Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v8n1p44

Double clustering for rating mutual funds

By Menardi, Lisi

Published: 26 April 2015

Electronic Journal of Applied Statistical Analysis
Vol. 08, Issue 01, April 2015, 44-56
DOI: 10.1285/i20705948v8n1p44

Double clustering for rating mutual funds

Lisi Francesco, Menardi Giovanna*

University of Padua, Department of Statistical Sciences
via C. Battisti 241, 35100, Padova, Italy

Published: 26 April 2015

Due to the increasing proliferation of mutual funds, in-depth evaluation of the available products for portfolio selection purposes is a difficult task. Hence, classification schemes giving quick information about which funds are worth to be monitored, are often provided. The aim of this work is to show an application of clustering methods to the mutual funds historical data. Starting from the monthly time series of the Net Asset Values of a specific style-based category, namely the Large Blend US mutual funds, we apply distance-based clustering methods twice on a set of return, risk and performance measures: firstly, with the aim of reducing data dimension, and secondly to cluster funds in homogeneous classes. The adopted procedure claims the feature of producing a partition of funds that are readily interpretable from a financial point of view and it is further possible to rank the identified groups, thus obtaining a rating of funds that turns out to account for different propensities toward the risk exposure.

keywords: cluster analysis, dimension reduction, mutual funds, performance measures, portfolio selection

1 Introduction

In the last decades, the mutual funds industry has experienced a period of tremendous growth, especially in the 1980s and 1990s. Despite the 2003 mutual funds scandal and the global financial crisis of the last few years, the number of available funds is still growing, amounting now to more than ten thousands in the US only.

*Corresponding Author: menardi@stat.unipd.it

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

Several reasons explain the spreading success of the mutual funds market: investing in funds allows the saver to have access to a diversified portfolio, while subscribing even a single share; usually transaction costs are lower than the corresponding costs for directly purchasing the same amount of investment; most of mutual funds feature a good level of market liquidity. In addition, one main reason for the increasing proliferation of mutual funds is that most of the savers do not have time nor sufficient skills to research, select, and monitor individual stocks: investing in funds allows for commissioning choices about timing and conditions of financial tools from the investment management companies.

Nonetheless, the wide range of available products still makes difficult evaluating and choosing a set of funds with good expected performance for investment purposes. National and international rating agencies and associations of mutual funds managers aim at helping such a choice, by providing a number of classifications that should give reliable information to the investors to quickly and easily identify funds that are worthy of be further kept watch over.

In the same perspective, this work aims at taking advantage of statistical methods as cluster analysis to classify mutual funds. Our work differs from other classification schemes well known and proposed in the literature (see the next section for further details and references). First, our goal is neither to identify mutual fund investment styles or to evaluate the consistency of the characteristics of funds with their stated objectives. Instead, by focusing on a given category of funds, whose components do not differ too much according to their risk factor exposure, we aim at distinguishing groups of funds with similar behaviour, so that funds in the same group can be considered reasonable substitutes for the purposes of portfolio construction. Secondly, we try to retain as much information as possible about different aspects of funds performance, thereby making use of a multivariate set of measures. Instead, the need of easy interpretation of groups has often motivated the use of low- or even uni-dimensional data to carry out the classification task.

To the end of comparing ex-post managed portfolios, financial literature provides several dozens of performance indexes, usually in the form of reward measures adjusted for their risk. The proliferation of these measures is motivated by the use of alternative ways of approximating the non-observable volatility component. Examples are the Sharpe ratio and its variants, indexes based on lower partial moments such as the Omega index, performance measures hinging on the drawdown-based evaluation of the risk as the Calmar ratio, and others (for a review, see, for instance, Eling (2008); Cogneau and Hübner (2009)). Even if several studies conjecture the redundancy and the virtual equivalence of some of these measures, most of them still provide useful information to evaluate the asset behaviour, by offering different perspectives to understand the past performance of financial assets.

These considerations suggest that producing classifications based on a plurality of ratios would be more appropriate than using one single measure, because different aspects of the past behaviour of the funds could be taken into account. On the other hand, using multidimensional data for classification often prevents from interpreting the features of the groups. Moreover, the use of multidimensional data does not allow for a subsequent ranking of groups.

Based on a set of US mutual funds, in this work we first perform a dimension reduction step where several associated measures of return, risk and performance are clustered and then combined. This preliminary step allows us to obtain a restricted set of latent measures whose meaning is immediate and easily interpretable. Subsequently, funds are clustered based on the identified measures. Notably, it is further possible to rank the groups, thus obtaining a rating of funds that turns out to account for different levels of risk aversion.

In the reminder of the paper, we review some schemes for classifications of funds that are described in the literature or proposed by the rating agencies (Section 1.1). After describing the data in Section 3, we present the proposed methodology for clustering the funds (Section 4) and the results (Section 5). Some concluding remarks are outlined in Section 6.

1.1 Review of mutual funds classifications

The wide diffusion of mutual funds has lead the rating agencies and associations of mutual fund managers to define several classification schemes in order to give the investors some information about the past performance of different categories of mutual funds. The Star Rating Morningstar (2007), for instance, is the classification of funds published by Morningstar. It is intended for use as the first step only in the fund assessment process, but although Morningstar warns against using the Star Rating as a basis for investment decisions, its strong influence on the investors behaviour has been widely reported. Classification is based on the Morningstar risk-adjusted return (MRAR) indicator, that accounts for variations of a fund monthly performance (including the effects of sales charges, loads and redemption fees) and whose expression is motivated by expected utility theory. The MRAR gives a fair indication about the fund behaviour, and it is easy to interpret. Moreover, using it as a basis for classify the funds into homogeneous groups has the further advantage of allowing for rating the funds, because one performance measure only is considered.

Financial literature also provides some hints to help the evaluation of performance of funds. Lisi and Otranto (2010), for instance, propose a two-steps clustering procedure that aims at bypassing the difficulties of interpreting groups obtained by using both return and risk measures.

Further classification schemes proposed in the literature follow more traditional approaches and use multidimensional data for clustering the funds. However, results turn out to be difficult to construe and the interpretation of cluster characteristics is based on the comparison to the stated objectives or the investment styles of the funds. An example is the work of Lytkin et al. (2008) that compare results of three clustering algorithms and conclude that the overlapping part of such partitions can be considered as the most robust and objectively consistent component in the existing classification of mutual funds by their management styles. Das (2005) uses nonhierarchical clustering for classifying hedge funds on the basis of asset class, size of the hedge funds, incentive fee, risk level, and liquidity. Comparing the groups with the existing classification of the ZCM/Hedge database, the author show that not completely consistent results arise. Similarly, Sidana and Acharya (2007) attempt to classify hundred mutual funds employing cluster analysis and using a limited set of performance measures. They find evidence of inconsistencies between the investment style/objective classification and the fund returns. Lajbcygier and Yahya (2008) compare two clustering methods for an out-of-sample prediction of mutual fund performance. Kim et al. (2000) approach the problem from a different perspective, and apply supervised methods of classification such as discriminant analysis instead of clustering, to understand the consistency of mutual fund behaviour with their stated objectives.

An alternative approach to classify the funds exploits temporal information to mine groups with similar behaviour and then predict their future evolution. However, this approach may result in groups with similar trend and evolution but not necessarily similar levels of risk-adjusted return. Examples of this approach can be found in Pattarin et al. (2004); Corduas and Piccolo (2008).

2 Material and methods

2.1 The data

We focus the attention on a specific style-based category, namely the Large Blend funds. The Morningstar classification of funds into the cross categories Large caps/Mid caps/Small caps and Growth/Blend/Value is well known and often used in determining how an investment fits into a particular portfolio from an asset allocation perspective. Large Blend funds tend to invest across the spectrum of U.S. Industries and have fairly representative portfolios of the overall stock market in both size, growth rates, and prices. Other style categories of funds could be taken into consideration but our choice of zooming on a specific style category is motivated by the idea that while the investor is able to make a preliminary rough selection of his propensity to different levels of risk exposure, he needs some tools to further discriminate between the assets within the chosen style-category.

The considered data refer to the Net Asset Values (that are the analogous for managed portfolios of prices for stocks) of all the Large Blend funds that were at least three years old at the moment of data collection. Starting from the monthly time series, the excess returns have been computed over time on the basis of the minimum risk free return[1]. The resulting data set amounts to 1436 funds observed from April 2007 to March 2010.

Funds younger than three years or no longer existing at time of data collection (April 2010) have been excluded from analysis. In principle, the resulting dataset suffers from incompleteness and survivorship bias (see, e.g., Elton et al., 1996). However, our purpose is not to provide an evaluation of fund performance, and the use of incomplete information on complete data could yeld to even more misleading results. Thus, the need of handle an unchanging sample of assets over the study period motivates our choice.

[1]Data have been drawn from the Treasury Constant Maturity Rates, source: http://www.treas.gov/offices/domestic-finance/debt-management/interest-rate/index.html

The features and the behaviour of funds over time have been, then, summarized by computing some measures of performance, return and risk over the considered three-years period. The estimated measures are listed below. For further details see Table 1. Refer to Cogneau and Hübner (2009) for a comprehensive review, and to Lönnbark et al. (2011), Cheridito and Stadje (2009) for more specific references.

- *Performance measures*: some indexes derived from the Capital Asset Pricing Model (CAPM) have been considered, namely the Treynor index (denoted as T in the rest of the paper), the Appraisal ratio (A); the Sharpe index (SH); some measures whose approach to evaluate the risk is based on partial moments as the Sortino ratio (SO), the Kappa 3 measure ($K3$), the Omega index (O) and a variant of such indexes introduced by Farinelli and Tibiletti, comparing an upper partial moment of order p with a lower partial moment of order q. Two different choices for the parameters, accounting for diverse investor style or preferences have been considered, namely ($p = 0.5$, $q = 2$, $FT1$) for a defensive investor and ($p = 3$, $q = 0.5$, $FT2$) for an aggressive strategy; two measures based on the Drawdown, namely the Burke index (denoted as B) and the Sterling ratio (ST).
- *Reward measures*: we have considered the expected excess return based on the monthly risk free returns (denoted, in the following as R) and the expected tail gain at a level $q = 0.05$ and $q = 0.10$ (corresponding to the expected gain of the top $100 \cdot q$ of the distribution of returns, denoted as $ETGq$). Moreover, the $1 - q$ empirical quantile of the returns distribution (*i.e.* the return level exceeded by the $100 \cdot q$ of the returns) has been considered as a further measure of reward, with $q = 0.05$ and $q = 0.10$ (TGq). Finally, being the Jensens Alpha (J) the intercept of the CAPM, it has been included among the measures of reward.
- *Risk measures*: probably the most common indicator to evaluate the risk is the standard deviation of returns (in the following denoted as σ); the semi-standard deviation (that is the standard deviation of the negative returns, also named lower partial moment of second order, $LPM2$), has been also considered, together with the expected tail loss at the 5 and 10% q level ($ETLq$), the Value at risk at the 5 and 10% q level ($VARq$), the maximum drawdown (MD).

The dataset so formed is then represented as a matrix with 1436 rows (corresponding to the mutual funds) and 24 columns (corresponding to the return, risk and performance measures).

2.2 Classification of mutual funds

The mainstream approach to pursue the goal of partitioning objects into a number of groups which share common characteristics, is typically achieved by defining a suitable measure of dissimilarity between the objetcs to be clustered. The objects are then partitioned to maximize the similarity within each group and/or the dissimilarity between the groups. The actual implementation of this idea is what characterizes different methods of cluster analysis.

Table 1: Reward, risk and performance indexes included in the application and notation. r_t denotes the return of the asset at time t, $r_{f,t}$ is the risk free rate, $r_{B,t}$ is the return of the market index.

Return			
Expected excess return	$\overline{R}$	$\sum_{t=1}^{T} \frac{r_t^*}{T}$	$r_t^* = (r_t - r_{f,t})$
Tail gain	$TG(q)$	$r_{t_0}^* : \frac{\#\{r_t^* \geq r_{t_0}^*\}}{T} = q$	$q := 0.05$, $q := 0.10$, T denotes the considered time window
Expected tail gain	$ETG(q)$	$\sum_t q \cdot r_t^* \mathbf{1}\{r_t^* \geq TG(q)\}$	$q := 0.05$, $q := 0.10$
Jensen's Alpha	α	α	estimated intercept of the CAPM line: $r_t^* = \alpha + \beta(r_{B,t} - r_{f,t}) + \varepsilon_t$
Risk			
Standard deviation of returns	σ	$\sum_{t=1}^{T} \frac{(r_t^* - \overline{R})^2}{T-1}$	
Lower partial moment	$LPM_k(\tau)$	$\sum_{t=1}^{T} \frac{max(\tau - r_t, 0)^k}{T}$	$k := 2$, semistandard deviation $\tau := 0$, minimum acceptable return
Value at risk	VAR(q)	$r_{t_0}^* : \frac{\#\{r_t^* \leq r_{t_0}^*\}}{T} = 1 - q$	
Expected tail loss	ETL(q)	$\sum_t q \cdot r_t^* \mathbf{1}\{r_t^* \leq VAR(q)\}$	
Maximum drowdown	MD	$\min_t r_t^*$	
Performance			
Treynor index	T	$\frac{\overline{R}}{\beta}$	β is the estimated coefficient of the CAPM line, as defined above.
Appraisal ratio	A	$\frac{\alpha}{\sigma_\varepsilon}$	α and σ_ε are the estimated intercept and the standard deviation of the residuals of the CAPM line, as defined above.
Sharpe ratio	SH	$\frac{\overline{R}}{\sigma}$	
Sortino ratio	SO	$\frac{\overline{R} - \tau}{\sqrt{LPM_2(\tau)}}$	
Kappa 3 index	$K3$	$\frac{\overline{R} - \tau}{\sqrt[3]{LPM_3(\tau)}}$	
Upside potential ratio	UP	$\frac{HPM_1(\tau)}{\sqrt{LPM_2(\tau)}}$	$HPM_k = \sum_{t=1}^{l} \frac{max(r_t - \tau, 0)^k}{l}$
Omega index	O	$\frac{\overline{r} - \tau}{LPM_1(\tau)} + 1$	
Farinelli-Tibiletti index	FT_1, FT_2	$\frac{\left[\frac{\sum_{t=1}^{T} max(r_t - \tau, 0)^p}{T}\right]^{\frac{1}{p}}}{\left[\frac{\sum_{t=1}^{T} max(\tau - r_t, 0)^q}{T}\right]^{\frac{1}{q}}}$	$(p = 0.5, q = 2)$, $(p = 3, q = 0.5)$
Burke ratio	B	$\frac{\overline{r} - r_f}{\sqrt{\sum_{s=1}^{S} \frac{MD_s^2}{S}}}$	MD_s denotes the s^{th} lowest return; $S := 3$, the nearest integer to $T/10$
Sterling ratio	ST	$\frac{\overline{r} - r_f}{\sum_{s=1}^{S} \frac{-MD_s}{S}}$	

In this work we follow a hierarchical agglomerative approach, which starts with assigning each observed object to a different cluster, i.e. there are as many clusters as objects. Then, clusters are aggregated sequentially, until only one cluster is left, to form a hierarchical structure called dendrogram. Aggregation of clusters is based on the minimization of some criterion of dissimilarity between clusters. According to the complete linkage criterion, adopted in this work, two clusters are aggreagted when the maximum dissimilarity between their elements is minimum.

The described technique could be, in principle, directly applied to our set of funds in order to identify a number of groups sharing some common characteristics. On the other hand, classification of the assets based on the whole set of measures is not unadvisable, as it is likely to lead to clusters that are not easily interpretable from a financial point of view. Funds partition is performed to get groups of funds that are essentially exchangeable for portfolio selection purposes. Then, a lack of interpretability of the cluster composition would make the classification pointless. This reason, together with the will of retaining as much information as possible about different aspects of funds performance, motivates the choice of adopting the following classification scheme:

1. *Measures clustering*
   Instead of clustering the funds, we first run cluster analysis on the measures of performance, return and risk. This allows us to obtain groups of measures which are homogeneous with respect to the information they explain. The similarity among the measures is evaluated by exploiting their structure of rank correlation, so that two measures will be considered as virtually equivalent if they produce the same ranking over the assets. The rank correlation matrix of the measures is converted to a dissimilarity matrix by applying the inverse cosine function. Next, the measures are clustered by following a complete-linkage criterion.

2. *Reduction of dimensionality*
   Subsequently, information is summarized by a suitable combination of the measures within each cluster. A simple approach for defining a new composite index in each group, consists of assigning the same weight to the single measures, that is producing a simple mean of the variables. However, this approach does not guarantee any property of optimality of the new composite index, that could in principle favour the preservation of redundant information. Instead, we perform the reduction of dimensionality of each group of measures by means of Principal Component Analysis (PCA), that find low-dimensional representations of the multivariate data by retaining as much as possible of the variability of the funds. It might be argued that Principal Component Analysis could be applied on the whole set of measures directly. However, the application of PCA separately on each group of variables allows us to preserve more information about the original data. Moreover, the more correlated the original measures are, the more variation in the data is explained by the first principal components. Lastly, combining the variables within each group (that is, combining the most correlated measures only) may help in providing a meaningful interpretation to the detected principal components.

   Principal Component Analysis has been performed on the standardized variables in order not to be affected by the scale of the measures. Next, the first component is extracted from each group, thus obtaining a number of new composite indexes corresponding to the number of selected groups of measures.

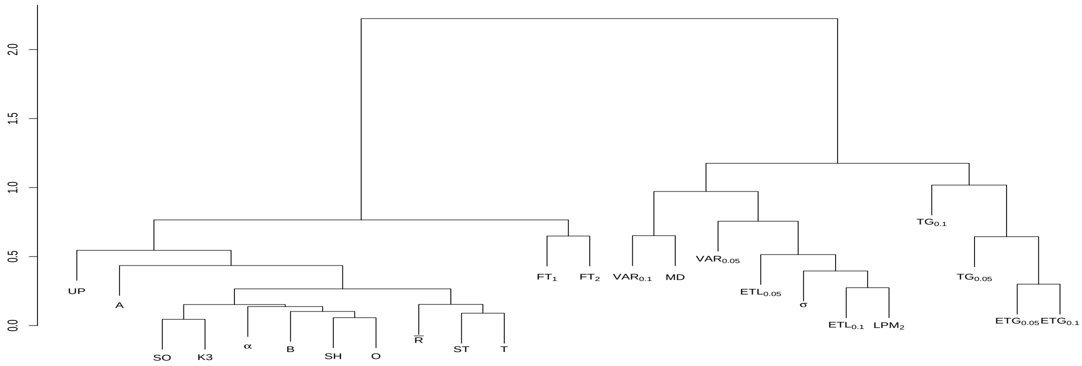

Figure 1: Dendrogram of the clustered measures of performance, return and risk.

3. *Fund clustering*
   Through the application of the previous step, we come to the definition of a new set of data having lower dimension than the number of return, risk and performance measures originally estimated. It is, then, possible to compute a distance matrix between the scores of the detected principal components at each observation (namely, the fund), thus gaining a remarkable computational saving. Next, clustering of the observations is performed by grouping the funds presenting minimum distance. In particular, the euclidean metric has been chosen to compute the distances and a hierarchical approach has been followed to cluster the groups, again according to the complete-linkage method.

3 Results

Results from the application of the described procedure on our set of mutual funds allow for drawing interesting considerations from a financial point of view.

In the first step of the classification scheme already, some configurations of the clustered variables appear particularly meaningful. In Figure 1 1 the dendrogram produced by the complete-linkage method on the estimated 1 the dendrogram produced by the complete-linkage method on the estimated measures of return, risk and performance is reported. The tree is upside-down with respect to the dissimilarity, meaning that the more similar the measures, the lower level of the hierarchy.

Two large clusters are clearly identifiable. The first one gathers most of the risk-adjusted return indexes at a very low level of distance. Since the measure of distance between variables is a decreasing function for positive values of the rank correlation index, this behaviour highlights that most of the performance measures provide very similar information in terms of ranking of the funds. Also two measures of reward (namely the expected excess return and the Jensen's Alpha) belong to this group of virtually equivalent variables, suggesting that the risk level of the considered funds does not change remarkably (this is consistent with the choice of working with funds belonging to the same style based category). A slightly different information in terms of rankings of funds is given by the Appraisal ratio, the Upside potential ratio, and the two parameterizations of the Farinelli-Tibiletti index, which aggregate into the same cluster at a higher level of the hierarchy.

A second large cluster includes the remaining measures, but two subgroups are clearly discernible, gathering the whole set of risk indexes and the measures of reward, respectively. Unlike the expected excess return and the Jensen's Alpha (allocated to the first large cluster), the latter measures provide information about the upper tail of the distribution of returns.

The described three clusters have been selected as a first phase, to be subsequently passed to the dimensionality reduction step of the procedure. The choice of cutting the dendrogram to form three groups has then mainly followed the need of interpretability of results instead of the goal of optimizing the intra-group homogeneity or the separation between clusters. By applying PCA within each group, we then obtain three composite indexes having an easily interpretable financial meaning: the first new measure is a combination of the return-to risk ratios and the two reward indexes R and α, thus providing indication about the location of the distribution of returns (in the rest of the paper it will be referred to, as the performance component). The second index describes the right tail of the distribution of returns, thus informing about the potentiality of gain that different funds offers (upper extreme return component). The third component can be legitimately considered as a risk measure, because it linearly combines the considered risk indexes only. In the rest of the paper it will be denoted as the risk component.

Table 2 lists the factor loadings of the three composite indexes, describing the correlation between each index and the measures by which it is formed: it can be noticed that the size of the contribution of each original variable to each composite index is approximatively the same. Among the performance measures, the lowest contribution is clearly given by those variables that in the first step of the procedure were clustered at a larger level of distance from the other components of the groups (that is, with a lower correlation), namely the Farinelli-Tibiletti indexes, the Appraisal ratio and then the Upside Potential ratio. Concerning the extreme return component, the 90% quantile of the distribution of returns ($TG0.10$) gets a slightly lower weight within the composite index while in the risk component a relatively higher weight is given to the semi-standard deviation and to the expected tail loss at a 10% level. The three extracted components explain a very large quote of the original variability of the data within the three groups, ranging from the 84% to the 93%.

The subsequent application of the complete-linkage clustering method allows to identify a final partition of funds. The ratio between the average distance within and between clusters, computed for different configurations of groups, suggest to partition the data into 6 clusters.

The distributions of the funds according to the three composite indexes, conditional to the cluster membership, are plotted in Figure 2, 2, where the cluster labels have been sorted according to the median value of the performance component. The observation of the performance component within the groups does not allow us to draw any conclusion about a clear separation between the six detected clusters: the distributions overlap almost completely both in groups 3 and 4, and distribution in clusters 1 and 2 overlap for almost the 50% of funds. On the other hand, when observing the distributions of the funds with respect to the upper extreme return and risk components, some remarkable differences can be noticed. Within each pair formed by groups $\{1, 2\}$, $\{3, 4\}$, and $\{5, 6\}$, one of the two groups presents an high level of risk and a high potential extreme return while the other group shows evidence of a lower risk and a lower extreme return. These considerations suggest to interpret the clusters as rated in a system having three classes, each of them corresponding to a pair of groups. Within each class, funds may be further partitioned into two twin sub-categories having similar performance but a different degree of risk and different extreme gains. In Figure 3 3 the distributions of funds with respect to the three composite indexes are reported, by distinguishing between funds having high risk and high extreme returns and funds having lower risk and lower extreme returns. Splitting the twin clusters into the higher and lower risk sub-categories allows for valuing more the three emerged rating classes, which winner, losers and intermediate funds belong to. Unlike the considerations drawn by jointly observing the six groups, when this further division is made explicit, it is possible to note that the three classes of funds are well separated with respect to their performance, especially in the low risk sub-category. Instead, the distributions of extreme returns and risk mainly overlap at least in two of the three classes.

Table 2: For each component, the loadings define the contribution of the original measures. The bottom line reports the proportion of the variance of the original data explained by each component.

Performance component		Extreme return component		Risk component	
$\overline{R}$	0.28	ETG_5	0.51	σ	0.39
α	0.28	TG_5	0.51	TL_5	0.37
SH	0.29	ETG_{10}	0.54	ETL_5	0.37
SO	0.29	TG_{10}	0.44	TL_{10}	0.36
K3	0.28			ETL_{10}	0.40
B	0.28			MD	0.37
UPR	0.27			LPM_2	0.40
FT_1	0.26				
FT_2	0.25				
O	0.29				
ST	0.29				
A	0.26				
T	0.28				
0.93		0.84		0.88	

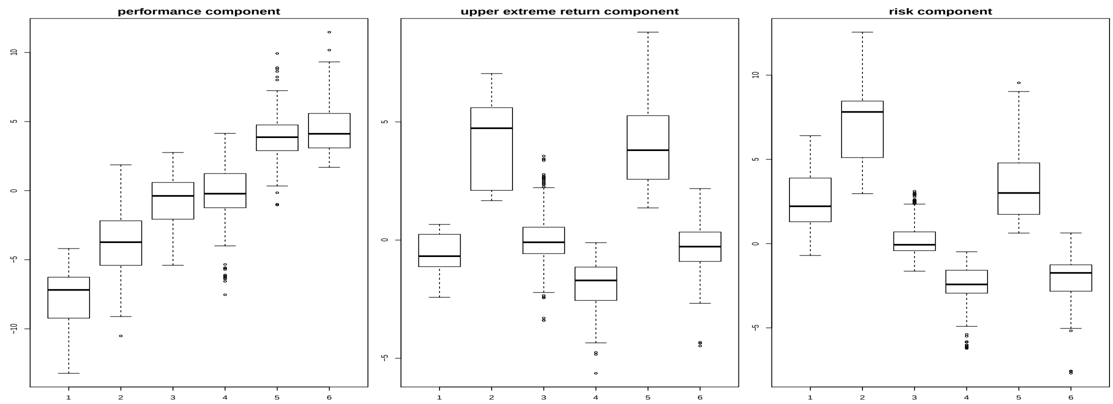

Figure 2: Boxplots of the performance, extreme return and risk components in the six detected groups. Label clusters are sorted with respect to the median value of the performance component at each group.

4 Discussion

In this work we have addressed the issue of classifying financial assets with respect to their return and risk level. To this aim several performance, return and risk measures have been estimated, based on the observation of historical data of changes in NAVs. The classification has been performed after suitably reducing the dimensionality of the data by clustering the variables into meaningful groups and then by combining the variables within each of these groups.

The adopted procedure does not claim to boast some novel theoretical contribution. Rather, it aims at showing how sound statistical methods may help to face the problem of partitioning objects when an amount of highly correlated proxy indicators are measured on them. While the procedure claims a general usability on any set of historical financial data, it has been applied on a set of US mutual funds. The application has resulted not only in producing a meaningful partition of funds into homogeneous clusters but also in defining a rating system that takes into account different levels of propensity to the risk.

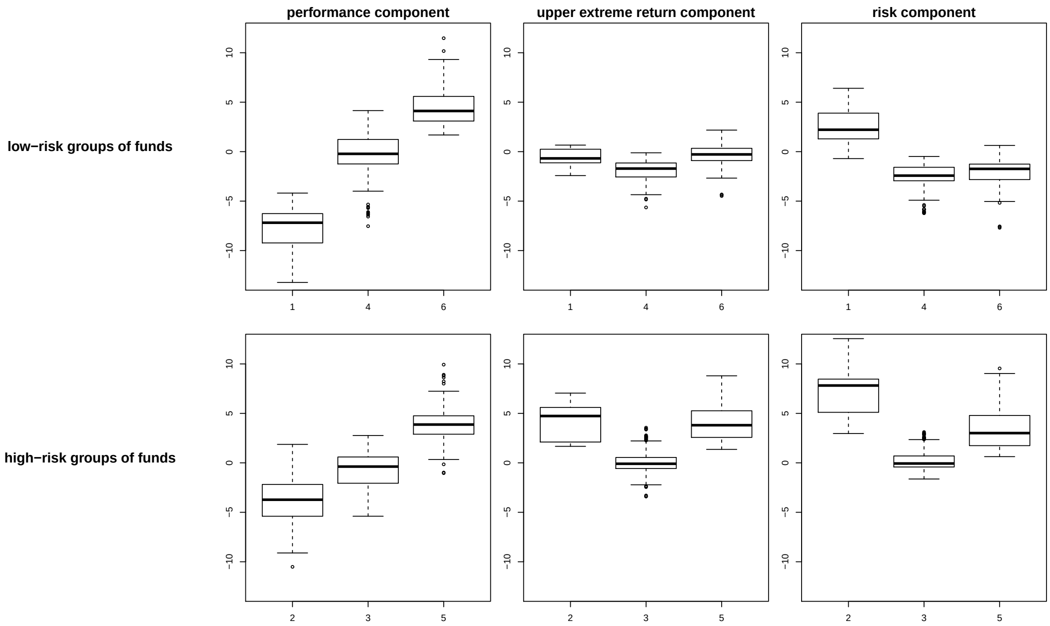

Figure 3: Distributions of funds in the low risk-low extreme return classes (top panel) and high risk-extreme return classes (bottom panel), with respect to the three composite indexes. On each panel, groups correspond to loser, intermediate and winner funds.

References

Cheridito, P., Stadje, M., 2009. Time-inconsistency of var and time-consistent alternatives. Finance Research Letters 6, 40–46.

Cogneau, P., Hübner, G., 2009. The (more than) 100 Ways to Measure Portfolio Performance. Part 1: Standardized Risk-Adjusted Measures. Journal of Performance Measurement 13, 56–71.

Corduas, M., Piccolo, D., 2008. Time series clustering and classification by the autoregressive metric. Computational Statistics and Data Analysis 52, 1860–1872.

Das, N., 2005. A new approach to hedge fund classification. Indian Journal of Economics and Business 4.

Eling, M., 2008. Does the measure matter in the mutual fund industry? Financial Analysts Journal 64, 54–66.

Elton, E., Gruber, M., Blake, C., 1996. Survivorship bias and mutual fund performance. Review of Financial Studies 9, 1097–1120.

Kim, M., Shukl, R., Tomas, M., 2000. Mutual fund objective misclassification. Journal of Economics and Business 52, 309–323.

Lajbcygier, P., Yahya, A., 2008. Soft Clustering for Funds Management Style Analysis: Out-of-Sample Predictability. Technical Report.

Lisi, F., Otranto, 2010. Clustering mutual funds by return and risk levels, in: Corazza, M., Pizzi C, E. (Eds.), Mathematical and Statistical Methods for Actuarial Sciences and Finance. Springer, pp. 183–191.

Lönnbark, C., Holmberg, U., Brännäs, K., 2011. Value at risk and expected shortfall for large portfolios. Finance Research Letters 8, 59 – 68.

Lytkin, N., Kulikowski, C., Muchnik, I., 2008. Variance-based criteria for clustering and their application to the analysis of management styles of mutual funds based on time series of daily returns. DIMACS Tech. Rep .

Morningstar, I., 2007. The New Morningstar Rating Methodology. Morningstar Research Report.

Pattarin, F., Paterlini, S., Minerva, T., 2004. Clustering financial time series: an application to mutual funds style analysis. Computational Statistics and Data Analysis 47, 353–372.

Sidana, G., Acharya, D., 2007. Classifying mutual funds in India: Some results from clustering. Indian Journal of Economics and Business 6, 71–79.