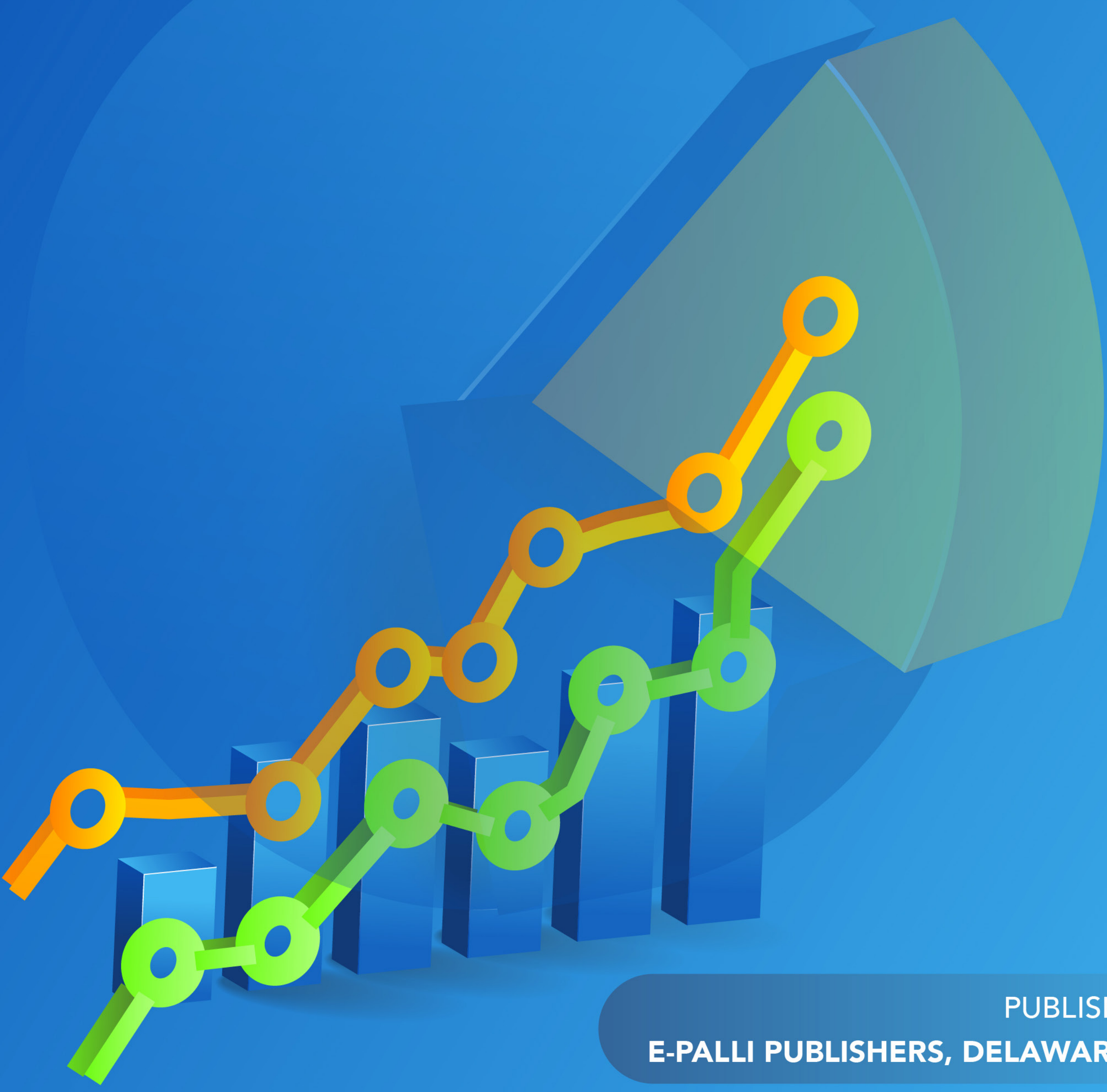
American Journal of Applied Statistics and Economics (AJASE)

ISSN: 2992-927X (ONLINE)

VOLUME 4 ISSUE 1 (2025)



PUBLISHED BY
E-PALLI PUBLISHERS, DELAWARE, USA

American Journal of Applied Statistics and Economics (AJASE)

Volume 4 Issue 1, Year 2025
ISSN: 2992-927X (Online)
DOI: https://doi.org/10.54536/ajase.v4i1.5514
https://journals.e-palli.com/home/index.php/ajase

Novel Neural Network State Switching Models for Returns Predicting with Regime Switching: A Monte Carlo's Simulation Approach

Oluwasegun Agbailu Adejumo[1*], Omorogbe Joseph Asemota[1], Samuel Olayemi Olanrewaju[1]

Article Information

Received: June 10, 2025

Accepted: July 17, 2025

Published: September 05, 2025

Keywords

Markov Process, MLP, Neural Network, Regime Switching, Returns

ABSTRACT

This study proposes new neural network state switching models that will improve the regime predictive performance of nonlinear time series returns, in small sample size contexts. In this study, we developed four new neural network state switching models namely Recurrent Neural Network State Switching Model (RNNSSM), Generalized Regression Neural Network State Switching Model (GRNNSSM), Radial Basis Function Network State Switching Model (RBFSSM) and Multilayer Perceptrons State Switching Model (MLPSSM). The study presents comparative results of the estimations of the novel models alongside the traditional Markov state switching model as well as the regime prediction performances of the models using the Akaike Information Criterion (AIC), Log-likelihood and prediction accuracy measures under a Monte Carlo simulation study. Evidence from the models' estimation results particularly the AIC and Log-likelihood statistics established the novel Multilayer Perceptrons State Switching Model (MLPSSM) as the parsimonious model for the simulated returns at small sample sizes. Results from the models' regime predictions evaluation, precisely the RMSE and MAE, evidently affirmed the novel MLPSSM model as superior in its ability to predict or forecast market returns, particularly the bull and bear regimes, at small sample sizes. Therefore, this study concludes that the novel MLPSSM is the best-fitted model for market returns at small sample sizes with excellent ability of market returns' regimes/states (i.e., bull and bear) prediction. This study recommends adoption of the novel Multilayer Perceptrons State Switching Model in modelling and regime predictions of time series returns.

INTRODUCTION

Forecasting returns is essential in financial modeling and decision-making. As a measure of volatility in time series, returns exemplify a stochastic process that indicates how much variables fluctuate over time. Accurate return prediction is vital for assessing risks in key economic activities such as value at risk, asset pricing, and exchange rate management (Liao *et al.*, 2020; Adejumo *et al.*, 2020; Roy & Sarkar, 2024). In finance, returns exhibit three important features: clustering property (Cont, 2007), asymmetry (Nelson, 1992), and nonlinearity (Maheu & McCurdy, 2002). Researchers often encounter difficulties when modeling and interpreting modern time series data from diverse fields because traditional assumptions—such as linearity, normality, and stationarity—are frequently inadequate. The origins of returns modeling trace back to Engle (1982), and Bollerslev (1986), who introduced discrete-time GARCH and stochastic return processes for autoregressive conditional heteroskedasticity. However, as noted by Hamilton (1989), Nguyen *et al.* (2014), Aliyu and Wambai (2018), Adejumo *et al.* (2020), Xiuqin *et al.* (2023), Al-Sulaiman (2024) and others, standard GARCH and stochastic models cannot fully capture all key features of financial markets. This results in complex nonlinear dynamics and irregular regime shifts. A promising approach to address these challenges is switching state modeling. Before switching state modeling was used in returns forecasting, the Hidden Markov Model (HMM) was the primary approach. The HMM is a bivariate discrete-time process, denoted as $\{S_t,Y_t\}_{(t\geq 0)}$, where $\{S_t\}$ is an underlying Markov chain, and $\{Y_t\}$ is a sequence of independent random variables. The conditional distribution of Y_t depends only on S_t. Since S_t is hidden, only the stochastic process Y_t is observable. In other words, the process's state is not directly visible, but the output, which depends on the state, is observable. Therefore, all statistical inference must be based solely on Y_t, as S_t cannot be directly observed (Rydén 2008). An HMM has a distinctive dependence structure, which is useful when analyzing financial time series. To better understand this dependence, it is illustrated here (Figure 1) with a graphical model.

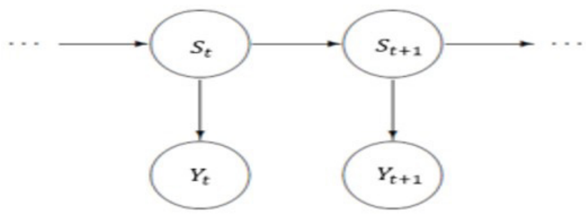


Figure 1: The HMM's Dependence Structure

According to Figure 1, the distribution of a variable $S_{(t+1)}$ conditional on the history of the process $S_0,S_1,\ldots,S_t$, is determined only by the value of the preceding variable S_t. This is all according to the Markov property, where future events are completely independent of the past,

[1] Department of Statistics, Faculty of Science, University of Abuja, Abuja, Nigeria
[*] Corresponding author's e-mail: agbailuoa@gmail.com

depending only on the present state. In addition, the distribution of Y_t conditionally on the past observations $Y_0, Y_1, \ldots, Y_{(t-1)}$ and the past values of the state, $S_0, S_1, \ldots, S_t$, is determined by S_t only (Rydén 2008). Putting this into mathematical terms, we get $f(S_{(t+1)} | S_t, S_{(t-1)}, \ldots, S_1) = f(S_{(t+1)} | S_t)$ and $f(Y_t | S_{(t-1)}, S_{(t-2)}, \ldots S_1) = f(Y_t | S_t$. Cconsequently, Markov State Switching Model (MSSM) was introduced by Hamilton to address the hidden states weakness of the HMM (Hamilton, 2005). To describe the MSSM, the number of states (or regimes) is assumed N, i.e. $S_t \, \Omega = \{1, \ldots, N\}$. This implies that; for instance the log returns of financial time series are drawn from N distinct normal distributions, depending on what state the HMM is currently in. This resolve to below models:

$$\begin{aligned} Y_{(t=1)} &= \mu_1 + \epsilon_t; \text{ where } \epsilon_t \sim (N, \sigma^2_1) \text{ for state 1} \\ Y_{(t=2)} &= \mu_2 + \epsilon_t; \text{ where } \epsilon_t \sim (N, \sigma^2_2) \text{ for state 2} \\ &\vdots \\ Y_{(t=N)} &= \mu_N + \epsilon_t; \text{ where } \epsilon_t \sim (N, \sigma^2_N) \text{ for state N} \end{aligned} \quad (1)$$

By implication, when the state of the HMM for time t is 1, then the expectation of the dependent variable is μ_1 with variance of innovations σ^2_1, similarly when the state of the HMM for time t is 2, then the expectation of the dependent variable is μ_2 with variance of innovations σ^2_2 and so on. Since the underlying Markov chain is hidden one cannot observe what state the HMM is in directly, but only deduce its operation through the observed behaviour of Y_t. In order to attain the probability law governing the observed data Y_t a probabilistic model of what causes the change from state $S_t = i$ to state $S_t = j$ is required. This can be specified using the transition probabilities of an N state HMM; $\varrho_{(i,j)} = P_r (S_t = j | S_{(t-1)} = i)$ $i, j \in \Omega = \{1, 2, \ldots, N\}$. The transition probability i.e. $\varrho_{(i,j)} = P_r (S_t = j | S_{(t-1)} = i) = P_r (S_t = j\} S_{(t-1)} = i, S_{(t-2)} = k, \ldots, S_1 = l)$ is dependent of the past only through the value of the most recent state. This is one of the central points of the structure of a Markov regime switching model, i.e. the switching of the states of the underlying HMM is a stochastic process itself. The state-switching modelling approach has proven to be more reliable in this aspect in that it models all observed salient features of returns. The approach has documented the distinctiveness and forecasting capabilities of regime switching models against the commonly used GARCH models. Thus, as a result its application has widely increased over time.

Furthermore, Machine Learning (ML) techniques also possess high capabilities in modelling all observed salient features of financial markets' returns. ML methods have been used in many other fields for many years. For example, Chen *et al.* (2019) exploit ML method in the estimation of stochastic discount factor and Gu *et al.* (2020) showed superior performance of ML models for empirical asset pricing. Also, is Christensen *et al.* (2023) employed ML in returns forecasting of index stocks. Similarly, Nelson *et al.* (2017), Shah *et al.* (2018), Yao *et al.* (2018), and Sunny *et al.* (2020) studied stock market returns based on the machine learning model. ML models especially the long-short term memory (LSTM) and recurrent neural network (RNN) types have been famous for most time series data. The machine learning models are well-known for predicting the time series data without considering the much assumptions of the parameters. However, despite both approaches significant values to returns and regime forecasting, they both suffer from either a severe modelling misspecification or a lack of effective identification of meaningful stochastic regimes especially for small sample sizes. One way to handle these problems is the switching state space modelling approach. Thus, there is a need for a well-designed modelling approach that allows the disturbance to be realistically represented, and at the same time does not lead to over frequently switching. Given the criticality of such regime identification, not only for its economic implications but also for the profound comprehension of underlying phenomena, this study aims to propose new neural network state switching models that will improve the regime predictive performance of nonlinear time series returns, in small sample size contexts. The specific objectives include to:

i Develop novel Neural Network State Switching Models (using Markov algorithms) for time series returns predicting in distinct regimes;

ii Examine the estimations of the Novel Neural Network State Switching Models towards modelling time series market returns;

iii Compare the predicting performances of the novel Neural Network State Switching Models using prediction accuracy measures under a Monte Carlo simulation study.

LITERATURE REVIEW

Regime-switching model is an unusual case of a more general framework called hidden Markov Model (Zucchini & MacDonald, 2009). Regime-switching models were early presented to econometric literature by Hamilton (1989) and have become very prevalent particularly in applied works. Applications of regime switching models range over a broad range of research areas, such as modeling shifts in inflation, exchange rates and interest rates (Piger 2013), Altug and Bildirici (2010), changes in government policy (Valente, 2003; Owyang & Ramey, 2004; Sims & Zha 2006) and shifts in exchange rate (Bekaert & Hodrick, 1992; Bollen *et al.*, 2000). The regime-switching modelling approach provides a completely new approach to the modeling of financial returns which it conceives as a multiplicative, hierarchically structured process (Frommel *et al.*, 2005).

Over the years, there have been several extensions to the state switching modelling by introducing nonlinear structures such include Chow and Zhang (2013), Johnson *et al.* (2024), Dong *et al.* (2020), and Farnoosh *et al.* (2021). The regime switching state space model was proposed by Chow and Zhang (2013), the model adopts a pre-specified nonlinear transition function. The Johnson *et al.* (2024) model called SVAE, parameterizes the emission function by neural networks, while the transition function remains linear. The Dong *et al.* (2020) switching non-linear dynamical systems (SNLDS) extension model parameterizes both the emission and transition functions with nonlinear neural networks. The deep switching autoregressive factorization (DSARF) proposed by Farnoosh *et al.* (2021), approximates high-dimensional time series with a multiplication of latent factors and latent weights, where the latent weights are modeled by a nonlinear autoregressive model, switched by a Markov chain of discrete latent variables. Most of the above-mentioned studies assumed that the discrete latent variables only influence the transition of state z_t. The discrete switching variables in the SLDS are assumed to be Markov, i.e. d_t depends on d_{t-1} only. The recurrent SLDS (rSLDS) proposed in Linderman *et al.* (2017) and Becker-Ehmck *et al.* (2020) extends the open-loop Markov dynamics and makes d_t depending on the hidden state z_{t-1}. In Nassar *et al.* (2018), a tree structure prior is imposed on the switching variables of rSLDS, where the dynamics of the switching variables behave similarly in the same sub-trees. The deep Rao-Blackwellised Particle Filter proposed in Kurle *et al.* (2020) also allow d_t to depend on z_{t-1}. The SNLDS model Dong *et al.* (2020) extended the open-loop Markov dynamics by making d_t depends on last observations. Such recurrent structures serve as a presence of disturbance to the switching dynamics.

Recently, Mari and Mari (2023) introduced a regime-switching model to analyze how market prices change over time. Their model features a mean-reverting diffusion process for the basic regime, while the alternate regime is driven by predictions from a deep neural network trained on market log-returns. To estimate the model using market data, they proposed a statistical method based on simulated moments. Xiuqin *et al.* (2023) developed the deep switching state space model (DSSSM), a new framework designed to tackle challenges in modeling, inference, and understanding stochastic processes. The DSSSM aims to deliver accurate forecasts and detect hidden regimes that carry significant economic implications and deepen understanding of market dynamics. It employs discrete latent variables for regimes and continuous ones for random influences, combining an RNN with a nonlinear switching state space model to capture nonlinear dependencies and regime shifts driven by a Markov process. Xiuqin *et al.* demonstrated DSSSM's effectiveness through forecasting tests on various simulated and real datasets across sectors such as healthcare, economics, traffic, meteorology, and energy. Later, Antulov-Fantulin *et al.* (2024) proposed a new method for regime detection using a deep learning architecture called the gated recurrent straight-through unit (GRSTU). Their extensive simulations showed that the GRSTU outperformed traditional statistical jump models, especially in regime classification on smaller datasets, while performing comparably on larger datasets. From the above, it is clear that many studies have focused on modeling nonlinear time series to address issues like significant modeling misspecification or difficulty in identifying meaningful stochastic regimes. As a result, the field of deep learning, especially recurrent neural networks with gate structures such as the Long-Short Term Memory (LSTM), Gated Recurrent Unit (GRU), Transformers, and temporal convolution networks, has become the new standard for modeling complex nonlinear dependencies. However, the small size of real-world data samples and, more importantly, the stochastic nature of regime switching, make traditional deep learning methods computationally challenging. In other words, these methods require large sample sizes for reliable estimation, which is often unrealistic because many disciplines do not have large volumes of time series data. In light of these challenges, Dong *et al.* (2020), Farnoosh *et al.* (2021), and Xiuqin *et al.* (2023) developed Switching Non-Linear Dynamical Systems (SNLDS), Deep Switching Autoregressive Factorization (DSARF), and Deep State Switching Model (DSSM), respectively, to better address issues of misspecification and the difficulty in identifying meaningful stochastic regimes in nonlinear time series. Nonetheless, computing the feasibility of these models, especially with small sample sizes, remains a challenge. Therefore, this research aims to propose new neural network-based regime switching models that can improve regime prediction for nonlinear time series data, such as returns, particularly when working with small sample sizes.

MATERIALS AND METHODS

The Novel Neural Network State Switching Models

To improve the regime prediction performance of famous state switching model described in equation (1) for nonlinear time series data, especially returns, in small sample size contexts, this study integrates deep neural networks these include; Recurrent Neural Network (RNN), Radial Basis Function Network (RBF), Generalized Regression Neural Network (GRNN) and Multilayer Perceptrons (MLP) with Markov two-State Switching modelling technique. The proposed neural network state switching modeling approach combines the strengths of deep neural networks (i.e. efficient in high complexity dataset) and Markov two-state switching capabilities of ensuring both interpretability and predictability of two financial state of market returns (i.e. bull and bear states). The models are developed in two phases.

1st Phase: Generative Network using RNN, GRNN, RBF and MLP

Given time series dataset of Y_t, the generative network procedure for the dataset include the following:

i Define of training and testing dataset of Y_t i.e. by setting training dataset at time step t_k, and testing dataset at time step $t_{(t-k)}$

ii At time step t_k, either RNN, GRNN, RBF or MLP is used to process the input data (i.e. training dataset) such as $h_t \sim f_h (y_t)$ where f_h is the function of RNN, GRNN, RBF or MLP.

2nd Phase: Integration of Generative Network h_t into the Markov Two-State Switching Model

Subsequently, the generative network h_t is integrated in the Markov two-state switching model alongside the defined training dataset to develop Recurrent Neural Network State Switching Model (RNNSSM), Generalized Regression Neural Network State Switching Model (GRNNSSM), Radial Basis Function Network State Switching Model (RBFSSM) and Multilayer Perceptrons State Switching Model (MLPSSM). This would give us the following model to work with:

$Y_{(tk\sim ht)}=g_t=c_{(st)}+B_1(g_{(t-1)}-c_{(st-1)})+B_2(g_{(t-2)}-c_{(st-2)})+\epsilon_t$ (2)

where

$c_{(st)}=c_0 S_{0t}+c_1 S_{1t}+c_2 S_{2t}$

$\sigma^2_{(st)}=\sigma^2_1 S_{1t}+\sigma^2_2 S_{2t}$

$\epsilon_t\sim i.i.d(0,\sigma^2_{(st)})$

$c_{(st)}$ is the state dependent mean, $\sigma^2_{(st)}$ is state dependent variance and the coefficients are B_1 or B_2; which could be different for different subsamples. The proposal will be to model the state S_t as the outcome of an unobserved two-state Markov chain with S_t independent of ϵ_t for all t. The transitions of the S_t, are presumed to be ergodic and intricate first order Markov-process. This means impacts of earlier observation(s) for the g_t and state(s) is/are completely captured in the recent g_t state(s) observations as represented in (3);

$\varrho_{ij}=Prob(S_t=j/S_{(t-1)}=i)\ \forall i, j=1, 2\sum^2_{(i=1)}\varrho_{ij}=1$ (3)

Matrix P captures the probability of switching which is known as a transition matrix;

$$\rho_{ij}=\begin{bmatrix}\rho_{11} & \rho_{12}\\ \rho_{21} & \rho_{22}\end{bmatrix} \quad (4)$$

where $P_{11}+P_{21}=1$, and $P_{12}+P_{22}=1$

The nearer the probability ϱ_{ij} is to one the longer it takes to shift to the next regime.

Consider the model given by equation (2), i.e. a Markov regime-switching model with 2-regimes. The estimation will be performed using Hamilton's filter, where the main idea is to calculate each state's filter probabilities by making inferences on each state's unknown probabilities based on the available information. When the filter probabilities are obtained, we have the probabilities one needs for calculating the log likelihood of the model. Subsequently, the estimation of the model Filter Probabilities are discussed.

The model's filter probabilities are calculated by utilizing the model's iterative relations by means of recursion. This can be done using a combination of the relation between observations and hidden states, and the endogenous relation between hidden states. Begin from the starting value in our recursion, i.e. with the probability of being in state i at time t = 1:

$$\alpha_{i,1}=P_r(S_1=i|y_1)$$
$$=\frac{f(S_1=i,y_1)}{f(y_1)}$$
$$=\frac{f(y_1|S_1=i)P_r(S_1=i)}{\sum_{j=1}^2 f(S_1=i,y_1)}$$
$$=\frac{f(y_1|S_1=i)P_r(S_1=i)}{\sum_{j=1}^2 f(y_1|S_1=j)P_r(S_1=j)}$$

The second element of the numerator is simply the previously mentioned initial probability of the Markov chain, i.e. $P_r(S_1=i)=\pi_i$, and it will henceforth be denoted as such. One can at this point notice that $\alpha_{(i,1)}$ is the normalized value of the product between the initial probability and the conditional probability function $f(Y_1|S_1=i)$, and can therefore be written as follows:

$$\alpha_{i,1}=\frac{f(Y_1|S_1=i)\pi_i}{\sum_{j=1}^2 f(Y_1|S_1=i)\pi_j}=\lceil(Y_1|S_1=i)\pi_i\rceil$$

Now, assume that we know the filter probability at time t-1, namely $\alpha_{(i,t-1)}$. Following the same strategy as for t=1 leads to the following recursion:

$$\alpha_{i,1}=P_r(S_1=i|y_{1:t})=\frac{f(S_t=i,y_t|y_{1:t-1})}{f(y_t|y_{1:t-1})}$$
$$=\frac{f(S_t=i,y_t|y_{1:t-1})}{\sum_{j=1}^2 f(S_t=j,y_t|y_{1:t-1})}$$
$$=\lceil f(S_t=i,y_t|y_{1:t-1})\rceil$$
$$=\lceil f(y_t|S_t=i)P_r(S_t=i|y_{1:t-1})\rceil$$

Data Source: Simulation Setups

To demonstrate the main idea behind the developed neural network state switching models towards enhancement of the prediction level of nonlinear time series data such as market returns particularly in small sample sizes, we consider a two states (regimes) model demonstrating the Bull and Bear regimes of daily market returns.

$Y_{(t=1)}=\mu_1+\epsilon_t$; where $\epsilon_t\sim(N,\sigma^2_1)$ for state 1

$Y_{(t=2)}=\mu_2+\epsilon_t$; where $\epsilon_t\sim(N,\sigma^2_2)$ for state 2

Parameters Settings

μ_1= 0.01; is the mean returns of bull regime,

μ_2= -0.02; is the mean returns of bear regime,

σ_1= 0.05; is the standard deviation of bull regime of the market returns,

σ_2= 0.1; is the standard deviation of bear regime of the market returns,

matrix(c(0.9,0.1,0.2,0.8)= is the transition matrix for the aforementioned regimes,

t = n is categories of small sample sizes sets at 30, 50, 70, and 100.

Number of replication r = 1000times.

Models' Estimation Procedure and Prediction Performance Evaluation

Prior to the estimation of proposed neural network state switching models (i.e. RNNSSM, GRNNSSM, RBFSSM and MLPSSM) against the traditional Markov State Switching Model (SSM), the simulated market returns (Y_t) are tested for nonlinear modelling suitability. Nonlinear models are employed where the financial system suggests nonlinearity in the system (Adejumo *et al.*, 2020; Mendy & Widodo, 2018). We utilized the most widely used tests known as BDS test by Brock, Dechert and Scheinkman. Also, the simulated returns were tested for stationarity and presence of volatility using Augmented Dickey-Fuller (ADF) test and ARCH test respectively.

Also, this study utilized the frequently used model selection principle - Akaike Information Criterion (AIC). AIC = T ln(residual sum of squares) + 2n, where T is the numeral of operational observations, while n is the number of parameters to be evaluated. The log-likelihood of the fitted model was also utilized. Hence, the outperform model is one with the smallest AIC value and highest log-likelihood.

Frequent error measures are available for model prediction evaluation; we evaluate the prediction ability of the proposed models against the Markov State Switching Model by means of several performance measures such as Root Mean Square Error (RMSE) and Mean Absolute Error (MAE).

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(A_t - F_t)^2}$$

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|(A_t - F_t)|$$

where A_t is the actual value in time t, and F_t is the prediction value in time t. The models' performances would be assessed using the earlier described Monte Carlos Simulation Study under four categories of small sample sizes.

RESULTS AND DISCUSSION

Simulated Data Presentation and Preliminary Assessment

This section presents and discusses the summary statistics and some time series features of simulated market returns data. Figure 2 and Fig 3 reveal the time series trend plots and the regimes plots of the simulated daily market returns. Explicitly, Fig 2 depicts the time series and regimes plots of simulated daily market returns of sample sizes 30 and 50 while Fig 3 shows the time series and regimes plots of simulated daily market returns of sample 70 and 100. As observed, all the sample sizes' time series trend plots depict relative stationarity i.e. constant means, variances and autocorrelation however relative to some external factors or trend. Also, the time series plots depict clustering feature within the plots indicating presence of volatility.

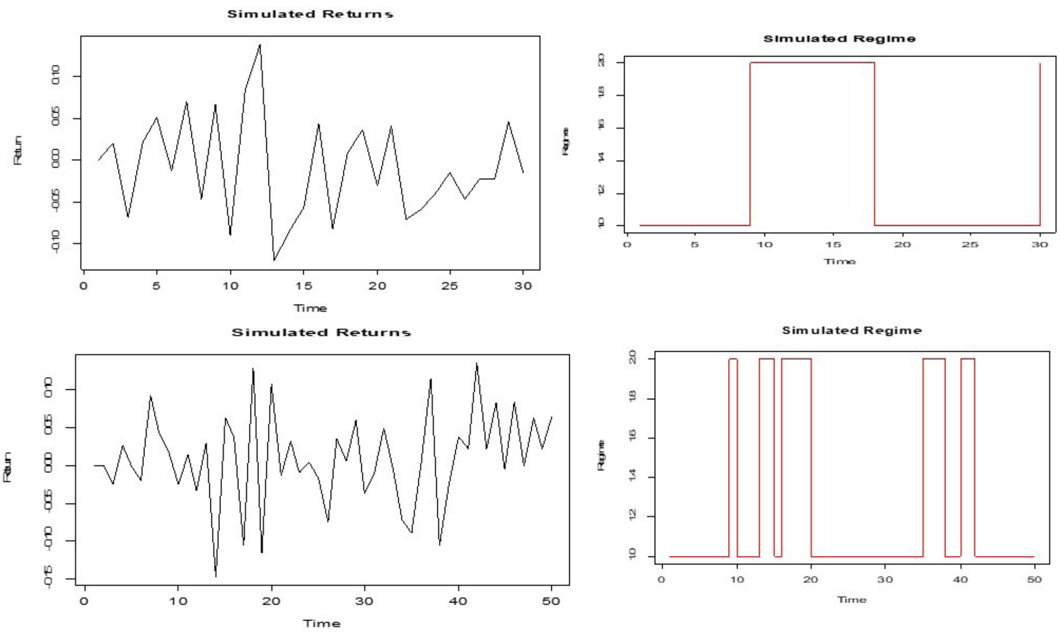


Figure 2: Time Series Plots of Simulated Daily Market Returns of Sample Sizes 30 and 50

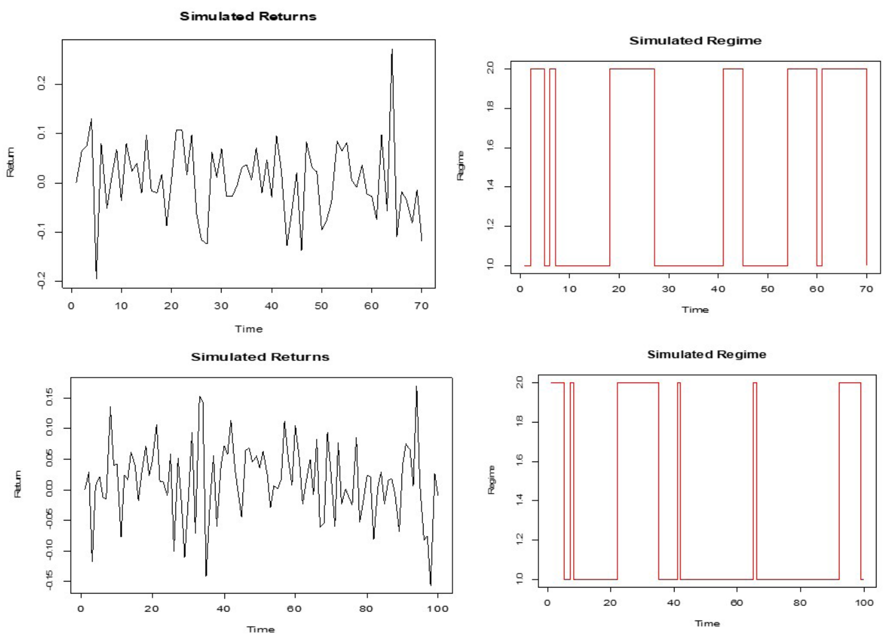


Figure 3: Time Series Plots of Simulated Daily Market Returns of Sample Sizes 70 and 100

Furthermore, Table 1 depicts the summary statistics, stationarity test, non-linearity test and volatility tests of the simulated market returns data over four (4) categories of small sample sizes. According to the table, the simulated market returns returned average returns of -0.0085±0.06, 0.0092±0.06, 0.0047±0.077 and 0.0153±0.06 for sample size 30, 50, 70 and 100 respectively. Based on the aforesaid, it can be observed that the simulated data depicts a relative uniform variation (i.e. standard deviation at around 0.06, which is close to induced 0.05) of the simulated returns across the sample sizes. Also, consequent to the observed features in the time series plots, the simulated data were tested for stationarity, non-linearity and presence of volatility. According to the ADF test results, all the returns series as expected are stationary at level. As observed in the BDS test results, the p-values depict significant at 0.05 level suggesting the non-acceptance of linearity assumption for the simulated returns series across all the sample sizes. Similarly, the ARCH test p-values results reveal significant at 0.05 level suggesting the non-acceptance of assumption of no presence of volatility in the simulated returns series across all the sample sizes. Based on the foregoing, is quite evident that the simulated returns at all the sample sizes significantly exhibit non-linearity and volatility features. Hence, the simulated returns data necessitate non-linear and volatility models such as the State Switching Model and novel neural network state switching models.

Table 1: Summary Statistics and Preliminary Tests of the Simulated Data

	30	50	70	100
Summary Statistics				
Mean	-0.0085	0.0092	0.0047	0.0153
Median	-0.0150	0.0053	0.0063	0.0174
Min.	-0.1203	-0.1470	-0.1935	-0.1567
Max.	0.1387	0.1349	0.2717	0.1699
Std. Dev.	0.0604	0.06199	0.0772	0.0618
Preliminary Tests				
ADF Test	I(0) 0.00*	I(0) 0.00*	I(0) 0.00*	I(0) 0.00*
BDS Test	0.0302*	0.0310*	0.0386*	0.0309*
ARCH Test	0.0068*	0.0272*	3.21e-06*	4.70e-06*

*Note: * denotes significant at 0.05 level*
Source: Researchers' Compilations from R-Output

Models Estimations

This section presents and discusses the estimations of the novel models and the traditional SSM. The models were fitted for training dataset that consisted of 80% of the simulated returns at different sample sizes. The diagnosis of the goodness of fit of the estimated models for the return series depicts, the Q–statistics p-values greater than 0.05, indicating that there is no statistically significant trace of dependency or autocorrelation left in the squared standardized residuals, indicating that all the estimated models are adequately specified. Table 2 present the summary of models' estimations for the simulated market returns at a sample size of 30. According to Table 2, among the estimated models the novel MLPSSM and RBFSSM returned with the least AIC of -314.6393 and -115.7039, respectively as well as higher log-likelihood of 161.3196 and 61.8519 respectively. Thus, MLPSSM returns to be the most parsimonious model among the estimated SSMs.

Table 2: Summary of Models Estimations at 30 Sample Size

		SSM	**RNNSSM**	**RBFSSM**	**GRNNSSM**	**MLPSSM**
Regime 1	Intercept	0.0461*	-0.4385	589.59*	0.0087	-1437.70
	Training_Network	-	rnn= 0.7062*	rbf=-1.18e+07*	grnn=1.1443	mlp=1443.53*
Regime 2	Intercept	-0.0626*	0.2849	-922.40*	0.0754	-1443.05
	Training_Network	-	rnn=-0.4932	rbf=1.84e+7*	grnn=0.6202	mlp=1448.90*
Transition Prob.	State 1 State 2	0.3484 0.6578	0.3461 0.3138	0.3423 0.6648	0.000 0.7007	0.9468 0.2294
	State 2 State 1	0.6516 0.3422	0.6539 0.6862	0.6577 0.3352	1.000 0.2992	0.0532 0.7705
AIC		-64.31498	-59.7527	-115.7039	-56.3426	-314.6393
logLik		34.1575	33.8764	61.8519	32.1713	161.3196

*Note: * denotes significant at 0.05 level*
Source: Researchers' Compilations from R-Output

In addition, Table 2 depicts the transition probabilities for the two identified states (namely bull and bear). The transition probabilities for MLPSSM model show that there is a high probability that the market returns' system remains in the same state hence implying limited switches in the state or regime. The results also indicate that the MLPSSM has an 95% probability of staying in the bull state and a 5% probability of switching to the bear state. When the system is in a bear state, it has a 77% probability of remaining in bear state and lower probability of 23% to switch to bull state. The transition probability results highlighted shows that only extreme/great events can switch the returns between states i.e. state 1(bull phase) to state 2 (bear phase), (Figure 4). It further indicates that not any of the state is lasting since all transition probabilities are below one.

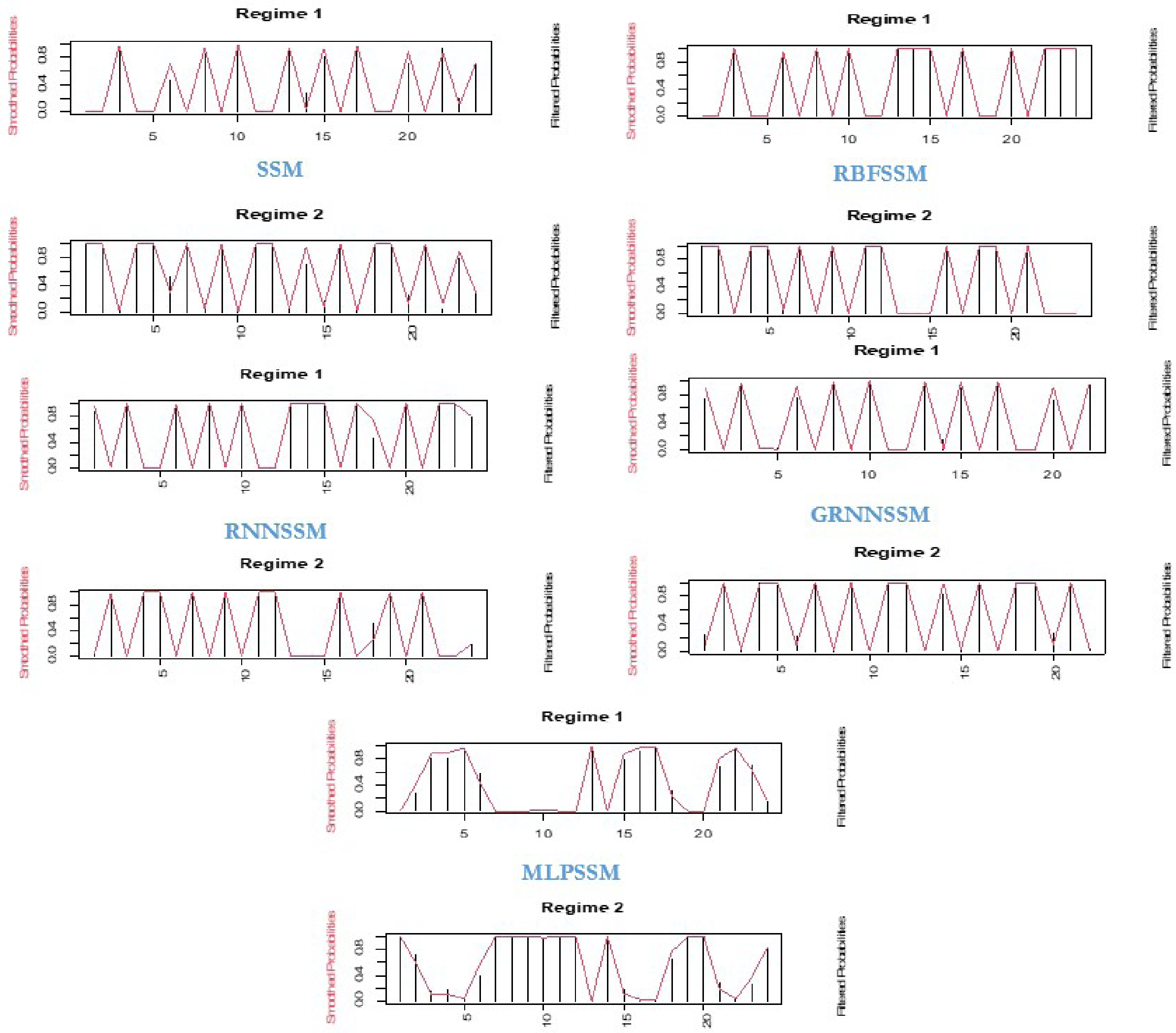


Figure 4: Models' Transition Probabilities of Returns at Sample Size 30

Furthermore, Table 3 depicts the summary of models' estimations for the simulated market returns at sample size of 50. The table reveals MLPSSM and RBFSSM returned with the least AIC values of -543.7162 and -190.5718, respectively as well as higher log-likelihood values of 275.8581 and 99.2859 respectively. Similarly, the results returned MLPSSM with the lowest AIC and highest log-likelihood values as the most parsimonious model among the estimated SSMs at sample size 50. Additionally, Table 3 illustrates the transition probabilities between the two identified states: bull and bear.

The transition probabilities for the MLPSSM model suggest a low likelihood of the market return system remaining stable, indicating frequent transitions between bull and bear states. Specifically, the results show that the MLPSSM has a 32% chance of remaining in the bull state and a 68% chance of transitioning to the bear state. In a bear state, there is a 38% probability of staying in that state, while the chance of switching to the bull state is lower at 62%. The highlighted transition probability results indicate that minor events can trigger changes in returns between states, such as from state 1 (bull phase) to state 2 (bear phase) or the other way around (see Figure 5). This suggests that neither state is permanent since all transition probabilities are less than one.

Table 3: Summary of Models Estimations at 50 Sample Size

		SSM	RNNSSM	RBFSSM	GRNNSSM	MLPSSM
Regime 1	Intercept	0.0032	0.0851	-505.86*	0.0342	-4218.13*
	Training_Network	-	rnn= -0.1698	rbf= 2.02e+07*	grnn=-0.1689	mlp=4231.02*
Regime 2	Intercept	-0.0036	-0.3103	-555.89*	0.0295	-4167.55*
	Training_Network	-	rnn=0.6503	rbf= 2.22e+07*	grnn=-1.4262	mlp=-4180.3*
Transition Prob.	State 1 State 2	0.7489 0.2135	0.5193 0.9275	0.7376 0.7779	0.522 0.9997	0.3202 0.3813
	State 2 State 1	0.2511 0.7865	0.4807 0.0725	0.2624 0.2221	0.478 0.0003	0.6798 0.6187
AIC		-110.943	-108.51	-190.5718	-98.6980	-543.7162
logLik		57.4715	58.2549	99.2859	53.3490	275.8581

*Note: * denotes significant at 0.05 level*
Source: Researchers' Compilations from R-Output

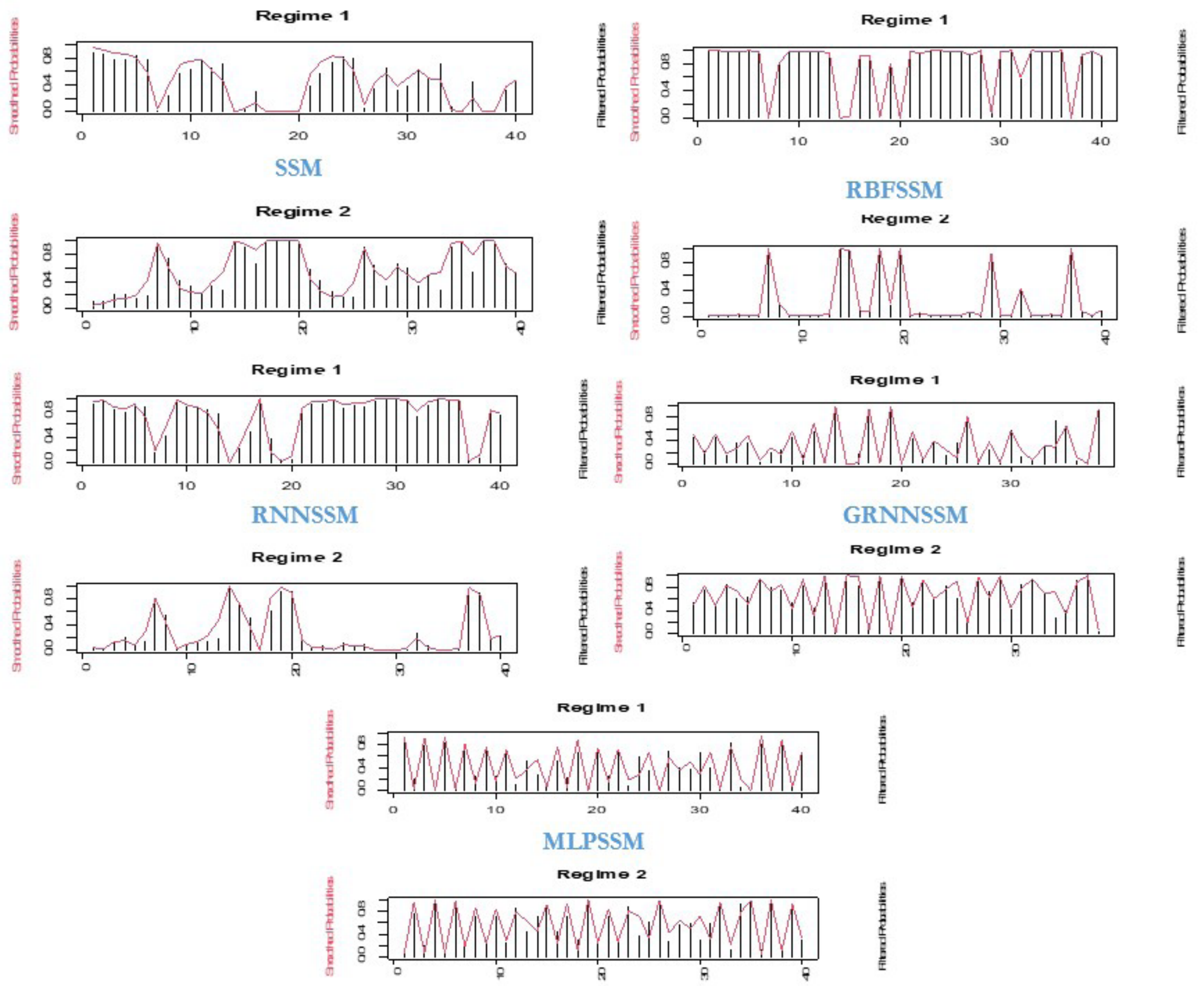


Figure 5: Models' Transition Probabilities of Returns at Sample Size 50

Moreover, Table 4 shows the summary of models' estimations for the simulated market returns at sample size of 70. The table reveals MLPSSM and RBFSSM returned with the least AIC values of -659.263 and -320.2873 respectively as well as higher log-likelihood values of 333.6315 and 164.1436 respectively. Similarly, the results returned MLPSSM with lowest AIC and highest log-likelihood as the most parsimonious model among the estimated SSMs at sample size 70.

Subsequently, Table 4 depicts the transition probabilities for the two identified states (namely bull and bear). The transition probabilities for MLPSSM model show that there is a high probability that the market returns' system remains in the same state hence implying low or limited switches in the bull state and bear state. Explicitly, the results indicate that the MLPSSM has an 86% probability of staying in the bull state and a 14% probability of switching to the bear state. When the system is in a bear state, it has a 65% probability of remaining in bear state and lower probability of 35% to switch to bull state. The transition probability results highlighted shows that only extreme/great events can switch the returns between states i.e. state 1(bull phase) to state 2 (bear phase) or vice versa, (see Figure 6). These results therefore, indicate that no any of the state is lasting since all transition probabilities are below one.

Table 4: Summary of Models Estimations at 70 Sample Size

		SSM	RNNSSM	RBFSSM	GRNNSSM	MLPSSM
Regime 1	Intercept	0.0330*	0.0509	4.32+02*	0.1226*	-12577.37*
	Training_Network	-	rnn= -0.0493	rbf= -1.7e+07*	grnn=-6.6674	mlp=12597.9
Regime 2	Intercept	-0.1010*	0.0469*	3.36e+02*	0.0117	-12094.81
	Training_Network	-	rnn=-0.3404*	rbf= -1.34e+07*	grnn=-4.0712	mlp=12114.5
Transition Prob.	State 1 State 2	0.8561 0.6539	0.9024 0.7165	0.5097 0.1724	0.0079 0.5004	0.8610 0.6472
	State 2 State 1	0.1439 0.3461	0.0976 0.2835	0.4903 0.8276	0.9921 0.4996	0.1390 0.3528
AIC		-140.2999	-153.0316	-320.2873	-133.6245	-659.263
logLik		72.1499	80.51579	164.1436	70.8122	333.6315

*Note: * denotes significant at 0.05 level*
Source: Researchers' Compilations from R-Output

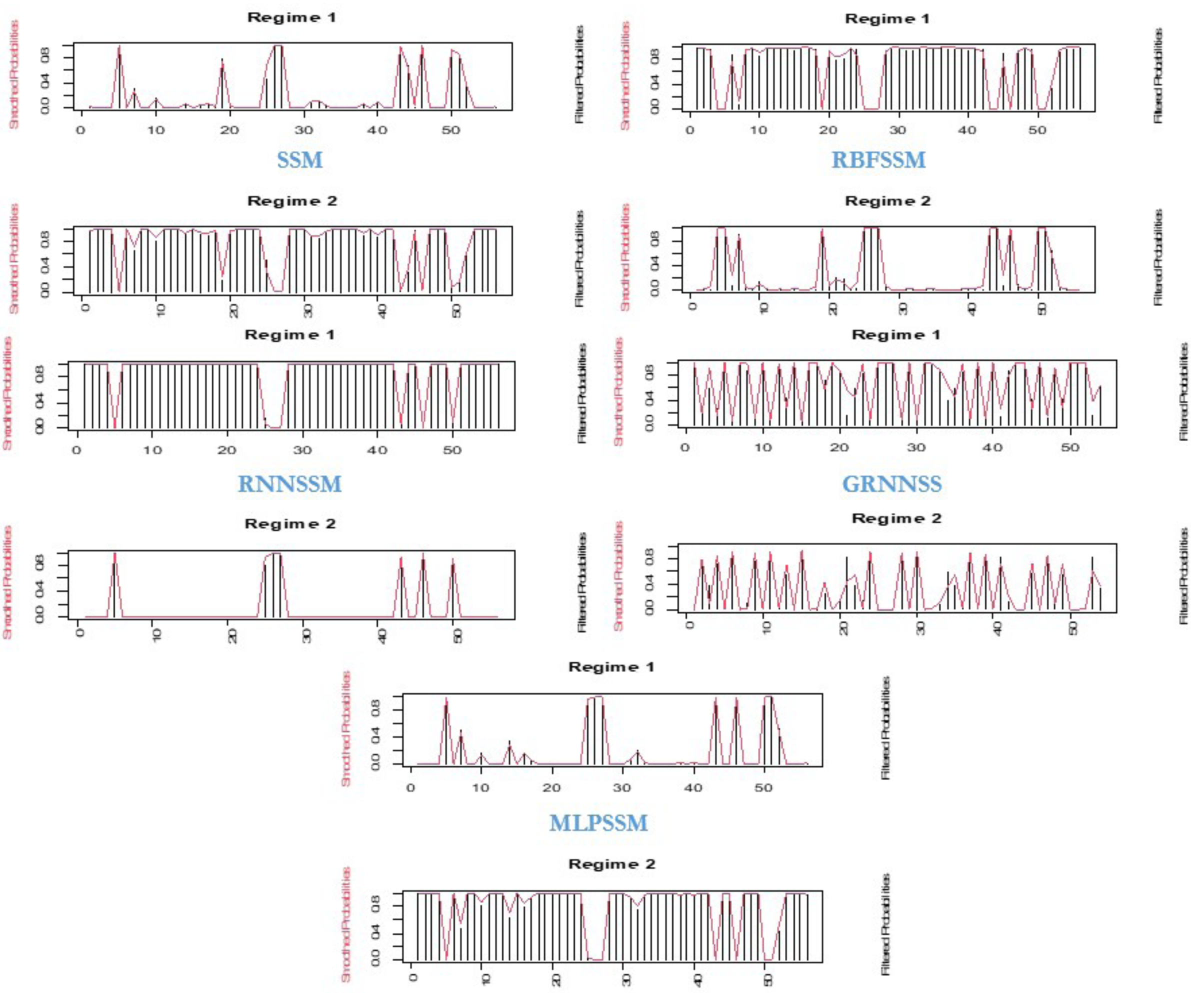


Figure 6: Models' Transition Probabilities of Returns at Sample Size 70

To conclude, Table 5 presents the summary of models' estimations for the simulated market returns at sample size of 100. Table 5 depicts MLPSSM and RBFSSM returned with the least AIC values of -594.7593 and -387.7917 respectively as well as higher log-likelihood values of 301.3796 and 197.8958 respectively. Similarly, the results returned MLPSSM with lowest AIC and highest log-likelihood as the most parsimonious model among the estimated SSMs at sample size 100.

Table 5: Summary of Models Estimations at 100 Sample Size

		SSM	RNNSSM	RBFSSM	GRNNSSM	MLPSSM
Regime 1	Intercept	0.0395*	-0.1989	6.13+02*	-0.2097*	-114443.95*
	Training_Network	-	rnn= 0.6103*	rbf= -2.5e+07*	grnn=3.6865*	mlp=114617.88
Regime 2	Intercept	-0.0100	0.1085*	-8.82e+02*	0.0395	-94142.78
	Training_Network	-	rnn=-0.2464	rbf= 3.53e+07*	grnn=-0.1143	mlp=9428.86
Transition Prob.	State 1 State 2	0.5108 0.7283	0.3991 0.3916	0.6214 0.7046	1.4e-07 0.1696	0.5705 0.5839
	State 2 State 1	0.4892 0.2717	0.6009 0.6084	03786 0.2954	0.9999 0.8304	0.4295 0.4160
AIC		-223.3181	-226.2772	-387.7917	-218.4201	-594.7593
logLik		113.659	117.1386	197.8958	113.2101	301.3796

*Note: * denotes significant at 0.05 level*
Source: Researchers' Compilations from R-Output

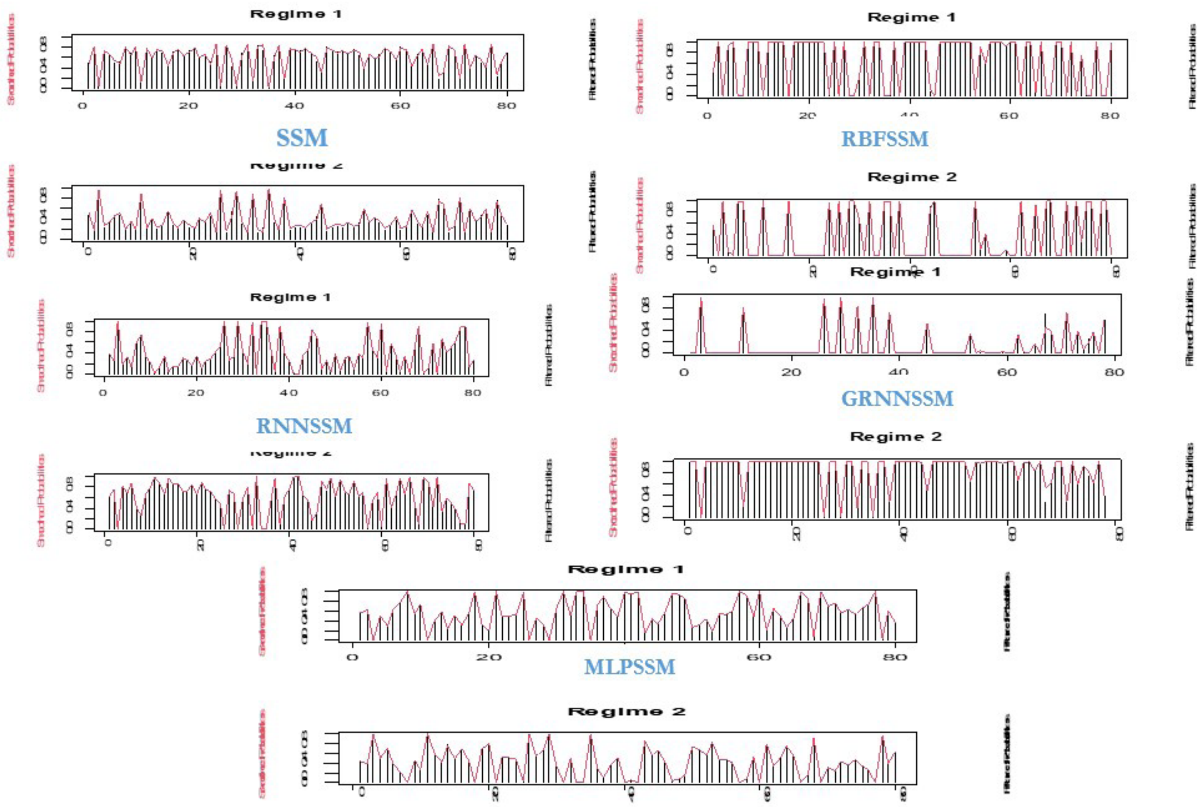


Figure 7: Models' Transition Probabilities of Returns at Sample Size 100

Subsequently Table 5 depicts the transition probabilities for the two identified states (namely bull and bear). The transition probabilities for MLPSSM model show that there is a fair probability that the market returns' system remains in the same state hence implying almost equal chances of state switching. Explicitly, the results indicate that the MLPSSM has an 57% probability of staying in the bull state and a 43% probability of switching to the bear state. When the system is in a bear state, it has a 58% probability of remaining in bear state and lower probability of 42% to switch to bull state. The transition probability results highlighted shows that either weak or extreme/great events can switch the returns between states i.e. state 1(bull phase) to state 2 (bear phase) or vice versa, (Figure 7). These results also, indicate that no any of the state is lasting since all transition probabilities are less than one.

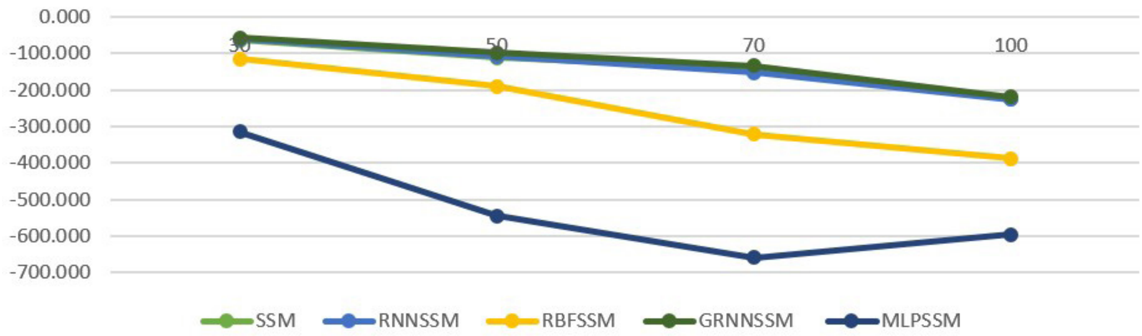


Figure 8: Models' AIC Estimates across the Sample Sizes

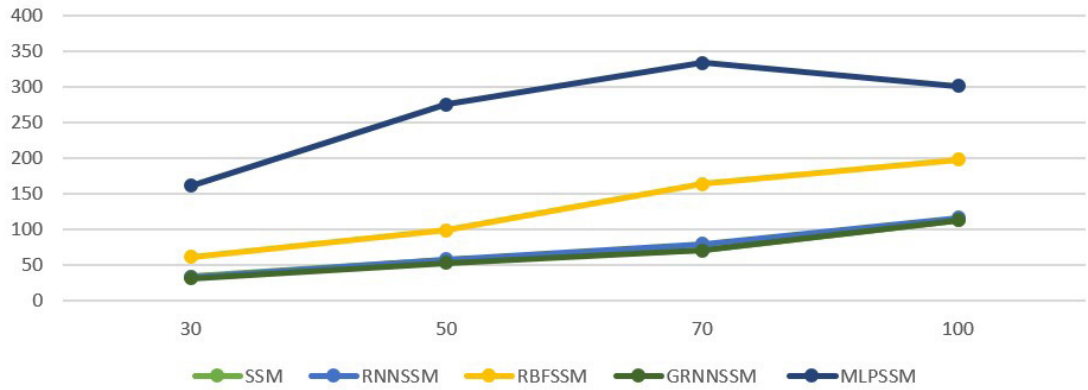


Figure 9: Models' Loglikelihood Estimates across the Sample Sizes

Consequently, Figures 8 and 9 present line plots for each of the models' AIC and Log-likelihood estimates, respectively. As observed from the figures, two out of four introduced novel models i.e. RBFSSM and MLPSSM outperformed the traditional SSM and the rest introduced models by returning with lower AIC and higher log-likelihood estimates across all four small sample sizes examined. Furthermore, according to Figure 8, the novel MLPSSM returned the least AIC estimates across all four small sample sizes examined. Similarly, as observed in Figure 9, the novel MLPSSM returned the highest Log-likelihood estimates across all four small sample sizes examined. Also, it is important to note that the AIC estimate inclines to increase and Log-likelihood estimate tends to decline when the sample size is 100. This result suggests that the novel MLPSSM is best fitted for sample sizes less than 100. Thus, the novel MLPSSM returned as the most parsimonious model among the estimated SSMs. Evidence from the AIC and log-likelihood line plots established Multilayer Perceptrons State Switching Model (MLPSSM) as the best-fitted model for the simulated returns at all the considered small sample sizes.

Evaluation of Models' Regime Prediction Performances

This section presents and discusses the performance of the fitted models for the training dataset. Each of the model performances was assessed based on in-sample predictions of the testing dataset at 20% of each sample size, i.e., 30, 50, 70, and 100. Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) were employed to assess the prediction accuracy errors of the models. Figure 10 depicts the models' prediction performances results. According to the figure, the novel MLPSSM returned the least RMSE and MAE for the in-sample predictions across all four sample sizes examined. Therefore, our introduced novel model, i.e., MLPSSM, returned as the optimal model for regime predictions of the stimulated returns.

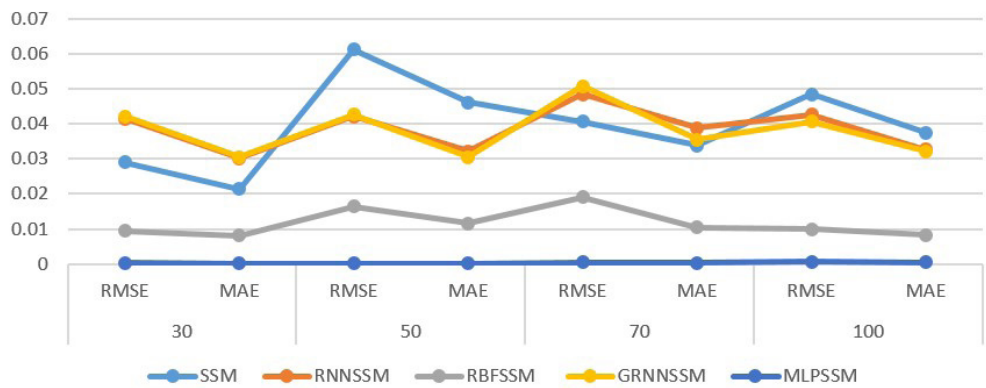


Figure 10: In-Sample Prediction Performances of the Fitted Models

Table 6: MLPSSM 20% In-Sample Daily Regime Mean Predictions of Small Sample Sizes Considered

	30 P_12=0.23,P_22=0.77	**50 P_12=0.38,P_22=0.62**	**70 P_11=0.86,P_21=0.14**	**100 P_11=0.57,P_21=0.43**
1	-0.00746	-0.00068	+0.007425	+0.01948
2	-0.00746	-0.00068	+0.007425	+0.01948
3	-0.00746	-0.00068	+0.007425	+0.01948
4	-0.00746	-0.00068	+0.007425	+0.01948
5	-0.00746	-0.00068	+0.007425	+0.01948
6	-0.00746	-0.00068	+0.007425	+0.01948
7		-0.00068	+0.007425	+0.01948
8		-0.00068	+0.007425	+0.01948
9		-0.00068	+0.007425	+0.01948
10		-0.00068	+0.007425	+0.01948
11			+0.007425	+0.01948
12			+0.007425	+0.01948
13			+0.007425	+0.01948
14			+0.007425	+0.01948
15				+0.01948
16				+0.01948
17				+0.01948
18				+0.01948
19				+0.01948
20				+0.01948

Note: + and - denote Bull and Bear Regime respectively
Source: Researchers' Compilations from R-Output

Furthermore, Table 6 presents the MLPSSM in-sample daily regime mean predictions of 20% of the simulated market returns sample sizes. The table reveals a stable or static regime means across all the sample sizes considered. Explicitly, for sample sizes 30 and 50, the novel MLPSSM predicted respectively a bear regime mean of -0.00746 and -0.00068 across the days predicted at each scenario. The predicted bear regime in sample size 30 and 50 returned high transition probability (i.e. 0.77 and 0.62) to bull regime. This implies the predicted bear regimes in sample size 30 and 50 has little probability to remain in the state, i.e., weak events can switch the bear regime to bull regime. On the other hand, for sample sizes 70 and 100, the novel MLPSSM predicted respectively a bull regime mean of +0.007425 and +0.01948 across the days predicted at each scenario. The predicted bull regime in sample size 70 returned a very low transition probability (i.e. 0.14) to bear regime while the predicted bull regime in sample size 100 returned a relatively moderate transition probability (i.e. 0.43) to bear regime. This implies the predicted bull regimes in sample size 70 has a great probability to remain in the state, i.e., only extreme events can switch the bull regime to bear regime. Meanwhile, for sample size 100, the result implies that the predicted bull regimes has a relatively fair or moderate probability of remaining in the bull state, i.e. not too weak or too extreme events can switch the bull regime to bear regime.

CONCLUSIONS

This paper centers on introducing novel Neural Network State Switching Models using the combination of strengths of deep learning neural networks and traditional Markov state switching approaches for regime predictions of time series returns. The paper shows the comparative results of the estimations of the novel models as well as the regime prediction performances of the novel models using prediction accuracy measures under a Monte Carlo simulation study.

Evidence from the models' estimation results particularly the AIC and Log-likelihood statistics established the Novel Multilayer Perceptrons State Switching Model (MLPSSM) as the best fitted model (against the traditional Markov state switching models and other introduced neural network state switching models) for the simulated returns at small sample sizes especially for sample sizes lower than 100.

Moreover, results from the models' regime predictions evaluation precisely the RMSE and MAE evidently affirmed the novel MLPSSM model as superior in its ability to predict or forecast market returns particularly the bull and bear regimes at small sample sizes. Base on the aforementioned, this study therefore concludes that the novel MLPSSM is best fitted model for market returns at small sample sizes with excellent ability of market returns' regimes/states (i.e. bull and bear) prediction.

In view of the aforementioned, this study recommends adoption of the novel Multilayer Perceptrons State Switching Model in modelling and regime predictions of time series returns.

REFERENCES

Adejumo, O. A., Albert, S., & Asemota, O. J. (2020). An Analysis of Stock Market Returns in Nigeria. *CBN Journal of Applied Statistics, 11*(2).

Aliyu, S. U. R. & Wambai, A. A. (2018). Economic regimes and stock market performance in Nigeria: Evidence from regime switching model. *Munich Personal RePEc Archive MPRA, 91430*. Retrieve from https://mpra.ub.uni-muenchen.de/91430/

Al-Sulaiman, T. (2024). Review of Recent Research Directions and Practical Implementation of Low-Frequency Algorithmic Trading. *American Journal of Financial Technology and Innovation, 2*(1), 1-14.

Altuğ, S., & Bildirici, M. (2010). *Business cycles around the globe: A regime switching approach* (Working Paper No. 1009).

Antulov-Fantulin, N., Cauderan, A., & Kolm, P. N. (2024). A Dynamic Regime-Switching Model Using Gated Recurrent Straight-Through Units. *Journal of Financial Data Science, 6*(4).

Becker-Ehmck, P., Karl, M., Peters, J., & van der Smagt, P. (2020). *Learning to fly via deep model-based reinforcement learning*. arXiv preprint arXiv:2003.08876.

Bekaert, G., & Hodrick, R. J. (1992). Characterizing predictable components in excess returns on equity and foreign exchange markets. *The Journal of Finance, 47*(2), 467-509.

Bollen, N. P., Gray, S. F., & Whaley, R. E. (2000). Regime switching in foreign exchange rates:: Evidence from currency option prices. *Journal of Econometrics, 94*(1-2), 239-276.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics, 31*(3), 307-327.

Chen, L., M. Pelger, & J. Zhu (2019). Deep learning in asset pricing. Available at SSRN3350138

Chow, S. M., & Zhang, G. (2013). Nonlinear regime-switching state-space (RSSS) models. *Psychometrika, 78*(4), 740-768.

Christensen, K., Siggaard, M., & Veliyev, B. (2023). A machine learning approach to volatility forecasting. *Journal of Financial Econometrics, 21*(5), 1680-1727.

Cont, R. (2007). Long memory in economics. In *Returns clustering in financial markets: Empirical facts and agent-based models* (pp. 289–309). Springer.

Dong, H., Chen, L., Zhang, X., Failler, P., & Xu, S. (2020). The asymmetric effect of volatility spillover in global virtual financial asset markets: the case of Bitcoin. *Emerging Markets Finance and Trade, 56*(6), 1293-1311.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the econometric society,* 987–1007. https://doi.org/10.2307/1912773

Farnoosh, A., Azari, B., & Ostadabbas, S. (2021, May). Deep switching auto-regressive factorization: Application to time series forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 35, No. 8, pp. 7394–7403). AAAI Press.

Frömmel, M., MacDonald, R., & Menkhoff, L. (2005). Markov switching regimes in a monetary exchange rate model. *Economic Modelling, 22*(3), 485-502.

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies, 33*(5), 2223–2273.

Hamilton, J. D. (1989). A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica, 57*(2), 357–384

Johnson, P., Szabó, D. Z., & Duck, P. (2024). Optimal trading with regime switching: Numerical and analytic techniques applied to valuing storage in an electricity balancing market. *European Journal of Operational Research, 319*(2), 611-624.

Kurle, R., Rangapuram, S. S., de Bézenac, E., Günnemann, S., & Gasthaus, J. (2020). Deep rao-blackwellised particle filters for time series forecasting. *Advances in Neural Information Processing Systems, 33,* 15371-15382.

Liao, R., Yamaka, W., & Sriboonchitta, S. (2020). Exchange rate volatility forecasting by hybrid neural network Markov switching Beta-t-EGARCH. *IEEE Access, 8,* 207563-207574.

Maheu, J. M., & McCurdy, T. H. (2002). Nonlinear features of realized FX returns. *Rev. Econ. Statist., 84*(4), 668-681.

Mari, C., & Mari, E. (2023). Deep learning based regime-switching models of energy commodity prices. *Energy Systems, 14*(4), 913-934.

Nassar, J., Linderman, S. W., Bugallo, M., & Park, I. M. (2018). *Tree-structured recurrent switching linear dynamical systems for multi-scale modeling*. arXiv.

Nelson, D. B. (1992). Filtering and forecasting with misspecified ARCH models I. *Journal of Econometrics, 52*(1–2), 61–90.

Nelson, D. M. Q., Pereira, A. C. M., & De Oliveira, R. A. (2017). Stock market's price movement prediction with LSTM neural networks. *International Joint Conference on Neural Networks* (IJCNN. https://doi.org/10.1109/ijcnn.2017.7966019

Nguyen, L. H., Chao, D., Lim, J. K., Ayoub, W., & Nguyen, M. H. (2014). Histologic changes in liver tissue from patients with chronic hepatitis B and minimal increases in levels of alanine aminotransferase: a meta-analysis and systematic review. *Clinical gastroenterology and hepatology, 12*(8), 1262-1266.

Owyang, M. T., & Ramey, G. (2004). Regime switching and monetary policy measurement. *Journal of Monetary Economics, 51*(8), 1577-1597.

Roy, G., & Sarkar, P. (2024). A Systematic Review on Impact of Political Crisis and Election on Capital Market Returns. *Available at SSRN 4939289*.

Rydén, T. (2008). *EM versus Markov chain Monte Carlo for estimation of hidden Markov models: A computational* *perspective*.

Shah, D., Campbell, W., & Zulkernine, F. H. (2018). A comparative study of LSTM and DNN for stock market forecasting. In *2018 IEEE International Conference on Big Data (Big Data)* (pp. 4148–4155). IEEE. https://doi.org/10.1109/BigData.2018.8622462

Sims, C. A., & Zha, T. (2006). Were there regime switches in US monetary policy?. *American Economic Review, 96*(1), 54-81.

Sunny, M. A. I., Maswood, M. M. S., &Alharbi, A. G. (2020). Deep learning-based stock price prediction using LSTM and bi-directional LSTM model. *2020 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES),* 87–92. https://doi .org/10.1109/ NILES50944.2020.9257950

Valente, G. (2003). Monetary policy rules and regime shifts. *Applied Financial Economics, 13*(7), 525-535.

Xiuqin, X., Hanqiu, P., & Ying, C. (2023). *Deep switching state space model (DSSSM) for nonlinear time series forecasting with regime switching*. arXiv preprint arXiv:2106.02329. https://arxiv.org/abs/2106.02329

Yao, S., Luo, L., &Peng, H. (2018). High-frequency stock trend forecast using LSTM model. *13th International Conference on Computer Science & Education* (ICCSE), 1–4. https://doi.org/10.1109/ICCSE.2018.8468703

Zucchini, W., & MacDonald, I. L. (2009). *Hidden Markov models for time series: An introduction using* R. Chapman and Hall/CRC.