CHAPITRE 2

RELEVANT LITERATURE SURVEY

As indicated in the previous chapter, the research efforts on the behavior of rate of return of equities or indices, specifically their volatility is recently dominated by the parametric statistical approaches, such as a variety of GARCH models. Contributions of Andersen *et al.* (1998, 2001, 2001b, 2003) gave a jump start to the theoretical and experimental research on the estimation and forecasting of the realized/integrated volatility and produced some decent results. However, there is need for further research to be done in order to improve the forecasting power of their methodology. For example, the one-day ahead forecasting of IV in a 25-day period generally lags the actual data by roughly two days, even though their results are considered to be already far superior to the GARCH results. To improve the forecasting accuracy, some researchers started to venture into different non-traditional exotic techniques however non-systematic the attempt might be at the current time. As indicated by Arino (1996) and Li (2003), the wavelet methodology, though it has not been used extensively, began to play a more important role in economic and/or financial time series studies because of its many unique features. A few researchers do make use of wavelets to analyze individual stock returns, high frequency stock index returns (Arino, 1996) and the foreign exchange rates (Kaboudan, 2005). Among the publications that apply wavelet analysis in either economics or finance applications (Li, 2003), there are still very few publications that deal with integrated/realized volatility regardless of the types of asset (Hog, 2003, Wang *et al.* 2005). As background information, Lee (1998) reviewed many of the applications of wavelets to provide general estimations of the output.

By applying GA, Fong and Szeto (2001) trained a group of rules based on predetermined format in order to extract patterns from a time series data set that is artificially generated with a short memory.

Those trained rules are first tested on one data set and then used to predict another similarly generated data set. With 100 simple 'IF/THEN' rules on a 4-lag recursive memory, the forecasting accuracy is shown to approach 50% ~ 60%. One could expect better result if more elaborate GA strategy is employed. Since Szeto's method was designed to forecast the rate of return of securities, it inherently conflicts with the well documented fact that rates of return usually do not exhibit straightforward autocorrelation (LeBaron, 1992). On the other hand, the corresponding equity, commodity and foreign exchange markets often exhibit volatility clustering, *i.e.* a strong autocorrelation in squared returns thus, the volatility variable. As a result, Szeto's algorithm could be considered as more useful in IV forecasting, and is used here as a basis for the deployment of a pattern recognition tool for the current research problem to find methods to forecast IV in different time horizons. Both wavelet transform and GA are just two of the critical components in the proposed integral approach presented in this thesis. In this chapter, we will concentrate on reviewing only those relevant published papers that attempt to deal firstly with volatility forecasting by applying EA's, and secondly with other type of methods which mainly deal with the non-linearity issues in volatility. Their strengths as well as shortcomings are indicated at the end of each subsection along with a commentary on the concepts that is brought in to play in this investigation. The title of each paper will be used as the heading of each subsection. In the subsequent chapters, other related literature will be further reviewed prior to the detailed discussion in introducing the IV forecasting methodology for the current investigation.

2.1 Volatility Forecasting with Genetic Programming – Key Publications

2.1.1 "Volatility forecasting using genetic programming" (Pictet et al., 2001)

This study uses GP to discover new types of volatility forecasting models for financial time series. The authors improve on the standard GP approach by introducing types in the GP trees, and by optimizing the program constants with a gradient search. These two modifications improve significantly the convergence properties of the algorithm. Moreover, the typing (type-casting) is used to impose a well-defined parity of the solutions so that only meaningful volatility models are built from the price time series. The volatility models are searched with data sampled at hourly frequency, and the optimization criterion is based on the in-sample forecasting quality of the average daily volatility. The heterogeneity of the financial markets is introduced into the models by price change information measured at different frequencies. Finally, the results are compared to standard models like GARCH(1,1). This research was further extended in the paper shown in the next subsection and thus shares the similar shortcomings.

2.1.2 "Genetic programming with syntactic restrictions applied to financial volatility forecasting" (Zumbach et al., 2001)

The use of data at higher frequency opens new avenues in volatility forecasting as the statistical uncertainty is decreased and intra-day effects must be taken into account. Undoubtedly, the use of high frequency data permits better short-term volatility estimations, but also implies more complexity in data treatment and volatility modeling. The main problems can be summarized as a direct question: what are the most important stylized properties that must be taken into account and brought in to play in any new methodology in order to obtain a good volatility forecast? In this referenced study, the authors use GP to discover new types of volatility forecasting models for foreign exchange (FX) rates at hourly frequency.

In the popular application to foreign exchange financial time series, there exists an important exact symmetry induced by the exchange of the two currencies, and this symmetry must be respected by any of the solutions proposed. For this purpose, the research investigators have used a strongly typed GP approach, where the typing system keeps track of the parity of the GP trees. In this way, they have been able to reduce their search space from all possible GP trees down to the subspace of trees that have the proper symmetry. However, it is critical to note that no such properties exist in equity index, so for our current research problem, we would need to look for other properties to be incorporated in the GP process, similar to ones used in the foreign exchange.

From a general viewpoint, volatility forecast is a function-fitting problem, where the realized value for the volatility is the function to be discovered using a causal information set. The time series of the realized volatility is dominated by randomness, and the actual amount of information contained in the information set about the future evolution is rather low. This is measured for example by the lagged correlation for the volatility, which is in the order of 3% to 15%, depending on the actual definition of volatility that is accepted. In short, this essentially means that the volatility forecast becomes a very difficult challenge for GP, and the algorithms need to be very efficient. In order to develop and test special tools, the authors have applied GP to a similar problem, namely the discovery of the transcendental function cosine, using polynomials. As for the volatility, the cosine function obeys a symmetry of parity since $\cos(-x) = \cos(x)$. The study with the cosine makes clear that the conventional GP cannot tackle the volatility forecast problem, and the main difficulty lies in the discovery of good workable values for the constants included in the GP trees. In order to speed up convergence to the optimal values of the constants, they have to use a local search algorithm, like a conjugate gradient. Only when using a mixed algorithm, are they able to obtain good solutions for the cosine problem, and to find volatility forecast that can compete with the standard GARCH(1,1) model. In this work the gradient of the cost function needs to be evaluated numerically, and they recommend use of the Broyden-Fletcher-Goldfarb-Shanno algorithm (BFGS) – a one dimension search quasi Newton method, for the local search (Press *et al.,* 1986).

Overall, the remarkable aspect of the result is that the GP solutions are consistently better than the benchmarks, including out-of-sample forecasting. This clearly shows that

GP, with the syntactic restrictions and the local optimization of the constants, can be considered as an efficient tool for discovering new forecasting models, without over fitting the data sets. Thus, we could in the future use such local optimization algorithms to improve the efficiency in predicting IV of equity index. Furthermore, the following facts can be considered for deployment when selecting the initial functions. The exponential moving averages (EMA) function is always used at least once in each tree, and most of the time an EMA operator is the root node. Clearly, a good forecast needs to have enough memory of the past historical behavior, and this is achieved through the use of EMAs. The EMAs have mostly a constant range z, and not a variable range given by the value of a subtree. All the good solutions do have to contain product of returns at different time horizons.

The key shortcoming of this recent research paper is that it did not explicitly account for the ever-present non-linearity that is inevitably encountered in every problem dealing with the analysis leading to the IV estimation and forecasting, and hence hampers the accuracy of the forecasting. Yet non-linearity is known to be inherent in every volatility time sequence, particularly when jumps, discontinuity or other structural changes in data flow arise. Moreover, the referenced paper's methodology focuses mainly on the characteristics of foreign exchanges, while our research focus lies mainly on handling the equity index.

2.1.3 "Using genetic programming to model volatility in financial time series: the cases of Nikkei 225 and S&P 500" (Chen and Yeh, 1997)

This paper proposes a time-variant and non-parametric approach to estimate volatility. This approach is based on the so-called recursive genetic programming (RGP). In simple terms, RGP is a method that applies Koza's basic GP concepts (Koza, 1998) to sliding windows of a time series and generates an improved sequence based on the evaluation of the average fitness. Such an approach can estimate volatility by simultaneously detecting and adapting to structural changes – which can, in a way, deal with non-linearity. Non-linearity is a key concern in estimating and predicting volatility, because high volatility is usually associated with discontinuity, abrupt breaks and other disruptive structural changes. It may be difficult to model the non-linearity, but they may possess certain distinct patterns that can be recognized. The main contribution of this paper is to employ a model-less adaptive approach to detect structural changes. Therefore, when the underlying structure experiences a certain change, RGP can detect it and, in the mean time, generates a population of volatility estimates under the new structure. As a result, one can avoid firstly, reliance on out-of-date knowledge and secondly the problem of overestimation of volatility.

The key shortcomings of this contribution are that the method given does not employ the most up-to-date estimation techniques such as those methods that are now based on the more accurate estimation of current volatility – IV. Please refer to Section 5.1 for more details. Another important shortcoming is that this technique does not forecast the future outcomes, which is important for equity index applications.

2.1.4 "Predicting exchange rate volatility: genetic programming vs. GARCH and RiskMetrics" (Neely and Weller, 2001)

It has been well-established that the volatility of asset prices displays considerable persistence. That is, large movements in prices tend to be followed by more large moves, producing positive serial correlation in squared returns. Because of this characteristic, the current and past volatility can then be used to predict future volatility. This fact is important to both financial market practitioners and regulators.

This article investigates the performance of a GP applied to the problem of forecasting volatility in the foreign exchange market. GP application here is a computer search and problem-solving methodology that can be adapted for use in non-parametric estimation.

It has been shown to detect patterns in the conditional mean of foreign exchange and equity returns that are not accounted for by standard statistical models (Neely and Weller 2001). This suggests that a GP may also be considered as a powerful tool for generating predictions of asset price volatility for developing our solutions in this thesis. The authors compare the performance of a GP in forecasting daily exchange rate volatility for the dollar-deutschemark and dollar-yen exchange rates with that of a GARCH(1, 1) model and a related RiskMetrics volatility forecast. These models are widely used both by academics and industry practitioners and thus are good benchmarks to which to compare the GP forecasts. While the overall forecast performance of the two methods is broadly similar, on some dimensions the GP produces significantly superior results. This is an encouraging finding, and suggests that more detailed investigation of this methodology applied to volatility forecasting would be warranted.

A core component of the RiskMetrics system is a statistical model — a member of the large ARCH/GARCH family as described in Section 1.2 and 1.3 — that forecasts volatility. Such ARCH/GARCH models are parametric. That is, they make specific assumptions about the functional form of the data generation process and the distribution of error terms. The key problems in this paper are that the functional forms do not explicitly account for non-linearity and structural change of the time series and IV has not been used. This drawback will make on a universal basis, its volatility applications less attractive.

2.1.5 "Extended daily exchange rates forecasts using wavelet temporal resolutions" (Kaboudan, 2005)

In this research investigation, by employing the natural computational intelligence (CI) strategies such as GP and Artificial Neural Networks (ANN), three exchange rate series were fitted and trained for the forecasts of one-step as well as 16-step-ahead exchange rates. Results seem to be generally more successful than the random walk predictions.

As concluded by the author, not all exchange rate series could be forecasted by this or any other method, and this conclusion agrees with the established efficient market theory. The work cited is among the first few that attempted to incorporate wavelet and CI methods such as GP to approach systematically the time series forecast problems. However, in using the Haar wavelet instead of Daubechies transformation, flexibility of finding the most suitable shape to match the data set is forfeited. Further work, therefore, is needed to extend the method to forecast volatility series.

2.1.6 "Forecasting high-Frequency financial data volatility via nonparametric algorithms -evidence from Taiwan financial market" (Lee, 2005)

This paper uses two CI algorithms ANN and GP for forecasting financial data volatilities of four residents in the Taiwan Stock Exchange (TAIEX) at high frequency with different horizons and compares the output sample forecasting performances with certain parametric volatility models such as HISVOL, GARCH(1,1), EGRACH(1,1) and GJRGARCH(1,1). Their results reveal that algorithms based on nonparametric CI are powerful for modeling high-frequency intraday financial data volatility and establishes the fact that this is the way for future research.

To compare the forecasting performance, several well known parametric volatility models, *e.g.* HISVOL, EGARCH, GARCH and GJR-GARCH model are used. The reason of taking parametric models as benchmark is that parametric model such as GARCH-type are easy to estimate and readily interpretable. It is also the least complex model that describes volatility clustering. Furthermore, it is widely used by both academic researchers and market practitioners and proves to be a good benchmark volatility model in majority of contemporary research works.

To sum up, the authors cited here find GP volatility forecasting model is clearly superior to all other models when judged by kernel density plot in most of different time horizons. One potential drawback of the approach is the calculation efficiency. The problem would become more acute if optimization is needed to determine the best forecasting horizons. The approach would be limited for medium to longer term applications and not for, say, day traders, if hours of CPU time in typical PC's are a constraint to complete one set of forecast.

2.2 Volatility Forecasting with Genetic Algorithms – Key Publications

2.2.1 "Explaining exchange rate volatility with a genetic algorithm" (Lawrenz and Westerhoff, 2000)

Traders evaluate and update their mix of rules from time to time. To be able to understand the dynamics, the authors referred here concentrate only on a limited number of trading rules. The agents are assumed to have the choice, between three technical and three fundamental trading rules. The selection process is modeled by GA's, which has proven to be a useful tool for describing learning behavior in a variety of earlier papers (Dawid 1999).

The chartists-fundamentalists approach is another research direction, which focuses on explaining speculative transactions. Of crucial importance in this class of models is the behavior of the so-called chartists and fundamentalists because the interaction between the two groups has the potential of generating the interesting non-linear dynamics into the problem at hand. Chartists are those who base their investment decisions purely on analyzing historical price data, while fundamentalists focus on the fundamental aspects of the underlying assets, *e.g.* company's cash flow, equity's price to earning ratio, *etc.* By analyzing only two groups, strong structural dynamic relations among the variables still remain. To overcome this problem the models have to introduce some stochastic features. More recently, some multi-agent models in the spirit of the chartists-fundamentalists models have emerged. See LeBaron (2000) for a full survey. Since these models allow for many interacting heterogeneous agents, the structure in the data declines endogenously.

The aim of the cited paper is to develop a realistic, yet simple exchange rate model to get a deeper understanding of the driving forces underlying the foreign exchange dynamics. Rather than deriving the results from a well-defined utility maximization problem, details from the market microstructure and psychological evidence are used to motivate the framework. They construct a model where heterogeneous boundedly rational market participants rely on a mix of technical and fundamental trading rules. The rules are applied according to a clever weighting scheme. Traders evaluate and update their mix of rules from time to time. To be able to understand the dynamics they concentrate only on a limited number of trading rules. The agents have the choice between three technical and three fundamental trading rules. The selection process is modeled by a GA, which has proven to be a useful tool for describing learning behavior in a variety of papers. For an overview, please refer to Dawid (1999). Thus, the authors here derive the dynamics endogenously from learning processes on individual level rather than imposing random disturbances.

In case one wants to derive a model that includes some driving factors that underlie the system of volatility, this could be a good starting point. In other words, this is a traditional approach such that a model is built to explain a phenomenon and to forecast the future event based on the built model. However, market driving forces are more than two groups of investors. Moreover, chartists and fundamentalists behave in ways that are far more complex than three trading rules can explain. As a result, this approach suffers a bias problem just like those parametric models.

2.2.2 "Investment analytics volatility report" (Kinlay et al. 2001)

A proprietary asset allocation and optimization model reported in this work enables the construction of a portfolio comprising different weightings in each long or short option positions that optimize the investment objectives. These weightings are computed using a **proprietary non-linear GA**. Not much public information is available regarding the internal composition of this GA method. In simplified terms, the authors of the paper find, from empirical data, that realized variances tend to be log-normally distributed and asset returns standardized by realized standard deviations tend to be normally distributed. This suggests that a lognormal-normal mixture may be considered as a good model for asset return forecasts. Based on such a model, the authors could estimate and forecast the volatility values that are crucial in calculating the price of the index options. As a result, one can be more confident about the strategies designed to capitalize on the corresponding option trading.

In essence, the report referred here may be characterized by the following investment strategy:

a) Asset class: Equity options.
b) Strategy: Volatility arbitrage.
c) Methodology: Statistical modeling.
d) Style: Market neutral.

Equity options are among the most popular derivative investment vehicles and provide investors with high level of liquidity. As indicated in Section 1.1, volatility arbitrage is more promising than concentrating on the rate of return investments, because of the more predictable nature of volatility. Moreover, market neutral investment strategies enable investors to make profits without taking significant directional risks, *i.e.* less relying on the forecast of the direction of the market. Up to now, forecasting the movement of the market itself is far more difficult than forecasting that of volatility of the market. Kinlay *et al's* proprietary asset allocation and optimization model enables the construction of a portfolio comprising different weightings in each long or short option positions that optimize the investment objectives.

As a result, the subject matter of this research dissemination is exactly the area where our research should target. The results reported in this reference are encouraging as it claims up to 75% of accuracy in forecasting the future direction of volatility of the stock and equity indices. The fact that their methodology is completely proprietary makes our research effort on this thesis subject even more compelling.

2.2.3 "Rules extraction in short memory time series using genetic algorithms" (Fong and Szeto, 2001)

On an artificially generated time series data with short memory, rules are extracted and tested in order to form a basis for the predictions on the test set. A simple GA based on a fixed format of rules is introduced to do the forecasting. The results are markedly improved over those derived based on the traditional approaches, *e.g.* random walk and random guess.

There are different methods to estimate and forecast volatility, as summarized in the current work. However, the methodology demonstrated in this paper, *i.e.* attaining a forecasting accuracy of 50% – 60% by using 100 simple IF/THEN rules on a 4-lag recursive memory (sliding window), seems to be more straightforward and feasible. Therefore, it would be interesting to build upon this method to analyze some time series data in the real world. This approach could thus be considered in our research when applying the GA to find the rules that can best forecast IV in different time horizons.

2.3 Volatility Forecasting with other Methods – Key Publications

2.3.1 "Nonlinear features of realized FX volatility" (Maheu, 1999)

In order to investigate the time series properties of the FX (foreign exchange) volatility, this paper implements a non-parametric measure of daily volatility that is estimated by using the sum of intra-day squared returns. The specification of a functional relationship between this estimate of *ex-post* (realized) volatility and a latent data generating process (DGP) for daily volatility, allows the latter to be parameterized in terms of realized volatility, other variables in the information set, and an error term. The author explores nonlinear departures from a linear specification using a doubly stochastic process under duration-dependent mixing. For example, the author evaluates the importance of time varying parameters and persistence. Furthermore, the author also finds the structure that parameterizes the conditional variance of volatility and can capture large abrupt changes in the level of volatility. The importance of non-linear effects in volatility is gauged by in-sample statistical tests and by out-of-sample forecasts. The volatility forecasts are also evaluated using a simulated trading exercise involving FX straddles. These results have implications for forecast precision, hedging, and pricing of derivatives. The author indicates that stochastic jumps in the conditional mean of the price process have a long tradition in the finance literature but that further research trying to establish the presence of jumps in the conditional variance was just getting underway (Maheu, 1999). This paper states that it intends to solve a similar problem as ours here, but with different type of approaches. Since it still employed stochastic analysis, the question of "which type of distribution should be used?" still remains. Therefore accuracy of their forecasts would become questionable. However, results of this paper could be used to compare with results from any subsequent research such as ours, for establishing thresholds of accuracy of a prediction.

2.3.2 "A Nonlinear structural model for volatility clustering", (Gaunersdorfer, 2000)

Gaunersdorfer's paper referred here attempts to explain the source of the volatility clustering, *i.e.* the interaction between two types of traders: fundamentalists and technical analysts. The beliefs of these two types of traders are driven by an adaptive, evolutionary dynamics according to the reported success of the prediction strategies in the recent past, conditional upon price deviations from the rational expected fundamental price. Such beliefs cause asset price to switch irregularly between *a*) the fundamental price fluctuations with low volatility and *b*) persistent deviations from fundamentals, which are triggered by technical trading, and thus create higher volatility.

The key feature of the nonlinear structural model is therefore, given by the following expression:

$$(1+r)p_t = n_{1t}\left(p^* + \nu\left(p_{t-1} - p^*\right)\right) + n_{2t}\left(p_{t-1} + s\left(p_{t-1} - p_{t-2}\right)\right) + \bar{d} + \varepsilon_t, \tag{4}$$

where $p^* = \bar{d}/r$ with r being the discount rate and $\bar{d}$ the expected dividend, whereas p_t is the asset price at time t, n_{1t}, n_{2t} are the fractions corresponding to fundamentalists and technical analysts, ν is fundamentalists' belief that tomorrow's price will move in the direction of the fundamental price p^* by a factor ν, s is an intensity factor that is specified for optimizing the past successful strategies, and ε_t is some IID variables representing the model approximation error or the dynamic noise. The modeling of n_{1t} and n_{2t} is where the non-linearity is introduced, *i.e.* they are estimated as $\tilde{n}_{h,t}(h,t)$ given by the relation:

$$\tilde{n}_{h,t}(h,t) = \exp\left(\beta U_{h,t-1}\right)/Z_t, \quad h \in \{1,2\}; \tag{5}$$

where β is the intensity of choice, measuring how fast the mass of traders will switch to the optimal prediction strategy. $U_{h,\ t-1}$ is a performance measure, which is the evolutionary fitness of predictor h in period t – 1 given by utilities of realized past profits. Z_t is a normalization factor such that the fractions do add up to unity or one.

This paper goes on to use the model to calculate the autocorrelations of the rate of returns, absolute rate of returns, return squares for an assumed asset to simulate the S&P 500 index during the last 40 years. It concludes that the rate of returns have the non predictability just like the real world data, while the absolute rate of return and return square (*i.e.* variance) do show to follow the trend closely with the index. However, the paper does not mention about the model's forecasting capability and did not intent to deal with it in the future. Moreover, the parametric nature of the model deems to exhibit a lack of flexibility and is difficult to test for validity. Please refer to APPENDIX 2 for more discussion about the parametric model of volatility forecasting.

2.3.3 "Forecasting and trading currency volatility: An application of recurrent neural regression and model combination" (Dunis, 2002)

This paper uses non-parametric Neural Network Regression (NNR) and Recurrent Neural Network (RNN) regression models to forecast and trade FX. Results based on this NNR model depend crucially on the choice of the number of hidden layers, the number of nodes and the type of non-linear transfer function retained. The use of NNR models enlarges the analysts' toolbox of available techniques by adding convenient models where no specific functional form is *a priori* assumed. RNN models are different from NNR models in that they include a loop back from one layer, either the output or the intermediate layer, to the input layer. The following factors have been applied to explain the exchange rate volatility: exchange rate volatilities (including the one to be modeled), the evolution of important stock and commodity prices, and the evolution of the yield curve as a measure of macro-economic and monetary policy expectations. Allowing for transaction costs, most of the trading strategies that are constructed based on the two neural network models produced positive returns. In other words, the volatility forecasting accuracy is in average, slightly over 50%. One might argue that a more elaborate model could produce better results. The contrary has been demonstrated by the authors when they combined the NNR with RNN models. Other recent work in the NNR field might shed more light in this regard. However, it is out of the scope of the current research, as we propose to concentrate mainly on analytical techniques in the EA area.

2.3.4 "An Empirical approach toward realistic modeling of capital market volatility" (Wang et al. 2005)

The authors referred here find that an all-over-the-time stationary generalized constant elasticity of variance (CEV) model will mismatch the mean reverting level θ as well as ignore the jump phenomenon. So they propose a jump-decaying CEV model to depict the realized IV process. The CEV model assumes the volatility process with a mean reverting level. Wavelet technique is introduced to verify that the volatility jumps are not natural to CEV modeling. Therefore, they find that the volatility series is essentially a stochastic jump-decay process rather than being all-over-the-time stationary in each market. These findings strengthen the theoretical foundation of our non-linear evolutionary approach towards IV forecasting.

2.3.5 "Detection and prediction of relative clustered volatility in financial markets" (Hovsepian *et al.* 2005)

This paper presents a methodology for detection and prediction of periods of relatively increased volatility in the time series data. It uses a synthesis of three concepts and methods derived from the field of computer science – *support vector classifiers* (SVC), statistics – GARCH, and signal analysis – *periodogram*. However, it still is GARCH based and only predicts a period of either volatile or non-volatile without specifying how volatile the future reading is. Moreover, it is verified with only simulated real-time cases and hence does not give the reliability expected for market applications.

2.3.6 "Volatility forecasting with sparse Bayesian kernel models" (Tino, *et al.* 2005)

These authors find that models built on quantized sequences *i.e.*, symbols such as signs + and - representing the sign of daily volatility differences, give superior results when compared to those constructed on the original real-valued time series. More specifically, their paper finds that quantization technique significantly improves the overall profit and quantization into just two symbols gives the best results. By trading option straddles on indices and achieving positive profit daily, it confirms the benefit of converting the real value time series into quantized sequence. It also finds that volatility patterns in durations of five and ten days could be detected and used to forecast the one-day-ahead volatility. The authors claim to achieve profitability by applying sparse Bayesian Kernel models based on neural networks. However, it works on daily volatility instead of IV, which has been demonstrated to be less accurate and appropriate. Moreover, this method is based on the premises that the price of the option straddle is strictly positively proportional to volatility. In reality many factors could affect their relationship. In cases where volatility switches signs, *e.g.* slightly negative to slightly positive, option prices will not change substantially and return of buying a straddle option could be negative after deducting the transaction cost. This is the other disadvantage of using just the sign instead of interval to classify the data.

2.3.7 "Boosting frameworks in financial applications: from volatility forecasting to portfolio strategy optimization" (Gavrishchaka1, 2005)

Based on several past experiences and trials, it is believed that accurate volatility modeling does not always warranty optimal decision making that leads to acceptable performance of a portfolio strategy. In this work, a boosting-based framework for a direct trading strategy and portfolio optimization has been introduced to strengthen such optimization procedure.

The author attempts to construct a model from the well-known market indicators for short-to-medium time horizons (from several months to 1-2 years). By calculating strategy returns on a series of intervals of length τ shifted with a step $\Delta\tau$ and encoding them as +1 (for $r >= r_c$) and -1 (for $r < r_c$), it is shown that one can obtain symbolically an encoded time series (distribution) of strategy returns. This is an incremental search algorithm without the aid of computation accelerating tools such as EA. Therefore, its efficiency is questionable. We can only conclude that it would be interesting to combine this approach with GP, in which different formulae could be evolved to find the ones best fit the time series. Moreover, this method could only distinguish positive from negative volatility, which has limitations similar to what was described earlier in section 2.3.6.

2.4 Summary and Discussion

We begin this subsection by presenting a table to summarize the important contributions and analytical characteristics of the reviewed papers discussed in the sections preceding this page. The information contained in the following table will provide us with important guidelines to further our discussion on the approach that we plan to take in this thesis research.

Based on the following summarized literature survey, one can only conclude that the contemporary research of volatility forecast has just begun to venture into the non traditional domain particularly in the CI area such as GA/GP in an attempt to seek better solutions on a backdrop of active IV research; each approach has its own advantages and weakness. IV provides a good starting estimation of the current volatility, and as indicated in previous sections, we as researchers could certainly apply a variety of powerful techniques including stochastic analysis to forecast future volatility more accurately. GA/GP on the other hand, could deal with non-linearity in an effective and progressively efficient manner, which opens up alternative application avenues besides the rigorous exercise in the traditional academic sense.

Table I

Summary of the papers reviewed in details

Author(s)	Approach	Goal	Comments
Pictet et al., 2001	GP	Discover new FX volatility models using "typed" GP trees.	Does not take into account non-linearity. Use FX symmetry to reduce considerably the search space.
Zumbach *et al.*, 2001	GP+LS	Use hybrid GP to forecast FX volatility.	Does not take into account non-linearity. Use FX symmetry to reduce considerably the search space.
Chen & Yeh, 1997	GP	Use a recursive GP to detect and adapt to structural changes of market volatility.	Explicit recognition of non-linearity but does not attempt to forecast. Integrated Volatility was not used.
Neely & Weller, 2001	GARCH, GP, RiskMetrics	Compare three approaches: Parametric, generalized parametric and non-parametric in FX forecasting.	In many instances, GP out-performed the other approaches. They were tested on FX volatility only. Non-linearity is not accounted for. No IV.
Kaboudan, 2005	GP, wavelet, NN	Apply an integrated approach to forecast one-step as well as 16-step-ahead exchange rate forecasting.	Does not deal with volatility. Other type of wavelet might improve the effectiveness.
Lee, 2005	ANN + GP, GARCH	Compare the computation intelligence method with GARCH models.	Better results are achieved at questionable calculation efficiency for medium to long forecasting horizons.
Lawrenz & Westerhoff, 2001	GA	Explore how trading rules can explain market volatility. Use GA to combine six simple trading rules using the chartists – fundamentalists point of view.	In real world there are more than two players (*i.e.* chartists and fundamentalists) and trading rules are much more complex.
Kinlay *et al.*, 2001	GA, SM	Asset allocation and optimization system based on a weighted sum technique. The weights are determined by statistical inference and aided by a GA.	Proprietary techniques with many undisclosed details. Best published results with 72% – 75% prediction accuracy.
Fong & Szeto, 2001	GA	Use GA to determine simple if – then – else rules in order to predict the behaviour of artificially generated time series.	Obtained 50% - 60% accuracy using only 100 simple if – then – else rules. Demonstrated the search power of GA applied to stochastic series.

Table I (continued)

Summary of the papers reviewed in details

Maheu, 1999	SM	Explores the nonlinear features of FX integrated volatility.	Found that stochastic jumps (structural changes) are a very determinant feature in IV.
Gaunersdorfer, 2000	SM	Attempt to define a nonlinear model that explains the volatility clustering phenomenon.	It concludes that the rate of return have non predictable behaviour while the variance does show trend that is close to the index measured. Thus confirming the usefulness of the integrated volatility approach.
Dunis & Huang, 2002	NN	Applied a non-linear non-parametric approach to forecast and trade FX	Achieved slightly above 50% of forecast accuracy. But elaborate models produced poor results.
Wang *et al.* 2005	CEV	Account for the non-linearity in volatility with a stochastic jump-decay process	Provides further theoretical foundation for the current research to deal with non-linearity.
Hovspian, *et al.* 2005	SVC, periodogram , GARCH	Detect and predict periods of relatively increased volatility by a synthesizing method.	It is still a GARCH based approach, *i.e.* parametric models and lacks verification with real data sets.
Tino *et al.* 2005	Sparse Bayesian Kernel	By quantizing real value time series, forecast the one-day-ahead volatility to generate profit.	Only forecasts the directions of volatility.
Gavrishchaka, 2005	Boosting framework	Make optimal investment decisions by forecasting the directions of volatility.	Other calculation methods could help improve efficiency. Only forecasts the directions of volatility.

GA – Genetic Algorithm, GP(+LS) – Genetic Programming (with local search), SM – Statistical Methods, NN – Neural Network, CEV – Constant Elasticity of Variance, SVC – Support Vector Classifier.