

# Adaptive Neuro-Fuzzy Inference System for Financial Trading using Intraday Seasonality Observation Model

A. Kablan

**Abstract**—The prediction of financial time series is a very complicated process. If the efficient market hypothesis holds, then the predictability of most financial time series would be a rather controversial issue, due to the fact that the current price contains already all available information in the market. This paper extends the Adaptive Neuro Fuzzy Inference System for High Frequency Trading which is an expert system that is capable of using fuzzy reasoning combined with the pattern recognition capability of neural networks to be used in financial forecasting and trading in high frequency. However, in order to eliminate unnecessary input in the training phase a new event based volatility model was proposed. Taking volatility and the scaling laws of financial time series into consideration has brought about the development of the Intraday Seasonality Observation Model. This new model allows the observation of specific events and seasonalities in data and subsequently removes any unnecessary data. This new event based volatility model provides the ANFIS system with more accurate input and has increased the overall performance of the system.

**Keywords**— Adaptive Neuro-fuzzy Inference system, High Frequency Trading, Intraday Seasonality Observation Model.

## I. INTRODUCTION

HIGH frequency trading is a new trend in financial trading. It is concerned with the making multiple trading decisions at extremely short time intervals, often using tick-data as input. This can be physically un-achievable with human traders, hence many algorithmic trading systems are being developed to implement high frequency trading strategies.

Financial investors and experts have always tried to trade and forecast the movement of markets. Current market information, news, and external factors affect the investors trading decisions of buying and selling. However, since the advent of modern financial markets, various financial theories, such as the Dow Theory, state that a financial market usually tends to follow a pattern [16]. This pattern is hard to recognize, notice, or to categorize and it applies to all financial markets [16]. Implementing a system that would provide a means of forecasting and trading of the markets would therefore help in improving an investor's financial trading record [15]. However, the *efficient market hypothesis* states that the current price contains all the available information in the market. This leads to the predictability of most financial

time series, such as stock prices or indices, being a rather controversial issue. Furthermore, financial time series are often very noisy and a filtering process should follow in order to remove such noise from the signal [2]. Moreover, traditional volatility measures have shown major drawbacks since they perform their calculations for the standard deviation in fixed times intervals, in contrast to the real life situation where volatility should be calculated from an event based perspective and observations made after a certain event has occurred.

Developing such a system which can outperform both human traders and also available automated traders would also eventually explain how the markets behave. The success rate of such a system would be determined by how accurate it is in predicting the movement of the next trading periods in seconds, minutes, hours, days, weeks, or months depending on the trading frequency. Taking the correct decisions after processing all the inputs from other blocks is also essential to a successful system.

Designing an expert system for such an unpredictable and unstable entity (financial markets) is a complicated task and various approaches can be followed. [1] has proposed the use of Adaptive Neuro-Fuzzy Inference Systems (ANFIS) for high frequency trading and forecasting. [1] displayed various promising results which allow further investigation in the area. However, it has been found that in order to fine tune this artificially intelligent trading system, volatility models must be extensively investigated. The trading system here is dealing with high frequency data, the data input to the system must be deseasonalised in a specific manner in order to eliminate unnecessary input in the training phase. However, in order to perform deseasonalisation, a new event based volatility model had to be proposed to observe specific events and seasonalities in the data and remove the ones that wouldn't be of use to the system. The resulting new volatility model that is proposed in this paper is the *Intraday Seasonality Observation Model* (ISOM). This model was inspired by the scaling laws in financial time series presented in [20,18]. The resulting new volatility model helped in maximising the efficiency of the training data set input to ANFIS and produced better results in its automated trading process

This paper is outlined as follows. In section II we briefly introduce the framework of the system, and how its various parts are interlinked. In section III we revisit the ANFIS

A. Kablan is with the centre of computational Finance and Economic Agents, University of Essex, Wivenhoe Park, Essex, CO4 3SQ, United Kingdom, (e-mail: akabla@essex.ac.uk).

structure implemented in [1] and used for high frequency trading. Section IV Introduces the state of art and ISOM and displays the empirical results of implementing ISOM. Section V provides a performance analysis of trading with ANFIS using data that has been deseasonalised using ISOM. Finally, concluding remarks are given in Section VI.

## II. FRAMEWORK

The main objective of this paper is to generate a very solid setup for high frequency trading using intelligent models. It combines artificial intelligence with volatility theory to come up with a framework that can produce high return from its trading strategies. Fig. 1 displays the general framework.

The first contribution is the use of Adaptive Neuro Fuzzy Inference System (ANFIS) for high frequency intraday trading. The second contribution is the introduction of the Intraday seasonality Observation Model (ISOM), a revolutionary new method of defining volatility using an event driven approach that takes in consideration directional changes within pre-specified thresholds inspired by the Scaling Laws of Finance [18,20]. An extension to ISOM is the Intraday Average Observation Model (IAOM), which provides an estimate of how the volatility of one day would look like. These distinct models will be interlinked to produce an optimized trading platform.

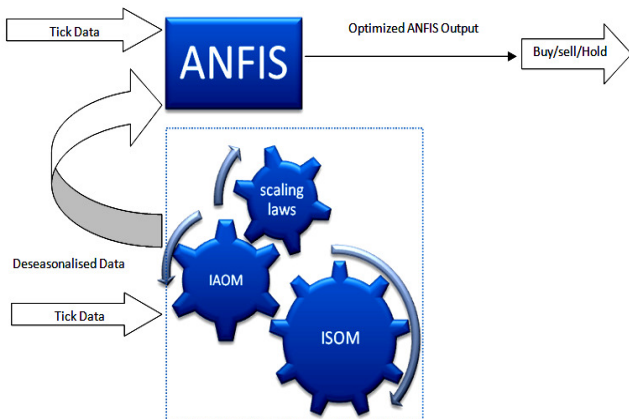


Figure 1. General framework of the system

## III. ADAPTIVE NEURO FUZZY INFERENCE SYSTEMS

Fuzzy inference systems are able to express aspects of human knowledge and reasoning processes in a linguistic fashion, avoiding the use of extensive quantitative analysis. Many types of fuzzy inference systems have been proposed in literature, however, in the implementation of an ANFIS for financial predictions and estimation the most suitable model is the Sugeno model. The Sugeno model makes use of if-then-rules to produce an output for each rule. Rule outputs consist of the linear combination of the input variables plus a constant term; the final output is the weighted average of each rule's output. The rule base in the Sugeno model has functional rules of the form:

$$\begin{aligned} \text{If } X \text{ is } A_1 \text{ and } Y \text{ is } B_1 \text{ then } f_1 &= p_1 * x + q_1 * y + r_1 \\ \text{If } X \text{ is } A_2 \text{ and } Y \text{ is } B_2 \text{ then } f_2 &= p_2 * x + q_2 * y + r_2 \end{aligned} \quad (1)$$

where X and Y are predefined membership functions,  $A_i$  and  $B_i$  are membership values, and  $p_i$ ,  $q_i$ , and  $r_i$  are the consequent parameters that are updated in the forward pass in the learning algorithm.

As we have already seen, fuzzy systems present particular problems to a developer:

- The if-then rules have to be determined somehow. This is usually done by 'knowledge acquisition' from an expert. It is a time consuming process that is weighed down by many problems.
- Membership functions. A fuzzy set is fully determined by its membership function. This has to be determined, for example if it is Gaussian then what are the parameters?

The ANFIS approach learns the rules and membership functions from data [7]. ANFIS is an adaptive network of nodes and directional links with associated learning rules. It is called adaptive because some, or all, of the nodes have parameters which affect the output of the node. These networks identify and learn relationships between inputs and outputs. ANFIS has been the adaptive network of choice to be investigated in detail and used for high frequency forecasting and trading due to its high learning capability and membership function definition properties.

The ANFIS architecture used in this paper is an improvement to ANFIS used in [1], where the root mean square error (RMSE) is used as a feed back to the network. This approach was adopted to avoid over filtration of original data. The architecture of ANFIS is shown in Fig. 2. The circular nodes represent nodes that are fixed whereas the square nodes are nodes that have parameters to be learnt. For the training of the network, there is a forward pass and a backward pass. We now look at each layer in turn for the forward pass. The forward pass propagates the input vector through the network layer by layer. In the backward pass, the error is sent back through the network in a similar manner to back propagation as shown in Table 1. Each layer is successively connected to the next layer and each layer represents a step of the fuzzy inference model explained in section II. The computational details of ANFIS at each layer are explained as follows:

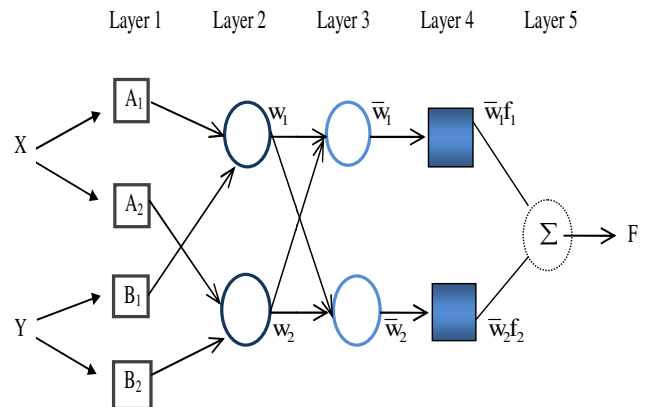


Figure 2. An ANFIS architecture for a two rule Sugeno system

Layer 1: The output of each node is:

$$\begin{aligned} O_{1,i} &= \mu_{A_i}(x) & \text{for } i = 1,2 \\ O_{1,i} &= \mu_{B_{i-2}}(y) & \text{for } i = 3,4 \end{aligned} \quad (2)$$

So, the  $O_{1,i}(x)$  is essentially the membership grade for  $x$  and  $y$ . The membership functions could be anything however experimental results lead to the conclusion that for the task of financial data training, the bell shaped membership function has to be used and is given by:

$$\mu_A(x) = \frac{1}{1 + \left| \frac{x - c_i}{a_i} \right|^{2b_i}} \quad (3)$$

where  $a_i, b_i, c_i$  are parameters to be learnt. These are the premise parameters.

Layer 2: Every node in this layer is fixed. This is where the t-norm is used to 'AND' the membership grades - for example the product:

$$O_{2,i} = w_i = \mu_{A_i}(x)\mu_{B_i}(y), \quad i = 1,2 \quad (4)$$

Layer 3: Contains fixed nodes which calculate the ratio of the firing strengths of the rules:

$$O_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad (5)$$

Layer 4: The nodes in this layer are adaptive and perform the consequent of the rules:

$$O_{4,i} = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \quad (6)$$

The parameters in this layer ( $p_i, q_i, r_i$ ) are to be determined and are referred to as the consequent parameters.

Layer 5: There is a single node here that computes the overall output:

$$O_{5,i} = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (7)$$

This is how the input vector is typically fed through the network layer by layer. We now consider how the ANFIS learns the premise and consequent parameters for the membership functions and the rules.

#### A. Learning Algorithm

Typically a gradient based method is used for the learning procedure of ANFIS, however this is known for its very slow

performance and its tendency to become trapped in a local minimum [11]. A hybrid learning algorithm has therefore been proposed for ANFIS.

This paper uses a standard hybrid learning algorithm proposed by [8], which uses a combination of Steepest Descent and Least Squares Estimation (LSE). It can be shown that for the network described, if premise parameters are fixed, the output is linear in the consequent parameters [13]. The learning algorithm consists of a forward pass and back propagation. In the forward pass, functional signals go forward till layer 4 and consequent parameters are identified by the least square estimate. In the backward pass, the error rates propagate backwards and premise parameters are updated by gradient descent. This approach guarantees the identification of global optimum point. Table 1 provides a summary of the learning methods.

#### B. ANFIS for Financial predictions and trading

The proposed system makes use of the above mentioned architectures, where the training and checking data to be learned will be high frequency financial data series. The system takes the price series as input. The values of 500 data points are first taken for system training and generation of the initial fuzzy inference system. This generates an ANFIS that has modified its parameters and membership functions. ANFIS is now ready to produce a prediction for the next data points, given the pattern (trend) that it has recognized.

##### 1) The Data

All simulations in this paper have been carried out on foreign exchange (FOREX), high frequency, EUR/USD, 5 minute, intraday data between 04/04/2006 and 01/09/2007. This dataset of 5-minute price data has been split into sets, each comprising 1000 data points. The first 500 data points in each set were used for system training; the remaining 500 points were used for checking the system's performance and updating the network structure using the output error. Once this was done, and also to test the system, another group of data was fed to the system. This out-of-sample group consists of 400 sets (1000 data points for each set) of data. The sets are widely spaced in the dataset so as to eliminate any possible data autocorrelation. Table 3 reports the overall average performance of the system compared against a traditional buy and hold strategy.

TABLE I  
SUMMARY OF TRAINING ALGORITHM

	Forward pass	Backward pass
Premise parameters	Fixed	Gradient descent
Consequent parameters	LSE	Fixed
Signals	Node o/p	Error rates

FOREX data was chosen for the system as it reflects most statistical properties found in financial time series [15]. Stationarity hypothesis states that past returns do not necessarily reflect future performance; this is true for FOREX data [14]. In addition, stylized facts such as gain/loss asymmetry and heavy tails are observed in FOREX returns distributions [14]. FOREX market is a 24-hour market where

there is high liquidity and volatility with three major centers in different parts of the world: New York, London, and Tokyo. It is highest in volatility during the early morning New York time, because both banks in London and New York are open and simultaneously trading. Commercial banks, corporate, funding and retail institutions from around the globe participate in FOREX trading. The price at the FOREX market is formed by buying and selling currencies to institutions, traders, exporters, importers, portfolio managers, and tourists. Each currency has two prices: a bid price at which a trader is willing to buy and an offer price at which a trader is willing to sell. If being in the major money centers banks traders deal in two way prices, for both buying and selling. In market-making banks worldwide much of the trading take place by direct dealing, while the rest takes place through brokers. Nowadays the buy and sell orders are matched via computerized services electronically using automated brokerage terminals. Grabbe quotes that around 85 percent of all FOREX trading happens between market makers [9]. It is a known fact that the most trading takes place between market makers which creates a space for speculative gains and losses. However, speculation in the FOREX market is potentially a zero-sum game: This means that the cumulative profits equal the cumulative losses. Dormael explains that today traders can trade against each other instead of trading against central banks as they did when currencies were not floating [10]. This paper presents the results of feeding ANFIS with EUR/USD intraday FOREX data for both training and checking a sample of this data is shown in Fig. 3. The resulting system will be able to perform predictions on the next move in the market.

### 2) Training and checking data sets

The aim of the proposed system is to produce one that is able to perform a prediction on the financial time series in high frequency using intraday data. The time series used in our simulation is 5-minutes intraday trading prices. The task is to use known values of the time series, which will be used as training data up to the point  $x=t$ , to predict the value at some point in the future  $x=t+S$ . Once that is achieved, a successful trade can be placed, e.g., a buy if the price is predicted to rise, or a sell if the price is predicted to go down.

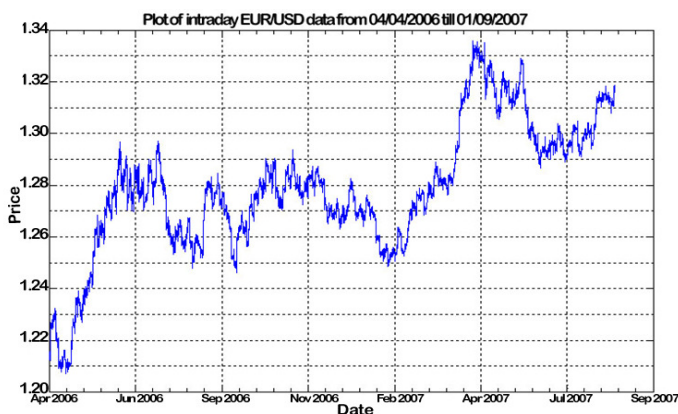


Figure 3. EUR/USD data set sample

The standard method for this type of prediction is to create a mapping from  $\tau$  points of the time series, spaced by  $\Delta$  time intervals, to predict the value at some point  $S$  in the future.

$$(x(t - (\tau - 1)\Delta), \dots, x(t - \Delta), x(t)) \rightarrow x(t + S) \quad (8)$$

In this particular case therefore, the system looks at the past three price observations in the market and the current observation to predict the next price observation. This is then used as a movement indicator (either up or down). Hence the value of  $\tau$  will be  $\tau = 4$  and  $S = 1$ . The extracted 1000 input-output data pairs are of the following format:

$$[x(t - 3), x(t - 2), x(t - 1), x(t); x(t + 1)] \quad (9)$$

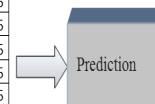
The idea here is to feed the system prices at each point as illustrated in Table 2. In order to make a prediction for  $t+1$ , the system will be fed the current price at time  $t$  plus the previous three price observations  $t-1$ ,  $t-2$ , and  $t-3$  respectively. This methodology will be used to feed the the Neuro Fuzzy system with EUR/USD intraday forex data for both training and checking. The resulting system will be able to perform predictions on the next move in the market. The next section shows the simulation results obtained using the above architecture and methodology

### C. Simulation Methodology

The developed Adaptive Neuro Fuzzy Inference System has been tested using EUR/USD high frequency tick data. The system has also been back tested on other sets of stock data to check whether the ANFIS approach would be efficient in performing data series predictions in order to place the right position. Now that a system to “predict” the movement of the market has been implemented (Fig. 4), a suitable position can be opened according to the indicator of this prediction. The implementation of a simple strategy was first proposed as in Listing 1 where, if it predicts an upwards movement it will buy, and if it predicts a downward movement it sells.

TABLE II  
 ANFIS TRAINING INPUT IS THE LAST FOUR PRICE OBSERVATIONS

1.2135	1.21365	1.21365	1.2135	1.2135
1.21345	1.21365	1.2135	1.2135	1.21345
1.2133	1.2135	1.2135	1.21345	1.2133
1.21325	1.2135	1.21345	1.2133	1.21325
1.21325	1.21345	1.2133	1.21325	1.21325
1.21305	1.2133	1.21325	1.21325	1.21305
1.21315	1.21325	1.21325	1.21305	1.21315
1.21305	1.21325	1.21305	1.21315	1.21305
1.21295	1.21305	1.21315	1.21305	1.21295
1.21295	1.21315	1.21305	1.21295	1.21295
1.21295	1.21305	1.21295	1.21295	1.21295
1.2131	1.21295	1.21295	1.21295	1.2131
1.21305	1.21295	1.21295	1.2131	1.21305
1.213	1.21295	1.2131	1.21305	1.213
1.213	1.2131	1.21305	1.213	1.213
1.2129	1.21305	1.213	1.213	1.2129





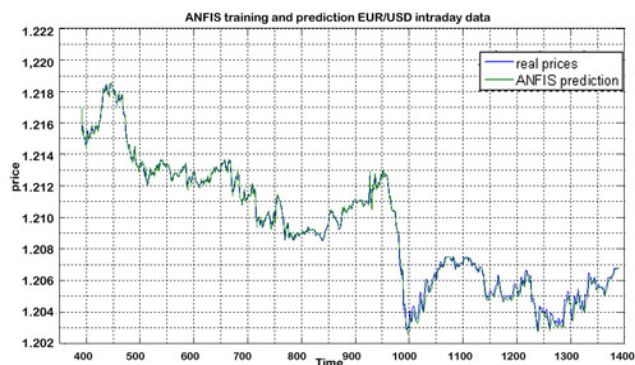


Figure 4. Initial prediction results

---

```

If Prediction is up then buy
else if prediction is down then sell
    
```

---

Listing 1: Simple buy/sell strategy

It was observed that having a very high accuracy rate and opening the right position didn't always translate into higher returns due to transaction costs. Hence the next modification to the system was to buy if the price was predicted to go up and hold the position as long as the price is going up as in Listing 2. This has improved the performance of the system; however it still did not provide a considerably high rate of return.

---

```

If Prediction is up then buy
    If next prediction is up then hold
else if prediction is down then sell
    If next prediction is up then hold
    
```

---

Listing 2: Introducing the hold position when prediction doesn't change in direction

Fig. 5 shows the above described strategy, where hold positions are introduced and the buy sell frequency is reduced. Finally in order to increase the return of the trading investment, a final prediction and trading strategy was introduced where a "trigger" value is used. Therefore, for a sequence of buy and hold positions, should the prediction of the next time sample fall below the set trigger, the position is

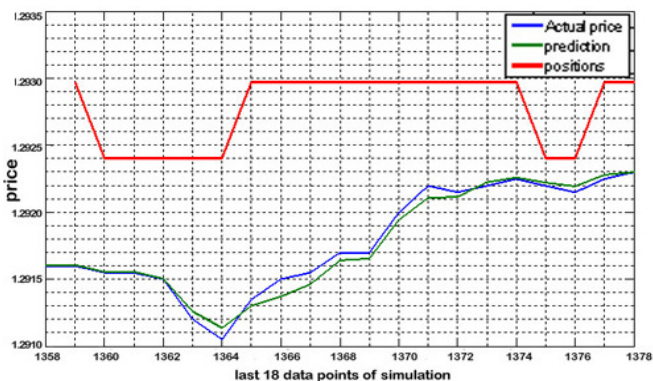


Figure 5. Introducing the hold position when the new prediction is equivalent to the previous

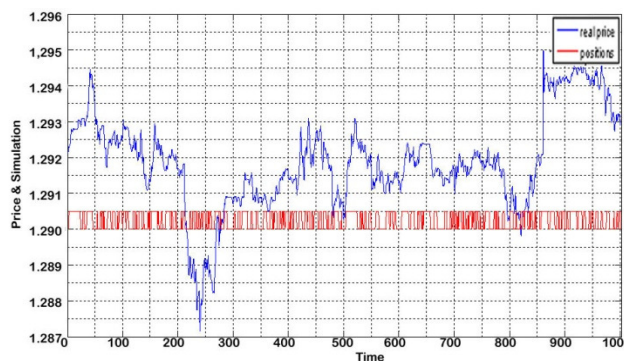


Figure 6. Positions as price evolves

closed and hence a sell position is opened. The trigger value is updated after each iteration as in Listing 3. Initially this trigger is set to the first value in the data set.

---

```

trigger = price (1)
if prediction is up
    and prediction > trigger
        Then trigger = prediction(now-1)
        position = buy;
else if prediction is down and prediction < trigger
        Then trigger = prediction(now-1)
        position sell;
    
```

---

Listing 3: introducing the trigger to track the prediction and detect directional changes to adjust position.

The initial performance of the above ANFIS system has been measured relative to some major benchmarks. The results are presented in the next section.

#### D. Initial Performance and analysis:

The benchmarks used for defining the ANFIS performance are defined as follows:

##### 1) Profit Factor

Profit Factor is an indication describing the profitability on an investment. The definition is:

$$\text{Profit Factor} = \frac{\text{Gross Profit}}{\text{Gross Loss}} \quad (10)$$

The profit factor mainly describes the historic profitability of a series of trades on an investment. The break-even of the profit factor is 1, referring to an investment that generates trades with a 50% chance of the gross sum of winning trades and a 50% chance of the gross sum of losing trades. Normally, investors pick investments that boast profit factors that are larger than one.

##### 2) Return On Investment (ROI)

ROI is used to evaluate the efficiency of an investment or compare returns on investments, that is, ROI is the ratio of profit gained or lost on an investment in relation to the amount of cost invested. ROI is defined as:

$$ROI = \ln \frac{V_f}{V_i} \quad (11)$$

Where:  $V_f$  is the initial investment value and  $V_i$  is the final investment value.

ROI is a very popular benchmark due to its flexibility and simplicity. In other words, ROI may be adjusted to suit different situations. ROI includes all the income results from the investment and the profit earned on selling the investment. If the final investment value, which includes sale price and any income, is higher than the initial investment value then ROI is positive. In the case of the opposite, ROI is negative.

### 3) Sharpe Ratio

The Sharpe ratio is used to measure the risk-adjusted return of an investment asset or a portfolio. This tells investors how well the return of an asset compensates investors for the risk taken. In other words, the Sharpe ratio can tell investors whether the returns of an asset or a portfolio come from a smart trading strategy or excess risk. The Sharpe ratio is defined as follows:

$$\text{Sharpe Ratio} = \frac{R_p - R_f}{\sigma_p} \quad (12)$$

where  $R_p$  is the expected return,  $R_f$  is the risk free rate and  $\sigma_p$  the standard deviation of a portfolio

Investors often pick investments with high Sharpe ratios because the higher the Sharpe ratio, the better its risk-adjusted performance has been.

### 4) Sortino Ratio

The main difference between the Sharpe ratio and the Sortino ratio is that the Sortino ratio only penalizes downside volatility while the Sharpe ratio penalizes both upside and downside volatility. The Sortino ratio is defined as:

$$\text{Sortino Ratio} = \frac{R_p - R_f}{\sigma_d} \quad (13)$$

where  $R_p$  is the expected return,  $R_f$  is the risk free rate and  $\sigma_d$  the standard deviation of the portfolio. Thus, the Sortino ratio measures the risk premium per each unit of downside risk in an investment asset or a portfolio.

The results of the system's performance have been compared to a traditional buy and hold strategy. Table 3 displays results that are very promising. Positive Sharpe and Sortino ratios show that the system has not taken high risk for the amount of return gained. Also ROI is 27% which is a high rate for a 1 year and a half investment described for this paper. Similarly, profit factor and winning rate show high accuracy in the prediction rate of the system.

TABLE III  
 ANFIS PERFORMANCE

	Winning Rate	Profit Factor	ROI	Sharpe Ratio	Sortino Ratio
Neuro Fuzzy system	0.651	2.3	0.27	0.19	0.201
Buy and Hold	0.42	1.1	0.09	-0.07	-0.05

The use of Adaptive Neuro-Fuzzy Technology for high frequency data trading is still an unexplored field. This paper presented an approach to trade intraday high frequency data using ANFIS. This has produced a promising application that encourages more development and expansion in the area, that is, more research has to be done. In addition, the training and learning algorithms can be improved and the use of filtration methods in the learning process has to be investigated, implemented and tested to observe more results.

## IV. VOLATILITY MODELLING

Analysing the results of the implemented ANFIS, it was concluded that the system design is theoretically very consistent in what it is delivering. However the question that can arise at this stage is what if the data set being fed to ANFIS is not as efficient as it is supposed to be. It is hence suggested that ANFIS would deliver much better results given that the data sets and the training sets were adjusted in a manner that would avoid seasonality patterns and volatility effects.

The concept of volatility of asset prices and returns is one of great weighting in financial markets. Hull [17] defines the volatility of a stock as a measure of our uncertainty about the returns provided by the stock. Generally, volatility measures the expected variability of the price of an asset or commodity over a specified period of time. This may be calculated as the standard deviation of the percentage price change over a 24 hour period. Multiplying by the square root of the number of trading days in one year (252 trading days) converts this daily deviation to an annual volatility.

Historical (or statistical) volatility measures past stock price movements and tells how variable a stock's price has been over a past time period such as 20 days or 100 days, etc. Historical volatility is measured by the standard deviation of stock prices. Using historical data to estimate volatility enables us to observe stock prices at monthly and daily intervals of time.

$$u_i = \ln \left( \frac{S_i}{S_{i-1}} \right) \quad (14)$$

The standard deviation of  $u_i$  is given by

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (u_i - \bar{u})^2} \quad (15)$$

where:  $u_i$  is the logarithmic return,  $S_i$  is the stock price at the end of  $i$ -th interval, and  $\bar{u}$  is the mean of  $u_i$ .

Depending on the range of data being used (daily, monthly, etc.), this must be converted to annual volatility data. The annual standard deviation (annual volatility) of  $u_i$  is:

$$\sigma = \sigma_{monthly} \times \sqrt{12}$$

(16)

or

$$\sigma = \sigma_{daily} \times \sqrt{252}$$

Higher historical volatility means that the stock has fluctuated more over the period in question.

Implied volatility is a concept used for the measurement of expected future stock price volatility and is derived from the current price of options. It provides information about how the options market expect the stock price to be in the future. Implied volatility is calculated from the current option price, using an option-pricing model (such as the Black-Scholes Model). The implied volatility is defined as the standard deviation that causes the pricing model to compute the current option price.

#### A. Redefining Volatility

A Major contribution in this work will be to hopefully redefine the way volatility is defined, where the usual approach in measuring the stock returns is to check observations at certain **fixed time intervals** in order to calculate the return. However this new approach does not actually cater for the current time interval as much as it caters for the events that might have occurred at that time. Once a new volatility model is implemented, it will be used as another input to ANFIS.

The idea is that everything will be looked at from a scaling law perspective of directional changes (events), where each up/down percentage changes within a pre-specified threshold are observed, the time stamps and the relative price are marked, and all data is iteratively stored in bins of time value which will then be analysed further.

The above takes inspiration from the fact that time itself is an entity that is hard to define, nor to explain. We have both Physical time and Intrinsic time, Professor Olsen [18] can be quoted speaking about different types of time by saying "Sometimes time goes by quickly, and sometimes it drags on and on," "When you see the market collapse 5% in a half-hour, for example, Physical time has little meaning". Here in this project we are not concerned about physical time, we are concerned about intrinsic time and the events that happen within a specified time, the impact these events might have had, and the volatility that arises because of these events.

#### B. Scaling Laws

As it has been shown in [18,20], there is great evidence that there are scaling laws that govern the way how the markets move. This has led to the observation that "for any given value of a pre-specified threshold (percentage of market movement)  $dx(\%)$ , a number of directional changes  $N(dc)$  can be observed, and as the value of the threshold increases, the

number of these directional changes would decrease forming a straight line". This relation can be best described as follows:

$$N(dc) = \left(\frac{\Delta x_{dc}}{C_{dc}}\right)^{E_{dc}} \quad (17)$$

With  $X_{\delta x}$  and  $E_{\delta x}$  being currency dependent constants. Fig. 7 shows the scaling laws being very visible in various currencies' price changes.

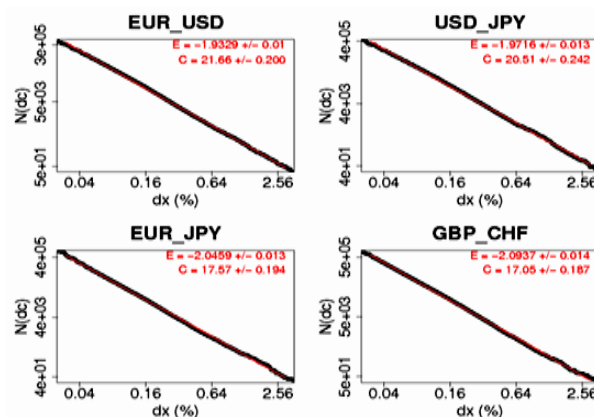


Figure 7. Scaling Laws for various currencies as shown in [18],[20]

#### C. Intraday Seasonality Observation Model (ISOM)

Given inspiration from the above Scaling Laws, one of the main contributions and findings of this paper is the development of a model for observing the seasonalities that would arise from the counting these directional changes have been designed.

In its simple definition, the ISOM is a model that takes in consideration a certain threshold  $dx(\%)$  and will observe the timings where the directional changes  $N(dc)$  occur, it would iteratively and consecutively parse through the whole data set of prices and will save the observations into their respective **time bins**, this would at the end give a horizon of seasonalities, pointing out exactly the times of day where these observations have been made, which would mean the times of day where the volatility was high or low. For the purposes of illustration, the threshold that was used in this paper was the 0.05% directional change. Hence, if we know that the scaling law holds, this will mean that the seasonality patterns will look similar irrespective to what threshold we are using in the experiment.

What is interesting about this approach is the fact that the time bins can vary to observe events at any desired frequency during the day. In other words, we can have 24 bins to observe at any hour during the day, or 48 bins to observe at every half an hour during the day, and so on and so forth we can even make observations for every five minutes during the day.

All the design and experiments on this model have been carried out using high frequency intraday data of EUR/USD prices as illustrated in section III. The resulting method and deseasonalisation process can apply to the price series of all stocks, commodities or even indices.

The ISOM for a particular time in the day at a particular threshold is equal to the total number of events that happened at that time in the whole data set.

$$ISOM_{t|dx(\%)} = \sum_{i=1}^n N(dc)_i$$

Where;  $t$  is the the particular time of day,  $dx(\%)$  is the particular threshold,  $n$  is total number of days in the data set and  $N(dc)$  is number of directional changes (events).

The ISOM model can be regarded as a proxy for volatility, due to the fact that it clearly maps the times of day with their respective volatility; which is looked at from an event based perspective, where each directional change with a specific threshold is an event. When applied for a 0.5% percentage change for a 24hrs bin for the EUR-USD data set between 04/04/2006 and 01/09/2007, the ISOM model resulted in the seasonality pattern shown in Fig. 8.

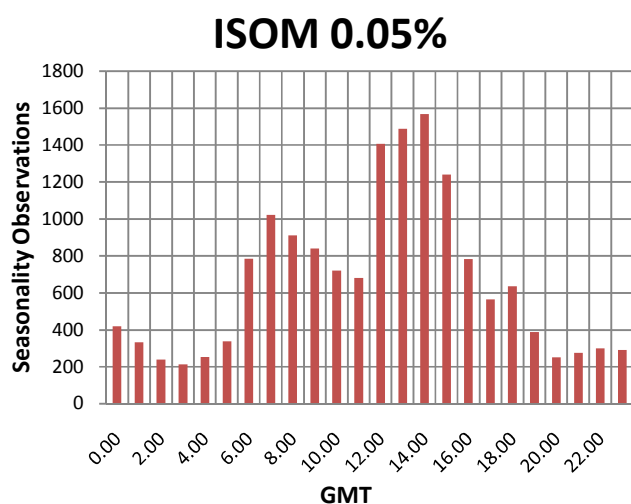


Figure 8. the intraday seasonality observation model for a threshold of 0.5% for the EUR-USD between 04/04/2006 and 01/09/2007

The analysis to the above plot is that most of the activity or **events** (i.e. volatility) happen between 12pm and 2pm GMT which confirms the fact that these are the times where the announcements are made and the markets reaction to these announcements take place, it also shows that even though it is lunch time in Europe, the few traders that are active tend to trade aggressively which drives the volatility to high peaks. Other times of high volatility are between 7am and 8am GMT which is usually the time before the markets open in europe.

An interesting observation takes place when the ISOM is applied to the same data set but the observations have been taken every 30 minutes which resulted in 48 bins of observations (24 x 2) the results are shown in Fig. 9.

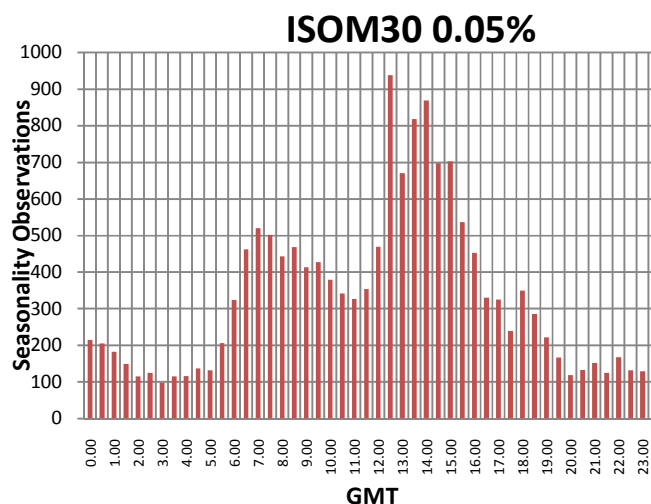


Figure 9. the intraday seasonality observation model for a threshold of 0.5% taken on half daily basis (every 30 minutes) for the EUR-USD between 04/04/2006 and 01/09/2007

Fig. 9 shows that when considering observations every 30 minutes the time with the highest volatility is between 12:30pm and 1pm GMT which is again the time when all the announcements that have been made at 12pm have been absorbed by the markets and the traders have started acting upon them. It can be seen that the period of the highest volatility is between 12pm and 4pm GMT i.e. the times that include the announcements, the open of the US markets until the close of the European Markets. The above results have been impressive in the sense that they have confirmed real life events that are known to increase markets volatility, Also they can help the trader or the system ignoring the periods with the low events happening. It has to be noticed that the ISOM model can be applied to any threshold and any time frequency (daily, half daily, quarter daily, five minutes etc.), we have taken a threshold of 0.5% and used it on daily and half daily observations just for the scope of illustration. More experiments on other thresholds have been carried out and confirmed the above, which also confirms that a scaling law does hold. Hence the ISOM concept can be applied freely to any threshold or any time frequency.

Finally the above results have to be averaged to provide an illustration of how volatility of a single day would look like, in order to come up with a standard model to capture the times with the highest or lowest volatility (number of events). This will make use of The Intraday Average Observations Model (IAOM).

#### D. The Intraday Average Observations Model (IAOM)

The Intraday Average Observations model (IAOM) is of the average observations category that was introduced by [19]. The method introduced in this paper differs from [19] in the fact that IAOM was based on ISOM which captures seasonalities using scaling laws. Hence now that ISOM have been proposed in this project, IAOM will be used to deseasonalise the patterns found by ISOM and to give the



intraday average volatility estimate using the following equation:

$$IAOM_{t|dx(\%)} = \frac{1}{M} \left( \sum_{i=1}^M N(dc)_i \right)$$

Where;  $t$  is the particular time of day,  $dx(\%)$  is the particular threshold,  $M$  is the total number of days in the data set, and  $N(dc)$  = number of directional changes (events).

The idea here is to take the average number of observations throughout the whole sample in order to produce a model that would estimate the average number of average observations or events (directional changes within a specified threshold) that would occur per day. It is basically a model that would average the ISOM to give an idea on how an average day would look like. Figure 10 shows the average number of observations for the EUR-USD between 04/04/2006 and 01/09/2007. Analysing this figure, it confirms the observations of ISOM and it gives an indication that the highest volatile time is 14:00-15:00 GMT, it also shows that sometimes such as in between 0.00-5:00 GMT can have negligible impact on the volatility since it did not include many events, hence they can be filtered out when feeding data for the training of ANFIS.

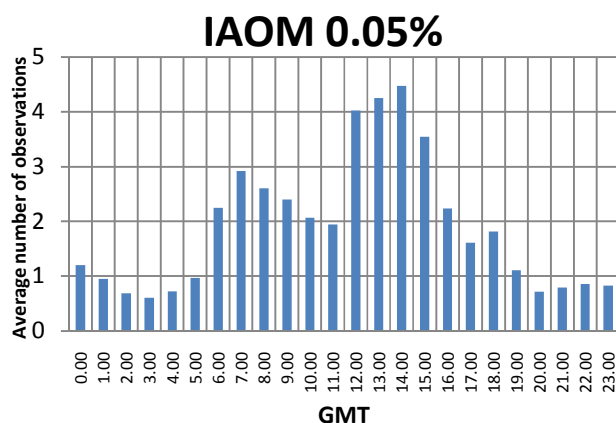


Figure 10. the intraday average observation model for a threshold of 0.5% for the EUR-USD between 04/04/2006 and 01/09/2007

The 30 minutes IAOM model shown in figure 11, shows that when we take in consideration 30 minutes time intervals, the maximum period where we have observations is again 12:30-13:00 GMT, the figure also shows the times with negligible events. ISOM and IAOM model provide a different measure of the volatility of financial time series by observing the events at specific times and capturing the ones that we are interested in by varying the threshold to any preferred value.

Next is the use of this methodology to feed the ANFIS with new cleaner training data and observe the results that are obtained.

## IAOM30 0.05%

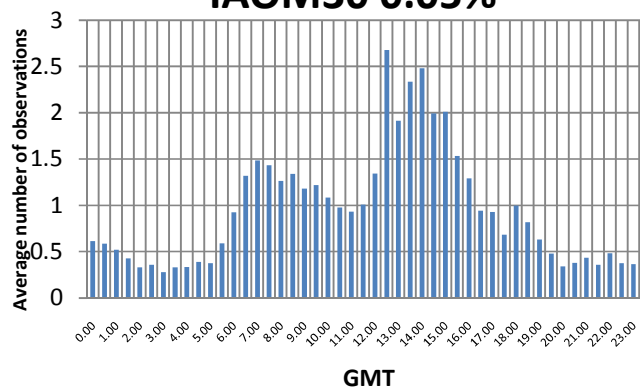


Figure 11. The intraday average observation model for a threshold of 0.5% taken on half daily basis (every 30 minutes) for the EUR-USD between 04/04/2006 and 01/09/2007

### V. OPTIMIZATION OF ANFIS USING ISOM:

It has been shown in section III, ANFIS is an inference system that had shown very good performance when used for predicting and trading high frequency data. ANFIS has been tested in previous sections on high frequency 5-minute EUR/USD data and produced very good results. The idea now is to use of the intraday seasonality observation model to filter and clean this data by pointing out periods of the day where the volatility has been above a certain range (number of events). This will be observed on a daily basis using the Intraday Average Observation Model.

In order to perform this final task, the ISOM model has been redesigned to cater for 5 minute data instead of hourly or 30 minute data as shown previously. We now have bins of 5 minute data and we will capture the directional change as they occur within these bins, where the counter of events will increase according of the number of events, and the number of times the threshold has been exceeded.

ANFIS is now fed data from times of day where the number of observations exceeded 3 events. This means that ANFIS will be trained on data with higher volatility (stress training) and will then be allowed to perform prediction on a set of checking data using the same methodology explained in section III.

Table 4 displays the new results of trading with ANFIS using deseasonalised data from ISOM and IAOM, and this is compared with both the initial ANFIS system and the traditional buy and hold strategy.

TABLE IV  
 ANFIS RESULTS USING FEEDBACK FROM ISOM MODEL

	Winning Rate	Profit Factor	ROI	Sharpe Ratio	Sortino Ratio
Neuro Fuzzy system	0.651	2.3	0.27	0.19	0.201
Buy and Hold	0.42	1.1	0.09	-0.07	-0.05
Neuro Fuzzy system using ISOM	0.72	2.7	0.33	0.22	0.20

The above results show a good improvement against the results in Table 3 in the winning rate, ROI, Sharpe ratio and Sortino ratio results. These results are a good indication that feeding ANFIS with “cleaner” data, would result in higher performance and better results. These results give indication that if ANFIS is fed with a large amount of training data (more than one year) of EUR/USD, and then we use the proposed volatility modelling system to filter out unwanted times of the day, the performance would increase significantly.

#### VI. CONCLUSION AND FUTURE WORK

This paper has investigated various areas of computational finance. The distinctive area of soft computing and artificial intelligence has been covered in this paper by developing the Adaptive Neuro Fuzzy Inference System ANFIS.

Secondly, a new promising volatility observation model has been proposed, implemented, and tested. It has displayed very good results which encourage further research in the area. The Intraday Seasonality Observation Model (ISOM) proposed in this paper was inspired by the new scaling laws of finance and has been tested on various levels of thresholds. The observation of a directional change within a threshold leads to the taking of the time stamp and its consequential addition to all the observations that have been made on that time. The power of this method is in the fact that any threshold can be used for any time frequency which leads to the observation of events for the whole data series from a new perspective. Such system, when combined with an already good performing automated trading system such as ANFIS, can provide very good results and can increase the returns of a trader's portfolio.

#### ACKNOWLEDGMENT

The Author would like to thank Professor Richard Olsen, for his kind support in providing guidance on the newly discovered scaling laws in finance and for providing access to the *Olsen Routes*<sup>®</sup> platform to implement ISOM and IAOM. To Author would also like to thank Professor Dietmar Maringer for his guidance during the process of implementing ANFIS.

#### REFERENCES

- [1] A. Kablan. Adaptive Neuro Fuzzy Inference Systems for High Frequency Financial Trading and Forecasting. Proceedings of The Third International Conference on Advanced Engineering Computing and Applications in Sciences. 2009.
- [2] E. F. Fama. Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, pages 383–417, May 1970.
- [3] T. Hellström and K. Holmstrom. Predicting the Stock Market. Technical Report Ima-TOM-1997-07, Center of Mathematical Modeling, Department of Mathematics and Physics, Malardalen University, Vasteras, Sweden, August 1998.
- [4] L.C. Ying and M.C. Pan. Using adaptive network based fuzzy inference system to forecast regional electricity loads", *Energy Conversion and Management*, Vol. 49, pp. 205-211, 2008.
- [5] T. Takagi and M. Sugeno. Fuzzy identification of systems and its application to modeling and control, *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 15, pp 116-132, 1985.
- [6] R. Wilson and R. Sharda. Bankruptcy prediction using neural networks, *Decision Support Systems*, Vol. 11, pp 545-557, 1994.
- [7] Y. Yoon, T. Guimaraes, and G. Swales. Integrating neural networks with rule-based expert systems, *Decision Support Systems*, Vol. 11 pp 497-507, 1994.
- [8] R. Jang. ANFIS: Adaptive network-based fuzzy inference system, *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 23 (3) pp 665-685, 1993.
- [9] O. J. Grabbe (1996): *International Financial Markets*, Englewood Hills, Prentice Hall Inc.
- [10] A. V. Dormale, (1997): *The Power of Money*, Macmillan Press, London.
- [11] S. Banik. Modeling chaotic behavior of Dhaka Stock Market Index values using the neuro-fuzzy model, 10th international conference on Computer and information technology ICCIT. Issue , 27-29 Dec. pp 1-6, 2007.
- [12] S. Chabaa and A. Zeroual. Predicting Packet Transmission Data over IP Networks Using Adaptive Neuro-Fuzzy Inference Systems. *Journal of Computer Science* Vol. 5(2), pp 123-130, 2009.
- [13] Y. Becerikli, "Optimal Estimation Theory" Lecture Notes, Kicaeli University, Turkey. 2005.
- [14] R. Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance* pp 223-236. 2001.
- [15] M. Dracogna, R. Gencay, U. Müller, and R. Olsen. 2001, *An Introduction to High-Frequency Finance*, Academic Press.
- [16] Murphy J: *Technical Analysis of Futures Markets*. New York, N.Y.: New York Inst. of Finance, 1986
- [17] J. Hull, Incorporating Volatility Updating into the Historical Simulation Method for VaR," *Journal of Risk* (Fall), Vol. 1, No. 1, 5-19, 1998.
- [18] R. Olsen. CCFEA Expert Seminar Series. University of Essex. United Kingdom. 6,13 and 20 March 2008.
- [19] Bauwens, L, Ben Omrane, W. Giot, P News Announcements, Market Activity and Volatility in the Euro/Dollar Foreign Exchange Market. *Journal of International Money and Finance* 24, 1108—1125. 2005.
- [20] J.B. Glattfelder, A. Dupuis, R.B. Olsen. An extensive set of scaling laws and the FX coastline. Unpublished manuscript. 2008.