Expert Systems With Applications 215 (2023) 119370

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

A new approach to portfolio selection based on forecasting

Ana Corberán-Vallet [a,*], Enriqueta Vercher [a], José V. Segura [b], José D. Bermúdez [a]

[a] Department of Statistics and Operations Research, University of Valencia, Dr. Moliner 50, 46100 Burjassot, Spain
[b] Center of Operational Research, University Miguel Hernández of Elche, Av. Ferrocarril s/n, 03202 Elche, Spain

ARTICLE INFO

Keywords:
Portfolio optimization
Finance
Time series analysis
Forecasting
Multi-objective genetic algorithm

ABSTRACT

In this paper we analyze the portfolio selection problem from a novel perspective based on the analysis and prediction of the time series corresponding to the portfolio's value. Namely, we define the value of a particular portfolio at the time of its acquisition. Using the time series of historical prices of the different financial assets, we calculate backward the value that said portfolio would have had in past time periods. A damped trend model is then used to analyze this time series and to predict the future values of the portfolio, providing estimates of the mean and variance for different forecasting horizons. These measures are used to formulate the portfolio selection problem, which is solved using a multi-objective genetic algorithm. To show the performance of this procedure, we use a data set of asset prices from the New York Stock Market.

1. Introduction

The portfolio selection problem is an important topic in financial mathematics. It focuses on allocating the investor's capital to a set of assets (which assets and in what proportion) so that the expected return is maximized assuming a minimum risk. The future return on an asset is unknown when the decision is made, therefore one has to deal with the uncertainty associated with the future portfolio return. This can be done by using probability-based stochastic tools such as the classical formulation provided by Markowitz (1952), where a mean–variance model is used to either maximize the investment return under certain risk level or minimize the investment risk under certain return level. The investment return and risk are quantified, respectively, by means of the expected value and the variance of the random variables representing the returns of individual assets. The standard mean–variance model assumes that the returns on the assets follow a multivariate Normal distribution, which does not necessarily hold in practice. In order to achieve suitable information regarding these random variables in absence of Normality, the incorporation of moments higher than the second plays an important role. Lai (1991) determined a portfolio selection with skewness from solving a polynomial goal programming problem that incorporates the investor's preferences. Alternatively, different risk measures have been proposed: for instance, the mean absolute deviation risk (Konno & Yamazaki, 1991) and the conditional value-at-risk (Rockafellar & Uryasev, 2000). These measures allow the portfolio selection problem to be formulated and solved using linear programming techniques. Within this framework, robust formulations, which use uncertainty structures, have been proposed for handling the sensitivity of optimal portfolios to statistical errors in the estimates of market parameters (Goldfarb & Iyengar, 2003). Robust optimization problems can be formulated as second-order cone programs. Models incorporating the multidimensional nature of the problem become relevant to portfolio management, and multiobjective optimization has also been used to address conflicting objectives related to return and risk (Ehrgott et al., 2004).

Time series forecasting has been applied to the finance industry in applications such as stock market price. Due to the nonlinearity and high volatility of financial time series, accurate stock price/return prediction is difficult (Rezaei et al., 2021). This has motivated the recent development of neural network and machine learning strategies, which can automatically learn the temporal dependencies and patterns present in this type of data (see, for instance, Kim & Kim, 2019, Wang et al., 2020, Zhang et al., 2020). In the context of portfolio management, these forecasts can be used to preselect the assets with higher potential returns and to apply, in a second stage, the mean–variance model to determine their weights (Wang et al., 2020). Freitas et al. (2009) describe a neural network prediction-based portfolio optimization model that have the same statistical foundation of the mean–variance model. The differences are that this model uses predicted returns as expected returns and the variance of the prediction errors as risk measure. Return forecasts have also been used to generate investor views in Black–Litterman asset allocation modeling (Kara et al., 2019). The evolution of electronic trading has also promoted

* Corresponding author.
E-mail addresses: Ana.Corberan@uv.es (A. Corberán-Vallet), Enriqueta.Vercher@uv.es (E. Vercher), Jvsh@umh.es (J.V. Segura), Jose.D.Bermudez@uv.es (J.D. Bermúdez).

https://doi.org/10.1016/j.eswa.2022.119370
Received 7 April 2022; Received in revised form 2 August 2022; Accepted 26 November 2022
Available online 5 December 2022

the development of stochastic learning algorithms for online portfolio selection in the context of high-frequency trading (Li et al., 2022).

The use of soft computing techniques based on fuzzy set theory has also been applied to the portfolio selection problem to model the uncertainty regarding the future performance of asset returns using possibility and credibility distributions (Liu, 2006, Tanaka & Guo, 1999). These techniques provide new procedures and strategies to build efficient portfolios under realistic goals and constraints (see, for instance, Arenas et al., 2001, Gupta et al., 2013, León et al., 2002, Li et al., 2010, Saborido et al., 2016, Vercher & Bermúdez, 2013, Watada, 1997). Within this approach, the decision-making process usually establishes the portfolio selection problem as a multi-objective optimization problem.

In this work we adopt a new procedure to portfolio management that models the time series corresponding to the portfolio's value. Specifically, we define the value of a particular portfolio at the time of its acquisition as the weighted average of the values of the assets included in the portfolio. Assuming that the amount of shares held in each asset has been kept back in time, the time series corresponding to the portfolio's value is then defined using the historical prices of the different assets. Hence, our procedure considers the historical data available for all the assets without making any particular assumption regarding the relationship between the price movements of the different assets. Any type of existing correlation among the assets will be implicitly included in the series corresponding to the portfolio's value. The expected return and risk of a given portfolio can both be derived from the prediction of the portfolio's future value, which can be obtained using time series forecasting techniques. In particular, we employ an exponential smoothing model. Forecasts of future values will be used by a multi-objective genetic algorithm to provide a set of efficient solutions.

The remainder of this paper is structured as follows. In Section 2 we define the problem from our perspective and introduce the main notation. Section 3 is devoted to the description of the methodology that will be used afterwards. Section 4 describes the portfolio decision making process and the inputs of the proposed algorithm. Numerical results are presented in Section 5. We conclude the paper with some final comments and possible future directions.

2. Problem description and notation

The classical portfolio selection problem has the double objective of maximizing the return while minimizing the risk. Hence, it can be formulated as a non-linear multi-objective optimization problem, where an optimal portfolio is selected from among the feasible ones lying on the efficient front based on the notion of Pareto optimality (see, for instance, Ehrgott et al., 2004).

In this work we analyze the time series corresponding to the value of given portfolios. Forecasts of the future values of the portfolios allow us to estimate their expected return and risk at different future time periods. These estimates can then be used to define the objective functions of suitable multi-objective optimization problems.

2.1. Portfolio's value definition

Let us consider a universe of n risky assets that have remained in the market from $t = 1$ to $t = T$, where the time unit can be days, weeks, months or quarters, and T represents the point in time when the decision regarding the portfolio composition has to be made. Portfolio selection aims at selecting a portfolio that fulfills the investor's goal with respect to return and risk in a context of uncertain returns on the assets due to changes in their future price and financial market volatility. In practice, this means that one has to select a vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)'$ of weights, where x_i is the proportion of the capital M that is invested in the ith asset at time T. This allocation vector must provide a good balance between return opportunities and risk.

For each financial asset i, $i = 1, 2, \ldots, n$, we have the corresponding time series of historical prices: $(y_{i1}, y_{i2}, \ldots, y_{iT})'$. So, for a particular portfolio allocation at time T, the amount of shares held in asset i is given by $s_i(\mathbf{x}) = M \times x_i / y_{iT}$; that is, the amount of the capital that is invested in asset i divided by the current price of the asset.

Assuming that the amount of shares held in each asset has been kept back in time, we can define the time series corresponding to the portfolio's value $\mathbf{v}(\mathbf{x}) = (v_1(\mathbf{x}), v_2(\mathbf{x}), \ldots, v_T(\mathbf{x}))'$ as:

$$\begin{pmatrix} v_1(\mathbf{x}) \\ v_2(\mathbf{x}) \\ \vdots \\ v_T(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1T} \end{pmatrix} \times s_1(\mathbf{x}) + \ldots + \begin{pmatrix} y_{n1} \\ y_{n2} \\ \vdots \\ y_{nT} \end{pmatrix} \times s_n(\mathbf{x}). \tag{1}$$

We can apply a forecasting technique to model this time series and forecast the vector of future values, represented by $\mathbf{v}_f(\mathbf{x}) = (v_{T+1}(\mathbf{x}), v_{T+2}(\mathbf{x}), \ldots, v_{T+h}(\mathbf{x}))'$, h being the forecasting horizon. The time series corresponding to the historical portfolio's value does not typically have seasonal patterns. Hence, we recommend using an additive damped trend model, which is appropriate when there is a trend in the time series, but the growth rate observed at the end of the series may not persist in the long term (McKenzie & Gardner Jr, 2010). As explained afterwards, the mean vector and the covariance matrix of the distribution of $\mathbf{v}_f(\mathbf{x})$ conditioning on $\mathbf{v}(\mathbf{x})$ can be used to derive the expected return and risk of the portfolio. Additionally, other alternative measures of interest based on either the distribution of $\mathbf{v}_f(\mathbf{x})$ or the prediction intervals can be included in the decision-making process.

2.2. Portfolio selection problem

In a portfolio selection problem, the decision variables are the vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)'$ of weights, which indicate the proportion of the capital allocated to each of the n assets. These decision variables have to fulfill the non-negativity constraint, $x_i \geq 0$, and investment of the total capital, $\sum_{i=1}^n x_i = 1$. Additionally, other constraints may be considered; for instance, upper and lower bounds can be included to ensure diversification of the investment ($x_i \leq u_i$) and to prevent small investments in a number of securities ($l_i \leq x_i$). The cardinality constraint is also commonly incorporated to restrict the number of assets composing the portfolio: $k_l \leq \sum_{i=1}^n I_{(0,1]}(x_i) \leq k_u$ (where $I_A(x)$ is the indicator function). These constraints define the decision space:

$$\mathcal{S} = \left\{ (x_1, x_2, \ldots, x_n) \in \mathcal{R}^n \ : \ 0 \leq l_i \leq x_i \leq u_i \leq 1; \ \sum_{i=1}^n x_i = 1; \ k_l \leq \sum_{i=1}^n I_{(0,1]}(x_i) \leq k_u \right\}.$$

The conditional mean and variance of the portfolio's value at future time point $T + h$, $v_{T+h}(\mathbf{x})$, are related to its expected return and risk. Hence, efficient portfolios can be found by solving the following non-linear bi-objective optimization problem:

$$\begin{aligned} \text{MV(h):} \quad & \text{Max} \quad f_1(\mathbf{x}) = E(v_{T+h}(\mathbf{x}) | \mathbf{v}(\mathbf{x})) \\ & \text{Min} \quad f_2(\mathbf{x}) = V(v_{T+h}(\mathbf{x}) | \mathbf{v}(\mathbf{x})) \\ & \text{s.t.} \quad \mathbf{x} \in \mathcal{S} \end{aligned} \tag{2}$$

where $h \geq 1$ represents the investment horizon and it is specified by the investor. Therefore, depending on the interests of the investor, one may have to solve different problems associated with different values of h. It is important to emphasize here that, from a mathematical and computational viewpoint, the analysis of the problem does not depend on the value of h. However, since the prediction error increases with h, long-term horizons are not recommended. Alternative objective functions, such as the skewness or the trend of the series, can also be incorporated into the multi-objective optimization problem if required.

The cardinality constraint incorporated into the problem makes us face an NP-hard problem (Moral-Escudero et al., 2006). To obtain efficient portfolios, we apply an adaptation of the genetic algorithm proposed by Bermúdez et al. (2012). This adaptation improves the construction of the Pareto front and reduces the computation time.

Portfolio rebalancing may be needed during the investment horizon due to changes in the market. In our procedure, this can be done by modifying the constraints that define the decision space; for instance, varying the box constraints associated with specific assets. The assets composing the efficient portfolios then generated will meet the specified requirements.

3. Methodology

In the next two subsections, we describe the forecasting model and the multi-objective genetic algorithm that will be later applied in the portfolio decision making process.

3.1. Damped trend model

We consider the approach for the incorporation of the damped trend as proposed in the original work by Gardner and McKenzie (1985). Gardner's damped trend model assumes that the observed data $\{y_t\}_{t=1}^T$ are described by the following equation:

$$y_t = a_{t-1} + \phi b_{t-1} + \epsilon_t, \quad t = 1, 2, \dots, T,$$

where a_t and b_t represent, respectively, the level and trend at time t, $0 \le \phi \le 1$ is the damping parameter and $\{\epsilon_t\}$ are independent homoscedastic Normal random variables with zero mean and unknown variance σ^2. Holt's linear model is obtained when $\phi = 1$.

The level and trend are updated through the transition equations:

$$\text{Level:} \quad a_t = \alpha y_t + (1-\alpha)(a_{t-1} + \phi b_{t-1}) = a_{t-1} + \phi b_{t-1} + \alpha\epsilon_t,$$
$$\text{Trend:} \quad b_t = \beta(a_t - a_{t-1}) + (1-\beta)\phi b_{t-1} = \phi b_{t-1} + \alpha\beta\epsilon_t,$$

$0 \le \alpha, \beta \le 1$ being the smoothing parameters.

For parameter estimation we use the alternative formulation proposed by Vercher et al. (2012), which eases the joint estimation of the initial conditions $\omega = (a_0, b_0)'$ and the smoothing parameters $\theta = (\alpha, \beta, \phi)'$ through maximum likelihood. Using recursively the observation equation together with the transition equations, the observation at time t can be stated as:

$$y_t = a_0 + \Phi_t b_0 + \alpha(1 + \beta\Phi_{t-1})\epsilon_1 + \alpha(1 + \beta\Phi_{t-2})\epsilon_2 + \cdots + \alpha(1 + \beta\Phi_1)\epsilon_{t-1} + \epsilon_t,$$

where $\Phi_t = \sum_{i=1}^t \phi^i = \frac{\phi}{1-\phi} \times (1 - \phi^t)$ if $\phi \ne 1$ (in the case $\phi = 1$, $\Phi_t = t$). Hence, the model can be formulated as a linear heteroscedastic model whose matrix form is:

$$\mathbf{Y} = A\omega + L\epsilon, \tag{3}$$

where A is the $T \times 2$ matrix whose first column is the identity vector 1_T and its second one the vector $(\Phi_1, \Phi_2, \dots, \Phi_T)'$; L is the $T \times T$ lower triangular matrix whose main diagonal is equal to the identity vector 1_T and $l_{ij} = \alpha(1 + \beta\Phi_{i-j})$ for $i = 2, 3, \dots, T$ and $i > j$; and ϵ is the error vector. The joint distribution of the data vector $\mathbf{Y}$ is multivariate Normal with mean $E(\mathbf{Y}) = A\omega$ and covariance matrix $V(\mathbf{Y}) = \sigma^2 LL'$. This covariance matrix depends on the smoothing parameter vector θ. It is always a positive definite matrix since $|L| = 1$, whatever the value of θ is, so no constraints on the vector θ are necessary. However, we will assume that $\theta \in [0, 1]^3$.

3.1.1. Parameter estimation

The log-likelihood function of the data vector $\mathbf{Y}$ is proportional to:

$$\begin{aligned} log(f(\mathbf{Y}|\theta, \omega, \sigma^2)) &\propto -\frac{T}{2}log(\sigma^2) - \frac{1}{2\sigma^2}(\mathbf{Y} - A\omega)'(LL')^{-1}(\mathbf{Y} - A\omega), \\ &\propto -\frac{T}{2}log(\sigma^2) - \frac{1}{2\sigma^2}(L^{-1}(\mathbf{Y} - A\omega))'(L^{-1}(\mathbf{Y} - A\omega)), \\ &\propto -\frac{T}{2}log(\sigma^2) - \frac{1}{2\sigma^2}(\mathbf{Z} - C\omega)'(\mathbf{Z} - C\omega), \end{aligned}$$

with $\mathbf{Z} = L^{-1}\mathbf{Y}$ and $C = L^{-1}A$. Let P_C be the orthogonal projection matrix on the vectorial space generated by the columns of matrix C, $P_C = C(C'C)^{-1}C'$, and let $\tilde{\omega}$ be the usual least squares estimator of ω, $\tilde{\omega} = (C'C)^{-1}C'\mathbf{Z}$. The log-likelihood function can then be expressed as:

$$-\frac{T}{2}log(\sigma^2) - \frac{1}{2\sigma^2}\mathbf{Z}'(I - P_C)\mathbf{Z} - \frac{1}{2\sigma^2}(\omega - \tilde{\omega})'C'C(\omega - \tilde{\omega}).$$

The second quadratic form in the above expression can always be annulled, whatever the value of θ is, while the first quadratic form only involves parameter θ, which appears in matrix L and $C = L^{-1}A$. Hence, the maximum likelihood estimator of θ is obtained by solving the following non-linear optimization problem with respect to three decision variables:

$$\min_{\theta \in [0,1]^3} \quad (L^{-1}\mathbf{Y})'(I - P_C)(L^{-1}\mathbf{Y})$$

Once $\hat{\theta}$ has been obtained, the maximum likelihood estimator of ω is given by the vector $\tilde{\omega}$ computed at $\hat{\theta}$; that is:

$$\hat{\omega} = (\hat{C}'\hat{C})^{-1}\hat{C}'\hat{L}^{-1}\mathbf{Y} = (\hat{A}'\hat{L}'^{-1}\hat{L}^{-1}\hat{A})^{-1}\hat{A}'\hat{L}'^{-1}\hat{L}^{-1}\mathbf{Y}.$$

Finally, the maximum likelihood estimator of σ^2 is:

$$\hat{\sigma^2} = \frac{1}{T}(\hat{L}^{-1}\mathbf{Y})'(I - P_{\hat{C}})(\hat{L}^{-1}\mathbf{Y}).$$

3.1.2. Forecasting

Let $\mathbf{Y}_f = (y_{T+1}, y_{T+2}, \dots, y_{T+h})'$ be the $h \times 1$ vector of future values. Assuming that the joint vector $(\mathbf{Y}', \mathbf{Y}_f')'$ follows the damped trend model (3), we obtain:

$$\begin{pmatrix} \mathbf{Y} \\ \mathbf{Y}_f \end{pmatrix} = \begin{pmatrix} A \\ A_f \end{pmatrix}\omega + \begin{pmatrix} L & 0 \\ L_{f1} & L_{ff} \end{pmatrix}\begin{pmatrix} \varepsilon \\ \varepsilon_f \end{pmatrix},$$

where matrices A and L are partitioned in a similar way to the vector $(\mathbf{Y}', \mathbf{Y}_f')'$. Hence, the joint vector $(\mathbf{Y}', \mathbf{Y}_f')'$ follows a multivariate Normal distribution and the conditional distribution of the future data vector $\mathbf{Y}_f$ given the historical data vector $\mathbf{Y}$ is also a multivariate Normal distribution with the following mean vector and covariance matrix:

$$E(\mathbf{Y}_f|\mathbf{Y}) = A_f\omega + L_{f1}L^{-1}(\mathbf{Y} - A\omega), \tag{4}$$
$$V(\mathbf{Y}_f|\mathbf{Y}) = \sigma^2 L_{ff}L_{ff}'. \tag{5}$$

Let $\mu_{2.1} = A_f\hat{\omega} + L_{f1}L^{-1}(\mathbf{Y} - A\hat{\omega})$. If the parameter vector θ were known, the distribution of $\mathbf{Y}_f - \mu_{2.1}$ would be the multivariate Normal distribution with mean 0 and covariance matrix $\sigma^2 S$, where:

$$S = (A_f - L_{f1}L^{-1}A)(A'(L')^{-1}L^{-1}A)^{-1}(A_f - L_{f1}L^{-1}A)' + L_{ff}L_{ff}'.$$

Then, for any constant vector $v \ne 0$, the distribution of the random variable defined as:

$$\sqrt{\frac{T-1}{T}}\frac{v'(\mathbf{Y}_f - \mu_{2.1})}{\hat{\sigma}(v'Sv)^{\frac{1}{2}}}$$

is the Student's t-distribution with $T - 1$ degrees of freedom. Since θ is unknown in practice, we can substitute L with $\hat{L}$ in the previous equations. This result allows us to build exact prediction intervals for different goals.

For more details on parameter estimation and forecasting see Vercher et al. (2012) and Bermúdez et al. (2007), where the calculation of prediction intervals is also presented.

3.2. A multi-objective genetic algorithm

Multi-objective problems (MOP) arise when the goal is to optimize simultaneously r objective functions: $f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_r(\mathbf{x})$, which are usually in competition with each other. The decision variable vector $\mathbf{x} = (x_1, x_2, \dots, x_n)'$ usually has to fulfill some constraints, which define the decision space (or feasible region) that is denoted by $\mathcal{S}$ ($\mathcal{S} \in \mathcal{R}^n$). In general, a MOP can be formulated as:

$$\begin{aligned} \text{Max} \quad & [f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_r(\mathbf{x})] \\ \text{s.t.} \quad & \mathbf{x} \in \mathcal{S} \end{aligned}$$

Note that if any objective function is to be minimized, we could define $f_i(\mathbf{x}) = -\tilde{f}_i(\mathbf{x})$, $\tilde{f}_i(\mathbf{x})$ being the objective function to be minimized.

The evaluation function of a MOP maps the decision variable vectors $\mathbf{x}$ to vectors $\mathbf{z} = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_r(\mathbf{x}))'$, which constitute the objective space $\mathcal{Z} \in \mathcal{R}^r$.

For non-trivial MOPs, there exists a (possibly infinite) number of solutions which are found through the use of the Pareto Optimality Theory (Ehrgott, 2005). A solution $\mathbf{x} \in \mathcal{S}$ is Pareto optimal (or efficient) if the corresponding objective vector $\mathbf{z}$ cannot be improved in any dimension without another one degrading. Hence, without additional subjective preference information, all Pareto optimal solutions are equally good. Let $\mathcal{P}$ be the set of all Pareto optimal solutions (also called the efficient set). Their corresponding vectors $\mathbf{z}$ are termed non-dominated and constitute the Pareto front.

Bermúdez et al. (2012) proposed a bi-objective evolutionary algorithm for generating efficient portfolios of restricted cardinality in the case of $r = 2$. Let $\{\mathbf{x}_i\}_{i=1}^N$ be a initial population of N points randomly generated in the decision space. In a general bi-objective optimization scenario, an improved version of the algorithm can be described as:

Repeat:

1. *Evaluation.*
   Compute the objective vectors $\mathbf{z}_i = (f_1(\mathbf{x}_i), f_2(\mathbf{x}_i))$, $i = 1, 2, \dots, N$. Define φ as the $N \times 2$ matrix whose first column is the vector $(f_1(\mathbf{x}_{(1)}), f_1(\mathbf{x}_{(2)}), \dots, f_1(\mathbf{x}_{(N)}))'$ whose elements have been sorted in decreasing order and the second one is the corresponding vector $(f_2(\mathbf{x}_{(1)}), f_2(\mathbf{x}_{(2)}), \dots, f_2(\mathbf{x}_{(N)}))'$.
2. *Selection of efficient solutions.*
   Eliminate the rows of φ that are dominated (row j_2 is dominated if there is a row j_1, with $j_1 < j_2$, such that $f_2(\mathbf{x}_{(j_1)}) > f_2(\mathbf{x}_{(j_2)})$), and define the current upper bound as the polygonal chain whose vertices are the rows of the resulting matrix φ.
3. *Elitism.*
   Select the N/c ($c > 1$ being a given integer) feasible points whose objective vectors are closest to the upper bound in the current generation. Here the distance is defined as the minimum of the orthogonal distances to each one of the vertices of the polygonal chain.
4. *Mutation.*
   Mutate $c-1$ times each one of the N/c feasible solutions selected in the previous step, obtaining $N - N/c$ new solutions. The mutation operator slightly perturbs a pair of randomly selected elements of $\mathbf{x}_i$ taking into account the constraints that define the decision space.
5. *Definition of the next generation.*
   Build the new generation as the union of the sets obtained by elitism and mutation (so that the population size remains equal to N).

until the termination condition is fulfilled.

A reasonable termination condition is that the distance between two successive upper bounds is smaller than a given constant d. The last upper bound provides a rough approximation of the Pareto front.

4. Portfolio decision making process

Once the time series corresponding to the portfolio's value has been defined (see Eq. (1)), we can apply the damped trend model previously described to forecast the vector of future values $\mathbf{v}_f(\mathbf{x})$. The mean vector and the covariance matrix of the distribution of $\mathbf{v}_f(\mathbf{x})$ conditioning on $\mathbf{v}(\mathbf{x})$ are given, respectively, by Eqs. (4) and (5) substituting $\mathbf{v}(\mathbf{x})$ for $\mathbf{Y}$. Namely:

$$\begin{aligned} E(\mathbf{v}_f(\mathbf{x})|\mathbf{v}(\mathbf{x})) &= A_f\omega + L_{f1}L^{-1}(\mathbf{v}(\mathbf{x}) - A\omega), \\ V(\mathbf{v}_f(\mathbf{x})|\mathbf{v}(\mathbf{x})) &= \sigma^2 L_{ff}L'_{ff}. \end{aligned}$$

Point forecasts for the portfolio's future value at time $T+1, T+2, \dots, T+h$ are given by an estimate of the mean vector:

$$\hat{E}(\mathbf{v}_f(\mathbf{x})|\mathbf{v}(\mathbf{x})) = \hat{A}_f\hat{\omega} + \hat{L}_{f1}\hat{L}^{-1}(\mathbf{v}(\mathbf{x}) - \hat{A}\hat{\omega}), \tag{6}$$

where the hat symbol means that the parameters have been substituted for their maximum likelihood estimates. Similarly, the covariance matrix can be estimated as $\hat{\sigma}^2\hat{L}_{ff}\hat{L}'_{ff}$. However, we consider here the alternative estimate of the predictive variance that takes into account the uncertainty about ω and σ; that is:

$$\hat{V}(\mathbf{v}_f(\mathbf{x})|\mathbf{v}(\mathbf{x})) = \hat{\sigma}^2\hat{S}. \tag{7}$$

Let $\mu_{T+h}(\mathbf{x})$ be the hth element of the estimated mean vector and $\sigma^2_{T+h}(\mathbf{x})$ the element in row h and column h of the estimated covariance matrix. Then, $\mu_{T+h}(\mathbf{x})$ and $\sigma^2_{T+h}(\mathbf{x})$ estimate, respectively, $E(v_{T+h}(\mathbf{x})|\mathbf{v}(\mathbf{x}))$ and $V(v_{T+h}(\mathbf{x})|\mathbf{v}(\mathbf{x}))$. Consequently, the bi-objective optimization problem that we have to solve to find efficient portfolios (see Eq. (2)) can be finally formulated as:

$$\begin{aligned} \text{MV(h):} \quad & \text{Max} \quad f_1(\mathbf{x}) = \mu_{T+h}(\mathbf{x}) \\ & \text{Min} \quad f_2(\mathbf{x}) = \sigma^2_{T+h}(\mathbf{x}) \\ & \text{s.t.} \quad \mathbf{x} \in \mathcal{S} \end{aligned} \tag{8}$$

Let:

- n be the number of risky assets,
- $\mathbf{y}$ the $T \times n$ matrix whose columns are the time series of historical prices of the assets,
- M the capital to be invested at time T,
- h the future time point for which we want to calculate the return and risk,
- $\mathbf{l} = (l_1, l_2, \dots, l_n)'$ and $\mathbf{u} = (u_1, u_2, \dots, u_n)'$ the vectors of lower and upper bounds,
- k_l and k_u the minimum and maximum number of assets to be included in the portfolio,
- N the size of the population used in the genetic algorithm,
- c the integer used in the elitism step to prevent the disappearance of the best individuals,
- q_0 the maximum quantity that a share in a portfolio will decrease in the mutation step, and
- d the maximum distance allowed between two consecutive upper bounds of the genetic algorithm to reach the termination condition.

The workflow diagram of our algorithm is described in **Algorithm 1** GAPS.

5. Case study

To assess the performance of our procedure in a real context, we have considered the daily closing price time series of $n = 74$ healthcare stocks listed on the New York Stock Exchange. Their quotes were observed from July 15, 2020 to April 29, 2021, so that the time series include a total of 200 observations. Based on these 74 price series,

Algorithm 1 GAPS

1: **procedure** GENETIC ALGORITHM FOR PORTFOLIO SELECTION$(n, \mathbf{y}, M, h, \mathbf{l}, \mathbf{u}, k_l, k_u, N, c, q_0, d)$
2: Set $\mathcal{I} \leftarrow \emptyset$ and $i \leftarrow 1$.
3: **while** $i \leq N$ **do** ▷ generation of the initial population.
4: Select k, an integer varying in the set $\{k_l, k_l+1, \ldots, k_u\}$.
5: Generate randomly k integers between 1 and n. ▷ representing assets.
6: Simulate a vector $\mathbf{x}_i'$ of dimension k using the Dirichlet distribution of order k, with all its parameters equal to 1. ▷ equivalent to the uniform distribution over the $(k-1)$-simplex.
7: Create the portfolio $\mathbf{x}_i$ using the components of $\mathbf{x}_i'$ in an orderly way and adding 0s in those positions corresponding to the assets that have not been selected in 5.
8: **if** $\mathbf{x}_i \in \mathcal{S}$ **then**
9: $\mathcal{I} \leftarrow \mathcal{I} \cup \{\mathbf{x}_i\}$; $i \leftarrow i+1$.
10: **end if**
11: **end while**
12: Set $\mathcal{I} = \{\mathbf{x}_i\}_{i=1}^N$ as the current population and $\text{UB}^o = \emptyset$.
13: **repeat** ▷ evaluation.
14: **for all** $i \in 1 : N$ **do**
15: Compute the portfolio's value time series $\mathbf{v}(\mathbf{x}_i)$ (see Equation (1).
16: Apply the damped trend model to the series $\mathbf{v}(\mathbf{x}_i)$.
17: Estimate $E(\mathbf{v}_f(\mathbf{x}_i)|\mathbf{v}(\mathbf{x}_i))$ and $V(\mathbf{v}_f(\mathbf{x}_i)|\mathbf{v}(\mathbf{x}_i))$ using Equations (6) and (7).
18: Compute the objective vectors $\mathbf{z}_i = (f_1(\mathbf{x}_i) = \mu_{T+h}(\mathbf{x}_i), f_2(\mathbf{x}_i) = \sigma^2_{T+h}(\mathbf{x}_i))$.
19: **end for**
20: Define the matrix φ whose first column is the vector $(f_1(\mathbf{x}_{(1)}), f_1(\mathbf{x}_{(2)}), \ldots, f_1(\mathbf{x}_{(N)}))'$ (elements sorted in decreasing order) and the second one $(f_2(\mathbf{x}_{(1)}), f_2(\mathbf{x}_{(2)}), \ldots, f_2(\mathbf{x}_{(N)}))'$.
21: Eliminate the rows of φ that are dominated. ▷ selection of efficient solutions.
22: Set the current upper bound UB^{curr} as the polygonal chain with vertices the rows of φ.
23: **for all** $i \in 1 : N$ **do**
24: Compute d_i^{ub} : distance from $\mathbf{z}_i$ to the current upper bound UB^{curr}.
25: **end for**
26: *Elistism.* Define $\mathbf{E}$ as the $n \times N/c$ matrix whose columns are the N/c portfolios $\mathbf{x}_i$ in $\mathcal{I}$ with smallest d_i^{ub}.
27: *Mutation.* Set $\mathcal{MP} \leftarrow \emptyset$.
28: **for all** $i \in 1 : (N/c)$ **do**
29: $j \leftarrow 1$
30: **while** $j \leq c-1$ **do**
31: $\mathbf{x}_{ij}^m \leftarrow i$-th column of $\mathbf{E}$ ▷ portfolio to be mutated.
32: Select randomly 2 integers k_1 and k_2 between 1 and n.
33: Define $q_1 = \min\{q_0, \mathbf{x}_{ij}^m[k_1] - \mathbf{l}[k_1]\}$ and $q_2 = \min\{q_1, \mathbf{u}[k_2] - \mathbf{x}_{ij}^m[k_2]\}$.
34: $\mathbf{x}_{ij}^m[k_1] \leftarrow \mathbf{x}_{ij}^m[k_1] - q_2$; $\mathbf{x}_{ij}^m[k_2] \leftarrow \mathbf{x}_{ij}^m[k_2] + q_2$.
35: **if** $\mathbf{x}_{ij}^m \in \mathcal{S}$ **then** ▷ the cardinality constraint may have been violated.
36: $\mathcal{MP} \leftarrow \mathcal{MP} \cup \{\mathbf{x}_{ij}^m\}$; $j \leftarrow j+1$.
37: **end if**
38: **end while**
39: **end for**
40: Set $\mathcal{I} = \mathbf{E} \cup \mathcal{MP}$ as the current population and name its elements $\{\mathbf{x}_i\}_{i=1}^N$.
41: **if** $\text{UB}^o \neq \emptyset$ **then**
42: Compute $dist$: the maximum distance from any vertex in UB^o to UB^{curr}.
43: **else**
44: $dist \leftarrow \infty$
45: **end if**
46: Set $\text{UB}^o \leftarrow \text{UB}^{\text{curr}}$.
47: **until** $dist \leq d$
48: **end procedure**

the time series of the portfolio's value is computed for each of the portfolios randomly generated by the GAPS procedure (see Step 15 of the algorithm).

Fig. 1. Time plot of the series corresponding to the value of a naive portfolio acquired at time $T = 199$ and assuming a capital $M = 80$.

Our objective here is to apply the proposed algorithm to find efficient portfolios based on the bi-objective optimization problem defined in Eq. (8) for the forecasting horizon $h = 1$. We will then analyze the financial behavior of some efficient portfolios that make up the last upper bound UB^o. The selected portfolios are associated with three different profiles in terms of the risk that the investor is willing to endure: *low*, *moderate* and *high*.

Due to the computational complexity of the analysis, we have run some of the processes in parallel. The analysis has been carried out on a server with 2 Intel Xeon E5-2650 2.3 GHz processors, with 10 physical cores and 20 threads (2 processes per core) each. The parallelization of the processes has been carried out using the *R parallel* library (R. Core Team, 2021).

5.1. Performance analysis of GAPS algorithm

In this experiment we use the first $T = 199$ observations of the time series corresponding to a given portfolio's value $\mathbf{v}(\mathbf{x}_i)$ as the test set. Applying the damped trend model, we can estimate $E(v_{T+1}(\mathbf{x}_i)|\mathbf{v}(\mathbf{x}_i))$ and $V(v_{T+1}(\mathbf{x}_i)|\mathbf{v}(\mathbf{x}_i))$, where $T+1$ corresponds to April 29, 2021 (see Step 17 of the algorithm). Let $\mu_{T+1}(\mathbf{x}_i)$ and $\sigma^2_{T+1}(\mathbf{x}_i)$ be the corresponding estimates. For the efficient portfolios selected by the GAPS algorithm, we can calculate their expected return as:

$$r_{T+1}(\mathbf{x}_i) = \frac{\mu_{T+1}(\mathbf{x}_i) - v_T(\mathbf{x}_i)}{v_T(\mathbf{x}_i)}.$$

The expected risk is given by $\sigma_{T+1}(\mathbf{x}_i)$ (the square root of $\sigma^2_{T+1}(\mathbf{x}_i)$). From now on, we will simply refer to it as risk.

Let us first describe the computation of these quantities for a naive portfolio acquired at time $T = 199$ and assuming $M = 80$. In particular, we have randomly selected $k = 8$ assets among the $n = 74$ considered and we have assumed that the proportion of the capital M allocated to each of these assets is $1/8$ (that is, vector $\mathbf{x}$ has all its elements equal to 0 except those associated with the selected assets, which are equal to $1/8$). Based on the price of the assets at time $T = 199$, y_{iT}, the amount of shares held in each of them is given by $s_i(\mathbf{x}) = \frac{M}{8}\frac{1}{y_{iT}}$ if asset i has been selected. For the non-selected assets, $s_i(\mathbf{x}) = 0$. Using the time series of historical prices of the assets and $s_i(\mathbf{x})$, the time series corresponding to the value of the naive portfolio has been computed using Eq. (1). Fig. 1 shows the evolution of this time series from $t = 1$ to $t = T = 199$. The expected return and risk for this naive portfolio at time $T+1 = 200$ are, respectively, $r_{200}(\mathbf{x}) = 0.101$ and $\sigma_{200}(\mathbf{x}) = 1.529$.

The first step to implement the GAPS algorithm is the definition of the constraints that define the decision space. In this experiment we consider the following: the cardinality constraint that restricts the number of assets composing the portfolio assumes $k_l = 6$ and $k_u = 9$; to ensure diversification of the investment, we set $0 = l_i \leq x_i \leq u_i = 0.3$. On the other hand, the GAPS algorithm has been implemented so

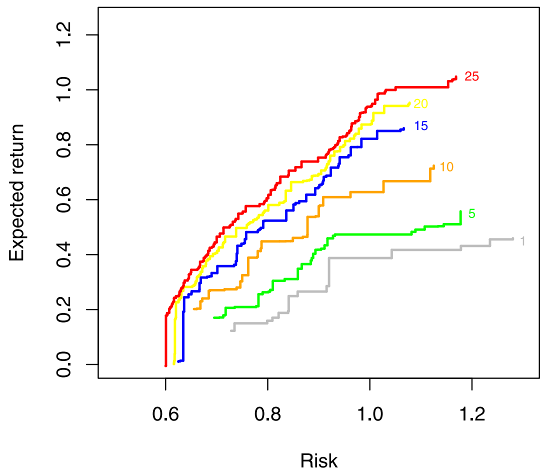

Fig. 2. Current upper bounds for generations 1, 5, 10, 15, 20, and 25 when the GAPS algorithm is applied to a population of $N_A = 600$ feasible portfolios.

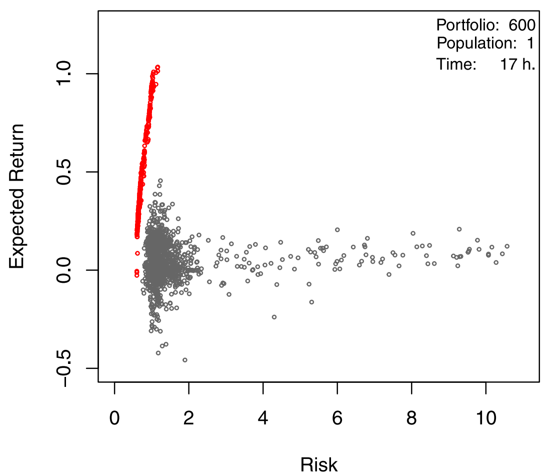

Fig. 3. Expected return and risk corresponding to the portfolios in the first (gray points) and last (red points) generations of the GAPS algorithm when it is applied to a population of $N_A = 600$ feasible portfolios.

that the elites contain 20% of the initial population size. Regarding the mutation parameter q_0, we have considered different values: $q_0 \in \{0.05, 0.02, 0.01, 0.0001\}$. Finally, the maximum distance allowed between two consecutive upper bounds to reach the termination condition is $d = 10^{-5}$.

Based on this definition of the decision space, we show now the results provided by the GAPS algorithm when it is applied to a population of $N_A = 600$ feasible portfolios. Fig. 2 shows the evolution of the upper bounds that the GAPS algorithm has generated at different iterations; namely, they correspond to generation 1, 5, 10, 15, 20, and 25, when the algorithm has reached the termination condition. This figure highlights both the improvement in the optimization process and the quasi continuity of the approximation of the Pareto front at generation 25.

To verify both that the initial random generation of the portfolios widely covers the objective space $\mathcal{Z} \in \mathcal{R}^2$ and that portfolios with a worse expected return–risk behavior are discarded by the optimization process, Fig. 3 shows the expected return and risk corresponding to the portfolios in the first (gray points) and last (red points) generations of the GAPS algorithm.

Finally, we assess if the performance of our procedure is affected by the size of the population used in the genetic algorithm. Fig. 4 shows the expected return and risk of the portfolios in the first (gray points)

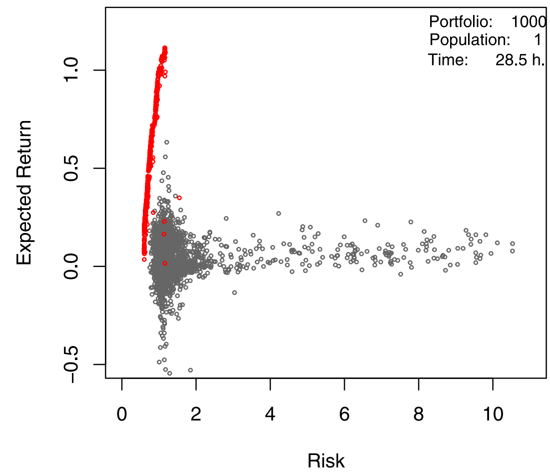

Fig. 4. Expected return and risk corresponding to the portfolios in the first (gray points) and last (red points) generations of the GAPS algorithm when it is applied to a population of $N_B = 1000$ feasible portfolios.

and last (red points) generations of the GAPS algorithm when it is applied to a population of $N_B = 1000$ feasible portfolios. As can be seen, the results provided in both cases are quite similar, which demonstrates the stability of our procedure. The biggest change is observed in the computation time, which increases from 17 to 28.5 h.

5.2. GAPS algorithm: A multi-start strategy

In this section, we evaluate the performance of our procedure when we implement a multi-start strategy that allows us to parallelize the generation of populations of feasible portfolios; that is, we can consider a single population of (for instance) $N_A = 600$ feasible portfolios or, alternatively, we can apply the GAPS algorithm to 6 different populations of size $N_{A_2} = 100$. In the latter case, an additional final step is included in the analysis to find efficient portfolios from among those that make up the different Pareto fronts obtained by applying the GAPS algorithm to each of the populations.

We consider here the following strategies:

- Strategy A2: The GAPS algorithm is applied to 6 different populations of size $N_{A_2} = 100$ feasible portfolios.
- Strategy B2: The GAPS algorithm is applied to 10 different populations of size $N_{B_2} = 100$ feasible portfolios.
- Strategy B3: The GAPS algorithm is applied to 5 different populations of size $N_{B_3} = 200$ feasible portfolios.

Fig. 5 shows the expected return and risk of the portfolios in the first (gray points) and last (red points) generations of the GAPS algorithm for these 3 strategies. The graphs shown in this figure combine the portfolios obtained when the GAPS algorithm is applied to each of the different populations considered. It is interesting to note that the multi-start strategy increases the time needed to solve the problem compared to using a single population of the same size. This increase in time is more noticeable for larger population sizes.

Fig. 6 shows the Pareto fronts (expected return–risk) generated by the GAPS algorithm in each of the strategies considered. Strategies A1 and B1 refer to the ones described in the previous section, when the GAPS algorithm is applied to a single population (of size $N_A = 600$ and $N_B = 1000$ respectively).

To better compare the final performance of these strategies, we have computed their hypervolume (HV), which is an indicator that evaluates the solution sets in terms of both convergence and diversity quality (Zitzler et al., 2003). So, the larger the value the better the computed Pareto front from a multiobjective perspective. We have

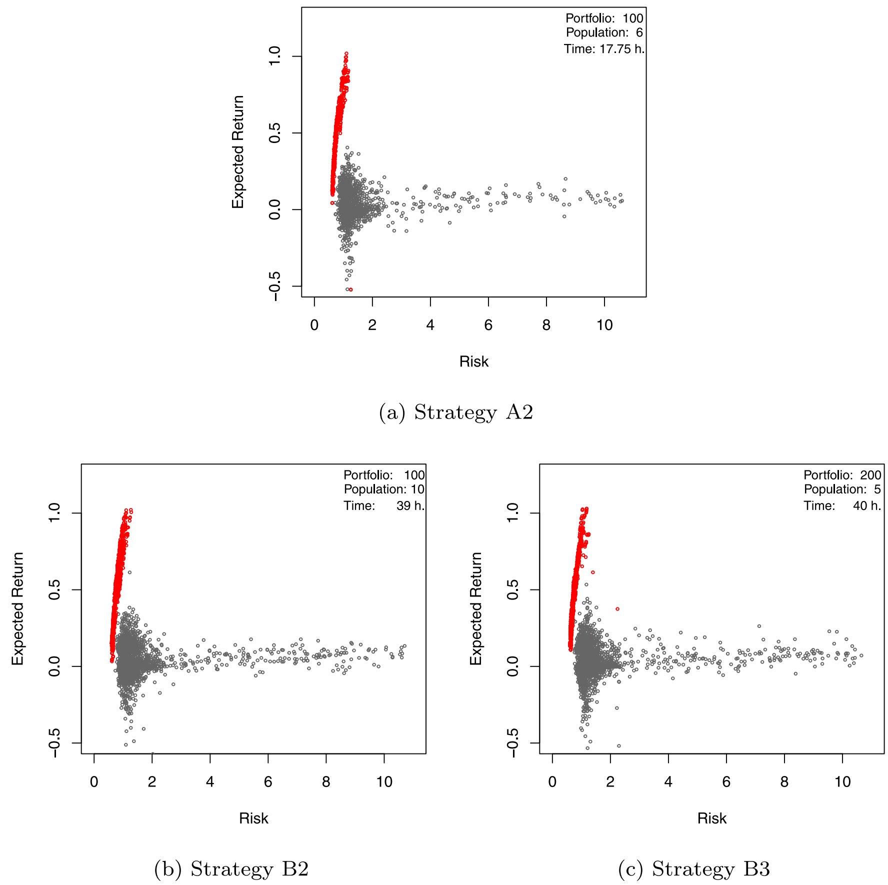

Fig. 5. Expected return and risk corresponding to the portfolios in the first (gray points) and last (red points) generations of the GAPS algorithm for the multi-start strategies A2, B2 and B3.

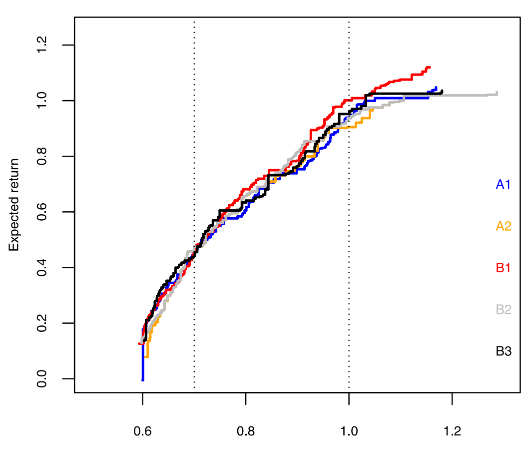

Fig. 6. Approximates of the Pareto front generated by the GAPS algorithm in each of the 5 strategies considered.

calculated the HV using a reference point which is worse than any point in the objective space. The coordinates of this nadir point are ER = 0.126 and risk = 1.156. The HV values corresponding to the 5 strategies are: $HV_{A1} = 0.378$, $HV_{A2} = 0.376$, $HV_{B1} = 0.403$, $HV_{B2} = 0.377$, and $HV_{B3} = 0.403$. These values together with Fig. 6 show the robustness of the results obtained with the proposed procedure.

5.3. Investment results: A comparative analysis

Finally, we describe in this section some features of the efficient portfolios provided by the GAPS algorithm and evaluate their ex-post performance.

For each strategy considered, Table 1 summarizes the number of efficient portfolios that are contained in the last upper bound, the range of the expected return and risk, the number of stocks (out of the $n = 74$ healthcare stocks considered) that compose those portfolios, and a match matrix indicating the number of assets that simultaneously appear in both strategies. It is worth emphasizing that there are 6 assets that are included in all strategies, with weights ranging between 60% and 75% in all portfolios. This proves that the different strategies coincide in the detection of the main investment assets.

We select now 3 efficient portfolios from each strategy, which correspond to a risk profile *low* (standard deviation below 0.7), *moderate* (standard deviation between 0.7 and 1) and *high* (standard deviation greater than 1). For the *low* and *high* profiles, we have chosen the portfolios with lowest and highest risk, respectively. For the *moderate* profile, the selected portfolios have both intermediate risk and expected return values. In almost all cases, efficient portfolios generated by the GAPS algorithm are made up of 9 assets, the maximum cardinality allowed. Table 2 shows the risk and expected return for these 15 selected portfolios at time $T + 1$ (April 29, 2021). As can be seen, strategy B1 provides slightly higher returns for a similar risk. Table 2 also shows the corresponding percentage of profit for the next 3 days, which allows us to evaluate their ex-post performance. This percentage has been calculated as the percentage change in the value of the portfolio with respect to time T (when the investment takes place), using new data on the price of the assets that make up the portfolio.

Table 1
Summary of some features of the efficient portfolios provided by the GAPS algorithm.

Strategy	# of efficient portfolios	Range of expected return	Range of risk	# of assets	Matching assets			
					A2	B1	B2	B3
A1	13	[0.00, 1.05]	[0.60, 1.17]	30	24	20	20	22
A2	10	[0.08, 1.02]	[0.61, 1.11]	31	–	18	22	21
B1	14	[0.13, 1.12]	[0.59, 1.16]	25	–	–	21	15
B2	10	[0.13, 1.03]	[0.60, 1.29]	30	–	–	–	17
B3	11	[0.14, 1.04]	[0.60, 1.18]	28	–	–	–	–

Table 2
Expected return and risk together with ex-post behavior of efficient portfolios provided by the GAPS algorithm for 3 risk profiles.

Risk profile	Strategy	Risk	Expected return	Percentage return		
				30/04	03/05	04/05
High	A1	1.17	1.05	0.68	2.02	2.43
	A2	1.11	1.02	0.58	1.38	1.97
	B1	1.16	1.12	0.79	2.18	1.61
	B2	1.29	1.03	1.21	3.54	3.44
	B3	1.18	1.04	0.30	1.61	1.72
Moderate	A1	0.76	0.58	−0.30	0.78	1.52
	A2	0.72	0.53	−0.13	0.98	0.73
	B1	0.77	0.62	0.13	1.14	0.78
	B2	0.80	0.66	0.07	1.19	0.42
	B3	0.75	0.61	−0.12	0.90	1.14
Low	A1	0.60	0.00	0.59	1.75	1.87
	A2	0.61	0.08	0.45	1.68	1.69
	B1	0.59	0.13	0.63	1.78	1.65
	B2	0.60	0.13	0.78	2.04	1.90
	B3	0.60	0.14	0.40	1.64	1.46
			Naive portfolio	−0.48	0.01	−0.57
			Dow Jones	−0.54	0.16	0.21

For comparative purposes, the return of the Dow Jones Index and a naive portfolio (where the proportion of the capital allocated to each of the $n = 74$ assets is $x_i = 1/74$) on those days is also indicated. These results prove the good ex-post investment performance of the GAPS algorithm, which provides efficient portfolios with larger percentage of profit than the Dow Jones Index. Finally, it is worth noting that portfolios corresponding to a risk profile *high* are preferable, the best results being those associated with strategy B2.

6. Conclusions

In this paper we have proposed a new method for portfolio selection that is based on the analysis of the time series corresponding to the value of given portfolios. For each particular portfolio, the series representing its value is computed using the time series of historical prices of the different assets. One of the main advantages of this approach is that possible correlations between the price movements of the individual assets will be implicitly incorporated into the portfolio value. So, there is no need to either assume independency to facilitate the analysis or estimate a covariance matrix.

Relatively easy to implement forecasting procedures can then be applied to model and predict these time series. Here we have implemented an alternative formulation of the damped trend model. The formulation of the model as a linear heteroscedastic model eases the joint estimation of the initial conditions and the smoothing parameters. This forecasting model allows us to predict the vector of future values of the time series corresponding to the value of the given portfolios. The mean vector and the covariance matrix of the predictive distribution can be used to derive the expected return and risk of each portfolio considered. In particular, a non-linear bi-objective optimization problem is defined using the estimated mean and variance of the portfolio's value at future time point $T + h$, where the value of h is specified by the investor.

This optimization problem allows us to take into account the multidimensional nature of the portfolio selection problem. Efficient portfolios of restricted cardinality are then found by applying a multiobjective genetic algorithm, the GAPS algorithm. As shown in the case study, this algorithm provides robust results in the different strategies considered regarding the number of populations and their sizes.

It is worth emphasizing that we have found here efficient portfolios when $h = 1$. However, the optimization problem can be extended to incorporate additional objectives corresponding to different investment horizons. Similarly, objectives of interest alternative to the mean and the variance can also be derived from the predictive distribution of the portfolio's value and incorporated into the optimization problem.

CRediT authorship contribution statement

Ana Corberán-Vallet: Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing, Supervision. **Enriqueta Vercher:** Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing. **José V. Segura:** Conceptualization, Methodology, Software, Validation, Formal analysis, Writing – review & editing. **José D. Bermúdez:** Conceptualization, Methodology, Software, Validation, Formal analysis, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors would like to thank the anonymous reviewers for their careful reading of our manuscript and their thoughtful comments and suggestions.

Funding

José V. Segura acknowledges the financial support from the Generalitat Valenciana, Spain under project PROMETEO/2021/063.

References

Arenas, M., Bilbao, A., & Rodríguez, M. V. (2001). A fuzzy goal programming approach to portfolio selection. *European Journal of Operational Research*, *133*, 287–297.
Bermúdez, J. D., Segura, J. V., & Vercher, E. (2007). Holt-winters forecasting: An alternative formulation applied to UK air passenger data. *Journal of Applied Statistics*, *34*, 1075–1090.
Bermúdez, J. D., Segura, J. V., & Vercher, E. (2012). A multi-objective genetic algorithm for cardinality constrained fuzzy portfolio selection. *Fuzzy Sets and Systems*, *188*, 16–26.
Ehrgott, M. (2005). Multicriteria optimization, (2nd ed.). Berlin: Springer.
Ehrgott, M., Klamroth, K., & Schwehm, C. (2004). An MCDM approach to portfolio optimization. *European Journal of Operational Research*, *155*, 752–770.

Freitas, F. D., De Souza, A. F., & de Almeida, A. R. (2009). Prediction-based portfolio optimization model using neural networks. *Neurocomputing*, *72*, 2155–2170.

Gardner Jr, E. S., & McKenzie, E. (1985). Forecasting trends in time series. *Management Science*, *31*, 1237–1246.

Goldfarb, D., & Iyengar, G. (2003). Robust portfolio selection problems. *Mathematics of Operations Research*, *28*, 1–38.

Gupta, P., Inuiguchi, M., Mehlawat, M. K., & Mittal, G. (2013). Multiobjective credibilistic portfolio selection model with fuzzy chance-constraints. *Information Sciences*, *229*, 1–17.

Kara, M., Ulucan, A., & Atici, K. B. (2019). A hybrid approach for generating investor views in Black-Litterman model. *Expert Systems with Applications*, *128*, 256–270.

Kim, T., & Kim, H. Y. (2019). Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data. *PLoS ONE*, *14*, Article e0212320.

Konno, H., & Yamazaki, H. (1991). Mean-absolute deviation portfolio optimization model and its applications to Tokyo Stock Market. *Management Science*, *37*, 519–531.

Lai, T. Y. (1991). Portfolio selection with skewness: A multiple-objective approach. *Review of Quantitative Finance and Accounting*, *1*, 293–305.

León, T., Liern, V., & Vercher, E. (2002). Viability of infeasible portfolio selection problems: a fuzzy approach. *European Journal of Operational Research*, *139*, 178–189.

Li, X., Qin, Z., & Kar, S. (2010). Mean-variance-skewness model for portfolio selection with fuzzy returns. *European Journal of Operational Research*, *202*, 239–247.

Li, Y., Zheng, X., Chen, C., Wang, J., & Xu, S. (2022). Exponential gradient with momentum for online portfolio selection. *Expert Systems with Applications*, *187*, Article 115889.

Liu, B. (2006). A survey of credibility theory. *Fuzzy Optimization and Decision Making*, *5*, 387–408.

Markowitz, H. M. (1952). Portfolio selection. *The Journal of Finance*, *7*, 77–91.

McKenzie, E., & Gardner Jr, E. S. (2010). Damped trend exponential smoothing: A modelling viewpoint. *International Journal of Forecasting*, *26*, 661–665.

Moral-Escudero, R., Ruiz-Torrubiano, R., & Suarez, A. (2006). Selection of optimal investment portfolios with cardinality constraints. In *IEEE congress on evolutionary computation* (pp. 2382–2388).

R. Core Team (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rezaei, H., Faaljou, H., & Mansourfar, G. (2021). Stock price prediction using deep learning and frequency decomposition. *Expert Systems with Applications*, *169*, Article 114332.

Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, *2*, 21–41.

Saborido, R., Ruiz, A. B., Bermúdez, J. D., Vercher, E., & Luque, M. (2016). Evolutionary multi-objective optimization algorithms for fuzzy portfolio selection. *Applied Soft Computing*, *39*, 48–63.

Tanaka, H., & Guo, P. (1999). Portfolio selection based on upper and lower exponential possibility distributions. *European Journal of Operational Research*, *114*, 115–126.

Vercher, E., & Bermúdez, J. D. (2013). A possibilistic mean-downside risk-skewness model for efficient portfolio selection. *IEEE Transactions on Fuzzy Systems*, *21*, 585–595.

Vercher, E., Corberán-Vallet, A., Segura, J. V., & Bermúdez, J. D. (2012). Initial conditions estimation for improving forecast accuracy in exponential smoothing. *TOP*, *20*, 517–533.

Wang, W., Li, W., Zhang, N., & Liu, K. (2020). Portfolio formation with preselection using deep learning from long-term financial data. *Expert Systems with Applications*, *143*, Article 113042.

Watada, J. (1997). Fuzzy portfolio selection and its application to decision making. *Tatra Mountains Mathematical Publications*, *13*, 219–248.

Zhang, Y., Yan, B., & Memon, A. (2020). A novel deep learning framework: Prediction and analysis of financial time series using CEEMD and LSTM. *Expert Systems with Applications*, *159*, Article 113609.

Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C. M., & Grunert da Fonseca, V. (2003). Performance assessment of multiobjective optimizers: An analysis and review. *IEEE Transactions on Evolutionary Computation*, *7*, 117–132.