Research Article

Long-Term Exchange Rate Probability Density Forecasting Using Gaussian Kernel and Quantile Random Forest

Samuel Asante Gyamerah ¹ and Edwin Moyo²

¹Pan African University, Institute for Basic Sciences, Technology, and Innovation, Nairobi, Kenya

²Mulungushi University, Kabwe, Zambia

Correspondence should be addressed to Samuel Asante Gyamerah; saasgyam@gmail.com

Received 26 August 2019; Revised 17 January 2020; Accepted 31 January 2020; Published 28 February 2020

Academic Editor: Hassan Zargarzadeh

Copyright © 2020 Samuel Asante Gyamerah and Edwin Moyo. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the midst of macro-economic uncertainties, accurate long-term exchange rate forecasting is crucial for decision-making and planning. To measure the uncertainty associated with exchange rate and obtaining additional information of future exchange rate, a hybrid model based on quantile regression forest and Gaussian kernel (GQRF) is constructed. Quarterly dataset of KSh/USD exchange rate and macro-economic variables from 2007 to 2016 are used. The forecast horizon spans from 2013 to 2016. With a prediction interval coverage probability and prediction interval average width of 95% and 29.6493%, the constructed model has a very high coverage probability. The method of determining the probabilistic forecasts is very significant to achieve forecasts with correct coverage. The probability density forecasting model for the exchange rate gave significant information—the probability distribution of the forecasted results. In this way, uncertainties around the forecast can be evaluated because the complete exchange rate distribution are forecasted. GQRF is efficient as it can uphold the uncertainty about the variance linked to each point, which is important for exchange rate forecasting. Using the constructed model, the probabilities of exceedance such as the probability of future exchange rate exceeding the average exchange rate for the year can be computed. This paper also adds to the scarce literature of exchange rate probability density forecasting using machine learning techniques.

1. Introduction

Exchange rate forecasting is truly a challenging task and continues to be a very vital research area for financial institutions and economists, foreign currency hedgers, speculators, traders, and all professionals in the foreign exchange (forex) market. Central banks, governments, and stakeholders in the forex market take into accounts exchange rate forecasting to make critical and important economic decisions. These decisions impact on the future movements of a country's economy. The expected value of exchange rate influences cash flows of all foreign transactions. Hence, forecasting the movement of exchange rate is beneficial to investors, trade, businesses, and policy makers. Most multinational corporations (MNCs) use exchange rate forecast to decide whether to hedge future payables and receivables which are in foreign currencies. MNCs also use exchange

rate forecasts for capital budgeting decisions, earning assessment, short- and long-term financing decisions. However, there is significant doubt about the likelihood of accurate forecasts because of the uncertainty of the major macro economic variables used in forecasting exchange rate. This has resulted in exchange rate uncertainty in the forex market leading to significant effects in the trade (international and domestic) sector and domestic investments of most countries.

As indicated by [1], exchange rate uncertainty for medium-term period affects trade flows of most industrial countries with the exception of the United States of America. Reference [2] examined the effects of exchange rate uncertainty on firm's growth rate to profits and concluded that an increase in the uncertainty of exchange rate results in a greater variability in the growth rate of the firm's profit. Using Croatia and Cyprus as a case study, [3] analysed the

effect of exchange rate uncertainty on sectoral exports. The authors indicated that there were significant negative effects of exchange rate uncertainty on exports. The relationship between Japanese firms' exposure to exchange rate uncertainty and their risk management practice was examined by [4]. It was concluded that firms with considerable dependency on sales in foreign markets have significant foreign exchange exposure that is determined by the market. Reference [5] studied the impact of exchange rate uncertainty on exports for some selected sub-saharan African countries (The Gambia, Ghana, Kenya, Madagascar, Mauritius, Mozambique, Nigeria, Sierra Leone, Tanzania, Uganda, and Zambia) from 2013 to 2014. Their study revealed that, there was a negative effect of exchange rate uncertainty on exports in the short-term and a positive effect in the long-term period.

Different from the above literatures are literatures that have examined the effects of exchange rate fluctuations on domestic investments [6–8]. Exchange rate uncertainties have two contrasting impacts on the economy and the gross future prospects of investment on the economy of a country [9]. That is, imports become comparatively inexpensive as in contrast with exported goods. Consequently, the marginal profit of investing an extra unit of capital is certain, owing to the lower revenue generated by domestic firms operating locally and internationally. In their studies, [6] found that exchange rate fluctuations indirectly affects domestic investment because of its effect on domestic and international trade. Reference [7] investigated the effects of exchange rate uncertainties on foreign direct investment in some selected countries in Sub Saharan Africa. They observed a strong negative and significant effect of exchange rate uncertainty on foreign direct investment. That is, when investors are uncertain about the performance of a country's currency in the long term, they are incline in moving their investment out of that country to other stable markets. In examining the impacts of exchange rate uncertainty on domestic investment for some selected emerging markets and developing economies (EMDE), [8] found that there was a positive and significant impact of exchange rate uncertainty on domestic investments for these EMDEs.

The above literatures confirm the effects of exchange rate uncertainties on domestic investment, foreign direct investment, trading activities of industries, growth of industries, international and domestic trade, and the economy of a country as a whole. It is therefore imperative to accurately forecast exchange rate in the midst of uncertainties to help in long- and short-term decision making for key stakeholders of the economy and for investors. Using quarterly exchange rate and other macro-economic variables data of Kenya as a case study, we propose a probability density forecasting model that is able to capture the uncertainties associated with exchange rate. The constructed model is a hybrid model of quantile regression forest and Gaussian kernel (GQRF). Quantile regression forest has been used in the area of health and agriculture. Reference [10] used quantile regression forest to predict drug response treatment of cancer patients and evaluate the prediction reliability. Using out-of-bag validation, they asserted that their QRF model attained a

higher prediction accuracy of the drug response. Recently, [11] applied a hybrid model of quantile regression forest and Epanechnikov Kernel function to capture weather uncertainties in crop yield forecasting. Their model performed well in predicting crop yields in the midst of extreme and uncertain weather conditions.

The contributions made in this paper are as follows: (1) a reliable, efficient, and accurate probability density forecasting model using quantile regression forest and Gaussian kernel is proposed and implemented to capture the uncertainty of exchange rate. The total conditional probability density curve of exchange rate is constructed and all the selected observed exchange rate values lies within the forecasted probability density curve. (2) Using the proposed model, the probabilities of exceedance can be computed, such as the probability of future exchange rate exceeding the average exchange rate for the year. To the best of our knowledge, the proposed probability density forecasting model is the first of its kind in literature.

The paper is structured as follows: Section 2 presents the theoretical concepts of the paper. An overview of quantile regression, random forest, and the proposed model (quantile regression forest and kernel density estimation) is presented in this section. Section 3 provides the evaluation metrics used to evaluate the performance of the point and interval predictions. In Section 4, a case study using exchange rate between United States dollars (USD) and Kenya Shillings (KSh) and selected macro-economic variables are presented. Data analysis and evaluation results are presented in this section. The conclusion is presented in Section 5.

2. Theoretical Concepts

2.1. Quantile Regression. Conventional regression techniques draw the inferences of the relationship between real-values response variable Y and a predictor variable X . It then finds an estimate $\hat{\mu}(x)$ of the conditional mean $\mathbb{E}(Y | X = x)$ of the predictor variable $Y|X = x$. In contrast to conventional regression techniques, Quantile regression (QR) does not assume a particular parametric distribution for the response variable, nor assume a constant variance for the response variable. Hence, the introduction of Quantile regression. Quantile regression (QR) is a statistical approach of estimating and making inferences on conditional quantile of the response variable [12].

The standard QR problem can be defined as in

$$Q_{\alpha}(\alpha | \mathbf{X}_i) = \mathbf{X}_i' \mathbf{b}_{\alpha}, \quad (1)$$

where $\mathbf{X}_i$ are the predictors/features, $\mathbf{b}$ is a vector of parameters for quantile α , and $Q_{\alpha}(\alpha | \mathbf{X}_i) = \inf\{y: f_{\alpha}(y | \mathbf{X}_i) \geq \alpha\}$ is the conditional α_{th} quantile of the exchange rate distribution (y_i). From equation (1), the conditional distribution function $F(y | \mathbf{X} = x)$ is explained as the probability that Y is smaller than or equal to $y \in \mathbb{R}$, that is

$$F(y | \mathbf{X} = x) = P(Y \leq y | \mathbf{X} = x). \quad (2)$$

By minimizing the loss function for a particular α_{th} quantile, the vector of parameters are evaluated,

$$\begin{aligned}
\min_{b_\alpha} \sum_{i=1}^n \rho_\alpha(y_i - \mathbf{X}_i' \mathbf{b}_\alpha) &= \min_b \left[\sum_{i: y_i \geq \mathbf{X}_i' \mathbf{b}_\alpha} \alpha |y_i - \mathbf{X}_i' \mathbf{b}_\alpha| \right. \\
&\quad \left. + \sum_{i: y_i < \mathbf{X}_i' \mathbf{b}_\alpha} (1 - \alpha) |y_i - \mathbf{X}_i' \mathbf{b}_\alpha| \right] \\
&= \min_{b_\alpha} \left[\sum_i |\alpha - 1_{y_i < \mathbf{X}_i' \mathbf{b}_\alpha}| (y_i - \mathbf{X}_i' \mathbf{b}_\alpha) \right],
\end{aligned} \tag{3}$$

where n is the total number of the sample data, 1 is the indicator function, and $\mathbf{X}_i = [x_{1i}, x_{2i}, x_{3i}, \dots, x_{ki}]$ are the independent variables.

2.2. Random Forest. Random forest (RF) is an example of machine learning (ML) ensemble algorithm called bagging (an abbreviation of “bootstrap aggregating”). In bagging, the learner comprises of an ensemble of distinct base learners and each base learner is trained with a bootstrap sample for every tree from the entire observation. That is, bagging takes M subsamples with replacement from the training data and train the model on those subsamples. This ensures that every tree in the forest is built on about 63% of the available training data, and the remaining data is used for testing. Hence, RF has no need of an explicit test sample. The remaining data is called out-of-bag (OOB) data. The performance of each model on the OOB data when averaged can give an estimated accuracy of the bagged models. Predictions made using the OOB data are called OOB predictions. The principal advantage of bagging is that it integrates the regularization in it and only needs an appropriate choice of parameters or the base algorithms. Bootstrapping techniques such as random forest is appropriate for small sample datasets having the likelihood to overfit [13].

Now assume that $T = \{y_i, \mathbf{x}_i\}_{i=1}^N$ is the set of training data and a technique for building a predictor $Q(\mathbf{x}, T)$ using the given training data, where $y_i \in \mathbb{R}$ is the target variable and $\mathbf{x}_i \in \mathbb{R}^d$ is the set of predictors. In bagging, a sequence of training datasets $T_{B,1}, T_{B,2}, \dots, T_{B,P}$ are created of equal size as T by bootstrapping from T . P predictors are built in a way that the p_{th} predictor $Q(\mathbf{x}, T_{p,B})$ depends on the p_{th} bootstrap training data. As long as these predictors are aggregated, the resultant predictor can be significantly more accurate than the actual predictor [14].

2.3. Gaussian Quantile Regression Forest (GQRF)

2.3.1. Quantile Regression Forest. Quantile regression forests (QRF) is a consistent algorithm that presents a non-parametric and explicit technique of evaluating conditional quantiles for high dimensional predictor variables [15]. Quantile regression forest is a Machine learning technique that is based on random forest and quantile regression. The prediction of random forest can be likened to the weighted mean of the actual response variables. Similar to random forest, trees are grown in quantile regression forests. Using

the weighted distribution of the actual response variable, the conditional distribution is evaluated. The assigned weights for the actual response variable are similar to random forest algorithm.

Consider the conditional distribution of the actual response variable Y given predictor variable $X = x$ based on a the tree $T(\vartheta)$. ϑ is defined as the random parameter that decides the growth of a tree. Now, assume that the leaf containing x is represented as $\ell_n(x, \vartheta)$. It follows that the weight $\omega_i(x, \vartheta)$ is specified as

$$\omega_i(x, \vartheta) = \frac{1_{X_i \in \ell(x, \vartheta)}}{\left\{j: X_j \in \ell(x, \vartheta)\right\}}. \tag{4}$$

Given that the T trees of the random forest can be represented as ϑ_T and $\omega_i(x)$ is the average of $\omega_i(x, \vartheta)$ over each and every single tree. It follows that

$$\omega_i(x) = \frac{1}{T} \sum_{t=1}^T \omega_i(x, \vartheta_t). \tag{5}$$

The estimate $\hat{F}(y | X = x)$ is defined as

$$\hat{F}(y | X = x) = \sum_{i=1}^n \omega_i(x) 1_{Y_i \leq y}. \tag{6}$$

The estimate of the conditional quantile $Q_\alpha(x)$ for $0 < \alpha < 1$ can be computed as

$$Q_\alpha(x) = \inf\{y: \hat{F}(y | X = x) \geq \alpha\}. \tag{7}$$

2.3.2. Kernel Density Estimation. The primary kernel density estimator can be written in a compact form as [16]:

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) = \frac{1}{n} \sum_{i=1}^n K_h(x - x_i), \tag{8}$$

where $K_h(\cdot) = K(\cdot/h)/h$ and it is called the scaled kernel. The statistical analysis of kernel estimators is more basic than histograms. Normally,

$$K(x_1, x_2, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x_1 - x_2)^2 / (2\sigma^2)}, \tag{9}$$

where $\sigma > 0$ is the bandwidth or smoothing parameter. A smaller bandwidth gives much more wobbly distribution and a substantial bandwidth loose much more details. Plug-in methods, Cross validation, rule of thumb are all types of bandwidth selectors. Different kernels such as triangular, triweight, cosine, uniform, Gaussian, Epanechnikov, and others can be used for kernel density estimation. In this study, we use the Gaussian kernel type of kernel density estimation because it can compute the bandwidth by a rule of thumb automatically [16]. To generate a total probability density, the density estimation is repeated for different values of x .

This paper constructs a hybrid model for probability density forecasting. First, for out-of-bag prediction, the keep.inbag parameter of the QRF model is set to “True” and node size was set to 5. The obtained conditional quantiles are used as the inputs of kernel density estimation. The Gaussian

kernel function is combined with the normal reference density (“nrd0”) bandwidth selection method to implement the probability density forecasting. Figure 1 gives the realization flowchart of the constructed hybrid model for the probability density forecasting.

3. Evaluation Metrics

3.1. Evaluation Metrics for Point Prediction Errors. Different evaluation metrics have been used to estimate the accuracy and efficiency of point forecast. Nonetheless, in this paper, we use the mean absolute percentage error (MAPE), root mean square error (RMSE), and R-squared (R^2). The smaller the values of MAPE and RMSE, the better the prediction model. The higher the value of R^2 , the higher the the goodness of fit. MAPE, RMSE, and R^2 are mathematically defined in equations (10)–(12) as follows,

$$\text{MAPE}(y_i, p_i) = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - p_i}{y_i} \right| \times 100\%, \quad (10)$$

$$\text{RMSE}(y_i, p_i) = \sqrt{\frac{\sum_{i=1}^n (y_i - p_i)^2}{n}}, \quad (11)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - p_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2}, \quad (12)$$

where i is quarterly exchange rate data, n is the aggregated quarterly exchange rate data over the forecasted period, y_i and p_i is respectively the i_{th} actual and the predicted exchange rate.

3.2. Evaluation Metrics for Prediction Intervals. Different metrics can be used to estimate the efficiency of prediction intervals. Prediction interval coverage probability and prediction interval normalized average deviation are the metrics that will be use to evaluate the reliability of prediction intervals in this paper.

Prediction interval coverage probability (PICP) is defined as in equation (13)

$$\text{PICP} = \frac{1}{n} \sum_{i=1}^n 1_i, \quad (13)$$

where 1_i is a Boolean variable determined as

$$1_i = \begin{cases} 1, & \text{if } y_i \in [L_i, U_i], \\ \text{otherwise,} & \text{if } y_i \notin [L_i, U_i], \end{cases} \quad (14)$$

and n is the total number of forecast. The higher the percentage of PICP, the better. That is, a larger PICP means a greater number of the actual values are covered by the forecasted prediction interval. However, the value of PICP can be increased by increasing the width of the prediction interval. The widths of prediction intervals should be measured to ensure the quality of prediction intervals. A narrow prediction interval gives a more factual information than a wider prediction interval. Prediction interval normalized average width (PINAW) is defined as follows:

$$\text{PINAW} = \frac{1}{n(\max(y_i) - \min(y_i))} \sum_{i=1}^n (U_i - L_i), \quad (15)$$

where $\min(y_i)$, $\max(y_i)$ are the minimum and maximum of the actual/target values.

4. Case Study

4.1. Data Selection and Description. The performance of exchange rate is controlled by different economic factors such as: stock price returns, consumer price index, interest rate, amount supply, foreign direct investment, international and domestic trade, real gross domestic product, consumer price index. There are complex interactions between these factors, hence it is very complex to accurately predict exchange rate. Following the work of [17], we selected the macro-economic variables (features) that consider several dimensions of interests like consumer price index, interest rate, money supply, foreign direct investment, real gross domestic product, and consumer price index. Table 1 presents the definition of the variables used in this study and the function of the the variable in the proposed model. Quarterly secondary data of money supply, foreign exchange rates, and weighted lending interest rates were sourced from the Kenyan Central Bank. Quarterly secondary data of foreign direct exchange, consumer price index (inflation), and gross domestic product were also taken from the Kenyan National Bureau of Statistics. Summary statistics for the sampled quarterly period from 2007 to 2016 is given in Table 2. From the table, the average of the foreign exchange rate, CPI, lending rate by commercial banks, money supply, foreign direct investments for the sample period under study were 83.81, 125.47, 15.79, 824287.6, 1513157, and 525046.5 respectively. From the nonparametric locally estimated scatterplot smoothing (loess) curve (dark grey line in the plots) fitted to the data in Figure 2, there is an upward trend in all the variables over the period of interest.

Anderson-Darling and Shapiro-Wilk test for normality are used to test the normality of the residuals of exchange rate data. From Table 3, the p value of EX residuals is less than the significance level of 0.05. As a result, we can reject the null hypothesis of normality and conclude that distribution of residuals data for EX are significantly different from the normal distribution. The departure of normality of EX is visually presented in Figures 3(a) and 3(b) where the density plot, box plot and normal quantile-quantile (QQ) plot (Figure 4) are presented. For instance, there were case of outliers in the boxplot and the residuals data is skewed. Using Non-Constant Error Variance (NCV) and Breusch Pagar tests for residuals homogeneity, the p value was found to be less than the significance level of 0.05 in both test. We can therefore reject the null hypothesis of constant error variance (homogeneity) and conclude that the variance are heteroscedastic (non-constant variance). From above, it is clear that some major assumptions of Ordinary least squares (OLS) regression model are violated. We can

FIGURE 1: The flow diagram of constructed model.

TABLE 1: Definition and function of economic variables.

Economic variable	Definition	Function
EX	Foreign exchange, a measure of exchange rate	Predictand
CPI	Consumer price index, a measure for inflation	Predictor
IR	Interest rate, a measure of the weighted average lending interest rate	Predictor
Real GDP	Real gross domestic product, a measure of economic growth	Predictor
MS	Money supply, as a measure of money in circulation	Predictor
FDI	Foreign direct investment, the natural log of foreign direct exchange inflows	Predictor

therefore employ a hybrid model using quantile regression technique and random forest to reveal the heterogenous effect of the features on the target variable (exchange rate) for different quantiles.

4.2. Data Normalization. The numeric values of the datasets are changed to a common scale without distorting the differences in the range of the values using a min-max normalization. The normalized values lie within the range

TABLE 2: Summary statistics of the variables.

	Min	Max	Mean	Std	Skewness	Kurtosis
EX	62.65	103.89	83.81	11.49	-0.04	-0.85
CPI	78.46	175.18	125.47	29.04	0.01	-1.30
IR	12.87	20.34	15.79	2.07	0.68	-0.52
Real GDP	633710	1094567	824287.6	134419.60	0.44	-0.98
MS	557650	2761800	1513157	691402	0.32	-1.04
FDI	77732.68	1766730	525046.50	541927.30	0.87	-0.64

FIGURE 2: Trend of the variables. (a) EX. (b) CPI. (c) IR. (d) Real GDP. (e) MS. (f) EDI.

TABLE 3: Normality test of the variables.

Feature	Anderson-Darling		Shapiro-Wilk	
	Test statistics	p value	Test statistics	p value
EX	2.5511	1.447×10^{-06}	0.7960	6.949×10^{-06}

FIGURE 3: (a) Density and (b) box plot of the quarterly exchange rate.

FIGURE 4: Normal QQ plot of exchange rate from January, 2007 to December, 2016.

$[0, 1]$, and is computed using equation (16). The normalized values are changed to the magnitude of actual/observed data using the antinormalization equation (see equation (17)).

$$x_{\text{norm}} = \frac{x - \min(x)}{\max(x) - \min(x)}, \quad (16)$$

$$x = x_{\text{norm}} (\max(x) - \min(x)) + \min(x), \quad (17)$$

where x_{norm} is the normalized input value, $\min(x)$ and $\max(x)$ are respectively the minimum and maximum value of the inputs.

R statistical software was used in implementing all the forecasting approaches.

4.3. Experimental Results and Discussion. The efficiency of the proposed Gaussian kernel quantile regression forest using the quarterly data described in Section 4.1 is determined in this section. The probability density of Kenya's exchange rate are predicted given the values of the selected macro-economic indicators from a scenario setting.

Figure 5 presents the visual prediction results for the out-of-bag (OOB) prediction. The OOB prediction values were obtained through the implementation of bootstrapping. The "quantregForest" package [18] of R statistical software was used. For optimal performance of quantile regression forest in this paper, we chose 1100 trees ($T=1100$). Figure 6 shows the mean square error and the r -squared plot. From the

FIGURE 5: Plot of actual/observed exchange rate and forecasts.

FIGURE 6: (a) Mean square error and (b) R -squared.

figure, mean square error and r -squared were at their optimum value of approximately 0.005 and 0.92 at 200 trees. This indicates that, the optimal number of trees was 200. The number of features selected from the entire set of features for each node of the regression tree was 2. The out-of-bag prediction and node size parameters were set to “True” and 5 respectively. Evaluation metrics for point forecast using the mean and median are presented in Table 4. The MAPE value for median prediction (0.0217) is smaller than the MAPE value of the mean prediction (0.0265). The RMSE value of the mean prediction is 3.119 as compared to 2.7146 for median prediction. With a R -squared value of 0.9428, the point forecast of the exchange rate using the median fits the data relatively better than the mean. From Table 4, it can be concluded that, point forecast of exchange rate using the median is significantly better than point forecast using the mean. Generally, the fitness of model to the data is very high in all cases. This is evident in the R -squared values of 0.9248 and 0.9248 recorded by both the point forecast of the mean and median.

Figure 7 shows the quarterly exchange rate for the sample period under study. For each quarterly period, the red dots represents observed quarterly exchange rate. The black line represents forecasted exchange rate at 99.5%

percentile. This is the line above which 0.5% of the forecasted exchange rate would fall and below which 99.5% of all forecasted exchange rate are expected to fall. The OOB predictions and prediction intervals based on the observed exchange rate shows that the median prediction always fall within the prediction intervals. 95% of observed exchange rate values also falls within the prediction interval. This shows that QRF can accurately forecast exchange rate.

The main conclusion that can be achieved by visualizing the prediction intervals in Figure 8 is the reliability of the prediction, as the length of the prediction intervals differs strongly. We can conclude that for majority of the observations, the prediction is more reliable than for others (only two observations). Table 5 (see appendix) presents the out-of-bag forecasting results, the relative percentage error and the forecasting error of GQRF using the Exchange rate quarterly data.

To evaluate the performance of the proposed Gaussian quantile regression forest, we computed the values of PICP and PINAW. Table 6 presents the values of PICP and PINAW. PICP recorded a high percentage value of 95% and PINAW 29.6493%. This indicates that the constructed prediction interval capture the actual exchange rate values very well. The small value of the PINAW also proofs the

TABLE 4: QRF point forecast evaluation metrics for out-of-bag predictions.

	MAPE (%)		RMSE (USD/KSh)		R-squared	
	Mean	Median	Mean	Median	Mean	Median
QRF	0.0265	0.0217	3.1119	2.7146	0.9248	0.9428

FIGURE 7: Prediction results and prediction intervals based on Gaussian kernel QRF for the period of study.

FIGURE 8: Prediction intervals ordered according to their length for 40 OOB predictions.

TABLE 5: Out-of-bag forecasting results of GQRF using the quarterly exchange rate data.

Quarterly actual EX (USD/KSh)	Forecasting EX	Median		Mean	
		Relative percentage error (%)	Forecasting error (USD/KSh)	Forecasting EX	Relative percentage error (%)
69.60	67.88	2.47	-1.72	67.86	2.50
69.16	68.35	1.17	-0.81	67.79	1.98
68.35	67.88	0.68	-0.47	67.71	0.93
67.45	68.35	1.33	0.9	68.30	1.26
65.93	67.45	0.63	-0.43	69.08	1.76
63.03	67.45	2.30	1.52	66.77	1.27
63.03	66.69	5.81	3.66	68.04	7.95
62.65	67.45	7.66	4.8	68.84	9.88
79.58	79.25	0.41	-0.33	79.03	0.69
79.81	78.94	1.09	-0.87	76.27	4.44
79.25	78.94	0.39	-0.31	77.46	2.25
78.45	79.25	1.02	0.8	75.51	3.75
76.49	79.58	4.04	3.09	82.98	8.48
76.98	77.58	0.78	0.6	78.69	2.23

TABLE 5: Continued.

Quarterly actual EX (USD/KSh)	Forecasting EX	Median		Forecasting EX	Mean	
		Relative percentage error (%)	Forecasting error (USD/KSh)		Relative percentage error (%)	Forecasting error (USD/KSh)
77.58	78.45	1.12	0.87	78.56	1.26	0.98
78.94	77.58	1.72	-1.36	77.89	1.33	-1.05
82.21	84.69	3.01	2.475	83.23	1.24	1.02
86.33	88.49	2.50	2.16	87.57	1.44	1.24
94.85	86.33	8.98	-8.52	86.01	9.31	-8.84
91.52	84.61	7.55	-6.91	84.59	7.57	-6.93
83.54	85.35	2.16	1.805	85.90	2.82	2.36
84.76	84.61	0.18	-0.15	85.22	0.54	0.46
84.61	84.98	0.44	0.37	85.97	1.60	1.36
85.71	84.76	1.11	-0.95	85.72	0.01	0.01
86.50	86.33	0.20	-0.17	87.44	1.09	0.94
84.98	86.50	1.79	1.52	86.90	2.26	1.92
87.17	85.71	1.67	-1.46	85.74	1.64	-1.43
86.15	86.33	0.21	0.18	86.29	0.16	0.14
86.33	90.04	4.30	3.71	92.44	7.08	6.11
87.43	88.49	1.21	1.06	89.01	1.81	1.58
88.49	87.43	1.20	-1.06	88.96	0.53	0.47
90.04	88.49	1.72	-1.55	88.59	1.61	-1.45
91.81	90.04	1.93	-1.77	91.24	0.62	-0.57
97.01	91.81	5.36	-5.2	94.23	2.86	-2.78
103.89	97.01	6.62	-6.88	97.02	6.61	-6.87
102.08	101.73	0.34	-0.35	101.31	0.75	-0.77
101.90	101.73	0.17	-0.17	101.30	0.59	-0.60
101.04	101.90	0.85	0.86	101.55	0.51	0.51
101.34	101.73	0.38	0.39	101.31	0.03	-0.03
101.73	101.34	0.38	-0.39	100.17	1.53	-1.56

TABLE 6: Prediction interval and point forecast evaluation metrics for quantile random forest Gaussian kernel.

PICP (%)	PINAW (%)	MAPE (%)		RMSE		R-squared	
		Mean	Median	Mean	Median	Mean	Median
95	29.6493	0.7136	0.7136	0.0056	0.0056	0.7473	0.7473

FIGURE 9: Probability density curve using GQRF. (a) 2013. (b) 2014. (c) 2015. (d) 2016.

efficiency of the the constructed prediction interval. For the purpose of illustration, the probability density curve plot for four years are presented in Figure 9. The probability density

curves gave full probability distribution of future exchange rate. For each density plot, the average of the quarterly observed data for the year are capture with a red line in the

plot. The average of the quarterly exchange rate always fall within the probability density curve. It can therefore be concluded that the proposed GQRF model can accurately depict exchange rate uncertainty.

The forecasted quarterly density exchange rate (see Figure 9) can be used to compute the probabilities of exceedance, for instance, the probability that the exchange rate will be more than the average exchange rate of 98.6975 USD/KSh for the year 2015. This exceedance probability can be computed from the density curve and it comprises of all exchange rate higher than 100 USD/KSh. The area of the density curve is calculated by finding the integral of the density function in Figure 9. The probability of exchange rate exceeding the observed average exchange rate 98.6975 USD/KSh in 2015 is 0.2899. The probability of exchange rate to exceed 101.5 USD/KSh in 2015 is 0.1584 and 0.7530 in 2016. This means it is less likely for exchange rate to exceed 101.5 in 2015 than in 2016.

5. Conclusion

Considering a long-term exchange rate forecasting problem, this study proposes a novel forecasting method which has the propensity to draw total conditional probability density curve of future exchange rate. Quantile regression forest is used to build a non-linear quantile regression probabilistic forecasting model. Direct plug-in bandwidth selector and Gaussian kernel selector are used to draw probability density curve. The significance of the proposed Gaussian quantile regression forest technique is its ability to capture information uncertainty of future exchange rate using probability density forecasting functions. Using a case study of the exchange rate between United States Dollar (USD) and Kenya Shilling (KSh), we were able to show that the proposed model performs accurately well. The values of PICP and PINAW revealed the high performance and accuracy of the prediction interval built by Gaussian quantile regression forest. The probability of exceedance for a given exchange rate value can also be computed using the proposed model.

Data Availability

Data for this work are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The first Author will like to thank Pan African University, Institute for Basic Sciences, Technology, and Innovation for funding this research.

References

- [1] E. Perée and A. Steinherr, "Exchange rate uncertainty and foreign trade," *European Economic Review*, vol. 33, no. 6, pp. 1241–1264, 1989.
- [2] C. F. Baum, M. Caglayan, and J. T. Barkoulas, "Exchange rate uncertainty and firm profitability," *Journal of Macroeconomics*, vol. 23, no. 4, pp. 565–576, 2001.
- [3] D. Serenis and N. Tsounis, "Exchange rate volatility and foreign trade: the case for Cyprus and Croatia," *Procedia Economics and Finance*, vol. 5, pp. 677–685, 2013.
- [4] T. Ito, S. Koibuchi, K. Sato, and J. Shimizu, "Exchange rate exposure and risk management: the case of Japanese exporting firms," *Journal of the Japanese and International Economies*, vol. 41, pp. 17–29, 2016.
- [5] B. Senadza and D. D. Diaba, "Effect of exchange rate volatility on trade in sub-Saharan africa," *Journal of African Trade*, vol. 4, no. 1-2, pp. 20–36, 2017.
- [6] L. S. Goldberg, "Exchange rates and investment in United States industry," *The Review of Economics and Statistics*, vol. 75, no. 4, pp. 575–588, 1993.
- [7] E. E. Asmah and F. K. Andoh, "Exchange rate volatility and foreign direct investment in sub-Saharan Africa," *Journal for the Advancement of Developing Economies*, vol. 2, no. 2, 2013.
- [8] D. R. Van Deventer, K. Imai, and M. Mesler, *Advanced Financial Risk Management: Tools and Techniques for Integrated Credit Risk and Interest Rate Risk Management*, John Wiley & Sons, Hoboken, NJ, USA, 2013.
- [9] T. M. Harchaoui, F. Tarkhani, T. Yuen et al., *The Effects of the Exchange Rate on Investment: Evidence from Canadian Manufacturing Industries*, Bank of Canada Ottawa, Ottawa, USA, 2005.
- [10] Y. Fang, P. Xu, J. Yang, and Y. Qin, "A quantile regression forest based method to predict drug response and assess prediction reliability," *PLoS One*, vol. 13, Article ID e0205155, 2018.
- [11] S. A. Gyamerah, P. Ngare, and D. Ikpe, "Crop yield probability density forecasting via quantile random forest and epa-nechnikov kernel function," 2019, <https://arxiv.org/abs/1904.10959>.
- [12] R. Koenker and G. Bassett Jr., "Regression quantiles," *Econometrica*, vol. 46, no. 1, pp. 33–50, 1978.
- [13] T. Shaikhina, D. Lowe, S. Daga, D. Briggs, R. Higgins, and N. Khovanova, "Decision tree and random forest models for outcome prediction in antibody incompatible kidney transplantation," *Biomedical Signal Processing and Control*, vol. 52, pp. 456–462, 2019.
- [14] L. Breiman, *Out-of-Bag Estimation*, University of California, Berkeley, CA, USA, 1996.
- [15] N. Meinshausen, "Quantile regression forests," *Journal of Machine Learning Research*, vol. 7, pp. 983–999, 2006.
- [16] D. W. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization*, John Wiley & Sons, Hoboken, NJ, USA, 2015.
- [17] M. Muchiri, *The Effect of Interest Rate and Inflation Rate on Exchange Rates in Kenya*, University of Nairobi, Nairobi, Kenya, 2015.
- [18] N. Meinshausen and M. Nicolai-Meinshausen: The quantreg-Forest Package, 2007.