Forecasting Tala-AUD and Tala-USD Exchange Rates with ANN

Shamsuddin Ahmed, M. G. M. Khan, Biman Prasad, and Avlin Prasad

Abstract—The focus of this paper is to construct daily time series exchange rate forecast models of Samoan Tala/USD and Tala/AUD during the year 2008 to 2012 with neural network. The performance of the models was measured by using various error functions such as Root Square Mean error (RSME), Mean absolute error (MAE), and Mean absolute percentage error (MAPE). Our empirical findings suggest that AR (1) model is an effective tool to forecast the Tala/USD and Tala/AUD.

Keywords—Neural Network Forecasting Model, Autoregressive time series, Exchange rate, Tala/AUD, winters model.

I. INTRODUCTION

Samoa is located half way between Hawaii and New Zealand in the South Pacific region with the geographic coordinates of 13° 35’ South and 172° 20’ west. It has a tropical climate. The total landmass area of Samoa is 2831 square kilometers with two main islands (Savaii and Upolu), and several smaller islands and uninhabited islets. Its major city is Apia. Its natural resources are forest, fish, and hydropower. Samoa gained independence on 1 January 1962 from New Zealand which was administered UN trusteeship.

The estimated population of Samoa in July 2012 is 194,320. This includes Samoan, Euronesian (persons of European and Polynesian blood), Europeans and other ethnicity. Their official language is Samoan (Polynesian) and English. In 2012 the estimated data shows that the Samoa’s population growth rate is 0.596%, birth rate is 22.1births/1,000, death rate is 5.34deaths/1,000 and Net migration is 10.81 migrant(s)/1,000.

Samoan exports. Its indigenous exports consist mainly of fish and agriculture products. A large proportion of the population works in subsistence agriculture or low-level commercial ventures. The gross domestic product (GDP) or value of all final goods and services produced within a nation for the 2nd quarter of 2012 shows 2.7%. Daily exchange rate is fixed by the central bank of Samoa in relation with a weighted basket of currencies, which includes United States of America dollar, New Zealand dollar, Australia dollar and the Euro. We compare autoregressive ANN model, and moving average to forecast daily exchange rate of Samoan Tala against AUD and USD.

II. RELATED PREVIOUS STUDIES

There are many ways to forecast exchange rate by using different models. The author in [1] compared the ability of different mathematical models, such as artificial neural networks (ANN) and ARCH and GARCH models, to forecast the daily exchange rates Euro/U.S. dollar (USD). The researcher used time series data of Euro/USD from December 31, 2008 until December 31, 2009. The anticipated to ANN developed to predict the trend of the exchange rate Euro/ USD up to three days ahead of last data available. He concluded ARCH and GARCH models, especially in their static formulations, are better than the ANN for analyzing and forecasting the dynamics of the exchange rates. According to the study, the suggested ARCH (2) model with a static approach showed the best predictive ability.

The study in [2] showed the exchange rate between Euro and USD using univariate model (autoregressive integrated moving average (ARIMA) and exponential smoothing method). He took 3202 observation of exchange rate between Euro and USD ranging from 4 January 1999 to 1 July 2011. He concluded that there was a presence of unit root in exchange rate between Euro and USD and it failed to give non-trivial confidence interval for forecasting of the exchange rate. When they differed the exchange rate between Euro and USD time series, this resulted in white noise, which could not be submitted to ARIMA or exponential smoothing method. However, the first difference followed Laplace distribution, and this distribution appeared different between two
independent variables each followed exponential smoothing distribution.

Many other authors such as studies in [3]-[10] also designed and tested the ANN model to predict exchange rate. The variable output designed in their techniques was either monthly or daily exchange rate. All these studies concluded that the ANN model is the better model to predict the exchange rate.

The study by done by authors in [11]-[13] investigated hybrid artificial intelligence method based on neural network and genetic algorithm for modeling daily foreign exchange rate. They used time series data of deutsche mark/USD, YEN/USD, USD/GBP. The targeted rates were the closing exchange rate at the end of the next day. They concluded in their study that out of sample genetic algorithm is better than ANN model and a statistical time series modeling approach.

Recently, study by author [14] predicted the monthly average exchange rates of Bangladesh using ANN and ARIMA models. A feed forward multilayer neural network namely, exchange rate neural network, has been developed and trained using back propagation learning algorithm. The effect of different network and tuning parameters was examined during training session. The ARIMA model was executed using Box-Jenkins methodology and obtained the appropriate model. The results showed that the ANN model has better predictability than the ARIMA model.

III. TIME SERIES DATA

The study uses the daily real exchange rate of Samoan Tala against AUD from 3 January 2008 to 28 September 2012. This excludes the weekends and public holiday in Samoa. There are 1180 observations. The figure below shows the trend of the Tala/AUD from 3 January 2008 to 28 September 2012. The Fig. 1 clearly shows that the Tala/AUD increased 21 October 2008 until 3 November 2008 and later Tala/AUD decreases. Thereafter, Tala/AUD maintained its fluctuating trend.

The average exchange rate of Tala against AUD over the study period is 0.44 with the standard deviation of 0.03. The table also reveals that the distribution is highly skewed to the right (skewness =1.30) and highly peaked (Kurtosis =4.00). This might imply that it is not advisable to use mean exchange rate value for business transactions. The coefficient of variance is 6.14%. Therefore, the Tala/AUD exchange rate is not highly volatile. The examples of exchange rate volatility are present in [15].

The preliminary test rejected stationary in AUD series. We carry out ADF test to check for the presence of unit root. The test suggested that there is a unit root present in AUD series with lag 1. To test the presence of unit root we consider (1). For this purpose, we considered the model with constant, intercept and trend.

In attempt to check whether the data is at least stationary [16]-[18] the data is analyzed yearly. The ADF test uses 22 lags of the variables to check for unit root for Tala/AUD. The test shows that the Null hypothesis (Null Hypothesis: AUD has a unit root) cannot be rejected. Thus, this states that the variables are non-stationary.
where \( Y_t = \text{AUD with time}; \ a_0 = \text{constant}; \ Y_{t-1} = \text{lag of AUD}, \ b_1, \ b_2 \) are the regression coefficients, \( \epsilon \) is the error term and \( n \) is the number of daily observations. The results are listed in table.

### TABLE I

**REGRESSION RESULT FOR TALA/AUD WITH LAG 1**

<table>
<thead>
<tr>
<th>Dependent Variable: AUD</th>
<th>Method: Least Squares</th>
<th>Included observations: 1179 after adjustments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>Std. Error</td>
<td>t-Statistic</td>
</tr>
<tr>
<td>C</td>
<td>0.004280</td>
<td>0.001815</td>
</tr>
<tr>
<td>TIME</td>
<td>-4.55E-07</td>
<td>3.07E-07</td>
</tr>
<tr>
<td>AUD(O1)</td>
<td>0.990837</td>
<td>0.003877</td>
</tr>
</tbody>
</table>

**R-squared** 0.987783  
**Adjusted R-squared** 0.987762  
**S.E. of regression** 0.002979  
**Sum squared resid** 0.010434

The regression result showed that the parameters of the model (constant and AUD (-1)) are significant (P values are less than 0.05) and R squared 0.994887. The Correlogram of forecasted values of Tala against AUD shows spikes at the 1st lag and slight spike at the 2nd lag. The Q-statistics are significant at all lags, indicating significant serial correlation in the residuals as shown in the corresponding figure.

### Fig. 4 Residual analysis from the regression of Tala/AUD

Due to non-stationary of the Tala/AUD daily exchange rate, the daily rate falls below the upper bound of the 95% confidence interval.

### Fig. 6 The 95% confidence interval of forecasted AUD

Following similar analysis, we briefly present the characteristics of the Tala/USD daily exchange rate behavior. The box plot observation and ADF test with 22 lags does not eliminate the probability of the existence of unit root. The test shows that the Null hypothesis (Null Hypothesis: Tala/USD exchange rate has a unit root) cannot be rejected. Hence, the series is non-stationary. The residual analysis of Tala/USD suggests that the model is adequate.

### Fig. 7 Box Plot of Tala/USD (-1) by Year
TABLE III
REGRESSION RESULT FOR TALA-USD

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.002366</td>
<td>0.001129</td>
<td>2.095055</td>
<td>0.0364</td>
</tr>
<tr>
<td>TIME</td>
<td>6.04E-07</td>
<td>3.00E-07</td>
<td>2.016273</td>
<td>0.0440</td>
</tr>
<tr>
<td>USD(1)</td>
<td>0.993328</td>
<td>0.003116</td>
<td>318.8001</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

R-squared: 0.994887
Adjusted R-squared: 0.994878
S.E. of regression: 0.002344
Sum squared residual: 0.006460

For the ANN model we estimate the neural network model by minimizing the square of the error term as shown below.

\[ \text{Min} \sum_{k=1}^{n} (y_k - \hat{y}_k)^2 \]  

(4)

The general autoregressive estimate model is expressed as:

\[ \hat{y}_t = [\phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \ldots + \phi_p y_{t-p}] \]  

(5)

A. ANN Model

The model is designed as multi-layer perception. The model consists of three layers: an input layer, a hidden layer, and an output layer. This study used Activation function “TANH” the performance and predictability. For the first lag of Tala/AUD, the input layer has 4 variables; the hidden layer had 5 neurons (the neurons in this layer can be changed depending on the performance of the result). The output layer has 1 variable- the exchange rate of Tala/AUD with 1179. The 2nd lag of Tala/AUD, the input layer has 3 variables; the hidden layer had 4 neurons (the neurons in this layer can be changed depending on the performance of the result). The output layer has one variable- the exchange rate of Tala/AUD with 1178. The 3rd lag of Tala/AUD, the input layer has five variables; the hidden layer had 5 neurons (the neurons in this layer can be changed depending on the performance of the result). The output layer has one variable- the exchange rate of Tala/AUD with 1178. Neural network is often pictured as layer of functional node. The graph below shows the back propagation neural network. To experiment with the exchange rate time series ANN model, the author developed Microsoft Excel neural network software. The computational method employs conjugate gradient algorithm but with additional bias term in ANN model. This is to structure the favorable error surface for computational convenience.

B. ANN Computations

Step 1. Evaluate the net input to the jth node and that to the kth node in the output layer as follows:

\[ \text{net}_j = \sum_{i=1}^{p} w_{ij} y_i - \theta_j \]  

(6)

\[ \text{net}_k = \sum_{j=1}^{p} W_{jk} \hat{y}_j \]  

(7)

USD is within the forecast interval except between 20-28 November 2008, 3-12 December 2008 and on 16 December 2008.
where $i$ is the input node, $j$ is the hidden layer node, $k$ is the output layer node, $w_{ij}$ is the weights connecting the $i^{th}$ input node to the $j^{th}$ hidden layer node, $w_{jk}$ is the weights connecting the $j^{th}$ hidden layer node to the $k^{th}$ output layer node, $\theta_j$ is the threshold between the input and hidden layers.

Step 2. Evaluate the output of the $j^{th}$ node in the hidden layer and the output of the $k^{th}$ node in the output layer as follows:

$$h_j = f_h(\sum_{i=1}^{n} W_{ij} Y_i - \theta_j)$$

(8)

$$k_j = f_k(\sum_{i=1}^{k} W_{jk} Y_j)$$

(9)

Define TANH activation function as:

$$\frac{e^y - e^{-y}}{e^y + e^{-y}}$$

(10)

where, $h_j$ is the vector of hidden-layer neurons, $f_h$ and $f_k$ are logistic activation functions from input layer to the hidden layer and from hidden layer to output layer respectively. The output of each neuron is obtained by applying an activation function $f_h$ and $f_k$. The nodes are used to perform the nonlinear input/output transformations using an activation function. Since the actual adaptation seems to be a non-linear form, the activation function influences strongly the ANN predictions in this study.

Step 3. For the calculation of errors in the output and hidden layers can be expressed as follows:

The output layer error between the target and the observed output is expressible as

$$\delta_k = -(d_k - y_k)f'_k$$

(11)

$$f'_k = \frac{4}{(e^y + e^{-y})^2}$$

(12)

where $\delta_k$ is the vector of errors for each output neuron ($y_k$) and $d_k$ the target activation of output layer. The term $\delta_k$ depends only on the error $(d_k - y_k)$ and $f'_k$ is the local slope of the node activation function for output nodes. The hidden layer error is expressible as

$$\delta_j = f'_h \sum_{k=1}^{m} W_{jk}\delta_k$$

(13)

where $\delta_j$ is the vector of errors for each hidden layer neuron, $\delta_k$ is a weighted sum of all nodes and $f'_h$ the local slope of the node activation function for hidden nodes.

Step 4. Considering the ANN error equation as 12, the Newton based optimization scheme is applicable to optimize neural network weights as discussed next. In this study, quasi Newton based algorithm such as Davidson (1959), Fletcher and Powell (1963) is employed. This method falls under the general class of quasi-Newton procedures, where the search directions are of the form:

$$d_j = -D_j \nabla f(Q)$$

(14)

The gradient direction is deflected by multiplying $-D_j$, where $D_j$ is $m \times m$ positive definite symmetric matrix. It approximates the inverse of the Hessian Matrix of the neural network error function of type as shown in (12).

Let $\epsilon > 0$ be the small scalar quantity set as termination criteria. Choose and initial $m$ dimensional weight $w_i$ ($i=1,2,...,m$) resulting from an ANN error function and the initial symmetric positive definite matrix $D_1$. Let $Q_1 = w_i$, $k = j = 1$, and go to the main step.

If $\|\nabla f(Q_j)\| < \epsilon$, stop, otherwise, let $d_j = -D_j \nabla f(Q_j)$ and let $\lambda_j$ be an optimal solution to the problem:

Minimize $f(Q_j + \lambda_j d_j)$; Subject to $\lambda \geq 0$.

The parameter $\lambda$ is obtained by line a suitable search method. Let $Q_{j+1} = Q_j + \lambda_j d_j$. If $j < m$, go to step b. If $j = m$, let $Q_m = Q_{m+1}$, replace i by i+1, let $j = 1$, and repeat step a.

Construct $D_{j+1}$ as follows:

$$D_{j+1} = D_j + \frac{P_j P'_j}{q'_{j+1} D_j q_j} - \frac{d_j q'_{j+1} D_j q_j}{q'_{j+1} D_j q_j}$$

where

$$P_j = \lambda_j d_j = Q_{j+1} - Q_j \text{ and } q_j = \nabla f(Q_{j+1}) - \nabla f(Q_j)$$

Repeat j by $j+1$ and repeat step a.
V. ANN MODEL SIMULATION

The importance of normal distribution of error vector is undeniable since the underlying assumption of many statistical procedures such as t-test, regression analysis, and analysis of variance (ANOVA). The Anderson-Darling test [20] is used to test if a sample of data came from a population with a specific distribution. The test makes use of the specific distribution in calculating critical values. This has the advantage of allowing a more sensitive test and the disadvantage that critical values must be calculated for each distribution. The Anderson darling test and Shapiro and Wilk test [21] exhibit similar results when the necessarily sample size is 500 or less. In our case, the data exceeds 1000 and therefore, Anderson Darling test is robust. Further, Anderson darling test is competitive with better known Shapiro and Wilk but has the advantage that the critical values of each sample size is not needed. It is an omnibus test, in sense that it is sensitive to all types of deviation to normality, but it is somewhat more sensitive to deviation in the tails of the distribution—which is that way non normality makes itself known. The microsoft excel based computational AD test and ANN computation software are developed by the author. The graph below shows the normal probability plot through Aderson Darling Normality Test AR model.

The Aderson Darling Normality Test exhibit the p-value less than 0.05 and it implies that the data is sampled from a population that is not normally distributed. The AR time series exchange rate is a multi-layer perception model.

VI. DIAGNOSTIC ERROR TERMS

To test the validity [19] of the exchange rate estimation model, the following diagnostics measures were considered: Mean Absolute Error:

$$M A E = \frac{1}{n} \sum_{t=1}^{n} |F_t - Y_t|$$  \hspace{1cm} (15)
Mean Absolute Percentage Error:

\[ MAPE = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{\hat{Y}_i - Y_i}{Y_i} \right| \]  

(16)

Root Mean Square Deviation:

\[ RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\hat{Y}_i - Y_i)^2} \]  

(17)

To develop a forecast accuracy measures, two types of errors are considered. The one is forecast bias where the direction of the error was considered. If the value of the error is negative, then forecasting method is overestimating. Positive values imply that forecasting is underestimating. By adding error term, if there is no bias the positive and negative error term will cancel each other out and the mean error term will be zero. Even if there is no bias, it is likely that the estimation result is significant variation from the actual value. The aim of identifying accuracy in estimation is to determine how well the forecasting method estimates the actual value without evaluating forecast bias. The utilization of mean square error (MSE) as a criterion in determining the accuracy of forecast has drawbacks. Some of the errors measures listed above are MAE, RSME, and MAPE. Variance indicates the ability of the forecasts to replication degree of variability in the variable to forecast. If the variance proportion is large then the actual series has fluctuated considerably whereas the forecast has not. Covariance is the proportion that measures unsystematic error. Ideally, this should have the highest proportion of inequality.

### TABLE IV

#### Model Analysis

<table>
<thead>
<tr>
<th>Currency</th>
<th>Model</th>
<th>RSME</th>
<th>MAE</th>
<th>MAPE</th>
<th>Bias Error</th>
<th>Variance Error</th>
<th>Co-variance Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tala/AUD</td>
<td>AR(1)</td>
<td>0.0209</td>
<td>0.0150</td>
<td>3.3398</td>
<td>0.0012</td>
<td>0.3853</td>
<td>0.6136</td>
</tr>
<tr>
<td></td>
<td>AR(2)</td>
<td>0.0209</td>
<td>0.0152</td>
<td>3.3635</td>
<td>0.0017</td>
<td>0.4039</td>
<td>0.5944</td>
</tr>
<tr>
<td></td>
<td>AR(3)</td>
<td>0.0209</td>
<td>0.0152</td>
<td>3.3736</td>
<td>0.0017</td>
<td>0.4033</td>
<td>0.5951</td>
</tr>
<tr>
<td></td>
<td>AR(4)</td>
<td>0.0209</td>
<td>0.0152</td>
<td>3.3668</td>
<td>0.0018</td>
<td>0.4069</td>
<td>0.5913</td>
</tr>
<tr>
<td></td>
<td>MA (1)</td>
<td>0.022367</td>
<td>0.016123</td>
<td>3.659854</td>
<td>0.000000</td>
<td>0.286987</td>
<td>0.713013</td>
</tr>
<tr>
<td></td>
<td>MA (2)</td>
<td>0.022365</td>
<td>0.016121</td>
<td>3.569492</td>
<td>0.000000</td>
<td>0.287079</td>
<td>0.712921</td>
</tr>
<tr>
<td></td>
<td>MA (3)</td>
<td>0.022363</td>
<td>0.016119</td>
<td>3.569118</td>
<td>0.000000</td>
<td>0.287981</td>
<td>0.712019</td>
</tr>
<tr>
<td></td>
<td>MA (4)</td>
<td>0.022356</td>
<td>0.016110</td>
<td>3.560900</td>
<td>0.000000</td>
<td>2.881840</td>
<td>0.711816</td>
</tr>
<tr>
<td>Tala/USD</td>
<td>AR(1)</td>
<td>0.0189</td>
<td>0.0149</td>
<td>3.8768</td>
<td>0.0022</td>
<td>0.2362</td>
<td>0.7615</td>
</tr>
<tr>
<td></td>
<td>AR(2)</td>
<td>0.0189</td>
<td>0.0149</td>
<td>3.8915</td>
<td>0.0026</td>
<td>0.2447</td>
<td>0.7527</td>
</tr>
<tr>
<td></td>
<td>AR(3)</td>
<td>0.0189</td>
<td>0.0149</td>
<td>3.8957</td>
<td>0.0025</td>
<td>0.2449</td>
<td>0.7525</td>
</tr>
<tr>
<td></td>
<td>AR(4)</td>
<td>0.0189</td>
<td>0.0149</td>
<td>3.8927</td>
<td>0.0026</td>
<td>0.2452</td>
<td>0.7521</td>
</tr>
<tr>
<td></td>
<td>MA (1)</td>
<td>0.02189</td>
<td>0.01649</td>
<td>4.33799</td>
<td>0.000000</td>
<td>0.14839</td>
<td>0.85161</td>
</tr>
<tr>
<td></td>
<td>MA (2)</td>
<td>0.02187</td>
<td>0.01649</td>
<td>4.33535</td>
<td>0.000000</td>
<td>0.14869</td>
<td>0.85131</td>
</tr>
<tr>
<td></td>
<td>MA (3)</td>
<td>0.02186</td>
<td>0.01648</td>
<td>4.33271</td>
<td>0.000000</td>
<td>0.14935</td>
<td>0.85065</td>
</tr>
<tr>
<td></td>
<td>MA (4)</td>
<td>0.02185</td>
<td>0.01647</td>
<td>4.33068</td>
<td>0.000000</td>
<td>0.14930</td>
<td>0.85070</td>
</tr>
</tbody>
</table>

#### VII. RESULTS

The result shown in table illustrates a comparison of different models. The Exchange rate of Tala/AUD and Tala/USD have been forecasted based on AR model’s performance measure and a comparison has been made with the different MA models as shown in the Table IV. The table represents that for Tala/AUD and Tala/USD daily exchange rate, AR (1) leads to minimum error in forecasts than the listed other models. In any model, the bias error should be close to zero, if the bias error is large, it indicates systematic over or under prediction. The SSR (sum of squared residuals) shows the degree of the discrepancy between data and an estimation model. The smaller the value the better the model is. Table shows AR (1) bias error is closer to zero when compared to AR (2), AR (3), and AR (4) models and MA models. Variance error should be small; if the error is large, it indicates the actual series has fluctuated considerably whereas the forecast has not. In this case, AR (1) also shows a smaller variance error compared to other AR models. Co-Variance error should have the highest proportion of inequality and this is higher in AR (1) as compared to AR (2), AR (3), and AR (4) and other MA models.

#### VIII. CONCLUSION

The Anderson Darling normality test for the error distribution of the ANN time series shows that the data comes from a population that is not normally distributed. This assumption deviated from the traditional central limit theorem for error distributions in statistical models. The ANN time series model forms a multi-layer perception with exponential activation function. The input layer of the ANN with four neurons receives the lagged daily exchange rate time series signals. The hidden layer has five neurons with trigonometric activation function. The output layer has one neuron and transforms the signal to daily forecasted exchange rate of Tala/AUD. The series is non-stationary and hence exhibit...
challenge to construct forecasting model. The same argument is true for Tala/USD exchanger model.

This paper compared the daily exchange forecast results derived from ANN time series, exponential smoothing and winters time series model. We conclude that the autoregressive time series exchange rate of Tala vs AUD and Tala vs USD is the best model to do daily exchange rate forecast.

REFERENCES

Avlin Pritishna Prasad is a Research Assistant of Research Cluster “Economic Growth, Trade and integration” at University of the South Pacific (USP), Suva, Fiji. The main research interests of Ms. Prasad are in the area of econometrics, forecasting, economic statistics, economic analysis, and economics of growth and Development.

Ms. Prasad graduated from University of the South Pacific in 2010 with BCom. She obtained her PGDip in Economics from USP in 2013. Earlier, she served as a Research Assistant-statistics from October 2010-April 2012 at Fiji Commerce Commission, Suva, Fiji. She also served Fiji National University as temporary student assistant and Graduate Assistant/Mentor from April 2010-September 2010.