Portfolio Construction Through Multi-Objective Risk Decomposition

Quantitative Research

Pierre-Henri Landrin

November 1, 2025

Contents

1 Introduction

The optimization of portfolios of automated trading systems (ATS) is a complex problem in quantitative finance. Unlike traditional assets, ATS have varied risk profiles that require an adapted approach. This report proposes a combinatorial optimization method based on equal weighting of strategies (1/N) and using a genetic algorithm to explore the solution space.

2 ATS Performance Metrics

2.1 Fundamental ratios

In the context of evaluating automated trading systems, three risk-adjusted performance ratios are used to characterize the quality of a portfolio.

2.1.1 Sharpe Ratio

For a portfolio P with periodic returns $\{r_t\}_{t=1}^{T}$, the Sharpe Ratio is:

$$SR(P) = \frac{\mu_P}{\sigma_P}$$

where:

- $\mu_P = \mathbb{E}[r_t]$ is the expected return (estimated by $\mu_P = \frac{1}{T}\sum_{t=1}^{T} r_t$)
- $\sigma_P = \sqrt{\mathbb{V}ar[r_t]}$ is the standard deviation of returns (estimated by $\sigma_P = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}(r_t - \mu_P)^2}$)

The Sharpe Ratio measures excess return per unit of volatility. A high ratio indicates better risk-adjusted performance. We use it here without the risk-free rate.

2.1.2 RoMaD (Return over Maximum Drawdown)

RoMaD adjusts return by the maximum loss risk incurred. The Maximum Drawdown of a capital curve $\{V_t\}_{t=0}^{T}$ is defined by:

$$\text{MaxDD} = \max_{t\in[0,T]} \left(\frac{\max_{s\in[0,t]} V_s - V_t}{\max_{s\in[0,t]} V_s} \right)$$

Hence the RoMaD formula:

$$\text{RoMaD}(P) = \frac{\mu_P}{\text{MaxDD}_P}$$

RoMaD penalizes strategies that have suffered significant losses, even temporary ones, regardless of the frequency of these losses.

2.1.3 Ulcer Ratio

The Ulcer Ratio integrates not only the magnitude but also the duration of drawdowns. The Ulcer Index measures the depth and duration of losses, offering a more comprehensive perspective than Maximum Drawdown alone. It is defined by:

$$\text{UI} = \sqrt{\frac{1}{T}\sum_{t=1}^{T} D_t^2}$$

where $D_t = \frac{\max_{s\leq t} V_s - V_t}{\max_{s\leq t} V_s} \times 100$ is the relative drawdown in percentage at time t.

Thus, we obtain the Ulcer Ratio:

$$\mathrm{UR}(P) = \frac{\mu_P}{\mathrm{UI}_P}$$

A high Ulcer Ratio indicates a favorable return/risk profile with limited drawdown periods.

2.2 Need for normalization

Aggregating these three ratios into a single objective function poses a scale problem. Indeed, σ_P, MaxDD_P and UI_P can have very different orders of magnitude, making their linear combination inadequate.

2.2.1 Problem of extreme values

Standard normalization by centering and scaling, defined by:

$$z_i = \frac{x_i - \mu}{\sigma}$$

where μ and σ are respectively the mean and standard deviation of the sample, is sensitive to outliers. In the context of ATS, some portfolios may exhibit extreme risks (e.g., $\mathrm{MaxDD} \to 1$ or $\mathrm{UI} \to \infty$), distorting the statistics μ and σ and thus the normalization of the entire distribution.

2.2.2 Robust normalization using MAD

To overcome sensitivity to extreme values, we adopt a robust normalization based on the *Median Absolute Deviation* (MAD), followed by a transformation via the cumulative distribution function (CDF) of the standard normal distribution.

Let $\{x_i\}_{i=1}^n$ be a sample of risks. We first define the median $m = \mathrm{median}(\mathbf{x})$, then the MAD:

$$\mathrm{MAD} = \mathrm{median}\big(|x_i - m|\big).$$

This measure of dispersion is robust to outliers and, for a Gaussian distribution, satisfies $\sigma \approx 1.4826 \times \mathrm{MAD}$.

The normalization procedure is as follows:

1. Infinite or undefined values (NaN) are temporarily excluded.
2. If $\mathrm{MAD} < \epsilon$ (with small $\epsilon > 0$, typically 10^{-8}), the data are quasi-constant: we then set $\mathbf{x}_{\mathrm{norm}} = 0.5 \times \mathbf{1}$.
3. Otherwise, each component is centered and scaled:

   $$z_i = \frac{x_i - m}{1.4826 \times \mathrm{MAD}},$$

   then normalized by the standard normal CDF:

   $$x_{\mathrm{norm},i} = \Phi(z_i).$$

4. Residual NaNs are finally replaced by 0.5, the median value of the target distribution.

This transformation produces a vector $\mathbf{x}_{\mathrm{norm}}$ with values in $[0, 1]$, resilient to anomalies and consistent with a probabilistic interpretation of risks.

Mathematical justification:

1. **Robustness:** The median and MAD have a breakdown point of 50%, meaning that up to 50% of observations can be outliers without significantly affecting these statistics. In comparison, the mean and standard deviation have a breakdown point of 0%.

2. **Standardization:** The robust score $z_i = \frac{x_i - m}{1.4826 \times \text{MAD}}$ is the analog of the classical z-score, but calculated with robust statistics.

3. **CDF transformation:** Applying the cumulative distribution function of the normal distribution $\Phi(z_i)$ transforms standardized scores into values in the interval $[0, 1]$, with the property that:

   - $\Phi(0) = 0.5$: the median is mapped to 0.5
   - $\Phi(z) \to 0$ as $z \to -\infty$: low risks are close to 0
   - $\Phi(z) \to 1$ as $z \to +\infty$: high risks are close to 1

   This transformation ensures a monotonically increasing normalized distribution, preserving the order of original values while bringing them into a standard interval.

Remark 2.1. This normalization is applied independently to the three risk measures: σ, MaxDD and UI, producing respectively σ_{norm}, MaxDD_{norm} and UI_{norm}, all in the interval $[0, 1]$.

3 Limitations of Markowitz Optimization

Before presenting the methodology, it is essential to understand why the classical Markowitz approach is not directly applicable to our problem.

3.1 Combinatorial problem with 1/N weighting

In the framework of equally weighted 1/N portfolios, finding the optimal portfolio equivalent to a Markowitz solution raises a combinatorial problem. Let $\mathcal{I} = \{1, 2, \ldots, N\}$ be the set of available instances. The space of possible 1/N portfolios is defined by:

$$\mathcal{P}_{1/N} = \{P \subseteq \mathcal{I} \mid P \neq \emptyset\}$$

with cardinality $|\mathcal{P}_{1/N}| = 2^N - 1$.

To identify the 1/N portfolio closest to an optimal Markowitz solution P^*_{MV}, it would be necessary to:

1. Solve the Markowitz problem in the continuous weight space $\mathbf{w} \in [0, 1]^N$ with $\sum_{i=1}^{N} w_i = 1$
2. Calculate the distance (in terms of performance or composition) between P^*_{MV} and each of the $2^N - 1$ possible 1/N portfolios
3. Select the 1/N portfolio minimizing this distance

This approach quickly becomes impractical: for $N = 100$ instances, exhaustive enumeration would require evaluating $2^{100} - 1 \approx 1.27 \times 10^{30}$ portfolios, which is computationally impossible.

3.2 Violation of Markowitz assumptions

Markowitz optimization relies on the assumption that the risk measure used is a risk norm, i.e., a function satisfying certain fundamental properties.

Indeed, we define a risk norm as a function $\rho : \mathbb{R}^n \to \mathbb{R}$ satisfying:

1. **Positivity:** $\rho(\mathbf{x}) \geq 0$ for all $\mathbf{x}$, with equality if and only if $\mathbf{x} = \mathbf{0}$
2. **Homogeneity of degree 1:** $\rho(\lambda \mathbf{x}) = |\lambda| \rho(\mathbf{x})$ for all $\lambda \in \mathbb{R}$
3. **Subadditivity:** $\rho(\mathbf{x} + \mathbf{y}) \leq \rho(\mathbf{x}) + \rho(\mathbf{y})$
4. **Continuity:** ρ is continuous on its domain

In our context, we use a composite risk measure defined by:

$$\mathcal{R}(\mathbf{x}) = \omega_1 \sigma_{norm}(\mathbf{x}) + \omega_2 \text{MaxDD}_{norm}(\mathbf{x}) + \omega_3 \text{UI}_{norm}(\mathbf{x})$$

where $\mathbf{x}$ represents portfolio returns, σ_{norm} is normalized volatility, MaxDD_{norm} is normalized Maximum Drawdown, and UI_{norm} is normalized Ulcer Index.

3.2.1 Non-homogeneity of Maximum Drawdown and Ulcer Index

Maximum Drawdown and Ulcer Index are not homogeneous of degree 1. Indeed, consider the Maximum Drawdown defined by:

$$\text{MaxDD}(\mathbf{x}) = \max_{t \in [0,T]} \left(\max_{s \in [0,t]} x_s - x_t \right)$$

For a factor $\lambda > 0$, the Maximum Drawdown of scaled returns $\lambda \mathbf{x}$ is:

$$\text{MaxDD}(\lambda \mathbf{x}) = \max_{t \in [0,T]} \left(\max_{s \in [0,t]} \lambda x_s - \lambda x_t \right) = \lambda \cdot \text{MaxDD}(\mathbf{x})$$

However, this relationship is only valid if returns are expressed in absolute terms. In financial practice, Maximum Drawdown is calculated on capital values (equity curve) V_t that evolve multiplicatively:

$$V_t = V_0 \prod_{i=1}^{t} (1 + r_i)$$

where r_i are periodic returns. In this case, for a portfolio of size λV_0, the Maximum Drawdown in relative terms remains the same, but the homogeneity relationship $\text{MaxDD}(\lambda V_0) = \lambda \cdot \text{MaxDD}(V_0)$ is only verified if we consider absolute drawdowns, not relative ones.

For the Ulcer Index, defined by:

$$\text{UI}(\mathbf{x}) = \sqrt{\frac{1}{T} \sum_{t=1}^{T} D_t^2}$$

where $D_t = \frac{\max_{s \leq t} V_s - V_t}{\max_{s \leq t} V_s} \times 100$ is the relative drawdown in percentage, we observe that $\text{UI}(\lambda \mathbf{x}) \neq \lambda \cdot \text{UI}(\mathbf{x})$ because relative drawdowns D_t are not affected by a multiplicative scaling factor of capital.

Consequently, neither Maximum Drawdown (in relative terms) nor Ulcer Index satisfy the homogeneity property of degree 1 required for a risk norm in the Markowitz sense.

This non-homogeneity implies that classical results of Markowitz optimization, notably the existence of a convex efficient frontier and the uniqueness of the tangent portfolio, are not guaranteed when using $\mathcal{R}(\mathbf{x})$ as a risk measure.

4 Fitness Function Construction

4.1 Initial objective and identified problem

The initial objective was to construct a fitness function as a weighted linear combination of the three normalized ratios:

$$F_{initial}(P) = \omega_1 \cdot SR_{norm}(P) + \omega_2 \cdot \text{RoMaD}_{norm}(P) + \omega_3 \cdot \text{UR}_{norm}(P)$$

By decomposing each ratio, we observe that the numerator is common: the expected return μ_P. Thus:

$$\begin{aligned} SR_{norm}(P) &= \frac{\mu_P}{\sigma_{norm}(P)} \\ \text{RoMaD}_{norm}(P) &= \frac{\mu_P}{\text{MaxDD}_{norm}(P)} \\ \text{UR}_{norm}(P) &= \frac{\mu_P}{\text{UI}_{norm}(P)} \end{aligned}$$

By factorization, we obtain:

$$F_{initial}(P) = \mu_P \left(\frac{\omega_1}{\sigma_{norm}(P)} + \frac{\omega_2}{\text{MaxDD}_{norm}(P)} + \frac{\omega_3}{\text{UI}_{norm}(P)} \right)$$

Which can be written in return/risk form:

$$F_{initial}(P) = \frac{\mu_P}{\mathcal{R}_{initial}(P)}$$

where the aggregated risk is defined by:

$$\mathcal{R}_{initial}(P) = \frac{1}{\omega_1/\sigma_{norm}(P) + \omega_2/\text{MaxDD}_{norm}(P) + \omega_3/\text{UI}_{norm}(P)}$$

4.2 Non-convexity problem

The function $\mathcal{R}_{initial}(P)$ is not convex, resulting in a non-convex distribution of portfolios in the return-risk space.

Empirical analysis: Graphical observation of the portfolio distribution in the space $(\mathcal{R}_{initial}, \mu)$ reveals "ridge" or "spike" structures, characteristic of a non-convex solution space. This non-convexity poses several problems:

1. **No guarantee of optimality:** In a non-convex space, an optimization algorithm may converge to local optima that are not global optima.
2. **Search difficulty:** Traditional optimization methods (gradient, quasi-Newton, etc.) lose their effectiveness in non-convex spaces, requiring metaheuristic approaches.
3. **Numerical instability:** The function $\mathcal{R}_{initial}$ contains terms in $1/\sigma_{norm}$ that can become very large when $\sigma_{norm} \to 0$, creating numerical discontinuities.

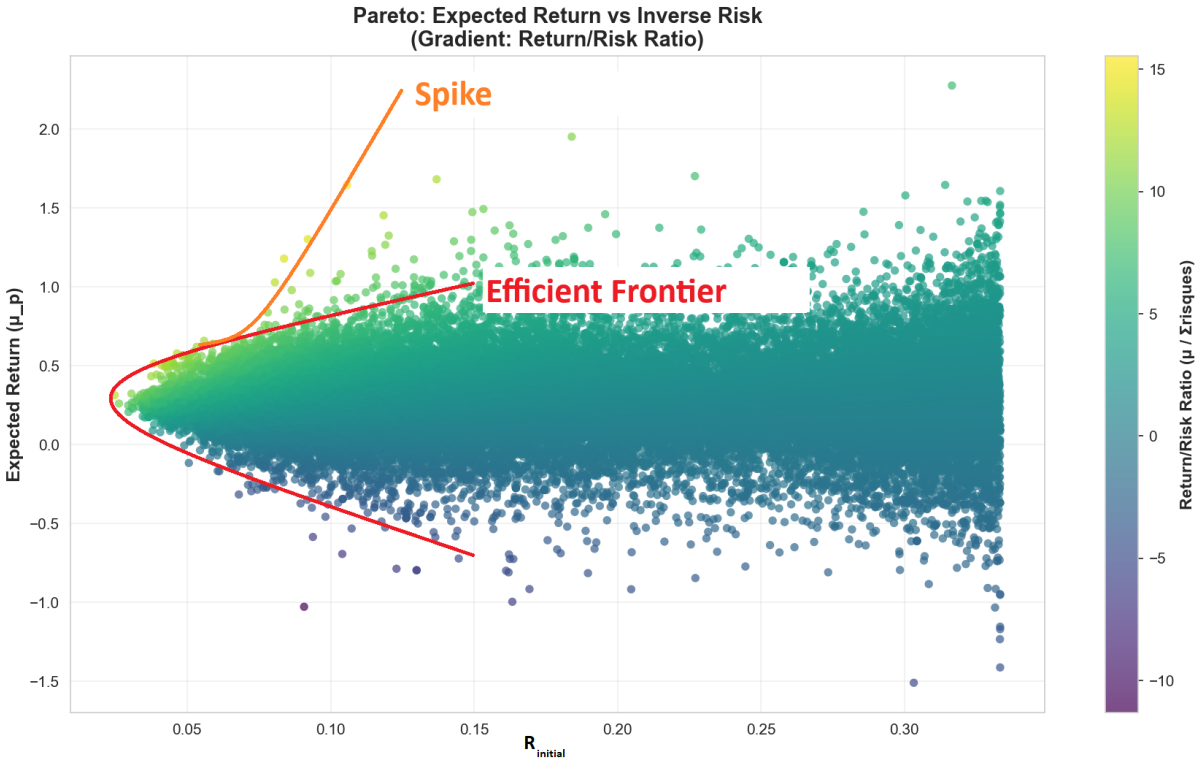


Figure 1: Graph of $\mathcal{F}_{initial}(P)$ for 100,000 randomly generated ATS

4.3 Corrected fitness function

To restore convexity of the solution space, we modify the definition of aggregated risk by using a linear combination rather than a harmonic combination:

$$F(P) = \frac{\mu_P}{\mathcal{R}(P)}$$

where the weighted risk is defined by:

$$\mathcal{R}(P) = \omega_1 \sigma_{norm}(P) + \omega_2 \text{MaxDD}_{norm}(P) + \omega_3 \text{UI}_{norm}(P)$$

with constraints $\omega_i > 0$ and $\sum_{i=1}^{3} \omega_i = 1$.

Indeed, if σ_{norm}, MaxDD_{norm} and UI_{norm} are convex functions of portfolio composition, then $\mathcal{R}(P)$ is convex.

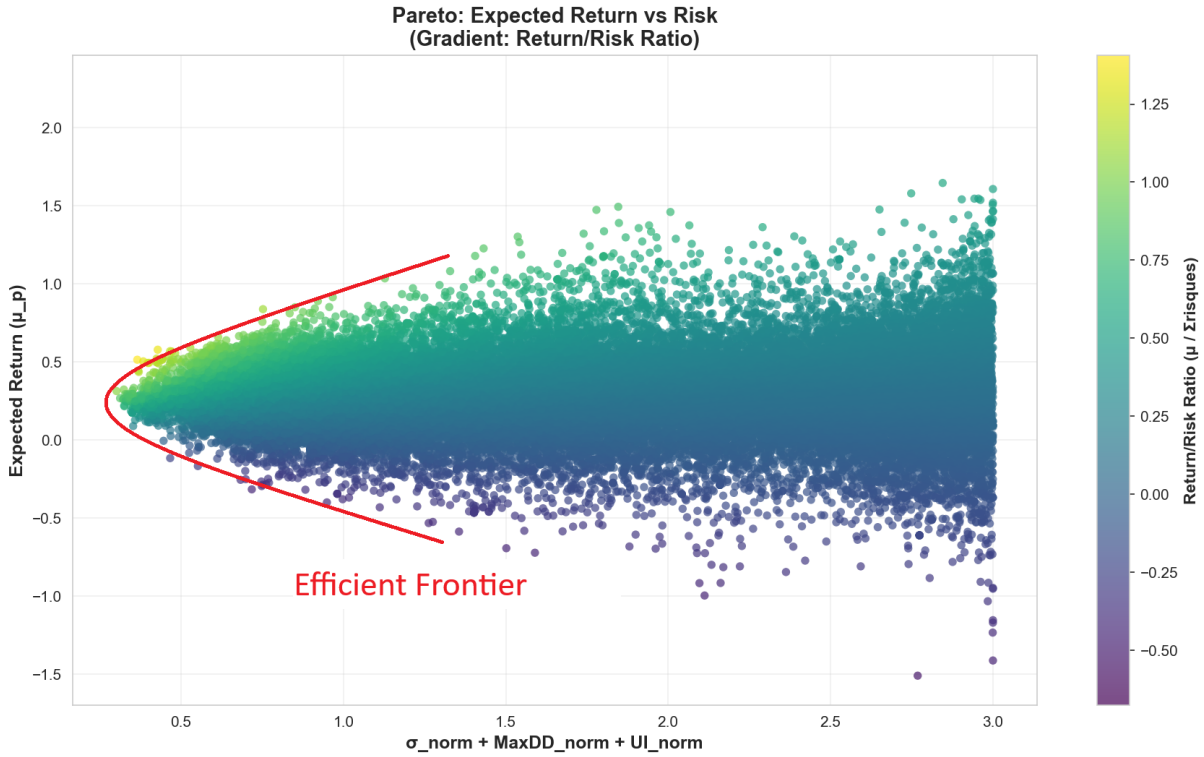


Figure 2: Graph of $\mathcal{F}_{final}(P)$ for 100,000 randomly generated ATS

Consequence: The solution space in the $(\mathcal{R}, \mu)$ plane becomes convex, guaranteeing the existence of an efficient frontier and facilitating the search for optima.

4.4 Importance of weights ω_i

The coefficients ω_1, ω_2 and ω_3 play an important role in balancing the fitness function. Without appropriate weighting, one risk measure could dominate the others, biasing the optimization. Manually setting weights is inadequate, as their relevance strongly depends on the initial dataset: certain choices may suit a particular set of instances, but prove ineffective for others.

If $\omega_1 = \omega_2 = \omega_3 = 1/3$ (uniform weighting), and if for example $\sigma_{norm} \approx 0.1$ while $\text{MaxDD}_{norm} \approx 0.8$, the MaxDD_{norm} term would dominate $\mathcal{R}(P)$, reducing the influence of volatility and Ulcer Index in optimization.

The determination of optimal weights ω_i dynamically is therefore essential, and is the subject of the following section.

5 Weight Determination by Pareto Optimization

5.1 Multi-objective optimization principle

The portfolio optimization problem is inherently multi-objective: we simultaneously seek to maximize return and minimize three distinct risk measures. The Pareto approach allows identifying non-dominated portfolios, i.e., those for which no improvement on one criterion is possible without deterioration on another.

5.2 Approach

Let two portfolios P and P'. We say that P dominates P' (denoted $P \succ P'$) if and only if:

$$\begin{cases} \mu_P \geq \mu_{P'} \\ \sigma_{norm}(P) \leq \sigma_{norm}(P') \\ \text{MaxDD}_{norm}(P) \leq \text{MaxDD}_{norm}(P') \\ \text{UI}_{norm}(P) \leq \text{UI}_{norm}(P') \end{cases}$$

with at least one strict inequality.

The Pareto set (or Pareto frontier) $\mathcal{F}_P$ is the set of non-dominated portfolios:

$$\mathcal{F}_P = \left\{ P \in \mathcal{P} \mid \nexists P' \in \mathcal{P} : P' \succ P \right\}$$

where $\mathcal{P}$ is the set of all considered portfolios.

5.3 Generation and selection procedure

Step 1: Random portfolio generation A set of N_{rand} portfolios is randomly generated with 1/N weighting. Each portfolio P_i is characterized by its normalized metric vector:

$$\mathbf{m}_i = (\mu_i, \sigma_{norm,i}, \text{MaxDD}_{norm,i}, \text{UI}_{norm,i})$$

Step 2: Pareto frontier identification The non-dominated sorting algorithm is applied to identify $\mathcal{F}_P$.

Step 3: Performance ratio definition For each portfolio $P \in \mathcal{F}_P$, three return/risk ratios are calculated:

$$R_1(P) = \frac{\mu_P}{\sigma_{norm}(P)} \quad \text{(Normalized Sharpe)}$$
$$R_2(P) = \frac{\mu_P}{\text{MaxDD}_{norm}(P)} \quad \text{(Normalized RoMaD)}$$
$$R_3(P) = \frac{\mu_P}{\text{UI}_{norm}(P)} \quad \text{(Normalized Ulcer Ratio)}$$

Step 4: Ideal point definition The ideal point I is defined as the vector of maxima of each ratio:

$$I = (R_1^*, R_2^*, R_3^*) \quad \text{where} \quad R_j^* = \max_{P \in \mathcal{F}_P} R_j(P), \quad j = 1, 2, 3$$

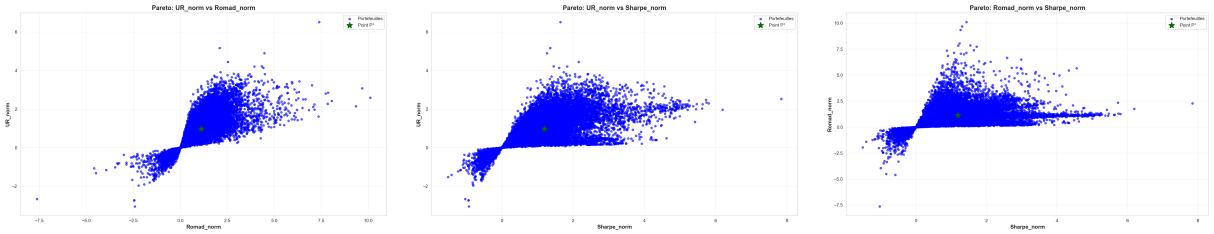
This point represents a theoretical optimal portfolio that simultaneously maximizes all three ratios (generally unattainable in practice).

Step 5: Optimal Pareto portfolio selection The optimal portfolio P^* is the one that minimizes the Euclidean distance to the ideal point:

$$P^* = \arg\min_{P \in \mathcal{F}_P} d_R(P) \tag{1}$$

where the distance in the ratio space is defined by:

$$d_R(P)^2 = (R_1^* - R_1(P))^2 + (R_2^* - R_2(P))^2 + (R_3^* - R_3(P))^2$$



(a) UR vs RoMaD (b) UR vs Sharpe (c) RoMaD vs Sharpe

Figure 3: Representation of Pareto fronts for normalized risk pairs (R1, R2, R3)

This distance measures the gap between portfolio P and the ideal point I in the three-dimensional ratio space. It represents a balanced compromise between the three objectives, avoiding arbitrarily favoring one ratio over the others.

5.4 Optimal weight calculation

Once the optimal Pareto portfolio P^* is identified, the weights ω_i are determined based on its normalized risk characteristics.

The weights ω_1, ω_2 and ω_3 are calculated according to:

$$\omega_1 = \frac{(1/\sigma_{norm}(P^*))^\kappa}{S_\kappa} \tag{2}$$

$$\omega_2 = \frac{(1/\text{MaxDD}_{norm}(P^*))^\kappa}{S_\kappa} \tag{3}$$

$$\omega_3 = \frac{(1/\text{UI}_{norm}(P^*))^\kappa}{S_\kappa} \tag{4}$$

where the normalization constant is: $S_\kappa = (1/\sigma_{norm}(P^*))^\kappa + (1/\text{MaxDD}_{norm}(P^*))^\kappa + (1/\text{UI}_{norm}(P^*))^\kappa$ and $\kappa > 0$ is a hyperparameter.

Interpretation:

1. **Inversion principle:** Weights are inversely proportional to the normalized risks of P^*. If a risk measure is low for P^*, its weight in $\mathcal{R}$ will be high, and vice versa.
2. **Normalization:** Division by S_κ ensures that $\omega_1 + \omega_2 + \omega_3 = 1$, guaranteeing a balanced contribution in the scale of $\mathcal{R}$.
3. **Role of κ:**
   - If $\kappa = 1$: linear weighting proportional to risk inverses
   - If $\kappa > 1$: amplification of differences between risks (low risks get disproportionately high weight)
   - If $0 < \kappa < 1$: attenuation of differences (more uniform weights)

Economic justification: This approach is based on the idea that the optimal Pareto portfolio P^* reveals the relative importance of each risk dimension. By assigning higher weights to risks that are low in P^*, we encourage the optimization algorithm to search for portfolios that excel on the most controllable criteria, while maintaining balance on others.

6 Optimization by Genetic Algorithm

6.1 Theoretical foundations

The genetic algorithm (GA) is a metaheuristic inspired by Darwinian evolution theory. It relies on mechanisms of natural selection, genetic recombination and mutation to efficiently explore complex solution spaces.

6.1.1 Formal definitions

Here is a brief introduction to the genetic algorithm:

- An **individual** $\mathbf{x} \in \{0,1\}^N$ is a binary vector representing a portfolio, where $x_i = 1$ if instance i is included, $x_i = 0$ otherwise.
- A **population** $\mathcal{P}(t)$ at generation t is a set of $|\mathcal{P}|$ individuals: $\mathcal{P}(t) = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{|\mathcal{P}|}\}$.
- The **fitness** function $F : \{0,1\}^N \to \mathbb{R}$ evaluates the quality of an individual. In our context: $F(\mathbf{x}) = \frac{\mu(\mathbf{x})}{\omega_1 \sigma_{norm}(\mathbf{x}) + \omega_2 \text{MaxDD}_{norm}(\mathbf{x}) + \omega_3 \text{UI}_{norm}(\mathbf{x})}$, the objective is to maximize $F(\mathbf{x})$.

6.1.2 Local search principles

The GA belongs to the class of stochastic local search methods. Unlike exact methods that guarantee optimality but become impractical on large instances, the GA proposes a compromise between solution quality and computation time.

For an individual $\mathbf{x}$, the neighborhood $\mathcal{N}(\mathbf{x})$ is the set of individuals reachable by a small modification (mutation or crossover). In the binary case: $\mathcal{N}_1(\mathbf{x}) = \{\mathbf{x}' \mid d_H(\mathbf{x}, \mathbf{x}') = 1\}$ where d_H is the Hamming distance (number of different bits).

Under certain conditions (diversity maintenance, elitism), a GA converges in probability to a global optimum when the number of generations tends to infinity.

Justification: The GA combines:

1. **Exploitation** (intensification): selection and elitism favor good solutions
2. **Exploration** (diversification): crossover and mutation allow exploring new regions of the search space

This balance between exploitation and exploration helps avoid local optima while converging to high-quality solutions.

6.2 Genetic operators

6.2.1 Selection by elitism

At each generation t, the K best individuals from $\mathcal{P}(t)$ according to F are systematically copied to $\mathcal{P}(t+1)$.

Advantage: Elitism guarantees monotonicity of the best fitness across generations: $\max_{\mathbf{x}\in\mathcal{P}(t+1)} F(\mathbf{x}) \geq \max_{\mathbf{x}\in\mathcal{P}(t)} F(\mathbf{x})$.

6.2.2 Crossover

Crossover combines genetic material from two parents to create offspring. Let two parents $\mathbf{p}_1, \mathbf{p}_2 \in \{0,1\}^N$ and a cutpoint $c \in \{1, \ldots, N-1\}$ chosen randomly. The two offspring are:

$$\mathbf{e}_1 = (p_{1,1}, \ldots, p_{1,c}, p_{2,c+1}, \ldots, p_{2,N})$$
$$\mathbf{e}_2 = (p_{2,1}, \ldots, p_{2,c}, p_{1,c+1}, \ldots, p_{1,N})$$

Let two cutpoints $c_1 < c_2$ be chosen randomly. The offspring are:

$$\mathbf{e}_1 = (p_{1,1}, \ldots, p_{1,c_1}, p_{2,c_1+1}, \ldots, p_{2,c_2}, p_{1,c_2+1}, \ldots, p_{1,N})$$
$$\mathbf{e}_2 = (p_{2,1}, \ldots, p_{2,c_1}, p_{1,c_1+1}, \ldots, p_{1,c_2}, p_{2,c_2+1}, \ldots, p_{2,N})$$

Crossover probability: Crossover is applied with probability $p_c \in [0,1]$. Typically, $p_c \approx 0.7$.

6.2.3 Mutation

For each bit x_i of an individual $\mathbf{x}$, with probability p_m (mutation rate), the bit is flipped: $x_i' = \begin{cases} 1 - x_i & \text{with probability } p_m \\ x_i & \text{with probability } 1 - p_m \end{cases}$

Role of mutation: Mutation introduces genetic diversity, allowing exploration of solution space regions not accessible by crossover. A typical mutation rate is $p_m \approx 0.05$.

6.2.4 Tournament selection

To select a parent:

1. Randomly choose k individuals from $\mathcal{P}(t)$ (tournament size)
2. Return the individual with the best fitness among the k selected

6.3 Cardinality constraints

In our implementation, we impose constraints on the number of instances included in a portfolio: $n_{min} \leq \|\mathbf{x}\|_1 \leq n_{max}$ where $\|\mathbf{x}\|_1 = \sum_{i=1}^{N} x_i$ is the number of selected instances.

After mutation, if $\|\mathbf{x}\|_1 < n_{min}$, random bits are set to 1 until satisfaction. If $\|\mathbf{x}\|_1 > n_{max}$, random bits are set to 0.

Justification:

- $n_{min} > 0$: avoids empty or overly concentrated portfolios
- $n_{max} < N$: limits complexity and promotes diversification

6.4 Algorithmic complexity

The time complexity of the genetic algorithm is: $\mathcal{O}(G \cdot |\mathcal{P}| \cdot C_F)$ where G is the number of generations, $|\mathcal{P}|$ the population size, and C_F the complexity of fitness evaluation.

7 Results and Analysis

7.1 Experimental setup

Experiments were conducted on the entire set of available instances, with the following parameters:

Parameter	Value
Data window	1 day
Population size ($\|\mathcal{P}\|$)	150
Number of generations (G)	100
Crossover rate (p_c)	0.7
Mutation rate (p_m)	0.05
Elitism (K_{elite})	15 (10% of population)
Tournament size	5
Cardinality constraints	$n_{min} = 1$, $n_{max} = 10$
Random portfolios (Pareto)	100,000
Parameter κ	1.0

Table 1: Genetic algorithm parameters

7.2 Pareto phase: Weight determination

The Pareto analysis phase identified the optimal portfolio P^* among 1,000 randomly generated portfolios. The determined optimal weights are:

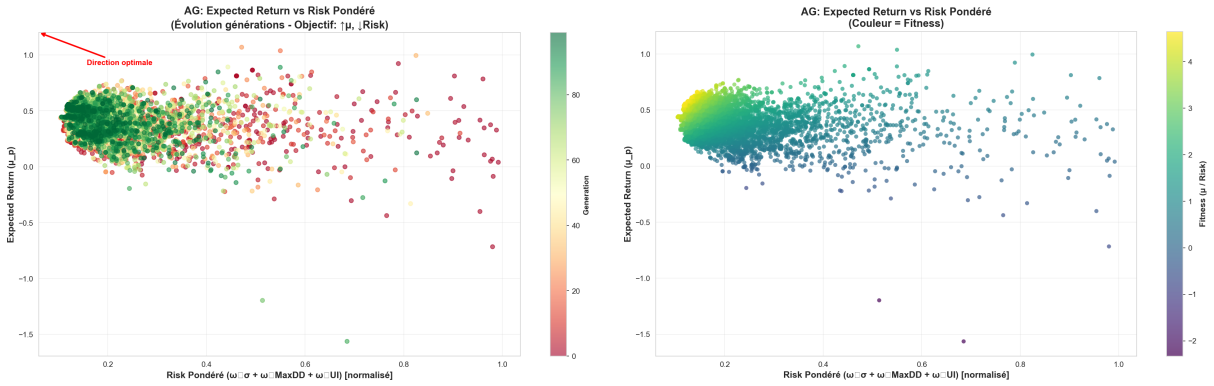
$\omega = (\omega_1, \omega_2, \omega_3)$

These weights reflect the relative importance of each risk dimension in constructing optimal portfolios. The observed Pareto frontier exhibits a convex structure in the normalized ratio space, confirming the validity of our approach.

```
CALCULATED OMEGA WEIGHTS
========================

KAPPA parameter: 1.0

omega_1 (Volatility): 0.362279
omega_2 (Max Drawdown): 0.341303
omega_3 (Ulcer Index): 0.296418

Selected Pareto point:
  Sharpe_norm: 1.193462
  UR_norm: 0.976495
  Romad_norm: 1.124360
  Mu (annualized return): 0.272398
  Sigma (annualized volatility): 1.291365
  Max Drawdown: 2.734823
  Ulcer Index: 19.357251
```

7.3 Genetic phase: Optimal portfolios

The genetic algorithm converged to high-quality solutions, with significant fitness improvement across generations.



(a) Generation Gradient (b) Fitness Gradient

Figure 4: Evolution of ATS Population During GA

The top 5 identified portfolios exhibit the following characteristics:

- **Diversification:** The number of instances per portfolio varies between n_{min} and n_{max}, with a median around 5-7 instances.
- **Performance:** The fitness of the best portfolios significantly exceeds that of the best random portfolio from the Pareto phase.
- **Convergence:** Fitness evolution shows rapid convergence in the first 20 generations, followed by a refinement phase.

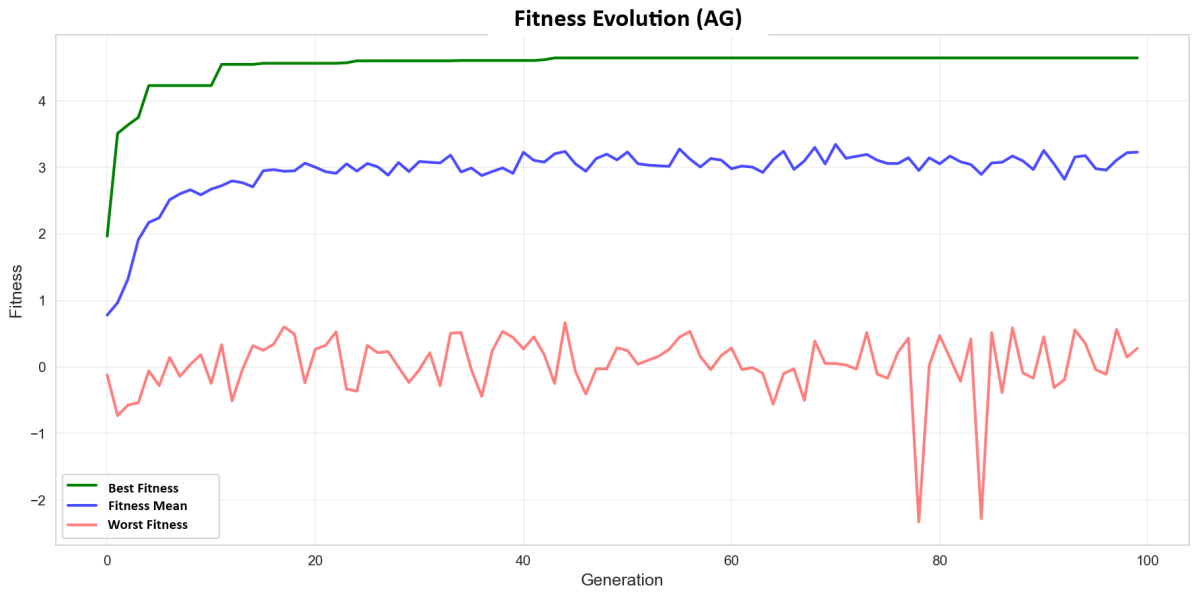


Figure 5: Fitness Evolution Over 100 Generations

7.4 Contribution analysis

For each optimal portfolio, we calculated the contribution of each instance to the total portfolio return. In the following example, we count the contribution of each instance only from the moment they are all active.

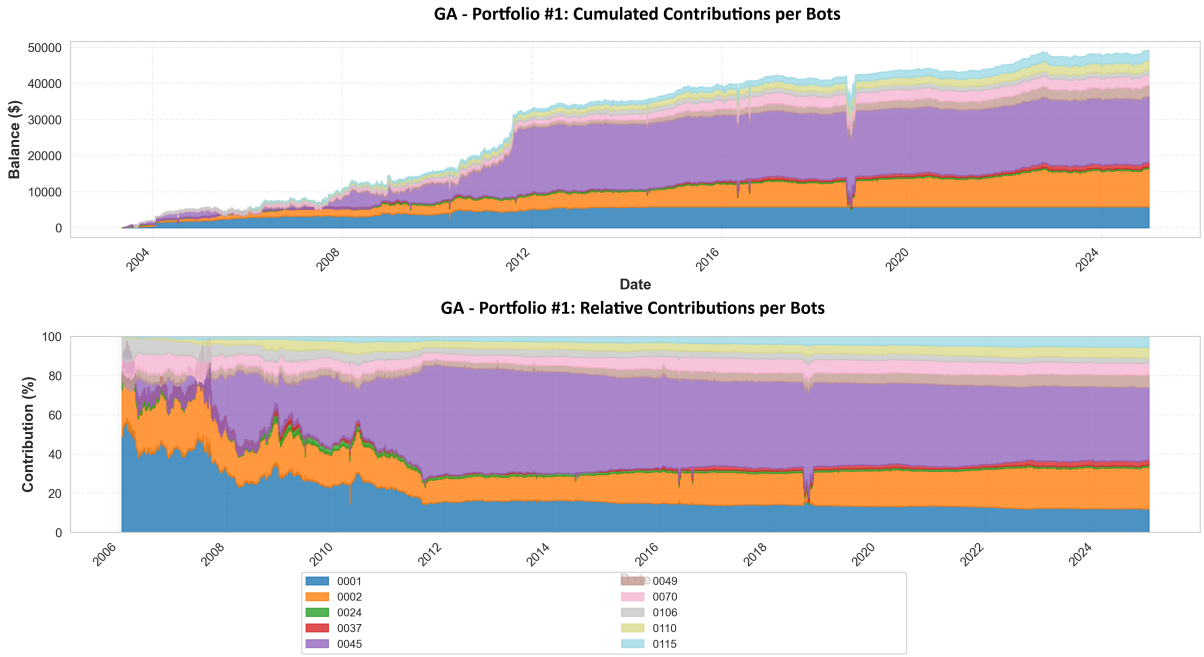


Figure 6: Cumulative and Relative Contributions of Instances in ATS #1

This analysis reveals:

1. **Partial domination:** In most portfolios, one instance contributes more than 40% of total return.
2. **Diversification effect:** Despite this concentration, minority instances play an important role in reducing overall risk, justifying their inclusion.
3. **Robustness:** Optimal portfolios exhibit a similar structure in terms of selected instances, suggesting the presence of a common core of high-performing instances.

7.5 Comparison of top 5 ATS

Table 2: Best ATS portfolios identified by the genetic algorithm

Rank	Fitness	μ	σ	MDD	UI	Sharpe	RoMaD	UR
1	5.569001	0.620	0.942	1.106	5.686	0.658	0.561	0.109
2	5.568218	0.601	0.901	1.110	5.579	0.666	0.541	0.108
3	5.566220	0.620	0.942	1.106	5.710	0.658	0.561	0.109
4	5.564527	0.600	0.901	1.110	5.559	0.666	0.540	0.108
5	5.562064	0.620	0.942	1.106	5.693	0.658	0.560	0.109

Robots (10 per portfolio)
#1: 0001 (L), 0002 (S), 0024 (L), 0040 (L), 0045 (S), 0053 (S), 0092 (L), 0104 (L), 0110 (L), 0111 (S)
#2: 0001 (L), 0002 (S), 0024 (L), 0045 (S), 0053 (S), 0069 (S), 0090 (L), 0092 (L), 0096 (L), 0104 (L)
#3: 0001 (L), 0002 (S), 0024 (L), 0040 (L), 0045 (S), 0053 (S), 0090 (L), 0104 (L), 0111 (S), 0115 (S)
#4: 0001 (L), 0002 (S), 0024 (L), 0042 (L), 0045 (S), 0053 (S), 0090 (L), 0092 (L), 0096 (L), 0104 (L)
#5: 0001 (L), 0002 (S), 0024 (L), 0040 (L), 0045 (S), 0053 (S), 0064 (L), 0090 (L), 0104 (L), 0111 (S)

8 Conclusion and Perspectives

8.1 Methodological contributions

This report presented a comprehensive methodology for ATS portfolio optimization, combining:

1. **Robust normalization:** The use of MAD and CDF transformation effectively handles extreme values present in trading data.
2. **Convex fitness function:** The reformulation of the risk measure guarantees convexity of the solution space, facilitating optimization.
3. **Automatic weight determination:** The Pareto approach provides an objective and reproducible method for weighting different risk dimensions.
4. **Metaheuristic optimization:** The genetic algorithm offers an optimal compromise between solution quality and computation time for a large-scale combinatorial problem.

To strengthen the robustness, generalization and operational performance of the proposed system, several quantitative improvement directions are envisioned. These paths are part of an iterative continuous optimization approach.

8.1.1 Hybridization with local search

The genetic algorithm, while effective at exploring a large-scale combinatorial solution space, may prematurely converge to local optima in dense regions of the search space, as shown by these 3 clusters in the following graph:

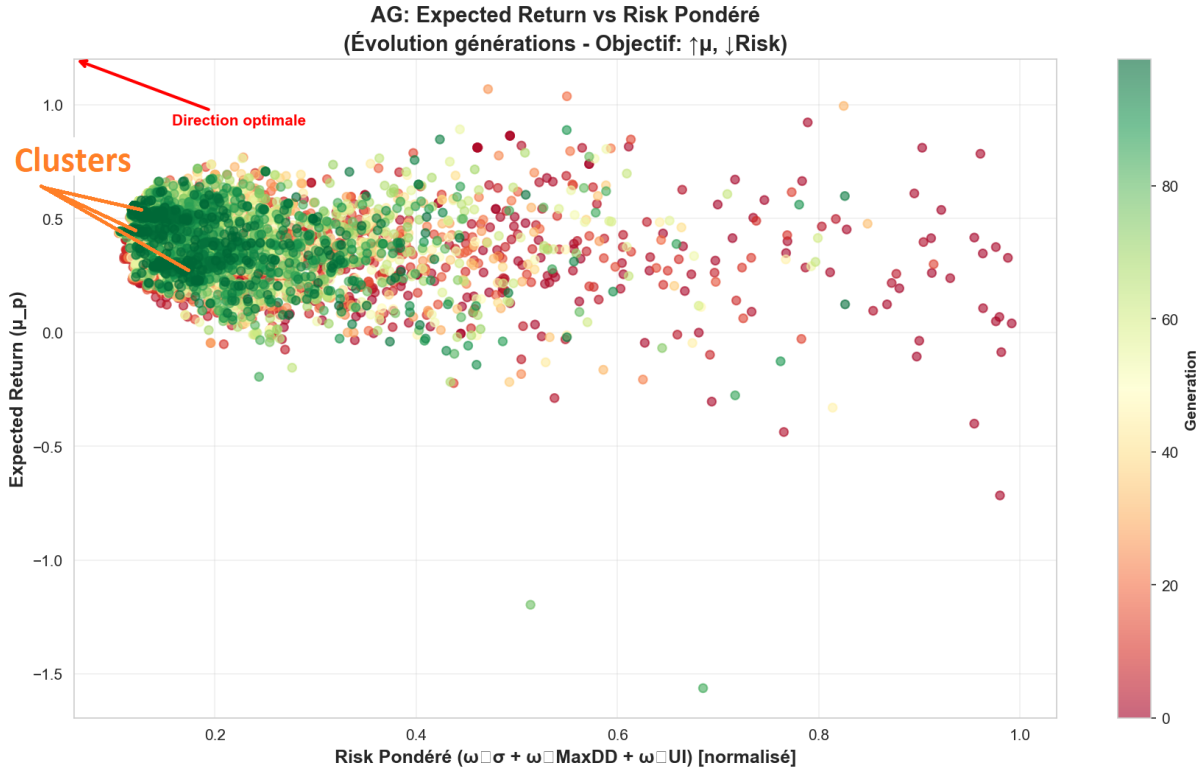


Figure 7: Visualization of Clustering Phenomenon

Hybridization with a *local search* phase applied to the best individuals of each generation would refine solutions with high precision. For example, a deterministic *hill-climbing* algorithm

or a stochastic *simulated annealing* could be activated on the K best portfolios (with $K = 5$ or 10) at each generation or only at the end of execution. This approach, known as a *memetic algorithm*, combines the global exploration of GA with local exploitation, which would potentially reduce the convergence speed of fitness but would certainly increase its convergence plateau.

8.1.2 Multi-period optimization

Current validation relies on a single training period, which exposes the model to overfitting risk. A natural extension consists of adopting *walk-forward analysis* with sliding time windows (for example, 12 months of training followed by 3 months of validation, repeated over 5 to 10 cycles). At each iteration, the GA is retrained on the current window, and portfolios are evaluated out-of-sample. This methodology not only measures the temporal robustness of strategies, but also identifies the most stable hyperparameter configurations (κ, mutation rate, etc.).

8.1.3 Sensitivity analysis

The behavior of the GA strongly depends on its hyperparameters, whose impact remains partially unexplored. A systematic sensitivity analysis, via a factorial design of experiments or a *Sobol* type method, would quantify the effect of κ, p_m, p_c, N_{pop} and $G_{\max}$ on final fitness, convergence time and Pareto front diversity. For example, a grid of $5^4 = 625$ configurations (5 values per parameter) could be evaluated in parallel on a cluster, with measurement of variance explained by each factor. This study, combined with meta-optimization (for example via *Bayesian Optimization*), could automate fine-tuning and adapt parameters to market regime (high/low volatility), thus reinforcing the overall system robustness.

8.2 General conclusion

The proposed methodology constitutes an efficient and rigorous solution for ATS optimization with 1/N weighting. It overcomes the limitations of classical approaches (Markowitz) while offering a flexible and extensible framework. Experimental results confirm the relevance of the approach, with optimal portfolios exhibiting an excellent return/risk compromise according to the three metrics considered.

The current implementation is operational and can be deployed in a production environment, with possibilities for methodological enrichment for future versions...