The Predictive Power of Volatility Pattern Recognition in Stock Market

Yue Li
Department of Financial Engineering
Stevens Institute of Technology
Hoboken, NJ, USA
Yue.Li@stevens.edu

Khaldoun M. Khashanah
Department of Financial Engineering
Stevens Institute of Technology
Hoboken, NJ, USA
Khaldoun.Khashanah@stevens.edu

***Abstract*—In this paper, we first show that there exists a daily pattern in volatility of equities and that the volatility pattern is different from the daily volume profile. To further emphasize the most important volatility change in a day, we decompose the matrix by using principal component analysis (PCA) and use eigenvectors as weights in our distance metrics so that it can help in the clustering step to generate and forecast volatility patterns.**

Clustering methods K-Means and Expectation Maximization (EM) for Gaussian mixture model with three different distance metrics are implemented, which enable us to optimize the clustering result. When new observations come in as a stream we compare the similarity with clustered volatility patterns under the distance measure to generate a predicted signal.

To examine the practicality of this pattern recognition in equity market volatility, we build a trading algorithm and design a back test to check the accuracy and profitability. The realized volatility calculated by SPY and a representative for implied volatility VXX are treated separately and compared throughout the paper. As a result, the test error, profit and loss and risk adjusted return are compared with the performance by using fixed volatility profile, as well as compared with GARCH(1,1) model for SPY realized volatility and ARMA(1,1) for VXX as implied volatility. The sharp ratio from back test of the trading algorithm outperforms both benchmarks.

I. INTRODUCTION

In a financial market, predication is always an attractive and challenging question to all participants. In this paper, we focus on volatility pattern recognition since it is becoming a commonly recognized fact that the importance of volatility is steadily rising with the fast development of the financial market. The application of volatility forecasting in the financial industry includes, but not limited to, risk management, portfolio optimization, derivatives pricing, and best execution algorithms. Newly invented instruments are making direct investment to volatility easier and more practical to participants. Exchange Traded Funds or ETFs such as VXX has increased exponentially in trading volume with estimated average daily turnover around 1 billion USD since it debuted in 2009.

Financial markets are generally not perfectly efficient markets, which means that traders can make a profit by analyzing historical data and past patterns. In other words, volatility in the stock market is not a Markov process, that is conditional only on the present state of the system, i.e., the market has some memory. Furthermore, the market is not fully pricing all market information at all time [1]. Park and Irwin [2] in their paper have shown that technical analysis is profitable though their research is only in foreign exchange and commodity futures market. Market efficiency does not occur by itself, which also depends on the time horizon [3] [4] [5]. We view the Adaptive Market Hypothesis (AMH) of Andrew Lo [32] as a more realistic approach to this concept.

The short term horizon prediction is a very popular topic. With the boost of Electronic Communication Network (ECN) and fast growing electrical trading, the high frequency intraday trading is quickly growing. Being an alternative of the macroeconomic fundamental analysis, the traders use technical analysis much more frequently and place more weight than fundamental analysis in short horizons (less than a week) in the surveys of technicians, according to the research conducted by Allen and Taylor [6]. Meese and Rogoff [7] show that macroeconomic models have failed to predict in the short term horizon, which indicates the financial market is more news or information-driven in the short term and affected by macroeconomics in the long term. Therefore, the fundamental analysis is not a good choice as predicting intraday volatility. As an alternative, pattern recognition, especially intraday pattern, could be used as a volatility trading guide.

In this paper, we focus on intraday volatility pattern because we are targeting high frequency environments; and the intraday volatility pattern can be integrated with most of the intraday trading strategies such as the Volume Weighted Average Price (VWAP) trading algorithm. However, the pattern could be significantly different for different time horizons.

The nature of volatility is dramatically different from the price return time series, known as stylized facts such as thick tails, volatility clustering and leverage effect [5] [6]. Effort has been focused on modeling these features, such as the famous ARCH and GARCH models [7] [8], as well

978-1-4799-7560-0/15 $31.00 © 2015 IEEE
DOI 10.1109/SSCI.2015.112

as stochastic volatility (SV) models [11] [13] such as Heston model and SABR model, etc. Besides predictions done by only using time series methods, the methods of time series combined with neural network/genetic programming are also adopted in the papers [17] [18] [19] [20], as well as using machine learning methods to combine multiple technical indicators in the papers [21] [22]. Florea compared the predictability in FX market, equity market and VIX by combing multiple indicators [24]. Most of the papers show certain kind of excess return and indicate the predictive power from historical data reflecting market memory or market inefficiency.

In this paper, firstly, we are going to show that volatility can be decomposed into a day pattern by using principal component analysis (PCA). PCA analyzes covariance matrix to pick most important factors. Here we can utilize decomposed PCA's eigenvectors as weights to emphasize the most volatility change in the distance metric, and the weight will be used in clustering patterns in next step. Secondly, we will generate distinguished patterns by using clustering method and finally match the new observation to a specific pattern as a prediction. The K-Means and Expectation-Maximization (EM) for Gaussian mixture clustering methods will be adopted.

Related research has been done by comparing the PCA with independent component analysis (ICA) in [5] [16], which will show the hidden factors of seasonality, holiday effect and the long term industrial growth trend. Also in [25], a predefined pattern is tested in the foreign exchange market shows excess return. In contrast to the well-known predefined patterns, we are going to generate patterns automatically by unsupervised clustering. But in a future work, we can also combine generated patterns with these existing and well adopted patterns to improve the result.

Empirical study is conducted by using stocks of SPY and VXX throughout the paper since SPY can represent broad stock market and VXX is the most active volatility ETF. As a result, we use GARCH(1,1) as benchmark as well as fixed volatility pattern as benchmark to confirm the improvement and predictive power of volatility pattern recognition.

II. Methodology

The data of SPY price is transferred into log return $r_t = \log(\frac{P_{t+\Delta t}}{P_t})$, and then realized volatility (RV) defined as $\sigma_t = \sqrt{\sum_{i=1}^{n} r_i^2}$ is calculated, which assumes zero mean. Practitioners use realized volatility in high frequency field by aggregating higher frequency returns to compute lower frequency volatility. Here, it is also known that the high frequency data suffer greatly from microstructure noise, which makes realized volatility statistically biased [28] [29]. The data of VXX price is also transferred into log return, then its cumulative return is used to represent intraday volatility pattern. In contrast to the realized volatility, implied volatility is not a statistical calculation of the direct observations but is derived from options by using the realized option market price. VXX, as a VIX index ETF, is not a perfect representative of VIX since it is constructed by using two near term VIX futures. But it is more practical to study VXX as it is a direct investment instrument while VIX index is not tradable.

To illustrate the volatility day pattern we fold the volatility time series into an *n*-by-*p* dimensional matrix, where *n* is the number of days and *p* is number of minutes during the trading hours. The prepared volatility matrix will be expressed as matrix below. After rearranging volatility to this *n* by *p* matrix, now each minute represents a dimension which reflects the feature of volatility within its own time interval across different days. For VXX, we use cumulative return $R_{d,t} = \Sigma_{i=0}^{t} r_{d,i}$ instead of realized volatility $\sigma_{d,t}$.

Volatility day pattern matrix: $\Sigma =$

$$= \begin{array}{c} \rightarrow Time\ within\ a\ day \\ \begin{bmatrix} \sigma_{d1,t-\tau} & \dots & \sigma_{d1,t-3} & \sigma_{d1,t-2} & \sigma_{d1,t-1} & \sigma_{d1,t} & day\ 1 \\ \sigma_{d2,t-\tau} & \dots & \sigma_{d2,t-3} & \sigma_{d2,t-2} & \sigma_{d2,t-1} & \sigma_{d2,t} & day\ 2 \\ \sigma_{d3,t-\tau} & \dots & \sigma_{d3,t-3} & \sigma_{d3,t-2} & \sigma_{d3,t-1} & \sigma_{d3,t} & day\ 3 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \sigma_{dn,t-\tau} & \dots & \sigma_{dn,t-3} & \sigma_{dn,t-2} & \sigma_{dn,t-1} & \sigma_{dn,t} & day\ n \end{bmatrix} \end{array}$$

A. Fixed pattern without clustering

Before performing the PCA decomposition, we would like to first illustrate the difference among SPY volume profile (left), realized volatility pattern from SPY (middle) and implied volatility represented by VXX (right) as shown in figure 1. In these charts, blue lines are the average of the curve from each dimension; while red lines are moving averages across time to smooth the line. The volume intraday pattern is similar to SPY realized volatility pattern. But volume profile is more like a "U" shape curve with volume going extremely high in the market closing period; while at closing time, realized volatility curve only goes up to half of the open high. A similar result is also shown by Almgren [30]. On the contrary, VXX pattern has significantly different shape, which could be due to the VIX future contango, or speculation by traders, or implied volatility suffers more from leverage effects. But this is not the focus of this paper. Therefore, the generated clusters and weights assigned by the distance metric should be treated separately based on this analysis.

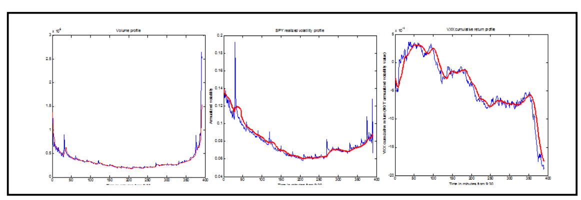

Figure 1. The average daily pattern of volume, volatility by SPY and VXX

Some conclusions by paper [31] suggest that the fixed pattern is good enough for intraday profile, especially for volume pattern curve. There is no significant difference when categorizing stocks by market cap, listed exchange, sector, region, or asset type, except for American Depositary Receipt (ADR)/non-ADR. To provide evidence of improved predictive power with using PCA and clustering methods for extracting the volatility pattern, fixed pattern is considered as a benchmark in the constructed trading strategy.

B. Principal component analysis (PCA)

By applying PCA to the volatility matrix, we decompose the covariance matrix into orthogonal principal components and corresponding loading factors. The PCA is a well-established method especially in face recognition, which can extract feature vectors by projecting value from the old space onto the rotated new space so that fewer principal dimensions in new space can represent most of the variance in the old space. Here we use PCA to help conclude which dimensions from the original space could explain the main variance of volatility change. The result of PCA decomposition is shown in figure 2.

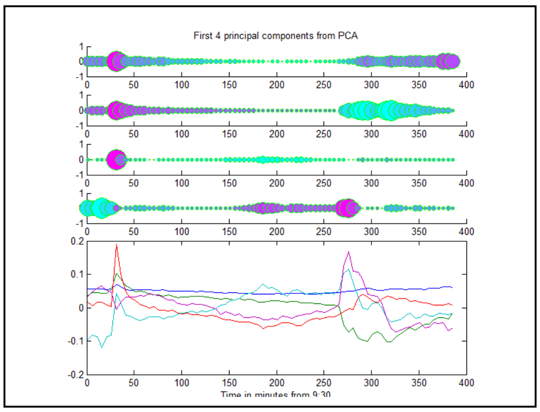

Figure 2. SPY realized volatility decomposition. Time series of first four principal components by PCA

In figure 2, horizontal axis is the time window of trading hour from 09:30AM to 16:00PM. Interval size is per minutes. In chart we can see there is a jump at around X=30 which is 10:00AM.

The first component is shown as in the first line, then the second component and so on. The circle size represents the importance at specified time period while the color represents the distance from 0. The last chart is the combination of the first five principal components which can explain more than 75% of the total variance in volatility change. From the line chart we can see an expansion in the beginning of the day and another expansion after 2pm then decrease a little until market close. The loading factors from PCA can be assigned as weights in the distance metric in the following step.

The result from PCA is consistent with the result from first step analysis, which indicates there is more volatility change during the beginning of trading session as well as the market closing period. But PCA suggests heavy weight starting from 2pm might be caused by special days such as Federal Reserve announcement. This can be seen from the clustering result in next step. However the jump around 10am is still an interesting phenomenon. One possible explanation may be due to some proprietary trading shops start intraday trading algorithm from 10am while leaving 9:30 to 10:00 off for trend analysis.

C. Distance metrics

The volatility from the original space is mapped to the new space by using PCA. An unsupervised clustering to categorize different volatility pattern is then deployed. To calculate, the metric which will be used in clustering is defined. As distance measure, a metric has to satisfy the following conditions:

Distance metric:

$d(x, y) \geq 0$ (non-negativity, or separation axiom)

$d(x, y) = 0$ if and only if $x = y$

$d(x, y) = d(y, x)$ (symmetry)

$d(x, z) \leq d(x, y) + d(y, z)$ (subadditivity)

The candidate distance measures we applied here:

$d = \sqrt{\sum w_i (X_i - Y_i)^2}$ (Weighted Euclidean distance)

$d = 1 - \rho$ (Weighted correlation distance)

$d = \sqrt{2(1-\rho)}$ (Modified weighted correlation distance)

Where in the last two distance metrics, weighted correlation is as:

$$\rho = \frac{\mathrm{cov}(x, y; w)}{\sqrt{\mathrm{cov}(x, x; w)\mathrm{cov}(y, y; w)}}$$

$$\mathrm{cov}(x, y; w) = \frac{\sum_i w_i (x_i - m(x; w))(y_i - m(y; w))}{\sum_i w_i}$$

$$m(x; w) = \frac{\sum_i w_i x_i}{\sum_i w_i}$$

By having the above distance measures, we apply the clustering method to generate most distinguished patterns. The purpose of clustering is to minimize the distance within cluster but maximize of the distance between clusters no matter under which distance measure. We can quickly review the methods adopted as following:

D. Clustering methods

The K-means clustering depends on partitioning observations into k clusters, so that each element belongs to the nearest cluster. The formula can be expressed as:

$$\arg\min_{s} \sum_{i=1}^{k} \sum_{x_j \epsilon S_i} d(x_j, m_i)$$

Where m_i is the i cluster centroid of S_i, and x_j is a data observation.

The algorithm of k-means has two steps. The first step assigns each point to the nearest cluster centroid by comparing the distance; then performs an update to reduce the sum of total distance; and cluster centroids are recalculated.

The iteration process to get K-Means cluster:

$$S_i^{(t)} = \{x_p : d(x_p, m_i^{(t)}) \leq d(x_p, m_j^{(t)}), \forall 1 < j < k\}$$

$$m_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{x_j \in S_i^{(t)}} x_j$$

Where $S_i^{(t)}$ stands for i-th cluster at t-th iteration; $|S_i^{(t)}|$ is the part for normalization to the sum of distance, which could be the count of points belonging to the cluster.

In general, if the iteration process of K-means can be viewed as special case of expectation–maximization (EM) algorithm, alternatively EM for Gaussian mixture model can be applied.

Expectation–maximization is a commonly used mixture decomposition method to decompose a mixture model. Mixture model can be seen as an overall population consisting of finite number of sub-populations, but without sub-population identity information. The EM algorithm is an iterative method for finding parameters. The algorithm has two steps to process, the E-step and M-step.

E step: compute the conditional expectation of the log likelihood ($logL$) at t-th iteration by given X and Z under current estimated parameter $\theta^{(t)}$, where Z is conditional distribution and X is observations.

Estimate $\theta^{(t)}$ given X and Z

$$Q(\theta|\theta^{(t)}) = E_{(Z|X,\theta^{(t)})}[\log L(\theta; X, Z)]$$

M step: update parameter estimation θ to the next iteration by maximize Q

Find θ to maximize Q

$$\theta^{(t+1)} = \arg\max_{\theta} Q(\theta|\theta^{(t)})$$

With two clustered methods and three distance measures, we can choose from the combination which gives the best result. Figure 3 is the demonstration from a training set that five volatility trajectories are clustered into two different groups.

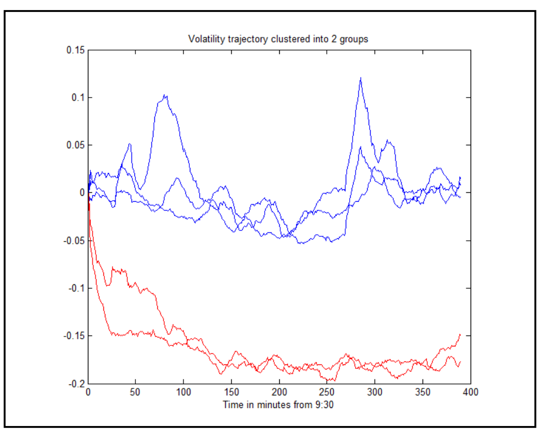

Figure 3. Example of clustered intraday volatility trajectory

Measuring by distance, the trajectories in the same cluster under given distance metric have similar pattern. The blue trajectories belong to one pattern which shows a jump at market open and in the afternoon. The red trajectories belong to another pattern which show a straight down trend at beginning then steadily reach the equilibrium. Since we assign the weight in distance metric from PCA, the clustering method is actually more sensitive to volatility change during the market open period and after 2 pm.

E. Trading strategy by utilizing volatility pattern

To examine the predictive power by using pattern recognition for intraday volatility, we build a trading strategy based on a proposed volatility forecasting method. Here VXX is a direct investable instrument, while realized volatility by SPY is not tradable. The strategy of trading VXX and pseudo "profit and loss (PnL)" by trading SPY's realized volatility is constructed as below:

Here we define:
p_i the i[th] clustered pattern
X new observed volatility data
+/- will be final output signal of long/short
w_i the weight corresponding to trading signal
R_t the return by $\Delta\sigma$ in time interval $[t, t + \Delta t]$

When matching a new observation to one of the existing patterns, we need to calculate posterior as

$$\text{posterior} = prior * likelihood$$

$$P(p_i|X) = \frac{P(p_i) * P(X|p_i)}{P(X)}$$

Trading signal $P(+|X) =$

Posterior * conditional probability given specific pattern

$$= \sum_i P(p_i|X) * P(+|X, p_i)$$

$$= \frac{\sum_i P(p_i) * P(X|p_i) * P(+|X, p_i)}{P(X)}$$

Here we have

$$\text{prior } P(p_i) = \frac{\#\text{ of cases in cluster pattern i}}{\#\text{ of Total Cases}}$$

$$\text{likelyhood } P(X|p_i) = \frac{\text{Distance from pattern i}}{\sum_i \text{Distance from pattern i}}$$

$$\text{conditional Probability given specific pattern}$$

$$P(+|X, p_i) = \frac{\#\text{ of up trend price movement}}{\#\text{ of Total cases in pattern i}}$$

$$\text{final trading signal will be} \begin{cases} +1 \text{ if } P(+|X) > P(-|X) \\ 0 \text{ if } P(+|X) = P(-|X) \\ -1 \text{ if } P(+|X) < P(-|X) \end{cases}$$

And we also construct a parameter as an investment weight in a portfolio, which will reflect our confidence on each trading signal:

$$\text{weight } w = |\, P(+|X) - P(-|X)|$$

In the end, the weight will be normalized so as to be comparable with non-weight adjusted investment

$$\text{weight } w_i = \frac{w_i}{\sum_i w_i} * \#\text{ of total trades}$$

Since our trading signal is generated at every time interval, as long as the investment period is known, the total number of trades will be known in advance. Hence the final cumulative profit and loss will be calculated as

$$\text{PnL} = \sum_t \text{trading signal}(\pm) * \text{return } R_i$$

Or with adjusted weight as

$$\text{PnL} = \sum_t w_i * \text{trading signal}(\pm) * \text{return } R_i$$

III. EXPERIMENT

A. Data

US stock of SPY and VXX are used throughout the paper since SPY can represent broad stock market in US and VXX is the most active volatility ETF. For SPY, realized volatility is calculated from tick data; For VXX, although we could use VIX to better represent implied volatility from market, but since VIX is not tradable and it only updates every 15 seconds we still decide to use VXX for intraday pattern.

Other information from stock market like traded volume and bid/ask spread are also available and could contribute to improve volatility forecasting; but those data are not used in this experiment. The integration with other data as extra dimensions will be conducted in future work. Here we will only use price/volatility information to do the pattern recognition and clustering.

To ignore the overnight effect and only focus on intraday volatility pattern, the first volatility points on each day is aligned to 0. The most recent four and half years data from 2010 are split into a training set and test set as 70/30.

B.Reuslt of clustered pattern

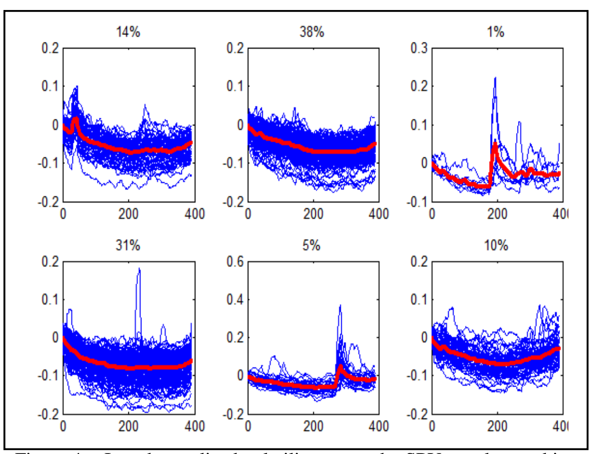

Figure 4. Intraday realized volatility pattern by SPY are clustered into 6 distinguish patterns

Figure 4 shows the clustered realized volatility patterns by SPY. The volatility jump around 2pm is clustered as a separated cluster; and the total number of trajectories in this cluster is about 5%. This can be explained as an event that happens once per month and is consistent with the Federal Reserve meeting announcement. One of the pros of proposed method is the capability of automatically filtering out special days. There is a 1% of population cluster whose volatility jump occurs at noon, but the shape of the volatility profile is significantly different from other patterns. Since we assign weight to emphasize the opening period, the first cluster is split out due to the small bump at beginning. The other clusters look similar by shape but are different on scale and how much it recovered back during market close time.

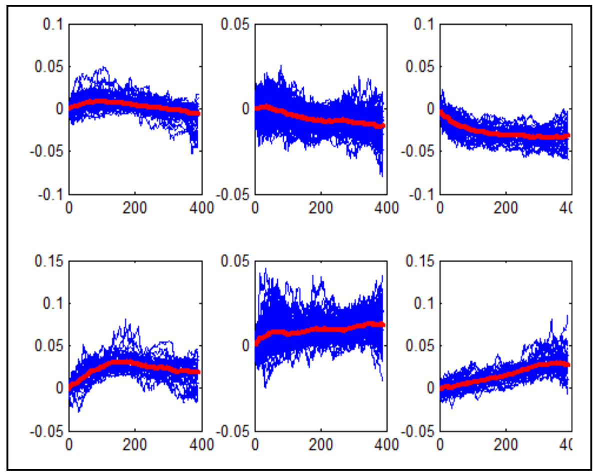

Figure 5. Intraday volatility pattern by VXX cumulative return are clustered into 6 distinguish patterns

By looking at the result chart above in figure 5, it shows dramatically different intraday volatility pattern represented by VXX cumulative return. The two clusters in the last column show that the volatility straightly increases or decreases; though it seems with a limitation when decrease. Compare to that, the two clusters in the middle column is very steady since scale is much smaller. The reason why cluster 1 (first row, first column) is separated from cluster 3 (first row, third column) is because we enforce more weight on market opening period. In the end, the forth cluster (second row, first column) shows volatility goes up with resistance and could turn direction from morning session to afternoon session.

The number of clusters to be categorized in the unsupervised clustering is given or known. If the training data are split into more clusters, the outliers in current pattern could be separated into a new pattern.

C. Result of back test statistics

To examine the power of predictability and profitability, we have built an algorithmic trading system based on the method described above. To avoid noise from market microstructure when calculating realized volatility from SPY, we extend the trading period longer to 15 minutes.

We use fixed intraday volatility pattern as benchmark as well as GARCH(1,1) as benchmark (ARMA(1,1) for VXX cumulative return as benchmark). A comparison to show the difference between three methods and result is listed in Table I.

TABLE I. RESULT OF COMPARISON WITH BENCHMARKS

Asset	Methods	Test Err	Total PnL	Annualized PnL	Stdev	Sharp ratio
SPY	Clustered Pattern	40.78%	28.79	19.19	6.32	
	Fixed Pattern	47.21%	18.31	12.20	6.39	
	GARCH(1,1)	49.24%	3.58	2.39	6.34	
VXX	Clustered Pattern	44.49%	109%	73%	0.44	1.64
	Fixed Pattern	48.88%	62%	42%	0.44	0.93
	ARMA(1,1)	48.77%	24%	16%	0.44	0.36

The simple statistics of result shows that the test result may not be perfect given that the test error is relatively high (comparing with other machine learning methods). However, it is still notable that we gain profits on all intraday volatility trading within test data set. When calculating profit and loss for SPY, it does not make too much sense except for comparison with benchmarks because the realized volatility calculated from SPY is not tradable. The profit we gain is in the unit of the annualized realized volatility; therefore it cannot calculate return in percent, and the sharp ratio does not exist.

From the table we can also see that our proposed method by assigning weights from PCA to cluster intraday volatility into base patterns beats the fixed pattern method and also beats GARCH(1,1) for intraday volatility forecasting from the aspects of test error, total return and sharp ratio.

.

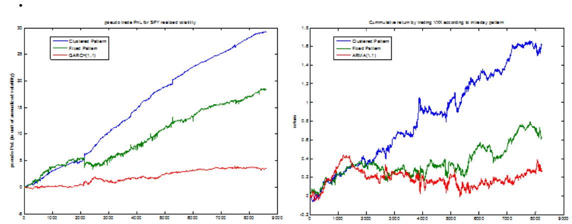

Figure 6. Cumulative PnL in test set (about one and half year)

To better understand the performance of the designed volatility trading strategy, a cumulative return is plotted as shown in figure 6. In the left chart, the pseudo PnL by clustered pattern, fixed pattern and GARCH(1,1) are showing amazing performance. But since it is not directly tradable, one explanation for the perfect performance may due to the market inefficiency. In the right chart is the performance of tradable VXX. Similarly, it shows PnL in the order of clustered pattern, fixed pattern and ARMA(1,1), where ARMA(1,1) is merely better than random.

To further compare the adjusted weight parameter, which reflects our confidence on each trading signal, we can use weight as a position size at each decision-making step.

D. Flow chart of automated trading system

We have built the system as the flow chart shows in figure 7. During the intraday trading experiment, we set trading signals to be generated every 15 minutes.

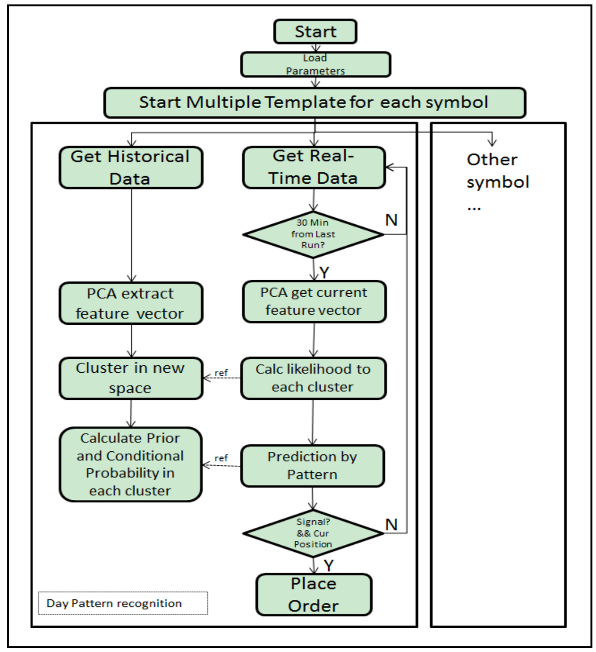

Figure 7. Flow chart of automated trading system

IV. CONCLUSION

In summary, the clustered patterns can help in finding trading signals and work as a hidden layer to calibrate the daily volatility profile. Results from the back test show a good profit and loss, which consistently beat the fixed pattern method and time series only methods. The cumulative return is also very stable which can be seen from the calculated sharp ratio.

The clustered patterns avoid the subjectivity in pre-defined patterns but still have some known issues. For example, it overemphasizes capturing certain patterns like straight trend up/down because the higher prior probabilities, which is not quite consistent with traders' human behavior in some circumstances. The parameter calibration issue is also notable in how many clusters to discriminate and the time interval and pattern time horizon. In this paper, we have shown a group of parameters that works well. In the future, a machine learning method with online adaptive learning can be considered to integrate with this pattern detection system.

V. Future Work

The additional adjusted weight which further reflects our confidence on the trading signal could improve the strategy and will be implemented in the future work. Also integration with spread data and trading volume data as extra dimension could potentially help in volatility forecasting.

ICA is targeting to remove the correlation of principal bases and is well applied in separating signals, which could be applied to improve PCA.

Pre-defined well known patterns could be combined with clustered patterns together and be used in pattern recognition procedure so that it is more compatible with idea of behavior finance. The pre-defined pattern could be explained as some herd effect from interaction between traders.

Reference

[1] Fama, E. F., (1991). Efficient capital markets. The journal of finance, 16(5).

[2] Irwin, Scott H. and Park, Cheol-Ho. (2007). "What Do We Know About the Profitability of Technical Analysis?" Journal of Economic Surveys, Vol. 21, No. 4, pp. 786-826. Available at SSRN. DOI: 10.1111/j.1467-6419.2007.00519.x.

[3] Akbar, M., Baig, H., (2010). Reaction of Stock Prices to Dividend Announcements and Market Efficiency in Pakistan. The Lahore Journal of Economics, 15(1), p. 103-25

[4] Bessembinder, H. and K. Chan, 1995, "The profitability of technical trading rules in the Asian stock markets", Pacific-Basin Finance Journal, 3, 257-284.

[5] Taylor, M.P. and H. Allen (1992), The Use of Technical Analysis in the Foreign Exchange Market. Journal of International Money and Finance, 11, 304-314.

[6] Allen, H., and M.P. Taylor (1990), Charts, Noise and Fundamentals in the Foreign Exchange Market, Economic Journal, 100, pp. 49-59.

[7] Meese and Rogoff, (1983) “Empirical exchange rate models of the seventies. Do they fit out of sample?” Journal of International Economics, Volume 14, Issue 1-2, February 1983, Pages 3-24

[8] Cont R., 2001, “Empirical properties of asset returns: stylized facts and statistical issues”, in: Quantitative Finance, Vol 1, No 2, 223-236.

[9] Mandelbrot B. B., "The variation of certain speculative prices", Journal of Business, XXXVI (1963), pp. 392–417.

[10] Engle, Robert F. (1982). "Autoregressive Conditional Heteroscedasticity with Estimates of Variance of United Kingdom Inflation". Econometrica 50 (4): 987–1008

[11] Bollerslev, Tim (1986). "Generalized Autoregressive Conditional Heteroskedasticity". Journal of Econometrics 31 (3): 307–327

[12] Enders, W. (2004). "Modelling Volatility". Applied Econometrics Time Series (Second ed.). John-Wiley & Sons. pp. 108–155

[13] Taylor S., "Modeling Financial Time Series", Wiley and Sons, 1986

[14] Taylor S.J., “Modelling stochastic volatility: a review and comparativestudy”, Mathematical Finance, 1994, 4, 183-204

[15] Hull J., White A., “The pricing of options on assets with stochastic volatilities”, Journal of Finance 42, 281-300.

[16] Andrew D. Back, Andreas S. Weigend, “A First Application of Independent Component Analysis to Extracting Structure from Stock Returns”

[17] Utans, J., Holt, W. T. and Refenes, A. N. (1997). Principal component analysis for modeling multi-currency portfolios, in A. S. Weigend, Y. S. Abu-Mostafa and A.-P. N. Refenes (eds), Decision Technologies for Financial Engineering (Proceedings of the Fourth International Conference on Neural Networks in the Capital Markets, NNCM-96), World Scientific, Singapore, pp. 359-368.

[18] M. I. Jordan and T. Petsche (eds), Advances in Neural Information Processing Systems 9 (NIPS*96),MIT Press, Cambridge, MA, pp. 995-1001.

[19] Bergerson, Karl, Wunsch Ii, Donald C. (1991) "A commodity trading model based on a neural network-expert system hybrid"

[20] Saad, E. W., D. V. Prokhorov, et al. (1998). "Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks." IEEE Transactions on Neural Networks 9(6): 1456-1470.

[21] Downs T. Hryshko, A. System for foreign exchange trading using geneticalgorithms and reinforcement learning. International Journal of Systems Science, 35(13-14):763–774, 2004.

[22] Germán Creamer,Yoav Freund (2006),"Automated Trading with Boosting and Expert Weighting",Quantitative Finance, Vol. 4, No. 10, pp. 401–420

[23] Pictet O.V.b Zumbach G.c Bhattacharyya, S.a. Knowledge-intensive genetic discovery in foreign exchange markets. IEEE Transactions on Evolutionary Computation, 6(2):169–181, 2002. cited By (since 1996) 23.

[24] Florea A. Bucur, L. Techniques for prediction in chaos - a comparative study on financial data. UPB Scientific Bulletin, Series C: Electrical Engineering, 73(3):17–32, 2011.

[25] Gene Savin, Paul Weller, Janis Zvingelis, The predictive Power of "Head-and-Shoulders" Price Patterns inthe U.S Stock Market

[26] Moody, J. E. and Wu, L. (199713). What is the "true price"? - State space models for high frequency Forex data, Proceedings of the IEEE/IAFE 1997 Conference on Computational Intelligence for Financial Engineering (CIFEr), IEEE Service Center, Piscataway, N J, pp. 150-156.

[27] Christopher J. Neely Paul A. Weller, 2011, Technical Analysis in the Foreign Exchange Market, July 24

[28] Lunde A., Hansen P.R., “Realized Variance and IID Market Microstructure Noise", Econometric Society, 2004 North American Summer Meetings 526

[29] Bandi F.M., Russell J.R., “Separating microstructure noise from volatility”, Journal of Financial Economics, 79 (2006), pp. 655-692

[30] Almgren R., Thum C., Hauptmann E. and Li H., "Direct estimation of equity market impact" in Risk, July 2005.

[31] Pragma Securities, 2012, "Constructing VWAP Curves", working paper

[32] Andrew W. Lo, 2004, "The Adaptive Markets Hypothesis", The Journal of Portfolio Management, 30th Anniversary, Vol. 30, No. 5: pp. 15-29, DOI: 10.3905/jpm.2004.442611