Optimizing the Magic Formula in Europe: Factors Driving Return, Risk and Risk-Adjusted Return

Hugo Santos[1], Catarina Ramos[2], Marta Faias[2], and Nuno C. Marques[1]

[1] NOVA LINCS, Universidade Nova de Lisboa, Portugal
nmm@fct.unl.pt
[2] DM and NOVA Math, Universidade Nova de Lisboa, Portugal

Abstract. Factor-based investment strategies provide a valuable approach for investors to allocate their resources effectively. However, stock markets are influenced by investor sentiment, leading to data scarcity relative to the vast universe of possible financial factors and their interactions. This paper presents an innovative study using AI-based techniques that optimizes and analyzes a well-known investment strategy called the Magic Formula (MF). To efficiently test the multitude of MF variants the techniques generate, we leverage on Qrumble, our recently developed Python framework for factor-based investment strategies. We employ Grid Search and Tabu Search algorithms to explore the space of MF variants and analyze how MF parameters and financial factors (used in ranking stocks) influence performance outcomes. Additionally, we conduct exploratory data analysis using the multiSOM, an advanced Self-Organizing Map, to identify the key factors driving return, risk, and risk-adjusted return. Our results reveal that factors used in ranking significantly outperform other MF parameters when optimizing the MF. Results from the exploratory factor analysis reveal interesting investment patterns prevalent in Europe during the sample period. Results, showcased here for the MF, shed light on the primary factors impacting diverse profiles for successful factor-based investment strategies. We believe our research provides a data-driven analysis that can offer valuable insights for refining investment strategies in the studied European financial sample data and beyond.

Keywords: Factor-based investment strategy · Local search optimization · Factor analysis · Self-organizing map.

1 Introduction

Factor-based investment strategies [2] drive numerous major investment funds worldwide. These investment strategies, some authors call factor-based, rank companies using specific criteria (e.g., a ranking formula) that incorporate financial factors derived from technical or fundamental analysis [19, 23]. They then invest in a number of top-ranked companies for medium to long-term periods. This process is repeated at set intervals over several years. However, analyzing the parameters—specifically which financial factors to use—that drive the performance of a factor-based investment strategy is challenging.

This study presents AI-based techniques to optimize and analyze factor-based investment strategies. For demonstration, we apply these techniques to the Magic Formula (MF) [11]—a renowned factor-based investment strategy developed by acclaimed investor and author Joel Greenblatt. A six-year financial sample consisting of the largest 600 European companies is used as a testbed, onto which more than 2,500 MF variants are run seeking to optimize key performance metrics of the MF.

Two major contributions result from the application of our AI-based techniques to the MF. The first contribution involves categorizing MF variants into two groups: those that modify the number of stocks being selected while maintaining the original ranking criteria, and those that keep the original stock numbers but alter the ranking criteria. This categorization and subsequent group analysis empirically demonstrates that modifying stock numbers alone improves MF performance only to a certain extent. A more effective approach, yielding significant performance improvements, involves adding more factors to the ranking formula and combining them in specific ways. The second contribution presents a method for effective factor analysis of factor-based investment strategies, training a Self-Organizing Map (SOM) [16] and identify crucial factors among the myriad of potential ranking formulas in the second category. The exploratory data analysis of the trained SOM reveals the key factors driving the metrics investors care most about: return, risk, and risk-adjusted return. While we focus on the MF as our target factor-based investment strategy for optimization and exploratory factor analysis, our AI-based techniques can be applied to other factor-based investment strategies beyond the MF.

2 Magic Formula

The Magic Formula (MF) [11] uses as the stock universe the largest 3,500 US stocks by market cap., excluding the financials and utilities sectors. From this universe, 30 stocks are selected by a 2-factor ranking system into an equal-weight portfolio and invested for one year. After one year, the portfolio is rebalanced: stocks currently in the portfolio are sold and 30 new stocks are selected for the equal-weight portfolio and invested again for one year, using the same ranking system. This process is repeated annually. Any gains or losses from the previous portfolio are reinvested into the next. The 30 stocks selected in each rebalance are those that score the highest according to the following ranking formula:

$$score = \mathrm{ROC}^{\langle + \rangle} + \mathrm{EarningsYield}^{\langle + \rangle}. \tag{1}$$

In this formula, stocks from the universe are scored using a 2-factor ranking system based on ROC and EarningsYield. ROC is the Return on Capital, which is the ratio between EBIT (Earnings Before Interest and Taxes) and Capital Employed, and EarningsYield is the ratio between EBIT and Enterprise Value. To calculate the score for a stock i ($i \in \{1, \ldots, U\}$, where U is the size of the stock universe), the following steps are applied:

1. Rank all U stocks based on their ROC ratios in ascending order: assign 1 (rank 1) to the stock with the lowest ROC, 2 to the stock with the 2nd lowest ROC, and so on, finally assigning U to the stock with the highest ROC; let i_{roc} be stock i's rank based on ROC;
2. Rank all U stocks based on their EarningsYield ratios in ascending order, using the same method; let i_{ey} be stock i's rank based on EarningsYield;
3. Sum both ranks: $i_{roc} + i_{ey}$, to yield a *score* for stock i ($2 \leq score \leq 2U$);
4. Repeat to get the score of every stock in the universe;
5. Select the top 30 stocks with the highest scores for investment.

The plus sign $^{\langle + \rangle}$ towering each factor indicates that ranking is done in *ascending* order, as used above. Conversely, if there was a term like $\text{ROC}^{\langle - \rangle}$ in the ranking formula, it would mean that stocks would be ranked by ROC in *descending* order. This reverses the above assignment, assigning 1 to the stock with the highest ROC, 2 to the stock with the 2nd highest ROC, and so on, finally assigning U to the stock with the lowest ROC.

To expand the ranking formula (1) beyond ROC and EarningsYield, we can include any sequence of k distinct factors:

$$score = F_1^{\langle +/- \rangle} + F_2^{\langle +/- \rangle} + \ldots + F_k^{\langle +/- \rangle}. \tag{2}$$

Here, k can be any number between 1 and K, where K represents the total number of factors under consideration for inclusion in the formula. Each of the k factors is assigned either an ascending or descending sorting order, based on whether higher or lower factor values should be given more weight in ranking the stocks.

3 Similar works employing the SOM for financial analysis

In previous research [7], the Self-Organizing Map (SOM) is used to analyze failures of small and medium-sized enterprises, providing insights without the need for complex financial manipulations. Additionally, the SOM is employed in risk analysis [5, 7], showcasing its potential for identifying companies likely to experience share price decreases within a year. Noteworthy applications of the SOM in finance include early studies by [7], where large datasets comprising 30 emerging markets are compressed into two-dimensional maps. These maps facilitate the analysis of investment opportunities and similarities between markets, aiding in asset allocation and benchmarking. However, the focus was primarily on relative volatility among emerging markets. In a more recent study [5], the SOM is used to cluster and visualize the temporal progression of financial indicators, assisting in predicting company development and bankruptcy risk. The dataset contained 29 financial ratios of over 110,000 companies from 2003 to 2006, with subsequent bankruptcy labeling. This study validates the SOM's effectiveness in financial analysis, especially in assessing bankruptcy risk. In another recent work [3], the authors employ the SOM for fundamental analysis of companies, uncovering new correlations among companies with similar fundamentals. The present work builds upon these foundations, extending the analysis to the importance of the distinct financial factors for investment strategy optimization and refinement.

4 Methodology

Fig. 1 outlines the methodology with the help of a diagram. In this diagram, data are represented by ovals and running code by rectangles. These are colored according to their functional roles: yellow for investment-related duties, pink for carrying out systematic search and green for analyzing and visualizing results. Two framed, miniaturized plots represent the results—essentially thumbnails of the full-sized plots included in the paper. Diagram components are briefly addressed below. Specific components are described in more detail in subsequent sections.

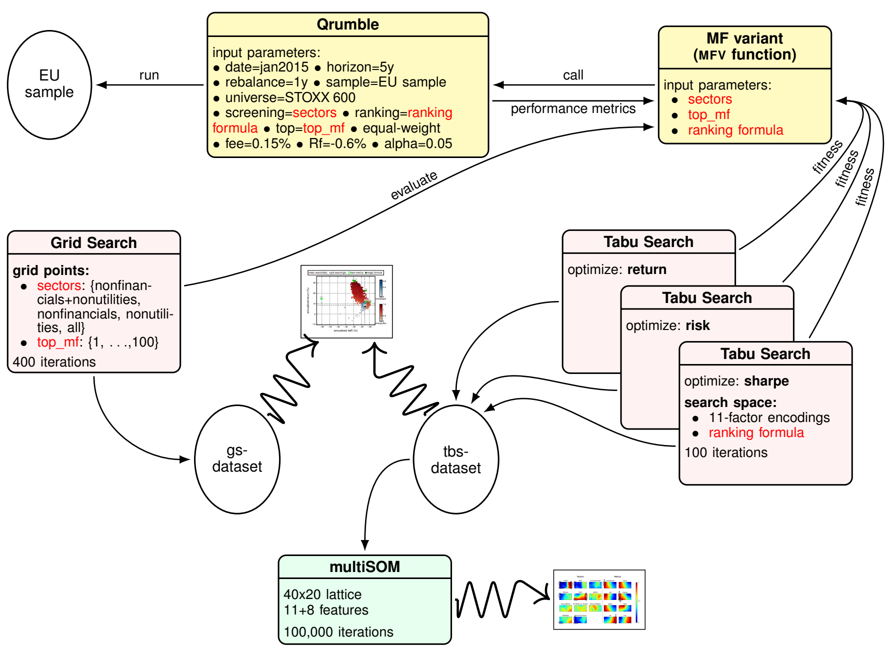

Fig. 1: Methodology overview.

MF variant is generated by breaking down the original MF into three parameters, as follows: 'sectors', 'top_mf' and 'ranking formula', highlighted in red throughout the diagram. The first two parameters affect the number of stocks chosen: 'sectors' expands the stock universe for selection, while 'top_mf' adjusts the number of stocks selected for investment. The 'ranking formula' modifies the criteria used to select stocks, incorporating additional factors beyond the two originally used.

The original MF operates based on fixed values of these parameters: 'sectors' (ordinal) is set to *nonfinancials+nonutilities*, 'top_mf' (number) is set to 30 and the 'ranking formula' (vector-encoded, described shortly) is defined by Formula (1). In this paper, we allowed parameters to vary, resulting in new MF-based investment strategies, which we refer to as MF variants. Each variant is identified by a unique combination of these three parameters. Except for any differences in the parameters, all other characteristics of the investment remain the same as the original MF. Each MF variant was evaluated in the same EU sample using Qrumble.

Qrumble is our Python framework for streamlined factor-based investing.[3] Its capacity to run a multitude of factor-based investment strategies [1] by adjusting a few input parameters, as illustrated in the diagram, made it an indispensable tool—without which this study would not have been possible. When tasked with running an investment strategy over the supplied sample, Qrumble calculates and delivers the corresponding performance metrics. Table 1 lists the eight performance metrics used by Qrumble to evaluate the MF variants in this study. For a detailed explanation of these metrics and their calculations, please refer to [12, 14], as well as the relevant section in Qrumble's documentation.

Table 1: Performance metrics used.

Metric	**Brief description**
annualized	Total return of the investment, annualized.
mean	Mean of the daily returns of the investment, scaled to an annual basis.
std	Standard deviation of the daily returns of the investment, scaled to an annual basis.
sharpe	Sharpe ratio of the investment, calculated on an annual basis, with -0.6% as the risk-free rate.
alpha	CAPM's α coefficient of the investment, with the STOXX 600 as the market reference.
beta	CAPM's β coefficient of the investment.
var	Value at Risk (VaR) of the investment, scaled to an annual basis, with a confidence level (α) of 0.05.
tvar	Tail Value at Risk (TVaR) of the investment, scaled to an annual basis, with the above confidence level.

MFV *function* is referred to as the *evaluation* (for Grid Search) or *fitness function* (for Tabu Search) in the three-part process of: 1) defining a specific MF variant

[3] Qrumble: A Python framework for streamlined factor-based investing, documented in `https://bit.ly/Qrumble`.

using a unique combination of three parameters, 2) running this variant over the EU sample, and 3) obtaining its performance metrics. This function enables a fair comparison of MF variants as they undergo systematic examination through one of the search methods.

EU sample is the European financial sample used to run the MF variants with Qrumble. It is based on the STOXX 600 index[4]. The original MF and its variants have been repurposed for this index, shifting the universe from the 3,500 largest US stocks to European companies. While the stock universe changed, we maintained all other characteristics of the MF, including the rebalancing period, equal-weight portfolios, fees and so forth. The financial sample covers the 2014–2019 period in Europe, and contains market data (daily) and fundamental data (half-yearly) for the companies listed in the STOXX 600 index across this 6-year period.

4.1 Grid Search

Grid Search [9] was used to explore 400 different MF variants of the first two parameters: 'sectors' and 'top_mf', while keeping the 'ranking formula' constant as per (1). These two parameters allow for defining varying *amounts of stock* that are invested by the MF-based strategy (MF variant), while keeping the criteria on how to select those stocks the same as in the original MF. The 'sectors' parameter (an ordinal variable), was allowed to vary between the more restricted universe of *nonfinancials+nonutilities* (like in the original MF, of around 400 European stocks on each rebalance, on average) to increasingly larger universe sizes of: *nonfinancials* (around 460 stocks), *nonutilities* (around 540 stocks) and *all* (around 600 stocks), respectively. The 'top_mf' parameter (a numeric variable), was permitted to vary within the range $\{1, \ldots, 100\}$, defining the number of top stocks invested in at each rebalance.

Grid Search went through 400 grid points formed by pairwise combinations of the two parameters. Each grid point corresponded to a unique MF variant, whose performance metrics were evaluated by the MFV function (Fig. 1). This function invokes Qrumble to execute the specific factor-based investment strategy over the common EU sample. Upon completion of Grid Search, all explored grid points (MF variants) were collected as datapoints into the gs-dataset. This dataset has 400 instances with 2+8 features having the following tabular structure: | sectors | top_mf | annualized | mean | std | sharpe | alpha | beta | var | tvar |. The first 2 features are the parameters that define one particular MF variant (out of 400), while the last 8 features are its performance metrics obtained from the sample.

4.2 Tabu Search

Tabu Search [10] was used to explore different variations of the 'ranking formula' while keeping the other two parameters, 'sectors' and 'top_mf', at their original

[4] `https://qontigo.com/index/SXXGR/`

values. Tabu Search was used instead of Grid Search due to the potentially vast search space involved, which grows exponentially with the number of factors K in (2). We conducted three distinct Tabu Searches, each optimizing a different performance metric through the fitness function (MFV function). To align with investor preferences, we aimed to optimize MF variants based on: 1) return, represented by the 'annualized' metric; 2) risk, represented by the 'var' metric (Value at Risk); and 3) risk-adjusted return, expressed by the 'sharpe' metric (Sharpe ratio). All metrics were annualized. For the actual implementation of Tabu Search, we reused the algorithm from [4]. Each Tabu Search started from the same *seed* (the MF itself), ran for 100 iterations (T=100), had a tabu size or tenure of 40 (τ=40) and used the same 11-factor base to pick factors for the ranking formulas (2). Parameter m is the performance metric that Tabu Search aims to optimize via three different calls to the same fitness function (MFV function in Fig. 1).

To explore different combinations of the ranking formulas through (2), we considered a total of eleven distinct factors (K=11). In addition to the two original MF factors used in ranking stocks—ROC and EarningsYield—we included nine more factors from other equally renowned factor-based investment strategies. While we could have included even more factors from other sources, we chose to limit our selection to 11 to demonstrate the techniques employed. Yield was taken from the "Dogs of the Dow" strategy [20], RS(6m) from 'Buying Winners and Selling Losers" strategy [13], and the remaining factors came from the "F-Score" strategy [21]. Table 2 lists all 11 factors, providing a brief description and their respective sources.

Tabu Search technicalities To perform a Tabu Search on ranking formulas across 11 factors based on Formula (2), we encoded each specific ranking formula as an 11-dimensional vector. Each dimension can have one of three values: $\{-1, 0, 1\}$. If factor F_i ($i \in \{1, \ldots, 11\}$) is present in a specific ranking formula, the i^{th} dimension of its vector is set to -1 if the factor is in descending order ($F_i^{\langle - \rangle}$), or 1 if it is in ascending order ($F_i^{\langle + \rangle}$). If factor F_i is absent in a specific ranking formula, the i^{th} dimension is set to 0. This encoding ensures a unique mapping between a ranking formula with k factors, each with a specific ascending/descending order, and its 11-dimensional vector representation. With 11 factors, the number of possible ranking formula combinations is 177,146 (i.e., $3^{11} - 1$). We subtract 1 because the all-zero vector is invalid (k must be at least 1). When doing local search-related tasks, the vector form of the ranking formula is used. When calling the fitness function, MFV, the ranking formula itself is converted back into (2). Starting with the *seed* (MF itself) in vector form, neighbors of any vector solution s are found by flipping a single value on one of its dimensions. The flip function performs this task: it changes the current value of s at dimension i to another value within the permitted range $\{-1, 0, 1\}$. The flip function must take care to not produce an all-zero vector solution as result. This produces 11 neighbor solutions in every iteration, from a given solution s. From this set of neighbors, any solutions that have since been designated as taboo

Table 2: Financial factors used in expanding the ranking formula of MF.

Factor	**Brief description**
ROC†	Return on Capital is used to measure the rate of return a business is making on its total capital. It is calculated as EBIT divided by Capital Employed (ttm+).
EarningsYield†	A measure of how much a company earns relative to its Enterprise Value. It is defined as EBIT divided by Enterprise Value (ttm).
Yield*	Stock's annual dividend payments to shareholders expressed as a percentage of the stock's current price.
RS(6m)††	The 6m Relative Strength measures a stock's price change over the last 6 months relative to the price change of a market index. It shows the relative outperformance or underperformance of the stock in that timeframe.
ROA‡	Return on Assets is a measure of how efficiently a company is using its assets to generate income. It is calculated by dividing a company's annual earnings by its average total assets (ttm).
ΔROA‡	Difference between current and last year's ROA (ttm).
AccrualRatio‡	A way to identify firms where Non-Cash or Accrual-Derived Earnings make up a significant proportion of Total Earnings. It is calculated as (Net Income - Free Cash Flow) divided by Total Assets (ttm).
ΔLTDetb-to-Assets‡	Difference between current and last year's Long Term Debt to Assets ratio. LTDetb-to-Assets is a measure of the level of the company's leverage.
ΔCurrentRatio‡	Difference between current and last year's Current Ratio. Current Ratio is a measure of the level of liquidity of a company. It is calculated as Total Current Assets divided by Total Current Liabilities (ttm).
ΔOpMgn‡	Difference between current and last year's Operating Profit Margin. OpMgn is a measure of how much income a company has left after paying its Operating Costs such as Rent and Salaries. It is calculated as Operating Profit divided by Revenue (ttm).
ΔAssetTurnover‡	Difference between current and last year's Asset Turnover ratio. Asset Turnover is a measure of how effectively a company is using its Assets to generate Revenue. It is calculated as Revenue divided by Total Assets (ttm).

† Magic Formula * Dogs of the Dow †† Buying Winners and Selling Losers ‡ F-Score
+ trailing twelve months (ttm)

must be excluded. The Tabu Search we ran implements a Short Term Memory tabu list (STM) to record the last τ solutions iterated as taboo (τ is the tenure parameter).

After completing the three Tabu Searches, all solutions (MF variants) whose performance metrics were made available through the fitness function, were collected as datapoints into the tbs-dataset. This dataset has 2574 instances with 11+8 features, with the following tabular structure: $\boxed{\text{Yield}}\boxed{\text{ROC}}\boxed{\text{EarningsYield}}$ $\boxed{\text{RS(6m)}}\boxed{\text{ROA}}\boxed{\Delta\text{ROA}}\boxed{\text{AccrualRatio}}\boxed{\Delta\text{LTDebt-to-Assets}}\boxed{\Delta\text{CurrentRatio}}\boxed{\Delta\text{OpMgn}}$ $\boxed{\Delta\text{AssetTurnover}}\boxed{\text{annualized}}\boxed{\text{mean}}\boxed{\text{std}}\boxed{\text{sharpe}}\boxed{\text{alpha}}\boxed{\text{beta}}\boxed{\text{var}}\boxed{\text{tvar}}$. The first 11 features represent the ranking formula of one particular MF variant (out of 2199)[5], as a 11-dimensional vector. Each feature can have one of $\{-1, 0, 1\}$ values. The

[5]Some datapoints in the tbs-dataset are duplicated due to the way the three Tabu Searches were run, all of which started from the original MF as *seed*. Consequently, the MF itself and its near variants in the search space can appear repeated up to three times.

last 8 features correspond to its performance metrics obtained from the sample. In total, only about 1.24% of the entire search space had been explored. However, the small portion of ranking formulas uncovered by Tabu Search proved to be exceptionally high-performing, as discussed in Section 5.1.

4.3 multiSOM

The multiSOM [17, 18], a Self-Organizing Map (SOM) package, was used to carry out exploratory factor analysis on the normalized tbs-dataset. This package retains the basic properties of the original Kohonen's SOM [16] while improving convergence in an interactive way. It also serves as an interactive visualization tool [18] during SOM training. However, for this particular study, we did not use such interactive capabilities.

Instead, we incorporated into the multiSOM the ability of using certain features (dimensions) to be a especial kind of *label*. These labeled features are unique in that they are not utilized in calculating the distances between the SOM's neurons (prototypes) and the input vectors during the training process, which are secured by the regular features. However, they still adapt and converge towards the input vectors, effectively representing the clustered feature in the map once training is complete. In [15], a similar approach is taken for what the authors call the *dummy pre-crisis* attribute. Such approach ensures that the SOM remains focused on the primary features used for clustering while still capturing and displaying important categorical information that can be used to interpret the results effectively.

In this study, we treated the 8 performance metrics as regular features and the 11 encoded factors as labeled features. Before training, we applied min-max normalization to the tbs-dataset, scaling every feature to the $[0, 1]$ interval. We chose a 40x20 neuron lattice for our map and fed the normalized tbs-dataset into the multiSOM for training. After approximately 100,000 training iterations (where one iteration processes a single datapoint), component planes for all 19 features were generated. We analyze these planes in Section 5.2.

It is crucial to understand that since only the 8 performance metrics were used to compute the SOM's distances, the trained map creates a topological representation of the performances observed in the population of MF variants (ranking formulas) included in the tbs-dataset. The trained map highlights zones of potential interest to investors, showcasing zones with high annualized returns, low risk, or high risk-adjusted returns. We identified these zones by examining the component planes of the 'annualized', 'var', and 'sharpe' metrics, respectively. We then conducted exploratory factor analysis in Section 5.2 to determine which financial factors were key contributors to achieving such overperformances, and how they contributed, through their ranking formulas.

5 Results

5.1 Risk/Return Performances

Fig. 2 displays the risk/return performances of the MF and its variants. Blue dots are variants from Grid Search (gs), and red dots are variants from Tabu Search (tbs), as detailed in Sections 4.1 and 4.2. The color intensity of both blue and red dots corresponds to their Sharpe ratios, as indicated by the colorbars on the right. The original MF is shown as a large green dot. Variants with the best metrics of annualized return, annualized VaR, and Sharpe ratio for each dataset are encircled in green, with their Sharpe ratios displayed above for easy comparison. This helps identify the best MF-based investment strategies uncovered. A dashed cross across the figure originates from the risk/return metrics of the STOXX 600 index, which serves as the benchmark index for this sample period in Europe. Variants to the right of the vertical dashed line outperform the market benchmark in terms of risk, while those above the horizontal dashed line outperform in terms of annualized return. The MF nearly matches the market benchmark in the period, though it slightly underperforms in return. Compact notations (gs) and (tbs) are used in the subsequent discussion to refer to MF variants from Grid Search and Tabu Search, respectively.

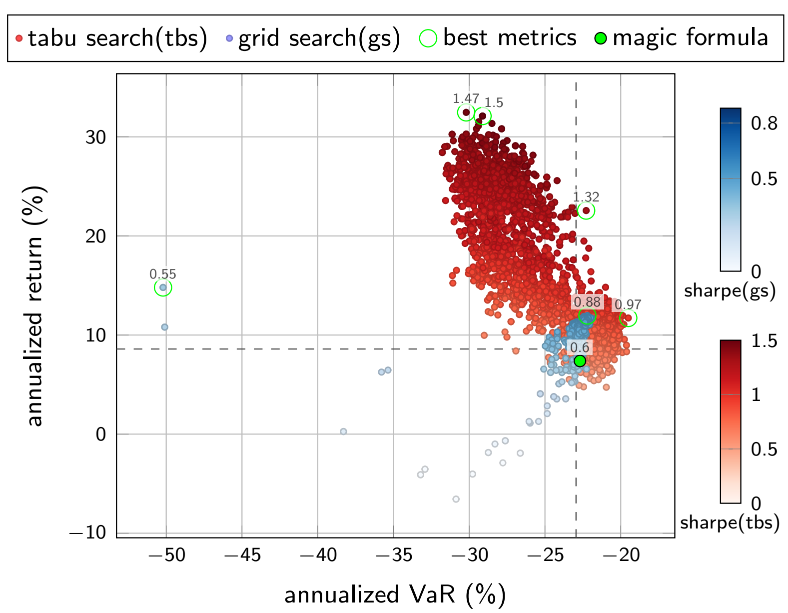

Fig. 2: Plot depicting returns, risks and Sharpe ratios of the MF variants explored through Grid and Tabu Search.

Referring to Fig. 2 and the two accompanying datasets, we can make the following observations. Modifying the ranking formula (tbs) by integrating additional factors (in our study, up to 11) significantly impacts both the return and risk-adjusted return of the MF. These factors should be thoughtfully combined into a Formula (2). This proves to be much more effective than merely changing the number of stocks to invest in (gs); be it by expanding the universe (considering more sectors beyond the original MF's), or choosing a higher (or lower) number of top stocks. Fig. 2 also shows that as MF variants in (tbs) hike in annualized returns, far surpassing the MF's comparatively modest +7.38% in the sample, they tend to also experience more and more downside risk. However, their returns greatly outweigh the associated risks, as indicated by their high Sharpe ratios. Based on the raw dataset (tbs) though, it is difficult to determine which factors have the most significant impact. Hence, we conducted a factor analysis using the multiSOM in the next section.

Risks in (tbs) are significantly concentrated due to the fixed investment in 30 stocks, compared to (gs). The latter displays a significant range of downside risks, from around −50% until it aligns with the MF's at approximately −22.5% annually. Some ranking formulas in (tbs) can even reduce the MF's downside risk by over 10%, which is not the case with (gs), even though the number of stocks to invest in can reach up to 100 in each rebalance. This risk concentration in (tbs) is tied to diversification, as consistently investing in 30 stocks has less diversifiable risk compared to large fluctuations in the number of stocks [22]. Additionally, all MF variants in (gs) with low Sharpe ratios (less than 0.3) invest in fewer than 10 stocks ('top_mf' parameter). In contrast, those variants that outperform MF in terms of Sharpe ratio or return usually do so by investing in an average of 60 stocks. Indeed, data suggests that the top quartile of (gs) in both returns and Sharpe ratio invest in around 80 stocks on average. Regarding the 'sectors' parameter in (gs), data suggests that incorporating additional sectors into the MF has no meaningful impact on the risks and returns.

These findings suggest that amassing stocks to invest in can improve the MF's performance metrics, but only up to a certain point. A much more effective strategy would be to modify the ranking formula, used to select stocks. This involves the inclusion of additional factors in thoughtful combinations, providing a significantly larger potential for improvement.

5.2 Factor Analysis

Fig. 3 shows the component planes after training the multiSOM with the normalized tbs-dataset (Section 4.3). The application of the Self-Organizing Map (SOM) technique provided valuable insights into the underlying structure of the dataset and facilitated the identification of distinct zones with varying investment performance characteristics. Here, we present a detailed analysis of the SOM results based on the three zones of interest discussed earlier.

Three main zones, labeled A, B, and C, were identified based on the criteria of superior returns, low risk, and superior risk-adjusted returns. Zones A and B both exhibited superior returns. Zone A had the highest returns and a high

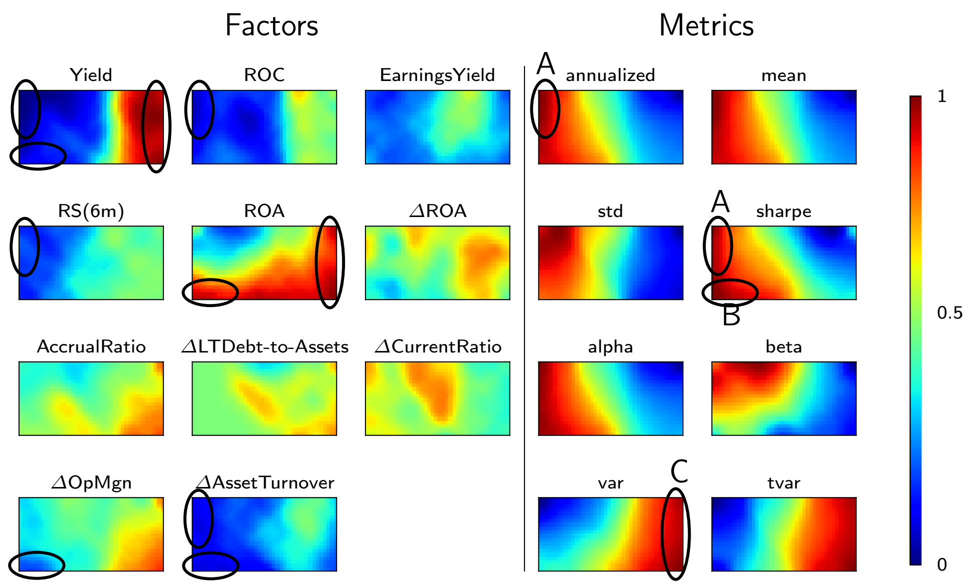

Fig. 3: Component planes of the normalized factors and performance metrics of MF variants after multiSOM training.

Sharpe ratio, though not as high as zone B. Zone B, carrying slightly less risk compared to zone A (as the var map clearly shows), attained the highest Sharpe ratios. Conversely, Zone C was characterized by the lowest risk.

Prior to multiSOM training, all feature values were normalized to the range [0, 1]. Hence, strong red tones on the maps correspond to values close to 1, while dark blue tones represent values close to 0. The intermediate colors cover the remaining spectrum between 0 and 1, e.g., a light green tone would be close to 0.5. Map color tones provide a visual topological representation of the dataset's normalized values. The component planes of the eight performance metrics show great consistency in themselves and align with the (tbs) in Fig. 2 and with the underlying dataset. Higher returns correspond to increased downside risk, as depicted in the VaR and TVaR. A higher Sharpe ratio makes VaR and TVaR less negative, suggesting a better balance between return and risk. Consistent color patterns are observed across profit metrics like annualized, mean, Sharpe, and alpha, and similarities are seen among VaR, TVaR, and standard deviation maps, which represent risk metrics (note that low risk carries a low positive std but a high, i.e. less negative, var and tvar, so the map colors between std and var/tvar appear reversed).

Upon visual inspection of the SOM planes in Fig. 3, factors sometimes show spots of either high or low values within specific map zones, as result of the SOM's convergence. Factors with strong red tones, indicating values close to 1, rank consistently in ascending order within their zones. This suggests that the factors were consistently used in ascending order in the ranking formulas of the MF variants that align performance-wise within that zone. Conversely, factors with dark blue tones, indicating values near 0 (or holding -1 before normalization), rank consistently in descending order within their zones. Zones with intermediate tones or mixed colors signal inconsistent factor behavior and thus make the factor less determinant in the respective zones. In other words, any color tones that deviate from the two extremes may imply that the factor was either absent in the ranking formulas (0 before normalization, 0.5 after normalization) and/or ranked intermittently between ascending and descending orders (alternating between -1 and 1 before normalization, between 0 and 1 after normalization), lacking a clear direction. This is especially true when the color approaches the mid tone (close to 0.5), and less so when the color moves towards the blue or red spectrum but without a strong tone. In these scenarios, without a strong blue or red color, the factor is used (or not) inconsistently in the ranking formulas and therefore is not determinant in driving the performance in their zones.

In summary, the SOM analysis provided a comprehensive understanding of the dataset's structure and facilitated the identification of zones with varying investment performance characteristics. This analysis serves as a crucial foundation for further exploration and refinement of investment strategies tailored to specific zones within the dataset.

Let's analyze the factors in zones A, B, and C. Zone A stands out with high annualized returns and high Sharpe ratios, revealing unexpected combinations of fundamental factors in *descending* order: Yield, ROC, ΔAssetTurnover, along with the descending momentum factor RS(6m). This is intriguing because the literature from which these four factors were taken, including the Magic Formula itself concerning ROC, advocates ranking companies by these factors and then investing in the top ones, i.e., those with the highest factor values. This suggests an ascending order in ranking formulas, contrary to the descending order implied by Fig. 3. Let's focus on Yield as an example, although similar stories could be told for each of the other three factors. Yield is puzzling because the stark distribution of strong tones of both blue and red indicates it is a convincing driver of both returns and risks in the three zones of the map. However, its use in the ranking formulas is unexpectedly reversed. It consistently drives high returns and high Sharpe ratios (zones A and B) when in *descending* order. Conversely, it drives low risk (and consequently low returns) when in *ascending* order. Given that the return metrics include dividends, it is surprising that investing in high Yield companies results in poor returns despite the high dividends. Investing in high Yield companies appears to lower risks, likely due to these companies being mature, blue-chip firms. Zone B, characterized by the highest Sharpe ratios and strong returns, again emphasizes *descending* Yield, ΔAssetTurnover, ΔOpMgn, ROC, and the momentum factor RS(6m). However, the latter two factors are less impactful (lighter blue tones) and are not emphasized in zone B. ROA is the key factor driving superior Sharpe ratios and risk-adjusted returns in zone B. While ROA (Return On Assets) could understandably be a driver of high Sharpe ratios due to its association with high operational efficiency, it fails in zone A, showing no clear direction. ROA appears to add a "low-risk effect" when used in ascending order in ranking formulas. As a clear driver of low-risk investments in zone C, ROA likely pushes ranking formulas toward lower risk, reducing returns (not contributing to zone A) but enhancing its effectiveness in zone B. The ΔOpMgn is also a surprising case. It shows a single dark blue tone around zone B, with no other dark tones seen elsewhere. This indicates it is consistently used in descending order in the ranking formulas that populate zone B. However, descending ΔOpMgn indicates a focus on companies with the largest year-on-year declines in margins, often troubled firms. Yet, investing in these companies appears to enhance risk-adjusted returns. Zone C exhibits the highest VaR (less negative) and therefore the lowest risks among Magic Formula (MF) variants. Three factors appear to drive low-risk investments: Yield and ROA, both discussed earlier, and to a lesser extent, Earnings Yield, in descending order. The remaining factors—ΔROA, AccrualRatio, ΔLTDebt-to-Assets, and ΔCurrentRatio—do not seem to contribute to any of the zones.

6 Conclusions

The original Magic Formula (MF) considers only two factors, but including additional factors can significantly improve results. This study explores 11 factors drawn from other well-known factor-based investment strategies. For the MF, our analysis shows that adding more top-ranked stocks improves performance only up to a certain point. A more effective approach, yielding significant performance improvements in the studied European data sample, is to add more factors to the ranking formula and combine them in specific ways. This not only increases returns and enhances Sharpe ratios but also minimizes value at risk. Given that only a limited number of factors were added to the experiments, this suggests that the optimal combination of factors may have been overlooked in previous MF-based studies (see e.g., [6, 8]).

Our AI-based factor analysis technique provides new insights into traditional theories of factor-based investing [2, 6, 11, 13, 20, 21]. It suggests that some factors might be effectively used in reverse, yielding new interesting results. Factors like Yield, ROC, and momentum RS(6m) seem to exhibit inverse functionality—stocks with the lowest values on these factors prove to be better investment selections than those with comparatively high values. These findings suggest the potential of our AI technique to uncover new insights and refine investment strategies in dynamic market environments.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

Bibliography

[1] A. Ramos, C., C. Marques, N., Faias, M., Santos, H.: Tailor-made Strategies through Different Weight Simulation of Factor-based Investing. Annals of Finance (in press)

[2] Bender, J., Briand, R., Melas, D., Subramanian, R.A.: Foundations of Factor Investing. SSRN Scholarly Paper 2543990, Social Science Research Network, Rochester, NY (Dec 2013). https://doi.org/10.2139/ssrn.2543990

[3] Carrega, M., Santos, H., Marques, N.: Data Streams for Unsupervised Analysis of Company Data. In: Marreiros, G., Melo, F.S., Lau, N., Lopes Cardoso, H., Reis, L.P. (eds.) Progress in Artificial Intelligence. pp. 609–621. Lecture Notes in Computer Science, Springer International Publishing (2021). https://doi.org/10.1007/978-3-030-86230-5_48

[4] Casado, A., Pérez-Peló, S., Sánchez-Oro, J., Duarte, A.: A GRASP algorithm with Tabu Search improvement for solving the maximum intersection of k-subsets problem. Journal of Heuristics **28**(1), 121–146 (Feb 2022). https://doi.org/10.1007/s10732-022-09490-8

[5] Chen, N., Ribeiro, B., Vieira, A., Chen, A.: Clustering and visualization of bankruptcy trajectory using self-organizing map. Expert Systems with Applications **40**(1), 385–393 (Jan 2013). https://doi.org/10.1016/j.eswa.2012.07.047

[6] Davydov, D., Tikkanen, J., Äijö, J.: Magic Formula vs. traditional value investment strategies in the finnish stock market. Nordic Journal of Business **65**(3-4), 38–54 (2016)

[7] Deboeck, G., Kohonen, T. (eds.): Visual Explorations in Finance: with Self-Organizing Maps. Springer Finance, Springer-Verlag, London (1998). https://doi.org/10.1007/978-1-4471-3913-3

[8] Domingues, C.H.S., Aronne, A., Pereira, F., Magalhães, F.: Piotroski, Graham and Greenblatt: an Empirical Approach to Value Investing in the Brazilian Stock Market. BBR. Brazilian Business Review **19**, 475–491 (Sep 2022). https://doi.org/10.15728/bbr.2022.19.5.1.en, publisher: Fucape Business School

[9] Dufour, J.M., Neves, J.: Chapter 1 - Finite-sample inference and nonstandard asymptotics with Monte Carlo tests and R. In: Vinod, H.D., Rao, C.R. (eds.) Handbook of Statistics, Conceptual Econometrics Using R, vol. 41, pp. 3–31. Elsevier (Jan 2019). https://doi.org/10.1016/bs.host.2019.05.001

[10] Glover, F., Laguna, M.: Tabu Search. In: Du, D.Z., Pardalos, P.M. (eds.) Handbook of Combinatorial Optimization: Volume1–3, pp. 2093–2229. Springer US, Boston, MA (1998). https://doi.org/10.1007/978-1-4613-0303-9_33

[11] Greenblatt, J.: The little book that beats the market. Wiley, Hoboken, N.J (2006)

[12] Huang, C., Litzenberger, R.H.: Foundations for Financial Economics. North-Holland (1988)

[13] Jegadeesh, N., Titman, S.: Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. The Journal of Finance **48**(1), 65–91 (1993). https://doi.org/https://doi.org/10.1111/j.1540-6261.1993.tb04702.x

[14] Jorion, P., GARP: Financial Risk Manager Handbook. John Wiley & Sons (Jun 2007), google-Books-ID: uGgrMwDfBRsC

[15] Kiang, M.Y., Kumar, A.: An Evaluation of Self-Organizing Map Networks as a Robust Alternative to Factor Analysis in Data Mining Applications. Information Systems Research **12**(2), 177–194 (Jun 2001). https://doi.org/10.1287/isre.12.2.177.9696, publisher: INFORMS

[16] Kohonen, T.: Self-Organizing Maps. Springer Series in Information Sciences, Springer-Verlag, Berlin Heidelberg, 3 edn. (2001). https://doi.org/10.1007/978-3-642-56927-2

[17] Marques, N.C., Silva, B.: Exploratory Cluster Analysis Using Self-Organizing Maps: Algorithms, Methodologies, and Framework. In: Philosophy of Artificial Intelligence and Its Place in Society, pp. 187–213. IGI Global (2023). https://doi.org/10.4018/978-1-6684-9591-9.ch010

[18] Marques, N.C., Silva, B., Santos, H.: An Interactive Interface for Multi-dimensional Data Stream Analysis. In: 20th International Conference Information Visualisation (IV). pp. 223–229. IEEE, Lisbon, Portugal (Jul 2016). https://doi.org/10.1109/IV.2016.72

[19] Murphy, J.J.: Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications. New York Institute of Finance (1999)

[20] O'Higgins, M.B., Downes, J.: Beating the Dow Completely Revised and Updated: A High-Return, Low-Risk Method for Investing in the Dow Jones Industrial Stocks with as Little as $5,000. HarperCollins e-books, revised, subsequent edn. (May 2011)

[21] Piotroski, J.D.: Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers. Journal of Accounting Research **38**, 1–41 (2000). https://doi.org/10.2307/2672906

[22] Tang, G.Y.N.: How efficient is naive portfolio diversification? an educational note. Omega **32**(2), 155–160 (Apr 2004). https://doi.org/10.1016/j.omega.2003.10.002

[23] Thomsett, M.C.: Mastering Fundamental Analysis. Dearborn Finanical Pub. (Aug 1998)