Applying market profile theory to forecast Taiwan Index Futures market

Chiu-Chin Chen *, Yi-Chun Kuo, Chien-Hua Huang, An-Pin Chen

Institute of Information Management, National Chiao Tung University, 1001 University Road, Hsin Chu 30010, Taiwan, ROC

A R T I C L E   I N F O

Keywords:
Market profile
Technical analysis
Neural network
TAIEX futures

A B S T R A C T

This research applies a market profile to establish an indicator to classify the correlation between the variation in price and value with the stock trends. The indicator and technical index are neural network architecture parameters that assist to extrapolate the market logic and knowledge rules that influence the TAIEX futures market structure via an integral assessment of physical quantities.

To implement the theory of market profile on neural network architecture, this study proposes qualitative and quantitative methods to compute a market profile indicator. In addition, the indicator considers the variation and relevance between long-term and short-term trends by incorporating the long-term and short-term change in market in its calculation. An assessment of forecasting performance on different calculation approaches of market profile indicator and technical analysis is conducted to differentiate their accuracies and profitability.

The experimental results show the qualitative market profile indicator outperforms the quantitative approach in a short-term forecast period. In contrast, the quantitative market profile indicator has a better trend-predicting ability, thus it is more effective in the long-term forecast period. The integration of market profile and technical analysis surpasses technical analysis as a neural network architecture parameter by effectively improving forecasting performance and profitability.

1. Introduction

The Taiwan Futures Exchange (TAIFEX) was established in September 1997. The Taiwan Weighted Stock Index Futures (TAIEX futures) was launched in July 1998, and declared the official start of Taiwan’s futures market. A number of commodity futures electronic futures, financial futures, small TAIEX Futures were launched over a ten year period. Investors can select the appropriate investment vehicles depending on the degree of risk. Futures accounts are growing year by year. The futures market has tended to improve, becoming an important hedging and arbitrage method for Taiwan stock market investors with tools.

However, the futures market has become increasingly volatile, with the legal entity participation in the futures market gradually increased (Lien, Lim, Yang, & Zhou, 2013). In addition, the financial tsunami of capital withdrawals during share transactions increasing year by year, has become the main force leading the Taiwan stock market ups and downs. Studies have shown that, in the Taiwan stock market, the average retail investor withstands losses of about 3.5% per year, with corporate investors, due to rare information and chip advantages, can obtain a 1% after-tax return (Barber, Lee, Liu, & Odean, 2004). Yu and Huarng (2008) proved fuzzy time series models can forecast TAIEX futures markets (Yu & Huarng, 2008).

Steidlmayer (1984) proposed the market profile theory; refute the efficient market and random walk theory (Steidlmayer, 1984). At different time intervals, participants in different prices bid passive or active, leading to price movement rather than random development. Different participants have different thoughts and behavior preferences for the same price, so the market cannot meet the needs of each participant, without any prices representing fair value. In other words, the market is not efficient. Steidlmayer also pointed out that the risk and reward in the market is not a linear relationship (Roll, Schwartz, & Subrahmanyam, 2007). The asymmetric opportunities, irrational human investment behavior cause market fluctuations through an understanding of long-term (artificial person) and short-term (retail investors) market behavior and logic is able to predict changes in the market structure to reduce investment risk.

In the face of these non-linear problems, artificial intelligence (AI) methods learn the knowledge (Lin, Hu, & Tsai, 2012; Won, Kim, & Bae, 2012) and rules that can be effective in predicting an environment of uncertainty without the need to rely on subjective judgment is better than the traditional model (Desai, Desai, Joshi, Juneja, & Dave, 2011; Roon, de Nijman, & Veld, 2000). The neural

The market profile concept has been widely used in the financial decision-making field (Canoles, Thompson, Irwin, & France, 1998); however, there has been little direct research (Dalton, Dalton, & Jones, 2007). Therefore, this study used the market profile principle and technical analysis (Taylor & Allen, 1992), as back-propagation neural network (BPNN) input variables. A better model than the old learning model is constructed using only the technical analysis and a new research model to explore the market logic and knowledge rules (Edwards & Magee, 1997).

The market contour theory with technical analysis is extracted from the relationship between price and value using the NN knowledge of rule changes learning using the market logic and market structure (Grudnitski & Osburn, 1993). How the market profile is used as the BPNN input variables is the focus of this study. The experimental design involved observed variables to determine technical analysis and market profile information. The impact on the future Taiwan stock market trend, to further assess and validate the predictive ability of the different intervals, to provide an innovative investment tools for investors and future researchers as a reference.

2. Methodology

2.1. Research design

The TAIEX futures’ tick data used the Windows Mobile cut five minutes of the K-bar. The K-bar used the opening price, high price, closing price and low price, as calculated BPNN input variables of the original value. The original value is then calculated and the item pre-treatment in order to determine technical analysis and market profile indicators.

In the control group model the MACD and KD of the technical indicators are used as the NN input variables. In order to compare the market profile model prediction mechanism, its performance is superior to the only technical indicators as the NN input variables. The input variables were used to calculate the technical indicators and market profile indicator values. Numerical regularization was used to avoid uneven numerical distribution.

The output variable reward punishment mechanism is used to calculate the relative change range and grouping. The input and output variables are then input into the back-propagation neural network to learn and predict the results (Kimoto, Asakawa, Yoda, & Takeoka, 1990).

2.4. Windows mobile cutting

TAIEX futures tick data use to move the window to be cut, calculating the required 5 min of data, and cutting shown in Fig. 1. Therefore, every five minutes of the opening price, closing price, highest price, lowest, opening tick transaction price is the point in time \( t - 5 \), the closing price for the time point \( t \) tick transaction price and the highest price and the lowest that the computation time point \( t - 5 \) to \( t \) transaction price between the maximum and minimum values.

2.5. Calculate the input variables

2.5.1. Moving average convergence divergence (MACD)

This MACD indicator indicates the big band trend. DIF said that the amount of a small band of fluctuation. If the market shows an upward trend, the deviation in the line speed is gradually expanded, while the MACD is still moving along the trend, resulting DIF and MACD cross situation, namely buy signal; contrary can sell signal. Using the following formula:

\[
EMA(m)_i = EMA(m)_{i-1} + \alpha_m \times (C_i - EMA(m)_{i-1}) \\
EMA(n)_i = EMA(n)_{i-1} + \alpha_n \times (C_i - EMA(n)_{i-1}) \\
EMA(m)_i: \text{The } i \text{ day's long term EMA value} \\
EMA(n)_i: \text{The } i \text{ day's medium term EMA value} \\
C_i: \text{The } i \text{ day's closing price} \\
\alpha_m = \frac{2}{1+m} \quad \alpha_n = \frac{2}{1+n} \\
\text{DIF}_i = EMA(n)_i - EMA(m)_i
\]

2.2. Subjects

The subject of this study is the TAIEX Futures. The data source is the TAIEX days Tick transaction data provided by the Taiwan Futures Exchange, including trading hours, the transaction price, number of transactions and information.

Experimental samples during the study period from August 10, 2009 to 2010. Screening and pre-processing a total of 48,080 pen five minutes of trading information, data, contains the opening price, closing price, highest price, lowest price.

During the experiment and verification inverted propagation neural network (Watanabe & lwata, 2009), the information should be divided into training and testing during the training period for the conduct of online learning, according to Kearns (1996) described the input data to 80% training period, 20% for the test period split for the ideal proportion (Kearns, 1996).

2.3. Data collection procedure

Data pre-processing, the first Windows Mobile the TAIEX days tick transaction information, cutting the required five minutes of data were calculated, and five minutes of data output and input variables. The input variables were used to calculate the technical indicators and market profile indicator values. Numerical regularization was used to avoid uneven numerical distribution.

The output variable reward punishment mechanism is used to calculate the relative change range and grouping. The input and output variables are then input into the back-propagation neural network to learn and predict the results (Kimoto, Asakawa, Yoda, & Takeoka, 1990).
MACD_t = MACD_{t-1} + \frac{a_k}{2} \times (DIF_t - MACD_{t-1}) \quad (4)

DIF_i: The _i_ day's DIF

MACD_t: The _t_ day's short term MACD

The study long-term EMA 26 units, 12 units of short-term EMA, and used to the 9 units DIF calculate MACD, so use \( a_m = 0.0741 \), \( a_s = 0.1538 \), \( a_t = 0.2 \), substituted into the formula to calculate the MACD and DIF.

2.5.2. Stochastics

Stochastics (KD) value of _K_ and _D_ two curves, which general nine days cycle. RSV called immature random value represents the current closing nine days of market strength of the position. The formula is as follows:

RSV = \frac{C - \text{\( L \)}}{9H - \text{\( L \)}} \times 100\%

The numerator represents the distance from the lowest point the closing price of nine days and on behalf price rose from the lowest point pushed to the closing price. RSV can be regarded as a multi-force strength in the nine days. The _K_ value MACD RSV take three days. The _D_ value takes the _K_ value of the three days MACD. The initial values are all 50. KD calculated input variables, the value of _K_ and _D_ values in the range of 0–100. Formula is as follows:

\[ K_t = K_{[t - 1]} \times \frac{2}{3} + \text{RSV} \times \frac{1}{3}, \quad K[0] = 50 \]

\[ D_t = D_{[t - 1]} \times \frac{2}{3} + K_t \times \frac{1}{3}, \quad K[0] = 50 \]

2.5.3. Market profile indicators

This study is based on market profile theory from the establishment of indicator values to judge the relationship between prices, changes in the trend value as the BPNN input variables. The plotted market contour 15 K-bar (75 min) is used to calculate the Time Price Opportunities (TPO). The number distribution is compared to the current five minutes of the K-bar of the opening price, closing price, highest price, lowest price, and the relative position of the previous 15 K-bar (first 75 min), with technical indicators to predict if the future stock price will rise or fall with the magnitude. The calculated market profile indicators are shown in Fig. 2.

In order to calculate the current price relative position of the first 75 min used to indicate the price to deviate from the range of values and the extent of the price range. Therefore, this study both qualitative and quantitative calculated price deviation values calculated.

2.6. Calculate the output variables

The NN output variable as the future of the fluctuation of the stock price, but consider that if the forecast period in the wrong direction must be given punishment. Therefore, the output value of NN is subject to appropriate amendments. Fig. 3 is shown the n-K bar to the m-K bar change range.

Change rate of n-K bar = \( L_n - L_m \)

Relatively change calculation certainty future prediction time interval, change the strength of the two parties, the up side is stronger than the power of the down side, the next time interval aggregate judgment as a rising trend.

2.7. Trend in the direction of learning

After these relatively change range of output variables of the study found that a very short-term ups and downs of small amplitude. To predict the change range of NN learning the outcomes must be very precise. Therefore, this study is relatively change range coupled with clustering, divided up, and not up not down, or a total of three groups. Their scheming as follows:

<table>
<thead>
<tr>
<th>Output variables</th>
<th>Relative price change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,</td>
<td>relative price change &gt; 0</td>
</tr>
<tr>
<td>0.5,</td>
<td>relative price change = 0</td>
</tr>
<tr>
<td>0,</td>
<td>relative price change &lt; 0</td>
</tr>
</tbody>
</table>

The magnitude of this study is relatively ups and downs for the 0 variables re-designated as a value of 0.5, called up not down. The variables relatively up a drop of greater than 0 is re-designated as a value of 1, is called up. The relatively up a drop of less than 0 variables reassign a value of 0, is called down. Use of clustering, NN only need to focus on the trend in the direction of learning.

2.8. Back-propagation neural network model

Vellido et al. (1999) referred to the NN parameter setting and no restrictions. General literature, advice or other methodology determined by experts. In the hidden layer, the number of settings, Zhang, Patuwo, and Hu (1998) proposed as long as a single hidden layer types of NN to achieve reliable learning level and a single hidden layer is currently the most popular architecture (Zhang et al., 1998), therefore this study was inverted pass the NN of the hidden layer is set to a layer. On the hidden layer node number decision, this study used Matlab software BPNN and the default parameters.

2.9. Performance appraisal model

Two evaluation methods used in this study to measure the performance of the model simulated trading, respectively for the accuracy and profitability. The different model between performance benchmark for comparison, its assessment methods are described below:

2.9.1. Accuracy

BPNN completed forecast, trading strategies based on the predicted direction, the predictive value is greater than the absolute value of the threshold, if the results of the transaction profit is greater than zero, the judgment of the transaction in the right direction. The accuracy of the forecast period profit is greater than
zero items divided by the total number of transactions. The formula is as follows:

\[
\text{Accuracy} = \frac{\text{The correct number of transactions}}{\text{The total number of transactions}}
\]

2.9.2. Profitability

For the profitability of the assessment, the study in the assumption of adequate margin TAIEX futures empirical trading and the futures profit points to calculate the profit and loss. Contains a bilateral transaction tax and fee total of 2 points to calculate each port of futures contracts on the transaction costs. The formula is as follows:

\[
\text{Average profits and losses (per port)} = \frac{\text{Total profits and losses}}{\text{The total number of transactions}}
\]

3. Experimental results and analysis

3.1. Experimental results

3.1.1. EG A – the qualitative the market profile indicators (long-term) + technical indicator

The EG A qualitative way to calculate the long-term 75-min market profile indicators, coupled with technical indicators as input variables. Respectively, to the prediction of 5, 15, and 30 min after the stock change range amplitude following experimental results shown in Table 1.

The experimental results show that the prediction time highest accuracy of 30 min, up to 76.43%, followed by 15 min. Corollary within five minutes of the market is likely to face short-term price shocks than 75 min the market profile and technical indicators as input variables, cannot accurately predict the short-term direction of the trend, EG A higher long-term forecast.
3.1.2. EG B – the quantitative the market profile indicators (long-term) + technical indicator

The EG B quantitative calculation of long-term 75 min market profile indicators, coupled with technical indicators as input variables. The results are shown in Table 2.

Experimental results show that: quantitative and qualitative computing market profile, its predictive ability is the same 30 min maximum. EG B is more than 83.38%. For five minutes after the change range of prediction accuracy rate of only 64.16%.

3.1.3. EG C – the qualitative the market profile indicators (long-term + short-term) + technical indicator

The EG C and EG D, in addition to the long-term 75-min market profile indicators and technical indicators, adding short-term 15-min market profile indicators as input variables. The EG C is calculated in a qualitative manner the market profile of the experimental results as shown in Table 3. EG D is based on a qualitative manner computing market profile, the experimental results as shown in Table 4.

EG C also change range after 30 min in order to predict accurately. Followed by 15 min, and no matter what the predicted time, are up to 70% accuracy rate.

3.1.4. EG D – the quantitative the market profile indicators (long-term + short-term) + technical indicator

EG D also 30 min after the change range forecast to predict the best accuracy rate of 81.84%. However, 5 min to 70% predictive ability without.

3.1.5. Control group – only technical indicator

The control group only uses technical indicators as input variables of the model, and the results are as follows in Table 5:

Experimental results show that the predictive ability of the 30-min change higher. However, its accuracy rates of 72.82%, a whole different prediction time difference is not high.

3.2. Performance evaluation and comparison

This study investigates the market profile of trading strength factors, whether a trading strategy based on technical indicators is better than simply using better investment performance. It is only in the control group, and the technical indicators as input variables, for the other four groups compare the model of the experimental group. The experimental group and the control group their experimental results are summarized in the Figs. 4 and 5.

Accuracy control results: In addition to the EG B at a predicted time of 5 min, the accuracy rate is lower than outside the control group, the control group, the prediction effect is generally worse than the experimental group, wherein the prediction time of 5 min and 15 min. The highest accuracy rate of the EG C; when the predicted time is 30 min, the highest accuracy rate of the EG B.

Profitability control results show: The predicted time of 5 min, the average profit points of the EG B and EG D are slightly lower than the control group, and the other experimental groups were higher than control group, the EG C is the highest. The predicted time of 15 min, only the average profit points of the EG B is lower than the control group, and the other are higher than the control group, the EG A is the highest. The predicted time of 30 min, the profitability of all experimental groups was significantly higher than the control group.

3.3. Statistical test

This study investigated the three directions, as follows:

(1) To investigate whether there are differences between quantitative and qualitative market profile calculation method.
(2) To explore the added short-term market profile variables and their performance is better than only with the long-term market profile.
(3) To prove the market profile can help predict the future stock price trends.

Therefore, if we use a statistical test further examination and comparison between the experimental group, the accuracy significant differences between the experimental and control groups, in order to enhance the credibility and reliability of the
experimental results can be summarized in a better market profile indicators method for future researchers, and finally verify the availability of market contours.

Each model of this study were repeated the experiment 30 times, according to the central limit theorem, the sample mean of the probability distribution similar to the normal distribution. However, because the population variance is unknown and the population’s normality cannot be confirmed, the population mean number differences between hypothesis testing, this study carried out by the $T$-test.

$T$ test can be divided into two population variances are equal or not equal, and use a different formula. Before conducting independent samples $T$-test using the $F$ distribution the conduct $\sigma_1^2, \sigma_2^2$ hypothesis testing to determine whether the two population variances are equal, thus decided to adopt what the T measurement formula (0.05 level of significance), if population variances are not equal, the individual variance statistic (Cochran & Cox), equal with integrated $T$-test, ANOVA (pooled-variance test).

3.3.1. Population variance

$F$-test to test the model $X$ prediction accuracy rate variance and model predicted $Y$ accurate rate variance equality between the results are shown in Table 6.

$F$ test results that: at the predicted time of 5 min, the EG B and the control group of the population variance is not the same ($0.02974 < 0.05$ reject $H_0$); prediction time for 15 min, the EG A and control group of population variance is not the same as ($0.034237 < 0.05$, reject $H_0$) EG C and the control group of the population variance is not the same ($0.001324 < 0.05$, reject $H_0$), so these three groups are made of individual variance $T$-test.

3.3.2. Predictive ability test

From the $T$-test results (Table 7): the predicted time of five minutes when the accuracy of the EG C was significantly better than the control group ($0.028337 < 0.05$ reject $H_0$); prediction time of 15 min while the EG A (0.030294 < 0.05, reject $H_0$) and EG C (0.039604 < 0.05, to reject $H_0$) accuracy significantly better than the control group. The prediction time is 30 min, the EG B (.001734 < 0.05, reject $H_0$), the EG C (0.035169 < 0.05, reject $H_0$), the EG D (.014197 < 0.05, reject $H_0$) the accuracy of the rates are
significantly better than the control group. Confirmed by the statistical test results the added market profile indicators does effectively enhance the accuracy of prediction, in particular, the predicted time for 30 min when the most significant.

4. Conclusion

This study used market profile theory to determine the price, value changes and trends through indicator values establishment as a basis for the relationship between the technical specifications for the aggregate evaluation of the physical forces. Different market profiles were compared and calculated with the simple use of the differences between the technical analysis models.

The experimental results are at the predicted time of 5 and 15 min. Qualitative computing market profile deviation value, the accuracy rate is better than quantitative way. At the predicted time of 30 min, calculated using quantitative method superior qualitative. The main cause of the qualitative methods can learn effectively break through and below the status of market pressure and support, but due to the qualitative calculation, take the first-order and second variable, cannot reflect the true market behavior and therefore judged that the long-term trend decreased ability. The quantitative method due to short-term correction in the case of pressure support caused by short-term forecasts cannot effectively enhance interpretation of ability but the trend is better than qualitative methods, so the long-term prediction is significantly improved.

Whether using qualitative or quantitative calculation, by adding a short-term 15-min market profile indicators help to improve the forecasting performance, especially with five minutes to predict the time of the most significant enhancement effects. Due to short-term 15-min market profile indicators for predicting stock prices of longer lines, judged weak. Therefore, when the forecast longer hours, the relatively small effect, and even negative growth.

Comparing the technical indicator with the neural network input variables, the experimental results confirm that adding the market profile indicators can effectively improve the forecast accuracy and profitability. To predict the change range after 30 min to enhance the most significant effect. From experiments designed to observe the market profile information, and the future trends of the Taiwan stock market, to further assess and validate the predictive ability of the different intervals, to provide an innovative investment tools for investors and future researchers.

Experimental results show qualitative market profile indicator performs quantitative approach in short-term forecast period. In contrast, quantitative market profile indicator has a better trend-predicting ability thus it is more effective in long-term forecast period. The results also manifest that both approaches considering the combination of long-term and short-term change in market enhance forecasting performance and are most effective in short-term time interval. In conclusion, the integration of market profile and technical analysis surpasses technical analysis as a parameter to neural network architecture by effectively improving forecasting performance and profitability.

References