Forex Market Prediction Using Multi Discrete Hidden Markov Models

José Pedro Alves

Abstract - This paper proposes a new Hidden Markov Model (HMM) approach to pattern discovery using MACD and RSI technical indicators to assist in the HMM forecast. This approach uses three discrete HMMs (DHMM) each of which is trained with different windows size. Having been trained differently, each HMM has a different sensibility to direction variations in the financial time series. With the assistance of the above technical indicators is possible to adapt each of the three different HMM to the market behavior. That is, whenever the indicators suggest a possible change in the market trend the developed approach will use the most sensitive HMM to quickly adapt and predict accordingly to this new trend. The developed approach was tested using FOREX EUR/USD historical data from 2002 to 2013 and is performed a sum of pips to assess the revenue of using the proposed strategy over the 12 years period.

Index Terms - Pattern discovery, forecasting, financial market, time series, Hidden Markov Model, MACD, RSI, multi-HMM time series analysis.

I. INTRODUCTION

Amongst all the research and models developed in recent years, a model that has shown good results is the hidden Markov model. This machine learning model is already widely used in gene prediction [1], protein folding [2], cryptanalysis [3], part-of-speech tagging [4] and speech recognition [5]. For the analysis of financial time-series is typically used their continuous variant (CHMM) in order to analyses and predict the exact value, or the closest possible approximation to the real value. Hassan and Nath [6] applied a continuous HMM to forecast some of the airline stocks. First, the CHMM is trained and then to forecast the next day’s closing price, the model computes a likelihood value “ϑ” for the day and locate from the past data set those instances that would produce the same “ϑ” or nearest to the “ϑ” likelihood value. Assuming that the next day’s stock price should follow the same pattern, from the located past day(s) is calculated the difference of that day’s closing price and next to that day’s closing price and the next day’s stock closing price is obtained adding the above difference to the current day’s closing price.

In [7] Gupta and Dhintra used a 3-dimensional vector in the CHMM with values computed from the open, high, low and close values from four different stock indices (TATA steel, Apple Inc, IBM Corporation and Dell Inc). For forecasting an approximated Maximum a Posteriori (MAP) is used to forecast future stock values. Likelihood values are estimated and from the selected past vector a prediction for the next day closing price is computed.

By using likelihood values, these two approaches do not take into account the pattern that led to the identified historic day, thus it is unlikely that the identified historic day has the same characteristics as the day that intends to forecast.

From [8] Hassan, Nath and Kirley developed a fusion model of HMM, Artificial Neural Networks (ANN) and Genetic Algorithms (GA). The GA optimize the initial parameters of HMM and ANN transforms the daily stock prices to independent sets of values that become input to HMM. For forecast the model computes a likelihood value and select those days that would produce the same or nearest likelihood values then is used a weighted average of the price differences. The result is added to the current price of the day in order to obtain a prediction of the next day price. The use of GA to optimize the initial parameters would not be necessary if the HMM used the Baum-Welch algorithm for training, gaining in simplicity and performance.

II. RELATED WORK

For the analysis of financial time-series is typically used the continuous variant (CHMM) in order to analyses and predict the exact value, or the closest possible approximation to the real value. Hassan and Nath [6] applied a continuous HMM to forecast some of the airline stocks. First, the CHMM is trained and then to forecast the next day’s closing price, the model computes a likelihood value “ϑ” for the day and locate from the past data set those instances that would produce the same “ϑ” or nearest to the “ϑ” likelihood value. Assuming that the next day’s stock price should follow the same pattern, from the located past day(s) is calculated the difference of that day’s closing price and next to that day’s closing price and the next day’s stock closing price is obtained adding the above difference to the current day’s closing price.

In [7] Gupta and Dhintra used a 3-dimensional vector in the CHMM with values computed from the open, high, low and close values from four different stock indices (TATA steel, Apple Inc, IBM Corporation and Dell Inc). For forecasting an approximated Maximum a Posteriori (MAP) is used to forecast future stock values. Likelihood values are estimated and from the selected past vector a prediction for the next day closing price is computed.

By using likelihood values, these two approaches do not take into account the pattern that led to the identified historic day, thus it is unlikely that the identified historic day has the same characteristics as the day that intends to forecast.

From [8] Hassan, Nath and Kirley developed a fusion model of HMM, Artificial Neural Networks (ANN) and Genetic Algorithms (GA). The GA optimize the initial parameters of HMM and ANN transforms the daily stock prices to independent sets of values that become input to HMM. For forecast the model computes a likelihood value and select those days that would produce the same or nearest likelihood values then is used a weighted average of the price differences. The result is added to the current price of the day in order to obtain a prediction of the next day price. The use of GA to optimize the initial parameters would not be necessary if the HMM used the Baum-Welch algorithm for training, gaining in simplicity and performance.

III. MULTI DESCRETE HMM

In this chapter is explained in detail the adjustments made in the development of Multi DHMM. First, it is explained how the chosen technical indicators interact with DHMM and the expected benefits of that interaction. After the different developed models are presented and indicated the choice of models to enter the Multi DHMM. In section 4.5 is explained
in detail the implementation process of the developed method.

A. DHMM Training: Baum-Welch Algorithm

- How can we estimate the model parameters given an observation set?

In order to answer the question it is important to use an algorithm capable of finding the unknown parameters \( \lambda = \{ \pi, A, B \} \) of a HMM having \( \pi \) as the initial matrix, \( A \) as the transition matrix and \( B \) as the emission matrix. Although exists some algorithms with capacity to address the question, due to the type of data on which it will be used it is necessary that the algorithm does not need any model initialization. This algorithm is called the Baum-Welch and use the EM algorithm to find the maximum likelihood estimate of \( \lambda = \{ \pi, A, B \} \) given the observation sequence \( O_1, O_2, ..., O_T \) and uses the production probability \( P(O|\lambda) \) as the optimization criterion.

For the training of DHMM it is crucial to have the transformed data as signals of declining, raising or maintaining \((0,1,2)\) of the market value and define the number of states of the HMM. The decision of the number of states appears as an unknown variable in the process but in the HMM process is often taken as a rule, though not required, having the number of states equal to the number of observations, that is, have a possible strategy (state) for each existing observation. This rule will be adopted in the development of the model.

The Baum-Welch algorithm is described as follow [9][10][11]:

Forward Procedure:

Having \( \alpha_t(i) = P(O_1, O_2, ..., O_t, s_t = i|\lambda) \), the probability of ending in state \( s_t \) given the observation sequence \( O_1, O_2, ..., O_t \) is recursively computed,

1. \( \alpha_t(1) = \pi_i b_j(O_1) \)
2. \( \alpha_{t+1}(j) = b_j(O_{t+1}) \sum_i \alpha_t(i) a_{ij} \)

Backward Procedure:

Having \( \beta_t(i) = P(O_{t+1}, O_{t+2}, ..., O_T|s_t = i, \lambda) \), the probability of the ending sequence \( O_{t+1}, O_{t+2}, O_T \) given the model \( \lambda \) and the \( s_{j,t} \) at time \( t \) is recursively computed,

1. \( \beta_T(i) = 1 \)
2. \( \beta_t(i) = \sum_j \beta_{t+1}(j) a_{ij} \)

Optimization:

It is now possible to compute the temporary variables:

\[ \gamma_t(i) = P(s_t = i|O, \lambda) = \frac{\alpha_t(i) \beta_t(i)}{\sum_i \alpha_t(i) \beta_t(i)} \]

This quantity \( \gamma_t(i) \) represents the probability of being in state \( s_t \) and time \( t \) having the observation set \( O_1, O_2, ..., O_t \) and the parameters from \( \lambda \).

\[ \xi_t(i,j) = P(s_t = i, s_{t+1} = j|O, \lambda) = \frac{\alpha_t(i) a_{ij} \beta_{t+1}(j) b_j(O_{t+1})}{\sum_i \sum_j \alpha_t(i) a_{ij} \beta_{t+1}(j) b_j(O_{t+1})} \]

This quantity \( \xi_t(i,j) \) represents the probability of being in state \( i \) and \( j \) in times \( t \) and \( t+1 \) respectively, having the observation set \( O_1, O_2, ..., O_t \) and the parameters from \( \lambda \).

With the computation of this two quantities it is now possible to update the model determining the expected quantities \( \hat{\lambda} = \{ \hat{\pi}, \hat{a}, \hat{b} \} \)

Update of the model \( \hat{\lambda} \):

1. \( \hat{\pi} = \gamma_t(1) \)
2. \( \hat{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T} \gamma_t(i)} \)
3. \( \hat{b}_j(k) = \frac{\sum_{t=1}^{T} \delta_t(k) \gamma_t(i)}{\sum_{i=1}^{N} \gamma_t(i)} \)

Termination:

If the quality measure \( P(O|\hat{\lambda}) \) was not improved by the updated model \( \hat{\lambda} \) compared to \( \lambda \), the process stops, however, if not, repeat all steps.

Figure 15 - Flowchart of the Baum-Welch algorithm for a discrete HMM

B. DHMM Testing: Viterbi Algorithm

The Viterbi algorithm is presented as the chosen option to forecast the direction of the market close price. This algorithm look at every state sequence and simply select the most likely sequence in a process assumed to be a finite-state and discrete in time Markov process. Like the forward or the backward algorithm, the Viterbi algorithm also have a variable represented by \( \delta_t(i) \). This new variable generate the segment from the observation sequence \( O_1, O_2, ..., O_T \) with maximum likelihood and ends in state \( s_t \).

\[ \delta_t(i) = \max_{s_{t+1}} P(O_1, O_2, ..., O_t, s_1, s_2, ..., s_{t-1}, s_t, s_{t+1} = i|\lambda) \]

This variable \( \delta_t(i) \) can be compared with the forward variable \( \alpha_t(i) \), except that the Viterbi algorithm uses maximization instead a summation over previous states.

The Viterbi algorithm is described as follows:

1. Select the most likely sequence in the process using the Viterbi algorithm [9]:
• **Initialization:** For all states \(i, j \in [1, N]\) in \(t = 1\) we have:

\[
\delta_1(i) = \pi_i b_i(O_1)
\]

\[
\psi_1(i) = 0
\]

**Recursion:** For all times \(t, 1 \leq t \leq T - 1\) and all states \(i, j \in [1, N]\) we have:

\[
\delta_{t+1}(j) = \max_i \{\delta_t(i) a_{ij}\} b_j(O_{t+1})
\]

\[
\psi_{t+1}(j) = \arg \max_i \{\delta_t(i) a_{ij}\}
\]

• **Termination:** For all states \(i, j \in [1, N]\) in \(t = T\) we have:

\[
P^*(O|\lambda) = P(O), s^*|\lambda) = \max_i \delta_T(i)
\]

\[
s^*_T = \arg \max_j \delta_T(j)
\]

2. Having the most likely sequence from \(t = 1\) to \(t = T\), the next step will be to assess the most likely state in \(t = T + 1\). This is calculated from the manipulation of the algorithm equations. It is created the matrix \(\varphi(T+1)\) that holds the probability of going to state \(i\) in case of having the observation \(O\) in \(T + 1\),

\[
\varphi(O_{T+1}) = \max_i \{\delta_T(i) a_{ij}\} b_j(O_{T+1})
\]

3. Next using (63) the most probable state in \(T\) is extracted from the results obtained with the Viterbi algorithm,

\[
\text{State} = \arg \max_i \delta_T(i)
\]

4. The state extracted the previous topic is used in \(\psi_T(j)\) to extract the most likely predecessor state, i.e., the most likely state at \(T + 1\) where,

\[
\text{Predecessor} = \psi_T(j = \text{State})
\]

5. Having the most probable predecessor state, it is now possible to compute the most probable observation in \(T+1\),

\[
\text{Forecast} = \arg \max_i \varphi_{\text{predecessor}}(O_{T+1})
\]

C. **Why Multi Hidden Markov Model**

The focus of the HMM is bounded by the size of the training window. Higher training windows give the HMM the capacity to perceive the formation of long-term trends but make the model less sensitive to detect changes in trends; as opposed, the use of reduced training windows gives the model the ability to identify the formation of short and transient patterns and greater sensitivity in detecting changing trends. Due to constant fluctuations in the financial markets, it is important that the developed model is able to adapt to such fluctuations. Thus, it is important that the model is capable of analyzing long term trends, while quickly adapts to these market changes.

To this end, it was decided to use three DHMM trained with different window sizes. The size of the windows has been obtained through test. The size was determined 90 days because this window size achieved the best results in all tested window sizes, nevertheless it was found that the 90-day DHMM slowly adapts to changing trends, and so it was decided that this size would be the maximum size to be considered. The choice of the 15-day size was again determined through tests that showed that the quality of forecasts and the sensitivity to fluctuations would eventually present itself as the most balanced choice. To bridge the gap between the maximum and minimum value of the window, it was decided to adopt a third DHMM with the window size that would be between the other two. Two cases were then tested, the 30 days size, equivalent to 1 month, and 90 days, equivalent to 2 months. The tests results showed the DHMM 30 days as the best among them.

D. **Development of a multi HMM strategy**

This first model intends to add three different DHMM in a single model. The idea centers on the use of the output of each model to generate the final prediction using the most predicted signal. For this decision to be possible, it was necessary to reduce the number of the DHMM output of observations for two, this because there was a risk of each of the three DHMM generates a different value having a halt in the forecast decision. Thus, it was decided to add the observation that there is a price drop with the observation concerning the existence of price maintenance. It can be concluded from the tests that the probability of the same repeating closing price on these two days is greatly reduced.
equal forecasts, the value of these forecasts is then chosen to be the value reported as final forecast from the model.

Figure 3 - Flowchart of the three DHMM model integration

E. Insertion of technical indicators with multi DHMM

Having three DHMM with different sensibility to direction variations enables a rapid and effective adaptation to the behavior of financial time series. Thus, stick to use on an analysis of the value predicted in higher number is to be under-utilizing the capabilities that this approach can offer. That said, the new objective focused on the use of DHMM 90 as the main model and the use of DHMM 30 and DHMM 15 each time a new trend is detected to accelerate the adaptation of the model to new market behaviors. To this end, were used three technical indicators to the detection and indication of possible overbought or oversold (RSI), new trends (MACD) or indication of the strength of those trends (ADX) and a fourth approach when the RSI and MACD indicators were used simultaneously. The above indicators are explained in detail in chapter 2.

To deal with the detection by the technical indicator of a trend change was developed an application controller of each HMM in an orderly manner. This controller applies the HMM 15 for the first 15 days following the indicator signal, the 15-20 day the HMM 30 and finally returns to the use of the HMM 90. These time intervals are chosen firstly such a way that the window of each HMM model did not stop contain the day when the indicator has detected a change in trend. Adaptation and final decision was made in the days presented from testing, which showed these are values that better obtained results.

Figure 4 - Flowchart of a multi DHMM model with technical indicators

In Figure 22 is shown the flowchart of the process where the Technical Indicator block is replaced by the analysis of the results of each indicator and Compute Indicator block compute the value of the indicator. The analysis process for each indicator can be found in Figure 23.

Figure 5 - Technical indicators used in the Technical Indicator box stated in Figure 4

The first stage of the model development was preceded by their respective testing which showed a significant improvement in the results, except for the ADX that due to its poorer performance this indicator will not be included in the second model development. The remaining cases were able to achieve the objectives. For the second stage the reduction of losses over the years has become as the primary goal, even if it means a decrease of profits, to increase the reliability and the investment safety of the model.

F. Fusion between different methods

For this second phase of development, as mentioned in the previous section the aim is to reduce the losses per year of the model even if the profit per year will eventually reduce. Thus, three of five models previously developed were added to this new model and a new "0" signal was added. Aggregation is made so that there is confirmation from two models of direction provided by DHMM, the signal will be transmitted whenever the prediction of both algorithms is contradictory. In these days of uncertainty, in which the two DHMM give a different forecast will not be done investment, avoiding possible losses.

From this second stage the signal is transformed into estimates of the new representations. In addition to adding the "0" signal, the down signal of closing market value is converted to "-1", and the increase to "1".
In Figure 24 is shown the flowchart of the process where the Aggregation of Developed Models block is replaced by the analysis of the results of each. The analysis process for each sub-model aggregation can be found in Figure 25.

The results show that there was a substantial reduction in losses despite having been a slight drop in earnings. So, to try to recover losses in annual earnings without further increase the damage was developed the third and final phase which is described in the next section.

G. Multi DHMM Automation

The development of the third and final stage were added the five developed models that showed the best results. Once again in order to have a confirmation of direction provided by the various models, the most predicted value will be chosen. Table 5 lists the selected models from phase one and two:

<table>
<thead>
<tr>
<th>Phase 1</th>
<th>Phase 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>DHMM MACD</td>
<td>DHMM RSI and DHMM MACD</td>
</tr>
<tr>
<td>DHMM MACD RSI</td>
<td>DHMM 15 30 90 AND DHMM MACD</td>
</tr>
<tr>
<td>DHMM 15 30 90 and DHMM RSI</td>
<td>DHMM MACD 15 30 90 and DHMM RSI</td>
</tr>
</tbody>
</table>

Table 1 - Selected models from each phase

Therefore, were chosen the first phase models with a higher total gain, but in order to mitigate the losses occurring in some of the years in these two models are used three models of the second phase, thus incorporating to the final model the signal "0" so that no investments were made when there are large uncertainties in the calculation of own forecast.

IV. EXPERIMENTS AND RESULTS

For the construction of the final model several tests were performed in order to detect the limitations of the model and carry out with possible solutions. The analysis and subsequent model improvements led to be held 7 case studies until reaching the final model. These tests were made using Forex EUR/USD historical data from 2002 to 2013 and a sum of pips is performed to assess the revenue of using the proposed strategy over the 11 years period. When a wrong prediction is obtained the pips difference between the closing value of the expected day and its former is subtracted from the pip total, in the event of a correct forecast this difference is added to the total. As mentioned previously, the approach uses a sliding window strategy for training and testing.

At the same time an analysis was made of the impact of the meetings of the European Central Bank (ECB) and the Federal Reserve System (FED) in predicting and behavior of the solutions developed, since the existence of such meetings insert a period of uncertainty in the FX pair EUR / USD.

A. Case Study I

The time detecting a change in direction depends of the size of the window used for training the DHMM. Smaller windows turn more sensitive the algorithm to slight changes but less sensitive to long-term trends, on the other hand, larger windows turn to have a low sensitivity to immediate changes because its focus will be on long-term trends.

Being said, it is important to find a balance between both cases. Thus arises the possibility of using a combination of three DHMM with different training window sizes, a DHMM sensitive to small changes, other sensitive to long-
term trends and a third who is relatively between both cases. The result will be three predictions from each one and will be chosen the more predicted direction of the three.

<table>
<thead>
<tr>
<th>Set [days]</th>
<th>Normal</th>
<th>W/out FED</th>
<th>W/out ECB</th>
<th>W/out FED+ECB</th>
</tr>
</thead>
<tbody>
<tr>
<td>15, 30 and 90</td>
<td>5936</td>
<td>5646</td>
<td>5167</td>
<td>4877</td>
</tr>
<tr>
<td>15, 60 and 90</td>
<td>4096</td>
<td>3072</td>
<td>3415</td>
<td>2391</td>
</tr>
<tr>
<td>30, 60 and 90</td>
<td>4230</td>
<td>3110</td>
<td>3467</td>
<td>2347</td>
</tr>
</tbody>
</table>

Table 2 - Results of different sets of DHMM training window sizes from 2002 to 2013

From the results in Table 11 it is possible to verify that the DHMM 15, 30 and 90 set has a much higher performance than the remaining sets, achieving in 2010 a total of 2461 pips. In Figure 29 it is noticeable the use of a long-term component between January and June when the algorithm does not change the direction of the forecast in April when the direction reversed temporarily. Logically that during this period the prediction proved wrong resulting in a loss of Pips. It is also easy to identify a greater sensitivity to changes since September the algorithm was able to follow a slight change in direction. In this approach, it was noted that an analysis without the days of the EDF or ECB meeting are even more harmful than a global approach.

<table>
<thead>
<tr>
<th>Year</th>
<th>Pip Values</th>
<th>Year</th>
<th>Pip Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>-446</td>
<td>2008</td>
<td>1455</td>
</tr>
<tr>
<td>2003</td>
<td>1075</td>
<td>2009</td>
<td>369</td>
</tr>
<tr>
<td>2004</td>
<td>222</td>
<td>2010</td>
<td>2461</td>
</tr>
<tr>
<td>2005</td>
<td>235</td>
<td>2011</td>
<td>1080</td>
</tr>
<tr>
<td>2006</td>
<td>22</td>
<td>2012</td>
<td>-763</td>
</tr>
<tr>
<td>2007</td>
<td>-46</td>
<td>2013</td>
<td>272</td>
</tr>
</tbody>
</table>

Table 3 - Results from 2002 to 2013 of the DHMM training using 15, 30 and 90 days window size

Although this approach has able to contemplate a macro and micro view of the patterns and their changes in direction, its use is somewhat random. To overcome this challenge, it is important to use a technical momentum indicator which will information the algorithm of a possible change in market behavior for a faster and more efficient adaptation.

Figure 9 - Results from multi DHMM in 2010

B. Case Study II

To attempt to reduce the time it takes the algorithm to detect a new direction transition resorted to the use of RSI. This technical momentum indicator attempts to determine overbought and oversold conditions and this information is expected to detect a change in behavior and market direction. That said, and from the results obtained with previous tests, especially with the results in 5.2.2 were used DHMM three algorithms:

- Training the algorithm using the previous 15 days:
  As mentioned in the previous chapter, when using a small dataset for training the algorithm makes it more sensitive to small changes of direction. So when the RSI indicate a possible change of direction, we expect that the algorithm is as quickly as possible to detect it.

- Training the algorithm using the previous 90 days:
  It is necessary that as soon as it detects and confirm the direction of change detected by the RSI, the algorithm again not be sensitive to noise and to focus on a detection medium and long term patterns.

- Training the algorithm using the previous 30 days:
  It was considered the need for a smooth transition between the last two points. So one can still discard false direction changes detected by RSI within a short period while carried forward to a stage where there is less sensitivity to noise.

It was necessary to define the time that each of the DHMM's presented in the previous 3 points should have after the RSI detect a possible change in market behavior. The first analysis was to use DHMM 15 during the first 20 days, the DHMM 30 during the following 15 days proceeding with the use of the remaining time until a new indication of RSI. Results are shown below:

<table>
<thead>
<tr>
<th>Year</th>
<th>Pip Values</th>
<th>Year</th>
<th>Pip Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>-18</td>
<td>2008</td>
<td>807</td>
</tr>
<tr>
<td>2003</td>
<td>-229</td>
<td>2009</td>
<td>-2761</td>
</tr>
<tr>
<td>2004</td>
<td>240</td>
<td>2010</td>
<td>3185</td>
</tr>
<tr>
<td>2005</td>
<td>1481</td>
<td>2011</td>
<td>-266</td>
</tr>
<tr>
<td>2006</td>
<td>-228</td>
<td>2012</td>
<td>825</td>
</tr>
<tr>
<td>2007</td>
<td>-1408</td>
<td>2013</td>
<td>-464</td>
</tr>
</tbody>
</table>

Table 4 - Results from 2002 to 2013 from Multi DHMM and RSI using 0, 20 and 45 days steps
The first results were lower than expected. The main reason was the use of DHMM 15 and 30 DHMM for too long. To analyze the impact, it was decided to reduce the use of each DHMM to the following intervals: DHMM 15 until 15 days after the statement of RSI, DHMM 30 within 5 days later and DHMM until next indication of RSI. The results in Table 18 show the positive effect of a small reduction of these ranges. It is thus possible to complete the utility of this approach in the detection and adapt to market changes in behavior. The supplementary analysis the presence or absence of the days of the ECB and Fed meetings show that for this case, their absence has no advantage after 14 years.

<table>
<thead>
<tr>
<th>Year</th>
<th>Pip Values</th>
<th>Year</th>
<th>Pip Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>-1172</td>
<td>2008</td>
<td>2563</td>
</tr>
<tr>
<td>2003</td>
<td>835</td>
<td>2009</td>
<td>1901</td>
</tr>
<tr>
<td>2004</td>
<td>606</td>
<td>2010</td>
<td>1899</td>
</tr>
<tr>
<td>2005</td>
<td>357</td>
<td>2011</td>
<td>1164</td>
</tr>
<tr>
<td>2006</td>
<td>-926</td>
<td>2012</td>
<td>927</td>
</tr>
<tr>
<td>2007</td>
<td>-298</td>
<td>2013</td>
<td>722</td>
</tr>
</tbody>
</table>

| 12 Years Total | 8579 |

Table 6 - Results from 2002 to 2013 from Multi DHMM using MACD and Divergence

The results presented in Table 15 show a great total of 8578 pips when combining Divergence and MACD with DHMM but by analyzing Table 16, which is present only the analysis of the combination of DHMM and Divergence one sees a decrease to half of the total achieved in the test of Table 15. Thus, the good result shown in the first case should be of the entire responsibility of the MACD.

C. Case Study III

This approach to try to improve the results of the model with the addition of the trend-following momentum indicator MACD. By using this indicator we are left with three possibilities at their junction, these are the indicator used in its entirety (MACD and divergence), only use the divergence or only use the values of the MACD. In each case were used the same strategy as the previous case study, corresponding to DHMM 15 until 15 days after the statement of RSI, DHMM 30 within 5 days later and DHMM until next indication of RSI.

<table>
<thead>
<tr>
<th>Year</th>
<th>Pip Values</th>
<th>Year</th>
<th>Pip Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>-1146</td>
<td>2008</td>
<td>883</td>
</tr>
<tr>
<td>2003</td>
<td>517</td>
<td>2009</td>
<td>1887</td>
</tr>
<tr>
<td>2004</td>
<td>416</td>
<td>2010</td>
<td>977</td>
</tr>
<tr>
<td>2005</td>
<td>37</td>
<td>2011</td>
<td>506</td>
</tr>
<tr>
<td>2006</td>
<td>-556</td>
<td>2012</td>
<td>473</td>
</tr>
<tr>
<td>2007</td>
<td>86</td>
<td>2013</td>
<td>644</td>
</tr>
</tbody>
</table>

| 12 Years Total | 4724 |

Table 7 - Results from 2002 to 2013 from Multi DHMM using Divergence

Confirmation can be obtained by analyzing Table 17, where the combination of DHMM and MACD was a 10038 pips profit far superior to the two previous cases. It is possible to verify that this result was due to the large profits between 2008 and 2011 and reduced losses over the 11 years ensuring a performance far superior compared to the two previous cases.

<table>
<thead>
<tr>
<th>Year</th>
<th>Pip Values</th>
<th>Year</th>
<th>Pip Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>806</td>
<td>2008</td>
<td>3169</td>
</tr>
<tr>
<td>2003</td>
<td>1575</td>
<td>2009</td>
<td>1873</td>
</tr>
<tr>
<td>2004</td>
<td>-848</td>
<td>2010</td>
<td>1889</td>
</tr>
<tr>
<td>2005</td>
<td>19</td>
<td>2011</td>
<td>1666</td>
</tr>
<tr>
<td>2006</td>
<td>458</td>
<td>2012</td>
<td>-121</td>
</tr>
<tr>
<td>2007</td>
<td>-500</td>
<td>2013</td>
<td>52</td>
</tr>
</tbody>
</table>

| 12 Years Total | 10038 |

Table 8 - Results from 2002 to 2013 from Multi DHMM using MACD

From the results is possible to conclude that the use of MACD greatly increases the algorithm’s performance. The divergence and the set, despite having an interesting result, it falls short of the results reported by isolated MACD. It is even possible to say that the use of divergence with the MACD makes the result decreases in 1566 Pips over the years.

D. Case Study IV

Trend Strength Indicator ADX was used to confirm that a new trend indicated by the RSI or MACD for would be strong or weak. The objective focused on ignoring false signals of new trends. Therefore, if the trend of the MACD or RSI was indicated by the ADX how strong would be used the analysis of 15, 30 and 90 days as referred to in the two previous case studies, if the trend is weak, the HMM continues with an analysis at 90 days. The results in Table 18 showed that due to the poor results there is no advantage in using this indicator in conjunction with other already analyzed. The result managing use MACD - ADX could even be considered, but this result alone does not reflect the performance of the ADX but the high-performance of the MACD indicator that had the same influence as in the previous section with the analysis of MACD with Divergence.
In Table 22 are the results obtained with the combination HMM 15, 30 and 90 and MACD HMM are plotted according to the purpose. On average this model managed to have an annual loss of -93.45 pips and an 842.91 pip profit. This combination also showed the best pip total in 11 years from the three analyzed in this section. In Table 23

Table 22 - Results from 2002 to 2013 combining HMM 15 30 90 and MACD

Table 23 shows the results obtained by combining HMM RSI and HMM 15 30 90. Although this sub-model has the lowest total of the three, has an average value much lower than the previous sub-model. This sub-model achieved an average loss of -49.1818 pips and annual earnings of 544.273 pips. Finally, Table 23 shows the results obtained from the HMM RSI and MACD HMM model. This model has the lowest average loss over the 11 years of the three achieving a total of -39.82 pips a very small result compared to all other models already developed. Their average gain is 698 pips, which presents itself as the second best of the three models analyzed in this section.

Table 23 - Results from 2002 to 2013 combining HMM RSI and HMM 15 30 90

With this approach, the aim of having losses below 500 Pips was reached, naturally in addition to the lower, the gains have also suffered from this new change. Nevertheless, the three cases studied had very positive results, specifically the pair HMM 15 30 90 and HMM MACD who got one of the best results of all simulations.

The junction of the various algorithms ultimately creates a stronger forecast removing large variations in earnings. Despite these variations can be seen as beneficial when
positive, is not the case with large negative changes that can completely disable the algorithm.

It is possible to identify Table 30 the complete elimination of years with a negative pip total. The year of 2007 that appeared negative in the previous case of states, recovered to a gain of 217 pips. Despite this gain can be considered small, also turns out to be the lowest value obtained over the 12 years analyzed. Furthermore, 2008 and 2010 shows an above the earning average of 2196 pip. The analysis without the days of FED and ECB press conferences suggests once again that despite these days represent days of instability in the market, their inclusion brings more benefits than one might think at first analysis being noticed a degradation of the results when those days are removed.

Over the 12 years analyzed the EUR / USD often changed behavior. Between 2002 and 2013 can be detected relatively stable periods and sharp falls, as the 2008 financial crisis and subsequent instability as large oscillation periods. From the graph we can see that the developed method detects and quickly adapts to new market trends, such as the rapid detection of the 2008 financial crisis where the current approach had the highest profit. These results suggest that the developed method is well prepared for fluctuations or different market trends that may arise in the future. Thus, the results show the multi DHMM a profitable method readily adaptable even in unpredictable market conditions.

G. Multi DHMM Automation

This section aims to analyze the result of the final adaptation of the Multi DHMM. This final adaptation had as main objective the substantial reduction of its losses each year. For this purpose, were introduced to the model some of the approaches analyzed in the previous case studies which have demonstrated better performance. A characteristic withdrawal of the model analyzed in the previous case study focuses on the addition of a STOP signal that tells that there is uncertainty in the forecast and therefore it will be better to discard the current forecast. By analyzing Table 28 and 29 is possible to verify the positive impact that this new signal had in the final results. When analyzed all possible outcomes from the model, the percentage of correct predictions reaches 52%, if the estimates that generated the STOP sign are discarded, since in these moments there were no losses nor gains, the percentage rises to 57%, one considerable difference of 5% which has its impact on the final results.

<table>
<thead>
<tr>
<th>Year</th>
<th>Pip Values</th>
<th>Year</th>
<th>Pip Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>1076</td>
<td>2008</td>
<td>5690</td>
</tr>
<tr>
<td>2003</td>
<td>1180</td>
<td>2009</td>
<td>2923</td>
</tr>
<tr>
<td>2004</td>
<td>2481</td>
<td>2010</td>
<td>4567</td>
</tr>
<tr>
<td>2005</td>
<td>2050</td>
<td>2011</td>
<td>2249</td>
</tr>
<tr>
<td>2006</td>
<td>754</td>
<td>2012</td>
<td>2398</td>
</tr>
<tr>
<td>2007</td>
<td>217</td>
<td>2013</td>
<td>764</td>
</tr>
</tbody>
</table>

| 12 Years Total | 26349 |

Table 14 - Resulting percentages from the final model

<table>
<thead>
<tr>
<th>False</th>
<th>True</th>
<th>Stop</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1210</td>
<td>1624</td>
<td>283</td>
<td>3117</td>
</tr>
<tr>
<td>Percentage</td>
<td>39%</td>
<td>52%</td>
<td>9%</td>
</tr>
</tbody>
</table>

Table 15 - Resulting percentages from the final model without stop signal

H. Conclusions

The first analysis of the DHMM demonstrated that this model is an excellent algorithm to perform the forecast of
the Forex values direction. In tests conducted the discrete version was more effective as the continuous version.

The HMM itself has limitations in adapting to new patterns and trends. The speed of adjustment depends on the window size which the DHMM is trained, i.e. the smaller the window more sensitive the algorithm will be to variations and as result, will be more vulnerable to noise.

To help overcome the difficulties encountered were used three indicators (RSI, MACD and ADX), and a joint between the DHMM 15, DHMM 30 and DHMM 90. Results showed that the use of RSI and MACD indicators could provide relevant information on market behavior changes. With the information provided by these indicators it is possible to make the adaptation of the algorithm to the new trend much faster.

The use of ADX, to confirm the existence of a trend turned out to be of no use, since the obtained results failed to achieve the intended objective.

The following objective focused on the limitation of the annual losses in -500 Pips. The objective was achieved by combining previous tested cases, such as, HMM 15 30 90 and HMM MACD, HMM 15 30 90 and HMM RSI and for last the pair HMM MACD and HMM RSI. Thus, values obtained from the best cases were selected for incorporation

the automated method, these are:

- DHMM MACD
- DHMM MACD RSI
- DHMM RSI and DHMM MACD
- DHMM 15 30 90 and DHMM RSI
- DHMM 15 30 90 and DHMM MACD

Having analyzed the automated method, we observed that the aim of reducing annual losses was largely achieved. This new approach could register 12 years of steady profits. In addition, the inclusion of a STOP signal largely expanded capacity to contain unnecessary losses and readily adaptable even in unpredictable market conditions.

V. FINAL CONCLUSIONS

After developing the model and analyzed the results is easy to conclude that the strategy used to forecast the direction of the daily closing value for the FX EUR / USD is presented as a great choice. The ease with which this strategy adapts to new trends and behaviors of the market and the quality of their predictions not only allowed faster adaptation but also a substantial reduction in losses that the DHMM alone had. These features were made possible by the merger of technical indicators (RSI and MACD), already widely used in technical analysis of financial markets, with the implementation strategy of three DHMM.

The results show that the inclusion of sub-models in the model developed during the process led to much higher gains compared to the individual gains from each sub-model, i.e., the final model has a gain of 26349 pips after 11 years (from 2002 to 2013), this value is 2.6 times higher than the best results obtained from the separate analysis of the sub-models for the same period. This is due to the need for confirmation of the estimate from the majority of the sub-models.

The confirmation of the adaptation of the model to new trends can also be confirmed from the model performance over the chosen years for analysis since between 2002 and 2013 it can be found diversified behaviors of the EUR / USD and despite this diversification, the model has de capability to adapt to each one of those cases.

VI. REFERENCES


