Procedia Computer Science

Volume 51, 2015, Pages 1593–1602

ICCS 2015 International Conference On Computational Science

Computational Visual Analysis of the Order Book Dynamics for Creating High-Frequency Foreign Exchange Trading Strategies.

Javier Sandoval[1] and Germán Hernández[2]

[1] Universidad Nacional de Colombia, Universidad Externado, Bogotá, Colombia
jhsandovala@unal.edu.co
[2] Universidad Nacional de Colombia, Bogotá, Colombia
gjhernandezp@unal.edu.co

Abstract

This paper presents a Hierarchical Hidden Markov Model used to capture the USD/COP market sentiment dynamics choosing from uptrend or downtrend latent regimes based on observed feature vector realizations calculated from transaction prices and wavelet-transformed order book volume dynamics. The HHMM learned a natural switching buy/uptrend sell/downtrend trading strategy using a training-validation framework over one month of market data. The model was tested on the following two months, and its performance was reported and compared to results obtained from randomly classified market states and a feed-forward Neural Network. This paper also separately assessed the contribution to the model's performance of the order book information and the wavelet transformation.

Keywords: Machine Learning, Price Prediction, Hierarchical Hidden Markov Model, Order Book Information, Wavelet Transform.

1 Introduction.

Learning profitable trading strategies requires the combination of expert knowledge and information extracted from data. Experts visually detect important patterns in financial charts and react accordingly. For this reason, the combination of a decision-making model and financial data should be the base for building up profitable trading strategies. In this context, this paper presents a trading strategy constructed using an HHMM that represents financial market interactions and wavelet-filtered order book information highlighting the most relevant features.

One of the first special cases of DBNs implemented in the price prediction problem were Hidden Markov Models (HMMs). HMMs assumed that the underlying modeled system exists in one of a finite number of states. The latter states are hidden and are responsible for producing a sequence of observable variables. Hassan [5] is one of the first authors who extracted HMMs from speech and image recognition problems and placed them in the stock price prediction domain.

Selection and peer-review under responsibility of the Scientific Programme Committee of ICCS 2015

doi:10.1016/j.procs.2015.05.290

As a natural extension to an HMM, a Hierarchical Hidden Markov Model (HHMM) was also used to represent financial markets and solve price prediction problems. An HHMM is a natural extension because experts identify different levels of time hierarchy when analyzing financial market information. For example, Jangmin et al. [7] presented a 5-state model that described market trends; strong bear, weak bear, random walk, weak bull and strong bull market phases. This work reported that the HHMM outperformed on average a simple buy-and-hold strategy and a trading strategy following a TRIX, a commonly used technical analysis indicator. HHMMs were also adapted to high-frequency financial data. In Sandoval's work [9], a 2-state model which captured runs and reversals was coupled with a second hidden variable layer which produced observable market features. This work implemented an asynchronous time model and recognized regime changes from uptrend to downtrend time periods. The input variable set went from historical prices to the first 10 orders from the market order book.

The application of DBNs to the price prediction domain is a recently explored study field. The main objective of this work is to assess the contribution of wavelet-transformed order book information to the design of a profitable trading strategy extending results found in Sandoval [9]. Next section will present the market information representation. Third and fourth sections discuss dataset, methods and experiment. Finally, model's performance and conclusions are provided.

2 Observed feature vector.

A feature vector was constructed to capture information from two valuable sources; transaction and order book dynamics. Transactions are defined as realized market trades. The order book can be defined as the group of orders that has not been executed yet but shows agents' intentions to trade at certain quantities and prices. This information was used to forecast future trend in financial prices because exact future prices are not needed to create a profitable trading strategy. In the specific case of the order book dynamics, several studies had shown that order books have relevant information to improve financial price direction prediction [1, 2, 3, 4, 10, 11]. Therefore, the feature vector will combine elements from the order book and transaction dynamics.

In order to understand the feature vector characterization, we made several definitions. Let $\{P_y\}, y \in \{1, 2, \ldots, Y\}$ be a series of market transactions not necessarily homogeneously distributed in time. Transaction durations have not been taken into account and have been left for future research. Let also define $\{E_m, I_m\}, m \in \{1, 2, \ldots, M\}$ as a sequence of local extrema from the series P_y at transaction index $I_m = y$, i.e. prices at which the transaction series changes direction. The index set I_m is a strictly increasing series recording positions of extremum transaction prices. The zig-zag process, $\{Z_x\}$, was constructed recording differences between n adjacent local extrema as $Z_x = E_m - E_{m-n},\ m \in \{2n, 3n, \ldots, M\}$ and $x \in \{1, 2, \ldots, X\}$. n controls the number of local extremum price differences accumulated. If $Z_x \geq 0$, the zig-zag is leading upward or is simply called positive and if $Z_x < 0$, the zig-zag is leading downward or is simply called negative. If $n = 1$, we do not accumulate zig-zags. This cumulative zig-zag is useful to reduce model's prediction instability.

The zig-zag and the extremum price series are complemented with information extracted from the order book dynamics. The order book is defined as a 2-dimensional process that describes changes in limit orders' volume for different price levels during each observed market event, i.e. order insertion, modification or elimination. The order book series, $\{B^2_y\}$, will be limited to book states when transactions occur. Because of the high sparsity of the order book, price levels are normally expressed as intervals over the price dimension.

Based on the raw limit order book process, two simple-smoothed exponential distance-weighted average volume series, $SEDWAV_x^{Bid}$, $SEDWAV_x^{Offer}$, are constructed to capture how volume was concentrated in the buy and offer sides of the order book for every zig-zag x. These two series corresponded in time with each zig-zag. The closer the volume to the best bid/offer best price, the higher the weight given to that volume. An exponential average was used to expressed this fact. Formally:

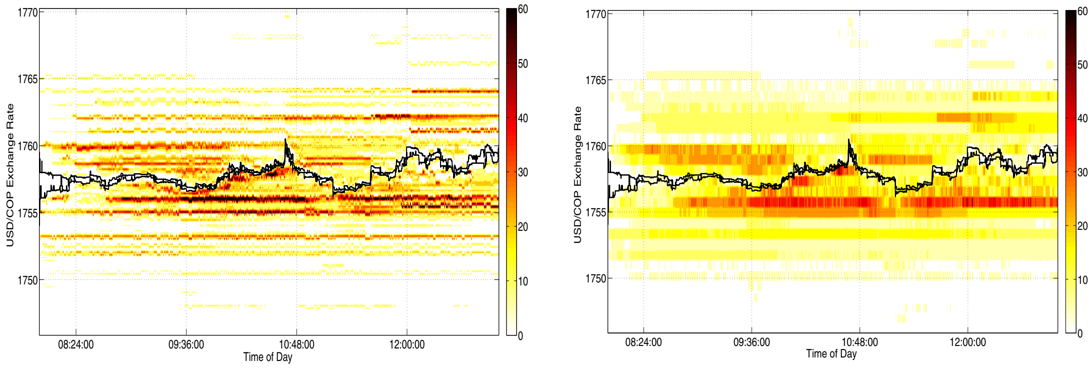

Figure 1: Raw (Left) and Denosed (Right) USD/COP order book evolution from March 21, 2012. White spaces correspond to zero volume. Every volume unit is equivalent to 250 thousand US dollars. Maximum volume, 60 in the color bar, is equal to 15 million US dollars. Solid black lines show best bid/offer price evolution.

$$
\begin{aligned}
SEDWAV_x^{Bid} &= \frac{1}{I_{m-1} - I_m} \sum_{y=I_{m-1}+1}^{I_m} EDWAV_y^{Bid}, \\
SEDWAV_x^{Offer} &= \frac{1}{I_{m-1} - I_m} \sum_{y=I_{m-1}+1}^{I_m} EDWAV_y^{Offer},
\end{aligned}
\tag{1}
$$

where $m = x \times n$ with $x = \{2, 3, \ldots, X\}$, I_m is the position of the mth extremum price. $EDWAV_x^{Bid}$ and $EDWAV_x^{Offer}$ show order book's strength and allow to identify what is commonly known as floors and caps in technical analysis. Figure 1 (Left) shows order book evolution for a certain day of the dataset. As described before, there are many price intervals with zero volume. Figure 1 (Left) also gives visual evidence of bid volume blocks around 1,756 pesos between 9:36:00 am and 10:48:00 am.

Wavelet transform of the SEDWAV Series. The raw order book series was denoised before calculating the SEDWAV series using a discrete $2D$-wavelet transform of the order book process with a Haar Wavelet over a daily window and a 2-level resolution. The wavelet transform was used to recover a filtered version of the order book series setting all detail coefficients at the second level in all three directions, horizontal, vertical and diagonal, to zero. Thus, we expect to capture relevant changes and leave aside the effect of noisy order book updates. The marginal contribution of the wavelet transform will be assessed on the validation data set. Figure 1 (Right) shows the denoised version of the order book information presented in Figure 1 (Left). The wavelet-filtered version of the order book clearly shows the bid volume block previously observed in the raw order book.

Next, we defined volume blocks as wavelet-transformed SEDWAV values that are greater than a threshold α_V. The same threshold is used for bid and offer SEDWAVs. Accordingly, a discrete feature

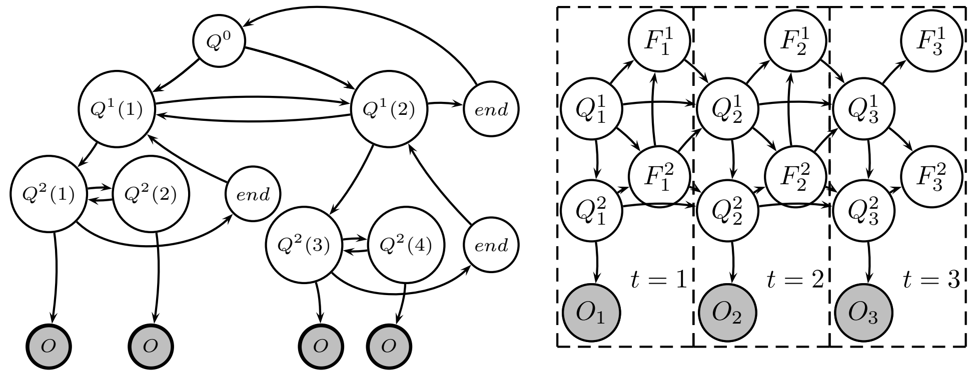

Figure 2: Left: Proposed 2-level automaton. Right: Proposed HHMM represented as DBN. Q_n^1 is the market regime and Q_n^2 is the feature producer. Elements of the observed variables are explained in section 2.

vector was created containing three elements that described zig-zag pattern types, transaction and order book dynamics. In particular,

$$O_x = (f_x^1, f_x^2, f_x^3) \quad \text{where,}$$

$$f_x^1 = \begin{cases} 1 & Z_x \geq 0 \text{ local maximum} \\ -1 & Z_x < 0 \text{ local minimum} \end{cases}$$
$$f_x^2 = \begin{cases} 1 & E_{x-4} + \alpha_P < E_{x-2} < E_x - \alpha_P \wedge E_{x-3} < E_{x-1} - \alpha_P \\ -1 & E_{x-4} - \alpha_P > E_{x-2} > E_x + \alpha_P \wedge E_{x-3} > E_{x-1} + \alpha_P \\ 0 & otherwise, \end{cases} \tag{2}$$
$$f_x^3 = \begin{cases} -1 & SEDWAV_x^{Offer} > \alpha_V \\ 1 & SEDWAV_x^{Bid} > \alpha_V \\ 0 & otherwise. \end{cases}$$

where α_P is a threshold used to differentiate significant transaction price movements.

Figure 3 summarizes feature vector interpretation. The first element in the observed feature vector is the zig-zag type, i.e. maximum or minimum. The second component captures price's momentum comparing current maximum or minimum with its recent historical values. Finally, the third element captures the existence of volume blocks on both sides of the order book. For example, $(1, 1, 1)$ means a local maximum, with a local uptrend and a volume block on the bid side of the order book. $D_{1:9}$ are observations exclusively produced by $Q^2(1)$ and $Q^2(4)$. $U_{1:9}$ are observations only produced by $Q^2(2)$ and $Q^2(3)$, see Figure 2. The HHMM's structure simulates a two-level market in which, first, it enters a market regime and then within each regime, positive and negative features are produced. This structure guaranteed that a positive feature is always followed by a negative featured and vice versa. This structure summarizes expert trader's knowledge of how financial prices evolve. We expect that $U_{1:4}$ and $D_{1:4}$ features are more probable observed during macro uptrends and $U_{6:9}$ and $D_{6:9}$ are more probably found during macro downtrends. U_5 and D_5, defined as no micro trend and no volume block existence are expected to be associated with no particular market state. These two states will represent what is commonly known as noise in price movement.

Using the previous model, this paper captured the dynamics of the USD/COP in order to predict its short-term future behavior. Next section will present the dataset and its characterization.

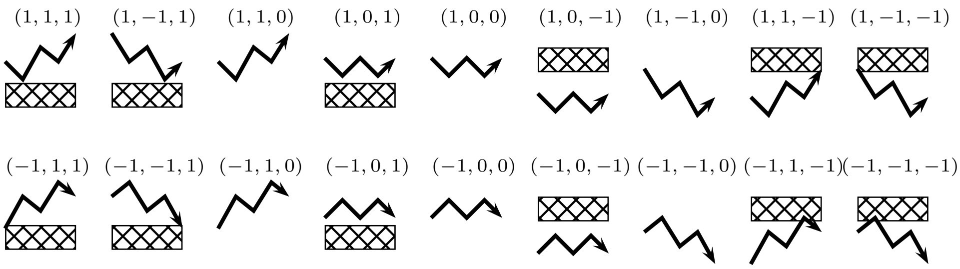

Figure 3: Graphical interpretation of the feature vector observations classified as positive, $U_{1:9}$ (upper) or negative, $D_{1:9}$ (lower). Crosshatched rectangles represent volume blocks.

3 DataSet

Dataset consisted of three months of tick-by-tick information from the limit order book and transactions of the USD/COP, the Colombian spot exchange rate starting in March 1, 2012. Data has been extracted from the Set-FX market, the local interbank FX exchange market. Dataset covered 43,431 transactions. Transactions with similar time stamp have been aggregated into one observation because it is very likely that the same agent executed them. Additionally, data included 658,059 order book updates for every order located 10 pesos above/under the best quoted prices aggregating orders' volume using 20 cent intervals. Volume was expressed in 250 thousand US dollar units. Due to liquidity issues, the first and last 10 minutes of the available data were not considered. USD/COP spot interbank exchange market opens at 8 am and closes at 1 pm. It is a semi-blind market, participants only know their counterparts after transactions are executed. USD/COP average daily turnover is 1 billion dollars. Figure 1 depicts the raw and filtered order book dynamics of a particular day.

4 Methods and Experiment.

Different from other studies [6, 8], the primary goal of this work was to predict market states instead of raw prices or price levels. Therefore, order book and transaction information series were transformed into a feature vector realization series. Though, market states were learned following an unsupervised framework, we expected that during the training phase, the DBN differentiated between two market regimes latter marked as uptrend or downtrend.

Transaction price and order book information were converted into an observed feature series following section 2. The zig-zag factor agregation, n, was set to 5 so 3 negative and positive basic zig-zags were aggregated to create unique zig-zags. Each aggregated zig-zag covers, on average, 10 transactions per observed feature realization. A lower n value would drive to a prohibitively high computational cost for real implementations. In contrast, a larger n would increase the number of missing entering points in price trends.

First, this paper explored potential values for price and volume threshold, α_V and α_P. To guarantee that every possible feature vector realization was observed during the training set, α_V took values between 12 and 18 volume units (3-4.5 million US dollars) and α_P took values between 5 and 30 cents.

Parameter values located out of these ranges produced feature vector observations with zero probabilities of occurrence. It is worth it to mention that the α_V value range may be influenced by the 20-cent volume interval chosen to represent the order book information. The feature vector observation series was divided into three disjoint sections. The first part was used to train the proposed DBN. The second part was the validation set used to test the model's generalization ability and to select the best parameter combination. Finally, the third part of the data was left to test the selected model using information not provided during the calibration phase. Unfortunately, because data has time structure, it is not possible to shuffle it to repeat the calibration and generalization testing procedure.

Training subset was the first 14 days. Validation set was the next 8 days. 38 days were left for testing the model. First, we used an Expectation Maximization (EM) algorithm to find the parameters that maximized model's likelihood function over the training set. After obtaining HHMM model's parameters, it was found the most probable states of $Q^1_{1:X}$ variables based on observed evidence over the training set. This process is known as forward-looking Viterbi inference over observations $1:X$ and it is formally defined as:

$$\underset{Q^1_{1:X}}{arg\ max}\ P(Q^1_{1:X} \mid O_{1:X}), \tag{3}$$

where $Q^1_{1:X}$ will be the state realizations of Q^1 variable during the training data. Learned states of $Q^1_{1:X}$ variables were marked as uptrend or downtrend based on the average returns obtained in similar states. Formally:

$$\bar{R}_{Q^1(j)} = \frac{1}{N_{Q^1(j)}} \sum_{n=1}^{N_{Q^1(j)}} \log P^F_n(Q^1(j)) - \log P^I_n(Q^1(j)),\ \ j = 1, 2, \tag{4}$$

where $P^F_n(Q^1(j))$ and $P^I_n(Q^1(j))$ mean mid-quote best bid-offer at the end (F) and beginning (I) of each consecutive Q^1 realizations in state j. $N_{Q^1(j)}$ is the total number of consecutive Q^1 realizations in state j over the training data set. if $\bar{R}_{Q^1(1)} > \bar{R}_{Q_1(2)}$, state 1 is called uptrend, state 2 is called downtrend and viceversa. After states $Q^1(j), j = 1, 2$ were marked as uptrend or downtrend in the training set, the forward-looking Viterbi inference problem was solved again in the validation set. Then, equation 4 was used to calculate conditional state returns used to compute the trading strategy, buy/uptrend and sell/downtrend, over the validation set. The validation set was chosen, so it contained roughly two easily separated up/down macro trends. Return/risk performance based on average over standard deviation of returns was calculated for each macro trend. Table 1 reports adjusted return/risk (ARR) performance measure for different combinations of α_V and α_P. Adjusted model's return/risk performance definition is:

$$ARR(\alpha_V, \alpha_P) = \frac{\bar{R}_U}{\sigma(R_U)} + \frac{\bar{R}_D}{\sigma(R_D)} - \left| \frac{\bar{R}_U}{\sigma(R_U)} - \frac{\bar{R}_D}{\sigma(R_D)} \right| \tag{5}$$

where, $\bar{R}_U$ and $\bar{R}_D$ correspond to average returns obtained from the simulated strategy during macro up/down trends observed during the first and last halves of the validation dataset. The ARR guarantees that the model's performance is balanced between both up and down macro trends. Small values of α_V and high values of α_P produced the poorest performance. The best parameter combination was $\alpha_V = 15$ and $\alpha_P = 9$. The wavelet transform of the bid and offer SEDWAV series helped to increase average performance from 0.18 to 0.37 measured using ARR over the validation set. Therefore, from now on SEDWAV will stands for wavelet-transformed SEDWAV. Standard performance measures were avoided because when creating a trading strategy based on a market state classification, the accuracy rate, and similar measures do not include the state conditional probability distribution of price movements. Though the mean over the standard deviation of returns calculated using a simulated trading strategy was selected to consider strategy's performance adjusted by risk, it is entirely possible to use

α_V/α_P	30	20	15	9	5
12	−1.178	−1.160	−1.326	0.089	0.089
13	0.053	0.095	0.086	0.397	0.158
14	0.397	0.397	0.397	0.585	0.89
15	0.607	0.825	0.212	**1.066**	1.017
16	0.659	0.757	0.647	0.968	0.956
17	0.250	0.698	0.569	0.538	0.803
18	−0.181	0.714	0.550	0.773	0.457

α_V/α_P	30	20	15	9	5
12	−0.536	−0.590	0.428	0.142	0.130
13	−1.269	0.268	−0.954	0.499	0.394
14	0.235	0.378	0.361	0.705	0.777
15	−1.177	0.390	0.378	0.616	0.262
16	0.176	0.373	0.368	0.216	0.565
17	−0.256	0.416	0.429	0.263	0.178
18	0.314	0.249	0.379	0.480	0.768

Table 1: Adjusted return/risk performance measure calculated over the validation set for different combinations of α_V and α_P using wavelet-transformed SEDWAV (Left) and raw SEDWAV series (Right). α_V is expressed in 250 thousand US dollars, and α_P is expressed in cents.

another performance measure.

Finally, the selected model was also assessed on the testing set. However, forward-looking Viterbi inference drives to a non-predictable trading strategy because information indexed $x+1 : X$ is not available at the moment we decide to trade. Therefore, when testing model's prediction ability, we implemented Viterbi inference without Forward-Looking. Formally, Viterbi inference is defined as,

$$\underset{Q_x^1}{arg\,max}\; P(Q_x^1|O_{1:x}), \;\; x = \{1, \cdots, X\}. \tag{6}$$

Viterbi inference is much slower than forward-looking Viterbi inference because it implies repeating the inference process every time a new zig-zag is completed. Equation 4 was recalculated over the testing set and the buy/uptrend and sell/downtrend strategy was also evaluated on the testing set using the average over the standard deviation of returns. Next section will present unconditional and Q^1 state conditional observed feature probability distributions over the training set. We will emphasize the model's ability to differentiate between uptrend and downtrend regimes and the possibility of migrating from a regime identification to a profitable trading strategy on the testing set. Model's performance using the mean over the standard deviation of returns (without differentiate between macro trends) is reported and compared to the performance results obtained from a random marking process and a two-layer feedforward neural network (FFNN).

5 Model Performance and Results.

Unconditional and conditional Q^1 probability distributions of observed features calculated over the training set using the selected model are shown in Figure 4. The first relevant observation is that $U5$ and $D5$ are the most common feature categories. This finding is an expected result because they capture what it is called noise in price movements, trans- actions that do not have support on price or order book dynamics. Because $Q^1(1)$ class was marked as an uptrend state (see Table 2 second row), it can be assessed which features are more frequent in each regime. Specifically, there is clear evidence to mark $D1$, $D2$, $D3$, $D4$, $U1$, $U2$, $U3$ and $U4$ as uptrend features and $D6$, $D7$, $D8$, $D9$, $U6$, $U7$, $U8$ and $U9$ as downtrend features. State classification perfectly fits what market experts believe based on common knowledge. Therefore, the market regime probability structure aligned with the traders' expert knowledge. Afterward, the selected model and the corresponding buy/uptrend, sell/downtrend strategy was run on the testing dataset using Viterbi inference. As shown in Table 2, left column, returns obtained from the $Q^1(2)$, the downtrend class are clearly skewed to the negative side. Likewise, returns obtained from the $Q^1(1)$ class, the uptrend regime, are skewed to the positive side. Moreover, the selected model also had the ability to find two distinct regimes based on comparing mean returns calculated in each market state. See Table 3, left column, for details. During the testing dataset, selected model 's total return was 4.35% and the return/risk performance was 0.1505. The proposed HHMM model was able to classify two high-frequency market regimes and this classification became a profitable trading strategy. Previous findings do not hold if order book volume information is not considered to construct the feature vector series. Using the same α_P and ignoring volume data in a simplified feature vector for the simulated trading experiment, total return decreased to 0.34% driving the return/performance measure to 0.012. See table 2 and 3, right column.

Random State Marking and feed-forward Neural Network results. Though total return and return/risk performance were positive, there is not a clear benchmark to assess proposed HHMM model's results. Therefore, zig-zag feature observations were randomly marked as uptrend or downtrend. Afterward, the buy/uptrend, sell/downtrend strategy was executed. This experiment was repeated 1,000 times. State change probability was set at 10% up to 50% for each experiment. Figure 5, left column, reported the $95th$ percentile upper limit of total return and return/risk performance measure calculated over the testing set. It has to be obvious that it is quite improbable to obtain the selected model's total return and return/risk performance value just by chance. However, when order book volume is not considered, it not possible to differentiate between the studied model's and random picking's results.

Moreover, a two-layer feed-forward neural network (FFNN) with sigmoid hidden and softmax output neurons was also used to classify future market states based on current and lagged transaction price and bid and offer SEDWAV series. Because an FFNN's calibration procedure can be trapped in a local minimum, network's parameters were estimated 100 times using different starting values. The $95th$ percentile upper limit was calculated for the total return and the return/risk performance measure using 2 to 25 neurons in the hidden layer. Though the proposed HHMM model managed zig-zags aggregating a variable number of transactions and order book volume information, the FFNN can not handle a variable amount of input data. In order to overcome these limitations, the FFNN was trained using different input information intervals ranging from 50 to 100 transactions and their corresponding wavelet-transformed SEDWAV values. On average, the proposed HHMM model aggregated 9.5 transactions per each feature observation and 9 feature observations per each micro trend. Figure 5, right column, summarizes FFNN-based trading performance results. As shown, the best testing results for the FFNN were obtained combining 20 neurons in the hidden layer and between 60 and 70 aggregated values in the transaction and SEDWAV series. However, the $95th$ percentile upper limit for the total return and the return/risk performance measure were lower than the ones obtained using the HHMM proposed model. The FFNN was not able to reproduce the HHMM-based model's results when using order book volume information.

6 Conclusion

This work proposed a 2-level HHMM that was converted to a DBN for training and assessment purposes. The proposed model assumed that a particular financial market could be viewed as a complex automaton which enters into two main regimes. Then, each regime cycled through two feature producers, throwing negative and positive observations calculated from transaction and wavelet-transformed order book volume data. The former model was tested over the USD/COP foreign exchange rate market using three months of high-frequency data covering transaction prices and tick-by-tick order book information. Data was divided into three groups that represented the training, validation, and testing sets. After training different models over the training set, the model that showed best generalization abilities using an adjusted return/risk performance measure over the validation set was selected.

	Complete Model		Using non-volume Info.	
	uptrend	downtrend	uptrend	downtrend
Observations	145	145	120	121
Mean	0.0138%	-0.0162%	$-.00218\%$	-0.00503%
Min	-0.20%	-0.41%	-0.35%	-0.38%
Max	0.51%	0.21%	0.50%	-0.39%
Std	0.09%	0.11%	0.11%	0.12%
Skewness	1.90	-1.34	0.869%	-0.881%
Kurtosis	10.11	5.57	6.43	5.042
Total Return	4.35%		0.34%	
return/risk perfom.	0.1505		0.012	

Table 2: Summary of return statistics found over out-sample data using Viterbi inference implemented as in equation 6. Using order book volume information (Left) and without using order book volume information (Right).

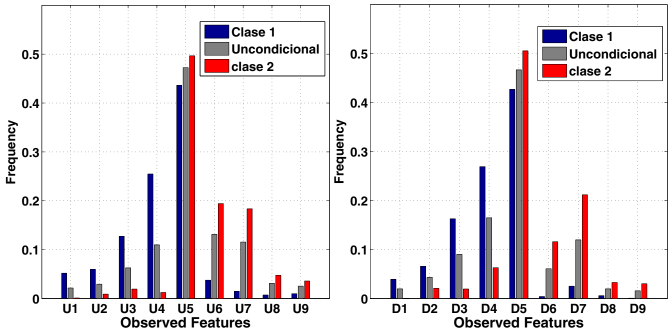

Figure 4: Observed conditional and unconditional probability distribution of observed market features calculated on the training set using the selected HHMM model. Left: Positive features. Right: Negative features.

This study used Viterbi inference to classify market regimes between uptrend and downtrend on the testing dataset. Based on the previous classification, the selected model produced a profitable buy/uptrend and sell/downtrend trading strategy that outperformed a simple random state marking process and a feed-forward neural network. Market regime classification ability of selected HHMM model was statistical significant. No future information was used during the execution of the studied trading strategy.

$H_0 : \mu_U - \mu_D = 0, H_1 : \mu_U - \mu_D \neq 0.$		
2-tailed t-test $\alpha = 5\%$		
	Complete Model	Using non-volume Info.
p-value	1.1%	84%
Action	Reject	No Reject
t-statistic	2.5591	0.1915
confident Interval	$[0.00692\%; 0.05304\%]$	$[-0.0265\%; 0.0323\%]$

Table 3: Hypothesis testing of the sample means of strategy's returns calculated on the testing set.

State Change Prob.	Return/risk perform.	Total return
10%	0.1476	2.65%
20%	0.1027	2.97%
30%	0.0873	3.05%
40%	0.0687	2.84%
50%	0.0603	2.85%

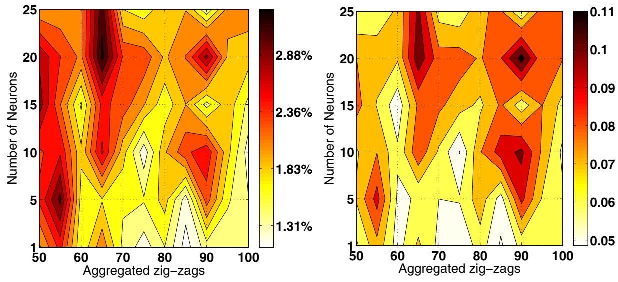

Figure 5: (Left) $95th$ percentile upper limit of return/risk performance and total Return obtained after executing the simulated buy/Uptrend and sell/Downtrend trading strategy when observed feature realizations have been randomly classified as positive or negative. The first column represents the transition probability between market states. (Right) $95th$ percentile upper limit of return/risk performance and total Return obtained after executing same strategy using the Feed-Forward Back Propagation Neural Network framework.

References

[1] Jean-Philippe Bouchaud, J. Doyne Farmer, and Fabrizio Lillo. How Markets Slowly Digest Changes in Supply and Demand. Quantitative Finance Papers 0809.0822, arXiv.org, September 2008.

[2] S.N. Dorogovtsev, J.F.F. Mendes, and J.G. Oliveira. Frequency of Occurrence of Numbers in the World Wide Web. *Physica A: Statistical Mechanics and its Applications*, 360(2):548 – 556, 2006.

[3] Zoltan Eisler, Janos Kertesz, and Fabrizio Lillo. The Limit Order Book on Different Time Scales. Quantitative Finance Papers 0705.4023, arXiv.org, May 2007.

[4] Gao-Feng Gu, Wei Chen, and Wei-Xing Zhou. Empirical Shape Function of Limit-Order Books in the Chinese Stock Market. *Physica*, 387(21):5182–5188, 2008.

[5] Md Hassan. Stock Market Forecasting Using Hidden Markov Model: A New Approach. In IEEE, editor, *Proceedings of the 2005 5th International Conference on Intelligent Systems Desing and Applications*, 2005.

[6] Md. Rafiul Hassan, Baikunth Nath, and Michael Kirley. A Fusion Model of HMM, ANN and GA for Stock Market Forecasting. *Expert Systems with Applications*, 33(1):171 – 180, 2007.

[7] O Jangmin, Lee Jae Won, Park Sung-Bae, and Zhang Byoung-Tak. Stock Trading by Modelling Price Trend with Dynamic Bayesian Networks. In *IDEAL'04*, pages 794–799, 2004.

[8] Luis Ortega and Khaldoun Khashanah. A neuro-wavelet model for the short-term forecasting of high-frequency time series of stock returns. *Journal of Forecasting*, 33(2):134–146, 2014.

[9] Javier Sandoval and Germán Hernández. Learning of natural trading strategies on foreign exchange high-frequency market data using dynamic bayesian networks. In Petra Perner, editor, *Machine Learning and Data Mining in Pattern Recognition*, volume 8556 of *Lecture Notes in Computer Science*, pages 408–421. Springer International Publishing, 2014.

[10] Gary Tian and Mingyuan Guo. Interday and Intraday Volatility: Additional Evidence from the Shanghai Stock Exchange. *Review of Quantitative Finance and Accounting*, 28(3):287–306, April 2007.

[11] P Weber and B Rosenow. Order Book Approach to Price Impact. *Quant. Finance*, 5(4):357–364, 2005.