G53IDS Dissertation

ARIMA-NN Based Exchange Rate Forecasting

Jianheng Qiu

4276906

Supervised by: Prof. Dr. Bai Li

I hereby declare that this dissertation is all my own work, except as indicated in the text.

School of Computer Science

University of Nottingham

April 24 2018

Abstract:

In this project, a model based on the combination of ARIMA and neural networks (PSONN and ANN) is investigated to forecast foreign currency exchange rates. The purpose of this research is to design a combination methodology benefiting from the strengths of both ARIMA and NN models to predict exchange rate series properly and accurately, since the exchange rate series contains characteristics of random and uncertainty. The models were trained from historical data of different time series. The linear part of the series set is processed through the ARIMA model, and the remaining nonlinear part is processed by ANN. The objective of the project is to predict three foreign exchange rates against USD (GBP, AUD and JPY). The performance of the model was evaluated using a set of widely accepted statistical metrics and compared with the performance of models accomplished by single methodology. With the analysis of the data, it is possible to conclude that the model designed with combination methodology gives a close prediction of exchange rates and contains better performance than models designed with single methodology.

Keywords: ARIMA, ANN, PSONN, Exchange Rates, Financial Forecasting with Combination Methodology.

Contents:

Introduction:

1. Motivation and Background:

Globalization has impacted nearly every aspect of modern life and continues to be a growing force in the global economy, in spite of a few drawbacks, such as protectionism and nationalism. Meanwhile, the financial risks have had the power to influence global market and been the most challenging force to destroy consistent economic growth around the world. The recent international economic crisis has given the warnings for banks that it is important to implement effective systems to control the risks of market. Particularly, international activities accomplished by the largest investment banks and the increasing variability of exchange rates have highlighted the significance of exchange rate risk. Besides, due to the introduction of new financial principles, such as floating exchange rates, and the dramatically developing process of global market of the past decades, the currency exchange market has developed unparalleled increments. Such changes contribute lots of new features and uncertainties to economic and financial environment of exchange rate. As a result, foreign exchange rates were not only determined by the balance of payments like before and become a complex and intractable issue for current time. Not only the active management of banks needs effective forecasting models, nations also require efficient techniques to avoid potential exchange rate risks. However, it is very difficult in practice, because the dynamic change of a financial time series data is affected by multiple economic variables, including economic growth, interest rate, inflation, deflation, political decision and psychological factors. So it turns to be more and more difficult to design and implement an effectively model to forecast foreign exchange rates.

2. Aims and Objectives:

The aim in this project is to develop a combining model to predict the trend and value of exchange rates up to a period of days ahead of the last data used for training. The model is creatively constructed by a combination of ARIMA, PSONN and ANN. Therefore, the performance of the model needs to be compared with other forecasting models that have been proposed by other literatures and have been proven to be effective. The paper will show detailed evaluation process and give an agreement that the model has advancing advantages and performs better. The paper will mainly concern technical data, which means the project only use the exchange rate data itself to analyze its own trend. The forecasting period will be short-term and further researches or improvements will be shown in the end.

The key objectives of this research:

1. To explore the intrinsic characteristics (linear or nonlinear) of exchange rate data.
2. To construct an adaptive and self-learning approach under the implementation of hybrid model to solve exchange rate prediction problem of particular currencies.
3. To show the performance results and evaluate the new model whether good or not.
4. To investigate the comparisons between hybrid models and single method models.

3. Related Work:

There are many methods to predict exchange rate. In general, those methods could be divided into four main representative methods: The theory of purchasing power parity, the structural equation of the determinants of the exchange rate, the time series analysis method of the historical data of the exchange rate itself and black box models.

- First, the purchasing power parity theory was useful during the 19th century, because foreign exchange rates were only determined by the balance of payments in the past [5]. That theory was merely a way of listing payments and receipts in international transactions for countries. For details, the payments contain a supply of the domestic currency and a demand for foreign currencies. Similarly, the receipts involve a demand for national currency and a supply of foreign currencies. The balance was mainly ensured by import and export trades. Thus, exchange rates trends were obvious according to the balance. However, as mentioned, the expanded global market and kinds of financial reforms had greatly changed currency exchange market. So such method is not suitable for current situation, especially, for floating exchange rates.
- Second, time series analysis method (linear ones), such as Auto-Regressive Integrated Moving Average (ARIMA), Exponential Smoothing Method and Random Walk, which through the learning of the information transmitted by the exchange rate itself, trying to analyze the reflection of the predictive value of the exchange rate[6]. These methods are flexible and easy to be implemented. Especially, the ARIMA method has been widely used for time series forecasting and been used as a benchmark to evaluate many new modelling approaches [11, 12]. However, the shortcomings are obvious, the financial time series (including the exchange rate) is mostly nonlinear, or is a complex system containing a nonlinear system [9, 10]. These methods are mainly based on linear relation assumption, so the predictions through the models contain large limitations in practical uses.
- Third, black box forecasting models (nonlinear ones), such as neural network, fuzzy models and genetic algorithms, which through the analyzing of the problem, attempting to identify and forecast the nonlinear and non-random dynamics of time series, but without clear explanations or logical functions that bind the variables in the research. Particularly, more and more attention has been paid to NN method because of its strong ability of learning and data processing. NN models can deeply explore complex and difficult internal relations, which are even difficult to be descripted in mathematics, to analyze the nonlinear characteristics behind data [13, 14]. However, as Sun [15], Maia [7] and Plasmans et al [16] found that when analyzing random and linear parts of time series, neural network shows worse prediction accuracy than ARIMA models.
- Fourth, the structural prediction models, such as Autoregressive Conditional Heteroscedasticity (ARCH) and Generalized Autoregressive Conditional Heteroscedasticity (GARCH), which are depended on the general point that all activities of traders could be understood by a model of explainable behavior and thus by a definite, clear equation or function that can combine with variables determinants of the phenomenon to be predicted. However, these models use explicit expressions to describe the relationship between data. But relationship between nonlinear data is generally extremely complicated, and the nonlinear modes among data are versatile. Therefore, it is impossible to use any particular structural models or functions to completely depict the nonlinear relationship [17, 18, and 19]. The results of these models always do not meet the expectations of researches.

As mentioned, the ARIMA method is based on the techniques of linear time series prediction, but the processing of nonlinear data is not reasonable, even the effectiveness is inferior. On the contrary, at the aspect of exploring the nonlinear relationship implied in the data, NN models are superior to other nonlinear models incomparably. However, the performance of neural networks in manipulating data with linear features is always not as well as the traditional linear techniques, like ARIMA model [7, 16 and 20]. In practice, most financial time series, usually include both linear temporal components, but also contain the nonlinear time series components, showing a complex linear and nonlinear characteristics, so the single linear or nonlinear forecasting model cannot capture commendably the intricate characteristics of time series [7, 8 and 20]. In the past until now, researchers still cannot determine which one of the NN models and ARIMA method is more suitable to predict exchange rate. For example, Jhee and Lee [21], Kamruzzaman and Sarker [22] and Wang and Leu [23] and many other researchers have shown that NN perform better than ARIMA methods, especially, for more irregular data and for multiple-period-ahead forecasting. In the contrary, Sun [15] and Zhang [4, 8] argued that ARIMA methods contain higher accuracy in forecasting than NNs in some circumstances. It is not wise to apply NNs blindly to any type of data. Those researches and arguments reveal that exchange rate forecasting problems cannot be solved appropriately by only single method. Therefore, researchers now are inspired to find a combination mechanism to predict the exchange rate.

For the purpose of improving forecasting performance, researchers began to combine ARIMA and NNs together to achieve mutual promotion and supplementation among the two methods. Since it I complicated to completely decide the characteristics in real problems, hybrid methodology that includes both linear and nonlinear modeling capabilities could be an appropriate strategy for practical use. Some representative researches ahold be reviewed:

- In 1996, Wedding and Cios [24] proposed an integrated model of the traditional ARMA model and RBF neural network to predict exchange rate. In the research, the RBF neural networks were trained to generate both time series forecasts and certainty factors. Their output was then combined with the ARIMA models to predict future values of data. This combination approach was shown to improve the overall reliability of time series forecasting.
- Tseng et al [25] studied a hybrid model (SARIMA-BP) with the seasonal time series ARIMA (SARIMA) and the neural network back propagation (BP) models. The research firstly manipulated the time series with ARIMA model to analyze the linear part of data, then the remained part was transmitted to neural network models for further analysis. The combination of the analysis results would be the forecasting of the time series. The performance of prediction was compared among four models, i.e., the SARIMABP, SARIMA models and the two neural network models, separately.
- G.P.Zhang [8] proposed a similar research to predict exchange rates of GBP/USD. The team implemented a hybrid model combined of ARIMA and ANN and the general procedure was same as Tseng's work. The ARIMA modeled the linear component while ANN modeled non-linear one. However, Zhang's paper also introduced some other related time series that came from different areas and had different statistical characteristics, to test the generalization of the proposed model.
- Michele [26] presented a large experiment to challenge the accuracy of the combination methodology with 3003 series. Researchers proposed a simple model-selection criterion to select among forecasts, and the experiments shown that, using this criterion, the accuracy of the selected combinations was significantly better and less variable than that of the selected individual forecasts. The results indicated that it was less risky in practice to combine forecasts

than to select an individual forecasting method.

- Hong et al [20] made improvements to the combination method. In the research, Random Walk methods was chose to model the linear part of exchange rate, this step made little changes to the whole model because RW method was also a branch of ARIMA. However, Hong's team both FANN and EANN to model the nonlinear component. The mean value of the outputs of the two networks were chose as results. The final report showed that such method contained even higher effectiveness.

4. Description of the work:

As the above analysis shows, the overall prediction performance of the hybrid methodology is significantly better than that of the individual characteristic model, and in practical applications, the risk of using the combined model is much less than that of the single model. There are also many literature studies that show that different integrated models do improve the forecasting performance, compared to individual ones [27, 28 and 29]. Based on these previous researches, this research will also implement a hybrid methodology of ARIMA and NN networks to predict the foreign exchange rates. Unarguably, both ARIMA and NN models have achieved significant successes in their own linear or nonlinear domains. This research will create a new model that combines of both advantages and try to achieve further improvements on forecasting accuracy. To establish such model, the ARIMA, PSO neural network and BP neural network will be used as modeling techniques to analyze the internal relations of exchange rates. The ARIMA will model the linear component and the nonlinear part will be modeled by both PSO and BP neural networks. Besides, one necessary research hypothesis should be concerned as pre-requisites: An exchange rate series is composed of linear and nonlinear autocorrelation components and the two parts can be stripped from the time series and can be modeled separately. Actually, many studies has proven that the financial time series are constructed in that way [4, 7, 20, 25]. If the hypothesis is invalid, it would assume that the exchange rate series only contain either linear or nonlinear characteristic. That means the series could be completely modeled by individual model methods. However, as those listed papers shown, the fact is not true. Therefore, the hypothesis is reasonable and valid.

The main aims are to analyze the ability of combination methodology models of both linear and non-linear nature, to highlight random, separable and predictable behavior in a highly volatile market. To this end, the model was developed and empirically tested with further evaluations for predicting currency exchange rates.

Methodology:

1. The ARIMA Model

Essentially, Autoregressive Integrated Moving Average model (ARIMA) is an extension of Autoregressive Moving Average model (ARMA). It was proposed by Box and Jenkin [30] and it has been widely used for time series modeling and forecasting. In most cases, ARMA model cannot be directly used, because the construction of ARMA models requires time series to satisfy stationarity. However, in practice, observed time series always presents some special characteristics, such as trend, cycle or heteroscedasticity, which do not meet the required condition of stationarity. To solve such problem, in Box and Jenkin's research, d order differencing and power transformation are often applied to the series data to remove the trend and stabilize the variance so that the manipulated data is stationary and fitted for ARMA (p, q) model to use. The above process is the general description of ARIMA modeling technique. In language, the ARIMA forecasting method can be introduced as the future value of a variable is assumed to be a linear function of several past observations and random errors. In mathematical expression, it is:

$$y_t = \theta_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p}$$
$$+ \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q},$$

In this equation, y_t are the actual value at time period t and $\mathcal{E}_t$ are random error for time t. These y_t have been stabilized by differencing and power transformation, if the original data was not stationary. While these $\mathcal{E}_t$ are mutually Independent white noise sequence, besides, they follow normal distribution with a mean of 0 and a constant variance of σ^2. $\emptyset_i$ (i = 1, 2, ..., p) and θ_j (j = 0, 1, 2, ..., q) are model parameters, respectively. The p and q are integers that referred to as orders of the model. One significant central task of ARIMA model construction is the determination of the proper model order (p, q).

The process for ARIMA modeling and forecasting includes 4 domain steps.

- First, as mentioned before, the stationarity is a necessary condition in building an ARIMA model, otherwise, the model could not manipulate and analyze the variable data. A stationary time series contains the capability that its statistical characteristics including the mean and the autocorrelation structure are constant while the time passing. Such stability is extremely important for the continuing steps.
- Second step is model identification, the basic idea of this stage is that if a generated time series comes from the process of ARIMA, then it should reflect certain theoretical autocorrelation properties. Afterwards, by comparing and matching the theoretical autocorrelation properties with empirical patterns, it is probable to identify one or several potential models for the given generated time series. In the research, the autocorrelation function and the partial autocorrelation function of the sample series are used as fundamental tools to identify the order (p, q) of the model of ARIMA part.
- Third, the step of parameter estimation, after the forwards two stages, a tentative model should be specified. The criterion for estimating parameters is based on minimizing overall measure of errors. By using a nonlinear optimization procedure, the estimation of parameters is easy and straightforward.
- Final step is diagnostic checking of model adequacy. This step tends to check whether the model assumptions about the errors are correct. Kinds of diagnostic statistics and residuals plots can be used to evaluate the performance of fit of the provisionally most suitable model to the historical data. If the model is checked to be inadequate, a new tentative model needs to be identified and selected. Then the process goes through the previous steps, model verification and parameter estimation, again.

Such three-step model constructing process was proposed by Box and Jenkins as three iterative steps for ARIMA modeling and forecasting. It would be typically repeated several times until a proper model is finally found. Then the final one can be used for predicting stage.

2. The NN Approaches

Since the linear characteristic of exchange rate data has been solved, the possible number of nonlinear models that can describe and model the rest parts of exchange rate data is large. According to the suggestion of Gooijer and Kumar [31], a good nonlinear model should suit the following evaluation: "generally enough to capture some of the nonlinear phenomena in the data." The neural networks are one of these models that are able to estimate versatile nonlinearities in the exchange rate data. The research of NN models was originally derived from the imitation of the human brain structure. Gradually, neural network technology has played an important role in many academic fields, such as classification, recognition and prediction [20]. With further in-depth studies, the neural network techniques had generated numerous derived models and they learnt the intrinsic relationship of data through a number of interconnected neurons distributed in different layers. These models always contain unique properties of non-parametric, non-assumable, noise-tolerant and adaptiveness. In addition, NNs are effective in analyzing the dynamics of non-stationary exchange rate data and can map any nonlinear function without a priori assumptions about the data [32]. Within these neural networks, multilayer feedforward neural network based on back propagation algorithm (BP) is a typical and classic model (BPNN) and has been used widely for decades [4, 5, 8 and 9]. Using BP algorithm or gradient descent algorithm to optimize model parameters, the model solutions are easy to get trapped in local optimal area of search space. Besides, the slow learning speed and long training time seriously reduce the searching efficiency and prediction effectiveness [33]. Recently, as a representation of intelligent swarm algorithm, the particle swarm optimization (PSO) algorithm has been widely concerned and highly valued in the NN forecasting (PSONN) due to its simple operation, straightforward implementation, rapid convergence and outstanding global performance [34]. PSONN is also a feedforward network and the only difference between the two networks is the type of learning algorithm. Obviously, ANN is based on gradient descent algorithm while PSONN decides on PSO to optimize the weights. Contrary to ANN, PSONN converge quickly during the initial stages of a global search, nevertheless, around global optimum, the search process becomes very slow. This disadvantage is always described as "precocity" by researchers [35]. It also implies that the PSONN still needs further more researches to be perfect. So for this paper, to obtain a good forecasting result, the designed model will combine the two networks together and try to fully utilize their advantages but avoid their respective defects. At the following stage, the similarities of two neural networks will be introduced firstly and their differences will be discussed separately.

To establish PSONN and BPNN, the first of all is to determine the network structure. As mentioned, the NN methods deal with data relations through interconnected neurons of different layers, since the input and output layers are fixed according to the requirements of research, the key to the network structures is the number of hidden layer and included nodes. In several applications, Yao et al [36], Zhang et al [37] and Hornik [38] have reported that increasing the number of inputs cannot necessarily promote the NN's performance. Although different literatures chose different structure optimization methods, single hidden layer is the most widely used structure form for feedforward network to model and predict time series [37]. So this research will also choose such structure to form PSONN and ANN.

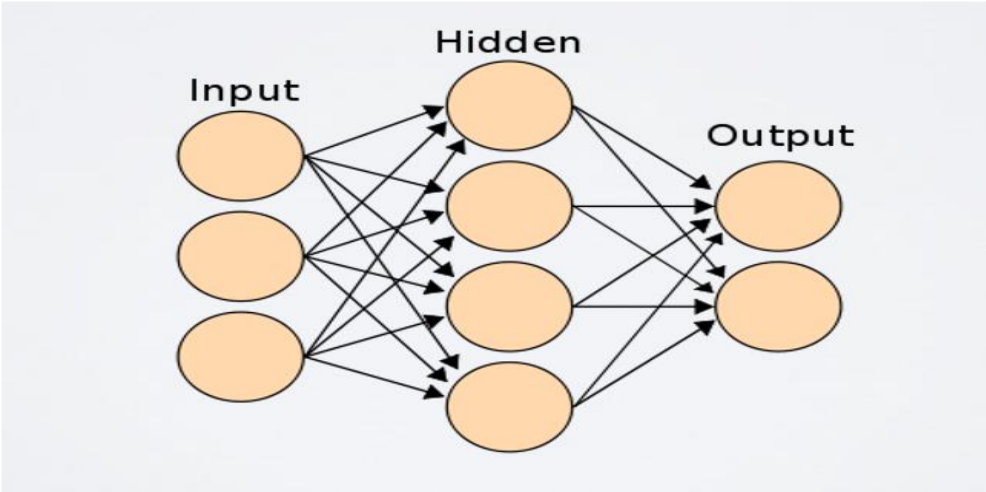


Figure 1. Structure for neural networks.

As a result, the model is designed as a network of three layers of simple processing units related by acyclic lines. The relationship between the inputs (y_{t-1}, y_{t-2}, … , y_{t-p}) and output y_t can be described as below.

$$y_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j g\left(\beta_{0j} + \sum_{i=1}^{p} \beta_{ij} y_{t-i}\right) + \varepsilon_t,$$

In the representation, the α_j (j = 0, 1, …, q) and β_{ij} (I = 0, 1, …, p; j = 1, 2, …, q) are parameters of model which are often called as connection weights. The p and q are the number of input nodes and hidden nodes, respectively. The g(x) is the transfer function of hidden layer and logistic function is often used at this place, that is:

$$g(x) = \frac{1}{1 + \exp(-x)}.$$

According to the relationship of y_t, the neural network model always shows a nonlinear functional mapping from the past observations (yt – 1, yt – 2, … , yt - p) to the future value y_t.

$$y_t = f(y_{t-1}, y_{t-2}, \ldots, y_{t-p}, w) + \varepsilon_t,$$

In such equation, f(x) is a function decided by weights of connection and network structure. The letter w represents a vector of all parameters. Hence, the PSONN and ANN models used in this research are equivalent to a nonlinear regressive model, so in fact there is only one output node involved in the output layer. The relationship expression also reflects such structure.

As the number of layers is certain, next necessary step is to determine the number of nodes. In theory, the above network structure should be able to approximate arbitrary function as the number of hidden neurons q is sufficiently large [38]. In practice, this simple network structure is surprisingly powerful to work greatly in out-of-sample forecasting with a small number of hidden nodes q. The reason may be the overfitting effect, a typical problem that always happens in the neural networks modeling. An over-fitted model has a great fitness to the sample data used for model building and training, but it has poor ability of generalization for the data out of sample. However, there is no systematic principle or most recognized approaches in determining the parameter q. Therefore, the choice is researches discriminated and data dependent. Similarly, another significant task is to select the number of the lagged observations, parameter p, the dimension of the input vector, or in other words the number of input nodes. Parameter p plays a domain role in deciding nonlinear autocorrelation structure of the time series. Again, there is no recommended rule to guide the choice of p. so in the experiment part of this paper, the q and p will be explained clearly for PSONN and ANN.

Once a model NN (p, q) is specified, the preparation for its training is ready. The training process is quite same as the parameter estimation process of ARIMA, it is that the parameters estimated are compared according to overall accuracy criterion to find minimizations. In NN models, efficient nonlinear optimization algorithms are responsible for this process, those are learning algorithms. As mentioned in the beginning of this part, ANN will choose SCG and PSONN will choose PSO optimizer. Nevertheless, the estimated NN model is frequent evaluated with a divided holdout sample set that is never shown to the training stage. On the contrary, ARIMA model will use one sample for the whole building process, including identification, estimation and evaluation. This is due to pre-specified manipulation of ARIMA model and then the order can be analyzed from exchange rate data. After being stabilized to stationary condition, the model best fitted to experiment data then is the optimal model for choosing. Back to NN models, if the testing set is exposed to training step, the generated model is more likely to overfit the current rate data and reduces the generalization on other exchange rates.

2.1. The ANN model

The choice of learning algorithm are dramatically important for ANN model building and forecasting, because this choice affects parameters estimation (p, q) directly and then influences learning speed and global performance. The Standard Backpropagation algorithm is the most popular method and has been used as benchmark for other algorithms. Many researchers chose this algorithm to build the hybrid model [4, 20]. However, the search space in Backpropagation may encounter long ravines with sharp curvature and smoothly sloping plain which makes slow convergence [5]. Thus, this project will choose a more effective algorithm called, Scaled Conjugate Gradient (SCG). In the Kamruzzaman's paper [5], three widely used learning algorithms, Standard Backpropagation (SBP), Scaled Conjugate Gradient (SCG) and Bayesian Regularization (BPR), they are compared together on the prediction of different exchange rates. As a result, the SCG was proved much more powerful than the other two algorithms. In conjugate gradient methods, a search is following along conjugate directions, which performs generally faster convergence than steepest decent directions [39]. The new direction of conjugate gradient methods takes full advantage of the minimization achieved by the previous direction and the step size is adjusted in every iteration. Thus, the new direction is determined by a combination of new steepest decent direction and the previous search direction so that the current and previous moving directions are conjugate. The changes of weights in successive steps are expressed as below.

$$\mathbf{w}_{t+1} = \mathbf{w}_t + \alpha_t \mathbf{d}_t$$

$$\mathbf{d}_t = -\mathbf{g}_t + \beta_t \mathbf{d}_{t-1}$$

With details:

$$\mathbf{g}_t \equiv \nabla E(\mathbf{w})\big|_{\mathbf{w} = \mathbf{w_t}}$$

$$\beta_t = \frac{\mathbf{g}_t^{\mathrm{T}} \mathbf{g}_t}{\mathbf{g}_{t-1}^{\mathrm{T}} \mathbf{g}_{t-1}} \quad \text{or} \quad \beta_t = \frac{\Delta \mathbf{g}_{t-1}^{\mathrm{T}} \mathbf{g}_t}{\mathbf{g}_{t-1}^{\mathrm{T}} \mathbf{d}_{t-1}} \quad \text{or} \quad \beta_t = \frac{\Delta \mathbf{g}_{t-1}^{\mathrm{T}} \mathbf{g}_t}{\mathbf{g}_{t-1}^{\mathrm{T}} \mathbf{g}_{t-1}}$$

In the equation, W_t represents the weights, d_t and d_{t-1} are conjugate directions in following iterations and the step size is identified by α_t and search direction is decided by β_t, where αt is calculated by this:

$$\alpha_t = -\frac{\mathbf{d}_t^{\mathrm{T}} \mathbf{g}_t}{\delta_t} \qquad \delta_t = \mathbf{d}_t^{\mathrm{T}} \mathbf{H}_t \mathbf{d}_t + \lambda_t \left\|\mathbf{d}_t\right\|^2$$

Where H_t is the Hessian matrix at iteration t and λ_t is the scaling coefficient at time t. λ is necessary to avoid the weight upgrade to be an increase of error function. The λ needs to be optimal in each iteration according to the measure of quadratic approximation accuracy, which is Δ_t. There are some suggestions proposed by Moller [40] to change the λ.

$$\Delta_t = \frac{2\{E(\mathbf{w}_t) - E(\mathbf{w}_t + \alpha_t \mathbf{d}_t)\}}{\alpha_t \, \mathbf{d}_t^{\mathrm{T}} \mathbf{g}_t}$$

For $\Delta_t > 0.75$, $\lambda_{t+1} = \lambda_t/2$; For $\Delta_t < 0.25$, $\lambda_{t+1} = 4\lambda_t$; Otherwise, $\lambda_{t+1} = \lambda_t$

2.2. The PSONN model

At the basis if the construction of ANN, PSONN mode also uses particle swarm optimization algorithm to optimize the model and determines the model parameters in the end. The general principles of PSO algorithm is that, assume in n-dimensional search space, a population consisting of M particles represented as **X**, where $X = (x_1, \ldots, x_i, \ldots, x_m)$. Among the **X**, the position of i-th particle is $x_i = (x_{i1}, x_{i2}, \ldots, x_{in})^T$ and its velocity is $v_i = (v_{i1}, v_{i2}, \ldots, v_{in})^T$. Individual peak of particle I is $P_i = (P_{i1}, P_{i2}, \ldots, P_{in})^T$, while the global peak value of the population is $P_g = (P_{g1}, P_{g2}, \ldots, P_{gn})^T$. in the searching process, particle keeps tracking two targets to update its position and velocity timely. One is the optimal solution that the particle has found at present, as individual peak value, the other one is the optimal solution that is found by global population at present, as global peak value. There is an Iterative calculation formula:

$$v_{id}^{k+1} = \omega v_{id}^{k} + c_1 \xi \left(P_{id}^{k} - x_{id}^{k}\right) + c_2 \eta \left(P_{gd}^{k} - x_{id}^{k}\right)$$

$$x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k+1}$$

In the expression, v_i^k is the velocity of particle I in the k-th iteration, v_{id}^k is the velocity component of v_i^k at d dimension; The x_i^k is the position of particle I in the k-th iteration, x_{id}^k is the position component of x_i^k. P_i^k is the individual peak of particle I in the k-th iteration and P_{id}^k is the individual peak component of P_i^k at d dimension; P_g^k is the global peak of particle I in the k-th iteration and P_{gd}^k is the global peak component of P_g^k at d dimension. ω is the Inertia weight, c_1 and c_2 is called learning factors that influence the learning effect and speed. ξ and η are two random numbers distributed in the interval of [0, 1]. The concrete steps of model building are as follow:

- First, one important step is settling down learning factors c1 and c2, the Inertia weight ω and the maximum evolutionary algebra K_{max}. The current maximum evolutionary algebra should be k = 1. Then, within the defined domain, particle swarm should be generated randomly as **X**(k), including the initial particle position x(k) and the velocity v(k), where k is the iteration number. At the beginning, the model will take this initial position as the optimal location of the individual's history, as individual peak value.
- Second step is to decide the global peak, the output error of each neural net corresponding to each particle is calculated, and the error is taken as the fitness value of the related particle. Thus, the best location of particle swarm (global peak) in the history is the position of the particle with the best fitness value.
- Third, by using the iterative calculation formula, the velocity and position of all particles should be updated, then a new swarm is generated as **X**(k + 1).
- Forth, the fitness value of position of each particles should be recalculated and compared with its own individual peak value. If the new fitness value is better, the best location memorized in current particle's history should be replaced with the new position.
- Fifth, the individual peak value of every particle will be compared with the global one in memory. If any of the individual value is better, the individual peak value of the particle is updated to the current group peak value. In other words, the current location of the particle with highest fitness value is taken as the best historical location of the swarm.
- Finally, if the output error satisfies the design requirements, the training is terminated and the output is generated; otherwise, return to the third step.

According to the mathematical expression and building process, the PSO algorithm contains three features. One is the part of inertia or momentum, it shows the habit of particle's movements and particles have the tendency of maintaining the previous velocity. Second one is cognition, it shows particles have memory and remembrance ability of its own historical experience and particles try to move to the best location of their own memory. Final one is the part of society, it shows the historical experience of group cooperation and knowledge sharing among particles, which represents particles contain another tendency of approaching to the best position in the history of a group or neighborhood.

3. The Design for Hybrid Model:

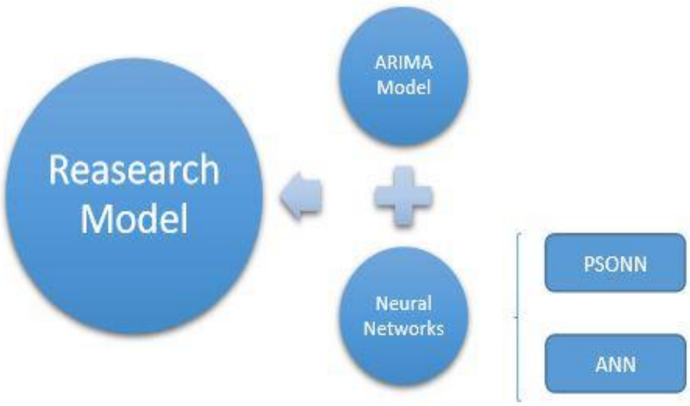


Figure 2. Structure for researching model

As the previous researches and analyses shown in the introduction, the actual time series data usually contains linear and nonlinear complex characteristics, a single ARIMA or NN model cannot depict this complex feature completely. Although, ARIMA and NN have obvious defects respectively, the defect of this model is the advantage of another model. In other words, there is explicit complementarity between ARIMA and NN models. Therefore, the fusion of the two is likely to produce better prediction results than a single model. As the literatures reviewed, many researchers had tried to combine one ARIMA model and one NN model, the prediction results of the new model were always better. However, as Markham and Rakes only one NN model cannot completely model the nonlinear part of time series data [41], so this paper tried to use a combined nonlinear fitted device composed of two NN models, PSONN and ANN to mutually promote and improve the modeling performance. Meanwhile, as the linear part of data is easy to analyze, the ARIMA model is powerful enough to model this part, so only one ARIMA model is sufficient to fit the linear characteristic [30]. So the general expression is:

$$y_t = L_t - N_t$$

where L_t means the linear component and N_t denotes the nonlinear part. These two components have to be estimated from the exchange rate data. First of all, ARIMA model is used to model the linear part, then the residuals from the linear model will have only the nonlinear feature. That is:

$$e_t = y_t - L_{ft}$$

Where e_t denotes the residual at time **t** and L_{ft} denotes the predicted value for time t from ARIMA model. After the residuals are separated out, PSONN and ANN will coordinate together to discover the nonlinear relationship.

$$e_t^{(1)} = f_1\,(e_{t-1}^{(1)},\ e_{t-2}^{(1)}, \ldots,\ e_{t-n}^{(1)}) + \varepsilon_t$$

$$e_t^{(2)} = f_2\,(e_{t-1}^{(2)},\ e_{t-2}^{(2)}, \ldots,\ e_{t-n}^{(2)}) + \varepsilon_t$$

The $e_t^{(1)}$ denotes the PSONN method and $e_t^{(2)}$ is ANN model, and f1 and f2 nonlinear functions determined by them separately. The n is the number of input nodes and ε_t is the random error. By calculating the average of the predicted values of PSONN and ANN models, the result of forecasting nonlinear component can be denoted as:

$$N_{ft} = \frac{1}{2}\ (e_t^{(1)} + e_t^{(2)})$$

So the final result for the whole model is:

$$Y_{ft} = L_{ft} + N_{ft}$$

In summary, this hybrid methodology exploits the unique characteristic and strength and ARIMA as well as NNs in determining different patterns. Therefore, it is advantageous to model linear and nonlinear patterns respectively with different models, then combine the forecasting results to improve the whole modeling and predicting performance. Besides, this paper creatively uses two NN models to learn more precisely about the nonlinear features implied in the data. It may be reasonable to promote the model effect with a more complete understand about the nonlinear peculiarity.

Experimental results and Discussion:

Technical and fundamental analyses are the two main financial forecasting methodologies. In details, the technical methodology tries to analyze time series' own cycle, trend, season and irregularity. That means technical methods only consider the data of target time series. On the contrary, in fundamental methods, both macro-economic variables and market data, from which it was assumed that the behavior of the exchange rate was conditional, will be considered. However, recently, technical methods have drawn more academic interest than fundamental ones, because of the evidences that analyzing variable markets data is less efficient than analyzing exchange rate data itself [42].

In this study, the hybrid model will use technical analyses methodology. As mentioned, first the ARIMA will be used to model the linear part of exchange rate data. After that, the residuals will be generated and be used as input variables for NNs. The number of residuals will be discussed later. Then, the non-linear part of the exchange rate data will be analyzed by NNs. In the following, the data collection, performance evaluation, construction process and simulator results will be described.

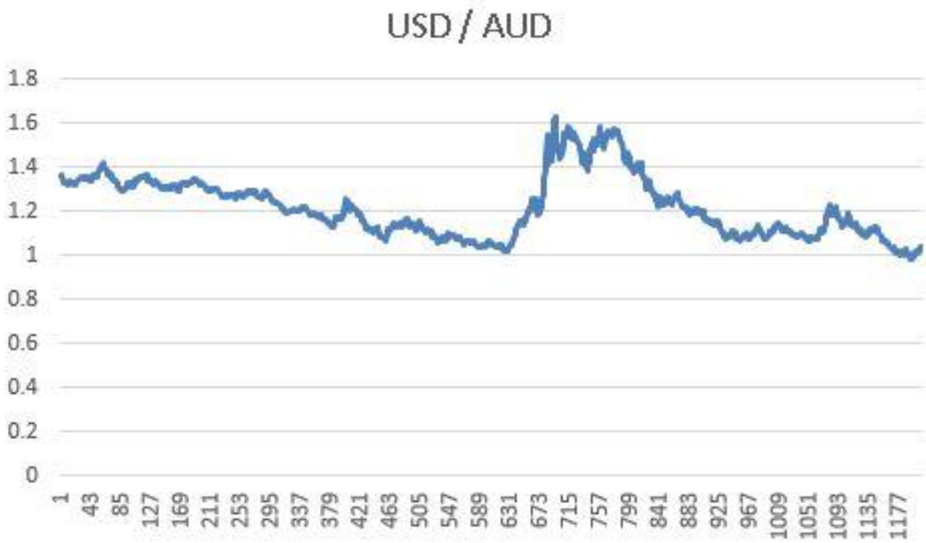
1. Data Collection:

The data used in this project is the foreign exchange rate of three different currencies against US dollar from 01/01/2006 to 30/11/2010. The British Pound (GBP), Japanese Yen (JPY) and Australian Dollar will be considered as three particular foreign currencies. This paper will mainly predict the buying prices of the exchange rates (ASK). In the given time interval, there will be 1208 daily ASK data for the model to analyze.

Figure 3. Exchange rate of USD against GBP:

Figure 4. Exchange rate of USD against JPY:

Figure 5. Exchange rate of USD against AUD:



In the sample data set, the first 1178 data are used as modeling set, the last 30 data are the set of test data, and the 30 data are divided into three forecasting periods (10, 20 and 30 days) to measure the prediction ability of the model under different forecasting periods.

Table 1.Sample compositions in the three exchange rates data:

series	Sample size	Training set	Test set
Exchange Rates	1208	1178	30

2. Model Implementation:

The prediction models of the three exchange rates are similar in construction process, due to the limited space of the paper, this part only takes USD to GBP exchange rate as an example to build and analyze the model. The explicit steps are shown as below:

2.1. ARIMA Model Implementation:

First, the ARIMA model should be constructed to predict the linear part of data composition and get the forecasting result L_{ft}. to the beginning, the exchange rates series should be changed to logarithmic yield sequence $R_t = \ln P_t - \ln P_{t-1}$. That means this research dose not study GBP sequence directly, but logarithmic yield sequence RGBP is the studying object. In fact, the process of RGBP is to take logarithms of sequence GBP first, and then do first order difference. In expression is:

$$RGBP_t = \ln(GBP_t) - \ln(GBP_{t-1})$$

The purpose for this step is to eliminate original sequence trend and reduce sequence fluctuation in modeling. Then the ARIMA model is ready to be constructed. The Matlab software is used as the platform. First, the input data series RGBP should be proven to be stationary. So the ADF detection results will be introduced as test standards.

Table 2. ADF Detection Results for Series RGBP:

Tested Type	Significant Level	Detection Results	Critical Value	Conclusion
Equation with Constant	1%	-32.56348	-3.568426	Stable
	5%	-32.56348	-2.874632	Stable
	10%	-32.56348	-2.569841	Stable

The above results show that the RGBP data is stationary series (the parameter d for model is 1) and is proper to be used as input to ARIMA model. By observing the sequence autocorrelation coefficient and partial autocorrelation coefficient, the initial values for parameter p and q could be： $p = 0 \sim 1$ and $q = 2 \sim 3$. However, according to AIC minimum criterion, the final model is determined as ARIMA(0, 1, 3). With previous empirical literatures, the estimated equation for RGBP is

$$RGBP = -0.00017 + v_t + 0.08036\, v_{t-3}$$

Then at the stage of diagnostic check, the error sequence was tested by ADF:

Table 3. ADF Detection Results for error sequences:

Tested Type	Significant Level	Detection Results	Critical Value	Conclusion
Equation without Constant and Trend Item	1%	-4.63217	-2.568123	Stable
	5%	-4.63217	-1.862317	Stable
	10%	-4.63217	-1.535449	Stable

Obviously, error sequence was stationary random sequence, which means the established ARIMA model was appropriate. Besides, because of $RGBP_t = \ln(GBP_t) - \ln(GBP_{t-1})$, the forecasting model for USD exchange rate can be expressed as:

$$\ln(GBP_t) = \ln(GBP_{t-1}) - 0.00017 + v_t + 0.08036\, v_{t-3}$$

By doing anti-logarithm transformation to the above equation, the estimated result for GBP exchange rate can be got easily and such result will be the forecasting L_{ft} for ARIMA model.

2.2. NN Model Implementation:

The second step is to use neural networks to model the nonlinear part of exchange rate data in order to get another forecasting results N_{ft}. This paper used two networks ANN and PSONN to model the data respectively. In the end, the two predicting results from the two networks were combined together and the mean value calculated from the combined result was determined as the final N_{ft} result.

All the neural networks were implemented on Matlab platform. As mentioned, all the networks contained only one hidden layer, the benefits and feasibilities had been proven earlier and such model had also been used by many other researchers, for example, Rao and Gabr [43] and McLeod and Hipel[44]. To implement the networks, currently the biggest challenge was to decide the number of nodes of hidden layer.

- For ANN, the neural model used is a 4 * 4 * 1 network which has been also employed by Groot and Wurtz [45] and Cottrell et al [46]. The network has four inputs e_{t-1}, e_{t-2}, e_{t-3}, e_{t-4} and one output e_t, the hidden layer contains four nodes. In this research, several network structures have been tried and tested and the 4 * 4 * 1 network was proven to be better than others. The project used "net = newff (PR, [S1 S2 ...SN], {TF1 TF2...TFN}, BTF, BLF, PF);" to build the network.

  PR: An R*2 matrix to define the min and max values for R inputs.

  Si: The number of neurons for i-th layer. **TFi:** Transformer Function.

  BTF: Training Function. **BLF:** Learning Function.

  PF: Performance Function.
- For PSONN, the related literatures are less, so it is hard to find an optimal structure among existing researches. This paper used the trial and error method, the general procedure is to determine a base value according to some experience, then to improve the convergence speed and fitting ability of the network, the number of nodes would be constantly changed to try new values until the outputs of network meet the requirements. After repeated trials, the structure was determined as 4 * 5 * 1, which means there are four inputs and one output (using e_{t-1}, e_{t-2}, e_{t-3}, e_{t-4} to predict e_t) and there are five nodes in the hidden layer. The following programs correspond to PSONN model building steps.

```
R=abs(G-pop);
vel=w(j).*vel+G+(-1+rand(N,D)*2).*R;
vel=(vel>Vmax).*Vmax+(vel<=Vmax).*vel;
vel=(vel<Vmin).*Vmin+(vel>=Vmin).*vel;
```

Figure 6. Updating velocity for PSO algorithm

```
pop=vel+pop;
pop=((pop>=VRmin)&(pop<=VRmax)).*pop...
    +(pop<VRmin).*(VRmin+0.25.*(VRmax-VRmin).*rand(N,D))+(pop>VRmax)
    .*(VRmax-0.25.*(VRmax-VRmin).*rand(N,D));
vel=((pop>=Vmin)&(pop<=Vmax)).*vel+(pop<Vmin).*0+(pop>Vmax).*0;
```

Figure 7. Updating position for PSO algorithm

```
t=(e<pbestval);
m=repmat(t',1,D);
pbest=m.*pop+(1-m).*pbest;
pbestval=t.*e+(1-t).*pbestval;
```

Figure 8. Updating global peak value.

```
for i=1:N;
 a7=a3(i,:);
[lbestval(i),m1]=min(pbestval(a3(i,:)));
a8=a7(m1);
lbest(i,:)=pbest(a8,:);
end
```

Figure 9. Updating local peak value.

2.3. Hybrid Method Implementation:

As the equation $Y_{ft} = L_{ft} + N_{ft}$ shows, the ARIMA model and NN model are combined together in this stage. Then Y_{ft} is the forecasting result for the hybrid model. The forecasting procedures for JPY and AUD are same, so the details will not be introduced in the paper.

Data Set	Sample size	Test Size	PSONN	ANN
USD/GBP	1178	30	4 * 5 * 1	4 * 4 * 1
USD/AUD	1178	30	4 * 5 * 1	4 * 4 * 1
USD/JPY	1178	30	4 * 5 * 1	4 * 4 * 1

3. Simulation Results:

As contrasts, this study has also constructed ARIMA, PSONN and ANN four single methodology model and two hybrid models combined of ARIMA &PSONN and ARIMA & ANN. By comparing them with the new model, the performance of this research's new model will be shown clearly.

3.1. Performance Metrics:

The forecasting performance of the mentioned models are evaluated against several widely used statistical metrics, In order to comprehensively describe the effects of different models to exchange rate prediction. Three criteria are used in the study, namely, Normalized Mean Square Error (NMSE), Mean Absolute Error (MAE) and Directional Symmetry (DS). NMSE and MAE measure the deviation between the estimated value and actual value. DS measures correctness in forecasting directions. The mathematical expressions are shown below:

$$NMSE = \frac{\sum_k (x_k - \hat{x}_k)^2}{\sum_k (x_k - \bar{x}_k)^2} = \frac{1}{\sigma^2 N} \sum_k (x_k - \hat{x}_k)^2$$

$$MAE = \frac{1}{N} |x_k - \hat{x}_k|$$

$$DS = \frac{100}{N} \sum_k d_k, \quad d_k = \begin{cases} 1 & \text{if } (x_k - x_{k-1})(\hat{x}_k - \hat{x}_{k-1}) \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

In above, the (x_1, x_2, …, x_k) are the actual values and ($\widehat{x1}$, $\widehat{x2}$, …, $\widehat{xk}$) are the forecasting values generated from the model. N is the sample number. To evaluate, the smaller the value of NMSE or MAE, the less the output error. In other words, the precision of model prediction is higher. In the contrary, the higher the DS value, the forecasting direction is more correct.

3.2. Results:

This paper used the constructed models to predict the three exchange rate series (DBP, AUD and JPY). Three forecast horizons of 10, 20 and 30 periods are chose as out-of-sample data within the last 30 trading days.

Table 4. Comparison of the prediction effect of each model to JPY sequence:

Model Type	10 Days			20 Days			30 Days		
	RMSE	**MAE**	**DS**	**RMSE**	**MAE**	**DS**	**RMSE**	**MAE**	**DS**
ARIMA	0.0381	0.0340	31.75	0.0375	0.0329	40.00	0.0349	0.0309	40.23
PSONN	0.0255	0.0237	60.00	0.0243	0.0227	55.15	0.0241	0.0223	53.33
ANN	0.0372	0.0330	60.00	0.0383	0.0352	50.46	0.0378	0.0351	46.67
ARIMA-PSONN	0.0095	0.0071	79.65	0.0091	0.0069	75.00	0.0089	0.0064	73.33
ARIMA-ANN	0.0165	0.0148	66.00	0.0156	0.0123	60.71	0.0103	0.0086	55.33
ARIMA-PSONN-ANN	0.0074	0.0069	85.23	0.0070	0.0057	83.33	0.0054	0.0037	80.17

Figure 10. Hybrid Prediction of exchange rate of USD/JPY:

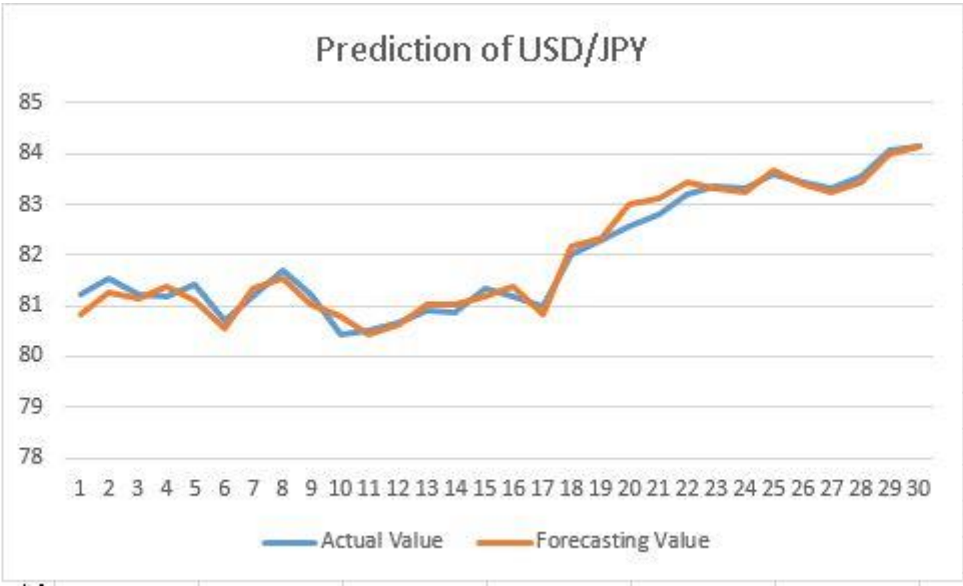


Table 5. Comparison of the prediction effect of each model to GBP sequence:

Model	10 Days	20 Days	30 Days

Type	RMSE	MAE	DS	RMSE	MAE	DS	RMSE	MAE	DS
ARIMA	0.0621	0.0452	50.46	0.1066	0.0783	50.46	0.2089	0.1584	46.67
PSONN	0.0431	0.0399	70.00	0.0695	0.0620	65.00	0.0928	0.0856	63.33
ANN	0.0524	0.0464	60.00	0.1273	0.0958	55.29	0.2227	0.1672	56.67
ARIMA-PSONN	0.0163	0.0107	80.00	0.0194	0.0069	80.00	0.0263	0.0249	76.67
ARIMA-ANN	0.0323	0.0274	71.45	0.0526	0.0486	70.00	0.0758	0.0710	65.69
ARIMA-PSONN-ANN	0.0153	0.0114	86.71	0.0189	0.0177	85.10	0.0206	0.0175	84.76

Figure 11. Hybrid Prediction of exchange rate of USD/GBP:

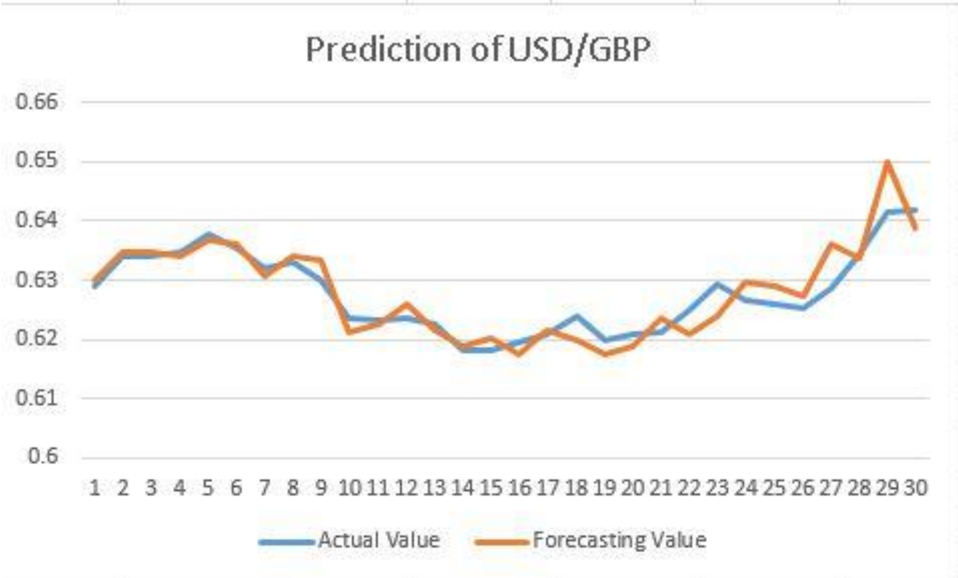
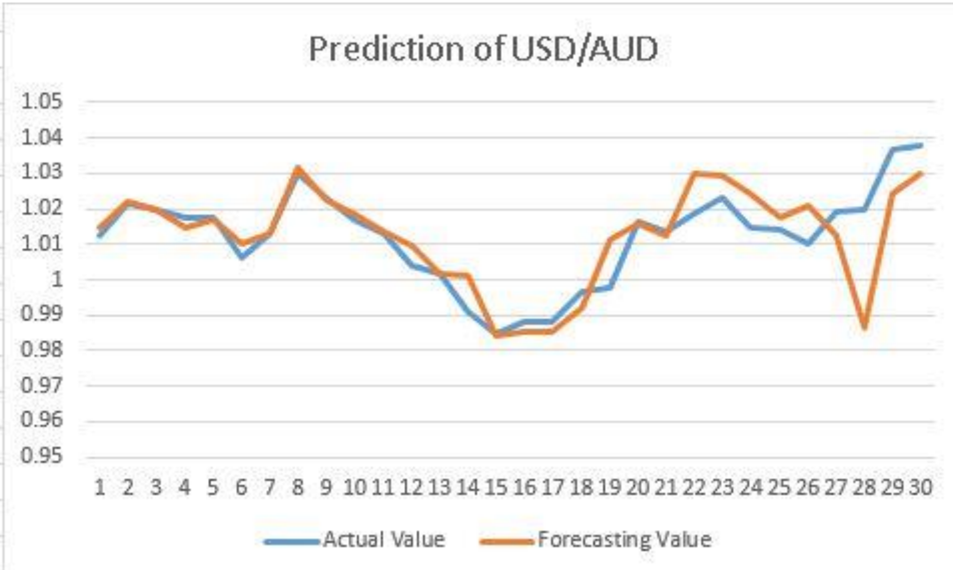


Table 6. Comparison of the prediction effect of each model to AUD sequence:

Model Type	10 Days			20 Days			30 Days		
	RMSE	MAE	DS	RMSE	MAE	DS	RMSE	MAE	DS
ARIMA	0.0643	0.0605	41.11	0.0822	0.0730	45.00	0.1233	0.1070	43.33
PSONN	0.0422	0.0383	70.00	0.0496	0.0445	65.00	0.0667	0.0601	63.33
ANN	0.0635	0.0479	70.00	0.0923	0.0858	60.57	0.1367	0.1168	52.75

ARIMA-PSONN	0.0157	0.0105	89.52	0.0162	0.0115	80.47	0.0214	0.0185	80.00
ARIMA-ANN	0.0385	0.0319	75.38	0.0423	0.0388	67.14	0.0886	0.0804	64.30
ARIMA-PSONN-ANN	0.0131	0.0107	92.17	0.0147	0.0101	88.52	0.0193	0.0177	85.00

Figure 12. Hybrid Prediction of exchange rate of USD/AUD:



4. Evaluation:

According to the table 5, in terms of forecasting GBP exchange rate sequence, different model has dramatically different performance. It is clear from the results that the hybrid model (ARIMA – PSONN - ANN) proposed from this paper has the best performance in terms of RMSE. The RMSE values under three different prediction periods are 0.0153, 0.018 and 0.0206, respectively. The hybrid model of ARIMA and PSONN contains the best performance among the remaining five models. The RMSE values are 0.0163, 0.0194 and 0.0263. Then next is hybrid method of ARIMA and ANN, however, the RMSE values are dramatically larger than the previous two models over the three periods. The single ARIMA model contains the worst performance compared with other models. Besides, the single PSONN model and single ANN model perform a little better than ARIMA model but still at a low level of accuracy. Such situation may show that neither the neural network nor the ARIMA model fully captures the whole patterns in the exchange rate data. It also proves the opinion proposed in the beginning of the paper that single neural network model or ARIMA model cannot fully learn the characteristics implied in time series data. It is reasonable to use hybrid method to improve forecasting performance since the three combined models can reduce significantly the overall forecasting errors. Furthermore, within the two neural networks, the PSONN model performs much better than ANN. This may suggest that the ANN method is easily to get trapped in local optima and leads to poor generalization ability. PSONN can avoid such problem to a certain extent. In three hybrid models, the results show that the performance of ARIMA-PSONN model is extremely close to the proposed model. This may suggest that the PSONN method is powerful enough to model the nonlinear part of the data. However, by combining PSONN and ANN together, certain defects existed in PSONN method can be avoided or weakened.

In terms of MAE values, the conclusions are same with RMSE. The ARIMA-PSONN-ANN model has the best performance, next is ARIMA-PSONN model and the ARIMA model still contains the highest forecasting errors. Hybrid models are still more advanced than single ones. In terms of direction correctness DS, the proposed model gets the highest values, next is ARIMA-PSONN model and then the ARIMA-ANN model. On the contrary, the DS values for ARIMA model are just around 50, which means the forecasting direction is not good. While the neural networks get values between 60 to 80. Such situation may show that the nonlinear feature takes an important role in the exchange rate data. That may also explain the reason that solving time series prediction problems with neural network methods got more attention in recent years. From the whole table, the trend for RMSE and MAE value is becoming higher, and for DS, the value is becoming lower. Such condition shows that the forecasting errors are increasing and the prediction precision is reducing. In conclusion, current models are suitable for short-time prediction. When the period is too large, the model cannot guarantee high precision.

In the table 4 and table 6, by comparing the effects of different models, the observed results accord to the above conclusion that the proposed hybrid model performs better than others. For GBP and AUD, the prediction performance of each model varies considerably within three different periods, and the prediction accuracy decreases while the length of prediction period increases. But from the table 1, it can be observed that the model has no significant change in the predictive performance of the JPY sequence within three terms, which is different from the other two exchange rate sequences. This phenomena shows that the fluctuation of the exchange rate of USD against JPY is relatively mild, while the exchange rates of AUD and GBP have more violent fluctuations. In other words, such volatility indicates complex features existing in the data. Furthermore, these conclusions could be also observed from figure 3 to 5. Obviously, the figure of JPY contains a smooth curve while the others have large fluctuations and the curves are more complicated.

Through the comparisons and analyses of the forecasting performance of the three exchange rate series, the optimal model is always the proposed model with a combination of ARIMA, PSONN and ANN. The new hybrid model is a tried and tested method and is reasonable for further study. Within these models, hybrid models always perform better than single methodology model. It indicates that hybrid models can truly integrate the advantages of the single methodology model and deeply explore the elusive linear and nonlinear characteristics implied in the USD exchange rate sequences. Such combination dramatically improve the performance on exchange rate prediction. Within the hybrid models, the model with only one neural network to capture nonlinear features, generates more forecasting errors than the model constructed by two different neural networks. In other words, the two networks complement each other and the model can study the nonlinear features more comprehensively.

Summary and Reflection:

1. Contributions and Reflections:

Exchange rates analysis and prediction has been an active academic field over the last few decades. The accuracy of exchange rate forecasting is fundamental and primary to many decision processes and financial analyses and hence the study for promoting the effectiveness of forecasting models will never stop and has generated amazing achievements. Beginning with the efforts of Box and Jenkins[30] the ARIMA has become one of the most popular methods in the prediction area. The ARIMA model has also been the benchmark for new proposed forecasting methods. More recently, neural networks have shown their good performance in exchange rate forecasting practice with their nonlinear modeling capability. As Medeiros et al [47] argues that favor linear or nonlinear models against random walk, and nonlinear models have a better chance when nonlinear capability is spread in time series. It is no doubt that the two models have their own unique advantages in forecasting exchange rates. Both of them have the flexibility in modeling a variety of applications, however, none of the two methods is the overall best model that can be used indiscriminately in every forecasting problems. Instead, the two models have strong complementarity. One's advantages are exactly what another model lacks.

1.1. Contributions:

In this paper, a combining approach is proposed to exchange rate forecasting. The linear ARIMA model and the nonlinear neural network model are used jointly, aiming to analyze different forms of relationships in the exchange rate data. As shown in the research, the nonlinear relationships are more difficult to be modeled. So it deserves more efforts to fully understand the nonlinear relationships. In this research, the forecasting processes contain 3 steps. First, the ARIMA model is used to model the linear part of exchange rate data. Then, PSONN and ANN are combined together to estimate the nonlinear residuals. Finally, the addition of the forward two steps is the result of the prediction. For complex problems just like exchange rate prediction, which have both linear and nonlinear correlation structures, the hybrid method can be a powerful way to improve the forecasting performance.

Kinds of combining methods haven been proposed during the development of forecasting techniques. However, in the early time, most of the combining methods are designed to combine similar models. For example, several linear models are often used together in the traditional forecasting literature. While in neural network researches, the combination always includes several neural networks and such combinations are called as neural committee or neural network ensemble. These combining methods truly improve the performance but the improvement is limited and unstable [4]. As Zhang[8] argues that a more effective way of combination should be based on quit different methods. Recently, such kind of method has become a rising concern. However, most of the current researches just simply combine one linear model with one nonlinear model. In the comparisons of this paper, we found that the nonlinear relationships in exchange rates are more complex. This has also been proven by Kamruzzaman[5], Tang and Fishwich[48] and Kamruzzaman and R.Sarker[22]. The new hybrid model act according to actual circumstances of exchange rate data and creatively use two neural networks to model the nonlinear part. The researching data and following discuss can show the idea is correct and the model is effective.

1.2. Achievements and Conclusions:

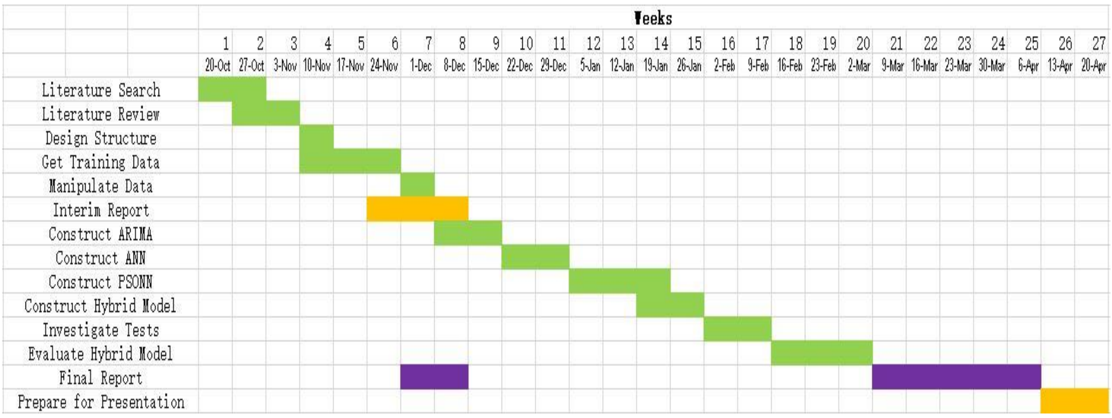
In the paper, the implemented model is used to forecast three different exchange rate sequences, namely, GBP, AUD and JPY against USD. Another five forecasting models are also implemented to compare with the proposed model with three evaluation indicators, RMSE, MAE and DS. The researching data are shown in table 4 to 6. Several conclusions can be detected from these results:

First, the hybrid models perform obviously better than single methodology models in USD exchange rate prediction. It explains the USD sequences contain both linear and nonlinear characteristics and the hybrid models are much more effective to capture such complex relationships. As a result, the forecasting performance is greatly improved. Additionally, due to the possible unstable or various patterns in the data, using hybrid method can reduce the model uncertainty, which especially occurred in statistical inference and time series forecasting. Besides, by fitting the ARIMA model first to exchange rate data, the overfitting problem that is more commonly related to neural networks can be relieved.
Second, the research also shows that the nonlinear relations existing in USD exchange rate data are harder to analyze, comparing with linear relations. Just like ARIMA model and NN model complement each other, this research uses both PSONN and ANN to coordinate with each other. The forecasting performance is also improved comparing with the model with only one neural network. To a certain extent, the neural network trained by PSO algorithm overcomes the ANN's shortcoming of poor generalization. At the same time, the ANN also helps to slow down the converge rate to avoid too fast fitting of PSO algorithm.

2. Further Researches:

The current research model is based on the time series data of exchange rate itself. It is suitable for short-term forecasting. For long-term prediction problems, On the basis of this study, some external factors including macro economy and international environment should be considered. How to improve the model by introducing these external factors is another subject to be studied further. In addition, the targets of future exchange rate prediction researches are about to explore the complementarity of different forecasting techniques, to study the combining models of different methods, and to improve the accuracy and reliability of the exchange rate prediction model.

3. Project Management Covering:



The project plan is designed in a linear style. All the major steps are listed on the table and the cost of time at right side. This schedule table is a little different with previous one in the proposal. Some details are added to the table and several time plans are changed. The reason for doing these is that I want to make sure the difficult parts will have enough time to be finished. Besides, learning from current achievements, some details of the project that had not been considered carefully will be took into account.
The arrangement of the project is reasonable and the progress follows the plan generally. Some

changes have been made to the whole plan. So there are some differences with the interim report and the initial plan. Initially, the project planned to only build an artificial neural network model to analyze two different methods, technical analyses and fundamental analyses. However, as the project went deeper, I found that the ANN model had been fully studied and analyzed. So the project would lack of novelty and creativity. Besides, the technical analyses suit for short-term forecasting while fundamental analyses suit for long-term forecasting. The comparability of the two methods was narrow. Therefore, after reviewing lots of literatures, I decided to build a new combining model. So the plan was changed to fit the new model. Fortunately, some preliminary work were quite useful and primary. Although, the plan was changed, it did not influence the whole work dramatically.

The purpose of the project is attractive and the efforts taken is strong enough to support the final result. This project stands at a good place at this critical halfway point, with a large amount of work already completed, including most of the work in the original brief. I also got lots of useful knowledge form those experience. They are good for my further study about neural network and machine learning. With more features and requirements, the models designed in the project could be engaged to software and any other services.

References:

[1] J.Kuepper. Globalization and Its Impact on Economic Growth. Available at: https://www.thebalance.com/globalization-and-its-impact-on-economic-growth-1978843.

[2] D.Dimitrova. The Relationship between Exchange Rates and Stock Prices: Studied in a Multivariate Model. Issues in Political Economy, Vol. 14, August 2005.

[3] T.Pettinger. Effect of the exchange rate on business. Available at: https://www.economicshelp.org/blog/9328/business/effect-exchange-rate-business/

[4] G.ZHANG. A neural network ensemble method with jittered training data for time series forecasting [J].Information Sciences, 2 007, 177(23):53295346.

[5] J.Kamruzzaman, ANN-Based Forecasting of Foreign Currency Exchange Rates. Neural Information Processing [J], Vol. 3, No. 2, May 2004.

[6] B.Liu and Z.Zhao. STAR Based Forecasting Nonlinear situation of real exchange rate in China [J]. Research on quantitative economy, technology and economy. Vol.6, 2008.

[7] Maia ALS, Carvalho FAT , Ludermir TB Forecasting models for interval valued time series [J] , Neurocomputing , 2008 , (71) : 3344～33521.

[8] G.P.Zhang. Time series forecasting using a hybrid ARIMA and neural network model [J], Ueurocomputing, 2003, (50): 159 – 175.

[9] C.Brooks, Testing for nonlinearity in daily sterling exchange rates [J] , Applied Finance Econ omics , 1996 , (6) : 307～317.

[10] C.Xie, N.Yang, B.Sun, Analysis and Empirical Study on the nonlinear characteristics of exchange rate time series [J], Systems Engineering, 2008, (10).

[11] H.B.Hwarng and H.T.Ang, A simple neural network for ARMA (p, q) time series, OMEGA: Int. Journal of Management Science [J], vol. 29, pp 319-333, 2002.

[12] J.Fan and Q.Yao, Nonlinear time series --- Modeling, prediction and Application [M], Higher

Education Press, 2005.
[13] C.Panda and V.Narasimhan, Forecasting daily foreign exchange rate in India with artificial neural network [J] , Singapore Economic Review , 2003 , Vol148 (2) : 181～1991.
[14] Y.Sun, Y.Liang, J.Jiang and C.Wu, Neural Network methods in forecasting financial time series [J], Journal of Jilin University, 2004, (1).
[15] Y.Sun, Exchange rate forecasting with an artificial neural network model: can we beat a random walk model? [D]New Zealand: Christchurch, Lincoln University, 2005.
[16] J.Plasmans, W.Verkonijen, H.Daniel, Estimating structural exchange rate models by artificial neural networks [J], Applied Financial Economics, 1998, (8): 541～5511.
[17] L.Anastasakis, N.Mort, Exchange rate forecasting using a combined parametric and non-parametric self-organizing modelling approach [J] , Expert Systems with Applications , 2009 , (36).
[18] L.Kilian, and M.P.Taylor, Why is it so Difficult to Beat the Random Walk Forecast of Exchange Rates? [J] , Journal of International Economics , 2003 , Vol160 (1) : 85～107.
[19] J.Wang and W.Li, Whether GARCH model can provide good prediction of floating rate? [J], Research on quantitative economy, technology and economy, 2005, (6).
[20] J.Hong, X.Li and B.Wang, Exchange Rate Forecast Based on Artificial Neural Network and Random Walk Mode [J]l, JOURNAL OF QUANTITATIVE ECONOMICS, 2016, vol.33, (1).
[21] W.C.Jhee and J. K.Lee, Performance of neural networks in managerial forecasting, Intelligent Systems in Accounting [J], Finance and Management, vol. 2, pp 55-71, 1993.
[22] J.Kamruzzaman and R.Sarker, Comparing ANN based models with ARIMA for prediction of forex rates [J], ASOR Bulletin, vol. 22, pp. 2-11, 2003.
[23] J.H.Wang and J.Y.Leu, Stock market trend prediction using ARIMA-based neural networks [J], Proc. of IEEE Int. Conf. on Neural Networks, vol. 4, pp. 2160-2165, 1996.
[24] H.D.K.Wedding and K.J.Cios, Time series forecasting by combining RBF networks, certainty factors and the Box-Jenkin model [J], Neurocomputing, 1996, (10), 149 – 168.
[25] F.M.Tseng, H.C.Yu and G.H.Tzeng, Combining neural network model with seasonal time series ARIMA models [J], Technological forecasting and social change, 2002, (69), 1 – 87.
[26] H.Michele, E.Theodoros, To combine or not to combine: selecting among forecasts and their combinations [J], International Journal of Forecasting, 2005 , (21) : 15 – 241.
[27] M.Khashei, S.R.Hejazi and M.Bijari, A New Hybrid Artificial Neural Networks and Fuzzy Regression Model for Time Series Forecasting [J], Fuzzy Sets and Systems, April 2008, Vol. 159, (7), pp. 769-786.
[28] W.K.Wong, M.Xia and W.C.Chu, Adaptive Neural Network Model for Time-Series Forecasting [J], European Journal of Operational Research, Vol. 207, No. 2, December 2010, pp. 807-816.
[29] M.Khashei, M.Bijari, A.R.Ardali, Improvement of autoregressive integrated moving average models using fuzzy logic and artificial neural network [J], Neurocomputing, 2009, (72), 956 – 967.
[30] G.E.P.Box and G.M.Jenkins, Time Series Analysis: Forecasting and Control, Holden-Day, San Francosco, CA, 1990.
[31] J.G.D Gooijer and K.Kumar, Some recent developments in non-linear time series modeling, testing and forecasting [J], Int. J. Forecasting, 1992, (8), 135 – 156.
[32] L.Cao and F.Tay, Financial forecasting using support vector machines [J], Neural Comput. & Applic, vol. 10, 2001, 184 - 192.
[33] J.R.Coakley, C.E.Brown, Artificial neural networks in accounting and finance: Modelling issues [J], International Journal of Intelligent Systems in Accounting, 2000 , Vol19, (2), 144 – 199.

[34] C.F.Juang, A hybrid of genetic algorithm and particle swarm optimization for recurrent network design [J], IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), Vol.4, (2), 2004, 997 – 1006.
[35] J.R.Zhang, J.Zhang, T.M.Lok and M.R.Lyu, A hybrid particle swarm optimization–back-propagation algorithm for feedforward neural network training [J], Applied Mathematics and Computation, Vol.185, (2), 2007, 1026 – 1037.
[36] J.Yao and C.L.Tan, A case study on using neural networks to perform technical forecasting of forex [J], Neurocomputing, vol.34, 2000, 79 – 98.
[37] G.Zhang, E.B.Patuwo and M.Y.Hu, Forecasting with artificial neural networks: the state of the art [J], Int.J. Forecasting, (14), 1998, 35 – 62.
[38] K.Hornik, M.Stinchcombe, H.White, Multilayer feedforward networks are universal approximators [J], Neural Networks, (2), 1989, 359～366.
[39] M.T.Hagan, H.B.Demuth and M.H.Beale, Neural Network Design, PWS Publishing, Boston, MA, 1996.
[40] A.F.Moller, A scaled conjugate gradient algorithm for fast supervised learning [J], Neural Networks, vol.6, 1993, 525 - 533.
[41] I.S.Markham and T.R.Rakes, The effect of sample size and variability of data on the comparative performance of artificial neural networks and regression [J], Comput. Oper. Res, (25), 1998, 251–263.
[42] B. LeBaron, Technical trading rule profitability and foreign exchange intervention [J], Journal of Int. Economics, vol. 49, pp. 124-143, 1999.
[43] T. Subba Rao, M.M. Sabr, An Introduction to Bispectral Analysis and Bilinear Time Series Models [J], Lecture Notes in Statistics, Vol. 24, Springer-Verlag, New York, 1984.
[44] K.W. Hipel, A.I. McLeod, Time Series Modelling of Water Resources and Environmental Systems [J], Amsterdam, Elsevier, 1994.
[45] C. D. Groot, D. Wurtz, Analysis of univariate time series with connectionist nets: a case study of two classical examples, Neurocomputing 3 (1991) 177–192.
[46] M. Cottrell, B. Girard, Y. Girard, M. Mangeas, C. Muller, Neural modeling for time series: a statistical stepwise method for weight elimination [J], IEEE Trans. Neural Networks 6 (1995) 1355–1364.
[47] M.C. Medeiros, A. Veiga and C.E. Pedreira, Modeling exchange rates: smooth transitions, neural network and linear models [J], IEEE Trans. Neural Networks, vol. 12, no. 4, 2001.
[48] Z. Tang and P. A. Fishwich, Backpropagation neural nets as models for time series forecasting [J], ORSA Journal on Computing, vol. 5, no. 4, pp 374-385, 1993.