Tail Point Density Estimation Using Probabilistic Fuzzy Systems

R. J. Almeida U. Kaymak

Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam
P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
Email: {rjalmeida,kaymak}@ese.eur.nl

Abstract— *Value at Risk (VaR) is a popular measure for quantifying the market risk that a financial institution faces into a single number. Due to the complexity of financial markets, the risks associated with a portfolio may vary over time. For accurate VaR estimation, it is necessary to have flexible methods that adapt to the underlying data distribution. In this paper, we consider VaR estimation by using probabilistic fuzzy systems (PFS). Contrary to previous publications, our focus is on the modeling of the tail points of the distribution of returns. We study two approaches to designing probabilistic fuzzy VaR models that take into account the extreme values of the data and compare their performance with the performance of a GARCH model. It is found that the VaR estimation process is simplified and improved by our proposed method.*

Keywords— Density estimation, extreme values, fuzzy histograms, probabilistic fuzzy systems, value-at-risk.

1 Introduction

Due to the volatile nature of the financial markets risk management is an important activity for financial institutions that operate in these markets. As a result of risk management, activities are undertaken to reduce the possibility of failure to an acceptable range. These activities may include portfolio adjustment, hedging or insurance [1, 2]. Nowadays, the financial sector operates under strict guidelines, which have been imposed through international agreements, partly due to various financial failures that have happened in 1990's. For example, due to the Basel Agreement, the financial institutes must have well documented procedures to manage the market risk, the credit risk and the operational risk that they are exposed to [3].

Managing risk is strongly dependent on the information available. When the amount of information grows beyond a specific level, there is a need for a concise representation of the risk a company or institution is facing. Due to the complex nature of financial markets in which many parties exchange information and interact through trading, the overall risk for a company is influenced by many internal and external factors. It is customary for management to classify different types of risk and develop models for dealing with each type of risk in order to keep the risk management problem tractable. One of the different types of risk that a financial institution has to deal with is the market risk, which is the exposure to the uncertain market value of a portfolio [4]. Value-at-risk (VaR) is a way to quantify the market risk. It is a single number for the senior management to express and summarise the total market risk of a portfolio with financial assets. Value at Risk measures the worst expected loss over a given horizon under normal market conditions at a given confidence level. Due to regulations, large banks must nowadays base their market risk capital requirements on the VaR estimate [3]. This drives the continued research into newer and better VaR models.

Different types of approaches have been proposed for VaR estimation, such as simulation and parametric approaches. In the parametric approaches, the risk is quantified in terms of volatility, which is expressed as the standard deviation σ of the portfolio. The simplest models of volatility assume that it does not vary over time, while more advanced models acknowledge that volatility varies dynamically over time. The dynamic aspect of volatility could be modelled in various ways. A popular model where volatility changes dynamically in time is the GARCH (Generalised Auto Regressive Heteroscedasticity) model [5]. For the GARCH $(1, 1)$ model, which is used quite often in practice, the variance is estimated using a first-order autoregressive model of the squared returns.

The disadvantage of the parametric approach is that, due to the complexity of financial markets, the data usually do not follow the parametric distributions that are assumed to underly the data generating process. For example, the returns are typically non-Gaussian, they have fat tails and volatility clustering is often observed in financial markets see *e.g.* [6, 7]. Therefore, flexible modelling approaches such as non-parametric modelling or semi-parametric modelling are needed in which the models can adapt themselves into the underlying actual data distribution. In this context, neural network models for VaR estimation have been studied by various researchers [8, 9, 10] as well as fuzzy set models [11, 12]. Probabilistic fuzzy systems [13] were also used to model VaR by estimating the whole density function, by following the distribution of the data and distributing more membership functions around the origin.

In this paper we focus on tail point density estimation using probabilistic fuzzy systems, to correctly estimate VaR. For this purpose we consider two approaches for estimating the conditional parameters of the PFS model, by giving more relevance to the extreme values, and comparing their performance in obtaining value-at-risk models. A Mamdami-type probabilistic fuzzy system [14] is used for this purpose. The location of the antecedent membership functions is determined by using fuzzy clustering and maximum likelihood parameter estimation is used for determining the probability parameters of the PFS. The validity of the VaR models is evaluated by using a statistical back-testing method based on failure rates.

The outline of the paper is as follows. In Section 2 we discuss the basics of probabilistic fuzzy systems and the concept of fuzzy histograms. In Section 3 we give a brief introduction to VaR modelling and VaR models. In Section 4 we present two methods for estimating the PFS parameters that give more relevance to the tail points of the distribution, where the VaR

ISBN: 978-989-95079-6-8

is estimated. The experimental setup for the empirical study using six different assets are given in Section 5, while the results are reported in Section 6. Finally, conclusions and future work are given in Section 7.

2 Probabilistic Fuzzy Systems

Probabilistic fuzzy systems (PFS) are based on the concept of the probability of a fuzzy event, as defined by Zadeh [15]. It is assumed that the input space is a subset of $\mathbb{R}^n$ and that the rule consequents are defined on a finite domain $Y \in \mathbb{R}$. The PFS consists of a set of rules whose antecedents are fuzzy conditions and whose consequents are probability distributions. In this study, we consider Mamdani PFS in which the rules have the following form [14].

$$\begin{aligned} \text{Rule } R_q\text{: If } \mathbf{x} \text{ is } A_q \text{ then} \\ & y \text{ is } C_{q1} \text{ with } \Pr(C_{q1}|A_q) \text{ and} \\ & y \text{ is } C_{q2} \text{ with } \Pr(C_{q2}|A_q) \text{ and} \ldots \text{ and} \\ & y \text{ is } C_{qN} \text{ with } \Pr(C_{qN}|A_q). \end{aligned} \tag{1}$$

Hence, a Mamdani PFS is a generalisation of a Mamdani fuzzy system in which the deterministic fuzzy rules are replaced with probabilistic fuzzy rules. These rules specify a probability distribution over a collection of fuzzy sets that partition the output domain. The rules of a PFS also express linguistic information and they can be used to explain the model behavior by a set of linguistic rules. This way, the system deals both with linguistic uncertainty as well as probabilistic uncertainty. The interpretation of the probabilistic fuzzy rules is as follows. Given the occurrence of a (multidimensional) antecedent fuzzy event A_q, which is a conjunction of the fuzzy conditions defined on input variables, each of the consequent fuzzy events C_j is likely to occur. The selection of which of the consequent fuzzy events occurs, is done proportionally to the conditional probabilities $\Pr(C_j|A_q)$, $(j = 1, 2, \ldots, N)$. This applies for all the rules R_q, $q = 1, 2, \ldots, Q$. Note that two conditional probabilities $\Pr(C_j|A_q)$ and $\Pr(C_j|A_{q'})$ will be different, in general.

Let

$$\beta_q(\mathbf{x}) = \frac{\mu_{A_q}(\mathbf{x})}{\sum_{q'=1}^{Q} \mu_{A_{q'}}(\mathbf{x})} \tag{2}$$

be the normalised degree of fulfillment of rule R_q, where μ_{A_q} is the degree of fulfillment of rule R_q. When $\mathbf{x}$ is n-dimensional, μ_{A_q} is determined as a conjunction of the individual memberships in the antecedents computed by a suitable t-norm, *i.e.* ,

$$\mu_{A_q}(\mathbf{x}) = \mu_{A_{q1}}(x_1) \circ \cdots \circ \mu_{A_{qn}}(x_n), \tag{3}$$

where x_n is the n-th components of $\mathbf{x}$ and $\circ$ denotes a t-norm. Then, it can be shown that the output of the above Mamdani PFS is a conditional probability density function if an additive reasoning scheme is used with multiplicative aggregation of the rule antecedents [16]. The conditional probability of the output given an input vector $\mathbf{x}$ can be computed as

$$f(y|\mathbf{x}) = \sum_{j=1}^{N} \frac{\sum_{q=1}^{Q} \beta_q(\mathbf{x}) \Pr(C_j|A_q) \mu_{C_j}(y)}{\int_{-\infty}^{\infty} \mu_{C_j}(y) dy}, \tag{4}$$

assuming that the output space is well-formed, *i.e.* the output membership values satisfy

$$\sum_{j=1}^{N} \mu_{C_j}(y) = 1, \qquad \forall y \in Y. \tag{5}$$

It is also possible to compute the crisp output of the probabilistic fuzzy system by taking the conditional expectation of the output according to

$$E(y|\mathbf{x}) = \int_{-\infty}^{\infty} y f(y|\mathbf{x}) dy. \tag{6}$$

However, we are not interested in the expected output of the system in this paper. We are primarily interested in the output of the PFS as a fuzzy histogram, by using (4).

Assuming Y is fuzzily partitioned in a set of N fuzzy classes C_j described by membership functions $\mu_{C_j}(y)$, then the (fuzzy) column $f_j(y)$ for fuzzy class C_j can be estimated according to

$$f_j(y) = \frac{\Pr(C_j)\mu_{C_j}(y)}{\int_{-\infty}^{\infty} \mu_{C_j}(y) dy}, \tag{7}$$

The numerator in (7) describes a probability weighted with membership function $\mu_{C_j}(y)$. The denominator of (7) is a scaling factor representing the fuzzified size of class C_j (which in the one-dimensional continuous case, equals the *fuzzy length* of the interval C_q). The complete pdf $f(y)$ is again approximated by a summation of the functions $f_j(y)$:

$$f(y) \approx f_{app}(y) = \sum_{j=1}^{N} f_j(y) = \sum_{j=1}^{N} \frac{\Pr(C_j)\mu_{C_j}(x)}{\int_{-\infty}^{\infty} \mu_{C_j}(y) dy}. \tag{8}$$

Due to the overlap of the fuzzy sets, fuzzy histograms approximate probability distributions better, in practice. In Fig. 1 a representation of this phenomenon is shown, where a normal probability density function is approximated using both a crisp and a fuzzy histogram. In both cases, seven classes have been used.

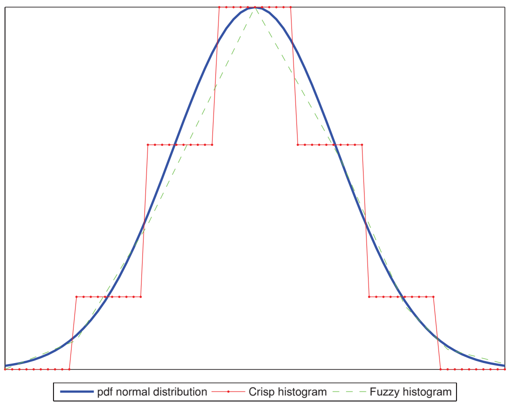


Figure 1: Fuzzy Histogram

3 VaR Estimation

In this section, we discuss value-at-risk estimation as an application example for probabilistic fuzzy systems. Value-at-risk is a single number for the senior management to express and

ISBN: 978-989-95079-6-8

summarise the market risk of a portfolio of financial assets. The VaR value of a portfolio is always calculated over a time horizon h at a significance level c. It indicates the maximum loss that a portfolio of assets will suffer over a horizon of h (days) with a confidence of c. An overview of the mainstream value at risk estimation methods can be found in [17]. Several methods are also discussed in [18]. Various building blocks of VaR measurement, methods for model validation as well as the differences between the parametric and nonparametric estimation approaches are discussed in [3].

3.1 *Value at risk*

Assume that a portfolio has value W_t at time t. Let r denote the one period percentage return of the portfolio. If $f(r)$ is the probability density function of the returns, define r_v such that

$$1 - c = \int_{-\infty}^{r_v} f(r)dr. \tag{9}$$

The value at risk V_t of the portfolio at time t is then defined as $V_t = -r_v W_t$. Assuming that the returns are distributed normally, the key step in the value at risk estimation can be formulated as determining the standard deviation σ of the returns distribution. This is also called *volatility estimation*.

Probabilistic fuzzy systems have been successfully applied in estimating the whole density distribution. In [13], good results were obtained by scaling the consequent membership functions, which were distributed with more membership functions are around the origin. The distribution of consequent functions alters the approximation of the density functions. Since VaR is calculated at the negative endpoint of the distribution function, it is logical to try to model the tail points more precisely. In Section 4 two approaches are presented that focus on the modeling of the tail points.

3.2 *Model validation*

Model validation is the process of checking whether a model performs adequately, and can be done in various ways. In this paper, we consider *exception based back testing*. Kupiec has developed a statistical test for assessing the validity of a VaR model [19]. Kupiec confidence regions are defined by the tail point of the log-likelihood ratio LR_{uc}

$$\begin{aligned} LR_{uc} &= -2\ln\left[c^{T-N}(1-c)^N\right] \\ &\quad +2\ln\left\{\left[1-\left(\frac{N}{T}\right)\right]^{T-N}\left(\frac{N}{T}\right)^N\right\}. \end{aligned} \tag{10}$$

In (10), N is the number of exceptions and T is the total number of observations. This ratio is shown to be asymptotically χ^2-distributed, with 1 degree of freedom, under the null hypothesis that the VaR model is valid [19]. Note that the Kupiec test statistic is two sided. Hence, the model is rejected both when there are too few exceptions, (the model is too conservative), as well as when there are too many exceptions, (the model underestimates the volatility). In this paper, we apply the Kupiec test with 95% confidence to assess the validity of the VaR models.

4 Parameter Learning for Tail Points

The parameters of the probabilistic fuzzy systems consist of the number of rules in the system, the parameters of the membership functions (*i.e.* number, type, location, etc.) and the probability parameters $\Pr(C_j|A_q)$ of the stochastic mapping between the antecedent and the consequents. The identification of all the parameters of the PFS simultaneously can be very time consuming and it suffers from the problem of multiple local minima. Thus, we use process knowledge to establish values of a subset of parameters. The other parameters are then optimised given the values of this subset of parameters.

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 [20] on the product space of the antecedent variables, and the spreads of the membership functions are derived from the distribution of the data, using a fuzzy covariance matrix, as proposed in [13]. Since the output membership functions must satisfy (5), it is convenient to use triangular membership functions for the output partition. These membership functions are combined with shouldered membership functions at the edges of the domain, as depicted in Fig. 2(a), to ensure that the domain is always covered by the fuzzy partition, no matter how extreme the values may be. The distribution of the membership functions can be uniform over the universe of discourse, or it can be varying. Two possible partitions for the output space are shown in Fig. 2. More membership functions placed around the edges of the universe of discourse, allows to better capture the variability in this region and to better model the tail points of the distribution. As one moves towards the edges of the universe of discourse, the membership functions have smaller supports and the separation between them decreases. Note that the triangular membership functions in this partition are not symmetric.

Assuming that the membership functions in the rule antecedents and the rule consequents have been defined, the optimal probability parameters $\Pr(C_j|A_q)$ can now be determined by using maximum likelihood parameter estimation, in which the log-likelihood function

$$J = \sum_{k=1}^{K} \ln\left(\Pr(y_k|\mathbf{x}_k)\right) \tag{11}$$

is maximised where K is the number of samples in the data set [21]. In (11), it is assumed that the samples in the data set are independent of one another. A suitable initialisation for iterative optimisation for maximum likelihood estimation is given by direct estimation from the data by using

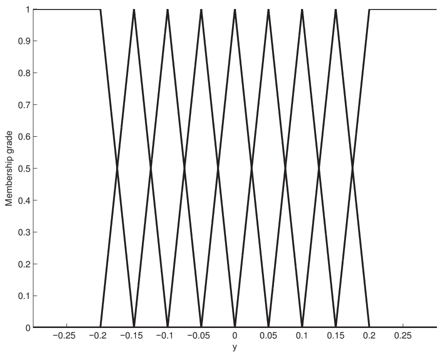
$$\Pr(C_j|A_q) = \frac{\sum_{k=1}^{K} \mu_{C_j}(y_k)\mu_{A_q}(\mathbf{x}_k)}{\sum_{k=1}^{K} \mu_{A_q}(\mathbf{x}_k)}. \tag{12}$$

During the parameter estimation of the optimal probability parameters $\Pr(C_j|A_q)$, more relevance can be given to the extremes of the universe of discourse. This can be accomplished by using an appropriate weight function $w(y_k)$ on the maximum likelihood parameter estimation, dependent on the distribution of the variable, that will give more weight to the tail points, in the form:

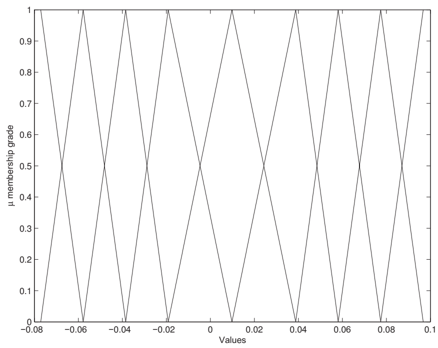
$$J = \sum_{k=1}^{K} \ln\left(w(y_k)\Pr(y_k|\mathbf{x}_k)\right). \tag{13}$$

The most simple kind of functions that can be used is an

ISBN: 978-989-95079-6-8



(a) Equally spaced.



(b) More sets towards the edges.

Figure 2: Two possible partitions for the output space.

affine function, such as

$$w(y_k) = \begin{cases} (\min_j(y_j)/2)^{-1}y_k & \text{if } y_k \leq 0, \\ (\max_j(y_j))^{-1}y_k & \text{if } y_k > 0. \end{cases} \quad (14)$$

Note that in this function more weight is given to the negative side, to better model the negative tail points and possibly obtain better VaR estimation.

5 Experimental Setup

We have studied the performance of the histogram-based probabilistic fuzzy systems to estimate VaR for six different stocks: KPN, ABN AMRO, JiaLing, BaoShan, COSCO and Merchant Bank. The performance of PFS models has been compared with the performance of the GARCH models.

In this paper, we consider value at risk models for one period ahead. In other words, the horizon over which the value at risk is computed is one day. Extensive models could be identified for the multiple day case (i.e. h-day horizon), but usually one suffices by using the simple $\sqrt{(h)}$ scaling, where the h-day value at risk is taken as $\sqrt{(h)}$ times the one-day value at risk.

The probabilistic fuzzy models that we consider use the returns r_t at period t to predict the distribution of the returns at period $t+1$. In our models, we have used nine antecedent membership functions and nine consequent membership functions. Hence, the fuzzy system had nine rules. The input space was partitioned using the FCM algorithm with nine clusters. In such a system, there are 81 probability parameters $\Pr(C_j|A_q)$ (nine for each rule). Since FCM has the tendency to place more clusters in regions covered with more data, there are more antecedent membership functions in the centre, where more samples are available.

The output space was partitioned with triangular membership functions. These triangular membership functions are combined with shouldered membership functions at the edges of the domain, to ensure that the domain is always covered by the fuzzy partitions, no matter how extreme the returns may be on a particular day.

Given the fuzzy membership functions whose parameters are determined as above, the conditional probability parameters for the PFS are determined by using maximum likelihood estimation. We used uniformly distributed triangular membership functions over the universe of discourse, combined with (13), and we name this model PFS_{UD}. For the case where we used the maximum likelihood parameter estimation as stated in (11), we used varying size of triangular membership functions over the universe of discourse, and we name this model PFS_V. Both approaches could be applied together. Nonetheless, in this work, we are interested in studying different ways of modelling the tail points of the probability density functions obtained with PFS, and thus we study them separately. Given the conditional probability distribution of one period returns, the value at risk of the portfolio is obtained by using (9).

The steps necessary for computing the one-period value-at-risk of a portfolio can now be summarised as follows for PFS_1 models.

1. Collect the price series regarding the portfolio and compute the one-period returns. Create training and validation data sets.

2. Determine antecedent membership functions by applying fuzzy c-means clustering to compute the locations of the membership functions and use cluster covariance to obtain the spreads of the membership functions.

3. Select the number of consequent membership functions and form a triangular partition, equally spaced in the case of PFS_{UD} and with more sets towards the edges on PFS_V.

4. Given the definitions of the antecedent and the consequent membership functions, determine the optimal probability parameters of the PFS by maximising (13) for the PFS_{UD} and (11) for the PFS_V.

5. Using the test set, compute the estimated conditional probability distribution function for the one-period returns for each observation in the test set.

6. Given the conditional probability distribution functions, compute the VaR by using (9).

7. Validate the model by using exception based back-testing as explained in Section 3.2.

6 Results

This section reports the application results for the proposed approaches to the stocks in study, for the validation data sets. Table 1 shows the obtained results of the exception-based back testing for the best GARCH and probabilistic fuzzy models.

ISBN: 978-989-95079-6-8

This table shows the number of exceptions that have occurred in the validation data for different levels of the confidence parameter c. The bold face numbers indicate that the model is not rejected according to the test statistic. The non-rejection region for the Kupiec test statistic is also shown.

Table 1: Failure rates for back testing

Asset	c	PFS_{UD}	PFS_V	GARCH	Non-Rejection Region
	95%	**26**	**30**	**19**	16 < N < 36
ABN	97.5%	**15**	**13**	**13**	6 < N < 20
	99%	**5**	**6**	**9**	1 < N < 10
	95%	**29**	**22**	11	16 < N < 36
KPN	97.5%	**14**	**15**	**8**	6 < N < 20
	99%	**10**	**9**	**4**	1 < N < 10
	95%	**30**	**35**	**22**	16 < N < 36
JiaLing	97.5%	**11**	**14**	**14**	6 < N < 20
	99%	**6**	**6**	6	1 < N < 10
	95%	**24**	**20**	12	16 < N < 36
BaoShan	97.5%	**14**	**13**	**8**	6 < N < 20
	99%	**9**	**4**	**6**	1 < N < 10
	95%	**26**	**25**	14	16 < N < 36
COSCO	97.5%	**9**	**9**	**11**	6 < N < 20
	99%	**5**	**6**	**5**	1 < N < 10
	95%	**27**	**27**	10	16 < N < 36
Merchant	97.5%	**15**	**14**	5	6 < N < 20
	99%	**6**	**3**	**4**	1 < N < 10

As can be seen in Table 1 the GARCH models are rejected for some data sets, while the PFS models are accepted for all data sets. This implies the presence of extreme values in the returns series that are not captured by the GARCH model [7]. Note that for KPN, the number of exceptions obtained for $c = 99\%$ is on the limit indicated by the Kupiec test.

Table 2 shows the initial probability parameters obtained with (12). It can be seen that all probability variables are positive according to this estimation. Table 3 and Table 4 shows the optimal probability parameters obtained after maximum likelihood estimation for PFS_{UD} and PFS_V, respectively.

Table 2: Initial probability parameters for ABN AMRO model.

Rule	Consequent 1	2	3	4	5	6	7	8	9
1	0.1003	0.1333	0.2066	0.0567	0.0441	0.1215	0.1012	0.1218	0.1143
2	0.0565	0.1351	0.1659	0.0792	0.0744	0.0972	0.1617	0.1489	0.0811
3	0.0459	0.1679	0.1495	0.1300	0.1024	0.0773	0.1463	0.1359	0.0448
4	0.0544	0.1683	0.1700	0.1105	0.0647	0.0802	0.1454	0.1579	0.0468
5	0.0516	0.1578	0.1800	0.1119	0.0692	0.0770	0.1547	0.1472	0.0506
6	0.0563	0.1648	0.1760	0.0956	0.0655	0.1206	0.1340	0.1227	0.0646
7	0.0700	0.1877	0.1659	0.0748	0.0405	0.0901	0.1147	0.2002	0.0562
8	0.0529	0.1625	0.1626	0.1122	0.0999	0.0930	0.1313	0.1286	0.0570
9	0.0539	0.1729	0.1624	0.1132	0.0746	0.0772	0.1476	0.1505	0.0476

Note that some of the probability parameters are now zero. As tables Table 3 and Table 4 show, more probability mass is centered around the middle consequents. A higher probability mass around the center may indicate that the extreme returns are much less frequent, but this is not always the case. The returns seem to revert to the average after a positive or negative extreme value, and the volatility seems to be rather different for the average returns and the extremes [16]. Hence these results capture the tail behavior in the returns distribution.

Table 3: Probability parameters for PFS_{UD} ABN AMRO model after optimisation.

Rule	Consequent 1	2	3	4	5	6	7	8	9
1	0	0	0.0397	0.2675	0.6928	0	0	0	0
2	0	0	0.1049	0.3454	0.5497	0	0	0	0
3	0	0.0090	0	0.3466	0.6225	0.0219	0	0	0
4	0.0562	0	0.0700	0.5111	0.1847	0.1224	0	0.0556	0
5	0	0.0121	0.0074	0.4865	0.4797	0	0.0142	0	0
6	0	0.0357	0	0.4496	0.4550	0.0085	0.0513	0	0
7	0.0145	0.0472	0	0.4219	0.4774	0.0226	0.0163	0	0
8	0	0	0	0.3355	0.5736	0	0.0908	0	0
9	0	0.0441	0	0.1853	0.6198	0.1286	0	0	0.0222

Table 4: Probability parameters for PFS_V ABN AMRO model after optimisation.

Rule	Consequent 1	2	3	4	5	6	7	8	9
1	0	0.0123	0.0326	0.2897	0.6654	0	0	0	0
2	0	0.0310	0	0.4509	0.4618	0	0.0563	0	0
3	0	0.0435	0	0.2054	0.6082	0.1208	0	0	0.0222
4	0.0596	0	0.0719	0.5362	0.1819	0.0909	0	0.0596	0
5	0	0.0122	0	0.3315	0.6317	0.0245	0	0	0
6	0	0	0.1068	0.4104	0.4828	0	0	0	0
7	0	0.0070	0.0244	0.4559	0.4989	0	0.0138	0	0
8	0.0229	0	0.0096	0.4270	0.5231	0.0071	0.0104	0	0
9	0	0	0	0.3320	0.5878	0	0.0801	0	0

When we compare these results with the ones obtained when distributing more membership functions are around the origin [13], we can see that the tails of the return distribution are more pronounced in the presented work. We believe that this is due to the fact the in the presented work, both proposed approaches better model and capture the extreme values of the data.

The obtained probability density functions for ABN-AMRO using the PFS_V are shown in Fig. 3. As can be seen, for some cases the distribution does not resembles a normal distribution. Specifically, the cases were the distribution exhibits fat tails, correspond to high volatility data.

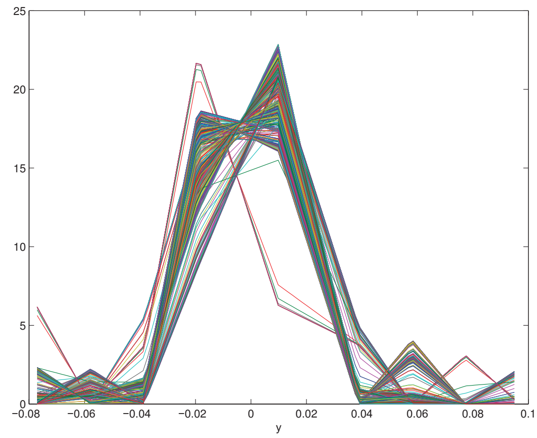


Figure 3: Obtained pdf ABN-AMRO using PFS_V

It is also interesting to consider how the VaR values estimated by the PFS compare to the values estimated by the GARCH models. Table 5 shows the sum of the differences between the VaR estimated and the actual losses in the periods where the VaR estimation is smaller than the actual losses, i.e., when exceptions occur. As can be seen in Table 5, the ex-

ISBN: 978-989-95079-6-8

Table 5: VaR Exceptions

Asset	c	PFS_{UD}	PFS_V	GARCH
ABN	95%	0.0339	0.0333	0.0446
	97.5%	0.0430	0.0452	0.0510
	99%	0.0586	0.0558	0.0587
KPN	95%	0.0591	0.0596	0.0887
	97.5%	0.0744	0.0724	0.1015
	99%	0.0902	0.0881	0.1166
JiaLing	95%	0.0301	0.0288	0.0405
	97.5%	0.0371	0.0348	0.0463
	99%	0.0469	0.0434	0.0533
BaoShan	95%	0.0439	0.0244	0.0350
	97.5%	0.0478	0.0314	0.0400
	99%	0.0531	0.0463	0.0460
COSCO	95%	0.0327	0.0330	0.0499
	97.5%	0.0428	0.0429	0.0571
	99%	0.0561	0.0570	0.0656
Merchant	95%	0.0264	0.0266	0.0362
	97.5%	0.0312	0.0318	0.0413
	99%	0.0371	0.0388	0.0475

pected losses are in most cases smaller in the PFS models, with the exception of BAOSHAN. In other cases where the GARCH model has smaller total expected losses than the PFS model, the GARCH model leads to a smaller number of exceptions, which may indicate a conservative model.

7 Conclusions

We have proposed two approaches for determining the model parameters of probabilistic fuzzy systems for value at risk modelling, where both approaches give more relevance to the tail points of the distribution of values. In one of the approaches we distributed more membership functions around the edges of the universe of discourse and used maximum likelihood to estimate the conditional parameters of the model. In the other approach we distributed the membership functions uniformly over the universe of discourse and used a weighted maximum likelihood to estimate the conditional parameters of the model.

The performance of the proposed models has been compared to the VaR estimation by using the popular GARCH (1,1) volatility estimation. It is found that PFS models are not rejected by back testing, while GARCH models are sometimes rejected. Furthermore, we show how VaR estimation is improved by using these approaches. This shows the added flexibility that comes through the use of the probabilistic fuzzy models, enabling them to adapt to the properties of the data and of the problem at study.

References

[1] R. A. Brealey and S. C. Myers. *Principles of Corporate Finance*. McGraw Hill, New York, 7th edition edition, 2001.

[2] J. Hull. *Options, Futures and Other Derivatives*. Prentice Hall, Upper Saddle River, 4 edition, 2000.

[3] Philippe Jorion. *Value at Risk : the new benchmark for managing financial risk*. McGraw–Hill, New York, 2 edition, 2001.

[4] Glyn A. Holton. *Value at Risk: thoery and practice*. Academic Press, Amsterdam, The Netherlands, 2003.

[5] T. Bollerslev. Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 31:307–327, 1986.

[6] Rama Cont. Volatility clustering in financial markets: empirical facts and agent-based models. In G. Teyssiére and A. P. Kirman, editors, *Long Memory in Economics*, pages 289–310. Springer, Berlin, 2007.

[7] N.S. Balke and T.B. Fomby. Large shocks, small shocks, and economic fluctuations: outliers in macroeconomic time series. *Journal of Applied Econometrics*, 9:181–200, 1994.

[8] J. W. Taylor. A quantile regression neural network approach to estimating the conditional density of multiperiod returns. *Journal of Forecasting*, 19(4):299–311, July 2000.

[9] N. Chapados and Y. Bengio. Cost functions and model combination for var-based asset allocation using neural networks. *IEEE Transactions on Neural Networks*, 12(4):890–906, July 2001.

[10] M. Jiang, Z. Q. Meng, and Q. Y. Hu. A neural network model on solving multiobjective conditional value-at-risk. In F. Yin, J. Wang, and C. Guo, editors, *Advances in Neural Networks*, number 3174 in Lecture Notes in Computer Science, pages 1000–1006. Springer, Berlin, 2004.

[11] Zdeněk Zmeškal. Value at risk methodology under soft conditions approach (fuzzy-stochastic approach). *European Journal of Operational Research*, 161:337–347, 2005.

[12] Umberto Cherubini and Giovanni Della Lunga. Fuzzy value-at-risk: accounting for market liquidity. *Economic Notes by Banca Monte dei Paschi di Siena SpA*, 30(2):293–312, 2001.

[13] Du Xu and Uzay Kaymak. Value-at-risk estimation by using probabilistic fuzzy systems. In *Proceedings of the 2008 World Congress on Computational Intelligence*, pages 2109–2116, Hong-Kong, June 2008.

[14] Uzay Kaymak, Willem-Max van den Bergh, and Jan van den Berg. A fuzzy additive reasoning scheme for probabilistic Mamdani fuzzy systems. In *Proceedings of the 2003 IEEE International Conference on Fuzzy Systems*, volume 1, pages 331–336, St. Louis, USA, May 2003.

[15] L. A. Zadeh. Probability measures of fuzzy events. *Journal of Mathematical Analysis and Applications*, 23(2):421–427, 1968.

[16] Jan van den Berg, Uzay Kaymak, and Willem-Max van den Bergh. Financial markets analysis by using a probabilistic fuzzy modelling approach. *International Journal of Approximate Reasoning*, 35:291–305, 2004.

[17] Darrell Duffie and Jun Pan. An overview of value at risk. *Journal of Derivatives*, 4(3):7–49, Spring 1997.

[18] Z. Wiener. Introduction to VaR (value-at-risk). In D. Galai, D. Ruthenberg, M. Sarnat, and B. Schreiber, editors, *Risk Management and Regulation in Banking*, pages 47–63. Kluwer Academic Publishers, Boston, 1999.

[19] Paul H. Kupiec. Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*, 2:73–84, 1995.

[20] J. C. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York, 1981.

[21] L. Waltman, U. Kaymak, and J. van den Berg. Maximum likelihood parameter estimation in probabilistic fuzzy classifiers. In *Proceedings of the 14th Annual IEEE International Conference on Fuzzy Systems*, pages 1098–1103, Reno, Nevada, USA, May 2005.

ISBN: 978-989-95079-6-8