International Journal of Innovative Computing, Information and Control
Volume **12**, Number **1**, February **2016**
ICIC International ©2016 ISSN 1349-4198
pp. 15–40

S-REITS' PERFORMANCE FORECAST USING A SMALL SAMPLE MODEL ASSOCIATING SUPPORT VECTOR MACHINE WITH VECTOR AUTO-REGRESSION MODEL

Xianning Wang[1], Xue Xiao[2] and Zhi Xiao[1,*]

[1]School of Economics and Business Administration
Chongqing University
No. 174, Shazheng Street, Shapingba, Chongqing 400044, P. R. China
*Corresponding author: xiaozhicqu@163.com

[2]Department of Real Estate
School of Design and Environment
National University of Singapore
4 Architecture Drive, 117566, Singapore
xiaoxuesnow.sg@gmail.com

Received June 2015; revised October 2015

Abstract. *Performance forecast in Singapore Real Estate Investment Trusts (S-REITs) is characterized by small sample and nonlinearity. And simultaneous analysis of commonly used performance variables is very important in practice, but few in research. To tackle it, we combined vector auto-regression model-AR(1) with support vector machine. The hybrid least square support vector regression with direct research was designed to make estimation. Then our model was explored to historical data of 22 S-REITs over the different sample ranges to forecast 5 performance variables including dividend per unit, return on equity, market to book value, gross revenue by asset under management and dividend yield. Prediction Error of the Last Year, Mean Percentage Error and Standard Deviation of Prediction Error were utilized to measure prediction accuracy and have validated the feasibility of the proposed method. Further BP Neural Network (BPNN) and Exponential Moving Average (EMA) were chosen to make comparative analysis. Results show that: (1) Performance forecast of S-REITs by new method owns virtues of simple operation and high precision even though the sample size is really small just as only 3, 5 or 7-year data or less; (2) The prediction accuracy of different sectors differs, and using 5 performance variables simultaneously gave us a comprehensive understanding at small computational cost; (3) It is also verified that our method is better than BPNN and EMA due to its feasibility of processing small sample and nonlinearities.*

Keywords: Performance forecast, Singapore Real Estate Investment Trusts, Small sample and nonlinearity, Support vector machine, Vector auto-regression

1. **Introduction.** Real Estate Investment Trusts (REITs) is designed to provide a real estate investment structure similar to the mutual funds in stocks. It is an important tool for the securitization of real estate market. As a recognized financial instrument and an investment channel, it is becoming the focus of capital markets to develop the real estate markets deeply and release liquidity for countries. And Singapore Real Estate Investment Trusts (S-REITs) is the second largest REITs market in Asia. It is very representative to provide practical reference value for REITs' design and operation in emerging regions such as China. Monetary Authority of Singapore (MAS) enacted tax transparency measures for REITs in 2001. Firms must obey the compulsory delivery of minimum 90% dividend, and the dividends to investors are tax-free. Dividend is only one index to measure the S-REITs' performance. However, S-REITs' performance is a direct,

but informative assessment, which can be measured from diverse angles. Comprehensive knowledge about the trend of S-REITs' performance could assist retail or institutional investors making profitable portfolio strategy with little cost in advance. Operators and managers in company could acquire useful tips of management or running problems of REITs. It is also necessary for policymakers to grasp the dynamic progress and extract auxiliary decision information for S-REITs' health development [28].

Along with rapid change of global economic nowadays, the S-REITs' performance forecast has new issues with many properties and less sample size objectively. Compared with the long history in the United States and Australia, Singapore's first REITs listed in July 2002. S-REITs has significant development only during the recent decades in Asia. Therefore, to the present its relative sample range is very small in company level. Many variables could be used to depict it, which also brings pressure in forecasting. Further firm performance affected by complicated internal and external factors does not display a simple linear relationship in time series. Obviously S-REITs' performance forecast is an important and specific problem with characteristic of small sample and complicated nonlinearity [13,14]. Traditional linear forecasting models cannot go through with current data collection. New methods should be explored to ensure credible performance forecast results.

Predictive ability of time series $\mathrm{VAR}(1,k)$ with fewer variables has been recognized by researchers in many fields. SVM has a unique advantage of small samples and being nonlinear in data mining. The organic combination of both is an exciting attempt. This paper aims to combine the first-order vector auto-regression models $\mathrm{VAR}(1,k)$ and Support Vector Machine (SVM) to tackle the small sample and complicated nonlinearity of performance forecast in S-REITs. It can also predict multiple performance variables of S-REITs simultaneously. Firstly, $\mathrm{VAR}(1,k)$ are introduced briefly. The principle of SVM is also explained by figures and formulas. And a novel forecast model will be established with Radial Basis Function (RBF) and Gaussian kernel function as the nonlinear relationship. Secondly, the hybrid least square support vector regression with direct research (LS-SVR-DS) is designed to make estimation of key parameters in the new model. Then, 5 variables including Dividend Per Unit (DPU), Return On Equity (ROE), Market Value to Book Value (MVBV), Gross Revenue by Asset Under Management (GR_AUM) and Dividend Yield (DY) are selected as firm performance. Our new model is explored to historical data of 22 S-REITs over the 11-year sample period (2002-2013) covering more than 70% of the REITs in Singapore. In order to manifest effectiveness the indices of the prediction accuracy (Prediction Error of the Last Year, Mean Percentage Error and Standard Deviation of Prediction Error) are computed and compared subsequently. Meanwhile, since the commonly used linear regression forecast model cannot play a good role, we choose BP Neural Network (BPNN) and the Exponential Moving Average (EMA) as reference to prove the reliability of our method. Finally, several important conclusions are derived.

We highlight and sum up the main properties and contributions of our paper as follows.

(1) A novel combination method was designed, which made mutual complementation of both traditional forecast model ($\mathrm{VAR}(1,k)$) and statistic learning theory (SVM), to tackle the nonlinearity (nonlinear kernel functions) and data characteristics of small samples when making prediction. In the novel model kernel functions such as Gaussian kernel function or k layers perception kernel function could be chosen to describe the nonlinear relationship. The hybrid LS-SVR-DS was designed to make estimation of key parameters in the new model above and show virtues of simple operation, fast convergence speed and high precision, which was manifested in our empirical part.

(2) It is an important problem from the company level to make performance prediction of SREITs. Data from the company level are quite small and limited, which bring difficulty in prediction accuracy. Instead of treating all the REITs in sample range as homogeneous data point, this paper made regression of performance of each REIT with the lagged values and established a specific forecast model by the suggested method, which is likely to be more close to realistic conditions of REITs. On basis of seizing characteristics of individual REIT effective performance forecast is meaningful to enhance the Singaporean competitiveness as a regional center of REITs and feasible to ensure the healthy development of S-REITs through high effective improvement of investment restrictions, management structure, dividend policies [9] and so on.

(3) DPU, ROE, MVBV, GR_AUM and DY are crucial financial indicators in decision making about the portfolio investment. They together reflected the company's market value and investment potential from different angles. And they also provide comprehensive financial information to understand the operational efficiency and future growth trends for investors and shareholders. To forecast the performance with DPU, ROE, MVBV, GR_AUM and DY systematically is also one advantage of the new model. After our suggested model was explored to relative historical data of 22 Singapore REITs over the 11-year sample period (2002-2013), results have shown that our method was tested to be good at the 5 variables' forecasting at small computational cost.

(4) Comparative analysis has validated the feasibility of the proposed method. BP Neural Network (BPNN) and Exponential Moving Average (EMA) are popular time-series forecasting methods. They were chosen as reference and comparing models. 3 indices – Prediction Error of the Last Year ($\boldsymbol{e_T}$), Mean Percentage Error ($\boldsymbol{mpe}$) and Standard Deviation of Prediction Error ($\boldsymbol{sde}$) were taken as 5 variables' performance accordance measure of each S-REIT. Calculation results validate the superiority of the proposed method. The comparison has also been used to increase the significance of the presented results here.

2. Related Work.

In this section, we review three lines of research work that are related to our work, and discuss the novelty of our work from them.

2.1. Performance forecast of S-REITs.

Literature is very fruitful on performance such as corporate governance [28], corporate shareholder revenue [8], returns and risk factors or characteristics [2,8], capital structure [24], and dividend taxes [12]. These achievements give us the following enlightenment.

(1) Most articles consider only one main indicator, and do not use more than one index to comprehensively reflect corporate performance of S-REITs. There are many methods to measure firm performance such as using stock return and Tobin's Q [17], absolute and relative stock performance to measure market performance [12,14], which are mostly market-based. The accounting measures, unlike stock performances, tend to be better performance measures as they are less influenced by external factors [35]. We choose five

Table 1. 5 performance variables in S-REITs

Name	Code	Definition
Dividend Per Unit	DPU	Distribution Per Unit or Per Staple Security in SGD
Return On Equity	ROE	Measuring profitability of REIT
Market Value to Book Value	MVBV	Measuring market perception on REITs
Scaled Gross Revenue	GR_AUM	Scale Gross Revenue (GR) by AUM
Dividend Yield	DY	Ratio of dividend per share and price per share

accounting-based performance measures comprehensively (DPU, ROE, MVBV, GR_AUM and DY) of S-REITs (shown in Table 1).

(2) REITs market is a complex, non-linear dynamical system. Its performance forecasting is characterized by data intensity, noise, non-stationary, unstructured nature, high degree of uncertainty, and hidden relationships. Many factors interact in S-REITs including political events, general economic conditions, and traders' expectations. Therefore, predicting performance movements is quite difficult. The paper focuses on firm performance of REITs listed on Singapore Exchange market (SGX). According to academic investigations and the following reasons in data collection, we consider that movements in S-REITs behave in a highly non-linear, dynamic manner and it is a small sample problem.

1) Each S-REITs firm has their own characteristics. We cannot treat them as the same sample points or with the same time interval. For example, corresponding sectors, corporate governance structure and operation mode of REITs are extinguished significantly. Just a piled-up time series is unreasonable to enlarge sample size to make estimation. However, there is a small sample when we concentrate on an individual firm.

2) S-REITs have been in different development stages and begun to develop formally since the late 1990s. The first listed REIT is dated from 2002, which happened also very late in some other Asian countries. As we can see, IPO years are different among S-REITs from 2002 to 2010. For example, Capita Mall Trust stepped to IPO in 2002 but Lippo Malls Indo Retail Trust in 2010, which is an unbalanced panel sample dataset. It is impossible to a same time interval for all the REITs, so a targeted forecast model on the basis of an individual firm is absolutely necessary.

3) As to January 2014, there are totally 31 REITs and stapled securities listed on SGX (excluding 5 existing business trusts). Among them 6 REITs went IPO in 2013 or 2014 and their financial or annual reports were not given out so that they cannot be chosen as research objectives[1]. We acquired 22 REITs, which covers more than 70% of the REITs in Singapore (see in Table 2).

(3) Research on tools of S-REITs performance forecast is very little, especially on its particular small sample characteristics and non-linear nature, which is due to its short time series available of company level data. VAR$(1,k)$ own good performance especially in fields of forecasting (discussed in Sections 2.2 and 3.1) and the classic statistical learning of SVM has advantages on dealing with small sample and nonlinear data patterns in prior literature (discussed in Sections 2.3 and 3.2). This paper aims to combine VAR$(1,k)$ with SVM together.

2.2. **Literature analysis of VAR.** VAR models, firstly introduced by Christopher Sims in 1980, present good performance especially in fields of forecasting. Compared to structural models in classic econometrics (such as simultaneous equation models), VAR models are unstructured. It is primarily through actual data rather than economic theory to reveal the dynamic structure of economic system [16], so information originated from data sets determines the success of modeling. Further, VAR models, including VAR$(1,k)$, are often limited by following conditions of data sets. (1) Take stationary time series and volumes of data as the premise. (2) The model specification is often traditional linear relationship or convertible linear relationships. (3) OLS is mainly the estimation method. Therefore, due to the existence of the non-stationary data or non-linear time series data with small sample in practice, original models could not show its universality.

[1]6 newly listed REITs and their IPO year: APLETREE GREATER CHINA COMMERCIAL TRUST (2013), OUE HOSPITALITY TRUST (2013), SOILBUILD BUS.SPACE REIT (2013), SPH REIT (2013), VIVA INDUSTRIAL TRUST (2013), OUE COMMERCIAL TRUST (2014).

Table 2. Data time range of detailed sample list of 22 REITs

No.	Name of REIT	Code	Data range	Years	IPO year	Sector
1	AIMS AMP CAP.INDL.REIT	AIMS_A	2008-2013	6	2007	Industrial
2	ASCENDAS HOSPITALITY TST.	ASCD_H	2013	1	2012	Hospitality
3	ASCENDAS REAL ESTATE IT.	ASCD	2003-2013	11	2002	Industrial/Mixed
4	ASCOTT RESIDENCE TRUST	ASCT_R	2006-2012	7	2006	Service Residence
5	CACHE LOGISTICS TRUST	CACH_L	2010-2012	3	2010	Industrial
6	CAMBRIDGE INDL.TRUST	CAMB_L	2006-2012	6	2006	Industrial
7	CAPITACOMMERCIAL TRUST	CCT	2004-2012	8	2004	Commercial
8	CAPITAMALL TRUST	CMT	2002-2012	10	2002	Commercial
9	CAPITARETAIL CHINA TRUST	CRCT	2007-2012	5	2006	Commercial
10	CDL HOSPITALITY TRUST	CDL_H	2006-2012	7	2006	Hospitality
11	FAR EAST HOSPITALITY TRUST	FAR_H	2013	1	2012	Hospitality
12	FIRST REIT	FIRST	2007-2012	5	2006	Healthcare
13	FORTUNE RLST.INV.TST.	FORT	2003-2012	9	2003	Commercial
14	FRASERS CENTREPOINT TST.	FRAS_CE	2007-2012	6	2006	Commercial
15	K-REIT ASIA	KREIT	2006-2012	7	2006	Commercial
16	LIPPO MALLS INDO.RET. TST.	LIPO_R	2008-2012	5	2007	Commercial
17	MAPLETREE INDUSTIAL TST.	MAPT_I	2010-2012	3	2010	Industrial
18	MAPLETREE LOGIST.TRUST	MAPT_L	2006-2013	8	2005	Industrial
19	PARKWAY LIFE REIT.TST.	PAKW	2007-2012	5	2007	Healthcare
20	SABANA SHARI'AH CMN. RLST.IT.	SABA	2011-2012	2	2010	Industrial
21	STARHILL GLOBAL REIT	STAR	2005-2012	8	2005	Commercial
22	SUNTEC RLST.IT	SUNT	2006-2012	7	2004	Commercial

Under the frame of large sample and linearity, VAR$(1, k)$ have been expanded widely in recent years. Structural VAR was proposed to explain variables contemporaneous relationship in the random error term [23]; Exogenous variables are incorporated by Corrected VAR to explain complex economic problems. To tackle lots of non-stationary economic variables, concepts about cointegration were also introduced to get the Vector Error Correction Model and Structural Vector Error Correction Model. Bayesian VAR or other error correction models also play a great role [6,27]. In the aspect of application, VAR$(1, k)$ and other VAR models are often used to predict time series of interconnected system and analyze dynamic impact of random disturbance on the variables system to explain the influence of various shocks. These mainly focus on business cycle [43], monetary policy [5], economic growth and environmental factors [45], specific market forecast [37] and so on.

It is found that model specifications above are based on traditional or convertible linear relationships. There is little attention on some complicated ones that may exist in real issues [49]. Although the linear modeling is simple and theoretical results are relatively perfect, Kendall pointed out that it seems more reasonable to use "nonlinear" instead of "linear" when considering the setting of time series model with economic data [18,49]. In fact, VAR models are mainly applied for economic forecasting, but confront the obstacle of methodology in other fields about the economic structure factors analysis, which leads to nonlinearity to some extent. In addition, few studies take the data set with small sample (which means a small information set) into consideration under the framework of forecasting. [38] points that using a small information set may imply that the VAR is mis-specified and may produce erroneous results such as the 'price puzzle' discovered by [44]. In addition, another problem is that sample size is very small or does not meet requirements of classic econometrics. For example, typically Corrected VAR are fitted to low frequency data on time intervals of daily, monthly or yearly data [38]; model estimation cannot be achieved by OLS when there are only 5 or much less numerical values in time series ($T = 3$ or 4). Similarly only using traditional VAR$(1, k)$ cannot perform well in conditions of small sample and unconvertible nonlinearity in the forecast of realistic economic issues. Thus, we intend to improve VAR by SVM.

2.3. **Literature analysis of SVM.** SVM, introduced firstly by Vapnik et al. and a popularized tool of statistic learning theory, has advantages on dealing with small sample and nonlinear data patterns [48]. Its diverse kernel function that presents types of linear and nonlinear relationship could expand specifications of traditional linear models. Its learning power with "Minimization of Structural Risk" allowed us to predict with not too many data. Forecasting on the basis of SVM has been extended to many fields. For example, in the economic forecasting many fruits are shown on financial time series forecasting [26], credit risk [14] and business default forecasting [19], management fraud and fake websites [1], warranty claim forecasting [51], customer churn prediction [11] and so on. Main literature about improved SVM can be roughly divided into classified ability, the parameter optimization algorithms, feature selection and how to choose the kernel function and other aspects. Most of mentioned achievements have further proved the feasibility of SVM on processing complex nonlinear and small samples.

(1) To enhance robustness and accuracy of the classification results by alleviating potential noise on the input data, robust and multiple output-SVR aroused scholars' interests. Pérez-Cruz et al. (2002) proposed multiple output SVR [36] and successfully applied it to complex biomedical problems, nonlinear channel estimation [42] and remote sensing biophysical parameter estimation [47]. Yi et al. (2011) have studied the structure selection for DAG-SVM based on misclassification cost minimization, which is a big improvement for multi-class SVM [55]. For Yan et al. (2013), the Fuzzy Support Vector Machines (FSVMs) were established to solve the over-fitting problems due to the fact that SVM is sensitive to outliers or noises [53]. Xiong et al. (2014) proposed a multiple-output support vector regression with a firefly algorithm for interval-valued stock price index forecasting [52].

(2) To date, considering the complex nonlinear mapping, a large number of evolutionary algorithms such as the Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Firefly Algorithm (FA) and Gravitational Search Algorithm (GSA) have been employed to optimize the parameters of SVM.

Wu et al. (2007) showed us a real-valued genetic algorithm to optimize the parameter SVM for predicting bankruptcy [50]. Huang et al. (2007) pointed that combining GA with SVM classifier is a promising data mining method, in which the proposed hybrid GA-SVM strategy can simultaneously perform feature selection task and model parameters optimization [21]. Pourbasheer et al. (2009) applied SVM to the Quantitative Structure-Activity Relationship (QSAR) for BK-channel activators, in which they adopted Genetic Algorithm (GA) to select variables that resulted in the best-fitted for models [39]. Based on GA and SVM, Li et al. (2011) designed an effective feature selection method for hyperspectral image classification [30].

Bazi and Melgani (2007) put forward a method that two criteria are empirical and structural expressions of the generalization capability of the resulting semisupervised PSO-SVM regression system [4]. Huang and Dun (2008) proposed a novel PSO–SVM model to improve the classification accuracy with a small and appropriate feature subset [22]. This optimization mechanism combined the discrete PSO with the continuous-valued PSO to simultaneously optimize the input feature subset selection and the SVM kernel parameter setting [32]. Lin et al. employed several public datasets to calculate the classification accuracy rate in order to evaluate the developed PSOSVM approach [31]. Experimental results demonstrate that the classification accuracy rates of the developed approach surpass those of grid search and many other approaches, and that the developed PSO-SVM approach has a similar result to GA-SVM. Nieto et al. (2015) described a hybrid PSO-SVM-based model for the prediction of the remaining useful life of aircraft engines [33].

Recently, Kazem et al. (2013) and Xiong et al. (2014) proposed a Firefly Algorithm (FA)-based approach, built on the established MSVR, for determining the parameters of MSVR [25,52]. To enhance parameter optimization and feature selection influence. Li et al. (2015) hybridized chaotic search and gravitational search algorithm (based on the law of Newtonian gravity) with SVM and present a new chaos embedded GSA-SVM hybrid system [29].

(3) Kernel functions are important tools to describe the complex nonlinear relationship between samples in high dimension space. Unsuitably chosen kernel functions or hyperparameter settings may lead to significantly poor performance. Most researchers use trial-and-error to choose proper values for the hyperparameters and adopt multiple-kernels to deal with these problems. Multiple Kernel Learning (MKL) aims at simultaneously learning a kernel and the associated predictor in supervised learning settings. Rakotomamonjy et al. (2008) solved a standard SVM optimization problem, where the kernel is defined as a linear combination of multiple kernels, and proposed a SimpleMKL algorithm to solve this MKL problem [40]. Szafranski et al. (2010) thought that the SVM's kernel is chosen prior to learning and lack of flexibility. Further they used "Composite Kernel Learning" to address the situation where distinct components give rise to a group structure among kernels [46]. Yeh et al. (2011) had developed a two-stage multiple-kernel learning algorithm by incorporating sequential minimal optimization and the gradient projection method, which was applied to stock price forecasting [54]. Rakotomamonjy and Chanda (2014) proposed a method to address the problem of scaling $l(p)$-norm multiple kernel for large learning tasks using low-rank kernel approximations, which takes advantage of the low-rank kernel approximations and a proximal gradient algorithm for solving that optimization problem [41].

As we can conclude, SVM owns advantages of dealing with small sample and nonlinear data patterns in [7]. These researches enlighten us the SVM models could provide more accurate estimates than the empirical equations and could tackle the complicated nonlinear relationship [3,10,16]. Therefore, it inspired us that combination of VAR$(1,k)$ and SVM may be a reasonable attempt in the aspect of more accurate forecasting of S-REITs. Considering the less data and big computation amount of above algorithms, we would use "Direct Research Algorithm" to find the optimal parameters and adopt the existing natural kernels to fit the nonlinearity in our paper.

3. **Preliminary.** In this section, we review two lines of models that are related to our new method, and explain why we choose both of them.

3.1. **VAR models.** As one main aspect of modern time series analysis, VAR(p) with k variables are shown as

$$Y_t = \mu + A_1 Y_{t-1} + \cdots + A_p Y_{t-p} + \varepsilon_t \quad t = 1, 2, \cdots, T \tag{1}$$

$$Y_{t-1} = \begin{bmatrix} Y_{1t-i} \\ Y_{2t-i} \\ \vdots \\ Y_{kt-i} \end{bmatrix} \quad i = 1, 2, \cdots, p \tag{2}$$

$$A_j = \begin{bmatrix} a_{11,j} & a_{12,j} & \cdots & a_{1k,j} \\ a_{21,j} & a_{22,j} & \cdots & a_{2k,j} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1,j} & a_{k2,j} & \cdots & a_{kk,j} \end{bmatrix} \quad j = 1, 2, \cdots, p \tag{3}$$

$$\mu = (\mu_1, \cdots, \mu_k)' \tag{4}$$

$$\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \cdots, \varepsilon_{kt})' \tag{5}$$

Y_t is a k dimensions endogenous variable vector, p is the lagged difference and T is the sample size; $A_1, \cdots, A_p$ are $k*k$ coefficient matrices; $\varepsilon_t \sim N(0, \Sigma)$ is a k dimensions disturbance term vector, in which they can be contemporaneous correlation but not related to the lag values or variables in the right of the first equation above. Σ is a covariance matrix of ε_t and a positive definite matrix.

VAR(1) with k variables is a special case [20] shown as

$$Y_t = \mu + A_1 Y_{t-1} + \varepsilon_t \quad t = 1, 2, \cdots, T \tag{6}$$

$$Y_{t-1} = \begin{bmatrix} Y_{1t-1} \\ Y_{2t-1} \\ \vdots \\ Y_{kt-1} \end{bmatrix} \tag{7}$$

$$A_1 = \begin{bmatrix} a_{11,1} & a_{12,1} & \cdots & a_{1k,1} \\ a_{21,1} & a_{22,1} & \cdots & a_{2k,1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1,1} & a_{k2,1} & \cdots & a_{kk,1} \end{bmatrix} \tag{8}$$

$$\mu = (\mu_1, \cdots, \mu_k)', \quad \varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \cdots, \varepsilon_{kt})'$$

VAR modeling need not take strict economic theories as the basis and only two quantities should be clarified in the process: one is the number of variables k that should be added to the VAR model; the other is selecting the maximum (reasonable) number of autoregressive lag order p to reflect the interactional relationship between variables and ensure that the random error term is white noise. Here the reasons why we choose VAR$(1, k)$ as targeted objects are listed below.

(1) It is the simplest and foundational form of VAR models, whose improvement and correction is meaningful for other complex models [20].

(2) Unrestricted VAR models are prone to over-fitting the data as more variables are added to the models; sometimes simplicity is helpful for a better explanation.

(3) Selecting the maximum (reasonable) number of autoregressive lag order p is of the utmost importance and should give consideration to both retaining as much variables information as possible and degrees of freedom in parameter estimation. In empirical research p mostly takes values of "1" or "2", so we set $p = 1$ to acquire interesting characteristics.

3.2. **SVM and LS-SVR.** Fundamental principles and corresponding simple explanation of SVM are displayed in the rectangular plane coordinate system by Figure 1. In Figure 1, a series of data points, $\{x_i, y_i\}$, $i = 1, 2, \cdots, m$, $x_i \in R^n$ for two different classes of data is shown. The blue squares represent the negative class and the white circles represent the positive class (Figure 1a). SVM firstly transfers the data points in to a multi-dimension space (Figure 1b). Next a linear boundary or a hyper plane is placed between the two classes of points to get the maximized margin which is represented by the dotted line (Figure 1b).

SVM is capable of acquiring the hyper plane in the middle of space of two classes so that distances from the points of two classes to the plane are the shortest. The distance from the margin of positive class to the plane and from the margin of negative class to the plane are marked as $+\varepsilon$ and $-\varepsilon$ by SVM respectively (Figure 1b). Correspondingly, SVM marks the distance from data point i of positive class to the margin of the positive as ξ_i^* and the distance from data point i of negative class to the margin of the negative

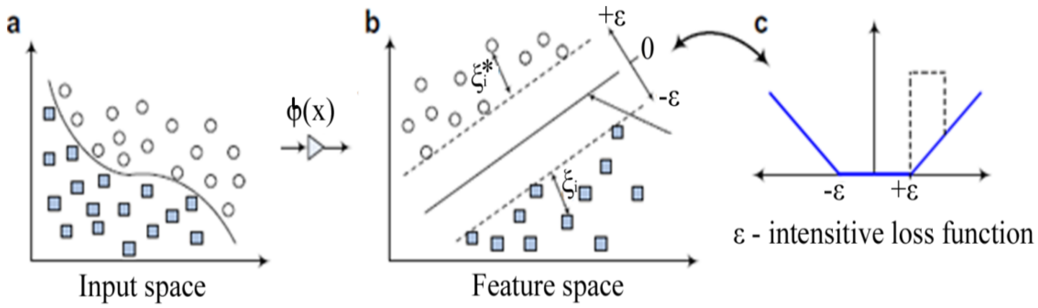

FIGURE 1. Explanation of SVM by the rectangular plane coordinate system

as ξ_i (Figure 1b and Figure 1c). And the nearest data points that are used to define the margins of classes are called support vectors (Figure 1b).

When the support vectors have been selected, the rest of the feature set is not required, as the support vectors contain all the information-based needed to define the classifier. If the data point falls between the margins $[-\varepsilon, +\varepsilon]$, the loss or error penalty c is zero. If not, according to distance from the data point to the nearer side of margin, c is set as ξ_i^* or ξ_i (Figure 1b and Figure 1c). According to the geometry, the geometrical margin is found to be $||w||^2$.

It is obvious that support vectors are critical because they are the data points only used to form the hyper plane which yields the decision function. Thus, the data points far away from the hyper plane can be redundant information, and removing them does not change the hyper plane as well as the decision function. This is one of the stronger points SVM has over other traditional data mining techniques [32].

Searching the optimal hyper plane can be transformed to finding a solution to an optimization problem by incorporating the error penalty c and the noise with slack variables x_i, which is written as the equation below.

$$Min_{w,\xi_i^*,\xi_i} R\left(w, \xi_i^*, \xi_i\right) = \frac{1}{2}||w||^2 + c\sum_{i=1}^{N}\left(\xi_i^* + \xi_i\right) \tag{9}$$

$$y_i - w^T\phi(x_i) - b \le \xi + \xi_i^* \tag{10}$$

$$-y_i + w^T\phi(x_i) + b \le \xi + \xi_i \tag{11}$$

$$\xi \ge 0, \quad \xi_i^* \ge 0, \quad i = 1, 2, \cdots, m$$

Solutions to this optimization problem are the so-called support vectors. Given the known support vectors, we can also calculate the corresponding value of w and further apply them to an optimized classification function or a non-linear decision function by the equations as below.

$$w = \sum_{i=1}^{N}\left(\beta_i^* - \beta_i\right)\phi(x_i) \tag{12}$$

$$f(x) = sign\left\{\sum_{i=1}^{m}\alpha_i y_i K(x \cdot x_i) + b\right\} \tag{13}$$

where β_i^*, β_i can be worked out with a quadratic programming and Lagrangian multipliers; α_i is the Lagrangian multiplier; b is a scalar.

Finally, we make the transformation of Equation (13) as

$$f(x) = \sum_{i=1}^{N}\left(\beta_i^* - \beta_i\right)K(x, x_i) + b \tag{14}$$

$K(x, x_i)$ is called the kernel function of SVM. And it is an inner product of x_i in the feature space. The kernel function $K(x, x_i)$ can be chosen as:

(1) $K(x, x_i) = x_i^T x$, which is called as the linear kernel and leads to a linear regression;

(2) $K(x, x_i) = [(x_i \cdot x) + 1]^d$, which is called as the polynomial kernel function;

(3) $K(x, x_i) = S[v(x_i \cdot x) + c]$, which is called as two layers perception kernel function;

(4) $K(x, x_i) = \exp(-||x - x_i||^2/2\sigma^2)$, which is the Radial Basis Function (RBF) called Gaussian kernel function with an extra parameter and leads to a nonlinear regression fitting. $|| \, ||$ is the Euclidean norms.

Basic ideas of SVM to solve regression estimation can be described as follows. First of all, input variables are mapped to high-dimensional feature space by a nonlinear mapping $\phi(x)$. Then construct the regression estimate function in the high-dimensional feature space. Finally with introduction of a kernel function, SVM could avoid complex calculation.

Given a training set $\{x_i, y_i\}$, $i = 1, 2, \cdots, m$, $x_i \in R^n$ is the input variable and $y_i \in R$ is the response variable. SVM is used to construct a regression function taking the following form:

$$y = f(x, w) = w^T \phi(x) + b \tag{15}$$

$\phi(x)$ is the nonlinear mapping that maps the input data into a higher dimensional feature space. Then LS-SVR involves equality constraints instead of inequality ones and works with the least squares loss function rather than ε-insensitivity loss. LS-SVR formulates the optimization problem as follows:

$$Min_{w,b,e} y = \frac{1}{2} w^T w + \frac{1}{2} \gamma e^T e \tag{16}$$

subject to the equality constraints with $e = (e_1, e_2, \cdots, e_m)^T$, where

$$y_i = w^T \phi(x_i) + b + e_i, \quad i = 1, 2, \cdots, m \tag{17}$$

$\frac{1}{2}\gamma e^T e$ in (16) is the squared loss measuring the regression function fitness. $\frac{1}{2} w^T w$ in (16) controls the function smoothness to avoid over-fitting problem. $\gamma \in R^+$ is the tuning parameter and the trade-off parameter between a smoothing solution and training error. This corresponds to a form of ridge regression.

Introduce the Lagrange multipliers $\alpha_i = (\alpha_1, \alpha_2, \cdots, \alpha_m)^T$, and the Lagrangian is

$$L(w, b, e, \alpha) = J(w, e) + \sum_{i=1}^{m} \alpha_i \left[y_i - \left(w^T \phi(x_i) + b + e_i \right) \right] \tag{18}$$

The optimality conditions are

$$\begin{cases} \dfrac{\partial L}{\partial w} := 0 \Rightarrow w = \sum\limits_{i=1}^{m} \alpha_i \phi(x_i) \\ \dfrac{\partial L}{\partial b} := 0 \Rightarrow \sum\limits_{i=1}^{m} \alpha_i = 0 \\ \dfrac{\partial L}{\partial e_i} := 0 \Rightarrow \alpha_i = \gamma e_i \\ \dfrac{\partial L}{\partial \alpha_i} := 0 \Rightarrow w^T \phi(x_i) + b + e_i = 0 \end{cases} \qquad i = 1, 2, \cdots, m \tag{19}$$

After eliminating e and w, we could have the solution by the following linear equations.

$$\begin{bmatrix} 0 & l^T \\ l & \Omega + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{20}$$

I is the identity matrix, $l = (1, 1, \cdots, 1)^T$, $y = (y_1, y_2, \cdots, y_m)^T$, Ω is the kernel matrix with element

$$\Omega_{jk} = \phi(x_j)\phi(x_k) = K(x_j, x_k) \quad j, k = 1, 2, \cdots, m \tag{21}$$

Set $A = \Omega + \gamma^{-1} I$. Since A is a symmetric and positive-definite matrix for all $\gamma > 0$, A^{-1} exists. Solving the linear equations of (20), we obtain the solution

$$\begin{aligned} \alpha &= A^{-1}(y - b_1) \\ b &= \frac{l^T A^{-1} y}{l^T A^{-1} 1} \end{aligned} \tag{22}$$

Substituting w in Equation (15) with the first equation of (19) and using (21), we can obtain

$$f(x|w) = y(x) = \sum_{i=1}^{m} \alpha_i K(x, x_i) + b \tag{23}$$

α_i and b are given by (22), which depends on γ and the kernel matrix. In this paper the solution of (17) is the function of two parameters γ and σ. In order to optimally determine the values of parameter, a hybrid LS-SVR-DS model is introduced.

4. **Model Design.**

4.1. **Novel models.**

In order to achieve our goal, we try to modify VAR$(1, k)$ with the Least Squares Support Vector Regression (LS-SVR) with Direct Search (DS) so that a new small sample model is put forward and can be applied into specific issues with simple operation, fast convergence speed and high precision.

$$g\left(Y_{i(t^*-1)}|w\right) = g^*\left(Y_{i(t^*-1)}\right) = \mu_i + \sum_{t=1}^{T} \alpha_t K\left(Y_{i(t^*-1)}, Y_{i(t-1)}\right) + \varepsilon_{it} \tag{24}$$

$i = 1, \cdots, k$, where $Y_{i(t^*-1)}$ present the variable value of Y_i at the time of $t^* - 1$, namely, the input values of SVM regression and $Y_{i(t-1)}$ present the variable value of Y_i at the time of t, which is given in the condition of $i = 1, \cdots, k$.

$K\left(Y_{i(t^*-1)}, Y_{i(t-1)}\right)$ is called the kernel function of SVM. And it is an inner product of x_i in the feature space. The kernel function $K\left(Y_{i(t^*-1)}, Y_{i(t-1)}\right)$ can be chosen as

(1) The linear kernel that leads to a linear regression $K\left(Y_{i(t^*-1)}, Y_{i(t-1)}\right) = {Y_{i(t-1)}}^T Y_{i(t^*-1)}$;

(2) The polynomial kernel function $K\left(Y_{i(t^*-1)}, Y_{i(t-1)}\right) = \left[\left(Y_{i(t^*-1)} \cdot Y_{i(t-1)}\right) + 1\right]^d$;

(3) Two layers perception kernel function $K\left(Y_{i(t^*-1)}, Y_{i(t-1)}\right) = S\left[v\left(Y_{i(t^*-1)} \cdot Y_{i(t-1)}\right) + c\right]$;

(4) The Radial Basis Function (RBF) or the Gaussian kernel function with an extra parameter that leads to a nonlinear regression fitting $K\left(Y_{i(t^*-1)}, Y_{i(t-1)}\right) = \exp\Big(-\left\|Y_{i(t^*-1)} - Y_{i(t-1)}\right\|^2 / 2\sigma^2\Big)$.

Suppose that $k = 1$,

$$g\left(Y_{1(t^*-1)}|w\right) = g^*\left(Y_{1(t^*-1)}\right) = \mu_1 + \sum_{t=1}^{T} \alpha_t K\left(Y_{1(t^*-1)}, Y_{1(t-1)}\right) + \varepsilon_{1t} \tag{25}$$

It is similar to the AR(1) model. The new model is also presented in the form of vectors below

$$G\left(Y_{t^*-1}|w\right) = G^*\left(Y_{t^*-1}\right) = \mu + A_2 K\left(Y_{t^*-1}, Y_{t-1}\right) + \varepsilon_t \tag{26}$$

$$Y_{t^*-1} = \begin{bmatrix} Y_{1(t^*-1)} \\ Y_{2(t^*-1)} \\ \vdots \\ Y_{k(t^*-1)} \end{bmatrix} \tag{27}$$

$$A_2 = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1T} \\ a_{21} & a_{22} & \cdots & a_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{iT} \end{bmatrix} \tag{28}$$

$$\mu = (\mu_1, \cdots, \mu_k)' \tag{29}$$

$$\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \cdots, \varepsilon_{kt})' \tag{30}$$

A_2 is strictly different from A_1 in Equation (8). The new model above can be considered as a simultaneous equation model is composed of k equations, in which each equation is viewed as an independent one. The commonly used OLS regression can be used for each equation one by one to get the consistent parameter estimation.

As we can see, the traditional linear relationship could be replaced by linear or nonlinear relationships that are represented by "kernels function of SVM". For example, if we choose the $K\left(Y_{i(t^*-1)}, Y_{i(t-1)}\right) = \exp\left(-\left\|Y_{i(t^*-1)} - Y_{i(t-1)}\right\|^2 / 2\sigma^2\right)$, the Radial Basis Function or the Gaussian kernel function, the model will be considered as a nonlinear regression fitting. Obviously this method gives us a big possibility to describe the nonlinear characteristics in practice, which is one advantage that motives great interest of our research.

4.2. **Estimation methods.** The generalization ability of SVR depends on adequately setting parameters, such as the penalty coefficient and kernel parameters, which is crucial to obtain good performance in handling forecasting task. The hybrid LS-SVR-DS can perform well with small sample and it incorporates "direct search" coined by Hooke and Jeeves in 1935 into traditional LS-SVR method. By this way it can be implemented easily and quickly by computer programming. And it is widely applied into intelligent computing and predicting and achieves good effects with simple operation, fast convergence speed and high precision.

Consider the solution of LS-SVR in (23). To emphasize its dependence on parameters σ and γ, we write the solution as

$$g^*\left(Y_{t^*-1}; \sigma, \gamma\right) = \sum_{i=1}^{T} \alpha_i(\sigma, \gamma) K\left(Y_{i(t^*-1)}, Y_{i(t-1)}\right) + b(\sigma, \gamma) \tag{31}$$

We want to optimally determine the values of parameters. Here the optimality means that the parameters minimize the average of squared errors. That is,

$$\begin{aligned} \underset{\sigma,\gamma}{Min}\, Z(\sigma, \gamma) &= \frac{1}{T} \sum_{j=1}^{T} \left[g^*\left(Y_{i(t^*-1)}\right) - g^*\left(Y_{i(t^*-1)}; \sigma, \gamma\right)\right]^2 \\ &= \frac{1}{T} \sum_{j=1}^{T} \left[g^*\left(Y_{i(t^*-1)}\right) - \sum_{i=1}^{T} \alpha_i(\sigma, \gamma) K\left(Y_{i(t^*-1)}, Y_{i(t-1)}\right) - b(\sigma, \gamma)\right]^2 \end{aligned} \tag{32}$$

$\alpha_i(\sigma, \gamma)$ and $b(\sigma, \gamma)$ are given by (22). The objective function Z is an implicit function of σ and γ without an explicit close form on σ and γ. However, function values can be easily evaluated. This motivates our efforts on integrating a direct search method into LS-SVR. To illustrate the search procedure, we give the following algorithm:

Step 1. Initialize a search point $B_0 = (\sigma_0, \gamma_0)$ and $k = 1$.

Step 2. Let $B_1 = (\sigma_0 + \lambda_\sigma, \gamma_0 + \lambda_\gamma)$ be an alternative point, where λ_σ and λ_γ are random step sizes generated from the uniform distribution on $(0, 1)$.

Step 3. Calculate $Z(\sigma_0, \gamma_0)$ and $Z(\sigma_0 + \lambda_\sigma, \gamma_0 + \lambda_\gamma)$ using (16).

Step 4. Update σ_0 by $\sigma_0 + \lambda_\sigma$ and γ_0 by $\gamma_0 + \lambda_\gamma$, if $Z(\sigma_0 + \lambda_\sigma, \gamma_0 + \lambda_\gamma) \le Z(\sigma_0, \gamma_0)$. Otherwise, $\boldsymbol{\sigma_0} = \sigma_0$ and $\boldsymbol{\gamma_0} = \gamma_0$.

Step 5. If $Z(\sigma_0, \gamma_0) \le \varepsilon$ or $k \ge N$, stop the iteration. Otherwise, let $\boldsymbol{k} \ge k + 1$ and go to Step 2. The iteration stops either when a desired accuracy is achieved or when the number of iterations exceeds a pre-specified limit N. After the algorithm stops, we obtain the "optimal" pair of (σ_0, γ_0) for SVR model, which minimizes the training error. The resulting SVR model describes most information about the original data.

Parameters in LS-SVR and the hybrid LS-SVR-DS do not represent marginal contribution, but they can make analysis about effect of the change of independent variables on the dependent variables (similar to the famous the impulse response or variance decomposition analysis) and predict the future values of the dependent variable when the independent value is given. LS-SVR and the hybrid LS-SVR-DS based on learning theory and the actual data information could overcome the incompetence to accurately predict economic variables with traditional VAR models under the structure factors.

5. Experiments Methodology.

5.1. **Empirical models.** In the empirical study there are two variables of Y and $k = 5$, that are presented respectively by the subscript capital letter DPU, ROE, MVBV, GR_AUM, DY. We rewrite the new model in Section 4.1 by

$$Y_{DPUt^*} = g^*\left(Y_{DPU(t^*-1)}\right) = \mu_{DPU} + \sum_{t=1}^{T} \alpha_t K\left(Y_{DPU(t^*-1)}, Y_{DPU(t-1)}\right) + \varepsilon_{DPUt} \tag{33}$$

$$Y_{ROEt^*} = g^*\left(Y_{ROE(t^*-1)}\right) = \mu_{ROE} + \sum_{t=1}^{T} \alpha_t K\left(Y_{ROE(t^*-1)}, Y_{ROE(t-1)}\right) + \varepsilon_{ROEt} \tag{34}$$

$$\begin{aligned} Y_{MVBVt^*} &= g^*\left(Y_{MVBV(t^*-1)}\right) \\ &= \mu_{MVBV} + \sum_{t=1}^{T} \alpha_t K\left(Y_{MVBV(t^*-1)}, Y_{MVBV(t-1)}\right) + \varepsilon_{MVBVt} \end{aligned} \tag{35}$$

$$\begin{aligned} Y_{GR\_AUMt^*} &= g^*\left(Y_{GR\_AUM(t^*-1)}\right) \\ &= \mu_{GR} + \sum_{t=1}^{T} \alpha_t K\left(Y_{GR\_AUM(t^*-1)}, Y_{GR\_AUM(t-1)}\right) + \varepsilon_{GRt} \end{aligned} \tag{36}$$

$$Y_{DYt^*} = g^*\left(Y_{DY(t^*-1)}\right) = \mu_{DY} + \sum_{t=1}^{T} \alpha_t K\left(Y_{DY(t^*-1)}, Y_{DY(t-1)}\right) + \varepsilon_{DYt} \tag{37}$$

They can be presented in the form of vectors below

$$G\left(Y_{t^*-1}|w\right) = G^*\left(Y_{t^*-1}\right) = \mu + DK\left(Y_{t^*-1}, Y_{t-1}\right) + \varepsilon_t \tag{38}$$

$$Y_{t^*-1} = \begin{bmatrix} Y_{DPUt^*} \\ Y_{ROEt^*} \\ Y_{MVBVt^*} \\ Y_{GR\_AUMt^*} \\ Y_{DYt^*} \end{bmatrix} \tag{39}$$

$$D = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1T} \\ a_{21} & a_{22} & \cdots & a_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{iT} \end{bmatrix} \tag{40}$$

$$\mu = (\mu_{DPU}, \mu_{ROE}, \mu_{MVBV}, \mu_{GR\_AUM}, \mu_{DY})' \tag{41}$$

$$\varepsilon_t = (\varepsilon_{DPUt}, \varepsilon_{ROEt}, \varepsilon_{MVBVt}, \varepsilon_{GR\_AUMt}, \varepsilon_{DYt})' \tag{42}$$

5.2. **Reference methods.** BPNN and EMA that show good performance in traditional forecasting are chosen as comparative methods. BPNN can be used as calibration method for its supervised learning ability. Back propagation is the generalization of the Widrow-Hoff learning rule to multiple-layer networks and nonlinear differentiable transfer functions. The input vectors and their corresponding target vectors are used to train a network until it can approximate a function, associate input vectors with specific output vectors, or classify input vectors in an appropriate way.

Exponential Moving Average is commonly applied to smooth data, as many window functions are in signal processing, acting as low-pass filters to remove high frequency noise. See Kolmogorov-Zurbenko filter for more information. Here we take the simplest form of exponential smoothing as the reverence because this kind of method is much more effective in volume of stationary data in previous literature.

Meanwhile, three measures are often used to assess the prediction accuracy of specific method in empirical research. Here we also adopt them in next section.

Prediction Error of the Last Year ($\boldsymbol{e_T}$)

$$\boldsymbol{e_T} = \frac{y_T - \widehat{y_T}}{y_T} \tag{43}$$

Mean Percentage Error ($\boldsymbol{mpe}$)

$$\boldsymbol{mpe} = \frac{\sum_{t=1}^{T} \frac{y_t - \widehat{y_t}}{y_t}}{T} \tag{44}$$

Mean Square or Standard Deviation of Prediction Error ($\boldsymbol{mse}$ or $\boldsymbol{sde}$)

$$\boldsymbol{mse} = \frac{\sum_{t=1}^{T} (y_t - \widehat{y_t})^2}{T} \tag{45}$$

$$\boldsymbol{sde} = \sqrt{\frac{\sum_{t=1}^{T} (y_t - \widehat{y_t})^2}{T}} \tag{46}$$

y_T and y_t present the known sample value in the last year and in the year t. $\widehat{y_T}$ and $\widehat{y_t}$ present the forecasted sample value in the last year and in the year t by our forecast model. $t = 1, 2, \cdots, T$. Generally $\boldsymbol{mse}$ or $\boldsymbol{sde}$ of prediction error measure forecast accuracy better than $\boldsymbol{mpe}$, and $\boldsymbol{e_T}$ is interesting to reflect the prediction accuracy of the latest time. Here we integrated all three methods into account.

6. **Application and Main Results.**

6.1. **Data description.** Seen as Table 2, the sample ranges from 2002 to 2013 (nearly 11 years). ROE and GR_AUM of the CACH_L in 2010, CAMB_L in 2006, CRCT in 2007, FIRST in 2007, LIPO_R in 2008, PAKW in 2007, SABA in 2011 cannot be obtained. In addition, FAR_H, SABA and ASCD_H have only 1 or 2-year data. After removing REITs with insufficient information missing data, the final sample consisting of 19 main representative REITs, so we are to apply their relevant data to the new method.

Sample information is collected from annual reports of 22 REITs, as they are disclosed directly as individual components in the financial statements. REIT sectors are also obtained from each REIT's IPO prospectuses. Key performance variables – Dividend

Per Unit (DPU), Return On Equity (ROE), Market to Book Value (MVBV), scale Gross Revenue (GR) by Asset Under Management (AUM) and REIT's stock excess return performance compared to market excess return are collected from Thomson DataStream. Definitions and descriptive statistics of variables are summarized in Table 3.

6.2. **Task description.** It is one key point of the hybrid LS-SVR-DS to search the optimal parameters (σ^*, γ^*), so we computed all the assembly parameters after a cycle of 5000 times. After the algorithm stops, the "optimal" pair of (σ^*, γ^*) for the empirical model minimizes the training error. By this simple method the resulting new model can describe most information about the original data quickly. In addition, α_t in all models of Y_{DPUt^*}, Y_{ROEt^*}, Y_{MVBVt^*}, $Y_{GR\_AUMt^*}$ and Y_{DYt^*} has no practical meanings and is strictly different from traditional auto-regression models, which is discussed before and not displayed in results.

Prediction Error of the Last Year ($\boldsymbol{e_T}$), Mean Percentage Error ($\boldsymbol{mpe}$) and Standard Deviation of Prediction Error ($\boldsymbol{sde}$) together help us judge whether it is suitable to apply our new model to forecast the 5 performance variables of the specific S-REIT in Tables 4-8, where *, ** and ***, which respectively mean the corresponding absolute values were falling in the interval $(10\%, 5\%)$, $(5\%, 1\%)$ and $(1\%, 0)$. It is convenient for us to observe and compare the prediction accuracy visually. $\boldsymbol{mpe}$ and $\boldsymbol{sde}$ were computed by sample observations y_t and values $\widehat{y_t}$ were predicted in every year of the whole time range. Fitting and forecast of the historical data could provide significant evidence for us to prove the advantages of the proposed method.

6.3. Prediction results.

6.3.1. *Results by new method.* As to one REIT, there are 5 performance measures in each of our new models. However, we categorize results by 5 performance variables with 5 tables in order to make comparison analysis of different sample sizes.

From Tables 4 to 8 we concluded that nearly all the REITs' prediction accuracy perform well under our new models. Their Prediction Error of the Last Year ($\boldsymbol{e_T}$) and Mean Percentage Error ($\boldsymbol{mpe}$) were very low, prediction accuracy went into the $(1\%, 0)$ even much smaller, so does the Standard Deviation of Prediction Error ($\boldsymbol{sde}$). What is more, CACH_L and MAPT_I with only 3-year data, LIPO_R, PAKW, FIRST and CRCT with 7-year data and the others with less than 11-year data have shown great performance through our new method. Since most traditional forecast models require a large amount of sample data, they could not put their ability to good use in small sample sizes.

Of course there were some special ones that should be explained carefully. In Table 4, 3 indices of prediction accuracy AIMS_A had not pass the stands of *, ** and ***. Prediction Error of the Last Year ($\boldsymbol{e_T}$) and Mean Percentage Error ($\boldsymbol{mpe}$) of CRCT, FORT and SUNT are feasible, although the Standard Deviation of Prediction Error ($\boldsymbol{sde}$) are a little bigger. CAMB_L only has a good performance of Mean Percentage Error ($\boldsymbol{mpe}$), which was not enough to be persuasive. Similar analysis in Table 5 and Table 8, CRCT, FORT, SUNT were considered to be heightened in the future study. In Table 6, ASCT_R and AIMS_A got our attention. All the REITs in Table 7 achieve good prediction accuracy.

There are different reasons to interpret the REITs' bad prediction accuracy. Firstly, DPU of AIMS_A (0.01) and FORT (0.01), ROE of FORT (111.29), MVBV of CRCT (0.30), DY of AIMS_A (187.45) may suffer from exogenous volatility of historical data because their variances that were listed in parentheses are relatively larger than other REITs, which can be seen in Table 3. Secondly, in reality a year of inflection point may bring big influence on calculating the $\boldsymbol{e_T}$, $\boldsymbol{mpe}$ and $\boldsymbol{sde}$ such as ROE of SUNT in 2009. Thirdly, some other unobserved factors such as management fee structure, scale, industrial

TABLE 3. Descriptive statistics of variables

REITs	**DPU**			**ROE**			**MVBV**			**GR_AUM**			**DY**		
Names	**Min/Max**	**Mean**	**Variance**	**Min/Max**	**Mean**	**Variance**	**Min/Max**	**Mean**	**Variance**	**Min/Max**	**Mean**	**Variance**	**Min/Max**	**Mean**	**Variance**
AIMS_A	0.1/0.35	0.19	0.01	(−28.83)/15.57	4.76	282.31	0.22/1.13	0.72	0.10	56.69/96.60	85.50	214.80	6.96/36.54	19.53	187.45
ASCD	0.07/0.15	0.12	0.00	2.87/27.76	10.48	55.68	0.64/1.63	1.16	0.07	34.77/90.20	76.71	310.99	4.82/11.74	6.32	3.55
ASCT_R	0.02/0.06	0.04	0.00	(−13.65)/17.79	6.80	115.07	0.39/1.02	0.84	0.05	85.27/110.05	101.34	66.41	1.15/9.19	3.70	6.94
CACH_L	0/0.08	0.05	0.00	8.56/9.52	9.04	0.46	1.03/1.29	1.13	0.02	73.36/75.10	74.23	1.51	0/8.45	5.04	19.86
CAMB_L	0/0.06	0.04	0.00	(−13.82)/19.32	8.19	140.05	0.36/1.08	0.82	0.06	70.14/80.62	74.23	22.47	0/22.25	10.10	44.76
CCT	0/0.09	0.04	0.00	(−18.27)/26.79	12.36	186.78	0.31/1.29	0.84	0.07	34.31/66.67	54.16	100.64	0/10.39	5.32	9.41
CMT	0/0.12	0.08	0.00	(−0.45)/15.62	7.66	22.48	0.65/1.5	1.21	0.06	63.05/84.38	72.15	43.26	0/8.09	4.72	3.89
CRCT	0/0.07	0.05	0.00	3.15/17.93	10.12	32.93	0.46/2.12	1.16	0.30	77.82/99.37	92.72	78.75	0/7.46	4.44	6.55
CDL_H	0/0.11	0.08	0.00	(−5.04)/11.72	5.24	29.78	0.49/1.47	1.16	0.12	20.00/71.42	58.31	323.76	0/15.11	5.66	22.38
FIRST	0.04/0.05	0.05	0.00	9.12/17.4	12.39	10.60	0.43/1.26	0.87	0.07	73.63/88.88	84.11	36.90	4.94/16.7	8.32	18.31
FORT	0/0.38	0.26	0.01	(−7.75)/28.37	10.78	111.29	0.24/1.00	0.66	0.05	33.37/75.37	57.77	174.07	0/20.19	7.45	29.53
FRAS_CE	0.07/0.1	0.08	0.00	5.33/18.32	12.25	19.80	0.50/1.28	0.95	0.07	65.97/90.00	76.31	60.07	4.71/11.95	6.41	7.51
KREIT	0/0.11	0.07	0.00	(−4.15)/24.37	8.88	107.27	0.31/1.24	0.77	0.09	23.33/45.90	34.32	84.05	0/18.65	7.18	35.31
LIPO_R	0.02/0.04	0.03	0.00	(−37.09)/13.13	−0.53	597.11	0.44/0.85	0.63	0.02	75.61/105.68	92.02	200.63	4.34/10.49	7.40	6.01
MAPT_I	0/0.08	0.05	0.00	12.60/28.56	18.42	77.72	1.08/1.21	1.14	0.00	89.56/95.15	91.44	10.36	0/6.1	4.05	12.32
MAPT_L	0.04/0.07	0.06	0.00	7.25/11.74	9.78	2.80	0.38/1.53	1.03	0.12	24.13/71.65	58.22	212.63	3.70/19.76	7.24	26.84
PAKW	0/0.08	0.06	0.00	3.95/11.98	8.33	8.35	0.55/1.36	0.99	0.09	16.33/63.17	51.85	398.91	0/9.72	4.57	10.14
STAR	0/0.07	0.04	0.00	(−16.9)/25.06	6.27	137.41	0.36/0.97	0.73	0.04	18.68/64.34	49.73	258.13	0/13.84	6.61	20.10
SUNT	0.07/0.12	0.09	0.00	(−2.54)/14.61	4.02	29.94	0.34/0.82	0.66	0.03	35.81/54.56	44.50	51.75	4.48/13.07	7.43	9.27

Table 4. Prediction accuracy of DPU by the new model

Names	VAR & SVM			BPNN			EMA		
	e_T	mpe	sde	e_T	mpe	sde	e_T	mpe	sde
AIMS_A	0.11795	0.17064	0.41901	0.32787	−0.06690	0.07246	−1.02199	−0.22876	0.12033
ASCD	0.00000***	0.00175***	0.03395**	0.01487**	0.00641***	0.00477***	0.12816	0.20155	0.28020
ASCT_R	0.00000***	0.00000***	0.00000***	0.01445**	0.16058	0.01221**	−0.21572	0.03768**	0.14291
CACH_L	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***
CAMB_L	0.29167	0.03333**	0.21093	−0.16377	−0.04406**	0.00878***	−0.36800	−0.10332	0.10483
CCT	0.00000***	0.00269***	0.05626*	−0.03445**	−0.16840	0.02001**	0.24085	0.39935	0.29046
CMT	−0.02660**	0.01706**	0.13456	0.03333**	0.01073**	0.01202**	0.15227	0.34817	0.35032
CRCT	0.00000***	0.10000*	0.20000	−0.13356	−0.09719*	0.00819	0.22674	0.23063	0.16267
CDL_H	−0.07972*	0.01900**	0.15427	0.32938	0.28168	0.02954**	0.36191	0.45472	0.42535
FIRST	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***
FORT	−0.00983***	0.01131**	0.10597	0.01341**	0.00714	0.03292**	0.11051	0.32866	0.10983
FRAS_CE	−0.13968	0.01125**	0.10208	0.10000	−0.03681**	0.00894***	0.26400	0.10880	0.11980
KREIT	0.08335*	0.00199***	0.04904**	−0.29870	−0.00045***	0.00947***	−0.01273**	0.44286	0.38912
LIPO_R	0.22736	0.03358**	0.17140	0.00000***	0.01724*	0.00408***	−0.60000	0.04211**	0.78145
MAPT_I	0.00000***	0.00000***	0.00000***	0.00035***	0.00017***	0.00002***	0.54500	0.56885	0.40375
MAPT_L	−0.04762**	0.01458**	0.12928	0.10277	0.06442*	0.00872***	0.12011	0.07256*	0.98317
PAKW	0.00000***	−0.03125**	0.06250*	−0.01869	−0.00402***	0.00414***	−0.02857**	0.02538**	0.45614
STAR	0.00296***	0.01039**	0.09897*	−0.10877	0.09038*	0.01374**	0.05780*	0.44421	0.23136
SUNT	−0.03280**	0.02670**	0.19318	−0.21420	−0.03991**	0.00876***	0.09362*	0.15074	0.19045

TABLE 5. Prediction accuracy of ROE by the new model

Names	VAR & SVM			BPNN			EMA		
	e_T	*mpe*	*sde*	e_T	*mpe*	*sde*	e_T	*mpe*	*sde*
AIMS_A	0.00000***	0.00000***	0.00000***	0.00000***	0.90767	18.19545	0.89204	1.09286	18.11453
ASCD	0.00000***	0.00000***	0.00000***	−0.50992	−0.10378	15.61159	−0.27209	0.48255	7.99910
ASCT_R	0.00000***	0.00000***	0.00000***	0.58773	0.36061	13.12334	0.56698	0.83659	10.95476
CACH_L	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.06050*	0.00839***	0.53018
CAMB_L	−0.00020	−0.00005***	0.00009***	−0.42830	0.23062	13.97460	−0.64487	0.45707	12.92253
CCT	0.00000***	0.00000***	0.00000***	−2.73477	−0.03430**	24.61625	0.27451	0.86870	14.25788
CMT	0.00000***	0.00000***	0.00000***	0.22250	0.21173	5.35827	0.19558	0.29856	4.95974
CRCT	0.00000***	0.46375	1.21650	0.48853	0.54287	7.39419	0.37700	0.52287	6.38972
CDL_H	0.00000***	0.00000***	0.00000***	0.91544	0.68254	6.24130	0.61797	0.65433	5.46802
FIRST	0.00000***	0.00000***	0.00000***	0.04634**	−0.06727*	3.71825	−0.25252	0.14449	3.38480
FORT	0.06470*	0.17107	0.92475	−0.25444	0.08543*	6.43169	0.21326	0.67982	10.99047
FRAS_CE	0.00000***	0.00000***	0.00000***	1.07326	0.83363	10.01026	0.13898	−0.03838	4.79489
KREIT	0.00000***	0.00000***	0.00000***	0.58433	0.91331	13.43749	0.11718	0.42046	10.73792
LIPO_R	0.00000***	0.00000***	0.00000***	−0.00001***	0.00000***	0.00005***	1.68948	0.98302	24.21075
MAPT_I	0.00000***	0.00000***	0.00000***	0.00000***	0.00001***	0.00022***	−0.48326	0.06565	8.20052
MAPT_L	0.00000***	0.00000***	0.00000***	0.00536***	−0.07266***	1.95058	0.11622	0.02377**	1.71454
PAKW	0.00000***	0.00000***	0.00000***	0.66747	0.44383	4.66035	0.41671	0.28542	3.12758
STAR	0.00000***	0.00000***	0.00000***	−0.65538	0.09736*	4.32569	0.34508	0.89191	12.25043
SUNT	−0.00152***	−0.33093	0.51879	0.58647	0.36550	7.98971	0.81905	0.76237	5.38885

TABLE 6. Prediction accuracy of MVBV by the new model

Names	VAR & SVM			BPNN			EMA		
	e_T	*mpe*	*sde*	e_T	*mpe*	*sde*	e_T	*mpe*	*sde*
AIMS_A	−0.07159*	0.02562**	0.19618	−0.03234**	0.13037	0.18304	0.46246	0.42425	0.35242
ASCD	0.02347**	0.04684**	0.24535	0.16748	0.21911	0.34043	0.01782**	0.10844	0.29730
ASCT_R	−0.16823	0.10497	0.46475	0.04807**	0.11039	0.19967	0.15070	−0.01170**	0.23515
CACH_L	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.12093	0.02471	0.14359
CAMB_L	0.00000***	0.00000***	0.00000***	0.16943	−0.00952***	0.22496	0.32284	0.21278	0.24246
CCT	0.00000***	0.00000***	0.00000***	0.24366	0.19496	0.29282	0.21164	0.10710	0.27789
CMT	0.00000***	0.00000***	0.00000***	0.06885*	0.03493**	0.22259	0.09186*	0.04595**	0.24859
CRCT	−0.11382	0.01215**	0.12415	0.25173	0.13760	0.20234	0.30634	0.20697	0.32000
CDL_H	0.00000***	0.00000***	0.00000***	−0.00081***	0.06343*	0.57316	−0.03551**	−0.06955*	0.38544
FIRST	−0.15129	0.01318**	0.12211	0.29135	0.26872	0.27427	0.41519	0.30535	0.31992
FORT	−0.06867*	0.08376*	0.36915	0.40872	0.08679*	0.13726	0.15112	−0.16306	0.24950
FRAS_CE	0.00000***	0.00000***	0.00000***	0.13226	−0.03479**	0.31055	0.32434	0.17624	0.27445
KREIT	0.00000***	0.00000***	0.00000***	0.00000***	−0.08870*	0.27868	0.22137	0.05267*	0.32041
LIPO_R	−0.00001***	0.00000***	0.00003***	−0.00006***	−0.00004***	0.00016***	0.27021	0.09039*	0.11783
MAPT_I	0.00000***	0.00000***	0.00000***	0.00030***	0.00028***	0.00032***	0.08760*	0.03176**	0.06522*
MAPT_L	−0.00001***	0.00001***	0.00003***	0.31527	0.09801*	0.25183	0.04495**	−0.05793*	0.38806
PAKW	−0.04893**	0.00369***	0.06690***	0.00018***	−0.04020**	0.07027*	0.35287	0.29903	0.35499
STAR	0.00000***	0.00000***	0.00000***	0.30287	0.16916	0.18202	0.16934	−0.09275*	0.22519
SUNT	−0.14201	0.08869*	0.39939	−0.02408***	−0.10270	0.27025	0.18945	0.05197*	0.17238

TABLE 7. Prediction accuracy of GR_AUM by the new model

Names	VAR & SVM			BPNN			EMA		
	e_T	*mpe*	*sde*	e_T	*mpe*	*sde*	e_T	*mpe*	*sde*
AIMS_A	0.00000***	0.00000***	0.00000***	−0.00948***	0.01941**	3.34928	0.04090**	0.09787*	15.25974
ASCD	−0.01570***	0.00186***	0.04374***	−0.00763***	−0.03719**	5.60284	0.12360	0.22350	20.70725
ASCT_R	0.00000***	0.00000***	0.00000***	0.01456**	0.02760**	11.87805	−0.03402**	−0.00823***	8.22169
CACH_L	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	−0.01423**	−0.00102***	0.96067*
CAMB_L	0.00000***	0.00000***	0.00000***	0.03252**	0.04821**	6.18739	−0.01485**	0.03299**	5.20159
CCT	0.00025***	0.00004***	0.00040***	−0.00266***	0.04026**	5.88794	−0.02585**	0.03978**	10.13471
CMT	0.00000***	0.00000***	0.00000***	−0.05135*	−0.01366**	5.93585	−0.14791	−0.02520**	6.63466
CRCT	0.00000***	0.00000***	0.00000***	−0.01090**	−0.00667***	3.52086	0.05781*	0.04515**	8.97117
CDL_H	0.00000***	0.00000***	0.00000***	0.52287	0.38838	28.16034	0.24942	0.29770	21.18257
FIRST	0.00000***	0.00000***	0.00000***	0.00000***	−0.04622**	7.96851	0.02750**	−0.03005**	6.39580
FORT	0.03741**	0.04990**	0.26441	−0.00372***	0.03053**	11.84288	−0.07410*	0.17146	14.76375
FRAS_CE	0.00000***	0.00000***	0.00000***	−0.01894**	0.00696***	3.47975	−0.00199***	−0.05350*	8.27641
KREIT	0.00000***	0.00000***	0.00000***	−0.90392	−0.01325**	10.31176	−0.62635	−0.15595	10.82377
LIPO_R	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	−0.08925*	0.04941**	13.66310
MAPT_I	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.05858*	0.02114**	3.21811
MAPT_L	0.00000***	0.00000***	0.00000***	0.08011*	−0.01821**	5.42230	0.22689	0.22278	16.30031
PAKW	0.00000***	0.00000***	0.00000***	0.06651*	0.03770**	3.10608	0.26056	0.29856	20.72045
STAR	−0.00004***	0.00000***	0.00003***	0.17281	−0.00227***	3.40605	−0.58938	0.17020	17.13937
SUNT	0.12628	0.01942**	0.14889	−0.11234	0.04362**	5.59256	−0.17879	−0.02913**	7.25977

TABLE 8. Prediction accuracy of DY by the new model

Names	VAR & SVM			BPNN			EMA		
	e_T	mpe	sde	e_T	mpe	sde	e_T	mpe	sde
AIMS_A	0.00003***	0.13670	0.30104	−1.15461	0.05170*	5.46650	−2.23028	−0.26171	15.89791
ASCD	0.00000***	0.00000***	0.00000***	−0.13230	−0.11473	3.25063	−0.02109**	0.13359	1.97491
ASCT_R	0.11955	0.00400***	0.06910***	−1.80213	−0.31480	3.88578	−0.47389	0.38300	2.71230
CACH_L	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	−0.15898	0.00331***	0.97751
CAMB_L	0.00000***	0.00000***	0.00000***	−0.63581	0.37099	7.07005	−0.99224	−0.08409*	5.89521
CCT	0.00000***	0.00000***	0.00000***	0.34195	0.48766	3.37654	−0.17137	0.44329	3.22172
CMT	0.00000***	0.00000***	0.00000***	−0.04884**	−0.04632**	1.61682	−0.07248*	0.31718	2.11426
CRCT	0.14680	0.01925**	0.13967	−0.20219	0.00922***	0.54477	−0.37199	0.01121**	1.48655
CDL_H	0.00000***	0.00000***	0.00000***	−2.00212	−0.96284	11.55845	0.21145	0.66080	5.32797
FIRST	0.00000***	0.00000***	0.00000***	−0.22928	0.00560***	0.73512	−1.09456	−0.15391	5.03105
FORT	0.76685	0.17671	0.35994	−0.29068	0.38925	4.93771	−0.77338	0.51315	5.79511
FRAS_CE	0.00004***	0.03472**	0.14917	−0.00005***	0.23751	2.23484	−0.38877	−0.04938**	2.98552
KREIT	0.00000***	0.00000***	0.00000***	−0.80525	0.21894	4.87490	−0.35958	0.64866	6.73768
LIPO_R	0.00000***	0.00000***	0.00000***	−0.00002***	0.00000***	0.00025***	−0.68194	0.13867	2.49414
MAPT_I	0.00000***	0.00000***	0.00000***	0.00000***	0.00000***	0.00002***	0.48341	0.54132	3.21957
MAPT_L	0.00000***	0.00000***	0.00000***	−0.19301	0.27936	4.74526	−0.09275*	0.44134	5.53397
PAKW	0.24962	0.01909**	0.16538	−0.00010***	0.14589	0.94170	−1.07442	−0.16798	2.73896
STAR	0.00000***	0.00000***	0.00000***	0.12447	0.46092	4.39369	−0.26084	0.53327	4.98136
SUNT	0.00000***	0.00000***	0.00000***	−0.34423	0.03507**	2.56957	−0.22793	0.31703	3.46999

sectors or other characteristics of company were not considered or distinguished in this paper, which provided the basis for the rationality of nonlinearity (such as Gaussian kernel function) from another angle.

Generally speaking, results show that performance forecast of S-REITs by the new method owns virtues of simple operation, fast convergence speed and high precision even though the sample size is really small just as only 3, 5 or 7-year data or less. It is also verified that our method is feasible of processing small sample and nonlinearities. Performance about DY, ROE, MVBV, GR_AUM and DPU are of current concern to investors, policymakers and researchers of S-REITs' performance. In practice our advanced method could provide an effective tool to make individual REIT forecast rather than macro forecasting. $\boldsymbol{e_T}$, $\boldsymbol{mpe}$ and $\boldsymbol{sde}$ focus on different forecasting periods, so empirical part can help us grasp the tendency of S-REIT.

6.3.2. *Comparison with BPNN and EMA.* Tables 4 to 8 also showed us forecasting accuracy of BPNN and EMA under the same sample. Compared to their results, the new proposed method is superior on the predictive performance.

Firstly, EMA displayed the much bigger error than BPNN, no matter with Prediction Error of the Last Year, Mean Percentage Error or the Standard Deviation of Prediction Error. It is explained by strict requirement of many stationary time trend data. 5 performance indices of S-REITs are affected by many uncertainties of dividend policy, the stock market, bonds, the confidence of investors and so on.

Secondly, except for a few other exceptions (3 accuracy measures of CAMB_L, FRAS_CE, LIPO_R are not consistent better in Table 4, so did SUNT in Tables 6 and 7), BPNN cannot be thought as good as the new method in most fields. The new method was shown to be very resistant to the over-fitting problem, eventually achieving a high generalization performance, and the solutions are always unique and globally optimal due to the advantage of SVM. Compared to the above, BPNN training requires nonlinear optimization with the danger of getting stuck at local minima. CACH_L with 3-year data and FIRST with 5-year data perform better in all methods in Table 4. CACH_L has high prediction accuracy by our methods and BPNN from Tables 4 to 8, so do LIPO_R and MAPT_I in Table 7.

Thirdly, from frequency and counting bigger value of 3 methods of $\boldsymbol{e_T}$, $\boldsymbol{mpe}$ and $\boldsymbol{sde}$, our new approach owns the dominant superiority. The total number of *, ** and *** are larger than two referred methods. Under the same condition of *, ** or *** the values are also bigger than the others. As we can see, our prediction results are much significant. It is also verified that our method is better than BPNN and EMA due to its feasibility.

6.3.3. *Comparison in different industries.* Forecasting volatility of different industries in the sample was made a comparison by calculating the average $\boldsymbol{sde}$ values of 5 key performance variables (see as in Figure 2). Table 2 showed us that S-REITs that belonged to different industrial sectors are distinct from each other. One performance variable with information from historical data affects the forecast accuracy to varying degrees. As a whole, GR_AUM and DY owned satisfactory forecast volatility (the average $\boldsymbol{sde}$ values is lower than 10%, 5% even 1%) regardless which industrial sector the REIT engaged in (see as in Figures 2e and 2f).

Except that the mean of ROE in commercial S-REITs is almost 30%, other sectors managed to pass the level of 1% even 0.01% when forecasting ROE. Similarly industrial/mixed and service residence sectors had a bigger volatility of forecasting MVBV. And commercial, industrial and hospitality sectors had a bigger volatility of forecasting DPU. Despite that some industrial bad performance could be explained by the total accounts of sample

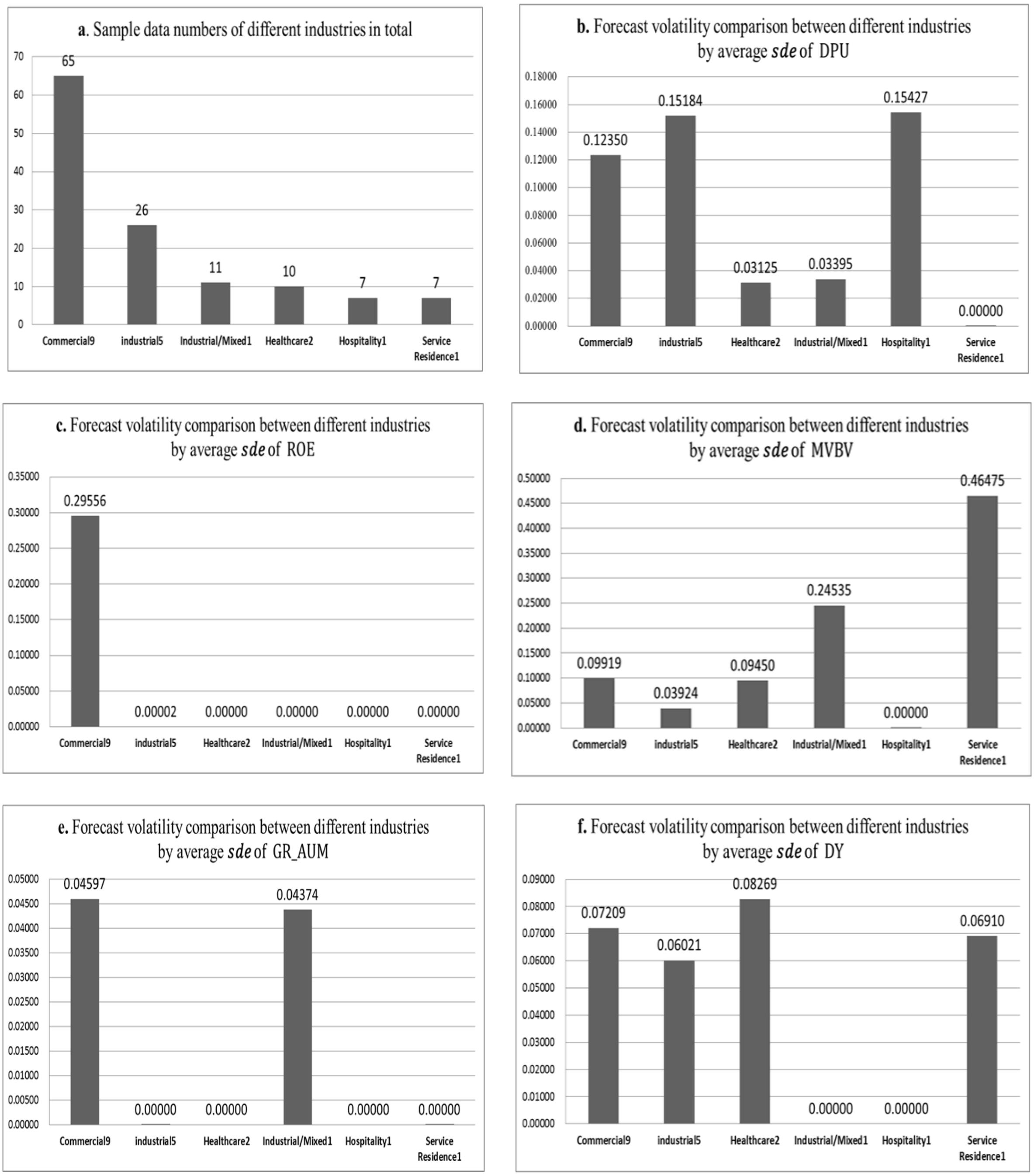

FIGURE 2. Forecast volatility comparison of different industries by calculating average $\boldsymbol{sde}$

data (see as in Figure 2a) to some extent, we still concluded that the proposed model provides the chance to analyze specific performance of S-REITs.

All in all, after a detailed discussion of empirical study about S-REITs, we can see that most S-REITs gained good forecast accuracy of 5 key performance variables with small sample. Compared to their results of **BPNN** and **EMA**, the new proposed method was superior on the predictive performance. And most sectors had an acceptable forecasting volatility in the industrial level on the whole. It also manifests that the nonlinearity like support vector kernels had better fitting abilities and forecasting performance. Meanwhile, simultaneous analysis of commonly used performance variables was useful and gave us a comprehensive understanding at smaller computational cost.

7. **Conclusions.** REITs are strong income vehicles because most of them must pay out at least 90 percent of taxable income in the form of dividends to shareholders. Performance prediction of S-REITs is of great importance to investors in practice. It is characterized by small sample size and nonlinear specification. As one of the most easily operated model on analyzing and forecasting of multiple related economic indicators, $\mathrm{VAR}(1,k)$ are very popular. We attempted to use SVM to modify it to obtain a novel model. The following conclusions could be drawn. (1) Data of S-REITs in company level is a specific and important economic problem with small sample. And relationship between performance indicators in different periods cannot be treated by simple linearity. (2) The novel combination of traditional forecast model $\mathrm{VAR}(1,k)$ and statistic learning theory (SVM) behave well on prediction accuracy of S-REITs' 5 performance index. Compared with BPNN and EMA, our new method is much better. In future study, during performance prediction of S-REITs, much more micro (corporate management fee structure, scale, industrial species, etc.) or macro (political policy, monetary policy, environmental conditions, etc.) factors should be given enough consideration.

Acknowledgements. This work is partially supported by National Natural Science Foundation of China (Study on Soft Sets Theories and their Applications faced on Multi-Attributes and Small Samples Decision Making: Grant No. 71171209). Thanks go to all my team members. The authors also gratefully acknowledge the helpful comments and suggestions of the associate editor and reviewers, which have improved the presentation.

REFERENCES

[1] A. Abbasi, Z. Zhang, D. Zimbra, H. Chen and J. F. Nunamaker, Detecting fake websites: The contribution of statistical learning theory, *MIS Quarterly*, vol.34, pp.435-461, 2010.

[2] J. Alcock, J. Glascock and E. Steiner, Manipulation in U.S. REIT investment performance evaluation: Empirical evidence, *Journal of Real Estate Finance and Economics*, vol.47, pp.434-465, 2013.

[3] G. Baudat and F. E. Anouar, Generalized discriminant analysis using a kernel approach, *Neural Computation*, vol.12, pp.2385-2404, 2000.

[4] Y. Bazi and F. Melgani, Semisupervised PSO-SVM regression for biophysical parameter estimation, *IEEE Trans. Geoscience and Remote Sensing*, vol.45, pp.1887-1895, 2007.

[5] B. S. Bernanke and A. S. Blinder, The federal-funds rate and the channels of monetary transmission, *American Economic Review*, vol.82, pp.901-921, 1992.

[6] P. T. Brandt and T. Sandler, A Bayesian poisson vector autoregression model, *Political Analysis*, vol.20, pp.292-315, 2011.

[7] J. P. Brooks, Support vector machines with the ramp loss and the hard margin loss, *Operations Research*, vol.59, pp.467-479, 2011.

[8] D. Brounen, N. Kok and D. C. Ling, Shareholder composition, share turnover, and returns in volatile markets: The case of international REITs, *Journal of International Money and Finance*, vol.31, pp.1867-1889, 2012.

[9] B. Case, W. G. Hardin and Z. H. Wu, REIT dividend policies and dividend announcement effects during the 2008-2009 liquidity crisis, *Real Estate Economics*, vol.40, pp.387-421, 2012.

[10] M. Chacon, C. Araya and R. B. Panerai, Non-linear multivariate modeling of cerebral hemodynamics with autoregressive support vector machines, *Medical Engineering & Physics*, vol.33, pp.180-187, 2011.

[11] Z. Y. Chen, Z. P. Fan and M. H. Sun, A hierarchical multiple kernel support vector machine for customer churn prediction using longitudinal behavioral data, *European Journal of Operational Research*, vol.223, pp.461-472, 2012.

[12] W. H. Chou, W. G. Hardin, M. D. Hill and G. W. Kelly, Dividends, values and agency costs in REITs, *Journal of Real Estate Finance and Economics*, vol.46, pp.91-114, 2013.

[13] R. Chung, S. Fung and S. Y. K. Hung, Institutional investors and firm efficiency of real estate investment trusts, *Journal of Real Estate Finance and Economics*, vol.45, pp.171-211, 2012.

[14] P. Danenas and G. Garsva, Support vector machines and their application in credit risk evaluation process, *Transformations in Business & Economics*, vol.8, pp.46-58, 2009.

[15] S. J. Dempsey, D. M. Harrison, K. F. Luchtenberg and M. J. Seiler, Financial opacity and firm performance: The readability of REIT annual reports, *Journal of Real Estate Finance and Economics*, vol.45, pp.450-470, 2012.

[16] V. Fernandez, Traditional versus novel forecasting techniques: How much do we gain? *Journal of Forecasting*, vol.27, pp.637-648, 2008.

[17] W. M. Gentry, D. Kemsley and C. J. Mayer, Dividend taxes and share prices: Evidence from real estate investment trusts, *Journal of Finance*, vol.58, pp.261-282, 2003.

[18] C. W. J. Granger and A. A. Weiss, Time series analysis of error-correction models, *Studies in Econometrics, Time Series, and Multivariate Statistics*, pp.255-278, 1983.

[19] Y. Gunduz and M. Uhrig-Homburg, Predicting credit default swap prices with financial and pure data-driven approaches, *Quantitative Finance*, vol.11, pp.1709-1727, 2011.

[20] K. Holden, Vector autoregression modeling and forecasting, *Journal of Forecasting*, vol.14, pp.159-166, 1995.

[21] C. L. Huang, M. C. Chen and C. J. Wang, Credit scoring with a data mining approach based on support vector machines, *Expert Systems with Applications*, vol.33, pp.847-856, 2007.

[22] C. L. Huang and J. F. Dun, A distributed PSO-SVM hybrid system with feature selection and parameter optimization, *Applied Soft Computing*, vol.8, pp.1381-1391, 2008.

[23] A. Hyvarinen, K. Zhang, S. Shimizu and P. O. Hoyer, Estimation of a structural vector autoregression model using non-Gaussianity, *Journal of Machine Learning Research*, vol.11, pp.1709-1731, 2010.

[24] J. F. Jaffe, Taxes and the capital structure of partnerships, REITs, and related entities, *Journal of Finance*, vol.46, pp.401-407, 1991.

[25] A. Kazem, E. Sharifi, F. K. Hussain, M. Saberi and O. K. Hussain, Support vector regression with chaos-based firefly algorithm for stock market price forecasting, *Applied Soft Computing*, vol.13, pp.947-958, 2013.

[26] K. J. Kim, Financial time series forecasting using support vector machines, *Neurocomputing*, vol.55, pp.307-319, 2003.

[27] G. M. Koop, Forecasting with medium and large Bayesian VARs, *Journal of Applied Econometrics*, vol.28, pp.177-203, 2013.

[28] P. Lecomte and J. T. L. Ooi, Corporate governance and performance of externally managed Singapore REITs, *Journal of Real Estate Finance and Economics*, vol.46, pp.664-684, 2013.

[29] C. S. Li, X. L. An and R. H. Li, A chaos embedded GSA-SVM hybrid system for classification, *Neural Computing & Applications*, vol.26, pp.713-721, 2015.

[30] S. J. Li, H. Wu, D. S. Wan and J. L. Zhu, An effective feature selection method for hyperspectral image classification based on genetic algorithm and support vector machine, *Knowledge-Based Systems*, vol.24, pp.40-48, 2011.

[31] S. W. Lin, K. C. Ying, S. C. Chen and Z. J. Lee, Particle swarm optimization for parameter determination and feature selection of support vector machines, *Expert Systems with Applications*, vol.35, pp.1817-1824, 2008.

[32] H. Moskowitz, P. Drnevich, O. Ersoy, K. Altinkemer and A. Chaturvedi, Using real-time decision tools to improve distributed decision-making capabilities in high-magnitude crisis situations, *Decision Sciences*, vol.42, pp.77-93, 2011.

[33] P. J. G. Nieto, E. Garcia-Gonzalo, F. S. Lasheras and F. J. D. Juez, Hybrid PSO-SVM-based method for forecasting of the remaining useful life for aircraft engines and evaluation of its reliability, *Reliability Engineering & System Safety*, vol.138, pp.219-231, 2015.

[34] J. Pahasa and I. Ngamroo, PSO based kernel principal component analysis and multi-class support vector machine for power quality problem classification, *International Journal of Innovative Computing, Information and Control*, vol.8, no.3(A), pp.1523-1539, 2012.

[35] A. K. Pennathur, O. W. Gilley and R. M. Shelor, An analysis of REIT CEO stock-based compensation, *Real Estate Economics*, vol.33, pp.189-202, 2005.

[36] F. Pérez-Cruz, G. Camps-Valls, E. Soria-Olivas, J. J. Pérez-Ruixo, A. R. Figueiras-Vidal and A. Artés-Rodríguez, Multi-dimensional function approximation and regression estimation, in *Artificial Neural Networks – ICANN 2002*, J. R. Dorronsoro (ed.), 2002.

[37] E. Pervukhina, J. F. Emmenegger, V. Golikova and K. Osipov, An optimization technique based on a vector autoregression model with state space representation: Application to Ukrainian cargo transport data, *Optimization*, vol.63, pp.93-108, 2014.

[38] G. W. Peters, B. Kannan, B. Lassock, C. Mellen and S. Godsill, Bayesian cointegrated vector autoregression models incorporating Alpha-stable noise for inter-day price movements via approximate Bayesian computation, *Bayesian Analysis*, vol.6, pp.755-791, 2011.

[39] E. Pourbasheer, S. Riahi, M. R. Ganjali and P. Norouzi, Application of genetic algorithm-support vector machine (GA-SVM) for prediction of BK-channels activity, *European Journal of Medicinal Chemistry*, vol.44, pp.5023-5028, 2009.

[40] A. Rakotomamonjy, F. R. Bach, S. Canu and Y. Grandvalet, SimpleMKL, *Journal of Machine Learning Research*, vol.9, pp.2491-2521, 2008.

[41] A. Rakotomamonjy and S. Chanda, L(P)-norm multiple kernel learning with low-rank kernels, *Neurocomputing*, vol.143, pp.68-79, 2014.

[42] M. Sanchez-Fernadez, M. de-Prado-Cumplido, J. Arenas-Garcia and F. Perez-Cruz, SVM multiregression for nonlinear channel estimation in multiple-input multiple-output systems, *IEEE Trans. Signal Processing*, vol.52, pp.2298-2307, 2004.

[43] S. Simkins, Forecasting with vector autoregression (VAR) models subject to business cycle restrictions, *International Journal of Forecasting*, vol.12, pp.569-570, 1996.

[44] C. A. Sims, Interpreting the macroeconomic time-series facts: The effects of monetary-policy, *European Economic Review*, vol.36, pp.975-1000, 1992.

[45] X. M. Sun, M. Zhou and M. Zhang, Empirical study on the relationship between economic growth and carbon emissions in resource-dependent cities based on vector autoregression model, *2010 International Conference on Energy, Environment and Development*, vol.5, pp.2461-2467, 2011.

[46] M. Szafranski, Y. Grandvalet and A. Rakotomamonjy, Composite kernel learning, *Machine Learning*, vol.79, pp.73-103, 2010.

[47] D. Tuia, J. Verrelst, L. Alonso, F. Perez-Cruz and G. Camps-Valls, Multioutput support vector regression for remote sensing biophysical parameter estimation, *IEEE Geoscience and Remote Sensing Letters*, vol.8, pp.804-808, 2011.

[48] V. Vapnik, S. E. Golowich and A. Smola, Support vector method for function approximation, regression estimation, and signal processing, in *Advances in Neural Information Processing Systems*, M. C. Mozer, M. I. Jordan and T. Petsche (eds.), 1997.

[49] N. Vlastakis, G. Dotsis and R. N. Markellos, Nonlinear modelling of European football scores using support vector machines, *Applied Economics*, vol.40, pp.111-118, 2008.

[50] C. H. Wu, G. H. Tzeng, Y. J. Goo and W. C. Fang, A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy, *Expert Systems with Applications*, vol.32, pp.397-408, 2007.

[51] S. M. Wu and A. Akbarov, Support vector regression for warranty claim forecasting, *European Journal of Operational Research*, vol.213, pp.196-204, 2011.

[52] T. Xiong, Y. Bao and Z. Hu, Multiple-output support vector regression with a firefly algorithm for interval-valued stock price index forecasting, *Knowledge-Based Systems*, vol.55, pp.87-100, 2014.

[53] D. Yan, X. Liu and L. Zou, Probability fuzzy support vector machines, *International Journal of Innovative Computing, Information and Control*, vol.9, no.7, pp.3053-3060, 2013.

[54] C.-Y. Yeh, C.-W. Huang and S.-J. Lee, A multiple-kernel support vector regression approach for stock market price forecasting, *Expert Systems with Applications*, vol.38, pp.2177-2186, 2011.

[55] H. Yi, X. Song and B. Jiang, Structure selection for DAG-SVM based on misclassification cost minimization, *International Journal of Innovative Computing, Information and Control,* vol.7, no.9, pp.5133-5143, 2011.