

Multivariate Time-Varying Volatility Modeling using Probabilistic Fuzzy Systems

Nalan Baştürk*, Rui Jorge Almeida[‡], Robert Golan[§] and Uzay Kaymak[‡]

*Department of Quantitative Economics, Maastricht University

P.O. Box 616, 6200 MD Maastricht, The Netherlands.

E-mail: n.basturk@maastrichtuniversity.nl

[‡]School of Industrial Engineering, Eindhoven University of Technology

P.O. Box 513, 5600 MB Eindhoven, The Netherlands.

E-mail: rjalmeida@tue.nl, u.kaymak@ieee.org

[§]DBmind Technologies Inc.

2829 Indian Creek Drive, Suite 1211, Miami, USA

E-mail: golan@dbmind.com

Abstract—Methods to accurately analyze financial risk have drawn considerable attention in financial institutions. One difficulty in financial risk analysis is the fact that banks and other financial institutions invest in several assets which show time-varying volatilities and hence time-varying financial risk. In addition, these assets are typically correlated and the correlation between different assets may change over time. Such changes in the multivariate volatility structure of the assets lead to substantial changes in the financial risk of a portfolio held by the financial institution. Therefore analyzing changes in the volatility of assets in a multivariate setting is essential to document changing risk properties of financial institutions. In this paper we propose a Probabilistic Fuzzy System (PFS) to model the unobserved time-varying correlation between a large set of financial returns. We define a parsimonious PFS where the current pairwise correlations between assets depend on two antecedent variables, namely the minimum and maximum past correlation in the market. We exemplify the proposed PFS model in six pairwise correlations for four industry portfolios in the US and show that the proposed method captures time-varying pairwise correlations while keeping the antecedent space parsimonious. Furthermore, we show that a portfolio investor that invests in these US industries calculates a lower risk for his/her portfolio when time-varying correlation estimates are not taken into account.

I. INTRODUCTION

Assessing measures of risk for financial returns has an important role in investment decisions, portfolio analysis and regulatory purposes [1], [2]. Accurate financial risk analysis has drawn considerable attention after the recent financial crisis [2], [3]. Measures of risk often need to be analyzed for more than one financial asset since most investment decisions are based on a selected portfolio of multiple assets, where the investor aims to diversify of risk [4].

For a univariate time series of financial returns, the associated risk changes considerable over time, which is denoted by time-varying conditional volatility in asset returns [5]. Methods which avoid strong distributional assumptions prove to be useful for estimating such time-varying volatility [6], [7], [8], [9], [10]. For multiple financial returns, rather than modeling assets in a univariate manner, several multivariate models have

been developed with specifications able to model and predict the temporal dependence in the second-order moments of many assets taking into account the correlated behavior of different assets [11], [12], [13], [14]. These models provide insight into stylized effects in assets returns such as time-varying correlations, portfolio diversification and contagion. For example, if the portfolio is based on two negatively correlated stocks, the portfolio is said to be ‘well-diversified’ with small risk [4]. Thus an accurate risk calculation, *e.g.* for a portfolio of financial assets, requires the accurate calculation of assets’ correlation at the given decision time.

Earlier research has shown that correlations between financial returns change over time, and these changing correlations are driven by common factors such as a financial crisis [15]. Different methodologies, with different distributional assumptions for correlation, have been proposed to calculate time-varying correlations for returns. Parametric models are proposed to estimate correlations between returns as well as other model parameters [16]. Alternative methodologies are based on moving window correlation estimates, where time-varying correlation at a given time is approximated by a proxy, namely the sample correlation at a selected time window. Moving window estimates have the advantage of avoiding strong distributional assumptions and are shown to perform well particularly in forecasting [17]. However, these estimates are also shown to be sensitive to the selection of a window size and there is a natural trade-off between capturing time variation in correlations and obtaining an accurate proxy for correlation at a given time. In [18], a PFS model is developed for modeling time-varying correlations of financial returns. It is also shown that the PFS model for conditional correlation captures time varying correlation and conditional volatility without an underlying restricted statistical model for the correlations for a single correlation between two stock market indices. PFS has also been shown to perform well for conditional volatility and risk estimation [19], [7], [9].

In [18] a PFS model is presented where the pairwise correlations between two specific assets depend on the pairwise

correlation of these assets in the last period. For multiple assets, it is intuitive that the correlation between two specific assets also depend on other pairwise correlations and the overall market properties in general. Such an extension of the PFS model requires a large antecedent space for the PFS model especially when the number of assets, hence the number of pairwise correlations between assets is high. In this paper, we extend the PFS model in [18] to model multivariate pairwise correlations between a large set of returns where correlations between assets can depend on the past correlation between all assets. The proposed model is refined to be parsimonious in the sense that past correlations are summarized using two antecedent variables, namely the minimum and maximum past correlation in the market. The proposed model is also different from the earlier PFS models since the antecedent and the consequent variables of the system are based on proxies, i.e. approximations of time-varying correlations instead of observed data.

We apply the proposed parsimonious PFS model to six pairwise correlations for four industry portfolios in the US and show that the proposed method captures time varying pairwise correlations while keeping the antecedent space parsimonious. We show that taking time-varying correlation into account leads to higher risk estimates for a portfolio investor. The proposed PFS model thus is useful to assess the portfolio risk accurately. Finally, we compare the proposed multivariate PFS with the univariate correlation model in [18]. The obtained univariate correlation results are in general in line with those from the proposed PFS. Despite this similarity, we show that the univariate PFS model results cannot be used to assess the portfolio risk due to the violation of the necessary conditions of a positive definite multivariate volatility matrix.

II. PROBABILISTIC FUZZY SYSTEMS

A probabilistic fuzzy system follows an idea similar to [20], [21] where the different concepts [22], [23], [24], [25] of fuzzy sets and probabilities are complementary [23]. In this work we consider that the probabilistic uncertainty relate to aleatoric variability, while fuzzy sets are used to represent gradualness, epistemic uncertainty or bipolarity [24], [26].

The PFS consists of a set of rules whose antecedents are fuzzy conditions and whose consequents are probability distributions. Assuming that the input space is a subset of $\mathbb{R}^n$ and that the rule consequents are defined on a finite domain $Y \subseteq \mathbb{R}$, a probabilistic fuzzy system consists of a system of rules R_q , $q = 1, \dots, Q$, of the type

$$R_q : \text{If } \mathbf{x} \text{ is } A_q \text{ then } f(y) \text{ is } f(y|A_q), \quad (1)$$

where $\mathbf{x} \in \mathbb{R}^n$ is an input vector, $A_q : X \rightarrow [0, 1]$ is a fuzzy set defined on X and $f(y|A_q)$ is the conditional pdf of the stochastic output variable y given the fuzzy event A_q . The interpretation is as follows: if fuzzy antecedent A_q is fully valid ($x \in \text{core}(A_q)$), then y is a sample value from the probability distribution with conditional pdf $f(y|A_q)$.

A PFS has been described with two possible and equivalent reasoning mechanisms, namely the fuzzy histogram approach

and the probabilistic fuzzy output approach [27]. We summarize these reasoning mechanisms, and refer to [27] for a detailed description. In both cases, we suppose that J fuzzy classes C_j form a fuzzy partition of the compact output space Y .

A. Fuzzy histogram model

In the fuzzy histogram model, the true unobserved pdf $f(y|A_q)$ of each rule in (1) is replaced by its fuzzy approximation $\hat{f}(y|A_q)$, yielding the rules for $q = 1, \dots, Q$

$$\hat{R}_q : \text{If } \mathbf{x} \text{ is } A_q \text{ then } f(y) \text{ is } \hat{f}(y|A_q), \quad (2)$$

where $\hat{f}(y|A_q)$ is a fuzzy histogram conform [28]

$$\hat{f}(y|A_q) = \sum_{j=1}^J \frac{\hat{\text{Pr}}(C_j|A_q) u_{C_j}(y)}{\int_{-\infty}^{\infty} u_{C_j}(y) dy}, \quad (3)$$

and fuzzy classes $C_1, \dots, C_J$ form a fuzzy partition of the compact output space Y . Fuzzy histograms have a high level of statistical efficiency due to the overlapping membership functions.

The output of a probabilistic fuzzy system of the form (2)-(3) has a proper pdf under the following necessary conditions for the probability parameters

$$\sum_{j=1}^J \hat{\text{Pr}}(C_j|A_q) = 1, \quad \hat{\text{Pr}}(C_j|A_q) \geq 0, \quad \forall j, q. \quad (4)$$

B. Probabilistic fuzzy output model

In the probabilistic fuzzy output approach, also referred to as Mamdani PFS [29], [19], [30], each rule output in (1) is defined with the following stochastic mapping between fuzzy antecedents and fuzzy consequents form

$$\begin{aligned} \text{Rule } \hat{R}_q : & \text{If } \mathbf{x} \text{ is } A_q \text{ then } y \text{ is } C_1 \text{ with } \hat{\text{Pr}}(C_1|A_q) \text{ and} \\ & \dots \\ & y \text{ is } C_J \text{ with } \hat{\text{Pr}}(C_J|A_q), \end{aligned} \quad (5)$$

where $\hat{\text{Pr}}(C_j|A_q)$ for $j = 1, \dots, J$ and $q = 1, \dots, Q$ are estimates of probability parameters satisfying condition (4), and PFS rules specify a probability distribution over a collection of fuzzy sets $C_1, \dots, C_J$, partitioning the output space Y .

C. Equivalence of reasoning mechanisms

Although the fuzzy rule bases (2) and (5) are different, it can be shown that the two corresponding probabilistic fuzzy systems implement the same crisp input-output mapping [27]. Let $\beta_q(\mathbf{x}) = u_{A_q}(\mathbf{x}) / \sum_{q'=1}^Q u_{A_{q'}}(\mathbf{x})$ be the normalised degree of fulfilment of rule R_q , where u_{A_q} is the degree of fulfilment of rule R_q . When $\mathbf{x}$ is n -dimensional, u_{A_q} is determined as a conjunction of the individual memberships in the antecedents computed by a suitable t-norm, i.e. ,

$$u_{A_q}(\mathbf{x}) = u_{A_{q1}}(x_1) \circ \dots \circ u_{A_{qn}}(x_n), \quad (6)$$

where $x_i, i = 1, \dots, n$ is the i -th component of $\mathbf{x}$ and $\circ$ denotes a t-norm. The output of the fuzzy rules (5) is a

conditional probability density function as presented in (2), if an additive reasoning scheme is used with multiplicative aggregation of the rule antecedents [27].

The conditional output probability distribution function $\hat{f}(y|\mathbf{x})$ given an input vector $\mathbf{x}$ is

$$\hat{f}(y|\mathbf{x}) = \sum_{j=1}^J \sum_{q=1}^Q \beta_q(\mathbf{x}) \hat{\Pr}(C_j|A_q) \frac{u_{C_j}(y)}{\int_{-\infty}^{\infty} u_{C_j}(y) dy}, \quad (7)$$

assuming that the output space is well-formed [27], *i.e.* that

$$\sum_{j=1}^J u_{C_j}(y) = 1, \quad \forall y \in Y. \quad (8)$$

Under these conditions, a crisp output using the expected value can be calculated based on the probability distribution function $\hat{f}(y|\mathbf{x})$

$$\hat{\mu}_{y|\mathbf{x}} = \hat{E}(y|\mathbf{x}) = \int_{-\infty}^{\infty} y \hat{f}(y|\mathbf{x}) dy \quad (9)$$

$$= \sum_{q=1}^Q \sum_{j=1}^J \beta_q(\mathbf{x}) \hat{\Pr}(C_j|A_q) z_{1,j}, \quad (10)$$

where $z_{1,j} = \int_{-\infty}^{\infty} y u_{C_j}(y) dy / \int_{-\infty}^{\infty} u_{C_j}(y) dy$ is the centroid of the j th output fuzzy set.

III. MULTIVARIATE VOLATILITY MODELING USING PFS

In this paper we extend the PFS model for a single correlation in [18] to a PFS model for multivariate volatility modeling. Specifically, we define a PFS model for the pairwise correlations for the multivariate volatility of n returns, denoted by $y_t = (y_{1,t}, y_{2,t}, \dots, y_{n,t})'$ where the returns are defined as

$$y_t = H_t^{1/2} z_t \quad (11)$$

for $t = 1, \dots, T$ time periods. In (11), the unobserved variable $z_t = (z_{1,t}, \dots, z_{n,t})'$ is such that $z_{i,t}$ for $i = 1, \dots, n$ are random variables with mean 0 and variance 1, H_t is a $n \times n$ positive definite matrix and $H_t^{1/2}$ denotes the Choleski decomposition of H_t . In addition, the $n \times n$ matrix $H_t = \text{Var}(y_t)$ represents the time-varying variance-covariances of y_t , which by construction are not observable.

Different models have been proposed to model the time-varying conditional variance-covariance matrix H_t in (11). Conventional methods for multivariate correlation analysis make use of the dependency of the current covariances H_t and past covariances $H_{t-1}, \dots, H_{t-p}$ leading to smooth changes in the variance-covariance structure over time [16]. Particularly for large n , ensuring a positive definite matrix H_t is cumbersome. This necessary condition may lead to additional parameter restrictions in models [16].

The following decomposition of the variance-covariance matrix is often used to identify variances and correlation coefficients [16]:

$$H_t = D_t R_t D_t \quad (12)$$

$$D_t = \begin{pmatrix} h_{1,1,t}^{1/2} & & 0 \\ & \ddots & \\ 0 & & h_{n,n,t}^{1/2} \end{pmatrix} \quad (13)$$

$$R_t = \begin{pmatrix} 1 & \rho_{1,1,t} & \dots & \rho_{1,n,t} \\ \rho_{1,1,t} & 1 & & \rho_{2,n,t} \\ \vdots & & \ddots & \vdots \\ \rho_{n,1,t} & \rho_{n,2,t} & \dots & 1 \end{pmatrix} \quad (14)$$

where the off-diagonal elements of R_t includes all pairwise correlations $\rho_{i,j,t}$ and $\rho_{i,j,t} = \rho_{j,i,t}$ by definition.

The covariance matrix decomposition in (12) ensures a positive definite matrix H_t in a univariate correlation case with $n = 2$ provided that the diagonal elements of D_t are positive and $\rho_{1,2,t} \in (-1, 1)$ in all time periods t . Ensuring that H_t is positive definite is more involved when $n > 2$. We illustrate this for $n = 3$ and the following numerical values

$$H_t = D_t R_t D_t \quad (15)$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1.0 & -0.4 & 0.6 \\ -0.4 & 1.0 & 0.9 \\ 0.6 & 0.9 & 1.0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (16)$$

where $\det(H_t) = \det(R_t) = -0.76$ and H_t is not a positive definite matrix. In this multivariate volatility setting, although the diagonal elements in D_t are positive and $\rho_{i,j,t} \in (-1, 1)$ for all t , the resulting multivariate correlation matrix may fail to satisfy the positive definiteness condition. The necessary condition of correlations instead is a positive definite correlation matrix R_t , with $\det(R_t) > 0$. In this paper, we do not explicitly deal with the conditions to satisfy a positive definite R_t . We instead check whether this condition is satisfied with the obtained parameter values.

The diagonal elements of the matrix D_t can be estimated using a given conditional volatility model for each series, as shown in [18]. For a crude estimation of the correlation coefficients $\rho_{i,j,t}$, moving-window (MW) correlation estimates $\hat{\rho}_t$ using window length m can be calculated using Pearson's linear correlation coefficient:

$$\hat{\mu}_{i,t} = \frac{\sum_{t'=t-m+1}^t y_{i,t'}}{m}, \quad \text{for } i = 1, \dots, n \quad (17)$$

$$\hat{\sigma}_{i,t}^2 = \frac{\sum_{t'=t-m+1}^t (y_{i,t'} - \hat{\mu}_{i,t})^2}{m-1}, \quad \text{for } i = 1, \dots, n \quad (18)$$

$$\hat{\rho}_{i,j,t}^{(m)} = \frac{\sum_{t'=t-m+1}^t (y_{i,t'} - \hat{\mu}_{i,t})(y_{j,t'} - \hat{\mu}_{j,t})}{(m-1)\hat{\sigma}_{i,t}\hat{\sigma}_{j,t}}. \quad (19)$$

The PFS model we propose makes use of the MW correlation estimates in (17)–(19) to obtain the antecedent and consequent variables. Specifically, we define the following PFS rules for $q = 1, \dots, Q$:

$$\hat{R}_q : \text{If } \min_{i,j} \hat{\rho}_{i,j,t-1}^{(m)} \text{ is } A_{q1} \text{ and } \max_{i,j} \hat{\rho}_{i,j,t-1}^{(m)} \text{ is } A_{q2} \text{ then}$$

$$f(\hat{\rho}_{i,j,t}^{(m)}) \text{ is } \hat{f}(\hat{\rho}_{i,j,t}^{(m)}|A_{q1}, A_{q2}), \quad (20)$$

where $i = 1, \dots, n-1$, $j = i+1, \dots, n$, denotes the

asset indices for pairwise correlations, t denotes the time period, the consequents are estimated using (19) with a pre-selected window size m , and $\hat{f}(\hat{\rho}_t^{(m)}|A_q)$ is a fuzzy histogram described as [28]

$$\hat{f}(\hat{\rho}_{i,j,t}^{(m)}|A_q) = \sum_{j=1}^J \frac{\hat{\Pr}(C_j|A_q) u_{C_j}(\hat{\rho}_{i,j,t}^{(m)})}{\int_{-\infty}^{\infty} u_{C_j}(\hat{\rho}_t^{(m)}) d\hat{\rho}_{i,j,t}^{(m)}}. \quad (21)$$

The equivalent probabilist fuzzy output model as defined in Section II-B is defined as:

$$\begin{aligned} \hat{R}_q : \text{If } \min_{i,j} \hat{\rho}_{i,j,t-1}^{(m)} \text{ is } A_{q1} \text{ and } \max_{i,j} \hat{\rho}_{i,j,t-1}^{(m)} \text{ is } A_{q2} \text{ then} \\ \underline{y} \text{ is } C_1 \text{ with } \hat{\Pr}(C_1|A_{q1}, A_{q2}) \text{ and} \\ \dots \\ \underline{y} \text{ is } C_J \text{ with } \hat{\Pr}(C_J|A_{q1}, A_{q2}), \end{aligned} \quad (22)$$

The antecedents variables

$$\hat{\rho}_{t-1}^{(m)} = \{\min_{i,j} \hat{\rho}_{i,j,t-1}^{(m)}, \max_{i,j} \hat{\rho}_{i,j,t-1}^{(m)}\}, \quad (23)$$

define only two antecedent variables from $n \times (n-1)/2$ pairwise correlations at each time period. This definition leads to three properties. First, similar to [18], the antecedent and the consequent variables are only approximations of the real variable of interest, correlations. Second, as opposed to the PFS model for correlation in [18], pairwise correlations between two specific assets does not specifically depend on the past correlation between these two assets, but instead depends on the extreme correlation values in the whole space of assets. Third, the antecedent space of PFS is parsimonious even when a large number of assets, hence a large number of pairwise correlations are modeled through PFS.

The parameters of the probabilistic fuzzy systems consist of the number of rules in the system, the parameters of the antecedent and consequent membership functions (*i.e.* number, type, location, etc.), and the probability parameters $\hat{\Pr}(C_j|A_q)$ of the stochastic mapping between the antecedent and the consequents. These parameters are estimated using a procedure similar to [9], where the optimization problem is divided in two parts. First we obtain the input membership parameters by using a fuzzy clustering heuristic, that uses the fuzzy c-means algorithm, set the output membership parameters as Gaussian, shouldered at the edges and finally optimize the probability parameters $\hat{\Pr}(C_j|A_q)$ using maximum likelihood estimation. This two step optimization problem follows the distinction between input and output present in the rule structure of (2).

In this work we determine the parameters of the antecedent membership functions by using a fuzzy clustering heuristic, that uses the fuzzy c-means algorithm [31] on the product space of the antecedent variables, to obtain a fuzzy partition matrix $U = [u_{qi}]$ for $p = 1, \dots, P$ samples, where $i = 1, \dots, n$. One dimensional fuzzy sets A_{qi} are obtained from the multidimensional fuzzy sets by projections onto the space of the input variables X . This is expressed by the point-wise projection operator $u_{A_{qi}}(x_{ip}) = \text{proj}_q(u_{qi})$. The point-wise

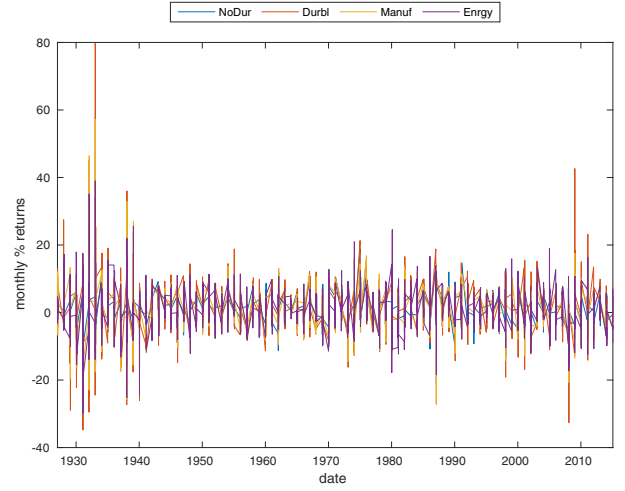


Fig. 1. Monthly percentage returns from equally weighted portfolios for four US industries.

defined fuzzy sets A_{qi} are then approximated by appropriate parametric functions. In this work we choose a combination of Gaussian membership functions of the form

$$f(x_{qi}; \sigma_{qi}^1, c_{qi}^1, \sigma_{qi}^2, c_{qi}^2) = e^{\left(\frac{-(x_{qi} - c_{qi}^1)^2}{2(\sigma_{qi}^1)^2} + \frac{-(x_{qi} - c_{qi}^2)^2}{2(\sigma_{qi}^2)^2} \right)}. \quad (24)$$

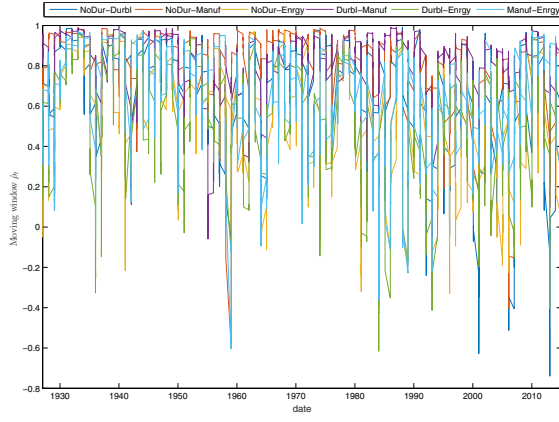
The output membership functions are triangular, as this is a convenient manner to satisfy (8) and fuzzy histograms built with these type of membership functions exhibit a high level of computational efficiency [32]. To satisfy (8) no matter how extreme the values may be and to ensure that the domain is always covered by the fuzzy partition, the membership functions are shouldered at the edges of the domain. The edges of the domain are set as the minimum and maximum of the output variables, which are the moving window correlation estimates. The output membership functions are uniformly distributed over the universe of discourse. Assuming that the membership functions in the rule antecedents have been defined, and the type of consequent membership functions and their distribution are known, the optimal probability parameters $\hat{\Pr}(C_j|A_q)$ and location of the output membership functions can be determined by using maximum likelihood estimation. Using (7) the log-likelihood for P samples of $\mathbf{x}_p$ can be written as [33]

$$L = \sum_{p=1}^P \ln \left(\sum_{j=1}^J \sum_{q=1}^Q \beta_q(\mathbf{x}_p) \hat{\Pr}(C_j|A_q) \frac{u_{C_j}(y_p)}{\int_{-\infty}^{\infty} u_{C_j}(y) dy} \right) \quad (25)$$

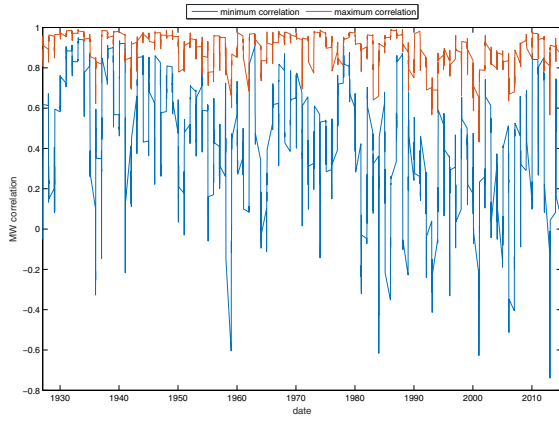
where u_{A_q} is calculated using (6). In (25) it is assumed that the samples in the data set are independent of one another. The probability parameters $\hat{\Pr}(C_j|A_q)$ must satisfy (4).

IV. APPLICATION TO US INDUSTRY PORTFOLIOS

We apply the proposed PFS model to monthly returns from four US industry portfolios between 1926 M7 and 2015 M7. The four selected industries are the non-durable goods, durable goods, manufacturing and energy industries. Each industry



(a) Six pairwise moving window correlation estimates, $\rho_{i,j,t}$.



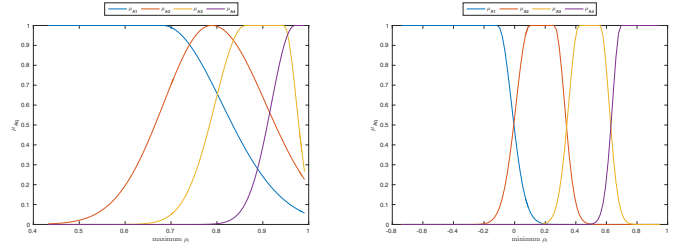
(b) Antecedent variables for PFS from pairwise moving window correlation estimates, $\min_{i,j} \rho_{i,j,t}, \max_{i,j} \rho_{i,j,t}$.

Fig. 2. Consequent and antecedent variables for PFS from moving window estimation.

portfolio is an equally weighted portfolio that invests equally on each stock within the industry. Portfolios constructed from these industry portfolios are analyzed extensively in the literature [34], [35].

Percentage returns from each industry portfolio are shown in Fig. 1. These four industry portfolios lead to 6 unobserved and possibly time-varying pairwise correlations, $\rho_{i,j,t}$, for $i = 1, \dots, n-1, j = i+1, \dots, n, n = 6$, at each time period $t = 1, \dots, T$. Monthly returns of each industry vary substantially over the considered time period, with clear extreme returns before 1950s and during the recent financial crisis.

In order to obtain the consequents, $\rho_{i,j,t}$, of PFS, we use moving window correlation estimates for each time period t using (19) with a window length of $M = 5$ for each pairwise correlation. The antecedent variables of PFS are defined as the the minimum and maximum values in the last period, denoted by $\min_{i,j}(\rho_{i,j,t-1}), \max_{i,j}(\rho_{i,j,t-1})$. The obtained correlation values are shown in Fig. 2a for each pairwise correlation.



(a) Maximum moving window correlation estimates. (b) Minimum moving window correlation estimates.

Fig. 3. Antecedent memberships for the PFS model.

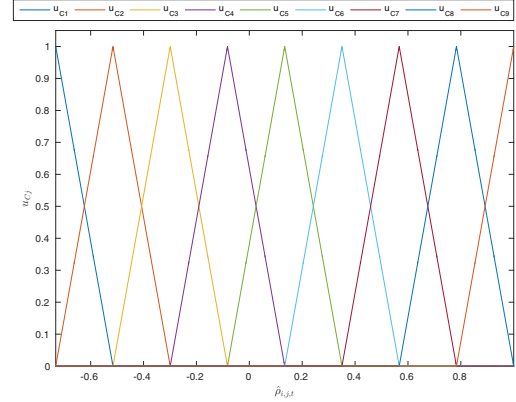
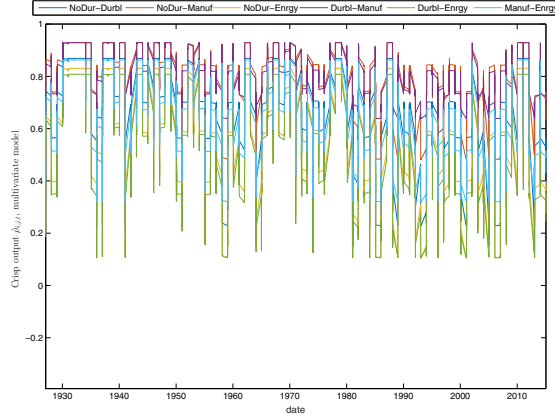


Fig. 4. Consequent memberships for the PFS model for the pairwise correlation between 'NoDur' and 'Durbl'.

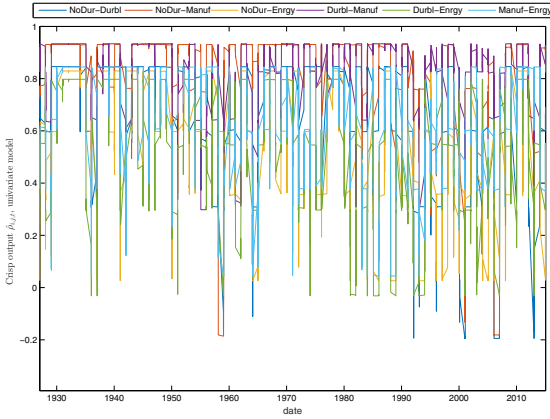
We note that the conventional method of obtaining correlation estimates from moving window averages leads to very volatile correlation estimates, and the obtained results are very sensitive to the choice of the window size, as discussed in [18]. In addition, the variables $\min_{i,j}(\rho_{i,j,t-1}), \max_{i,j}(\rho_{i,j,t-1})$ used as input for the PFS model are shown in Fig. 2b. Similarly to the pairwise correlations obtained from moving windows, these variables minimum and maximum correlation estimates at each time period, which are used as input for the PFS model, are also volatile as Fig. 2b shows. We note that this definition of minimum and maximum values of past correlation as PFS input variables provides a parsimonious antecedent space with two antecedent variables, and all consequent variables are explained by these two antecedent variables.

The proposed PFS model is estimated with 4 antecedent and 9 consequent membership functions. The obtained input membership functions are shown in Fig. 3 while the output membership functions are shown in Fig. 4. The fuzzy membership functions for maximum past correlation shown in Fig. 3a, are concentrated around high correlation values above 0.5. On the other hand the membership functions for minimum past correlation shown in Fig. 3b, are concentrated around positive correlation values above 0.

We next present the correlation estimates from the PFS model in Fig. 5a. The estimates for each pairwise correlation and time period are obtained using the crisp output of a



(a) Crisp output of the proposed PFS model.



(b) Crisp output of the univariate PFS model in [18].

Fig. 5. Six pairwise correlation estimates from the multivariate and univariate PFS models.

PFS model given by (10). Estimated pairwise correlations vary substantially over time. A particularly interesting high-correlation pattern is observed during the last financial crisis and a decrease in correlation values after 2011. We note that the obtained PFS correlations in Fig. 5a are smoother than those moving window estimates of PFS antecedent and consequent variables in Fig. 2a and Fig. 2b. This is a desired result as we do not expect correlation between different assets to change drastically at each time period, but rather vary gradually. The proposed PFS model captures time-varying pairwise correlations between assets without an underlying restricted statistical model for the correlations.

Other properties of pairwise correlations can be inferred from the PFS output, using the probabilistic fuzzy output approach. For example, different consequent variables and PFS probabilities are obtained for each pairwise correlation in the system, although all pairwise correlations have the same antecedent variables. Due to space limitations, we report the optimized PFS probabilities for two pairwise correlations in

TABLE I
OPTIMIZED PFS PROBABILITIES FOR SELECTED PAIRWISE CORRELATIONS
FOR Q INPUT AND J OUTPUT MEMBERSHIPS.

‘NoDur’ and ‘Energy’ Consequent									
Rule	1	2	3	4	5	6	7	8	9
1	0.00	0.18	0.29	0.17	0.20	0.11	0.05	0.00	0.00
2	0.00	0.00	0.04	0.09	0.38	0.24	0.22	0.03	0.00
3	0.00	0.00	0.00	0.02	0.05	0.34	0.41	0.18	0.00
4	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.70	0.24

‘Manuf’ and ‘Energy’ Consequent									
Rule	1	2	3	4	5	6	7	8	9
1	0.01	0.00	0.11	0.13	0.20	0.24	0.19	0.12	0.00
2	0.00	0.00	0.00	0.01	0.07	0.44	0.28	0.20	0.00
3	0.00	0.00	0.00	0.00	0.02	0.04	0.49	0.43	0.02
4	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.47	0.51

Table I. In general, high correlation values, corresponding to high J values, have a higher probability compared to low correlation values regardless of the antecedent values. Even when both antecedent variables have small values, corresponding to low Q values, the next period’s correlation is expected to be high. Still, there are clear heterogeneities between pairwise correlations. For example, the top panel of Table I shows that the pairwise correlation between manufacturing and energy sectors are in general higher than the pairwise correlation between the non-durables and energy sectors, independent of the antecedent values.

An interesting question is how the inferred portfolio risk using the PFS time-varying correlations compares with a simple equal-weighted portfolio for the four industry assets. The risk of an equal-weighted portfolio, PR_t for each time period t is given by:

$$\hat{PR}_t = w_t \hat{H}_t w_t' = w_t \hat{D}_t \hat{R}_t \hat{D}_t w_t' \quad (26)$$

where w_t is an $1 \times n$ vector with equal weights, $1/n$, $\hat{R}_t$ is the estimated correlation matrix from the crisp output for pairwise correlations, and $\hat{D}_t$ is the diagonal volatility matrix for each industry, as defined in the covariance decomposition in (12).

We compare the obtained portfolio risk to that of a zero-correlation portfolio, where the volatility matrix $\hat{D}_t$ in (26) is obtained from moving window estimates as in the PFS model, and the correlation matrix $\hat{R}_t$ is replaced by an identity matrix in (26). We denote this zero-correlation portfolio by $\hat{PR}_t^0$. Fig. 6 presents the portfolio risk calculated by the PFS output and the portfolio risk calculated under a zero correlation assumption. It can be seen that the zero correlation assumption leads to a substantially lower risk at each time period. *i.e.* an investor who disregards the time-varying correlation properties estimates a much lower risk for his/her portfolio. This result is also in line with the crisp output in Fig. 5a. Almost all pairwise correlations between industry assets are positive, hence a portfolio investing in these assets is not diversified. We note that the validation of the portfolio risk in comparison to the realized risk, for example using backtesting measures,

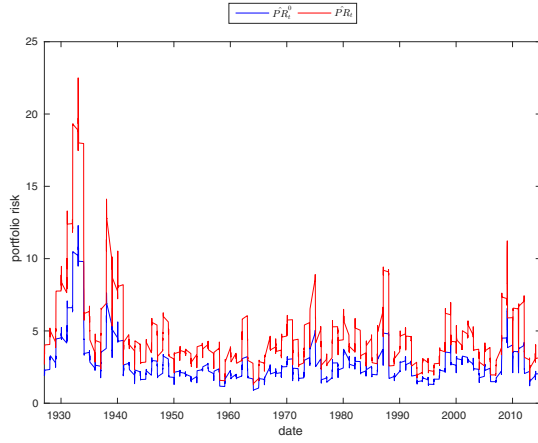


Fig. 6. Portfolio risk using PFS correlation ($\hat{P}R_t$) and under the assumption of zero correlation ($\hat{P}R_t^0$).

is not straightforward without further assumptions on the error terms z_t .

Finally, we compare the proposed multivariate PFS model with the univariate correlation model as presented in [18], where we apply the univariate correlation model separately to six pairwise correlations. In the univariate model, the antecedent variable for each pairwise correlation for assets (i, j) is defined as the past pairwise correlation between assets (i, j) and potential relations between pairwise correlations of different assets are not taken into account.

The crisp outputs of the univariate PFS model for six pairwise correlations are shown in Fig. 5b. Similar to the multivariate PFS results in Fig. 5a, estimated pairwise correlations are volatile with generally higher correlation values in 1930s and relatively lower correlation values starting from 1980s. Despite this similarity, univariate model correlations are much more volatile. In addition, pairwise correlations in the univariate model in Fig. 5b tend to go up and down together, although such a co-movement in pairwise correlations is not taken into account in the univariate PFS model. We therefore conclude that the proposed multivariate PFS model which explicitly deals with co-movements of pairwise correlations through common antecedent variables for all pairwise correlations is more appropriate for these data.

The univariate PFS model results also violate the positive definiteness condition in H_t and R_t in several time periods. We present the obtained determinants of R_t in (12) for both univariate and multivariate PFS models in Fig. 7. The proposed PFS model leads to positive definite covariance matrices at all time periods, while the univariate PFS violates this necessary condition for several time periods where $\det(R_t) \leq 0$. The univariate PFS model results cannot be used to assess the portfolio risk due to the violation of this condition.

V. CONCLUSIONS

In this paper we show that a parsimonious PFS can be used to model unobserved time-varying pairwise correlations

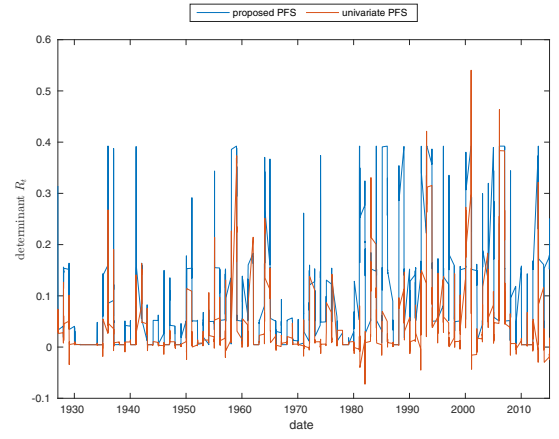


Fig. 7. $\det(R_t)$ from multivariate and univariate PFS models, obtained using the crisp output.

between a set of financial returns, where the obtained pairwise correlations lead to a multivariate volatility model with time-varying properties. The proposed method avoids strong distributional assumptions on the correlation process and uses the conventional approximation of time-varying correlation, namely sample correlations from moving windows, as antecedent and consequent variables. The antecedent space of the PFS model is reduced through defining the minimum and maximum past correlation as indicators of the overall correlation pattern in the market. The method is applied to monthly returns of four US industry portfolios where we show that the PFS application captures time-varying correlations for all pairwise correlations despite the small number of antecedent variables. It can be seen that the zero correlation assumption leads to a substantially lower risk at each time period. I.e. an investor who disregards the time-varying correlation properties estimates a much lower risk for his/her portfolio.

REFERENCES

- [1] Basel Committee on Banking Supervision, "Principles for effective risk data aggregation and risk reporting," January 2013, Bank for International Settlements report. [Online]. Available: <http://www.bis.org/publ/bcb239.pdf>
- [2] —, "Regulatory consistency assessment programme (RCAP)—analysis of risk-weighted assets for credit risk in the banking book," July 2013, Bank for International Settlements report. [Online]. Available: <http://www.bis.org/publ/bcb256.pdf>
- [3] P. Embrechts, G. Puccetti, L. Rüschendorf, R. Wang, and A. Beleraj, "An academic response to Basel 3.5," *Risks*, vol. 2, no. 1, pp. 25–48, 2014.
- [4] H. M. Markowitz, *Portfolio selection: efficient diversification of investments*. Yale University Press, 1968, vol. 16.
- [5] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, vol. 31, no. 3, pp. 307–327, 1986.
- [6] T. G. Andersen, T. Bollerslev, and F. X. Diebold, "Parametric and non-parametric volatility measurement," in *Handbook of Financial Econometrics, Vol 1: Tools and Techniques*, Y. Ait-Sahalia and L. P. Hansen, Eds. Elsevier, 2010, pp. 67–138.
- [7] R. J. Almeida, N. Baştürk, U. Kaymak, and V. Milea, "A multi-covariate semi-parametric conditional volatility model using probabilistic fuzzy systems," in *Computational Intelligence for Financial Engineering Economics (CIFER)*, 2012 IEEE Conference on, march 2012, pp. 1–8.

- [8] R. J. Almeida, N. Baştürk, U. Kaymak, and J. M. da Costa Sousa, "Conditional density estimation using fuzzy GARCH models," in *Synergies of Soft Computing and Statistics for Intelligent Data Analysis*, ser. Advances in Intelligent Systems and Computing, R. Kruse, M. R. Berthold, C. Moewes, M. A. Gil, P. Grzegorzewski, and O. Hryniewicz, Eds. Springer Berlin Heidelberg, 2013, vol. 190, pp. 173–181. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-33042-1_19
- [9] R. J. Almeida, N. Baştürk, and Kaymak, "Probabilistic fuzzy systems for seasonality analysis and multiple horizon forecasts," in *Computational Intelligence for Financial Engineering Economics (CIFEr), 2014 IEEE Conference on*, March 2014, pp. 497–504.
- [10] R. J. Almeida, N. Basturk, U. Kaymak, and J. M. da Costa Sousa, "Estimation of flexible fuzzy GARCH models for conditional density estimation," *Information Sciences*, vol. 267, pp. 252–266, 2014.
- [11] C. R. Harvey, "Time-varying conditional covariances in tests of asset pricing models," *Journal of Financial Economics*, vol. 24, no. 2, pp. 289–317, 1989.
- [12] Y. K. Tse and A. K. C. Tsui, "A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations," *Journal of Business & Economic Statistics*, vol. 20, no. 3, pp. 351–362, 2002.
- [13] R. Engle, "Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models," *Journal of Business & Economic Statistics*, vol. 20, no. 3, pp. 339–350, 2002.
- [14] C. Krishnan, R. Petkova, and P. Ritchken, "Correlation risk," *Journal of Empirical Finance*, vol. 16, no. 3, pp. 353–367, 2009.
- [15] Y. Hamao, R. W. Masulis, and V. Ng, "Correlations in price changes and volatility across international stock markets," *Review of Financial Studies*, vol. 3, no. 2, pp. 281–307, 1990.
- [16] L. Bauwens, S. Laurent, and J. V. Rombouts, "Multivariate GARCH models: a survey," *Journal of applied econometrics*, vol. 21, no. 1, pp. 79–109, 2006.
- [17] R. Becker, A. E. Clements, M. B. Doolan, and A. S. Hurn, "Selecting volatility forecasting models for portfolio allocation purposes," *International Journal of Forecasting*, vol. 31, no. 3, pp. 849–861, 2015.
- [18] N. Baştürk and R. J. Almeida, "Time-varying correlation estimation using probabilistic fuzzy systems," 2016, forthcoming in proceedings of IPMU-2016 International Conference, Communications in Computer and Information Science (CCIS) with Springer.
- [19] D. Xu and U. Kaymak, "Value-at-risk estimation by using probabilistic fuzzy systems," in *Proceedings of the 2008 World Congress on Computational Intelligence*, Hong-Kong, Jun. 2008, pp. 2109–2116.
- [20] Z. Liu and H.-X. Li, "A probabilistic fuzzy logic system for modeling and control," *IEEE Transactions on Fuzzy Systems*, vol. 13, no. 6, pp. 848–859, 2005.
- [21] G. Zhang and H.-X. Li, "An efficient configuration for probabilistic fuzzy logic system," *Fuzzy Systems, IEEE Transactions on*, vol. 20, no. 5, pp. 898–909, oct. 2012.
- [22] L. A. Zadeh, "Probability measures of fuzzy events," *J. Math. Anal. Appl.*, vol. 23, pp. 421–427, 1968.
- [23] —, "Discussion: Probability theory and fuzzy logic are complementary rather than competitive," *Technometrics*, vol. 37, no. 3, pp. 271–276, 1995.
- [24] D. Dubois and H. Prade, "The three semantics of fuzzy sets," *Fuzzy Sets and Systems*, vol. 90, no. 2, pp. 141 – 150, 1997, fuzzy Sets: Where Do We Stand? Where Do We Go?
- [25] C. Bertoluzza, M. A. Gil, and D. A. Ralescu, Eds., *Statistical Modeling, Analysis and Management of Fuzzy Data*, ser. Studies in Fuzziness and Soft Computing. Heidelberg: Physica Verlag, 2002.
- [26] D. Dubois and H. Prade, "Gradualness, uncertainty and bipolarity: Making sense of fuzzy sets," *Fuzzy Sets and Systems*, vol. 192, no. 0, pp. 3 – 24, 2012, fuzzy Set Theory Where Do We Stand and Where Do We Go?
- [27] J. van den Berg, U. Kaymak, and R. J. Almeida, "Conditional density estimation using probabilistic fuzzy systems," *Fuzzy Systems, IEEE Transactions on*, vol. 21, no. 5, pp. 869–882, 2013.
- [28] J. van den Berg, W.-M. van den Bergh, and U. Kaymak, "Probabilistic and statistical fuzzy set foundations of competitive exception learning," in *Proceedings of the Tenth IEEE International Conference on Fuzzy Systems*, vol. 2, Melbourne, Australia, Dec. 2001, pp. 1035–1038.
- [29] U. Kaymak, W.-M. van den Bergh, and J. van den Berg, "A fuzzy additive reasoning scheme for probabilistic Mamdani fuzzy systems," in *Proceedings of the 2003 IEEE International Conference on Fuzzy Systems*, vol. 1, St. Louis, USA, May 2003, pp. 331–336.
- [30] R. J. Almeida and U. Kaymak, "Probabilistic fuzzy systems in Value-at-Risk estimation," *Intelligent Systems in Accounting, Finance & Management*, vol. 16, no. 1–2, pp. 49–70, 2009.
- [31] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York: Plenum Press, 1981.
- [32] L. Waltman, U. Kaymak, and J. van den Berg, "Fuzzy histograms: A statistical analysis," in *EUSFLAT-LFA 2005 Joint 4th EUSFLAT 11th LFA Conference*, 2005, pp. 605–610.
- [33] —, "Maximum likelihood parameter estimation in probabilistic fuzzy classifiers," in *Fuzzy Systems, 2005. FUZZ '05. The 14th IEEE International Conference on*, May 2005, pp. 1098–1103.
- [34] T. J. Moskowitz and M. Grinblatt, "Do Industries Explain Momentum?" *The Journal of Finance*, vol. 54, pp. 1249–1290, 1999.
- [35] H. L. Hühn, "Industry momentum: The role of time-varying factor exposures and market conditions," October 2015. [Online]. Available: <http://ssrn.com/abstract=2650378>