Non-parametric Approach in Modelling Effects of Remittances on Household Credit in Kenya

Caspah Lidiema[1,*], Anthony Waititu[1], Thomas Mageto[1], Anthonyn Ngunyi[2]

[1]Department of Statistics and Actuarial Sciences Jomo Kenyatta University of Agriculture and Technology
[2]Department of Statistics and Actuarial Sciences Dedan Kimathi University of Technology
*Corresponding author: lidiema.research@gmail.com

Abstract Generalized Additive Models for Location, Scale and Shape (GAMLSS)is a very flexible model class, extending the classical Generalized Additive Model (GAM) framework. Not only the mean, but all distribution parameters are regressed to the predictors. It is suitable for fitting linear or non-linear parametric models using the distributions. Artificial Neural Networks (ANN) are biologically inspired computer programs designed to simulate the way in which the human brain processes information. ANNs gather their knowledge by detecting the patterns and relationships in data and learn (or are trained) through experience. The main advantages of using ANN is that, it has the ability to implicitly detect complex nonlinear relationships between dependent and independent variables and also has ability to detect all possible interactions between predictor variables. Given all the dynamic nature of these two models are their outlined merits, it's important to test and see which of this model estimates parameters better and which of them a better model in forecasting financial data. To test and compare this models an application of effect remittances on household credit was be used. The study employed monthly data for period January 2005-December 2017 in Kenya. Our findings showed that mixed results where, GAMLSS performed better than ANN in estimation while ANN provided a better model in prediction than GAMLSS. Our results confirm that the surge in Remittances leads to increase credit uptake due to increased resource mobilization by financial institutions and also resource availability for loan repayment. The research recommends Banks and Financial institutions should also carry out their assessment using GAMLSS and ANN and come up with ways of tapping into remittances not only to boost their deposits but also increase their funds for issuing credit and hence increase interest income, and also boost financial inclusion in Kenya through increased consumer loans.

Keywords: GAMLSS, neural network, remittance, credit

1. Introduction and Literature Review

1.1. Background of the Study

A remittance is a transfer of money by a foreign worker to an individual in his or her home country. Money sent home by migrants competes with international aid as one of the largest financial inflows to developing countries. Workers' remittances are a significant part of international capital flows, especially with regard to labour-exporting countries, Al-Assaf and Al-Malki [1].

The Kenyan financial system has had a myriad of development and challenges in equal measure. Foremost is the growth of technology, especially mobile telephony that has revo- lutionized not only payments platform but mobile money transfers as well. Secondly is the paradigm shift in savings mobilization and credit extension which has increased the level of financial inclusion and the intermediation space. But this Financial Intermediation has greatly been influenced by remittances from foreign countries. Therefore, remittances can be expected to pave the way for a sound development of a formal financial sector, critical for sustaining high economic growth, Ebeke et al [2]. However, Singer [3] believed that remit- tances tend to be stable and increase during periods of economic downturns. In addition, Aggarwal and Peria [4] also confirmed that remittances tend also increase during economic slowdowns and natural disasters as compared to private capital flows. Furthermore, while a surge in inflows, including aid flows, can erode a country's competitiveness, remittances do not seem to have this adverse effect.

Remittances contributes to economic development of a country. Sharma [5] argued that increased remittances leads to reduction in economic insecurity. The author further elaborates on economic insecurity as factors related to country vulnerability, being affected by negative economic shock like unemployment, low incomes and poverty and the general economic volatility. With Kenya's, the Financial sector one of the key drivers of the economy others being Agriculture and Tourism, remittances through formal channels, leads to money supply hence providing a mechanism of savings and credit which eventually boast investments. While savings from

remittances will mostly depend on the consumption rates of the recipients, credit uptake on the other hand will depend on the use of the received money. In the second scenario (credit uptake), the recipient would likely be involved in some business or otherwise loan uptake would be a residue or expected receipt of foreign remittance. This implies that if its business related there is a possibility of economic growth not only for the receiving individuals but for the country as well.

This argument is supported by Mbaye [6] who argued that foreign remittances increases the probability of recipients securing loans from local financial institutions since, remittances provide an alternative way of repaying loans and hence cushion against household economic shocks. The author further gives the slip side causality argument of the view that, not only does remittances enable households to access to credit but they could also be sent to repay loans that were taken by their relatives before the receiving date or alternatively to repay loans that were taken to finance their migrations to foreign countries. Similarly, research by Hernandez et al [7], states that remittances changes borrowers expected marginal returns to credit and/or their creditworthiness, as perceived by lenders.

Given the dynamic nature of foreign remittances and credit uptake which includes nonlinearity due to changing nature of micro-economic environment, there is need to employ methods than can model this data without overfitting and without losing so much information. Two of the methods that can perform have these attributes include Generalized Additive Models for Location, Scale and Shape (GAMLSS) and Artificial Neural Networks (ANN). The main reason for employing GAMLSS and ANN is because, Non-Parametric relaxes the usual assumption of linearity and enables you to uncover relationships between the independent variables and the dependent variable that might otherwise be missed.

GAMLSS is a very flexible model class, extending the classical GAM framework not only the mean, but all distribution parameters are regressed to the predictors. It has been applied extensively in real estate, environment and climatology, partly in finance industry and health. GAMLSS allows fitting of linear or non-linear parametric models using the distributions, Stasinopoulos and Rigby [8]. The author further states that the model is also flexible in specifying the distribution of the response variable including highly skewed and/or kurtotic distributions and can also allow modelling of all parameters for all distribution parameters.

Artificial neural networks (ANNs) are biologically inspired computer programs designed to simulate the way in which the human brain processes information. ANNs gather their knowledge by detecting the patterns and relationships in data and learn (or are trained) through experience. ANN represents a promising modeling technique, especially for data sets having non-linear relationships which are frequently encountered in finance and economics. The main advantages of using Artificial Neural Networks (ANN) include: it can handle large amount of data sets; it has the ability to implicitly detect complex nonlin- ear relationships between dependent and independent variables; it has ability to detect all possible interactions between predictor variables. Given all the

dynamic nature of these two models are their outlined merits, it's important to test and see which of this model is better in estimating and forecasting financial data.

1.2. Review of Previous Studies

Takagi et al [9] compared the performance of Generalized Additive Model with Artificial Hierarchical Neural Network in the Analysis of Pharmaceutical Data. The authors employed two types of datasets. First, they determine the relationships between the molecular properties and intestinal absorption rates of some different drugs. Second, they used esophageal cancer data set from their many data sets to assess the relationships between some risk factors, such as alcohol and tobacco consumption, and esophageal cancer morbidity. Their results showed that Generalized Additive Models (GAM) and ANN showed similar regression/classification/prediction powers.

Villarani et al [10] analyzed Non stationary records of data and temperatures using GAMLSS with data period 1862-2004. The results showed that GAMLSS is able to model rainfall and temperatures in Rome using magnitude and spread of oscillations from different seasons.

Tan [11] using Artificial Neural Networks developed an early warning predictor for Credit Union financial distress. The author used quarterly financial data for 191 New South Wales Credit Unions from 1989 to 1991. The data were 'cleaned' by Hall and Byron to exclude all credit unions with total assets less than $60,000. The total number of observations obtained for the study was 2144, of which 66 were classified as in distress. The input (independent) variables were financial ratios derived from the financial data used by Hall and Byron. The author employed artificial neural network (ANN) model and Probit Model. The author found that The ANN model has performed better than the Probit model as an early warning model for predicting Credit Unions in distress. The overall accuracy of the ANN model vs. the Probit model was between 90% - 92%.

Setodji et al [12] used Generalized Additive Modeling to Empirically Identify Thresholds within the ITERS (Infant/Toddler Environment Rating Scale) in Relation to Toddlers' Cognitive Development. They compared the GAM results with the simple linear regression and found out that analyzed performance of GAM was better in depicting the relation- ships between ITERS scores and Toddlers' Cognitive Development and that GAM-derived thresholds were more effective at differentiating among children's performance levels on the BMDSF-R than the simple linear models.

Stasinopoulos and Rigby [8] modeled CD4 counts from uninfected children born to HIV-1 mothers and the age in years of the child using GAMLSS. They modeled using age as the independent variable. The author fitted orthogonal polynomials of different orders to the data and chose the best using a GAIC criterion. They found out that the model with a polynomial of degree 7 was the best model.

Serinaldi [13] Studied distributional modeling and short-term forecasting of electricity prices by Generalized Additive Models for Location, Scale and Shape. The author tested this model using two data sets from California Power Exchange (CalPX) market, and the

Italian Power Exchange (IPEX). CalPX data allow comparing the GAMLSS forecasting performance with published results obtained by different models. The study found out that out that the GAMLSS framework can be a flexible alternative to several linear and nonlinear stochastic models especially in enable the short-term (one-day ahead) forecast of the entire distribution of prices.

Mbaye [6] examined the impact that remittance on household's access to credit using, ordinary least square regression (OLS). The author employed household fixed effects model in order to control for problems of selection effects and potentially time invariant characteristics of households in the sample. The results indicated that, recipients of foreign remittances increase their likelihood of receiving a loan in a household by about 4.9 percentage points.

1.3. Statement of the Problem

A number of studies in the world have developed a fairly sizeable literature on effects of credit on economic growth. A recent paper by the OECD [14] intimates that even though Finance is a vital ingredient of economic growth, too much credit expansion can hurt the economy by likely slowing long-term growth and raise inequality rather than make it grow. Therefore the impact of the impact of remittances on this credit will be key area of study to determine what role it plays on credit uptake. Brown and Carmignani [15] argued that the inflow of remittances modifies the credit supply and demand schedule and thus requires the bank to adjust the rates in order to achieve profit maximization. However, this argument has not empirically been tested. But there is also little evidence that the existing literature have included macro-economic variables (Lending/interest rate and Foreign exchange rate) as in intervening variable in the study relationship. To be able to draw conclusions on the remittance-household credit interplay, there is need to study the effect macroeconomic variables in a bid to establish better policies along this study line. Additionally, financial time series data is non-linear in nature due to market dynamics and any model that is used to model this kind of data must be able to capture this dynamics of both linear and non-linear and also relax the assumption of normality and parametric nature. The traditional regression models like Multiple Linear regression or traditional time series models like Autoregressive model (AR), Auto regressive Distributed Lag model (ARDL) and the Vector error correction model (VECM) and the vector Autoregressive model (VAR) usually assume linearity and parametric nature of any financial data. The traditional methods have several deficiencies for example over-fitting and over-parameterization. Due to these limitations this study employs Generalized Additive Model for Location, Scale and Shape, Artificial Neural Network (ANN) so as to capture and allow fitting a variety of distributions whose parameters change according to covariates via a number of linear and nonlinear relationships.

1.4. Justification of the Study

The Kenyan financial system is dominated by commercial banks and therefore any developments of either market or policy nature can adversely influence the financial inclusion and Finance (savings and credit) in either positive or negative way. Since Financial Inclusion is highly affected by Credit it is prudent to investigate factors that affect Credit uptake in Kenya. One such factor is remittances from foreign countries. To a large extent remittances act as source of external finance in many developing and emerging countries and as such, this may greatly influence internal financing (which is basically credit from Commercial banks and other Non-Banking Financial Institutions (NBFIs)). This research is therefore justified in the sense that it will assist a number of stakeholders in answering several questions surrounding the true effect of remittances on credit. The choice of modelling will also play an important role in determining which model gives the best results and trying to respond to the nature of this financial data described above. Some of the stakeholders who will benefit include.

1. To assist researchers and policy makers on the best model for modeling and forecasting remittances and household credit.
2. This information will enable policy makers to understand the effect of remittances on credit and the general Financial inclusion agenda and hence make wise monetary policies to boost investment and the economy.
3. The study will also be of great interest to as they seek to know the effect of monetary policy on Finance and hence able to use this research to forecast their credit levels and hence their profitability.

1.5. Objectives of the Study

1.5.1. General Objective

To Model Non-parametric approach of effects of Remittances on Household credit in Kenya.

1.5.2. Specific Objectives

1. To investigate the performance of Generalized Additive Model for Location, Scale and Shape (GAMLSS)
2. To Examine the performance of Artificial Neural Network (ANN)
3. To determine the predictive ability of GAMLSS and ANN

2. Methodology

In this study, two models will be employed, namely; Generalized Additive Model for Location, Scale and shape (GAMLSS) and Artificial Neutral Network (ANN). The general model equation will take the form:

$$Y = \alpha_0 + \sum_{i=1}^{k} \alpha_i (X_i) + \varepsilon \qquad (1)$$

2.1. Generalized Additive Model for Location, Scale and Shape (GAMLSS)

Generalized Additive Models for Location, Scale and Shape (GAMLSS) is framework for fitting regression

models. It is a special type of generalized additive models. GAMLSS were introduced by (Rigby & Stasinopoulos, 2001, 2005) in order to overcome limitations associated with Generalized Linear Models (GLM) and Generalized Additive Models (GAM) (Hastie & Tibshirani, 1990), Stasinopoulos et al [16]. Some of these limitations include overfitting due to the flexibility of nonparametric and additive regression. There is also a risk of interpreting spurious relationship in the overfitted data. In using GAMLSS the exponential family distribution assumption relaxed and replaced by a general distribution.

Stasinopoulos et al [16] stated that GAMLSS assumes independent random observations y_i for $i = 1, 2, ..., n$ with probability density function $\left(f\left(y_i \mid \theta^i \right) \right)$ which is conditional on θ^i where $\theta^i = \left(\theta_{i1}, \theta_{i2}, ..., \theta_{ip} \right)$ is a vector/matrix of p parameters which are related to the explanatory variables.

Stasinopoulos et al [16] propose that this parameters should be at most four ($p = 4$) denoted by $\left(\mu_i, \sigma_i, v_i, \tau_i \right)$. Where mu_i and σ_i represent location and scale respectively while v_i and τ_i represent the shape parameters.

Therefore, let $y^T = \left(y_1, y_2, ..., y_m \right)$ be a vector of response variable with length n and let $k = 1, 2, 3, 4$. We can define a function $g_k(.)$ as a monotonic link function relating to k^{th} parameter denoted as θ_k. Using the parameters defined above and the vector definitions, the model can therefore be written ad follows:

$$g_k\left(\theta_k\right) = \eta_k = \mathbf{X}_k \beta_k + \sum_{j=1}^{jk} \mathbf{Z}_{jk} \gamma_{jk}. \tag{2}$$

Expanding the above equation we get a set of equations as below

$$
\begin{aligned}
g_1\left(\mu\right) &= \eta_1 = \mathbf{X}_1 \beta_1 + \sum_{j=1}^{j1} \mathbf{Z}_{j1} \gamma_{j1} \\
g_2\left(\sigma\right) &= \eta_2 = \mathbf{X}_2 \beta_2 + \sum_{j=1}^{j2} \mathbf{Z}_{j2} \gamma_{j2} \\
g_1\left(v\right) &= \eta_3 = \mathbf{X}_3 \beta_3 + \sum_{j=1}^{j2} \mathbf{Z}_{j3} \gamma_{j3} \\
g_4\left(\tau\right) &= \eta_4 = \mathbf{X}_4 \beta_4 + \sum_{j=1}^{j4} \mathbf{Z}_{j4} \gamma_{j4}
\end{aligned} \tag{3}
$$

where γ_{jk} have an independent normal distribution. That is, $\gamma_{jk} \sim N_{qjk}\left(0, \mathbf{G}_{jk}^{-1}\right)$ is the generalized inverse of $q_{jk} \times q_{jk}$ vector system $G_{jk} = \mathbf{G}_{jk}\left(\lambda_{jk}\right)$. If $\mathbf{G}_{jk}$ is singular, then λ_{jk} will have a prior density function equivalent to $\exp\left(-\frac{1}{2} \lambda_{jk}^T \mathbf{G}_{jk} \lambda_{jk} \right)$.

Assume $\mathbf{Z}_{jk} = \mathbf{I}_n$ where $\mathbf{I}_n$ is an identity matrix of order $n \times n$ and $\lambda_{jk} = h_{jk}\left(x_{jk}\right)$ then our GAMLSS equation will be

$$g_k\left(\theta_k\right) = \eta_k = \mathbf{X}_k \beta_k + \sum_{j=1}^{jk} h_{jk}\left(x_{jk}\right). \tag{4}$$

The set of equations will also transform and become:

$$
\begin{aligned}
g_1\left(\mu\right) &= \eta_1 = \mathbf{X}_1 \beta_1 + \sum_{j=1}^{j1} h_{j1}\left(x_{j1}\right) \\
g_2\left(\sigma\right) &= \eta_2 = \mathbf{X}_2 \beta_2 + \sum_{j=1}^{j2} h_{j2}\left(x_{j2}\right) \\
g_1\left(v\right) &= \eta_3 = \mathbf{X}_3 \beta_3 + \sum_{j=1}^{j2} h_{j3}\left(x_{j3}\right) \\
g_4\left(\tau\right) &= \eta_4 = \mathbf{X}_4 \beta_4 + \sum_{j=1}^{j4} h_{j4}\left(x_{j4}\right)
\end{aligned} \tag{5}
$$

where $\mu_i, \sigma_i, v_i, \tau_i$ are vectors of length n and h_{jk} is a non-parametric additive function of the independent variable x_{jk}. β_k are parameter vectors.

2.1.1. Back Fitting Algorithm for GAMLSS

The model parameters can be estimated using three algorithms. The first is CG algorithm by Cole and Green introduced in 1992. This algorithm uses the first derivative and expected values of the second and cross derivatives of the likelihood function with respect to the parameters *μ, σ, υ, τ*.

The second algorithm will is the RS algorithm by Rigby and Stasinopoulos introduced in 1996. This method uses the mean and Dispersion Additive Models denoted MADAM.

The third method is the Mixed method which combines the RS and the CG. This means that the *j* iterations of RS, followed by *k* iterations of CG.

2.1.2. Best Model for GAMLSS

Usually the best model using the algorithms above can be selected using the generalized Akaike Information Criteria (GAIC). GAIC usually penalize over fitting. This is achieved through addition of a fixed penalty k to the global deviance for each degrees of freedom in the model.

$$\text{GAIC} = -2\hat{l} + k.df \tag{6}$$

where $\hat{l}$ is the log-likelihood of the fitted function and k is a required penalty. $k = 2$ for the AIC.

Therefore our AIC will be given by

$$\text{AIC} = -2\hat{l} + 2.df. \tag{7}$$

2.2. Artificial Neural Network (ANN)

Artificial neural networks (ANNs) is computing systems inspired by the biological neural networks similar to the human/animal brains. ANN is based on a collection of connected units called artificial neurons. Generally, we can say that Artificial Neural Networks, is a biologically

inspired network of artificial neurons configured to perform specific tasks.

From equation 1 we can re-write it as follows

$$Y = f\left(net\right) = f\left(\sum_{i=1}^{n} w_i x_i\right) \tag{8}$$

where Y is the Output neuron, w_i is the weight of the vector and the function $f\left(net\right)$ is the activation function. net is a a scalar product variable of the weight and input vectors given by;

$$net = w^t x = w_1 x_1 + \ldots + w_n x_n \tag{9}$$

where t is the transpose matrix.

Y can be computed by the equation below

$$Y = f\left(net\right) = \begin{cases} 1 \; if \; w^t x \geq \theta \\ 0 \; Otherwise \end{cases} \tag{10}$$

where θ represents the linear threshold node.

A typical neural network contains a large number of artificial neurons called units arranged in a series of layers as shown below.

- Input layer: it contains those units (artificial neurons) which receive input from the outside world on which network will learn, recognize about or otherwise process.
- Output layer: it contains units that respond to the information about how it's learned any task.
- Hidden layer: these units are in between input and output layers. The job of hidden layer is to transform the input into something that output unit can use in some way.

The similarity of Artificial Neural networks and the human brain lies in two instances, namely; (i) A neural network acquires knowledge through learning and (ii) The Networks knowledge is stored within inter-neuron connection strengths known as synaptic weights. Artificial neural networks can therefore be said to be a system in which the artificial neurons are nodes and directed by the weights (connections) between neuron outputs and neuron inputs. These inputs are mathematically designated by the notation $x(n)$ for n number of inputs. Every input is multiplied by its corresponding weights (the information used by the neural network to solve a problem). These weights represent the strength of the interconnection between neurons inside the neural network. Usually the weighted inputs are summed up within the artificial neuron. When weighted sum is zero, bias is added to make the output not-zero or to scale up the system response. Bias has the weight and input always equal to '1'. The sum always goes through the activation function. The activation function which is either be binary, linear (sigmoidal) or non-linear (tan hyperbolic sigmoidal) is set of the transfer system used to get the required output.

2.2.1 Feed-Forward Neural Network (FFNN)

A Feed-Forward Neural Network (FFNN) has all the information moving in only one direction, forward from the input nodes, through the hidden nodes and to the output nodes. There are no cycles or loops in the network. These implies that, the feed forward neural network does not have any connections from output to input. All inputs with variable weights are connected with every other node. A single layer feed forward network has one layer of nodes, whereas a multi-layer feed forward network has multiple layers of nodes. The structure of a feed forward single layer network is shown above in Figure 1.
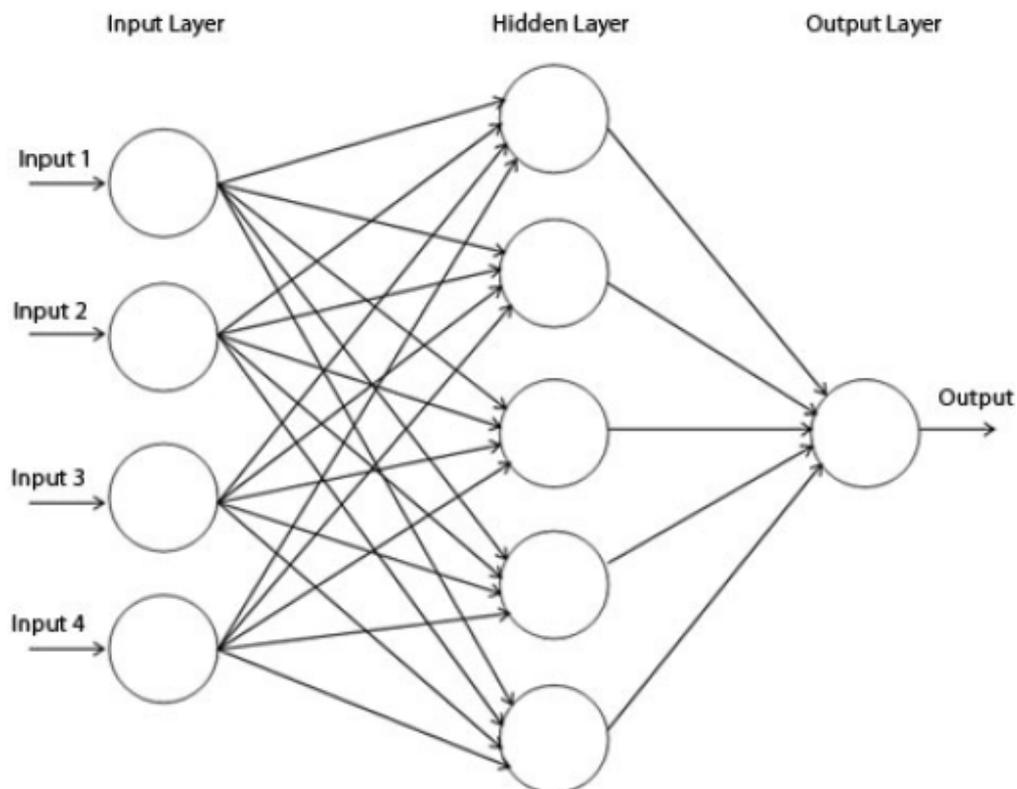

Figure 1. Feedforward neural network with one hidden layer (p =4 and H =5)

The above ANN structure has 4 inputs, one layer of 5 hidden nodes and an activation function $\psi(x)$. The input at hidden layer nodes are connected by weights W_{hj} for $h \in (0)$.

Feed-forward networks have the following characteristics:

1. Perceptrons are arranged in layers, with the first layer taking in inputs and the last layer producing outputs. The middle layers have no connection with the external world, and hence are called hidden layers.
2. Each perceptron in one layer is connected to every perceptron on the next layer. Hence information is constantly "fed forward" from one layer to the next., and this explains why these networks are called feed-forward networks.
3. There is no connection among perceptrons in the same layer.

The advantages of a neural network are: firstly, it can detect nonlinear relationship between the dependent and independent variable. Sarle [17] stated that Multi Layer perceptron (MLP) strength lies in its ability to model problems of different levels of complexity, ranging from a simple parametric model to a highly flexible, nonparametric model. The second advantage is that, one can efficiently train large datasets using the parallel architecture. Thirdly, it is a nonparametric model so that one can eliminate errors in the estimation of parameters. The main disadvantages of neural network are that it often converges to the local minimum rather than the global minimum. Also, it might over-fit when the training process goes on for too long. Some of the studies carried out using Artificial neural networks include, Forouzanfar et al [18] and Maliki et al [19].

2.2.2. Training the Network

Our model estimation will use the Broyden-Fletcher-Goldfarb-Shanno (BFGS) training algorithm with the use of sum of squared errors (SSE) method as the stopping rule. The BFGS method belongs to quasi-Newton methods which are essentially much faster than steepest descent and the Back Propagation training methods.

Like any Quasi-Newton methods, BFGS uses quadratic Taylor approximation of the objective function in a d-vicinity of x.

BFGS solves for the roots of a nonlinear equation by providing a linear approximation to the nonlinear equation at a given point x_k. The iterative method is expressed as

$$x_{k+1} = x_k + \frac{f(x_k)}{f'(x_k)}. \tag{11}$$

The objective is to find a stationary point x^* such that $f'(x^*) = 0$. if we differentiate twice the objective function we get and calculate iteratively we get

$$x_{k+1} = x_k + \frac{f'(x_k)}{f''(x_k)} \tag{12}$$

we shall interchangeably replace $f'(x)$ and $f''(x)$ with the gradient, $\nabla f(x)$ and the Hessian, $\nabla^2 f(x)$,

respectively. We can approximates $f''(x_k)$ with a finite differencing scheme as

$$f'(x_k) \approx \frac{f'(x_k) - f'(x_{k-1})}{x_k - x_{k-1}}. \tag{13}$$

This BFGS quasi-Newton approach can therefore be thought of as a generalization of the secant method. Since BFGS optimization is, given a point x_k, we try to solve for p_k by approximating the Hessian vector β_k at x_k. We can then get the direction p_k by solving $\beta_k p_k = -\nabla f(x_k)$ for $k = 0,1,....$. We then follow this search direction to obtain a new guess $x_{k+1} = x_k + \alpha_k p_k$, where α_k is the step-size, often determined through a scalar optimization or a line search procedure. The descent direction then becomes

$$p_k = -\nabla f(x_k). \tag{14}$$

The BFGS update formula falls into a broader class of rank two update formulas that maintains both symmetry and positive definiteness (SPD) of the Hessian approximation. Preservation of SPD ensures that the search direction is always a descent direction. The general rank two update formula is given by

$$\beta_{k+1} = \beta_k - \frac{(\beta_k s_k)(\beta_k s_k)^T}{s_k^T \beta_k s_k} + \frac{y_k y_k^T}{y_k^T s_k} + \phi(s_k^T \beta_k s_k) v_k v_k^T, \tag{15}$$

where ϕ is a scalar and $s_k = x_{k+1} - x_k$, $y_k = \nabla f(x_{k+1}) = \nabla f(x_k)$, $v_k = \frac{y_k}{y_k^T s_k} - \frac{\beta_k s_k}{s_k^T \beta_k s_k}$. The BFGS Hessian update is obtained by setting $\phi = 0$, and therefore simplifies to

$$\beta_{k+1} = \beta_k - \frac{(\beta_k s_k)(\beta_k s_k)^T}{s_k^T \beta_k s_k} + \frac{y_k y_k^T}{y_k^T s_k}. \tag{16}$$

The The BFGS quasi-Newton algorithm can be summarized by the following steps:

1. Specify an initial β_0 and x_0.
2. For $k = 0,1,2,...$
   (a) Stop if x_k is optimal
   (b) Solve $\beta_k p_k = -\nabla f(x_k)$ for search direction p_k.
   (c) Use a line search to determine the step-size α_k.
   (d) Update $x_{k+1} = x_k + \alpha_k p_k$.
   (e) Compute β_{k+1} using the BFGS update.

2.3. Model Comparison

According to Pitt et al [20] and Myung & Pitt [21], Models are often compared based on their goodness of fit. There are various methods to evaluate the goodness of fit. This include they are the Akaike Information Criterion (AIC); the Bayesian Information Criterion (BIC); minimum

description length (MDL); cross-validation (CV) and finally, Bayesian model selection (BMS).

This research will employ the AIC for model performance and root-mean-square error (RMSE), will be used for forecasting purposes.

Generally AIC is given by:

$$AIC = 2 \times \ln(Linklihood) + 2 \times k \qquad (17)$$

where ln is the natural logarithm, k is the number of parameters in the mode.

RMSE is generally given by

$$RMSE = \sqrt{\frac{1}{n}\sum(y - \hat{y})^2} \qquad (18)$$

where: $\hat{y}$ is the forecast values (Expected values) and y is the observed values (known Values)

RMSE measures the differences between values predicted by a hypothetical model and the observed values. In other words, it measures the quality of the fit between the actual data and the predicted model. For the two measures above, the smaller the value, the better the fit of the model

3. Results and Discussions

3.1. Data

The empirical data analysis is based on a set of monthly data for the period January 2005 to December 2017 for Household Credit as a dependent variable, foreign remittances into Kenya as independent variable and lending rates as a moderating variable. This data was obtained from Kenya Bankers Association (KBA) and the Central Bank of Kenya (CBK) website. The data was divided into two. The first part was used for training and the second part was used for testing accuracy of the prediction. All objectives of this study will be analyzed using R Project for Statistical computing available utilizing all possible packages from R Comprehensive Archive Network (CRAN) at https://cran.r-project.org/mirrors.html. The packages used are gamlss and neuralnet.

3.2. Fitting the GAMLSS

Table 2 provides estimated results for the relationship between Household Credit, Remittances and Lending rates. Four family distributions were tested namely, the normal (NO), gamma (GA), power exponential (PE) and a Box-Cox power exponential (BCPE) distribution. According to Stasinopoulos and Rigby [8], normal with gamma explores checks whether there is a positive skewness in the data, while normal with power exponential evaluates the possibility of kurtosis and finally BCPE confirms if both skewness and kurtosis are exhibited in the data. AIC is then used to pick the best distribution for use in further analysis. Among the four distributions Normal distribution had the best AIC and hence used. Results show that there is a direct relationship between Diaspora remittances and Household credit. As expected, lending rates have an inverse relationship with the household credit since

unfavorable pricing of loans is expensive for common borrowers. Remittance and Lending rates were all found to be significant at 5% level of significance the model AIC is -272.58.

Table 1. Mu Coefficients

	Estimate	Std. Error	t value	Pr (>|t|)
Intercept	-4.87	0.088	-55.13	0.00***
Remittance	1.09	0.005	208.78	0.00***
Lending rates	-0.01	0.00	-68.03	0.00***

***Significant Level is 5% Level.

Table 2. Sigma Coefficients

	Estimate	Std. Error	t value	Pr (>|t|)
Intercept	24.90	1.71	14.53	0.00***
Remittance	-1.92	0.13	-13.75	0.00***
Lending rates	0.15	0.04	3.6	0.00***

***Significant Level is 5% Level.

3.3. Diagnostic Checks

The diagnostic checks were also performed using Figure 2, Figure 3 and Figure 4. The upper panel in Figure 2, shows quantile residuals against the fitted values and the two lower panels shows the kernel density estimate and the normal Q-Q plot. From Figure 2, results indicate residuals are random and the Q-Q lot confirm normality. Figure 3 and Figure 4 on the other hand shows the worm plots. Worm plot is a de-trended representation of the normal Q-Q plot, Zhang et al [22]. The worm plots confirm the choice of the family of distribution. Usually the shape of the worm supports the distribution family selection. For our case the shape is at meaning the shape of the worm is in agreement with our selection of Normal distribution family at 95% confidence level.

3.4. Artificial Neural Network

In this research, Multi-layer, feed-forward ANNs consisting of two input neuron, four hidden neurons under one hidden layer, and one output neuron were used. In reference to Katz (1992), Maciel & Ballini [23] stated that optimal number of hidden neurons can be anything between one-half to three times the number of input neurons. This research therefore trained the network using several hidden neurons from 1 to 6. Network with 2 hidden neurons provided the stable model with smallest error as shown in Table 2. Sheela and Deepa [24] argued that hidden neuron can influence the error on the nodes to which their output is connected. Since the stability of the neural network is estimated by error, the minimal error after testing several models with different hidden layers reflects better stability while higher error reflects worst stability. The learning rate was set as 'Null' so that the system picks the optimal learning by itself.
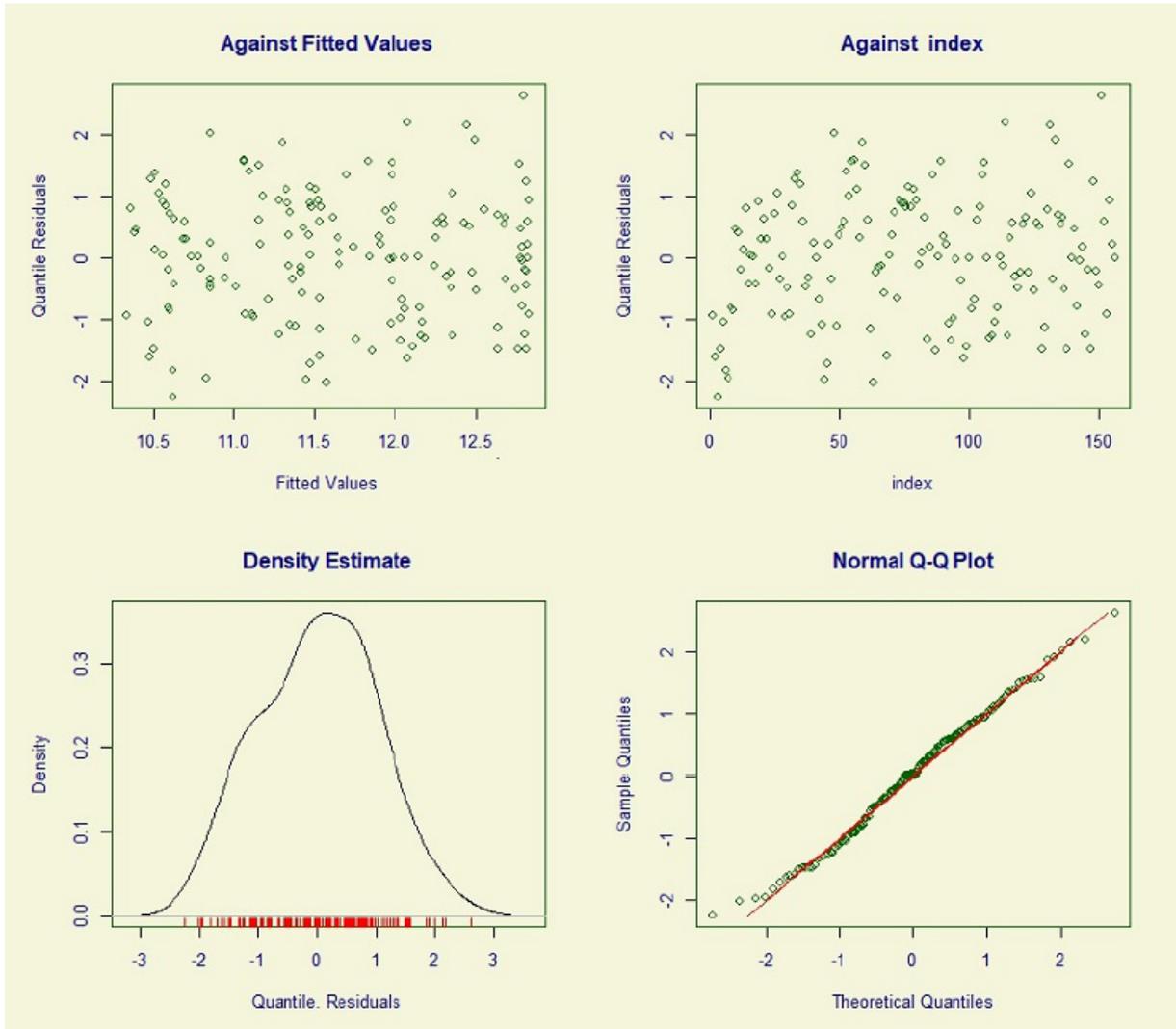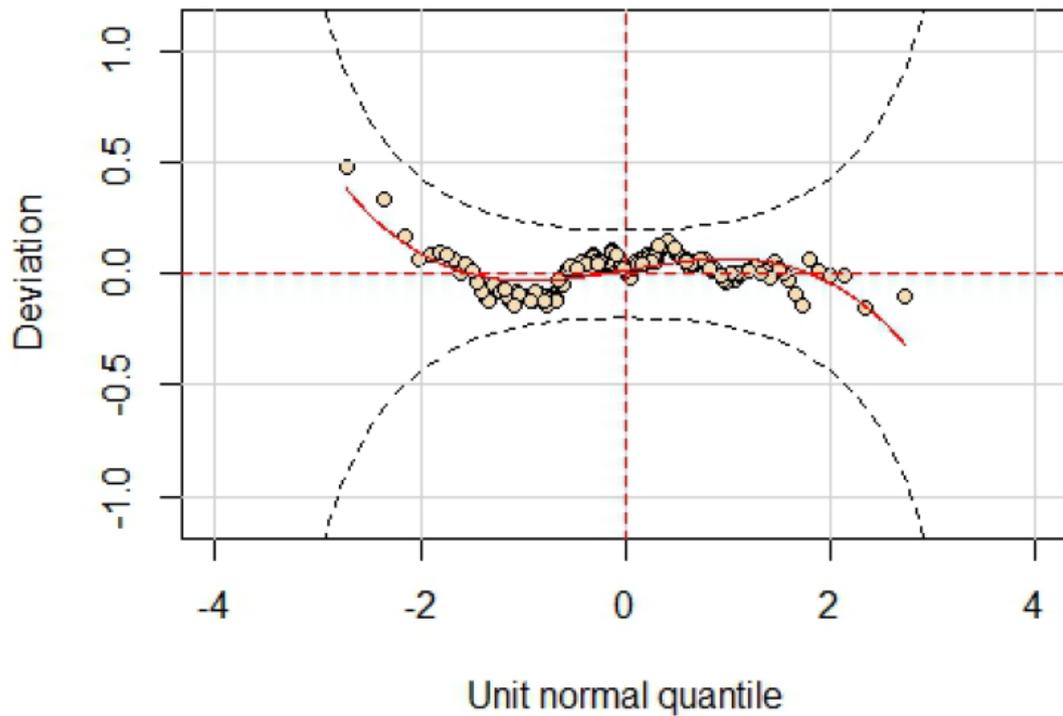
Figure 2. Diagnostic Plots
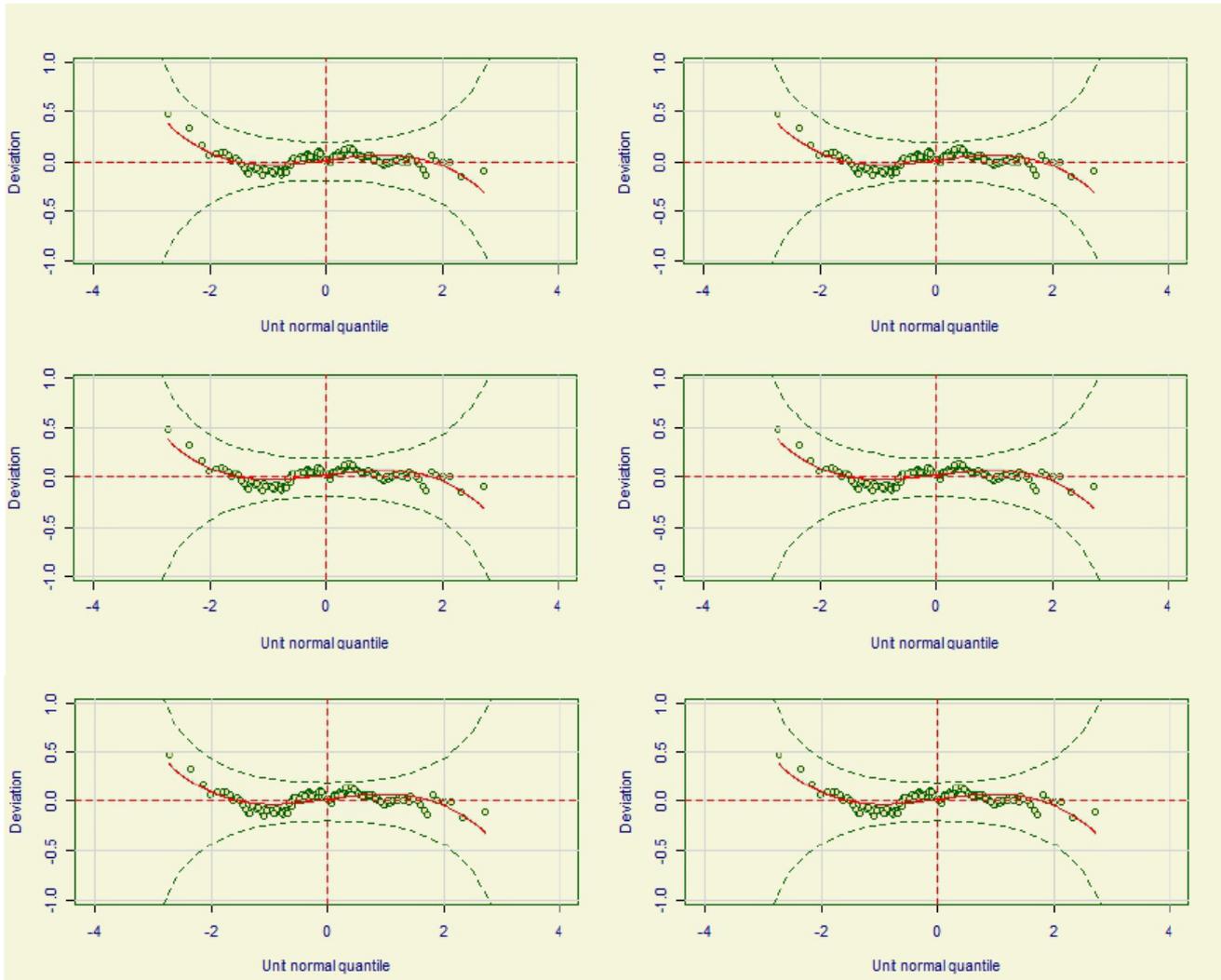

Figure 3. Worm Plots.

Figure 4. Residual Plots

3.4.1. Training of the Network

Our network was trained with a default threshold of 0.01 which network requires the model partial derivative error to change at least 0.01. The network was also trained using 100,000 iterations. Another item that we included in our network in the rep function which allows the network to repeat training with different starting weights since we did not define the starting weights.

Table 3. Hidden Neurons choice

Hidden Neurons	Error	Steps
6	3.576	67545
5	3.573	87352
4	3.571	88979
3	3.59	13259
2	38.219	66
1	38.219	93

By selecting a model with 4 hidden neurons our results are presented in Table 4 and our Network plot is presented in Figure 5.

Table 4. Estimate results

| No.	ANN Item	Estimate
1	error	3.57086
2	reached.threshold	0.00995
3	steps	88979
4	aic	-41.14171
5	bic	-91.62854
6	Intercept.to.1layhid1	1.76634
7	Remt.to.1layhid1	1.08061
8	Lenrate.to.1layhid1	0.90688
9	Intercept.to.1layhid2	0.29537
10	Remt.to.1layhid2	1.41799
11	Lenrate.to.1layhid2	-0.39197
12	Intercept.to.1layhid3	2.8423
13	Remt.to.1layhid3	-0.2446
14	Lenrate.to.1layhid3	0.02196
15	Intercept.to.1layhid4	3.42628
16	Remt.to.1layhid4	0.83948
17	Lenrate.to.1layhid4	1.68956
18	Intercept.to.credit	0.6504
19	1layhid.1.to.credit	1.45754
20	1layhid.2.to.credit	1.81282
21	1layhid.3.to.credit	-9.40812
22	1layhid.4.to.credit	1.15977

Error: 3.570856 Steps: 88979
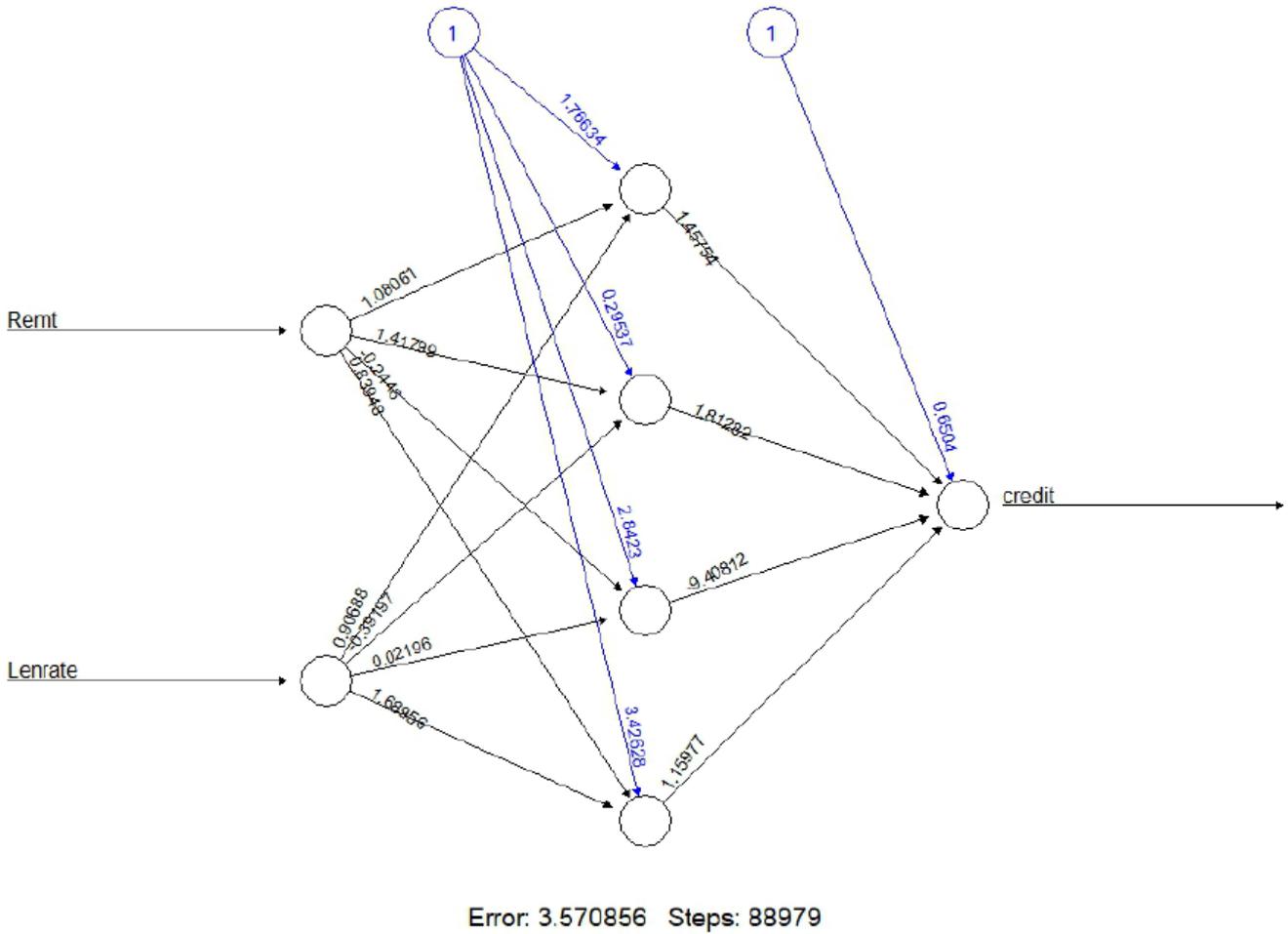
Figure 5. Network plot with 4 hidden neurons

3.5. Model Estimation Comparison

This research study employs Akaike Information Criteria (AIC) to select the best model in estimation. AIC is chosen because it is easier to interpret. AIC is generally used to estimate the information loss and the best model is the one the loses the lowest information. AIC is generally given as

$$\text{AIC} = = -2\hat{l} + 2.D \qquad (19)$$

where $\hat{l}$ is the maximum likelihood for the model, and D is the the free parameter (otherwise known as degrees of freedom).

The author also compared the AIC and BIC with a view to find out if the results would be different. According to Wagenmakers and Farrell [25], AIC neglects the sampling variability of the estimated parameters. Further AIC becomes generally inconsistent especially in situations where the sample size is very large. This is why the author decided to include the BIC to compare the two.

BIC is denoted as

$$\text{BIC} = -2\log\hat{l} + D\log n \qquad (20)$$

where $\hat{l}$ is the maximum likelihood for the model, and D is the the free parameter (otherwise known as degrees of freedom). The main advantage of BIC is its consistent as $n \to \infty$.

From estimation results both AIC and BIC performed the same in model selection as indicated in Table 5. GAMLSS outperformed ANN and BVAR in estimation.

Table 5. BVAR Lag Selection Results

Model	AIC	BIC	RMSE (out- sample)	RMSE (in-sample)
GAMLSS	-272.58	-176.14	0.007487	0.017
ANN	-41.14	-91.62854	0.000058	0.0000659

3.6. Predictive Ability

The last objective of this research is to compare the models in terms of prediction performance in order to make a decision on the forecasting accuracy of the models. The forecasting performance is evaluated root mean square errors (RMSE).

The RMSE is given by:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum(y - \hat{y})^2} \qquad (21)$$

Where y is the actual value and $\hat{y}$ is the forecasted one. This is essentially the sample standard deviation of the forecast errors (without any degrees of freedom adjustment).

The results indicate that the Artificial Neural Network model outperforms the GAMLSS in predictive accuracy

as indicated in Table 5 above. These results are interesting considering that GAMLSS had a better estimation in terms of AIC and BIC but ANN has a better predictive accuracy. These findings provide mixed results where one model is good in estimation and one is better in prediction. Wooldridge [26] on page 651 confirmed that a model might provide a good fit to a response variable in the sample used to estimate the parameters, but this need not translate to good forecasting performance. This results also confirms that purely Non-Parametric Models like ANN performs better than semi-Parametric models like GAMLSS in prediction and forecasting. We also found that the out of sample prediction performed better than the in sample prediction.

4. Conclusion and Recommendations

The main objective of the study was to establish the best model between GAMLSS and ANN in estimating and predicting effects of remittances on Household credit. GAMLSS provides a way to solve non-parametric relationships since it does not assume the exponential family of distributions but rather allows for a wide spectrum of family of distributions. The two models employed in this study showed almost similar results, in which the relationship between remittances and household credit is positive and inverse relationship between the Lending rates and household credit, their main difference was in the accuracy if estimation and Prediction. While GAMLSS was a better model in estimation, ANN provided a better model in prediction. Our findings indicate mixed results and as such, further research should be carried out to establish which among the good models in estimation and good models in prediction is the better model overall. With these findings, Banks and financial institutions should assess extent to which remittances affect household credit and their future loan uptake using either of the models depending on the exact purpose they wish to achieve. Increased consumer loans, could possibly mean increased access to financial services from banks and other financial institutions. Therefore, other studies should be carried out to establish if remittances impact on financial inclusion in Kenya. Banks and Financial institutions can also adopt this study and assess the remittance situation and come up with ways of tapping into remittances not only to boost their deposits but also increase their funds for issuing credit and hence increase interest income. Further studies can be done in the same area using other Non-Parametric models like Kernel Smoothing, K-Nearest Neighbor Estimates (K-NN) and compare with the results of this study.

References

[1] Al-Assaf, G., & Al-Malki, A. M. (2014). Modelling the Macroeconomic Determinants of Workers' Remittances: The Case of Jordan. *International Journal of Economics and Financial Issues*, 4(3), 514.

[2] Ebeke, M. C., Loko, M. B., & Viseth, A. (2014). *Credit Quality in Developing Economies: Remittances to the Rescue?* (No. 14-144). International Monetary Fund.

[3] Singer, D. A. (2008). Migrant Remittances, Financial Globalization, and Exchange Rate Regimes in the Developing World. *Massachusetts Institute of Technology, Cam- bridge, MA*.

[4] Aggarwal, R., & Peria, M. S. M. (2006). *Do workers' remittances promote financial development?* (Vol. 3957). World Bank Publications.

[5] Sharma, K. (2010). The impact of remittances on economic insecurity. *Journal of Human Development and Capabilities*, 11(4), 555-577.

[6] Mbaye, L (2016). Remittances and credit markets: complimentarities and evidence from senegal. *international Growth Center Blog*.

[7] Hernandez-Hernandez, E., Sam, A. G., Gonzalez-Vega, C., & Chen, J. (2009). *Impact of Conditional Cash Transfers and Remittances on Credit Market Outcomes in Rural Nicaragua* (No. 49319). Agricultural and Applied Economics Association.

[8] D. Mikis Stasinopoulos and Robert A. Rigby (2007).Generalized Additive Models for Location Scale and Shape (GAMLSS) in R*Journal of Statistical software*.

[9] Takagi T., Kurokawa E, Miyata K., Okamoto, K., Tanaka, Y. & Kurokawa, K. (2002). The Comparison of Generalized Additive Model with Artificial Hierarchical Neural Network in the Analysis of Pharmaceutical Data Vol.3, ISSN 1345-8647. *Journal of Computer Aided Chemistry*, 56-62.

[10] Villarini, G., Smith, J. A., & Napolitano, F. (2010). Nonstationary modeling of a long record of rainfall and temperature over Rome. *Advances in Water Resources*, 33(10), 1256-1267.

[11] Tan, C. N. (1997). An Artificial Neural Networks Primer with Financial Applications Examples in Financial Distress Predictions and Foreign Exchange Hybrid Trading System. *Bond University*. 50-78

[12] Setodji, C. M., Le, V. N., & Schaack, D. (2013). Using generalized additive modeling to empirically identify thresholds within the ITERS in relation to toddlers' cognitive development. *Developmental psychology*, 49(4), 632.

[13] Serinaldi, F. (2011). Distributional modeling and short-term forecasting of electricity prices by generalized additive models for location, scale and shape. *Energy Economics*, 33(6), 1216-1226.

[14] OECD (2015), "How to restore a healthy financial sector that supports long-lasting, inclusive growth?", *OECD Economics Department Policy Notes, No. 27*, June 2015.

[15] Brown, R. P., & Carmignani, F. (2015). Revisiting the effects of remittances on bank credit: a macro perspective. Scottish Journal of Political Economy, 62(5), 454-485.

[16] Stasinopoulos, M., Rigby, B., & Akantziliotou, C. (2008). Instructions on how to use the gamlss package in R Second Edition.

[17] Sarle, W. S.(1994). "Neural Networks and Statistical Models", Proceedings of the Nineteenth Annual SAS Users Group International Conference, Cary, NC: SAS Institute, USA, pp. 1538-1550.

[18] Forouzanfar, M., Dajani, H. R., Groza, V. Z., Bolic, M., & Rajan, S. (2010, July). Comparison of feed-forward neural network training algorithms for oscillometric blood pressure estimation. In *Soft Computing Applications (SOFA), 2010 4th International Workshop on* (pp. 119-123). IEEE

[19] Maliki, O. S., Agbo, A. O., Maliki, A. O., Ibeh, L. M., & Agwu, C. O. (2011). Comparison of regression model and artificial neural network model for the prediction of electrical power generated in Nigeria. *Advances in Applied Science Research*, 2(5), 329-339.

[20] M.A. Pitt, IJ Myung and S.Zhang, 2002, *Psychol.rev* 109, 472.

[21] IJ Myung and M.A. Pitt, (2003), *Numerical Computer Methods, Part D(A volume of Methods in Enzymology)*.

[22] Zhang, D. D., Yan, D. H., Wang, Y. C., Lu, F., & Liu, S. H. (2015). GAMLSS-based nonstationary modeling of extreme precipitation in Beijing-Tianjin-Hebei region of China. *Natural Hazards*, 77(2), 1037-1053.

[23] Maciel, L. S., & Ballini, R. (2010). Neural networks applied to stock market forecast- ing: An empirical analysis. *Journal of the Brazilian Neural Network Society*, 8(1), 3-22.

[24] Sheela, K. G., & Deepa, S. N. (2013). Review on methods to fix number of hidden neurons in neural networks. *Mathematical Problems in Engineering*.

[25] Wagenmakers, E. J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic bulletin & review*, 11(1), 192-196.

[26] Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach.* Fourth Edition.Nelson Education.