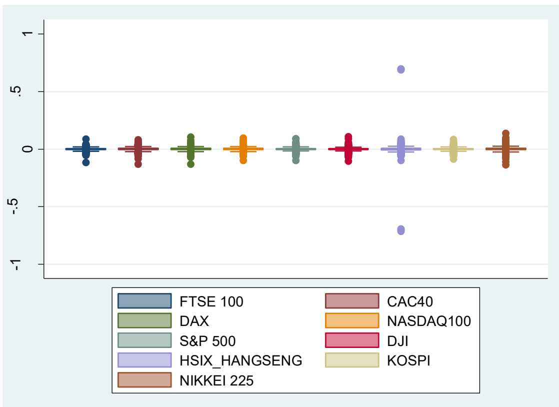

Figure 2. Box plot of the out-African's assets returns

Note. Similar to African stock markets, this figure also shows the presence of outliers in the distributions of returns in markets located outside Africa.

To correct the data, taking into account the observed aberrations (outliers), we first proceeded by fitting them to the AR-GJR-GARCH model, which has capacity to filter both the mean and variance components so that the optimization is ultimately focused only on the residual part.

Estimation of the AR-GJR-GARCH model

Below we present the results of the estimations of the AR-GJR-GARCH models by sector of activity or by asset grouped in the major geographical region. Calculations were performed using Stata 4.0 software.

Table 2. AR-GJR-GARCH parameters of Johannesburg Stock Exchange

South Africa: JSE						
Parameters	**FTSE-JSE-ALSI**		**FTSE-JSE-FINANCIALS**		**FTSE-JSE-INDUSTRIALS**	
	Coefficient	**P>z**	**Coefficient**	**P>z**	**Coefficient**	**P>z**
Phi	-0.443659	0,000	-.0370336	0,000	-0.4826523	0.000
Constant	0.0014524	0.041	.0028104	0,000	.000034	0.027
Omega	0.0005392	0.000	.0001271	0,000	.0000806	0.000
Alpha	0,0037722	0.568	12.19767	0,000	.5642084	0.000
Gamma	0,1952377	0.000	-9.88303	0,000	27.98133	0.000
Beta	-0.023159	0.503	.0011795	0,121	.1710691	0.000

Table 2. (Continued)

South Africa: JSE						
Parameters	**FTSE-OIL-GAS-COAL**		**FTSE-JSE-Precious-Metals-Mining**		**FTSE-JSE-TOP40**	
	Coefficient	**P>z**	**Coefficient**	**P>z**	**Coefficient**	**P>z**
Phi	-.04033329	0.000	-0,2727767	0,000	-0.481575	0.000
Constant	0.005572	0.046	-.0003912	0.284	.0003500	0.763
Omega	0.013585	0.000	.000021	0.000	.0016075	0.000
Alpha	0.102759	0.000	.0777194	0.000	.1227218	0.023
Gamma	0.135468	0.009	.1719998	0.000	-.0440618	0.506
Beta	-0.012894	0.000	.8531662	0.000	.0856956	0.411

Source: Authors' calculations based on Stata 4.0 software

Note. This Table presents on the row the estimates of each parameters of AR (1)-GJR-GARCH (1, 1) model, namely: Phi (the coefficient of lag 1 of return), the Constant of the AR (1) part; for the volatility part, we have Omega, Alpha, Gamma and Beta coefficients. In the column, we have the coefficient values and their p_value associated with the assets: FTSE-JSE-ALSI; FTSE-JSE-FINANCIALS; FTSE-JSE-INDUSTRIALS; FTSE-JSE-OIL-GAS-COAL; FTSE-JSE-Precious-Metals-Mining and FTSE-JSE-TOP40.

Table 3. AR-GJR-GARCH parameters on the Northeast Africa Stock Market

Northeast Africa						
Parameters	**FTSE NSE KENYA 25**		**EGX30**		**NSE All Share NIGERIA**	
	Coefficient	**P>z**	**Coefficient**	**P>z**	**Coefficient**	**P>z**
Phi	-0,332837	0,00	-0,4766116	0,000	-0,434660	0,000
Constant	-.0001451	0.003	-.0000517	0,949	-.0025886	0.076
Omega	.0001763	0,000	.0031449	0,000	.0000232	0.000

Parameters	Northeast Africa					
	FTSE NSE KENYA 25		**EGX30**		**NSE All Share NIGERIA**	
	Coefficient	**P>z**	**Coefficient**	**P>z**	**Coefficient**	**P>z**
Alpha	.9714100	0,000	.1287752	0,000	.0206668	0.000
Gamma	3.106513	0,000	.0004838	0,992	-.018737	0.000
Beta	.1168281	0,000	-.0188825	0,270	.9825417	0.000

Source: Authors' calculations based on Stata 4.0 software

Note. This Table presents on the row, the estimates of each parameters of AR (1)-GJR-GARCH (1, 1) model, namely: Phi (the coefficient of lag 1 of return), the Constant of the AR (1) part; for the volatility part we have Omega, Alpha, Gamma and Beta coefficients. In the column, we have the coefficient values and their p_value associated with the assets: FTSE, NSE KENYA25, EGX30, and NSE All Share NIGERIA.

Table 4. AR-GJR-GARCH parameters on the Western Europe Stock Market

Parameters	Western Europe					
	FTSE 100		**CAC 40**		**DAX**	
	Coefficient	**P>z**	**Coefficient**	**P>z**	**Coefficient**	**P>z**
Phi	-0,111688	0,000	-0,274361	0,000	-0,2825554	0,000
Constant	.0001700	0.247	.0007191	0,004	.0005775	0,018
Omega	9.26e-06	0,000	.0001125	0,000	.0002361	0,000
Alpha	.1770251	0,000	.400958	0,000	.1545798	0,000
Gamma	-.1285987	0,000	-.3416183	0,000	-.1154461	0,000
Beta	.8065994	0,000	.3941671	0,000	.3303556	0,000

Source: Authors' calculations based on Stata 4.0 software

Note. This Table presents on the row the estimates of each parameters of AR (1)-GJR-GARCH(1, 1) model, namely : Phi (the coefficient of lag 1 of return), the Constant of the AR(1) part ; for the volatility part we have Omega, Alpha, Gamma and Beta coefficients. In the column we have the coefficient values and their p_value associated for the assets FTSE100; CAC 40; DAX.

Table 5. AR-GJR-GARCH parameters on the United States of America Stock Market

Parameters	United States of America					
	NASDAQ 100		S & P 500		DJI	
	Coefficient	P>z	Coefficient	P>z	Coefficient	P>z
Phi	-0,4848593	0,000	-0,4919358	0,000	-0,4654284	0,000
Constant	.0010664	0.261	.0038207	0.038	.004404	0,000
Omega	.0024859	0.000	.0041710	0.000	.0001038	0,000
Alpha	.1361891	0.001	.0323813	0.008	15.23345	0,000
Gamma	.0152688	0.737	.6473376	0.000	-15.10985	0,000
Beta	-.014642	0.745	.1123504	0.004	.0031654	0,000

Source: Authors' calculations based on Stata 4.0 software

Note. This Table presents on the row the estimates of each parameters of AR (1)-GJR-GARCH (1, 1) model, namely : Phi (the coefficient of lag 1 of return), the Constant of the AR(1) part ; for the volatility part we have Omega, Alpha, Gamma and Beta coefficients. In the column we have the coefficient values and their p_value associated for the assets: NASDAQ 100; S&P 500; DJI

Table 6. AR-GJR-GARCH parameters on the Asia Pacific Stock Market

Parameters	Asia Pacific					
	HSIX-HANG SENG		KOSPI		NIKKEI 225	
	Coefficient	P>z	Coefficient	P>z	Coefficient	P>z
Phi	-0,2141867	0,000	-0,352905	0,000	-0,387813	0,000
Constant	.0008411	0.156	-.0005008	0.000	.001167	0.012
Omega	.0002552	0,000	.0000359	0.000	.0002806	0,000
Alpha	.0708871	0,000	.5296441	0.000	.3475915	0.000
Gamma	.0555494	0,000	1.674365	0.000	-.3161793	0,000
Beta	.5980517	0,000	.3090407	0.000	.3263242	0,000

Source: Authors' calculations based on Stata 4.0 software

Note. This Table presents on the row the estimates of each parameters of AR (1)-GJR-GARCH (1, 1) model, namely: Phi (the coefficient of lag 1 of return), the Constant of the AR (1) part; for the volatility part we have Omega, Alpha, Gamma and Beta coefficients. In the column we have the coefficient values and their p_value associated for the assets: HSIX-HANG SENG; KOSPI; NIKKEI 225.

This modeling step is crucial because it prepares for the optimization phase, which will be carried out on the residues from each of these models. This will make it possible to eliminate the predictable part of the returns and limit ourselves to the stochastic part.

The results of the empirical study

In this analysis, five models are estimated, three of which are classical portfolio optimization models, namely the Equally Weighted (EW), Minimum Variance (MinVar) and Mean-Variance strategies; and the models including the Kullback-Leibler (KL) penalty measure by associating the mean (M-KL), and associating the mean and variance (MV-KL).

The results were obtained by implementing the appropriate Python code. This gave the optimal values of risk aversion rho = 0.1 and the divergence risk aversion between the target portfolio and the optimal portfolio represented by the KL–penalty parameter $\kappa = 0.0001$.

Table 7. Optimal portfolio weights by strategies

Assets	EW	Variance min.	Moyenne–Variance	Moyenne–KL	Moyenne–Variance–KL
Cac40	0,055556	0.0839	0.1232	0.0560	0.0585
dax	0,055556	0.0671	0.1081	0.0560	0.0584
dji	0,055556	0.0563	0.0000	0.0555	0.0578
egx30	0,055556	0.0515	0.0765	0.0560	0.0585
ftse100	0,055556	0.1181	0.1517	0.0560	0.0586
nse_kenya25	0,055556	0.1082	0.1380	0.0560	0.0586
Alsi	0,055556	0.0696	0.1066	0.0560	0.0583
jse_financials	0,055556	0.0547	0.0914	0.0560	0.0582
jse_industrials	0,055556	0.0000	0.0000	0.0560	0.0206

Assets	EW	Variance min.	Moyenne–Variance	Moyenne–KL	Moyenne–Variance–KL
jse_oil_gas_coal	0,055556	0.0000	0.0000	0.0560	0.0561
jse_prcious Metals Mining	0,055556	0.0908	0.0000	0.0556	0.0582
jsetop40	0,055556	0.0335	0.0000	0.0548	0.0570
hsix_hangseng	0,055556	0.0926	0.0000	0.0529	0.0553
Kospi	0,055556	0.0997	0.1348	0.0560	0.0585
nasdaq100	0,055556	0.0000	0.0000	0.0534	0.0555
nikkei225	0,055556	0.0739	0.0687	0.0559	0.0583
nseallshare_nigeria	0,055556	0.0000	0.0000	0.0560	0.0553
sp500	0,055556	0.0000	0.0010	0.0560	0.0581

Note. The first column of this table contains the 18 Assets under study (CAC40, …, S&P 500); the other columns give the optimal weightings according to the optimization strategies EW, Minimum Variance, Mean-Variance, Mean-Kullback-Leibler, and the Mean-Variance-Kullback-Leibler

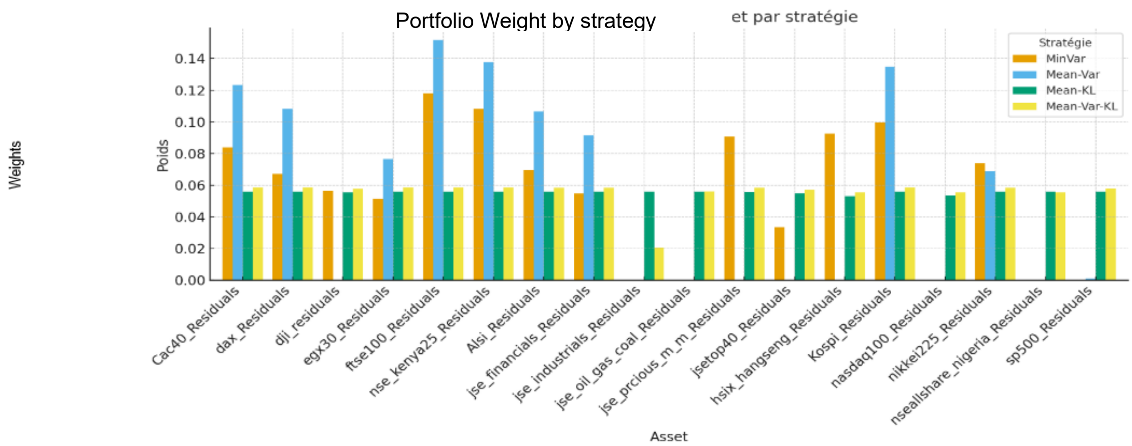

Figure 3 Portfolio Weight by strategy

The Figure 3 graph illustrates how the various optimization frameworks translate the underlying asset characteristics into distinct allocation patterns. The differences across strategies highlight their underlying assumptions about risk, return, and robustness.In Figure 3, we observe that the Mean–Variance-KL-divergence strategy occupies an intermediate position between return-seeking and diversification robustness. The graph shows that this approach tempers the concentration of the pure Mean-Variance solution while preserving some of its return-orientation. The resulting allocation structure generally appears more balanced, suggesting that the KL penalty helps control estimation error and weight instability while still allowing exploitation of expected return signals.

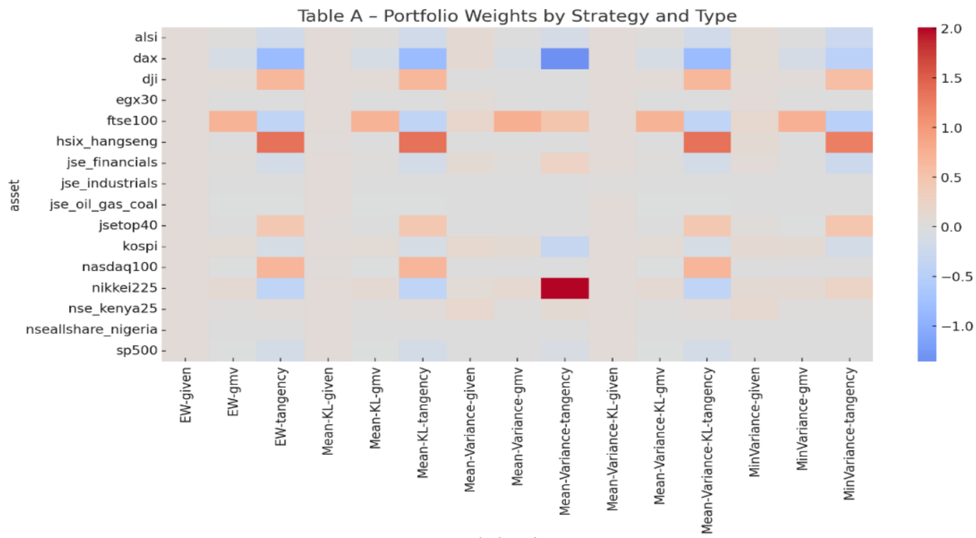

Figure 4. Portfolio weights by strategy and type

Figure 4 illustrates how different optimization frameworks
allocate capital across asset classes and categories, thereby revealing the theoretical priorities and risk–return assumptions embedded within each strategy, namely : Equal Weights, Minimum Variance, Mean-Variance, Mean-KL divergence and the Mean-Variance-KL divergence strategies.

This Figure 4 confirms the observations of the previous figure, that is to say that the Hybrid Mean-Variance-KL-divergence strategy occupies an intermediate position between aggressive return seeking and robust diversification. In the graph, the corresponding weight distribution is neither as concentrated as the Mean-Variance portfolio nor as uniform as the KL-divergence portfolio. Instead, it demonstrates controlled variability across asset types, indicating that the KL penalty effectively stabilizes the mean–variance solution by reducing sensitivity to input uncertainty while preserving its return-oriented structure. This balanced allocation pattern supports the theoretical rationale for combining classical optimization with distributional robustness techniques.

Figure 5. Cumulative Return

The cumulative-return graph (Figure 5) provides a dynamic view of how each portfolio strategy performs over time, capturing the compounding effect of gains and losses rather than relying solely on point-in-time statistics.

In our case, the cumulative-return graph underscores the fundamental trade-offs between risk, return, and robustness across portfolio strategies. While Minimum-Variance prioritizes stability and the Mean-Variance strategy emphasizes return, the KL-based and hybrid strategies deliver more balanced performance over time. The hybrid MV-KL portfolio often achieves the most attractive cumulative-return trajectory, supporting the literature on distributionally robust optimization, which argues that regularization can meaningfully improve out-of-sample performance by reducing sensitivity to estimation error. In summary, considering all these tables and figures, we conclude that :

Equal Weights (EW) strategy allocates the same weight to each asset, ensuring full diversification with no optimization; in this case, every asset receives an identical weight (0.0556), implying a completely diversified but non-optimized strategy. It serves as a neutral benchmark. But the Variance Minimum (MinVar) minimizes total portfolio variance; it focuses on risk reduction but not expected returns. Heavy allocations go to assets with historically lower volatility e.g., FTSE100 (0.1181), NSE_Kenya25 (0.1082), and KOSPI (0.0997). High-risk assets like JSE_Industrials, Nasdaq100, and NSE_AllShare_Nigeria receive zero weight. This suggests the portfolio strongly avoids risky or poorly correlated assets, leading to extreme concentration in stable markets.

We know that the Mean-Variance MV) balance expected return and variance; it is the fundamental Markowitz model. In this study, allocations are more spread but still concentrated in assets offering better risk-return trade-offs (e.g.,FTSE100,NSE_Kenya25). Some assets (Nasdaq100, JSE Industrials, JSE_Oil_Gas_Coal) are excluded, likely due to poor Sharpe contributions or negative covariance. Regarding the hybrid model offered here, the Mean-KL (M-KL) maximizes expected return while accounting for the information gain (or divergence) from a reference distribution using KL divergence. The inclusion of KL divergence enforces a smoother distribution ; almost all assets receive similar weights (~0.055–0.056). This shows that KL regularization penalizes extreme allocations, favoring diversification and distributional robustness. In addition, the Mean-Variance-KL (MV-KL) model combines traditional mean-variance optimization with KL divergence, seeking balance between performance, risk, and informational robustness. Similar to Mean–KL but slightly differentiated; minor variations (0.0206–0.0586) indicate that variance still influences weights but without major concentration. The portfolio achieves a balance between risk efficiency and robust diversification.

However, regarding the theory, we observe that the MinVar portfolio is highly concentrated in low-volatility assets; in our case, the portfolio exhibits the lowest volatility (0.0136), as expected. However, KL-based portfolios (Mean–KL, Mean–Variance–KL) have higher volatilities (0.1288 and 0.0555), indicating broader diversification and less concentration in low-risk assets.

Portfolio performance measures

In this section, we introduce the various portfolio performance measures and the rolling window procedure to evaluate the performance of the M-KL and MV-KL relative to the EW, MinVariance, and Mean-Variance Models. To justify the better performance of our model, including the Kullback-Leibler measure, we retained the following indicators: the Sharpe ratio, the Sortino ratio, the value at risk (VaR, 95%), and the conditional value at risk (CvaR, 95%).

The Sharpe ratio and its variant, the Sortino ratio, show the efficiency between return and risk for a given portfolio of assets, while the VaR and CvaR show the distribution of extreme losses for a portfolio of assets.

Table 8. Performance of Portfolio Strategies

Portfolio strategies	Expected Residual Return	Residual Volatility	Sharpe ratio (rf=0)	Sortino ratio	VaR 95 %	CVaR 95 %
EW	-0.000800	0.1278369	-0.0063	-0.0073	-0.02744	-0.1293
MinVariance	-0.000724	0.013631	-0.0531	-0.0662	-0.0164	-0.0341
mean-variance	-0.000016	0.014571	-0.0011	-0.0014	-0.0152	-0.0334
Mean–KL	-0.000773	0.128827	-0.0060	-0.0061	-0.0275	-0.1301
Mean–Variance–KL	-0.000806	0.055501	-0.0145	-0.0152	-0.0282	-0.0910

Note. This Table presents in columns the performance measures (Sharpe ratio, Sortino ratio, VaR at 95% and CVaR at 95% of confidence), the residual means and residual volatility for the five optimization strategies namely the EW, Min-Variance, Mean-Variance, Mean-KL and Mean-Variance-KL

.As indicated in this table, we observed that:

Sharpe and Sortino Ratios are negative due to small residues from fitting the return assets. However, Mean–Variance has the best (least negative) Sharpe and Sortino ratios (≈ -0.001), but Min Variance has the worst (-0.05), showing that extreme risk aversion sacrifices return. Value at Risk (VaR) and Conditional VaR (CVaR) indicate that, the Min Variance strategy again performs best with the lowest losses at 95% confidence (VaR = -0.0164; CVaR = -0.0341), moreover, KL-based portfolios (especially Mean–Variance–KL) show moderate tail risk improvement compared to EW and Mean–KL, reflecting better downside protection due to diversification.

In short, we conclude that, the minimum variance metric yields the lowest volatility, but the Markowitz mean-variance metric provides the most efficient and stable measure. The hybrid model, incorporating the Kullback-Leibler divergence measure and mean–variance returns, results in greater portfolio diversification and robustness. This is the main finding of this study.

The spider diagram below summarises the key information on the performance according to optimization strategies used in this study, namely: Sharpe ratio, Sortino ratio, VaR and CVaR at 95%.

Portfolio performance radar chart

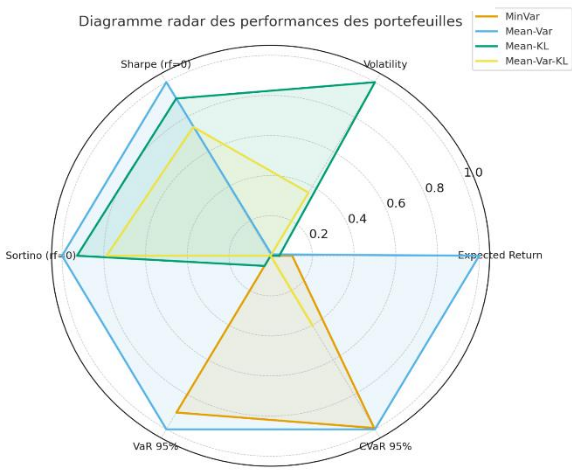

Figure 6. Portofolio performance radar chart

The radar chart synthesizes the strategies' performance across multiple dimensions-typically including realized return, volatility, drawdown, risk-adjusted performance, and possibly additional robustness metrics. Its role is to reveal the multidimensional trade-offs that are less visible when looking at single metrics in isolation.

Figure 6 shows that the hybrid Mean–Variance-KL-divergence strategy appears as one of the best-rounded performers. In this radar chart, its footprint often spans more area simultaneously across return, volatility, drawdown, and risk-adjusted performance metrics. This validates the theoretical premise that incorporating distributional robustness (via KL divergence) within a return-sensitive objective improves portfolio stability, reduces overfitting, and enhances multi-metric efficiency. The shape of this radar chart typically captures the ability of the hybrid strategy to balance performance and robustness more effectively than any of the single-criterion strategies.

Our study focuses on the returns of shares of companies listed on the stock exchanges of African, European, American, and Asian countries. Initial analyses show that:

When we look across the different markets, a few patterns start to stand out. Most of them show small but positive average returns, suggesting that, over time, they tend to move upward. Some markets, like the NASDAQ 100 and Egypt's EGX30, show slightly stronger average performance than others such as the FTSE100 or the CAC40, hinting at better growth potential.

Volatility, however, tells a different story. Emerging markets Nigeria and Egypt in particular are far more volatile than developed markets. This means their returns move around a lot more, making them riskier but also offering the possibility of bigger gains. In contrast, markets like the FTSE100 and CAC40 show much steadier behavior. The broader message is the familiar one: higher potential returns usually come with higher risk.

The shape of the return distribution adds another layer of insight. Several markets show negative skewness, meaning they are more prone to sudden, sharp declines than unexpected surges. Kenya, Hong Kong, and Japan fall into this category, which suggests that investors in these markets need to be prepared for occasional but severe downturns. A few markets display slight positive skewness such as the Dow Jones, indicating the possibility of rare but sizable gains. For investors who dislike large losses, positive skewness is generally more appealing.

Kurtosis brings the most striking results. All markets show very high values, well above what we would expect in a normal distribution. This tells us that extreme events, both positive and negative, occur far more often than standard models assume. South African indices like the JSE Top40 and Industrials are particularly affected, showing that large,unexpected shocks play a big role in shaping returns. Because of this, relying solely on traditional mean-variance models can be misleading, which is why we incorporate more robust measures such as the Kullback–Leibler divergence.

So, the picture is consistent: developed markets such as the S&P 500, DAX, and FTSE100 tend to be more stable, with lower volatility and fewer extreme movements. Emerging and sector-focused markets, on the other hand, carry more risk due to higher volatility, skewness, and kurtosis. These features reflect their exposure to economic shocks and structural vulnerabilities. The AR-GJR-GARCH models

Table 2 presents the parameter estimates of the AR-GJR-GARCH model carried out on the asset series for companies listed on JSE. In these models, the phi coefficients are all negative, which confirms the fact that the returns series evolve in a sawtooth pattern. It is in the industrial sector (phi = -0.4826), companies listed on JSE-Top 40 (phi = -0.4815) and on JSE-ALSI (phi = -0.4436) where the dependence of returns on time is greater.

In cases where drift, alpha, and omega are all positive, the models had better capture persistent volatility; this is the case for FTSE-JSE-INDUSTRIALS and FTSE-JSE Top 40. Furthermore, some coefficients are negative but are not significant in the models, this is the case for FTSE-JSE-ALSI (beta = -0.023) and FTSE-JSE-Precious-Metal-Mining (Constant = -0.0039). The significantly large and positive gamma coefficient for FTSE-JSE-INDUSTRIALS (Gamma = 27.98) means that negative shocks have a larger impact on volatility than positive shocks of the same magnitude. The same is true for FTSE-JSE-ALSI, FTSE-JSE-Precious-Metal-Mining, and FTSE-JSE-Oil-Gas-Coal, whose gammas are positive. On the other hand, the negative gamma indicates that negative shocks led to a decrease in volatility. This is the case for FTSE-JSE-FINANCIALS (Gamma = -9.88). Moreover, when the value of gamma is high as well as that of beta, the model is very able to capture the leverage effect, this is the case for FTSE-JSE-INDUSTRIALS. Models whose sum of alpha and beta is less than 1, such as FTSE-JSE-ALSI, FTSE-JSE-Oil-Gas-Coal, FTSE-JSE-Precious-Metal-Mining, and FTSE-JSE-TOP40, are stationary, and this sum reflects the persistence of the entire financial series. While the FTSE-JSE-FINANCIALS and FTSE-JSE-INDUSTRIAL series are not stationary and their trend often experiences sudden shocks.

In Table 3 the results are similar to those in Table 1. Indeed, for the Autoregressive (AR) part of the model the phi coefficients are all negative and between -0.5 and -0.3, that is to say that the series of returns evolves with consecutive peaks and troughs and is stationary. Concerning the volatility part or GJR-GARCH, the variance of the previous period is recovered to the current period from a positive beta; for example, for NSE-All Share NIGERIA, beta = 0.9825, and for NSE-KENYA 25, beta = 0.1168; for EGX30, beta is not significant. Taking into account the asymmetry of volatility, NSE-KENYA 25 gives the largest value of gamma = 3.1065, which shows that negative shocks are very important in the volatility of returns, but they are negligible in the EGX30 model, as they become negative in NSE-ALL Share NIGERIA.

In Table 4, all the gamma coefficients are negative; in other words, negative shocks reduce volatility. Furthermore, volatility and previous shocks have a positive impact on current volatility. A similar study conducted by Marisetty (2024) gave the following results for the FTSE 100 index for the GJR-GARCH model: constant = 0.0242; Omega = 0.0256; Alpha = 0.0495; Gamma = 1.0200 and beta = 0.8717. This author compares several variants of the GARCH model and it is the TGARCH (1, 1) model that performs well. The leverage effect is present in all the models since their respective Gammas are all negative; the volatilities of all three assets, FTSE100, CAC40, and DAX, are stationary because the sum of Alpha with Beta remains less than 1. This volatility is positively impacted by shocks and volatility at the time prior to the observation time. The AR-GJR-GARCH model is therefore effective in capturing both overall volatility and the return.

Table 5 confirms the opposite relationship between consecutive returns, the Phi coefficient is negative everywhere. Regarding volatility, it is stationary for the NASDAQ 100 and the S&P 500 (Alpha + Beta less than 1), but it diverges for the DJI. For the NASDAQ 100, current volatility negatively influences next volatility, while negative residuals positively affect next volatility. This proves the relevance of using the GJR-GARCH model, to which we associated the statistically significant AR part. The DJI nevertheless presents a particularity insofar as its Alpha=15.23345 and Gamma=15.10985 are excessive; these parameters mean that the current residuals have a greater impact of the order of 15.23345 times on the next volatility than the negative residuals negatively impact this volatility. Concerning the series of returns of S&P500, we estimated the parameters Phi=-0.4919; constant=0.003820; Omega=0.004171; Alpha=0.0323813; Gamma=0.6473376 and Beta=0.1123504. These estimates reaffirm that volatility is stationary, residuals at time t-1 have a positive impact on volatility, even negative residuals, and also prior volatility has a positive effect on current volatility. For his part, Marisetty (2024), estimating the GJR-GARCH (1, 1) model on the series of S&P500 returns from 2004 to 2023, found that the constant = 0.0550; Omega = 0.0196; Alpha = 0.0549; Gamma = 1.0049 and Beta = 0.8713. These results are close to ours insofar as they conclude in the same direction on the behavior of volatility as what we observed. Furthermore, Chen (2023) studied the series of returns of S&P500 from 2002 to 2020 by separating the periods before the financial crisis from those after this crisis and taking into account the period of COVID-19 and that without Cvid-19. The author was interested in comparing different variant methods of GARCH, he arrived at the results that it is the GjR-GARCH which gives the best results. However, his results do not corroborate with ours, because he obtains negative Gamma on all periods, which means that there is a leverage effect, and the sum of Alpha and Beta being in all cases greater than 1, the volatility is not stationary.

Table 6 gives the parameter estimates of the AR-GJR-GARCH model for the stock markets located in Asia Pacific. In this table, only the NIKKEI225 has a leverage effect because its Gamma is negative; all three series, HSIX-HANG SENG and NIKKEI 225, and KOSPI, are stationary, and therefore the behavior of their volatility over time is stable. According to Marisetty (2024) concerning the NIKKEI225 he obtained the following results: constant = 0.0520; Omega = 0.0623; Alpha = 0.0800; Gamma = 0.5425 and Beta = 0.8591. These results confirm that the volatility of NIKKEI225 is stationary but there is no leverage effect. The difference with our model lies in the introduction of the AR part in the model. The same study gives the following estimates of the HSIX-HANG SENG returns series: constant = 0.0337; Omega=0.0175; Alpha=0.0510; Gamma=0.3656 and Beta=0.9314. We observe that for this asset (HSIX-HANG SENG), we reach the conclusion, but the estimates differ because we added the AR part to the GJR-CGARCH model.

The optimal portfolio weights

Overall, integrating the Kullback-Leibler divergence measure into the optimization strategy resulted in more diversification and robust portfolios. Furthermore, measures such as minimum variance led to high concentrations of weightings on European assets like the CAC40, DAX, and FTSE. Also, the Sharpe ratio favors the traditional Markov metric, which balances the advantages of mean returns with the disadvantages of their volatility.

In, short, this study highlighted the benefit of incorporating more penalty elements into the utility function to better guide investors. In fact, MV-KL outperforms traditional models in diversification and performance metrics (Table 8), validating its robustness; the analytical solutions (see Equations 20, 24 and 29) show how KL divergence regularizes weights, mitigating overfitting. Finally, the model (MV-KL) adapts to non-normal returns and investor-defined targets, addressing the limitations of Markowitz's framework

We observe a paradox concerning portfolios from models including KL whose Sharpe ratios are low compared to portfolios from traditional models, and yet they are better than these traditional models. The reason is that the Sharpe ratio is not the only metric. For an investor whose primary goal is to track a specific target distribution (e.g., for ESG reasons, specific income needs, or regulatory constraints), the KL divergence is a more direct objective. The M-KL model excels in this and also provides better diversification than the highly concentrated Mean-Variance portfolio.

CONCLUSION

This paper introduced a novel Mean-Variance-Kullback-Leibler Divergence (MVKL) portfolio optimization framework. Our key contribution is the derivation of an analytical solution for this hybrid model and the empirical demonstration that its simplified variant, the M-KL model, effectively balances the pursuit of return with the adherence to a flexible, non-parametric target distribution, leading to superior diversification and drawdown characteristics compared to traditional concentrated portfolios. In fact, we observed, according to descriptive statistics, that companies registered in the metals and mining sector in South Africa showed a negative average return for this study period, unlike all other sectors. As for skewness, it varies depending on the sectors of activity in Africa. For example, it is positive in the oil and gas sector, in the mining sector in South Africa and in all companies listed on the Nigerian stock exchange. But it is negative in all the stock markets of countries in regions outside Africa. Moreover, kurtosis is increasing everywhere outside the African stock market, and in Africa it is high where skewness is positive. Being also a measure of risk, the increase in kurtosis means that investments in such sectors are much riskier. Thus, considering all these elements in terms of risk and profit, traditional portfolio optimization methods such as minimum portfolio variance, maximum average portfolio return to its variance and others are thus ineffective in providing a fair portfolio. This is why we propose a new approach based on a target portfolio taking into account the relevant constraints, on the optimization of a multi-objective model.

This paper therefore presents a multi-objective model that includes mean and variance of the portfolio, as well as the Kullback-Leibler divergence measure of portfolio weights, and compares its performance with traditional models, namely the minimum variance, the minimum variance with the target return, the mean-variance with risk-averse and equally weighted models.

This comparison is made using two different empirical datasets. As a result, we find that the M-KL performs better than the other considered models. This model gives results close to those of Markowitz mean-variance and also takes into account the investor's target. The portfolio is more diversified, which confirms the work of Usta et al. (2011).

In summary, we found that by combining the mean-variance model with the Kullback and Leibler measure, the model improves the portfolio well through its greater diversification. The results obtained indicate that it is more interesting to invest in the African market than elsewhere because we observe more favorable (positive) skewness and returns and lower kurtosis. In addition, the proposed model ensures greater portfolio flexibility considering the target portfolio and the level of aversion that the investor has to this target portfolio. Finally, the approach of an exact solution developed for this model greatly simplifies the calculations compared to the algorithmic or genetic approach.This study opens a new era in the study of investors' wishes in the development of investment models or strategies. We propose that future research in this area should focus on how to integrate psycho-sociological variables into portfolio optimization models. We propose that such research can first incorporate qualitative variables.

REFERENCES

Ahmadi-Javid, Amir and Fallh-Tafti, Malihe. (2019). Portfolio optimization with entropic value-at-risk, European Journal of Operational Research, Elsevier, vol. 279(1), pages 225-241.

Bera, A. and Park, S. (2008). Optimal Portfolio diversification using maximum entropy principle. Econometric Reviews, 27(4-6), 484-512.

Black, F. and Litterman, R. (1992). Global portfolio optimization. Financial analysis journal, 48(5), 28-43.

Britten-Jones, M. (1999). The sampling error in estimates of mean-variance efficient portfolio weights. The Journal of Finance, 5(2), 655-671.

Chalabi, Y. and Wurtz, D. (2014). Portfolio optimization based on divergence measures. ETH Econohysics Working and White Papers Series.

Chamakh, L. (2021). Quantifying uncertainty in asset management: Kernel methods and statistical fluctuations. Institut Polytechnique de Paris PhD, thesis.

Chunhachinda, P., Dandapani, K., Hamid, S. and Prakash, A. J. (1997). Portfolio selection and skewness: Evidence from internationalstock markets. Journal of Banking & Finance, 21(2), 143-167.

Clark, B., Edirisinghe, C. and Simaan M. (2022). Estimation risk and the implicit value of index-tracking. Quantitative Finance, 22(2), 303-319.

Daniel B. Nelson, (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach, Econometrica, Volume 59, Issue 2, 347-370.

Daniel P. Palomar; (2025). Portfolio Optimization, Theory and Application. The Hong Kong University of Science and Technology, CAMBRIDGE UNIVERSITY PRESS.

Engel, R. (1982). The Use of ARCH/GARCH Models in Applied Econometrics, Journal of Economic Perspectives, 15, 157-168. https://dx.doi.org/10.1257/jep.15.4.157.

Frank K. Reilly and Keith C. Brown (2012). Investment Analysis and Portfolio Management, 10th Edition.

Cenage Learning. 9780538482387.

Glasserman, P. and Xu, X. (2013). Robust risk measurement and model risk. Quantitative Finance, 14(1), 29-58.

Glosten, L.R., Jagannathan, R. and Runkle, D. (1993). On the relation between the Expected Value and the Volatility of Nominal Excess

Returns on Stocks, Journal of Finance, 48, 1779.1801.

Harvey, C. R., Liechty, J. C., Liechty, M. W., & Müller, P. (2010). Portfolio selection with higher moments. Quantitative Finance, 10(5), 469-485. https://doi.org/10.1016/j.ejor.2023.02 014

Ilhan Usta and Yeliz Mert Kantar. (2011). Mean-Variance-Skewness-Entropy Measures: A Multi-Objective Approach for Portfolio Seletion, Entropy 2011, 13, 117-133; doi: 10.3390/e13010117. ISSN 1099-4300; www.mdpi.com/journal/entropy.

Jiang, L., Wu, K., & Zhou, G. (2018). Asymmetry in stock comovements: An entropy approach. Journal of Financial and Quantitative Analysis, 53(4), 1479-1507.

Kullback S. and Leibler R. (1951) On information and sufficieny, Annals of Mathematical Statistics, vol. 22, p. 229-318.

Lassance, N. and Vrins, F. (2019). Portfolio selection with higher-order moments: A target-distribution approach. Paris, France. 9th General AMaMeF Conference. 3, 24, 31, 39, 50, 65, 69.

Lassance, N. and Vrins, F. (2023). Portfolio selection: A target-distribution approach. European Journal of operational Research, 310 (2023), 302-314.

Lleo, Sébastien. (2010). "Risk Management: A Review". In Risk Management, Foundations for a Changing Financial World. Charlottesville, VA: Researh Foundation of CFA Institute: 73-111.

Luisa Martinez-Nieto, Francisco Fernandez-Navarro, Mariano Carbonero-Ruz, and Teresa Montero-Romero. (2021). An experimental study on diversification in portfolio optimization. Expert Systems with Applications, Vol. 181, Pages 115203.

Mahalanobis, P. (1936). On divergences and informations in statistics and information theory. Proceedings of the National Institute of Sciences of India, 2:49–55.

Marhfor A. (2016). Portfolio performance measurement: Review of literature and avenues of future research. American Journal of Industrial and Business Management, 6(4), 432-438.

Markowitz, H. (1952). Portfolio selection. J. Finance, 7, 77-91.

Masri, H.; Abdelaziz, F.B. and Meftahi, I. (2010) A multiple Objective Stochastic portfolio selection program with partial information on probability distribution. In 2010 second international conference on computer and network technology, pp. 536-539, IEEE.

Olivier Le Courtois and Xia Xu. (2023). Efficient portfolios and extreme risks : a Pareto-Dirihlet approah, Annals of Operations Researh. https://doi.org/10.1007/s10479-023-05507-y.

Pézier, J. (2004). « Risk and Risk Aversion », in Alexander C. and Sheedy, E. (2004) eds : The Professional Risk Manager's Handbook. Comprehansive Guide to Current Theory and Best Practices, vol 3, PRIM Publications.

Rama, Cont, (2001). Empirical properties of asset returns: Stylized facts and statistical issues, Quantitative Finance, Volume 1: 223-236.

Rényi, A. (1951). On measures of entropy and information. Fourth Berkeley

Symposium on Mathematical Statistics and Probability, (1):547–561. 24, 26

Rockafellar, R. and Uryasev, S. (2000). Optimization of caonditional value-at-risk. Journal of Risk, 2(3), 21-41.

Schmidt, A.B. (2019). Managing portfolio diversity within the mean variance theory. Annals of operations Research, 28(1-2), 315-329.

Sharpe, W.F. (1966) Mutual fund performance. The Journal of Business, 39, 119-138. http://dx.doi.org/10.1086/294846.

Sinon, M. B. A., & Mba, J. C. (2024). The analysis of diversification properties of stablecoins through the Shannon entropy measure. Knowledge and Information Systems, 66(9), 5501-5540.