Fuzzy Logic-based Portfolio Selection with Particle Filtering and Anomaly Detection

Masafumi Nakano
Graduate School of Economics, The University of Tokyo

Akihiko Takahashi
The University of Tokyo

Soichiro Takahashio
Graduate School of Economics, The University of Tokyo

February 2017

Fuzzy Logic-based Portfolio Selection with Particle Filtering and Anomaly Detection

Masafumi Nakano, Akihiko Takahashi, Soichiro Takahashi

Graduate School of Economics, University of Tokyo,

7-3-1 Hongo Bunkyo-ku, Tokyo, Japan, 113-0033

February 16, 2017

Abstract

This paper proposes a new fuzzy logic (FL)-based expert system with particle filtering and anomaly detection to create high-performance investment portfolios. In particular, our FL system selects a portfolio with fine risk-return profiles from a number of candidates by integrating multilateral performance measures including Sharpe, Sortino and Sterling ratios. The candidates consist of various mean-variance portfolios with multiple time-series models estimated by a particle filter and anomaly detectors. In an out-of-sample numerical experiment with a dataset of international financial assets, we demonstrates our expert system successfully generates a series of mean-variance portfolios with satisfactory investment records.

Keywords: expert system, fuzzy logic, performance measure, mean-variance portfolio, particle filtering, anomaly detection

1 Introduction

It is well-recognized both in academia and industry that financial markets are highly complex and non-linear systems with large noise, which are affected by economic, political, geopolitical and psychological factors. Therefore, in the analysis of financial investment problems, machine learning (ML) techniques based on sophisticated computer science are widely used since they are effective to deal with non-linearity and uncertainty (e.g., Cavalcante, Brasileiro, Souza, Nobrega, & Oliveira, 2016).

Fuzzy logic (FL) is considered to be one of the ML techniques, which is applied with great success in financial investment problems. Especially, fuzzy set theory, introduced by Zadeh (1965), is used in portfolio optimization problems since it enables to represent imperfect knowledge or ambiguity for the future asset return. For example, although a mean-variance (MV) portfolio (Markowitz, 1952) has been one of the most famous strategies, there is a well-known serious problem that the direct MV optimization amplifies the effects of estimation errors (e.g., Michaud, 1989). Consequently, many researchers introduce the fuzziness in portfolio optimization problems from various perspectives (e.g., Chen & Huang, 2009; Fang, Lai & Wang, 2006; Huang, 2008; Jalota, Thakur, & Mittal, 2017; Li, Guo & Yu, 2015; Li & Xu, 2009, 2011; Liu, 2011; Mehlawat, 2016; Mehlawat & Gupta, 2014; Nguyen & Gordon-Brown, 2012; Nguyen, Gordon-Brown, Khosravi, Creighton, & Nahavandi, 2015; Pai, 2016; Wang, Wang, & Watada2011; Zhou, Li, & Pedrycz, 2016).

Also, a fuzzy logic (FL) system (e.g., Mamdani & Assilian, 1975) is applied to the construction of trading strategies based on technical indicators, fundamental analysis and text data (e.g., Dourra & Siy, 2002; Dymova, Sevastianov & Bartosiewicz, 2010; Geva & Zahavi, 2014; Goumatianos, Christou & Lindgren, 2013; Kosaka, Mizuno, Sasaki, Someya, & Hamada, 1991; Lam, 2001; Lincy & John, 2016; Simutis, 2000). In trading practice, domain expert knowledge is often expressed in the linguistic form, which can be incorporated into trading strategies through IF-THEN rules of FL systems.

For instance, Chourmouziadis and Chatzoglou (2016) have designed a short-term trading fuzzy system based on novel technical indicators. Also, Yunusoglu and Selim (2013) have developed a fuzzy rule-based expert system (ES) for stock evaluation and portfolio construction by using technical and fundamental analysis with an application to Istanbul Stock Exchange.

This paper provides a new FL-based expert system (ES) to create a high-performing portfolio. Specifically, our ES consists of three stages: estimation, portfolio construction and FL-based selection. That is, we first estimate expected return and volatility with several time-series models, and then calculate MV optimal portfolio weights based on the predictors. Finally, our FL system, which combines multiple investment criteria, evaluates the historical performances of each portfolio and selects the best one. We remark that our FL system is applicable to fund of funds investment since it provides a general framework to portfolio evaluation.

In the estimation step, we introduce state space models to obtain the estimates of expected return and volatility for MV portfolio construction. State space models are commonly used in various fields to represent the dynamic dependence between the latent state and observed variables. Their dynamics are described as stochastic processes called system and observation models. Moreover, applying filtering methods, we can partially observe latent state variables with noises through the observation model.

In the current work, we regard expected return and volatility as state variables, which are observed with noise as asset return. For model estimation, we resort to a particle filtering (PF) method applicable to non-linear and non-Gaussian models (Gordon, Salmond, & Smit, 1993; Kitagawa, 1996). PF is an on-line estimation algorithm to a time-series data under the state-space representation of models, which takes much less computational time than repeated

implementation of off-line algorithms with sliding windows.

In particular, we assume an exponential moving average (EMA) and generalized EMA as system models. The former is often used both in investment practice and previous ML researches (e.g., Huang, Yang, & Chuang, 2008; Patel, Shah, Thakkar, & Kotecha, 2015a, 2015b; Rather, Agarwal, & Sastry, 2015), and the latter is introduced in Nakano, Takahashi, and Takahashi (2017) to obtain marginally better predictions based on EMA.

In addition, we exploit three anomaly detection (AD) methods to refine investment universe (Nakano et al., 2017). Since we can obtain the model likelihoods or the distributions of state variables for each time step, PF is easily combined with on-line detection schemes. In our investment problem, realized asset returns sometimes largely deviate from the models, whence it is inappropriate to implement predictions. Hence, by excluding the assets for which anomalies are detected from investment universe, we can enhance portfolio performance.

Our investment universe is designed by international equities, REITs and bonds with cash. Besides, we employ various investment criteria, i.e., compound return (CR), standard deviation (SD), downside deviation (DD), maximum drawdown (MDD), Sharpe ratio (ShR), Sortino ratio (SoR) and Stering ratio (StR), which enable multilateral assessment.

The remainder of this paper is organized as follows. Section 2 presents composition of our ES: PF-based estimation with AD scheme, MV portfolio simulation and FL-based selection. Section 3 shows the results of out-of-sample numerical experiments. Finally, Section 4 concludes.

2   Expert system

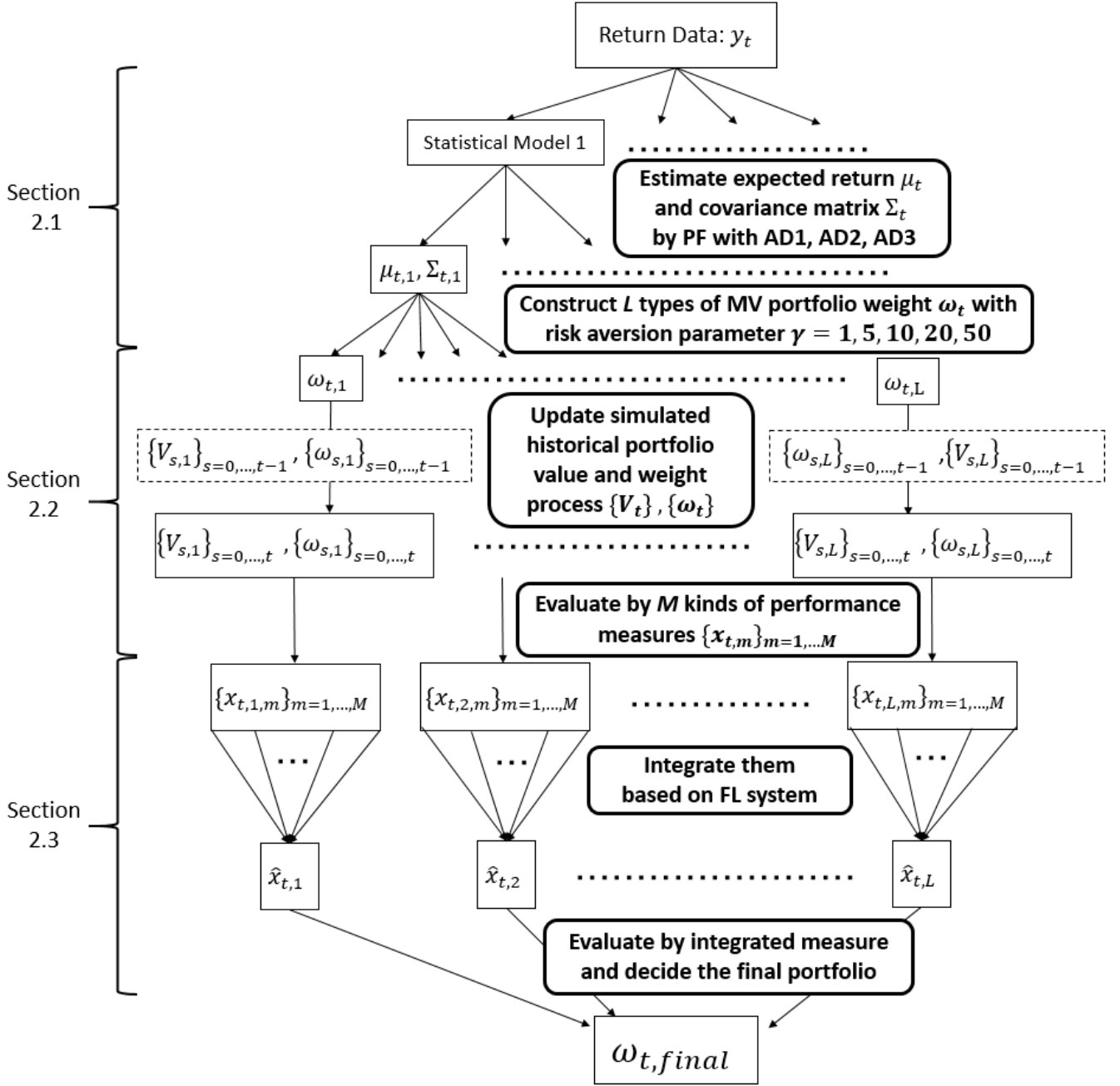Our expert system (ES) is summarized by the following flowchart.

Fig. 1: Flowchart of our expert system

Here, we shortly describe the composition of our ES. Firstly, we apply particle filtering with anomaly detectors to various time-series models, which provides expected return, volatility and investment universe in real-time. Secondly, at each time, we simulate MV portfolios based on each estimation result for several risk aversion parameters. At this stage, we run L kinds of MV portfolio simulations. Thirdly, we assess these simulated MV portfolios with M kinds of investment performance measures commonly used in practice. Lastly, we exploit a FL system

for integrating these evaluations, which gives us the final portfolio.

2.1 Estimation

2.1.1 State space modeling

In constructing mean-variance (MV) portfolios, we need to estimate expected return and volatility. For estimation, this paper employs a state space model that consists of the following system and observation model.

$$
\begin{aligned}
Y_t &= H(Z_t, u_t), \quad \text{[observation model]} \\
Z_t &= F(Z_{t-1}, v_t), \quad \text{[system model]}
\end{aligned}
\tag{1}
$$

where Z_t and Y_t denote a n-dim state and m-dim observation vector at time t, respectively. In state space modeling, we suppose that there exist unobservable latent/state variables Z_t driven by the system model $Z_t = F(Z_{t-1}, v_t)$, which are observed with noise as $Y_t = H(Z_t, u_t)$. Here, $H : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ and $F : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ are non-linear functions, and the observation noise u_t and system noise v_t are random variables which do not necessarily follow Gaussian distributions.

In general, unobservable state variables are estimated by the statistical on-line algorithm called filtering. Hence, we can sequentially estimate expected return and volatility by regarding them as state variables, which are observed with noise as realized asset return.

2.1.2 Particle filtering

The objective of filtering is to sequentially estimate the unobservable state variables Z_t through the observation variables $Y_{1:t} \equiv (Y_1, \cdots, Y_t)$ until time t, that is, to estimate the density $p(Z_t|Y_{1:t})$ called the filtering density.

Especially, particle filtering (PF) resorts to Monte Carlo simulations for state estimation, whereby it is applicable to non-linear and non-Gaussian settings. We first explain the PF algorithm in the case that all parameters are known (Gordon et al., 1993; Kitagawa, 1996).

First of all, Bayes formula shows

$$
p(Z_t|Y_{1:t}) \propto p(Y_t|Z_t)p(Z_t|Y_{1:t-1}).
\tag{2}
$$

Therefore, we exploit a sampling important resampling (SIR) method to obtain the samples $\{\hat{Z}_t^{(r)}\}_{r=1,\cdots,L}$ from the posterior $p(Z_t|Y_{1:t})$. That is, by regarding the prior $p(Z_t|Y_{1:t-1})$ as an importance function, we first draw $\{Z_t^{(r)}\}_{r=1,\cdots,L}$ from the distribution $p(Z_t|Y_{1:t-1})$, then resample from them with weights proportional to the likelihood $p(Y_t|Z_t)$, which gives us $\{\hat{Z}_t^{(r)}\}_{r=1,\cdots,L}$.

Here, we get the samples $\{Z_t^{(r)}\}_{r=1,\cdots,L}$ from the prior $p(Z_t|Y_{1:t-1})$ by the following way. Note that

$$
\begin{aligned}
&p(Z_t|Y_{1:t-1}) \\
&= \int \int p(Z_t, Z_{t-1}, v_t|Y_{1:t-1}) dZ_{t-1} dv_t \\
&= \int \int p(Z_t|Z_{t-1}, v_t, Y_{1:t-1}) p(v_t|Z_{t-1}, Y_{1:t-1}) p(Z_{t-1}|Y_{1:t-1}) dZ_{t-1} dv_t \\
&= \int \int p(Z_t|Z_{t-1}, v_t) p(v_t) p(Z_{t-1}|Y_{1:t-1}) dZ_{t-1} dv_t \\
&= \int \int \delta(Z_t - F(Z_{t-1}, v_t)) p(v_t) p(Z_{t-1}|Y_{1:t-1}) dZ_{t-1} dv_t,
\end{aligned}
\tag{3}
$$

where $\delta(z)$ is Dirac delta function. Then, for given the samples of $\{v_t^{(r)}\}_{r=1,\cdots,R}$ and $\{\hat{Z}_{t-1}^{(r)}\}_{r=1,\cdots,R}$ from $p(v_t)$ and $p(Z_{t-1}|Y_{1:t-1})$, we obtain the samples $\{Z_t^{(r)}\}_{r=1,\cdots,R}$ from $p(Z_t|Y_{1:t-1})$ by setting $Z_t^{(r)} = F(\hat{Z}_{t-1}^{(r)}, v_t^{(r)})$.

Particle filtering (PF) algorithm

1. Generate the initial state vector $\{\hat{Z}_0^{(1)}, \cdots, \hat{Z}_0^{(R)}\}$.

2. Apply the following steps (a)$\sim$(d) to each time $t = 1, \cdots, T$.

    (a) Generate system noise $v_t^{(r)}$, $r = 1, \cdots, R$.

    (b) Compute for each $r = 1, \cdots, R$ $\quad Z_t^{(r)} = F(\hat{Z}_{t-1}^{(r)}, v_t^{(r)})$.

    (c) Evaluate the weights of particles $\{Z_t^{(1)}, \cdots, Z_t^{(R)}\}$ as $\delta_t^{(r)} \equiv p(Y_t|Z_t^{(r)})$, $r = 1, \cdots, R$ by using the likelihood function.

    (d) Resample $\{\hat{Z}_t^{(1)}, \cdots, \hat{Z}_t^{(R)}\}$ from $\{Z_t^{(1)}, \cdots, Z_t^{(R)}\}$. More precisely, resample each $\hat{Z}_t^{(r')}$, $r' = 1, \cdots, R$ from $\{Z_t^{(1)}, \cdots, Z_t^{(R)}\}$ with the probability given by

    $$\text{Prob.}(\hat{Z}_t^{(r')} = Z_t^{(r)}|Y_t) = \frac{\delta_t^{(r)}}{\sum_{k=1}^{R} \delta_t^{(k)}}, \quad r = 1, \cdots, R.$$

Here, the likelihood at time t, $p(Y_t|Z_t)$, is approximately calculated by

$$p(Y_t|Z_t) \simeq \frac{1}{R} \sum_{k=1}^{R} \delta_t^{(k)}. \tag{4}$$

Next, we discuss the case that there exist unknown parameters. The unknown parameters' vector θ_t is sequentially estimated by augmenting the state vector as $\tilde{Z}_t = (Z_t, \theta_t)$. If the transition of θ_t follows

$$\theta_t = \theta_{t-1}, \tag{5}$$

this algorithm will degenerate in the sense that almost all of the particles quickly reach zero weight. Besides, parameter estimation does not work when the true values are not included in particles generated by initial distributions. Then, we add an artificial noise ζ_t to Eq. (5):

$$\theta_t = \theta_{t-1} + \zeta_t. \tag{6}$$

This framework is called "self-organizing state space model" by Kitagawa (1998), which enables us to estimate states and parameters simultaneously.

For implementing the self-organizing method, it is necessary to specify the distribution of the artificial noise ζ_t conditioned on θ_{t-1}. In this paper, we use a kernel smoothing (KS) method developed by Liu and West (2001). In the KS method, the distribution of ζ_t conditioned on θ_{t-1} is given by

$$p(\zeta_t|\theta_{t-1}) \sim N((a-1)(\theta_{t-1} - \bar{\theta}_{t-1}), (1-a^2)s_{t-1}^2), \tag{7}$$

i.e. the conditional distribution of θ_t is

$$p(\theta_t|\theta_{t-1}) \sim N(a\theta_{t-1} + (1-a)\bar{\theta}_{t-1}, (1-a^2)s_{t-1}^2), \tag{8}$$

6

where $a = (3d-1)/2d$. $\bar{\theta}_t$ and s_t^2 represent the mean and variance of the particles $\{\theta_t^{(r)}\}_{r=1,\cdots,R}$, respectively. d is a shrinkage factor which usually takes a value between 0.95 and 0.99. Here, we set $d = 0.98$.

In this paper, we execute PF individually for each asset using 1,000,000 particles. In other words, we do not take into account the correlations among asset returns, which reduces estimation errors and computational complexity.

2.1.3 Asset return model

In the current work, we assume that the dynamics of observed asset returns $y = \{y_t;\ t = 0, 1, \cdots, T\}$ are represented by the following observation model:

$$y_t = \mu_t + \sigma_t \epsilon_t, \quad \epsilon_t \sim i.i.d.\ N(0,1), \quad t \geq 0, \tag{9}$$

where $\mu = \{\mu_t;\ t = 0, \cdots, T\}$ and $\sigma = \{\sigma_t;\ t = 0, \cdots, T\}$ are state variables which stand for expected return and volatility processes, respectively.

With regard to system models, we use the next two time-series models for μ.

- EMA model:

$$\mu_t = \beta_\mu y_{t-1} + (1 - \beta_\mu)\mu_{t-1}, \quad t \geq 1,$$
$$\mu_0 = \alpha_\mu \in \mathbb{R}, \tag{10}$$

- generalized EMA (GEMA) model:

$$\mu_t = \beta_\mu y_{t-1} + (1 - \beta_\mu)\mu_{t-1} + \sigma_\mu \eta_t, \quad t \geq 1,$$
$$\mu_0 \sim N(\alpha_\mu, \sigma_\mu^2/(1 - \beta_\mu^2)), \quad \eta_t \sim i.i.d.\ N(0,1), \quad Cov(\epsilon_t, \eta_t) = 0, \tag{11}$$

where $\sigma_\mu(> 0)$ is a constant unknown parameter and $\beta_\mu \in (0, 1)$ is the so-called smoothing factor that represents the degree of weighting decrease in EMA. Note that the GEMA model is a stochastic model which aims to obtain marginally better estimation than EMA.

As well, we assume EMA and GEMA model for volatility process σ_t.

- EMA (IGARCH(1,1)) model:

$$\sigma_t^2 = \beta_\sigma(y_{t-1} - \mu_{t-1})^2 + (1 - \beta_\sigma)\sigma_{t-1}^2,$$
$$\sigma_0 = \alpha_\sigma \in \mathbb{R} \tag{12}$$

- GEMA model:

$$\log \sigma_t^2 = \log\{\beta_\sigma(y_{t-1} - \mu_{t-1})^2 + (1 - \beta_\sigma)\sigma_{t-1}^2\} + \sigma_\sigma \xi_t, \quad t \geq 1,$$
$$\log \sigma_0^2 \sim N(\log \alpha_\sigma^2, \sigma_\sigma^2), \quad \xi_t \sim i.i.d.\ N(0,1), \quad Cov(\epsilon_t, \xi_t) = Cov(\eta_t, \xi_t) = 0, \tag{13}$$

where $\beta_\sigma \in (0, 1)$ is the so-called smoothing factor of EMA and $\sigma_\sigma(> 0)$ is a constant unknown parameter.

Now, let us summarize the system models, i.e. the combinations of μ and σ processes. Namely, "EMA+GEMA", "GEMA+EMA" and "GEMA+GEMA" are stochastic models for improving estimation of simple EMA in volatility, expected return and both expected return & volatility, respectively. Here, we employ the notation "model I + model II" representing model I for μ and model II for σ.

Let us remark that as the estimates of expected return and volatility for time t, we take averages of one-step-ahead predictions at time $t - 1$ for state variables μ_t and σ_t, that is the

particles $\{Z_t^{(r)}\}_{r=1,\cdots,R}$ obtained from the step (b) of PF algorithm, described in Section 2.1.2. With regard to correlation, we employ the standard EMA model, i.e.,

$$\rho_{t,i,j} = \frac{\sigma_{t,i,j}}{\sigma_{t,i,i}\sigma_{t,j,j}}, \tag{14}$$

where $\sigma_{t,i,j} = \beta_\rho(y_{t-1,i} - \mu_{t-1,i})(y_{t-1,j} - \mu_{t-1,j}) + (1 - \beta_\rho)\sigma_{t-1,i,j}$.

Here, we describe the setting of model parameters. First, we notice that in the above models the smoothing factors of EMA/GEMA are fixed. In detail, we consider the case $\beta_\mu = \beta_\sigma = \beta_\rho$ ($\equiv \beta$) for simplicity and set $\beta = 0.2, 0.4, 0.6, 0.8$ in advance. Although this parameter strongly affects the investment results, it is very difficult to know the optimal level in advance. It may seem natural to estimate it by a self-organizing framework, but its estimates are optimal in terms of likelihood, not investment performance. Therefore, we do not estimate β in this stage. Instead, our FL system, introduced later, selects the best one in terms of investment record.

Second, as for constant α_μ and α_σ in EMA/GEMA, we set as the sample mean and standard deviation of asset returns in the first two years, $\{y_t\}_{t=0,\cdots,23}$, respectively. With regard to an unknown parameter σ_μ in GEMA, we guess the value to utilize the data during the training period. That is, we use 0.5 times the sample standard deviation of EMAs in the first two years, Different from σ_μ, it is difficult to guess a reasonable value of σ_σ. Therefore, we estimate it with a self-organizing approach. For their initial distributions, we draw σ_σ from the uniform distribution $U(0,1)$ following the previous research (Nakano et al., 2017).

2.1.4 Anomaly detection

We also adopt three types of anomaly detectors in order to judge whether the models really capture asset returns' dynamics, introduced in Nakano et al. (2017). If an anomaly is detected for an asset return at time t, y_t, we exclude the asset from investment universe at time t, which makes it possible to enhance our investment performances.

The first type utilizes the log-likelihoods. Here we define the log-likelihood at time t by $\ell(y_t) \equiv \log p(y_t|Z_t)$ which is approximately calculated with taking log of the right hand side of Eq. (4). In our PF algorithm, we are able to obtain $\{\ell(y_t)\}_t$ for each asset. Then, if $\ell(y_t)$ takes a lower value than a predetermined threshold at time t, we regard y_t as an anomaly.

In the second type, we employ the traditional Hotelling approach since the observational noise ϵ_t follows the standard normal distribution. In this approach, we define an anomaly indicator $a(y_t)$ by the negative log-likelihood $-\ell(y_t)$. We rewrite $a(y_t)$ by using only the term which includes y_t as follows.

$$a(y_t) = \left(\frac{y_t - \mu_t}{\sigma_t}\right)^2. \tag{15}$$

Note that in our models $a(y_t)$ coincides with the square of observation error, ϵ_t^2. Again, this can be approximately calculated in our estimation algorithm. Then the remaining task is to determine a threshold as in the first approach. Here, we make use of the fact that this $a(y_t)$ asymptotically follows a chi-squared distribution with one degrees of freedom $\chi^2(1)$. Although the number of our data may not be large enough to utilize this asymptotic property, we approximately use the 95 percentile of $\chi^2(1)$ as a threshold.

In the third type, we test a method by the one-step-ahead predictive distribution of asset returns, $p(y_t|y_{1:t-1})$ where $y_{1:t-1} \equiv (y_1, \cdots, y_{t-1})$. In PF algorithm, we can obtain the approximation of predictive distribution $p(y_t|y_{1:t-1})$ based on Monte Carlo simulation. Then, we are able to detect an anomaly by calculating the 2.5 and 97.5 percentiles of $p(y_t|y_{1:t-1})$. If the realized return y_t of some asset is outside these percentiles, we exclude this asset from our investment universe.

In the following, we call those three AD methods AD1, AD2 and AD3, respectively. Note that all the methods are implemented in our PF algorithm quite easily. In contrast to Nakano et al. (2017) not offering a useful way for deciding in advance which AD is the best, our ES proposes a method for its decision.

2.2 Mean-variance portfolio and performance measure

Suppose that there exists a risk-free asset in the financial market and the risk-free rate is zero. We also put no-short-sale constraint. Then, the weight of MV portfolio at each time step $t = 0, 1, \cdots, T-1$ is given by the solution of the following optimization problem for given estimates of an expected return vector $\mu_t = (\mu_{t,i})_{i=1,\cdots,N}$ and a covariance matrix $\Sigma_t = (\sigma_{t,i,j})_{i,j=1,\cdots,N}$.

$$
\max_{\omega_t} \quad \omega_t' \mu_t - \frac{\gamma}{2} \omega_t' \Sigma_t \omega_t,
$$
$$
s.t. \quad \omega_{t,i} \geq 0 \ , \ \sum_{i=1}^{N} \omega_{t,i} \leq 1,
\tag{16}
$$

where N denotes the number of risky assets composing the investment universe, and $\omega_t = (\omega_{t,1}, \cdots, \omega_{t,N})'$ is a portfolio weight vector of risky assets in the time interval $[t, t+1)$. Besides, the parameter γ indicates degree of risk aversion and we test $\gamma = 1, 5, 10, 20, 50$. Although the investment result of MV portfolio largely depends on risk aversion parameter, our ES can select its optimal value in terms of portfolio performance as well as the smoothing factor. We notice that there are sometimes abnormal periods when the investment universe is not large enough to construct MV portfolio due to the anomaly detectors, that is $N = 0, 1$. In this case, we do not invest risky asset at all. Remark that the weight of risk-free asset $\omega_{t,N+1}$ is determined by $\omega_{t,N+1} = 1 - \sum_{i=1}^{N} \omega_{t,i}$.

The portfolio values $\{V_t\}_{t=0,\cdots,T}$ and portfolio returns $\{R_t\}_{t=1,\cdots,T}$ are defined as follows.

$$
V_{t+1} = V_t \left(1 + \sum_{i=1}^{N} \omega_{t,i} y_{t,i} \right) - \sum_{i=1}^{N} c_i |\omega_{t,i} V_t - \omega_{t-1,i} V_{t-1}(1 + y_{t,i})| \ , \quad V_0 = 1,
$$
$$
R_{t+1} = \frac{V_{t+1}}{V_t} - 1,
\tag{17}
$$

where c_i and $y_{t,i}$ denote a transaction spread and a return of i-th risky asset, respectively.

The penalty term $\sum_{i=1}^{N} c_i |\omega_{t,i} V_t - \omega_{t-1,i} V_{t-1}(1 + y_{t,i})|$ of Eq. (17) is the total transaction cost arising from the portfolio re-balance at time t. Since $\omega_{t-1,i}$ and $\omega_{t,i}$ are portfolio weights of the i-th risky asset during $[t-1, t)$ and $[t, t+1)$, $\omega_{t-1,i} V_{t-1}(1 + y_{t,i})$ and $\omega_{t,i} V_t$ indicate the values of the i-th risky asset before and after the position change at time t, respectively. That is, $|\omega_{t,i} V_t - \omega_{t-1,i} V_{t-1}(1 + y_{t,i})|$ represents the necessary amount of money for the position change of asset i at time t. Hence, the total transaction cost at time t equals to the summation of $c_i |\omega_{t,i} V_t - \omega_{t-1,i} V_{t-1}(1 + y_{t,i})|$ for all i. In this paper, we set $c_i = 10$ *bps* for all risky assets.

In the following, we briefly describe the well-known performance measures used in our fuzzy inference system.

- Compound Return (CR): We define CR as the annualized geometric average of the portfolio returns $\{R_t\}$ defined in Eq. (17), which is a standard measure of investment returns.

$$
CR \equiv \left\{ \prod_{t=1}^{T} (1 + R_t) \right\}^{12/T} - 1.
\tag{18}
$$

- Standard Deviation (SD), Downside Deviation (DD): SD is a well-known investment risk measure defined as the annualized standard deviation of $\{R_t\}$, while DD only regards negative returns as risk.

$$SD \equiv \left\{\frac{12}{T}\sum_{t=1}^{T}(R_t - \bar{R})^2\right\}^{1/2}, \ DD \equiv \left\{\frac{12}{T}\sum_{t=1}^{T}\min(0, R_t)^2\right\}^{1/2}, \ \bar{R} \equiv \frac{1}{T}\sum_{t=1}^{T}R_t. \ (19)$$

- Maximum Drawdown (MDD):

$$MDD \equiv \max_{1 \leq t \leq T}\frac{M_t - V_t}{M_t}, \quad M_t \equiv \max_{0 \leq s \leq t}V_s. \tag{20}$$

The drawdown is the decline from the past peak value M_t to the present value V_t. In general, portfolio performance depends on the investment timing. The MDD contributes to the performance analysis because it is independent of the investment timing given the horizon $[0, T]$.

- Sharpe Ratio (ShR): ShR is usually defined as portfolio excess average returns divided by portfolio standard deviation. Since interest rates on cash are assumed to be zero, we define ShR as follows.

$$ShR \equiv AR/SD, \quad AR \equiv 12\bar{R}. \tag{21}$$

Here, AR denotes the annualized arithmetic average of $\{R_t\}$, which corresponds to a simple return.

- Sortino Ratio (SoR): SoR does not regard upside volatility as a risk while ShR penalizes both upside and downside volatility, which is often pointed out as a weakness of ShR.

$$SoR \equiv AR/DD. \tag{22}$$

- Sterling Ratio (StR): StR is a measure of risk-adjusted return that uses drawdown measures as denominator. We adopt the following definition:

$$StR \equiv AR/MDD. \tag{23}$$

2.3 Fuzzy logic system

In the above sections, there exists much variety for time-series modeling, that is EMA+GEMA, GEMA+EMA and GEMA+GEMA for each $\beta = 0.2, 0.4, 0.6, 0.8$ and anomaly detectors, AD1, AD2 and AD3. Moreover, we assume three cases of risk aversion parameter $\gamma = 1, 5, 10, 20, 50$. In practice, it is important to select the "best" case from these 180 candidates ($= 3 \times 4 \times 3 \times 5$), though we cannot know in advance which is the best.

One of the effective approaches to this problem is to choose a portfolio whose past performance is the best in terms of some measures introduced in Section 2.2. Then, we apply a FL system to integrating the practically well-known performance measures, which is meaningful since a unified evaluation is difficult due to the variety of these measures.

Suppose that there are L kinds of portfolios, which are evaluated with M kinds of performance measures. Our fuzzy system is a nonlinear mapping from $\mathbb{R}^{M \times L}$ into $\mathbb{R}^L$, that is, the inputs are M well-known performance measures $\{x_{t,l,m}\}_{l=1,\cdots,L, \ m=1,\cdots,M}$ and the output is a new integrated performance measure $\{\hat{x}_{t,l}\}_{l=1,\cdots,L}$. More concretely, our fuzzy system implements the following procedures at each investment time $t = t_s, \cdots, t_e$, where $t_s > 0, \ t_e = T$.

10

(i) Firstly, performance measures $\{x_{t,l,m}\}_{l=1,\cdots,L,\ m=1,\cdots,M}$ are calculated by time-series of the past portfolio values $\{V_{s,l}; 0 \le s \le t\}$ for each portfolio $l = 1, \cdots, L$ as the inputs.

(ii) Secondly, these inputs are fuzzified by the following Gaussian membership functions associated with three kinds of fuzzy sets $X_k, k = 1, 2, 3$, i.e. High, Medium and Low.

$$MF_k(x_{t,l,m}) = \exp\left(-\frac{(x_{t,l,m} - a_{m,k})^2}{2b_m^2}\right), \quad m = 1, \cdots, M, \tag{24}$$

where $k = 1, 2, 3$ correspond with the three fuzzy sets, high, medium and low, respectively. Here, for each $m = 1, \cdots, M$, we set $a_{m,1}, a_{m,2}, a_{m,3}$ to be maximum, median and minimum of $x_{t,l,m}$ over $l = 1, \cdots, L$ and b_m to be a quarter of the difference between the maximum and the minimum.

(iii) Thirdly, we employ the following form of IF-THEN rule:

- The case of return or risk-adjusted return measures (CR, ShR, SoR, StR):
  "IF $x_{t,l,m}$ is $X_1/X_2/X_3$, i.e. High/Medium/Low, THEN $\hat{x}_{t,l,m,k} = 2/1/0$."
- The case of risk measures (SD, DD, MDD):
  " IF $x_{t,l,m}$ is $X_3/X_2/X_1$, i.e. Low/Medium/High, THEN the output $\hat{x}_{t,l,m,k} = 2/1/0$."

Here, the strength of each IF-THEN rule is evaluated by using grades of the membership function:

$$g_k(x_{t,l,m}) = \frac{MF_k(x_{t,l,m})}{\sum_{k=1,2,3,\ m=1,\cdots,M} MF_k(x_{t,l,m})}. \tag{25}$$

Note that we use $3 \times M$ number of IF-THEN rules.

(iv) Lastly, the output $\hat{x}_{t,l}$ is calculated by the Takagi-Sugeno-Kang inference method (Sugeno & Kang, 1988; Takagi & Sugeno, 1985):

$$\hat{x}_{t,l} = \sum_{k=1,2,3,\ m=1,\cdots,M} g_k(x_{t,l,m})\hat{x}_{t,l,m,k}. \tag{26}$$

Then, by comparing the values of this integrated measure $\{\hat{x}_{t,l}\}_{l=1,\cdots,L}$, the final investment portfolio at time t is created. Specifically, we use the MV portfolio which attains the highest performance in terms of the new measure at each time.

Remark that since this FL system enables to evaluate a number of portfolios in a unified framework, it is applicable to fund of funds investment. For instance, by assessing various investable funds instead of the MV portfolios, we can construct a new portfolio of the hedge funds with weights proportional to the new integrated measures $\{\hat{x}_{t,l}\}_{l=1,\cdots,L}$.

3   Numerical experiment

3.1   Data

We use monthly total returns of 8 indexes corresponding to stocks, bonds, and REITs as listed in Table 1. Hereafter we employ the abbreviations of the index names in this table. The time period of the return data is 156 months, from April 2003 to March 2016. As the PF estimation during the first several periods are strongly affected by the initial distribution, we discard the estimation results over the first 24 periods, which means $t = 0$ is April 2005. Moreover, $t = t_s$ is set to be April 2007 for the FL system that requires historical portfolio value processes, as described in

Section 2.3. Further, the threshold of AD1 is set to be the 5 percentile of $\{\ell(y_t)\}_{t=6,\cdots,47}$. We omit the first six months of the log-likelihoods because this period seems to be strongly affected by initial distributions.

An asset monthly return y_t is given by $y_t = 100 \times (P_t/P_{t-1} - 1)$ where P_t denotes the asset price at time t. Our data are downloaded from Bloomberg in JPY (Japanese yen)-denominated form so that we consider the global investment with no currency hedging, where the initial investment is made in JPY. Table 2 shows the descriptive statistics of the asset returns.

Table 1: Data

Index name	Ticker (Bloomberg)	Abbreviation
Tokyo Stock Price Index	TPXDDVD.Index	JPE
Tokyo Stock Exchange REIT Index	TPXDREIT.Index	JPR
S&P500 Index	SPTR.Index	USE
Morgan Stanley REIT Index	RMS.G.Index	USR
FTSE Developed ex North America Net Tax (US RIC) Index	TGPVAN33.Index	DE
FTSE Emerging Total Return Index	FTS5ALEM.Index	EE
Barclays US Treasury 10 Year TERM Index	BCEY4T.INDEX	USB
JPMorgan Emerging Market Bond Index	JPEIGLBL.INDEX	EB

JPE = Japanese equity index, JPR = Japanese REIT index, USE = US equity index, USR = US REIT index, DE = developed countries' equity index, EE = emerging countries' equity index, USB = US bond index, EB = emerging countries' bond index.

REIT, or Real Estate Investment Trust, is a security that invests in real estate through property or mortgages, and often trades on major exchanges like equities. REITs provide investors with an extremely liquid stake in real estate.

Table 2: Descriptive statistics

	Mean	Std. Dev.	Skew	Kurtosis
JPE	0.629	5.273	-0.407	0.855
JPR	0.953	5.721	-0.206	3.531
USE	0.853	5.279	-0.724	1.764
USR	1.193	7.338	-0.941	6.002
DE	0.825	5.903	-0.933	2.384
EE	1.216	7.188	-0.833	2.598
UB	0.430	2.572	-0.181	0.720
EB	0.726	3.486	-1.528	8.993

Skew $= \frac{1}{T}\sum_{t=1}^{T}\left(\frac{y_t-\hat{\mu}}{\hat{\sigma}}\right)^3$, Kurtosis $= \frac{1}{T}\sum_{t=1}^{T}\left(\frac{y_t-\hat{\mu}}{\hat{\sigma}}\right)^4 - 3$, where $\hat{\mu}$ and $\hat{\sigma}$ denote Mean and Std. Dev., respectively.

3.2 Out-of-sample investment result

Table 3 compares the investment results of our proposed ES with the traditional strategies such as buy-and-hold (BH) and equally-weighted (EW) strategy.

Table 3: Performance comparison with traditional strategies (%)

	CR	SD	DD	MDD	ShR	SoR	StR
our ES	12.23	11.62	4.64	8.76	105.39	264.05	139.70
EW strategy	3.91	17.26	12.61	53.63	31.25	42.76	10.05
BH strategy							
JPE	-0.66	19.77	14.25	56.23	6.67	9.26	2.35
JPR	1.88	21.98	15.34	67.56	19.62	28.12	6.38
USE	5.96	20.58	14.56	59.79	38.77	54.80	13.35
USR	-0.05	23.07	17.19	63.04	11.80	15.83	4.32
DE	4.37	28.58	20.75	73.94	30.22	41.63	11.68
EE	1.37	27.17	19.67	67.67	19.29	26.65	7.75
USB	5.72	9.16	5.71	15.25	65.44	104.95	39.32
EB	6.18	13.30	9.68	33.00	52.12	71.63	21.00

First of all, Table 3 shows that our proposed ES substantially outperforms the traditional strategies in all performance measures other than SD, which is also confirmed by Fig. 2. Although our proposed ES is slightly worse than BH strategy of USB in terms of SD, this is not problem at all because the result of DD suggests that the higher SD results from the upside deviation, i.e., the positive portfolio return.
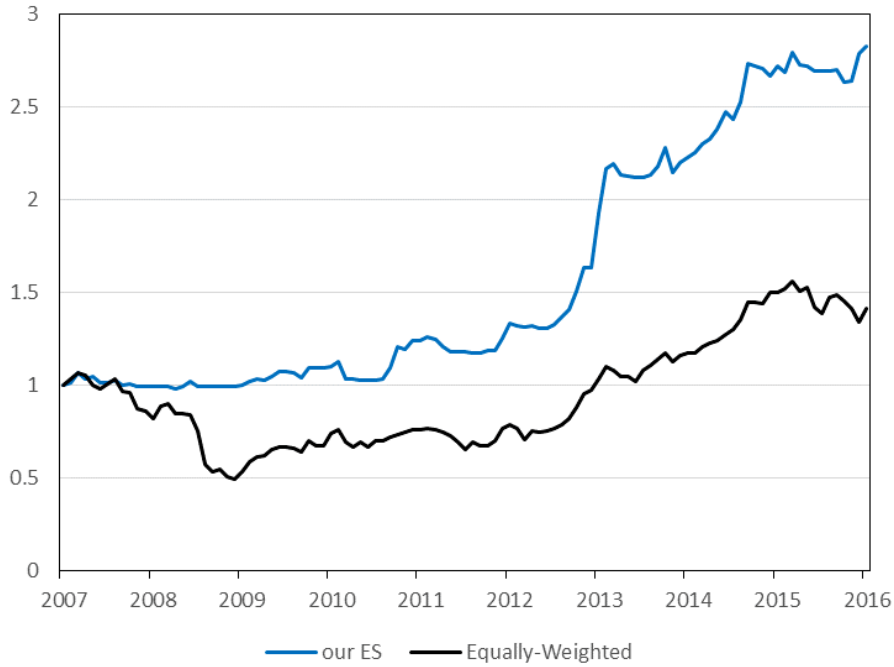

Fig. 2: Comparison of portfolio values V_t: our ES vs EW strategy

Table 4: Performance comparison with simulated MV portfolios (%)

	CR	SD	DD	MDD	ShR	SoR	StR	FL-based measure
0th percentile	3.51	6.57	3.43	7.43	27.95	45.29	12.47	0.03
25th percentile	8.32	12.68	5.76	14.05	69.99	132.87	50.68	1.00
50th percentile	12.68	14.97	6.88	17.03	90.70	195.85	80.31	1.17
75th percentile	15.15	16.13	7.79	19.50	98.95	220.71	96.17	1.27
100th percentile	17.81	19.19	11.90	42.66	109.74	272.46	145.40	1.67
our ES	12.23	11.62	4.64	8.76	105.39	264.05	139.70	1.62
ranking (out of 181)	95	35	24	9	10	3	5	3

We summarize $180 + 1$ pattern of investments by using their quantiles in each performance measure, i.e., CR, SD, DD, MDD, ShR, SoR, StR and FL-based integrated measure.

We also evaluate the final output portfolio compared to those 180 simulated MV portfolios, which are too many to present all the records. Therefore, we summarize those investment results as the quantiles for each performance measure in Table 4. It also shows ranking of the final portfolio.

From Table 4, it is clear that our ES can successfully construct a MV portfolio achieving fine risk-return profiles, though the investment performance substantially changes depending on the models, smoothing factors, anomaly detectors and risk aversion parameters. Particularly, all risk adjusted return measures (ShR, SoR and StR) of our proposed ES rank in top 10. Moreover, we can see from this table that our proposed ES in fact achieves a considerably high ranking in terms of FL-based measure. This is also clearly observed by the red line in the Fig. 3, which indicates the ranking of output portfolio in the histogram of FL-based measure.
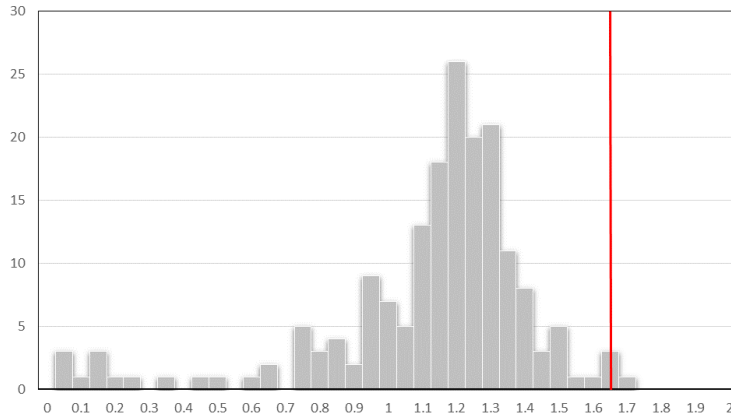

Fig. 3: Histogram of FL-based performance measure

Lastly, to test the validity of our FL system, we also construct portfolios by using just each performance measure (CR, SD, DD, MDD, ShR, SoR and StR), instead of our FL-based criteria.

Table 5: Comparison with single performance measure based evaluation

	CR	SD	DD	MDD	ShR	SoR	StR
our ES	12.23	11.62	4.64	8.76	105.39	264.05	139.70
ranking (out of 181)	95	35	24	9	10	3	5
CR-based	11.25	16.47	8.61	26.54	72.94	139.50	45.26
ranking (out of 181)	107	144	167	166	133	134	144
SD-based	3.56	6.59	3.65	14.09	56.48	101.99	26.39
ranking (out of 181)	180	3	7	48	166	166	174
DD-based	3.93	7.08	3.78	14.43	58.04	108.71	28.46
ranking (out of 181)	176	10	10	49	162	159	170
MDD-based	7.73	9.22	4.13	12.72	85.49	191.07	62.01
ranking (out of 181)	144	16	16	30	109	95	123
ShR-based	10.72	15.07	6.94	14.66	75.00	162.95	77.07
ranking (out of 181)	109	93	94	53	129	127	98
SoR-based	9.63	11.30	4.86	12.05	87.07	202.32	81.66
ranking (out of 181)	122	31	26	27	103	77	87
StR-based	11.35	11.68	4.72	12.86	98.06	242.50	89.04
ranking (out of 181)	103	36	24	30	50	12	63

Table 5 shows the investment records of single measure-based portfolios compared with our FL system. From Table 5, it is obvious that the single measure-based assessment does not work at all, especially in return measures and risk adjusted returns. That is, in the cases of CR, ShR, SoR and StR, the investment performances drastically get worse than our FL-based evaluation. With regard to the risk measures, i.e., SD, DD and MDD, even if portfolio risk becomes relatively low, investment return and risk-adjusted return (CR, ShR, SoR and StR) are much worse than our FL case. As a result, our FL system, which combines various investment criteria, substantially contributes to the selection of a well-performed MV portfolio.

4 Conclusion

In this paper, we have proposed a novel fuzzy logic (FL)-based expert system (ES) to select a portfolio with fine risk-return profiles from a number of investment candidates. In particular, our ES directly links to high performance because it chooses the best portfolio in terms of various evaluation criteria frequently used in practice.

Concretely, our ES consists of three stages: estimation, portfolio construction and selection with a FL system. First, we assume EMA-based models for multiple smoothing factors to prevent model misspecification, which are estimated by particle filtering with anomaly detectors. Then, we calculate MV portfolio weights with regard to various models, anomaly detectors and risk aversion parameters. Lastly, we assess historical records for each portfolio based on a FL system through integrating well-known performance measures, and select the best portfolio. Further, an out-of-sample numerical experiment has confirmed that our ES generates a satisfactory investment record.

Although the current paper has considered the limited cases, i.e. only EMA-based models and MV portfolios, the proposed ES is effective in more general settings such as different time-series models and investment strategies, which will be shown in our future researches. For example, our ES is applicable to fund of funds investment, since it provides a unified scheme to comprehensive portfolio evaluation.

Acknowledgement

References

[1] Cavalcante, R. C., Brasileiro, R. C., Souza, V. L., Nobrega, J. P., & Oliveira, A. L. (2016). Computational intelligence and financial markets: A survey and future directions. Expert Systems with Applications, 55, 194-211.

[2] Chen, L. H., & Huang, L. (2009). Portfolio optimization of equity mutual funds with fuzzy return rates and risks. Expert Systems with Applications, 36(2), 3720-3727.

[3] Chourmouziadis, K., & Chatzoglou, P. D. (2016). An intelligent short term stock trading fuzzy system for assisting investors in portfolio management. Expert Systems with Applications, 43, 298-311.

[4] Dourra, H., & Siy, P. (2002). Investment using technical analysis and fuzzy logic. Fuzzy sets and systems, 127(2), 221-240.

[5] Dymova, L., Sevastianov, P., & Bartosiewicz, P. (2010). A new approach to the rule-base evidential reasoning: Stock trading expert system application. Expert Systems with Applications, 37(8), 5564-5576.

[6] Fang, Y., Lai, K. K., & Wang, S. Y. (2006). Portfolio rebalancing model with transaction costs based on fuzzy decision theory. European Journal of Operational Research, 175(2), 879-893.

[7] Geva, T., & Zahavi, J. (2014). Empirical evaluation of an automated intraday stock recommendation system incorporating both market data and textual news. Decision support systems, 57, 212-223.

[8] Gordon, N. J., Salmond, D. J., & Smith, A. F. (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. In Radar and Signal Processing, IEEE Proceedings F (Vol. 140, No. 2, pp. 107-113). IET.

[9] Goumatianos, N., Christou, I., & Lindgren, P. (2013). Stock selection system: Building long/short portfolios using intraday patterns. Procedia Economics and Finance, 5, 298-307.

[10] Huang, X. (2008). Risk curve and fuzzy portfolio selection. Computers & Mathematics with Applications, 55(6), 1102-1112.

[11] Huang, C. J., Yang, D. X., & Chuang, Y. T. (2008). Application of wrapper approach and composite classifier to the stock trend prediction. Expert Systems with Applications, 34(4), 2870-2878.

[12] Jalota, H., Thakur, M., & Mittal, G. (2017). Modelling and constructing membership function for uncertain portfolio parameters: A credibilistic framework. Expert Systems with Applications, 71, 40-56.

[13] Kitagawa, G. (1996). Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. Journal of computational and graphical statistics, 5(1), 1-25.

[14] Kitagawa, G. (1998). A self-organizing state-space model. Journal of the American Statistical Association, 1203-1215.

[15] Kosaka, M., Mizuno, H., Sasaki, T., Someya, R., & Hamada, N. (1991, October). Applications of fuzzy logic/neural network to securities trading decision support system. In Systems, Man, and Cybernetics, 1991.'Decision Aiding for Complex Systems, Conference Proceedings., 1991 IEEE International Conference on (pp. 1913-1918). IEEE.

[16] Lam, S. S. (2001). A genetic fuzzy expert system for stock market timing. In Evolutionary Computation, 2001. Proceedings of the 2001 Congress on (Vol. 1, pp. 410-417). IEEE.

[17] Li, X., Guo, S., & Yu, L. (2015). Skewness of fuzzy numbers and its applications in portfolio selection. IEEE Transactions on Fuzzy Systems, 23(6), 2135-2143.

[18] Li, J., & Xu, J. (2009). A novel portfolio selection model in a hybrid uncertain environment. Omega, 37(2), 439-449.

[19] Li, J., & Xu, J. (2013). Multi-objective portfolio selection model with fuzzy random returns and a compromise approach-based genetic algorithm. Information Sciences, 220, 507-521.

[20] Lincy, G. R. M., & John, C. J. (2016). A multiple fuzzy inference systems framework for daily stock trading with application to NASDAQ stock exchange. Expert Systems with Applications: An International Journal, 44(C), 13-21.

[21] Liu, S. T. (2011). A fuzzy modeling for fuzzy portfolio optimization. Expert Systems with Applications, 38(11), 13803-13809.

[22] Liu, J., & West, M. (2001). Combined parameter and state estimation in simulation-based filtering. In Sequential Monte Carlo methods in practice (pp. 197-223). Springer New York.

[23] Mamdani, E. H., & Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller. International journal of man-machine studies, 7(1), 1-13.

[24] Markowitz, H. (1952). Portfolio selection. The Journal of Finance, 7(1), 77-91.

[25] Mehlawat, M. K., & Gupta, P. (2014). Fuzzy chance-constrained multiobjective portfolio selection model. IEEE Transactions on Fuzzy Systems, 22(3), 653-671.

[26] Mehlawat, M. K. (2016). Credibilistic mean-entropy models for multi-period portfolio selection with multi-choice aspiration levels. Information Sciences, 345, 9-26.

[27] Michaud, R. O. (1989). The Markowitz optimization enigma: is' optimized'optimal?. Financial Analysts Journal, 45(1), 31-42.

[28] Nakano, M., Takahashi, A., & Takahashi, S. (2017). Generalized Exponential Moving Average (EMA) Model with Particle Filtering and Anomaly Detection. Expert Systems with Applications, 73, 187-200.

[29] Nguyen, T. T., & Gordon-Brown, L. (2012). Constrained fuzzy hierarchical analysis for portfolio selection under higher moments. IEEE Transactions on Fuzzy Systems, 20(4), 666-682.

[30] Nguyen, T. T., Gordon-Brown, L., Khosravi, A., Creighton, D., & Nahavandi, S. (2015). Fuzzy portfolio allocation models through a new risk measure and fuzzy Sharpe ratio. IEEE Transactions on Fuzzy Systems, 23(3), 656-676.

[31] Pai, G. V. (2016). Fuzzy Decision Theory based Metaheuristic Portfolio Optimization and Active Rebalancing using Interval Type-2 Fuzzy Sets. IEEE Transactions on Fuzzy Systems.

[32] Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock market index using fusion of machine learning techniques. Expert Systems with Applications, 42(4), 2162-2172.

[33] Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. Expert Systems with Applications, 42(1), 259-268.

[34] Rather, A. M., Agarwal, A., & Sastry, V. N. (2015). Recurrent neural network and a hybrid model for prediction of stock returns. Expert Systems with Applications, 42(6), 3234-3241.

[35] Sugeno, M., & Kang, G. T. (1988). Structure identification of fuzzy model. Fuzzy sets and systems, 28(1), 15-33.

[36] Simutis, R. (2000). Fuzzy logic based stock trading system. In Computational Intelligence for Financial Engineering, 2000.(CIFEr) Proceedings of the IEEE/IAFE/INFORMS 2000 Conference on (pp. 19-21). IEEE.

[37] Takagi, T., & Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. IEEE transactions on systems, man, and cybernetics, (1), 116-132.

[38] Wang, B., Wang, S., & Watada, J. (2011). Fuzzy-portfolio-selection models with value-at-risk. IEEE Transactions on Fuzzy Systems, 19(4), 758-769.

[39] Yunusoglu, M. G., & Selim, H. (2013). A fuzzy rule based expert system for stock evaluation and portfolio construction: An application to Istanbul Stock Exchange. Expert Systems with Applications, 40(3), 908-920.

[40] Zhou, J., Li, X., & Pedrycz, W. (2016). Mean-Semi-Entropy Models of Fuzzy Portfolio Selection. IEEE Transactions on Fuzzy Systems, 24(6), 1627-1636.