A Variable-Order Regime Switching Model to Identify Significant Patterns in Financial Markets

Philippe Chatigny*, Rongbo Chen*, Jean-Marc Patenaude† and Shengrui Wang*
*Department of Computer Science, Université de Sherbrooke, Sherbrooke, Québec, Canada
†Laplace Insights, Sherbrooke, Québec Canada
Email: {Philippe.Chatigny, Rongbo.Chen, Shengrui.Wang}@Usherbrooke.ca, JeanMarc@Laplaceinsights.com

***Abstract*—The identification and prediction of complex behaviors in time series are fundamental problems of interest in the field of financial data analysis. Autoregressive (AR) model and Regime switching (RS) models have been used successfully to study the behaviors of financial time series. However, conventional RS models evaluate regimes by using a fixed-order Markov chain and underlying patterns in the data are not considered in their design. In this paper, we propose a novel RS model to identify and predict regimes based on a weighted conditional probability distribution (WCPD) framework capable of discovering and exploiting the significant underlying patterns in time series. Experimental results on stock market data, with 200 stocks, suggest that the structures underlying the financial market behaviors exhibit different dynamics and can be leveraged to better define regimes with superior prediction capabilities than traditional models.**

I. INTRODUCTION

The basic idea that a time series exhibits certain behaviors and transits between different states has been at the core of many time-series models. In finance, studying the market behaviors is of great interest since it allows to establish links between market dynamics and the general state of the market. Many phenomena observed in the financial market, such as fat tails, volatility clustering and co-movement, have been studied to explore the financial behaviors in various financial markets [1][2][3]. Regime switching models [4][5] have been used to characterize these complex behaviors in a wide range of applications. These studies suggest that structural breaks in the time series lead to a regime switch, where a regime expresses some kind of behaviors that explains the market dynamic over a period of time.

In practice, certain characteristics of time series can be hard to observe when time series exhibit non-linearity, mixing or noise. In such cases, choosing the right RS model is difficult. This is true especially when the dynamic of the time series is unknown. Adoption of a model, such as MSGARCH, requires completion of the difficult tasks of model specification, while a wide variety of variance models exists [7]. Moreover, model evaluation is also a challenge since the evaluation functions in these models do not provide consistent assessment of model quality [15]. Finally, the existing models do not provide easy-to-interpret patterns representing the dynamics of the discovered regimes. Given these difficulties, the current RS models fail to provide a satisfactory solution to regime modelling.

In this paper, we propose a novel pattern-based regime switching model that evaluates regimes based on explicit patterns discovered using a variable order Markov chain created from a time series. The key contributions of this paper can be summarized as follows:

1) Our novel clustering-based approach allows discovering regimes with explicit representations using a variable-order Markov chain.
2) Our approach is able to learn the process governing regime switches by regime without specifying it explicitly.
3) Our experimental evaluations show that our model is superior in forecasting volatility.

The remainder of this paper is organized as follows: In Section II, we discuss the rationale for approaching the problem differently than using traditional RS models. Section III presents the proposed model. In Section IV, we present an empirical evaluations of our proposed model using time series of 200 stocks from the financial market. Finally, Section V is devoted to the discussion of our results and future work.

II. BACKGROUND AND MOTIVATION

The widely used RS model, introduced by Hamilton [6], characterizes a time series behaviors in different regimes. It is assumed that regime switches are governed by an unobservable state variable that follows a fixed-order Markov chain process. The process governing regime switch is usually characterized by multiple parameters such as the distribution's type and the function's parameters of the process itself. Techniques such as maximum likelihood are used to estimate the model parameters. The main drawbacks of these conventional RS models are as follows:

1) Conventional RS models assume that the time series is stationary. This assumption is known to be false for financial time series [10].
2) The Markov chain, determining regimes transition probabilities, is assumed to be completely independent from all other parts of the model, which is unrealistic in many cases [13].
3) To the best of our knowledge, there is no published work on the use of a variable order Markov model to study regime switch and transition probabilities between regimes. When using a fixed-order Markov process, the model assumes that the number of significant past observations is always the same for a market regime, which is not necessarily true.

978-1-5386-9159-5/18/$31.00 ©2018 IEEE
DOI 10.1109/ICDM.2018.00106

4) It is often difficult to interpret the output of a conventional RS model given the challenge of identifying market's state robustly as noted in [11].

The goal of this work is to propose an alternative way of evaluating regimes in time series. Instead of modeling regime switch by exploiting dynamics of a latent variable, we model regimes from underlying patterns in the time series. To this end, we propose a novel regime switching framework with a variable-order Markov chain, which generates behavioral patterns for each regime, allowing us to better understand their characteristics and to predict changes between regimes.

III. The WCPD-RS model

In this section, we present the weighted conditional probability distribution regime switching model (WCPD-RS) in detail. Unlike conventional RS models, we analyze the regime switch by investigating the categorical sequence obtained by transforming the time series. The overall pipeline of WCPD-RS, shown in Fig. 1, is divided into three main parts. In the first part, we transform the time series into categorical sequences that will be used for regime detection. In the second part, we detect regimes using Model-based Categorical Sequence Clustering algorithm (MCSC) [14]. The final part is the regime prediction framework derived from MCSC, which highlights the properties of each regime, such as the regime switch probabilities.

A. Time series transformation

We transform a time series $X = \{X_1 \dots X_t\}$, where t is the number of time interval observed, into a categorical sequence $S = \{s_1 \dots s_t\}$ to extract significant patterns of categories. This transformation allows discovery of more explicit patterns to explain regime switch drivers. By identifying statistically significant patterns, we will be able to find why certain dynamics are more likely to arise based on the market's state. We introduce a categorical classifier that has been developed jointly with an investment industry expert to transform the time series into a categorical one.

Our classifier transforms a daily OHLCV time series, i.e., Open, High, Low, Close Prices and Volume indicators, into a categorical time series for volatility forecasting with eight possible categorical values, $\Omega = \{A, B, C, D, E, H, I, L\}$ such that each data point s_i where $i \in [1, t]$ takes one of the possible values. Each value of Ω corresponds to a type of particular day: A = unexciting days, C, D = "clear up and down days" and E, H, I, L = "abnormal days". The classifier is summarized in algorithm (1) where we denote $X_{i,features}$ as the value of the features of the time series at time interval i. The following features are computed before classification:

1) The log daily return (Log_return).
2) A measure of price fluctuation: $Intraday\_swing = \frac{2\times(High_t - Low_t)}{High_t + Low_t}$ where $High_t$ is the highest price reached during the trading day t, whereas Low_t is the lowest price reached during that same day. We identify the 50th quantile of the $Intraday\_swing$ given the distribution of all $Intraday\_swing$ calculated $Q50_{IS}$.
3) A normality measure for volume ($Normal\_volume$) evaluated as the volume divided by the daily volume mean of the last 21 days period. From this measure, we identifiy the 50th quantile ($Q50_{NV}$), and the 80th quantile ($Q80_{NV}$) given the distribution of all $Normal\_volume$ calculated.

Algorithm 1 Stock classifier

```
Input: X = {X_1 ... X_t}
Output: S = {s_1 ... s_t}
  for i = 1 to |X| do
    if X_{i,volume} < Q50_NV then
      s_i = A
    else if X_{i,volume} < Q80_NV then
      s_i = B if Log_return ≥ 0 else C
    else
      if X_{i,Log_return} ≥ 0 then
        if X_{i,intraday_swing} ≥ Q50_IS : s_i = D
        else if X_{i,intraday_swing} < Q50_IS : s_i = E
        else : s_i = L
      else
        if X_{i,intraday_swing} ≥ Q50_IS : s_i = H
        else if X_{i,intraday_swing} < Q50_IS : s_i = I
        else : s_i = L
```

B. Modeling regime behavior

To model each regime, we use the MCSC hierarchical clustering algorithm from Xiong & al. work [14] to cluster subsequences of S retrieved by using a sliding window technique. MCSC is a top-down divisive algorithm implemented through a two-tier cascade optimization framework. The first tier of the framework is built from the first-order Markov model implemented via a weighted fuzzy indicator (WFI) matrix for the purpose of cluster splitting and optimization, while the second tier is built from the WCPD model for the purpose of cluster refinement and pattern discovery. The latter is a variable order Markov model and allows the discovery of statistically significant patterns of variable lengths for estimating the similarity between a sequence and a cluster. These patterns are retrieved natively in the WCPD model when building the generative model for each cluster, i.e., the conditional probability distribution (CPD) of each cluster by means of the probabilistic suffix tree (PST). We refer to the Xiong & al. work [14] for technical explanation of the algorithm to model individual regimes.

Contrary to the indication in [14], two adjustments are made to the WCPD model and are summarized in algorithm (1). First, we reconstruct the PST at every iteration of the WCPD but select the memory subsequences for the optimization procedure only at the beginning of the optimization. The reason for this strategy is that we have observed that the number of memory subsequences may change greatly after pruning, leading to convergence instability in the WCPD model. Second, we do not enforce the constraint that the PST is a full probabilistic tree in which every node other than

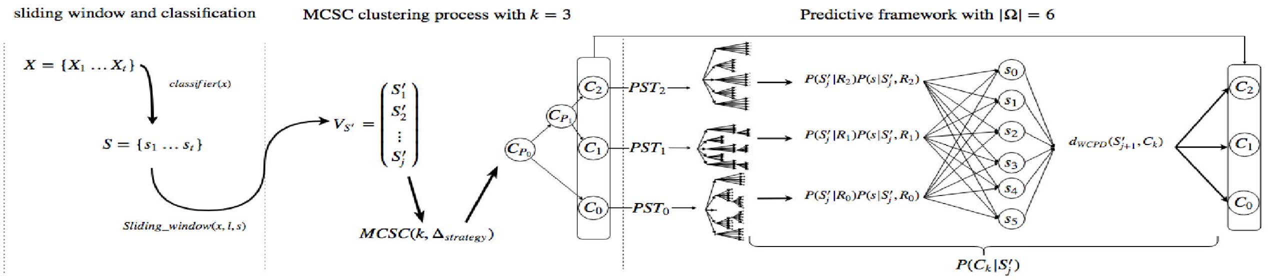

Fig. 1. Pipeline of the WCPD-RS model: The process starts by transforming the time series into a set of categorical sequences using a sliding window technique and a classifier: S' is a subsequence of S and j is the number of subsequences produced from S. Then the WCPD model is used with the MCSC algorithm to cluster the sequence. Here $\Delta_{strategy}$ is the cluster-splitting strategy, which is applicable only if more than two clusters are created. The prediction framework is then derived from the clusters to evaluate the likelihood of next categorical class and predict the next regime.

the leaves has $|\Omega|$ children because we wish our model to be built solely from observed transitions, without having to induce passing probabilities to account for unobserved state transitions. This has no impact on the WCPD model for pruning, but adds complexity to the predictive framework which is addressed in the next section.

The resulting clusters from the WCPD model are calculated by evaluating the statistical center parameter for each cluster C_k as follows:

$$\rho_{\lambda_k}(s|\sigma) = \left(\frac{\sum_{S_i \in C_k} |S_i| \bar{P}^2_{S_i}(s|\sigma)}{\sum_{S_i \in C_k} |S_i|}\right)^{1/2} \quad (1)$$

where $\bar{P}_{S_i}(s|\sigma) = oc(\sigma s, S_i)/|S_i|$, $oc(\sigma s, S_i)$ is the number of occurrences of the memory subsequence σ followed by s in sequence S_i and $|S_i|$ is the length of sequence $|S|$. To measure the dissimilarity between S_i and C_k, we use the following measure:

$$d_{WCPD}(S_i, C_k) = \sum_{\sigma} \sum_{s \in \Omega} \left(\frac{\bar{P}^2_{S_i}(s|\sigma)}{\rho_{\lambda_k}(s|\sigma)} + \rho_{\lambda_k}(s|\sigma) \right) \quad (2)$$

where σ is a statistically significant memory subsequence shared by all models λ, Ω is the alphabet of categorical value, $\bar{P}_{S_i}(s|\sigma)$ is the weighted conditional probability of occurrence of symbol s_i in sequence S_i given memory subsequence σ and $\rho_{\lambda_k}(s|\sigma)$ is the statistical parameter of the WCPD model λ_k on σ defined in (1). These two functions are necessary for the WCPD model and are used in the predictive framework.

C. The Predictive Framework

The basic idea behind the predictive framework is that by using the PST of each cluster produced with WCPD, we can predict the behavior of the associated regime. The PST contains the observed CPD and can be used as an estimate of the "real" CPD of this regime, such that we can model the regime as a stochastic process. To measure how likely a sequence of events is being produced by a regime, we combine both the dissimilarity of the WCPD model (2) between the sequence and the regime and the probability that the sequence of events was produced by the CPD of the regime.

The probability that a sequence S_i occurs given a regime R_k can be calculated as $P(S_i|C_k) = \prod_{j=1}^{|S_i|} P(s_{ij}|\lambda_k)$ where $|S_i|$ is the length of sequence S_i, λ_k is the statistical model of cluster C_k and $P(s_{ij}|\lambda_k)$ is the probability of generating symbol s_{ij} given λ_k, i.e., the CPD of C_k. However, this approach is sensitive to noise caused by statistically non-significant transitions. Alternatively, the dissimilarity measure in equation (2) can be used to determine how a sequence is similar by considering the occurrences of statistically significant patterns only. We combine the two measures to define a similarity function for estimating the likelihood that S_i belongs to regime R_k as follows:

$$sim(S_i, R_k) = P(S_i|C_k) \times \frac{1}{d_{WCPD}(S_i, C_k)} \quad (3)$$

The inverse of $d_{WCPD}(S_i, C_k)$ acts as a confidence level for S_i being in R_k. Finally, by converting the similarity measures in 3, we define the probability $P(S_i|R_k)$ that sequence S_i belongs to regime R_k as follows:

$$P(S_i|R_k) = \frac{sim(S_i, R_k)}{\sum_{i=0}^{k-1} sim(S_i, R_i)} \quad (4)$$

To forecast the following state given a sequence, we consider the CPD of each regime and evaluate which state is most likely to occur. In the context of market uncertainty, it is expected that a sequence is likely to be produced by more than one regime. To weight the importance of each regime's PST, we used the probability of a sequence being generated by a regime defined in (4). However, because we do not enforce the PST to be a full tree, an additional step of normalization is required prior weighting the PST to ensure that the sum of probability of each PST is equal to one. Thus, the probability of a state occurring after a given sequence is evaluated as follows:

$$P(s|S_i) = \frac{\sum_{k \in K} P(S_i|R_k) P(s|S_i, R_k)}{K} \quad (5)$$

where $s \in \Omega$, $P(S_i|R_k)$ is defined in (4), $P(s|S_i, R_k)$ is retrieved by evaluating the probability that s occurs given preceding subsequence S_i from the CPD of regime R_k. The state s retained for prediction is the one most likely to occur next.

To predict whether a regime switch is likely to occur, we evaluate the probability that the predicted subsequent sequence, which we will call "the next sequence" for simplicity, will be allocated to a different regime. There can be a number of next sequences because there is a limited amount of possible states s. By measuring the likelihood of each state occurring with (5) and by considering that a sequence is likely of being associated to a cluster C_k if its WCPD distance $d_{WCPD}(S_i, C_k)$ is the smallest compared to other clusters, we are able to measure the probability that the next sequence will be associated to a regime. This evaluation is done as follows:

$$P(Cj|S_i) = \sum_{s \in \Omega} f(s, S_i, C_j) \tag{6}$$

$$f(s, S_i, C_j) = \begin{cases} P(s|S_i), & \text{if } min(d_{WCPD}(S'_{i+1}, C_j)) \\ 0, & \text{Otherwise} \end{cases} \tag{7}$$

where S_i and C_j are the sequence and cluster of interest, S'_{i+1} is the 1-day ahead forecast sequence and $min(d_{WCPD}(S'_{i+1}, C_j))$ is true if its WCPD distance to C_j is the smallest compared to other clusters. Equation (6) and (7) are illustrated in Fig. 1

When splitting S using a sliding window technique, it is possible that multiple subsequences overlap at t_i but belong to two different regimes. This can be ambiguous as it is unreasonable to use one subsequence for representing t_i and evaluate the probability of being in a regime at t_i. To solve this, we calculate the probability of being in a regime at a given time t_i by considering the probability of all subsequences that crosses t_i. We give more importance to the sequence that is centered at t_i by assigning weights to each sequence based on the distance between the middle point of the sequence and t_i. Then we calculate a weighted average of the probabilities for each regime given this set of subsequences. The probability of being in a regime at time t_i, i.e. $P(t_i, R_k)$, is evaluated as:

$$P(t_i, R_k) = \frac{\sum_{S \in t_i} \frac{1}{dist(S)+1} P(S|R_k)}{\sum_{S \in t_i} \frac{1}{dist(S)+1}} \tag{8}$$

where $\sum_{S \in t_i}$ indicates summation on all the sequences crossing t_i, $\frac{1}{dist(S)+1}$ is the weight attributed to the sequence and $dist(S)$ is the length of time interval between t_i and the center of the sequence. Compared to traditional regime switching models, this approach is equivalent to calculating the smoothing probabilities.

To predict a numerical value instead with the WCPD-RS model, we can approximate the volatility based on the previous realized volatility observed and the state predicted. The strategy consists of using the predicted state from the WCPD-RS predictive framework and selecting the realized volatility based on the mean of the previous observed volatility in that window. If the predicted state s is observed in the last available window, we use the mean of the realized volatility for each days associated with this state. If s was never observed, we use the mean of the last window. In the case where s was previously observed but was not present in the last window, we use the mean of all previous day classified as s that was used to train the model.

IV. Empirical results: A study of the stock market

In this section, we describe experiments conducted to verify the effectiveness of the proposed WCPD-RS model on financial datasets.

A. Datasets

To evaluate our model, we used 200 stocks from the SP500 divided into two datasets. The first dataset, composed of daily OHCLV (open, high, close, low, volume) data from 2000-01-03 to 2018-02-16, was used only for training the models. The second, composed of intra-day market hours OHCLV data from 2017-09-11 to 2018-02-16, was used for validation of the models. The daily data is available to the public on Yahoo Finance website[1] and the intra-day data was retrieved from a Kaggle dataset[2]. Both datasets are adjusted for splits and dividends.

The value of interest to predict is the implied daily volatility of the stock. Since the true volatility is unobservable, we estimated its value with an estimator based on the realized volatility. We used the classical volatility estimator, defined as follows, on the intra-day data:

$$\sigma_t = \sqrt{\sum_{t=1}^{n} (r_t)^2} \tag{9}$$

where $r_t = ln(c_t/c_{t-1})$ and c_t is the closing price at time t. We follow the recommendation of Liu & al. [16] and used 5-minute intra-day data to measure the volatility.

B. Experimental Methodology and Evaluation Criteria

To verify the effectiveness of the proposed WCPD-RS model, we evaluated the volatility prediction performance of WCPD-RS against three different RS models implementation available in D. Ardia & al. [17]: MSGARCH, MSGJR-GARCH, MSTGARCH. A Student distribution was selected for all RS models to take into account the skewness of the return observed in financial time series. We used the predictive framework of the WCPD-RS model to predict a state from which we approximated the volatility based on the past realized volatility. For all models, we used daily OHLCV data of stocks from 2001-01-03 to 2017-10-09 for the initial training and evaluated their predicting accuracy from 2017-10-09 to 2018-02-16 (90 days). We compared all the models by measuring the 1 day ahead volatility forecast accuracy. All the models were initially trained from 2001-01-03 to 2017-10-09 and incrementally updated after each forecast by re-training

[1] https://ca.finance.yahoo.com/

[2] https://www.kaggle.com/borismarjanovic/daily-and-intraday-stock-price-data

the model. To complement our analysis, we calculated the state prediction accuracy of our model and showed a sample of patterns identified in the stock of Apple Inc. (NASDAQ: AAPL).

To train the WCPD-RS model, we split the sequence using a sliding window of 21 days, along with 3-days overlapping. Considering the limited amount of intra-day data available (6 months) and the strategy used for volatility forecasting, using a longer sliding window would reduce the amount of data available for evaluation. On the other hand, a sequence shorter than 21 days sequence might yield more sporadic regimes which are expected to be modeled by smaller patterns. The purpose of the 3-days overlap was to reduce the number of sequences used for training. The same evaluation parameters were used for all datasets.

We used 4 error functions and 1 directional accuracy measure to measure the performance on 1-day-ahead out-of sample prediction for all models. The error functions used are the maximum square error (MSE), the maximum absolute error (MAE), the rooted mean square error (RMSE), the mean absolute percentage error (MAPE) and the mean directional accuracy (MDA). The definitions of all the error functions are as follows: $MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$, $MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_i - \hat{Y}_i|$, $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}$, $MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|Y_i - \hat{Y}_i|}{\hat{Y}_i}$, $MDA = \frac{1}{n}\sum_{i=1}^{n} 1$ if $sign(\hat{Y}_i - \hat{Y}_{i-1}) == sign(Y_i - Y_{i-1})$. Y_i is the realized volatility value, $\hat{Y}_i$ the predicted value and n is the number of predictions; $sign$ a sign function and 1 is an indicator function.

C. Results:

First, we show by a concrete example how patterns are represented in each regime and what the advantages such a representation provides. Let's look at patterns in two regimes for a financial asset: the Apple stock is presented in Table I. We observe that each regime has its own particular patterns that form some easy-to-interpret characteristics. Regime 1 contains longer patterns composed of *A*s compared to Regime 2. An example of such patterns is *AAAAAAAAB* versus *AAAAAAB*. According to the definition of *A* in subsection III.A, this means that Regime 1 tends to have longer "unexciting" days. If we examine short patterns in both regimes, we can notice that sub-patterns *BC* and *CB* rarely occur in Regime 1. Moreover, they never occur at the beginning of a pattern in Regime 1. However, these two sub-patterns are omnipresent in Regime 2. Recall that the sub-patterns *BC* and *CB* imply two clear up-and-down and clear down-and-up days respectively. Regime 1 contains few patterns involving transitions to or from rarely occurring days represented by *D,E,H,I* and *L*, while Regime 2 contains plenty of the patterns involving one or two such days. Recall that *D,E,H,I* and *L* correspond to rarely occurring days involving either an abnormally high volume or high volatility. All these suggest that Regime 2 is a high volatility regime as compared to Regime 1. The three main observations above suggest that the dynamics of the time series vary greatly over the time series (In this example, regime 1 is composed of 597 subsequences and regime 2 is composed of 917 subsequences).

Regime 1:	**Regime 2:**
AAAAAAAAA, AAAAAAAAB, AAAAAAAAC, AAAAAAAAD, AAAAAAAAE, AAAAAAAAI, AAAAAAAL, ABA, ABB, ABC, ABD, ABE, ABI, ABL, ACA, ACB, ACC, ACD, ACE, ACH, ACI, ACL, BAA, BAB, BAC, BAI, CAAA, CAAB, CAAC, CAAD, CAB, CAC, CAE, CAL, CCA, CCB, CCC, CCE, CCH, CCI, CCL	AAAAAAA, AAAAAAB, AAAAAAC, AAAAAAD, AAAAAAH, AAAE, AAAI, AAL, ABA, ABB, ABC, ABD, ABE, ABH, ABL, ACA, ACB, ACC, ACD, ACH, ACI, ACL, BAA, BAB, BAC, BAH, BAI, BBA, BBB, BBC, BBD, BBE, BBH, BBI, BCA, BCB, BCC, BCD, BCE, BCH, BCI, BCL, CAA, CAB, CAC, CAD, CAE, CAH, CAL, CBA, CBB, CBC, CBD, CBE, CBH, CBI, CBL, CCA, CCB, CCC, CCD, CCE, CCH, CCI, CCL, DA, DB, DC, DD, DE, DH, DI, HA, HB, HC, HD, HE, HH, HI, HL

TABLE I
PATTERNS FOUND IN AAPL STOCK DATA FROM 2001-01-03 TO 2018-02-16

The forecasting performance on the 1-day ahead volatility for the 200 stocks is presented in Figure 2 . The figure shows that, on average, WCPD-RS is better at forecasting the actual value most of the time, with a lower variance. For each error function, the median of WCPD-RS is lower than with any other model except in the case of MAPE, where GJRGARCH (30.13%) won by 0.13%. The number of outliers forecast in WCPD-RS is smaller than for any of the other models. Our model is more consistent and in general similar to, if not better than, other RS models. The results are encouraging given a classes prediction of 44% with a standard deviation of 6% for 1 day-ahead forecast while the set of top-3 most probable state contains the actual state at least 75% of the time. Given that there are 8 possible discrete states to predict, such a level of accuracy confirms existence of predictive power of the WCPD-RS framework although GARCH-models are better suited to predict a numerical value of interest.

V. DISCUSSION

Market behaviors are constantly studied to assess systemic risks, financial stresses and other financial behaviors to make better investing decisions. State-of-the-art RS models have been used to model the dynamics of time series but fail to provide easy-to-interpret patterns representing regimes. By identifying the patterns that describe behaviors, we have developed a novel approach for modeling regimes by exploring the underlying structure of time series and by providing a more explicit interpretation of changing dynamics in time series.

In this paper, we have presented a new RS model (WCPD-RS) that models regimes explicitly by extracting the underlying patterns of the time series. The model operates by clustering segments of a time series and extracting significant underlying patterns to build a conditional probability distribution for each regime. Using 200 time series for the validation, we observed that our model outperformed traditional RS models in volatility forecast accuracy except in directional accuracy.

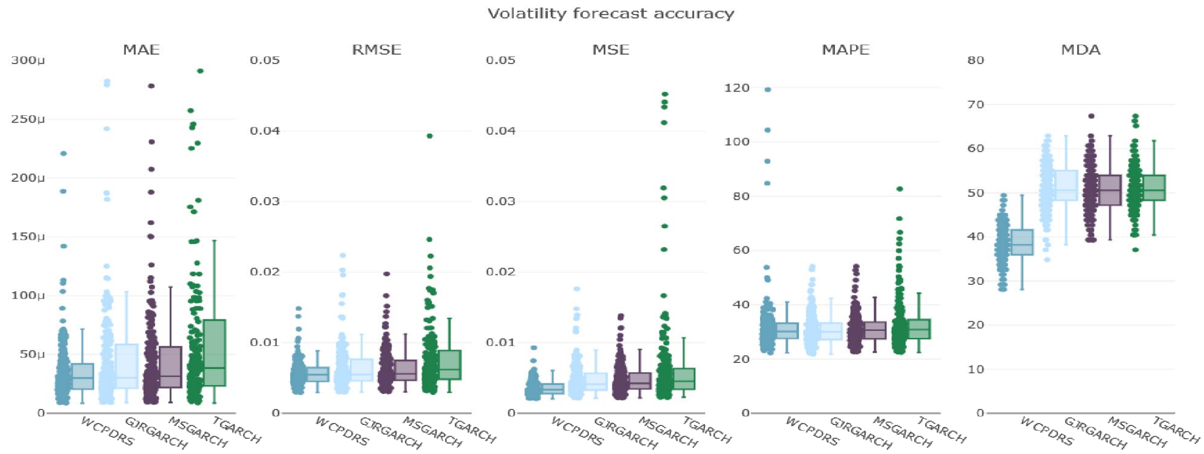

Fig. 2. Volatility forecasting accuracy for 200 stocks from the SP500.

The patterns observed explicitly show that the dynamics of a time series change over time.

The approach proposed here has multiple advantages over traditional regime switching models:

- WCPD-RS extracts natively significant patterns from time series and makes it easy to distinguish regime's dynamics.
- The transition probabilities are evaluated using a variable order Markov chain and are directly dependent on the current and significant past states.
- The model does not require to provide a model specification of a non-observable variable to model regimes.

WCPD-RS thus addresses the main difficulties of traditional RS models, as the complexity is reduced to defining what days are of similar interests. It is worth mentioning that the computational cost of this methodology can be greater than that of some traditional regime switching. Furthermore, traditional RS models may be more appropriate for characterizing different statistical properties of a value of interest but our model excels at extracting explicit behaviors involving multiple variables.

The encouraging results of this work suggest that the behavior of a time series can be learned from the underlying structure of that time series. Future work is needed in incorporating wild-card patterns into the WCPD-RS model. Furthermore, more experimental work is needed to investigate the method's performance with other models beyond two regimes. This is an interesting avenue for future work since gaining insights into the intricacies of more complex regimes is of interest to financial practitioners.

ACKNOWLEDGMENT

We would like to thank NSERC and Laplace Insights for their financial support and inputs throughout the project.

REFERENCES

[1] Bidarkota, Prasad V., and Brice V. Dupoyet. "The impact of fat tails on equilibrium rates of return and term premia." Journal of Economic Dynamics and Control 31.3 (2007): 887-905.

[2] Longin, Francois, and Bruno Solnik. "Extreme correlation of international equity markets." The journal of finance 56.2 (2001): 649-676.

[3] Connolly, Robert A., and F. Albert Wang. "International equity market comovements: Economic fundamentals or contagion?." Pacific-Basin Finance Journal 11.1 (2003): 23-43.

[4] Hardy, Mary R. "A regime-switching model of long-term stock returns." North American Actuarial Journal 5.2 (2001): 41-53.

[5] Mount, Timothy D., Yumei Ning, and Xiaobin Cai. "Predicting price spikes in electricity markets using a regime-switching model with time-varying parameters." Energy Economics 28.1 (2006): 62-80.

[6] Hamilton, James D. "A new approach to the economic analysis of nonstationary time series and the business cycle." Econometrica: Journal of the Econometric Society (1989): 357-384.

[7] Zopounidis, Constantin, and Emilios Galariotis. Quantitative Financial Risk Management: Theory and Practice. John Wiley & Sons, 2015.

[8] Diebold, Francis X., and Atsushi Inoue. "Long memory and regime switching." Journal of econometrics 105.1 (2001): 131-159.

[9] Chang, Yoosoon, Yongok Choi, and Joon Y. Park. "A new approach to model regime switching." Journal of Econometrics 196.1 (2017): 127-143.

[10] Schmitt, Thilo A., et al. "Non-stationarity in financial time series: Generic features and tail behavior." EPL (Europhysics Letters) 103.5 (2013): 58003.

[11] McNeil, Alexander J., Rdiger Frey, and Paul Embrechts. Quantitative Risk Management: Concepts, Techniques and Tools-revised edition. Princeton university press, 2015.

[12] Chen, Dipeng, and Derek Bunn. "The Forecasting Performance of a Finite Mixture RegimeSwitching Model for Daily Electricity Prices." Journal of Forecasting 33.5 (2014): 364-375.

[13] Chang, Yoosoon, Yongok Choi, and Joon Y. Park. "A new approach to model regime switching." Journal of Econometrics 196.1 (2017): 127-143.

[14] Tengke Xiong, Shengrui Wang, Qingshan Jiang, and Joshua Zhexue Huang. A novel variable-order markov model for clustering categorical sequences. IEEE Transactions on Knowledge and Data Engineering, 26(10):23392353, 2014.

[15] Chuffart, Thomas. "Selection criteria in regime switching conditional volatility models." Econometrics 3.2 (2015): 289-316.

[16] Liu, Lily Y., Andrew J. Patton, and Kevin Sheppard. "Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes." Journal of Econometrics 187.1 (2015): 293-311.

[17] Ardia, D., Bluteau, K., Boudt, K., Catania, L., & Trottier, D. A. (2016). Markov-switching GARCH models in R: The MSGARCH package.