Price Efficiency Forecasting in Financial Markets Using Continuous-Continuous Hidden Markov Model

Ali Shiri†, Mohammad Hossein Moghaddam‡, Ashkan Rahimi-Kian† and Behrouz Maham†
†SNL/CIPCE, School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran, Iran.
‡Department of Electrical Engineering, K.N. Toosi University of Technology, Tehran, Iran
Email: {ali.shiri@ut.ac.ir and arkian@ut.ac.ir}
§School of Engineering, Nazarbayev University, Astana, Kazakhstan
Email: {behrouz.maham@nu.edu.kz}

Abstract—Financial market analysis and prediction is one of the interesting areas in financial signal processing. This area is of high importance for professionals in Foreign Exchange market and stock exchange market as they can lead and direct future trends or manage crisis over time. In this paper we introduced price efficiency as a prediction parameter and proposed a forecasting model based on a continuous-continuous Hidden Markov Model and non-stationary time series. We have examined our model according to IFC Markets Ltd. real data in a ten-year period. Simulation results show that the proposed model has a very good efficiency and expresses low error rate which could be considered in reliable risk management strategies.


I. INTRODUCTION

Forecasting financial markets has been one of the biggest challenges to signal processing and artificial intelligence communities. There are many difficulties to process and analyze financial market signals due to complex and volatile characteristics. Efficient Market Hypothesis states that the market price history up to current market price assimilates all the information relevant to a market, which could be used in forecasting [1]. On the other hand, market reaction to information which could pop up from any sources is completely complex and volatile. There are many parameters affecting the market which arise from various events among the world such as economical crisis, change of seasons, political decisions and even social networks. Whenever a new event occurs, the market makes a chain of reactions to it which decreases the chance of efficient prediction. Nevertheless, several studies have been performed on the financial markets in order to predict future price and also forecast the market behaviour in future [2].

There are many works in literature in which researchers attempted to forecast the financial time series in exchange markets. There are various forecasting models and classifiers proposed to predict financial time series using signal processing and computational tools such as Autoregressive moving average(ARMA) [3], Neural Networks (NN) [4], Support Vector Machines (SVM) [5], [6], Hidden Markov Model (HMM) [7] and other Hybrid Models [8], [9]. Authors in [3] considered and analyzed one special stock market signal using ARMA model with different number of poles and zeros, in order to estimate the values for the next days’ prices. ARMA model is a very useful prediction model for linear dynamic signals but, in case of complex (non-linear) volatile signals it doesn’t have favorable performance. In [4] recurrent neural nets (RNN) is used as the main prediction model to predict daily trading of straddles on financial indexes. Neural Network is a very useful tool in predicting different kinds of complex signals, but it’s complexity grows exponentially with growing layers of network. It has also higher error rates in comparison with SVM and HMM in similar situations. One of prediction methods which could compete with Neural Network is SVM. Its performance in comparison with Multi-layer back propagation (BP) neural network and the regularized radial basis function (RBF) neural network is remarkable [5]. In [5], authors investigated The variability in performance of SVM with respect to the free parameters and show that, there are comparable generalization performance between SVM and the regularized RBF neural network which means that, the free parameters of SVM have a great effect on the generalization performance. Although SVM shows very good performance in volatile time series even in comparison with other methods, finding proper kernels, push great amount of difficulty to its procedure. Another prediction method which is used widely in financial market signals is HMM [7]. HMM is originally a statistical model which is used for prediction in different procedures with different kinds of complexity [10]. There are two HMM procedures commonly used in different forecasting strategies, continuous-discrete [11], [12] and discrete-discrete [13], [14]. Both of these methods gives probable points in future which could be evaluated according to different time series data sets. Spot price estimation give a point in future but, in most cases finding points with high probability needs so many input data and also need to have access to some information about market which in real world cases is not possible or in best cases could be done in a very high cost. On the other hand, Interval estimation gives us a reliable probability of occurring future events in different intervals by affordable input data.

978-1-4673-8789-7/16/$31.00 2016 IEEE
and information which could help us to have better decision making and risk management.

In this paper we have proposed a method based on continuous-input, continuous-output HMM (CHHMM) which instead of predicting points in future, predicts the output probability distribution function and gives probable intervals which the point would lie on in future. We have done this by defining price efficiency as the prediction parameter and used CCHMM to predict it. Interval prediction based on CCHMM could be used efficiently as a reliable method for risk management strategies since it has lower error rates in comparison with other HMM procedures that try to propose just a point in the future.

The remainder of this paper is organized as follows: Section II provides a brief overview on HMM. Then, we describe the details for the proposed algorithm in section III. In section IV, we illustrate the effectiveness of the proposed algorithm by numerical experiments. Finally, section V conclude the paper.

II. HIDDEN MARKOV MODEL

A Hidden Markov Model (HMM) is a finite state machine which has some fixed number of states. It provides a probabilistic framework for modelling a time series of observations. Rabiner explained the basics of HMM which has some fixed number of states. It provides a probabilistic framework for modelling a time series of observations. 

1) A set of N states:

\[ S = \{s_1, s_2, ..., s_N\} \]  (1)

which are hidden from observation point. HMM could work in any state \( s_i \) at time \( t \) which could be illustrated as \( q_t = s_i \) that could be defined as a hidden state sequence

\[ Q = \{q_1, q_2, ..., q_t\} \]  (2)

2) A set of possible observations in output:

\[ O = \{o_1, o_2, ..., o_1\} \]  (3)

in case of discrete observations, the outputs could be represented as a set of symbols \( V = \{v_1, v_2, ..., v_K\} \), which at time \( t \) \( o_t \in V \).

3) The initial state probability distribution sequence:

\[ \pi = \{\pi_1, \pi_2, ..., \pi_N\} \]  (4)

where \( \pi_i = P(q_1 = s_i) \) represents the probability that the hidden state \( i \) is the initial state.

4) The state transition probability distribution matrix:

\[ A = \{a_{ij}, 1 \leq i, j \leq N\} \]  (5)

where \( a_{ij} = P(q_{t+1} = s_j | q_t = s_i) \) represents the transition probability from the hidden state \( i \) to state \( j \), and the sum of the entries in each row of matrix \( A \) must be 1:

\[ \sum_{j=1}^{N} a_{ij} = 1, \ 1 \leq i, j \leq N \]  (6)

5) The output observation probability distribution matrix in discrete observations:

\[ B = \{b_j(k), 1 \leq j \leq N\} \]  (7)

where \( b_j(k) = \{o_t = v_k | q_t = s_i\} \) gives the probability of the observation symbol \( v_k \) with the given hidden state \( j \). We also have

\[ \sum_{k=1}^{K} b_j(k) = 1, \ 1 \leq j \leq N \]  (8)

but in case of continuous output observations, we should use a continuous probability density function instead of discrete probability for symbols. Typically this probability density function is estimated by weighted sum of \( M \) normal distribution functions:

\[ b_j(\bar{O}) = \sum_{m=1}^{M} c_{j,m}N(\bar{O}, \mu_{j,m}, \Sigma_{j,m}) \]  (9)

\[ N(\bar{O}, \mu_{j,m}, \Sigma_{j,m}) = \frac{1}{2\pi^{\frac{K}{2}} |\Sigma_{j,m}|^{\frac{1}{2}}} \exp\left( -\frac{(x-\mu_{j,m})^T \Sigma_{j,m}^{-1} (x-\mu_{j,m})}{2} \right) \]  (10)

\[ \sum_{m=1}^{M} c_{j,m} = 1, \ 1 \leq j \leq N \]  (11)

where \( \bar{O} \) is the vector of observations being modeled, \( c_{j,m} \) mixture coefficient for the m-th mixture in state \( j \), \( \mu_{j,m} \) mean vector for the m-th mixture component state \( j \) and \( \Sigma_{j,m} \) covariance matrix for the m-th mixture component in state \( j \).

An HMM in discrete case according to (4)-(7) could be denoted by \( \lambda \) as

\[ \lambda = (\pi, A, B) \]  (12)

and in continuous case as

\[ \lambda = (A, c_{j,m}, \mu_{j,m}, \Sigma_{j,m}, \pi) \]  (13)

according to (4)-(11), respectively.

To work with HMM in continuous case, the following three fundamental questions should be resolved [15]. [7]:

1) Given the model \( \lambda = (A, c_{j,m}, \mu_{j,m}, \Sigma_{j,m}, \pi) \) how do we compute \( P(\bar{O} | \lambda) \).

2) Given the observation sequence \( \bar{O} \) and a model \( \lambda \), how do we choose a state sequence \( q_1, q_2, ..., q_t \) that best explains the observations.

3) Given the observation sequence \( \bar{O} \) and a space of models found by varying the model parameters \( A, c_{j,m}, \mu_{j,m}, \Sigma_{j,m} \) and \( \pi \), how do we find the model that best explains the observed data.

There are established algorithms to solve the above questions. In our task we have used the forward-backward algorithm to
compute the $P(\tilde{O} \mid \lambda)$, Viterbi algorithm to resolve problem 2 and, Baum-Welch algorithm to train the HMM. The details of these algorithms are given in the tutorial by Rabiner [15].

III. PROPOSED MODEL

As we stated before, HMM is an intelligent method which could be used to predict time series. In this section we explain our proposed model based on continuous-input, continuous-output HMM to produce the output distribution function. First consider that the input distribution is normal, then we try to compute the $P$ function. Rising in number of inputs in each group also makes the clustering sequentially in a queue. Suppose that we have $N$ data amounts as $d_1, ..., d_N$ and wanted to divide them in to $X$ groups with $K$ data amounts in each group. We put the data amounts $d_i, ..., d_i+K-1$ in group $i$ for $i = 1, ..., X$. Then in each group, mean and variance of corresponding data sets are computed and used as inputs to HMM which finally give us a pair of mean and variance values which we use them to represent the output normal distribution function. Our system model is depicted in Figure 1. Predicting Normal distribution function for each day and computing the next day price interval probability according to the estimated pair of mean and variance values, gives us proper information about the corresponding price efficiency interval and also related probabilities about that interval. We have considered the probability and price efficiency according to 4 different standard deviations as below:

$$P(o_1) = \frac{n_1}{n_T}, \text{ for } \mu - \sigma \leq r(n_1(i)) \leq \mu + \sigma$$

$$P(o_2) = \frac{n_2}{n_T}, \text{ for } \left\{ \begin{array}{ll} \mu + \sigma < r(n_2(i)) \leq \mu + 2\sigma \\ \mu - 2\sigma \leq r(n_2(i)) < \mu - \sigma \end{array} \right.$$ (16)

$$P(o_3) = \frac{n_3}{n_T}, \text{ for } \left\{ \begin{array}{ll} \mu + 2\sigma < r(n_3(i)) \leq \mu + 3\sigma \\ \mu - 3\sigma \leq r(n_3(i)) < \mu - 2\sigma \end{array} \right.$$ (17)

$$P(o_4) = \frac{n_4}{n_T}, \text{ for } \left\{ \begin{array}{ll} r(n_4(i)) > \mu + 3\sigma \\ r(n_4(i)) < \mu - 3\sigma \end{array} \right.$$ (18)

where $n_i$, $r(n_1(i))$ and $n_T$ are number of price efficiencies located in $i$th interval, price efficiency in first interval and total number of experiments ($n_T = n_1 + n_2 + n_3 + n_4$). In HMM block, the best parameters for number of inputs, number of states and number of gaussian mixtures are computed in an iterative manner. Our criterion to compare different number of parameters is the standard gaussian distribution, in which 68% of variables lie in the first standard deviation, 95% of variables lie in the second standard deviation and 99.7% of variables lie in the third standard deviation. Best parameters change from one financial signal to another according to behaviour of financial markets and considering the frequency of deviations in volatile situations. For example, the frequency of deviations in currency exchange markets is lower than oil price or stock markets. After finding the best parameters, the HMM gives us mean and variance values of price efficiency as outputs which then we use them to represent the gaussian distribution function. By having these information, we could find that the next day price could lie in the resulted distribution function. Next day, the yesterday price falls into our examining groups and the procedure continues during the time while the sliding window moves towards it with different probabilities according to gaussian distribution function.

IV. SIMULATION RESULTS

In this section we represent our simulation results according to real data, provided from IFC Markets Ltd data base. we have used EUR/USD spot price, Standard and Poor's index (S&P) and oil price according to IFC Markets Data base in 12 years from January first 2000 to the end of December 2012. Our model’s best parameters include number of states, number of gaussian mixtures and the size of window according to the number of inputs. In each simulation round, we have used a window of inputs and predicted the price intervals for next day of last input value in each window to compare it with data sets for evaluation. At the end of each simulation round, the sliding window steps forward one day, and simulation starts to predict price intervals for next day. In all simulations we have supposed that the output distribution function is zero outside the third standard deviation.

We made our experiments by dividing the input data in each window to different group sizes from 3 to 7 inputs in each group as explained in previous section. Rising the number of inputs in each group leads to increase the propagation in each group and more distributed standard deviation accordingly. This makes the corresponding intervals wider and increases the probability of next day to be lied in output distribution function. Rising in number of inputs in each group also makes the mean and standard deviation values more uniform among
different groups. With more uniform mean and standard deviation values we need less states in HMM. This fact is favorable since increasing the HMM states makes more overlaps between states which leads to less reliable estimation. We have also compare the results of our model with Neural Network Model (NNM) results in table I through III, to evaluate the results in different models. According to table I to III, it could be found that although the average of standard deviations in EUR/USD spot price time series is lower than of those for S&P index and oil price, but number of prices lied in first standard deviation for EUR/USD is greater than other two series. On the other hand, by increasing the number of inputs in each group, number of HMM states in EUR/USD time series decreases with a greater slope in comparison with S&P and oil price time series which means that the frequency of deviations in EUR/USD spot price is lower than S&P index and oil price.

Output distribution functions for different intervals which are predicted for the last month of 2012 are depicted in figures 1 through 3, using 5-day input grouping.

**TABLE I. HMM & NNM ACCURACY EVALUATION FOR EUR/USD SPOT PRICE**

<table>
<thead>
<tr>
<th>Number of data in each group</th>
<th>State</th>
<th>Mixture</th>
<th>Input</th>
<th>Mean of Standard deviations</th>
<th>1st Intervals</th>
<th>2nd Intervals</th>
<th>3rd Intervals</th>
<th>Out</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>81</td>
<td>NNM 1.1732 -65.6 &lt; 25.09 &gt; 68.3</td>
<td>-2.51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>4</td>
<td></td>
<td>HMM 1.1832 -66.28 &lt; 24.36 &gt; 66.6</td>
<td>-2.26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>6</td>
<td>104</td>
<td>NNM 1.1904 -66.68 &lt; 24.35 &gt; 66.46</td>
<td>-2.59</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>3</td>
<td>155</td>
<td>NNM 1.1985 -67.77 &lt; 24.00 &gt; 59.72</td>
<td>-2.26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>3</td>
<td></td>
<td>HMM 1.2045 -68.27 &lt; 23.91 &gt; 58.24</td>
<td>-1.98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>2</td>
<td>150</td>
<td>NNM 1.2047 -68.13 &lt; 23.93 &gt; 57.56</td>
<td>-2.18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td>HMM 1.2087 -68.07 &lt; 23.84 &gt; 57.54</td>
<td>-1.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>1</td>
<td>140</td>
<td>NNM 1.2200 -69.30 &lt; 23.69 &gt; 56.42</td>
<td>-1.99</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>1</td>
<td></td>
<td>HMM 1.2209 -69.62 &lt; 23.45 &gt; 55.92</td>
<td>-1.93</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TABLE II. HMM & NNM ACCURACY EVALUATION FOR OIL PRICE**

<table>
<thead>
<tr>
<th>Number of data in each group</th>
<th>State</th>
<th>Mixture</th>
<th>Input</th>
<th>Mean of Standard deviations</th>
<th>1st Intervals</th>
<th>2nd Intervals</th>
<th>3rd Intervals</th>
<th>Out</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>69</td>
<td>NNM 1.6613 -63.65 &lt; 27.16 &gt; 72.71</td>
<td>-2.28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>4</td>
<td></td>
<td>HMM 1.6696 -63.85 &lt; 27.04 &gt; 66.32</td>
<td>-2.19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>6</td>
<td>104</td>
<td>NNM 1.7579 -65.42 &lt; 26.41 &gt; 60.07</td>
<td>-2.09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>4</td>
<td></td>
<td>HMM 1.7615 -66.15 &lt; 26.16 &gt; 59.53</td>
<td>-1.76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>3</td>
<td>135</td>
<td>NNM 1.8061 -66.29 &lt; 25.89 &gt; 60.63</td>
<td>-1.79</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>3</td>
<td></td>
<td>HMM 1.8175 -67.56 &lt; 24.66 &gt; 59.60</td>
<td>-1.79</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>2</td>
<td>136</td>
<td>NNM 1.8492 -67.21 &lt; 25.12 &gt; 59.97</td>
<td>1.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>2</td>
<td></td>
<td>HMM 1.8522 -68.38 &lt; 24.44 &gt; 59.52</td>
<td>-1.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>1</td>
<td>161</td>
<td>NNM 1.8504 -67.07 &lt; 25.94 &lt; 53.54</td>
<td>1.65</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td>HMM 1.8677 -67.77 &lt; 25.51 &gt; 52.12</td>
<td>-1.4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TABLE III. HMM & NNM ACCURACY EVALUATION FOR S&P INDEX**

<table>
<thead>
<tr>
<th>Number of data in each group</th>
<th>State</th>
<th>Mixture</th>
<th>Input</th>
<th>Mean of Standard deviations</th>
<th>1st Intervals</th>
<th>2nd Intervals</th>
<th>3rd Intervals</th>
<th>Out</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>72</td>
<td>NNM 1.4188 -62.5 &lt; 27.85 &gt; 72.28</td>
<td>-2.37</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>4</td>
<td></td>
<td>HMM 1.4241 -64.55 &lt; 26.45 &gt; 74.51</td>
<td>-2.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>5</td>
<td>108</td>
<td>NNM 1.4999 -64.21 &lt; 26.16 &gt; 68.68</td>
<td>-2.39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>4</td>
<td></td>
<td>HMM 1.5033 -65.37 &lt; 26.16 &gt; 62.27</td>
<td>-2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>4</td>
<td>105</td>
<td>NNM 1.5684 -65.00 &lt; 26.23 &gt; 67.47</td>
<td>-1.96</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td>HMM 1.5761 -66.11 &lt; 25.95 &gt; 67.27</td>
<td>-1.72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>4</td>
<td>100</td>
<td>NNM 1.5833 -67.07 &lt; 25.68 &lt; 59.95</td>
<td>1.90</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td>HMM 1.6057 -68.72 &lt; 24.05 &gt; 55.51</td>
<td>1.72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>3</td>
<td>154</td>
<td>NNM 1.6322 -67.16 &lt; 25.51 &gt; 55.56</td>
<td>1.76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td></td>
<td></td>
<td>HMM 1.6982 -68.87 &lt; 25.21 &gt; 53.51</td>
<td>1.62</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In these figures green intervals represent the first standard deviation, black intervals represent second standard deviation and blue intervals represent third standard deviation of output normal distribution functions. Analysing the outputs, shows that the number of next day prices which lie in second standard deviation for oil price is greater than same numbers in EUR/USD rate and S&P index which indicates that the oil price time series had a higher frequency of deviations in comparison with other time series in last month of 2012. Figures show that most of next day prices lie in first standard deviation, so we can distinguish maximum increase or decrease in price by producing continuous outputs which could help a trader to have a proper understanding of market behaviour and to make proper risk management strategies.

V. CONCLUSION

We have proposed a method for price efficiency forecasting by using continuous-input, continuous-output and continuous Hidden Markov Model. In our method, we tried to compute the best parameters and produce the output distribution function which next day price lies on that. Our algorithm is examined with real data sets, and simulation results show that by increasing the number of averaging inputs, the output standard deviation becomes wider and made the placement of next day price on output distribution function more probable. Results also indicate that most of next day prices lie in first standard deviation of output normal distribution function, a fact which could make our algorithm a reliable criterion for market evaluation and risk management decisions.

ACKNOWLEDGMENT

The authors would like to thank IFC Markets Ltd for open access to their financial data base.

REFERENCES