Journal of Statistics Applications & Probability
*An International Journal*

http://dx.doi.org/10.18576/jsap/110109

# Prediction of Foreign Direct Investment: an Application to South African Data

*Nyawedzeni Netshivhazwaulu, Caston Sigauke* and Alphonce Bere*

Department of Statistics, University of Venda, Private Bag X5050, Thohoyandou 0950, Limpopo, South Africa

Received: 30 Oct. 2020, Revised: 21 Dec. 2020, Accepted: 16 Jan. 2021
Published online: 1 Jan. 2022

**Abstract:** Foreign direct investment is considered as a vehicle for transferring new ideas, capital, superior technology and skills from developed countries to developing countries. Kernel quantile regression is used in this study to estimate the relationship between foreign direct investment and the factors influencing it in South Africa, using data for the period 1996 to 2015. Using the least absolute shrinkage and selection operator technique, all the variables were selected to be in the models. The developed kernel quantile regression models were used for forecasting the future inflow of foreign direct investment in South Africa. The forecast evaluation was done on all the models and the model based on the ANOVA radial basis kernel was selected as the best in terms of the accuracy measures (mean absolute percentage error, root mean square error and mean absolute error). The forecasts from the individual models were then combined using linear quantile regression averaging. The kernel quantile regression model using an ANOVA radial basis kernel was found to be the best model for forecasting foreign direct investment in South Africa. Accurate forecasts of FDI aid in economic planning. Identification of key drivers of FDI inflow can assist in crafting strategies to attract more FDI.

**Keywords:** Foreign direct investment, kernels, kernel quantile regression, linear quantile regression, predictive ability.

## 1 Introduction

### *1.1 Context*

Foreign Direct Investment (FDI) involves an investment from one country to another ([1]). FDI plays an important role in the transfer of skills, technology and capital from mature economies to underdeveloped countries. Whether or not FDI has a positive impact on the economy of a country depends on several conditional impact factors. These include government regulations and policies on investment, cost and ease of doing business, availability of raw materials, trade facilitation instruments and appropriate human resources, economic growth, political stability and security of the country ([2]). South Africa (SA) has been pursuing a set of economic development strategies to promote national economic development. The driving force behind the strategies is to reduce the most challenging problems facing the SA economy; employment creation, economic growth, promoting the program of radical economic transformation and reducing poverty ([3]).

The transition to multi-racial democracy in 1994 has presented difficult political, social, and economic challenges since the new government was concerned with the establishment of a credible and prudent fiscal stance, efforts to reduce inflation, and the needed reunification of the dual exchange rate system ([4]). SA's interesting achievements in overcoming these challenges have been widely recognised and the policy perseverance shown over the past years has yielded tangible macro stabilisation success and enhanced policy legitimacy. The growth and employment challenges facing SA are a cause for concern since the investment rates are low, FDI inflows disappointing, leaving SA at a disadvantage within an increasingly competitive global environment. Many statistical models on determining the factors that explain the variation in FDI and the impact of FDI on economic growth have been suggested. Some of the researchers go further to make forecasts using the best fit models.

* Corresponding author e-mail: caston.sigauke@univen.ac.za

This study uses kernel quantile regression (KQR) modelling to determine the important determinants of FDI in SA and the nature of their impact on FDI. The variables considered as potential determinants of FDI inflow are economic growth (GDP), Consumer price index (CPI), political stability (PS), labour productivity index (LPI), gross fiscal capital formation (GFCF). SA has implemented various economic reforms to restructure the economy to archive higher economic growth and development. The proposed kernel quantile regression model will be used to forecast the FDI inflow for the years to come. The forecasts obtained from this study will provide the SA government with information on the basis of which it can craft strategies for attracting future FDI. Accurate FDI forecasts play an important role in assisting policymakers and decision-makers to come up with good policies and suitable strategic plans to promote FDI.

### *1.2 An overview of the literature on foreign direct investment forecasting*

In the middle of the 1980s, the world FDI started to increase sharply and has received a lot of attention ([5]). For many years different models for modelling and forecasting FDI inflows were developed, applied, reviewed and published. This section provides an overview of FDI, factors affecting it as well as the summaries of some studies that used the proposed methodology to forecast FDI inflow in different countries.

FDI is one of the most dynamic international resources flows to developing countries. In [6], it is highlighted that FDI can affect growth and development by complementing domestic investment, transfer of knowledge and technology, and facilitating trade. Each country needs FDI to leap itself to sustainable growth levels and fill the saving and foreign exchange gaps.

Previous work has looked at the correlation of FDI with several determinants. Some that might be thought to have a connection to FDI flows are political stability and security of a country, market size and growth potential, government regulations and policies on investment, economic stability, clustering effect, openness to trade (OT), availability of appropriate human resources, wage rates, availability of raw materials including commodities among others. It is noted by [7] that countries with larger populations have a high affinity for attracting more FDI. It is suggested in [8] that large economies attract more FDI.

It has been found that political stability is one of the key determinants of potential FDI inflows. In [9] it is argued that political instability significantly affects FDI inflows. A weaker exchange rate is expected to increase FDI inflow, as firms take advantage of low prices in host markets to purchase facilities ([10]). Policies on OT might produce a significant impact in attracting FDI.

Different models have been developed and used to forecast the future inflow of FDI in different countries. Among these are the Box-Jenkins methodolody ([11], [12]) and the quantile regression (QR) models ([13]).

The QR models have been used in many application areas including forecasting. In a study by [14], a QR model was applied to determine the impact of FDI on technology innovation from 1987 to 2009 in China. The main aim was to analyze the correlation between FDI and technology innovation. They found that FDI exerts significant positive effects only for low technological innovation level at different locations of the conditional distribution. The instrumental variable quantile regression (IVQR) model is used by [15] on country-level data to examine the role of absorptive capacity in contributing to the growth effects of FDI at different quantiles of the income distribution. They found that the IVQR estimates positive growth effects of FDI at higher quantiles of the distribution, while the estimates of FDI for QR are insignificant and small in magnitude across different quantiles.

A more recent study is that of [16], who developed a QR method for bilateral FDI panel data. The study aimed to estimate the individual firms' effect on FDI inflows among 161 countries from 2003 to 2012. They concluded that FDI's determinants vary across quantiles, where the effect of individual projects on FDI flow increases in the upper quantiles. In another study, [17], used QR to model the effects of FDI at different quantiles of the distribution of productivity. Empirical results from this study showed benefits derived from absorptive capacity and distance matter for productivity spillover benefits. It was also noted that there was significant heterogeneity across sectors and quantiles.

The ordinary QR model is designed to estimate the coefficients of linear or non-linear models to obtain the estimated conditional quantiles. This model setting restricts the QR model. This can be overcome by using kernel quantile regression (KQR) which is a nonparametric quantile regression method. KQR uses the ideas of nonparametric kernel density estimation, nonparametric kernel regression and quantile regression ([18]). This proposed method is known to provide a better fit to the data ([18], among others).

### *1.3 Research highlights*

In view of the literature survey presented in Section 1.2 the present study presents the following highlights.

1. To the best of our knowledge this is the first paper to use kernel quantile regression in forecasting South African FDI inflows.
2. Gross domestic product, gross fixed capital formation and exports were found to be the major determinants of foreign direct inflows into South Africa.
3. The kernel quantile regression model with ANOVA radial basis kernel function was found to be the best-fitting model.
4. QRA forecasts result in valid prediction interval coverage probabilities and narrow prediction interval widths.

The rest of the paper is organised as follows: Section 2 presents the models. The empirical results and discussions are presented in Section 3 and Section 4 concludes.

## 2 Methodology

### *2.1 Kernel regression*

Kernel regression is a non-parametric technique which is used to estimate the conditional expectation of a random variable using a weighted filter to the data. Its objective is to find a non-linear relation between a pair of random variables.

$$y_t = r(x_t) + \varepsilon_t, \qquad E(\varepsilon_t) = 0, \tag{1}$$

where $r$ is a smooth function, $x_t$ are the covariates, $y_t$ is the response variable and $\varepsilon_t$ is an error term. We estimate $r(x_t)$ by averaging nearby values of $y_t$.

$$\hat{r}(x) = \sum_{t=1}^{n} w_t(x) y_t, \qquad \sum_{t=1}^{n} w_t(x) = 1. \tag{2}$$

$\hat{r}(x)$ is a kernel smoother if weights are defined by a kernel function. If the weights are assigned a kernel function we get:

$$w_t(x) = cK\left(\frac{x_t - x}{h}\right), \tag{3}$$

where $c$ is a constant and $h$ is the bandwidth. This leads to:

$$\hat{y}_t = \frac{\sum_{t=1}^{n} K\left(\frac{|x_t - x|}{h}\right) y_t}{\sum_{t=1}^{n} K\left(\frac{|x_t - x|}{h}\right)} + \varepsilon_t. \tag{4}$$

We need $h$ which minimises

$$CV(h) = \frac{1}{n}\sum[\hat{r}_t(x_t - y_t)]^2, \tag{5}$$

where $CV(h)$ is the cross validation of the bandwidth and $\hat{r}_t(x_t)$ uses all the data except $(x_t, y_t)$.

#### 2.1.1 Bandwidth

The bandwidth $h$ is the maximum distance from the target point of any observation receiving weight. The quality of the curve estimates depends sensitively on the choice of $h$. This study will use $h$ which minimises

$$CV(h) = \frac{1}{n}\sum[\hat{r}_t(x_t - y_t)]^2. \tag{6}$$

#### 2.1.2 Different types of kernel functions

The idea is to approximate the results locally with series of quantile regressions that are estimated using a subset of the observations that are close to a set of target values, with more weight placed on observations that are close to the target points. For each of the target points, we define a set of weights that decline with distance, up to some maximum. At larger distances, the weight is set to zero. Then any kernel weight function can be used. A kernel function $K$ calculates the inner product of two vectors $x, x^{'}$, where $x$ is the training input and $x^{'}$ is the unlabeled input. This study will use the following kernel functions, discussed in [19].
The linear kernel is defined as:

$$K(x,x^{'}) = \langle x,x' \rangle. \tag{7}$$

The linear kernel is the simplest kernel function, which is useful specially when dealing with large sparse data vectors of $x$.
The Gaussian kernel is defined as:

$$K(x,x^{'}) = \exp(-\sigma \|x-x'\|^2). \tag{8}$$

The Gaussian kernel is the mostly widely used due to the fact that its linear combinations can approximate any continuous function. Its feature space has a finite dimension.
The Bessel kernel function is defined as:

$$K(x,x') = \frac{Bessel^{n}_{(\nu+1)}(\sigma\|x-x'\|)}{(\|x-x'\|)^{-n(\nu+1)}}. \tag{9}$$

Bessel kernel function is normally used when no prior knowledge is available.
The Laplace radial basis kernel is defined as:

$$K(x,x') = \exp(-\sigma\|x-x'\|). \tag{10}$$

It is the radial basis function kernel that is less sensitive for changes in the sigma parameter. The observations made about the sigma parameter for Gaussian kernel also apply to Laplace kernel.
The ANOVA radial basis kernel function is defined by:

$$K(x,x') = \left( \sum_{k=1}^{n} \exp(-\sigma(x^k - x'^k)^2 \right)^d, \tag{11}$$

where $x^k$ is the $k^{\text{th}}$ component of $x$ and $d$ is the degree. It is also a radial basis kernel function as the Gaussian and Laplace kernels. It is said to perform well in multidimensional regression problems.

### *2.2 Quantile regression and model specification*

Quantile regression (QR) was first introduced by [20], as the consolidated statistical methodology for estimating models of conditional quantile functions. QR is the extension of the ordinary least square regression (OLSR). The method of OLSR estimates the relationship between one or more predictor variables and the conditional mean of the response variable, while QR models the relationship between predictor variables and the conditional quantiles of response variable rather than the conditional mean of the response variable. A QR model gives a more comprehensive picture of the effect of the predictor variables on the response variable and contains more information than can be presented in OLSR ([21]).

QR is flexible in modelling data with a heterogeneous conditional distribution. It is richer in characterisation and description of the data, as it can show different effects of predictor variables on the response variable across the spectrum of the response variable. QR estimates are robust to non-normal errors and against outliers in the response measurements, but the main interest of QR goes beyond that. The QR model adapts readily to non-parametric estimation procedures, as the non-parametric estimation turns out to be easier to implement in a QR framework and the results can be presented in a straight forward way in a set of graphs ([21]).

The $\tau$ - quantile of a random variable $Y$, with cumulative distribution function $F_{Y|\mathbf{X}}$ is the conditional $Q_\tau$ of order $\tau \in (0,1)$ of $Y$ knowing $\mathbf{X}$ is defined as the generalized inverse of $F_{Y|\mathbf{X}}$ ([13]) and is given in equation (12).

$$Q_\tau(Y|\mathbf{X}) = F^{-1}_{Y|\mathbf{X}}(\tau) = inf\{y \in \mathbb{R} : F_{Y|\mathbf{X}}(y) \geq \tau\}. \tag{12}$$

QR assumes a model in the form of a deterministic distribution, as well as the distribution of the error term isn't assumed, because it is non parametric estimation. A QR model is as given in equation (13).

$$Q_{Y|X}(\tau) = \beta_0(\tau) + \beta_1(\tau)x_1 + ... + \beta_p(\tau)x_p + \varepsilon_\tau, \tag{13}$$

where $\beta_i(\tau)$ represents the parameter that corresponds to $\tau$ and $i = 0,1,2,...,p$. Due to the impact of different quantiles to error term, the model parameters will change when $\tau$ changes. The interpretation of the coefficients in the model is similar to the general linear model, but is not limited to the conditional mean, and different quantile positions are considered. In this study we are going to focus on kernel quantile regression. The conditional quantile $Q_{Y|X}(\tau)$ given in equation (13) is a solution to

$$\hat{\beta}_\tau = \arg \min_{\beta \in \mathbb{R}^{p+1}} \sum_{i=1}^{n} \rho_\tau \left( y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_j \right), \tag{14}$$

where $\rho_\tau(.)$ is the quantile loss also known as the pinball loss defined as $\rho_{\tau(u)} = u(\tau - \mathbf{I}(u < 0))$ and $\mathbf{I}(.)$ is an indicator function.

### *2.3 Kernel quantile regression*

We follow the work discussed by [22] who developed a KQR model which can be be used to estimate the parameters of a nonlinear conditional quantile function. This work is extended by [23]. The kernel ression model given in [23] is given in equation (15).

$$f(\mathbf{x}) = \beta_0 + \sum_{i=1}^{n} \beta_i K(\mathbf{x}, \mathbf{x}_i). \tag{15}$$

The $\tau^{\text{th}}$ conditional function is then given as ([23]):

$$f_\tau(\mathbf{x}) = \omega^{\mathrm{T}} \phi(\mathbf{x}) + \beta_0 + F^{-1}(\tau), \tag{16}$$

where $f_\tau(\mathbf{x})$ is a nonlinear quantile function, $F^{-1}(\tau)$ is the $\tau^{\text{th}}$ quantile of the error term $\varepsilon$. Let $d_\tau = \beta_0 + F^{-1}(\tau)$, then equation (16) reduces to

$$f_\tau(\mathbf{x}) = \omega^{\mathrm{T}} \phi(\mathbf{x}) + d_\tau. \tag{17}$$

The KQR model is then given as ([23]:

$$\min \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^{n} \rho_\tau \left( y_i - \omega^{\mathrm{T}} \phi(\mathbf{x}_i) - d_\tau \right), \tag{18}$$

where $\rho_\tau(.)$ is the pinball loss function as defined in equation (14)

The function in equation (15) can now be written as

$$f_\tau(\mathbf{x}) = \sum_{i=1}^{n} \beta_{i,\tau} K(\mathbf{x}_i, \mathbf{x}_j) + d_\tau. \tag{19}$$

Therefore the KQR model given in equation (18) can now be written as

$$\min \sum_{i=1}^{n} \rho_\tau \left( y_i - \sum_{j=1}^{n} \beta_{j,\tau} K(\mathbf{x}_i, \mathbf{x}_j) - d_\tau \right) + \lambda \beta_\tau^{\mathrm{T}} K \beta \tau, \tag{20}$$

where $\beta_\tau = (\beta_{1,\tau}, ..., \beta_{n,\tau})^{\mathrm{T}}$ denotes the kernel coefficient vector for the $\tau^{\text{th}}$ conditional quantile function and $\lambda$ is a regularisation parameter (smoothing parameter).

## 2.4 Variable selection

There are many different methods available that can be used to select important variables in a study. The most common used method is the subset selection which includes techniques such as the stepwise criterion and many more. The shrinkage selection methods includes elasticnet, least absolute shrinkage and selection operator (Lasso), ([24], [25]). This study will use the Lasso, which is given in equation (21).

$$\min_{\beta_0,\beta}\left(\sum_{t=1}^{n}(y_t-\beta^T x_t)+\lambda\|\beta\|_1\right), \tag{21}$$

where $\lambda$ is a tuning parameter. Lasso method regards absolute coefficient function as a penalty term to compress the coefficients of the model and coefficients whose absolute value is relative smaller than others are compressed to zero, so as to achieve the purpose of variable selection and parameter estimation.

## 2.5 Prediction intervals

A prediction interval (PI) is a useful tool for uncertainty modeling by its nature. It consists of lower and upper bounds that cover the future unknown target value with a certain probability $(1-\alpha)\%$ called the confidence level. PIs are more suitable and and produce more useful information than point forecasts for decision makers ([26]). Prediction interval width (PIW) is defined by:

$$\text{PIW}_t = UL_t - LL_t, \tag{22}$$

where $\text{UL}_t$ and $\text{LL}_t$ are the upper and lower limits respectively. The performance of PIs of the models are normally evaluated by the prediction interval coverage probability (PICP), prediction interval normalised average width(PINAW) and the prediction interval normalised average deviation (PINAD). PICP examines the reliability of the constructed PIs ([27]). PICP measures the probability with which targets lie in the constructed PIs. PICP is calculated as follows ([28]):

$$\text{PICP}_t = \frac{1}{m}\sum_{t=1}^{m} I_t, \tag{23}$$

where $m$ is the number of samples and

$$\text{I}_t = \begin{cases} 1 & \text{if} \quad y_t \varepsilon [LL_t; UL_t] \\ 0 & otherwise. \end{cases}$$

PICP is valid, if it is greater or equal to the nominal confidence level of PIs (PICP$\geq 1-\alpha$). PICP=100% if the target values are within the PIs. PINAW is the significant evaluation index of the PIs. PINAW is the higher quality of PIs, which evaluates the overall width of PIs. PINAW is given as ([27]):

$$\text{PINAW} = \frac{1}{mR}\sum_{t=1}^{m}(UL_t - LL_t), \tag{24}$$

where R is the range of the underlying targets. Prediction interval normalised average deviation (PINAD) state the deviation degree from the real values to the PIs. PINAD is given as ([26]):

$$\text{PINAD} = \frac{1}{mR}\sum_{t=1}^{m} d_t, \tag{25}$$

where $\text{d}_t$ denotes the deviation degree and is given by:

$$\text{d}_t = \begin{cases} LL_t - y_t & \text{if} \quad y_t < LL_t \\ 0 & \text{if} \quad LL_t \leq y_t \leq UL_t \\ y_t - UL_t & \text{if} \quad y_t > UL_t. \end{cases}$$

## 2.6 Forecast error distribution

The measures of central tendency (mean and the median) together with the lower and upper quartile of the residuals are used in the characterisation of the forecast error distribution. They provide important information about the distribution. The skewness and kurtosis can also provide more information about the distribution. Skewness measures the probability distribution's asymmetry, while the kurtosis describes the magnitude of the distribution's peak. The summary statistics will be obtained for the residuals of the best models, to emphasise the model that best forecast the FDI.

## 2.7 Tests for predictive accuracy

### 2.7.1 Diebold-Mariano test

The predictive ability of two competing models can be tested using statistical tests such as the Diebold-Mariano (DM) test ([29]). Under this test, the null hypothesis is that the two competing sets of forecasts have equal predictive accuracy. The DM test is known to take into account the sampling variability in the average losses ([30]).

Let $y_t, t = 1,...,m$ be the FDI values and two forecasts $\hat{y}_{i,t}, \hat{y}_{j,t}, \forall i \neq j, i, j = 1,2,...,K$. Assuming the errors from the forecasts are defined as $\varepsilon_{i,t} = \hat{y}_{i,t} - y_t, i = 1,2$. If $g(\varepsilon_{i,t})$ is an error loss function, then a loss function which penalises heavily underprediction than overprediction is given as ([31]):

$$g(\varepsilon_{i,t,\tau}) = e^{\lambda \varepsilon_{i,t}} - 1 - \lambda \varepsilon_{i,t}. \tag{26}$$

### 2.7.2 Giacomini-White test

A generalisation of the DM test is the Giacomini-White (GW) test. It tests the conditional predictive ability of two competing forecasting methods ([32]). This test takes into consideration the uncertainty in the estimation of parameters ([32]). Under this test, the null hypothesis is that of equal conditional predictive ability between two competing models.

# 3 Empirical results and discussion

## 3.1 Exploratory data analysis

The study uses data, obtained from the South African Reserve Bank (SARB) and the International Country Risk Guide (ICRG). The ICRG is part of the political risk services group which provides ratings affecting political risk, economic risk and financial risk. The ICRG provides political risk indices which are given into five bands summarised as follows ([33]):

- Band 1: Very low risk (80% - 100%)
- Band 2: Low risk (70% - 79.9%)
- Band 3: Moderate risk (60% - 169.9%)
- Band 4: High risk (50% - 59.9%)
- Band 5: Very high risk (0.0% - 49.9%)

The data used was for the period years 1996 to 2015. This study will use the R package 'kernlab' developed by [34] for estimating parameters of the locally linear quantile regression models under each of the discussed kernel functions.

### 3.1.1 Description of variables

In this study, the dependent variable is the foreign direct investment (FDI), which is an investment made by a company or individual in one country in business interests in another country. The inward FDI stock is the value of foreign investors' equity in and net loans to an enterprise resident in the reporting economy. FDI stocks measure the total level of direct investment at a given point in time, usually at the end of the quarter or of the year. This study will use the FDI measured at the end of each year from 1996 to 2015.
The independent variables are as follows:

- GDP - is the gross domestic product and is considered to be one of the best ways to measure a country's economy.
- CPI - is the consumer price index.
- LP - is labour productivity.
- GFCF - is the gross fixed capital formation.
- Exp - is the exports.
- PS - is political stability.
- noltrend - is the nonlinear trend variable.

3.1.2 Summary statistics

Table 1 presents the summary statistics for all the variables that are used in this study. From the year 1996 to the year 2015, we can see that SA has received the average FDI inflow of R3,856,486,223.00, having the maximum of R9,885,001,293.00 which was in 2008 when the political risk index was 0.69 and the minimum inflow of R5,500,338,596.00 in 1998 with a political risk index of 0.72 in that year. The average political risk index during the sampling period was 0.71 with a minimum of 0.65 in 2012 and a maximum of 0.79 in 2004.

The Exp, LP, noltrend and GDP have negative skewness, while other variables are positively skewed. The kurtosis of the CPI is 3.92 which is more than 3, indicating that its distribution is leptokurtic, meaning its central peak is higher and sharper compared to the normal distribution. All other variables have a kurtosis of less than 3, indicating that their distributions are specified as platykurtic when compared with a normal distribution.

**Table 1:** Summary statistics.

| Variable | Mean | Q1 | Q2 | Q3 | Min | Max | Skewness | Kurtosis |
|---|---|---|---|---|---|---|---|---|
| FDI | 3.9e+09 | 9.3e+08 | 3.8e+09 | 6.5e+09 | 5.5e+08 | 9.9e+09 | 0.426 | 1.83 |
| GFCF | 436734 | 289816 | 435706 | 562394 | 262458 | 639383 | 0.05 | 1.34 |
| GDP | 50015 | 44990 | 50333 | 54484 | 43720 | 56469 | -0.003 | 1.33 |
| Exp | 736213 | 658529 | 738834 | 827752 | 542552 | 911366 | -0.109 | 1.72 |
| CPI | 6.025 | 4.925 | 5.700 | 7.100 | 1.400 | 11.500 | 0.47 | 3.92 |
| LP | 90.33 | 79.58 | 91.70 | 101.10 | 64.8 | 110.50 | -0.18 | 1.80 |
| PS | 0.71 | 0.69 | 0.69 | 0.73 | 0.65 | 0.79 | 0.45 | 2.50 |
| noltrend | 0.948 | 0.564 | 0.565 | 0.517 | 0.274 | 0.519 | -0.374 | 1.000 |

3.1.3 Graphical description of the data set

The data set contains data of 8 variables with 20 observations. The data was initially standardized. First we will analyze the FDI data set using visualization plots. Looking at Figure 1(a) on the time series plot we can see that our dependent variable (FDI) has a seasonal pattern. Figure 1(a) shows that since 1994, there is a positive upward trend in FDI inflows into SA. This major increase in the attraction of FDI was due to democratisation of SA, where there was subsequent openness to trade. In 1997, however, there was a significant increase of FDI, due to the partial privatisation of Telkom and South African Airlines ([35]). A remarkable increase occurred in 2001. In 2002, FDI inflows decreased and continued to fall in 2004. In 2005 there was another increase since SA was the largest FDI recipient in Africa. In 2006, the inflows declined drastically but made an increase in 2007.

There was a steady increase until 2008 and 2009. In 2010, the inflows decreased as the world was experiencing a contraction in demand due to the 2008 financial crisis, even though there was an increase in the tourism industry's FDI inflows as a result of the 2010 World Cup, hosted by South Africa. In 2012, the country faced one of the biggest situations of labour unrest in its history from the mining industry, and as a result, FDI inflows slumped. In 2014, FDI inflows decreased again towards 2015.

Figure 2 shows the correlations between FDI and the covariates in the top panel, with the middle and bottom panels showing FDI plotted against the covariates, respectively. The correlations between the FDI and the covariates are all positive except for PS (political stability) which is negative. This shows that an increae in instability of a country resiluts in a decrease in FDI inflows.

## *3.2 Qunatile regression results*

3.2.1 Variable selection using Lasso

The data is split into two sets: a training set, from the year 1996 to the year 2011, and a testing set which is from the year 2012 to the year 2015. In the training data set we want to build a model to predict the dependent variable FDI. To choose

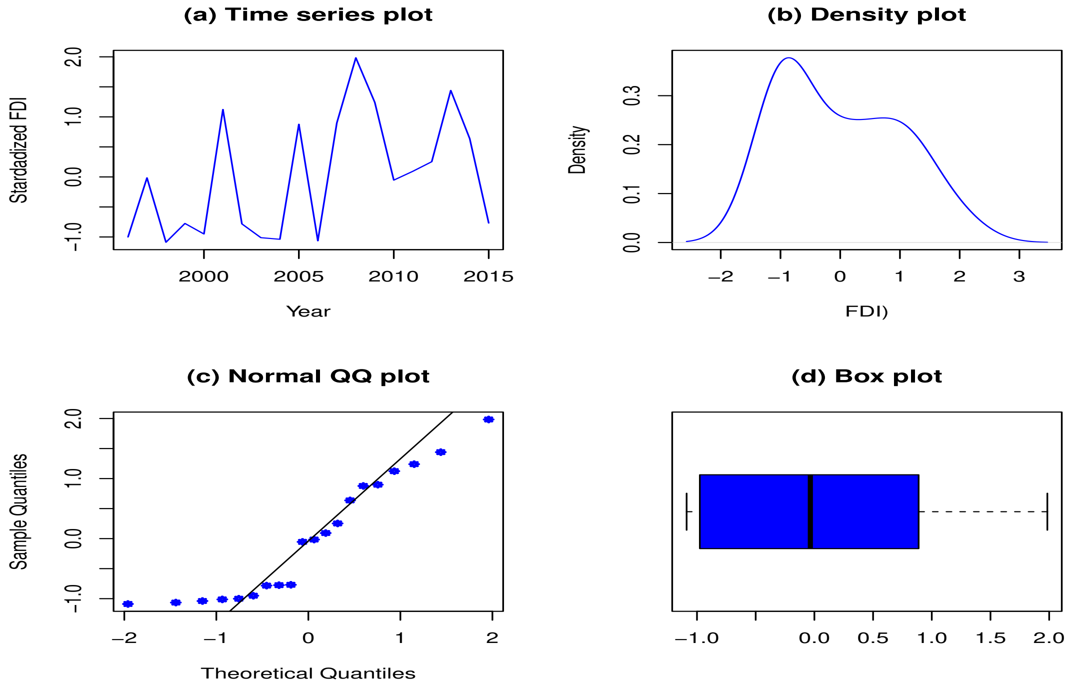


**Fig. 1:** FDI time series, density, normal and box plots.

| | GFCF | GDP | Exports | CPI | LP | PS |
|---|---|---|---|---|---|---|
| FDI | 0.55 | 0.53 | 0.51 | 0.34 | 0.47 | -0.26 |

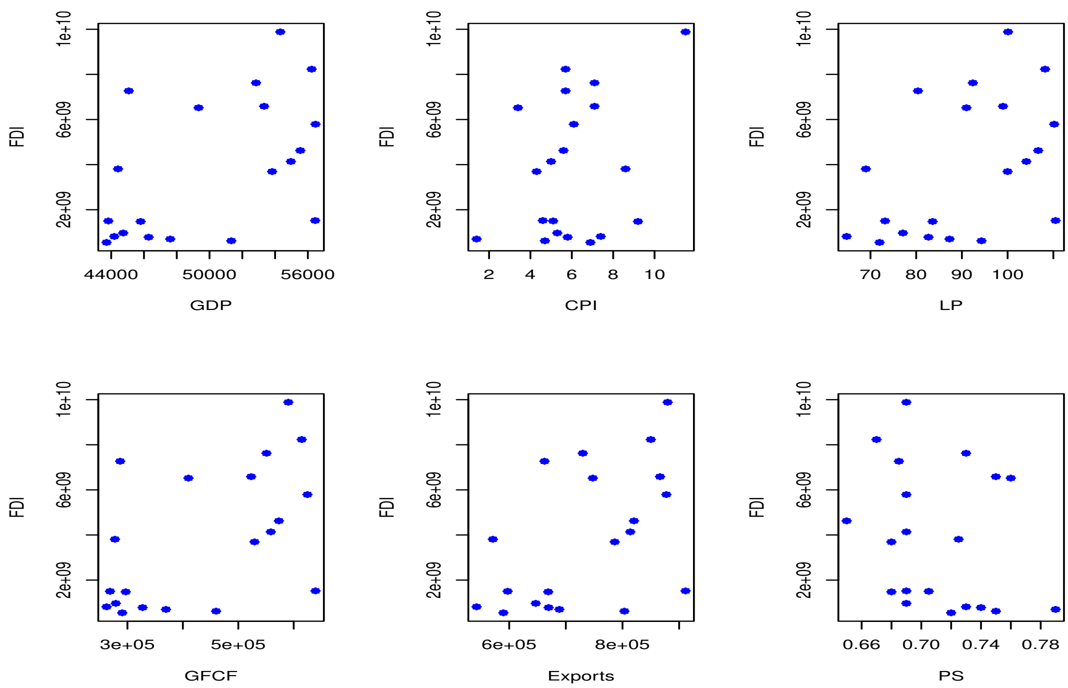


**Fig. 2:** (a) Top panel: Correlations between FDI and the covariates. (b) Middle and bottom panels: FDI plotted against the covariates.

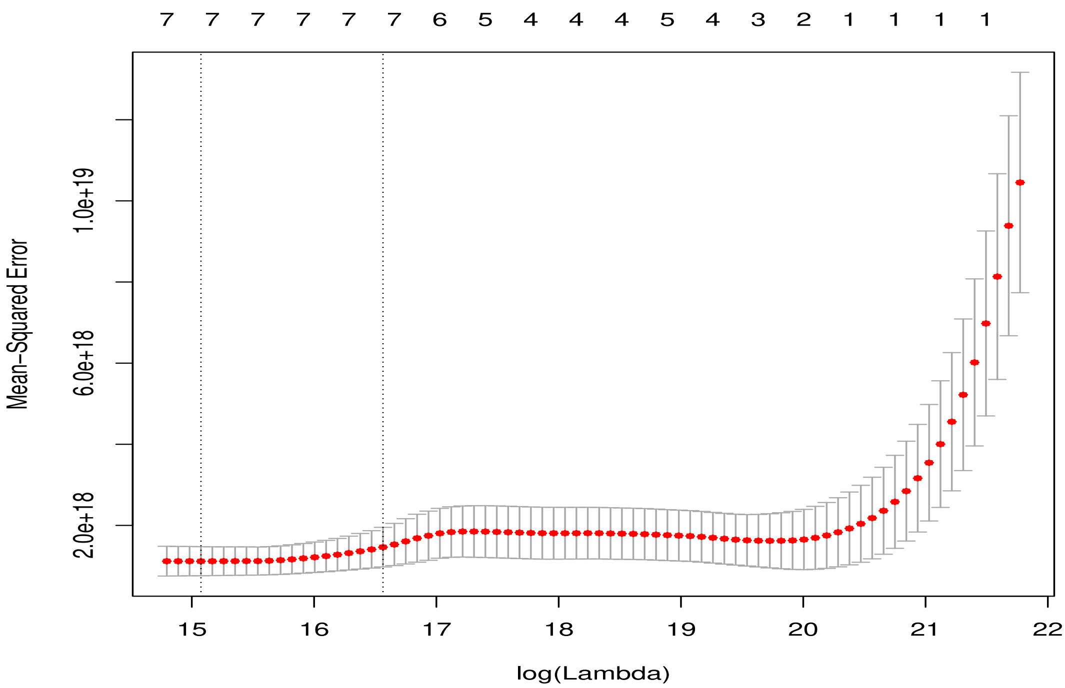


**Fig. 3:** Cross-validation.

which variable to use in the model we use the Lasso method making use of the “glmnet” package in R developed by [25]. Glmnet returns a sequence of different models for different values of $\lambda$.

On the Lasso variable selection, all seven variables were selected. All solutions were found after 9 iterations. It only took several iterations before the results of the optimal solution. Variable selection results can only be presented after the selection of the optimal $\lambda$ (see paragraph below).

Table 2 shows variables selection using Lasso. From Table 2 all the variables were selected which shows that all the independent variables have an influence on the model.

**Table 2:** Variables selected using Lasso.

| | |
|---|---|
| (Intercept) | 1.1888 |
| GFCF | 3.4081 |
| CPI | 8.8239 |
| LP | -6.6944 |
| Exp | 1.5286 |
| GDP | -1.0792 |
| PS | 4.1839 |
| noltrend | 1.9708 |

The next step is to choose which value of $\lambda$ to consider. The R function ‘cv.glmnet’ helps us to select the most appropriate value of $\lambda$, choosing the number of validations. The relevant plot is shown in Figure 3. To choose the most appropriate value for $\lambda$ the Lasso method extracts different values of $\lambda$, such as $\lambda_{min}$ (first vertical dotted line) that gives minimum mean cross-validated error and $\lambda_{1se}$ (second vertical dotted line), that gives a model such that the error is within one standard error of the minimum. For both the values of $\lambda$ all seven variables are selected to be in the model.

### 3.2.2 Kernel quantile regression models

We present results of the kernel quantile regression models based on the following kernel functions: Linear (M3), Gaussian (M4), Bessel (M5), Laplace radial basis (M6) and ANOVA radial basis (M7). The Linear kernel is denoted as M3. Gaussian kernel is denoted as M4, Bessel kernel is denoted as M5 and Laplace radial basis and ANOVA radial basis are denoted as M6 and M7, respectively. Table 3 gives the 95% prediction intervals for all kernel quantile regression models.

The forecast for the lower bound of the prediction interval was obtained with the quantile of $\tau$ of 0.25. The forecast for the upper bound of the prediction interval was obtained with the quantile of $\tau$ of 0.975.

**Table 3:** Prediction intervals of Linear, Gaussian, Bessel, Laplace and ANOVA radial basis kernel functions.

| Linear kernel function | | | | |
|---|---|---|---|---|
| Date | Actual FDI | $Q_{0.25}$ | $Q_{0.5}$ | $Q_{0.975}$ |
| 2011 | 4,139,289,123 | 3,601,483,119 | 4,457,091,910 | 6,312,521,517 |
| 2012 | 4,626,029,122 | 3,459,780,039 | 4,295,305,779 | 6,754,326,509 |
| 2013 | 8,232,518,816 | 4,727,292,115 | 6,291,891,470 | 8,232,518,819 |
| 2014 | 5,791,659,020 | 4,277,708,536 | 5,791,659,046 | 7,645,834,982 |
| 2015 | 1,521,139,945 | 1,521,139,931 | 2,884,559,217 | 4,246,327,276 |
| Gaussian kernel function | | | | |
| Date | Actual FDI | $Q_{0.25}$ | $Q_{0.5}$ | $Q_{0.975}$ |
| 2011 | 4,139,289,123 | 2,327,241,043 | 4,139,289,150 | 7,493,224,082 |
| 2012 | 4,626,029,122 | 2,392,803,728 | 4,626,029,130 | 7,896,707,171 |
| 2013 | 8,232,518,816 | 2,509,221,920 | 6,280,533,976 | 8,232,518,850 |
| 2014 | 5,791,659,020 | 2,360,959,675 | 5,791,659,018 | 8,101,312,070 |
| 2015 | 1521139945 | 1521139935 | 1521139954 | 7192630913 |
| Bessel kernel function | | | | |
| Date | Actual FDI | $Q_{0.25}$ | $Q_{0.5}$ | $Q_{0.975}$ |
| 2011 | 4,139,289,123 | 2,487,178,389 | 4,362,045,845 | 7,503,534,252 |
| 2012 | 4,626,029,122 | 2,505,392,846 | 4,626,029,123 | 8,005,525,466 |
| 2013 | 8,232,518,816 | 2,618,284,541 | 6,152,270,374 | 8,232,518,843 |
| 2014 | 5,791,659,020 | 2,461,073,480 | 5,791,659,014 | 8,138,562,732 |
| 2015 | 1,521,139,945 | 1,521,139,935 | 1,521,139,949 | 7,125,854,659 |
| Laplace radial basis kernel function | | | | |
| Date | Actual FDI | $Q_{0.25}$ | $Q_{0.5}$ | $Q_{0.975}$ |
| 2011 | 4,139,289,123 | 4,139,289,119 | 4,139,289,124 | 7,658,408,711 |
| 2012 | 4,626,029,122 | 4,615,340,098 | 4,626,029,126 | 7,965,387,527 |
| 2013 | 8,232,518,816 | 5,246,630,120 | 7,253,495,221 | 8,232,518,860 |
| 2014 | 5,791,659,020 | 4,862,831,701 | 5,791,659,021 | 8,135,666,559 |
| 2015 | 1,521,139,945 | 1,840,239,977 | 1,521,139,968 | 7,652,734,546 |
| ANOVA radial basis kernel function | | | | |
| Date | Actual FDI | $Q_{0.25}$ | $Q_{0.5}$ | $Q_{0.975}$ |
| 2011 | 4,139,289,123 | 4,139,288,915 | 4,139,289,107 | 6,359,577,977 |
| 2012 | 4,626,029,122 | 4,626,029,093 | 4,626,029,127 | 6,486,153,500 |
| 2013 | 8,232,518,816 | 5,002,751,144 | 8,232,518,379 | 8,232,518,826 |
| 2014 | 5,791,659,020 | 4,859,770,963 | 5,823,624,433 | 7,577,278,212 |
| 2015 | 1,521,139,945 | 1,521,139,935 | 1,521,139,942 | 3,594,033,866 |

### 3.2.3 Linear quantile regression averaging

The linear quantile regression averaging model was built using the linear quantile regression with all the non-parametric quantile regression models (M3, M4, M5, M6 and M7) as the independent variables. The linear quantile regression averaging model which we denote as M8 is given by equation (27).

$$Q_{\tau}(FDI_t) = \beta_{0,\tau} + \beta_{1,\tau} fM3 + \beta_{2,\tau} fM4 + \beta_{3,\tau} fM5 + \beta_{4,\tau} fM6 + \beta_{5,\tau} fM7 + \varepsilon_{t,\tau} \quad (27)$$

where $fM$ is the forecast of the models M3 to M7. Table 4 presents a summary of the estimates of the parameters. The coefficient of models M3, M4, M5 and M6 are all zero, this shows that they are all not necessary in model M8. M7 has the coefficient of one and the p-value of 0.03684, this makes it to be the only variable that is significant on predicting FDI in model M8. Table 5 is showing the 95% prediction intervals for model M8. The forecast for the lower bound of the prediction interval was obtained with the quantile of $\tau$ of 0.25. The forecast for the upper bound of the prediction interval was obtained with the quantile of $\tau$ of 0.975. All the forecasts of model M8 are inside the prediction intervals.

**Table 4:** Coefficients of linear quantile regression averaging model.

| | Coefficients | Std. Error | t value | P-value |
|---|---|---|---|---|
| Intercept | 0.0000 | 0.0002 | 0.0000 | 1.0000 |
| M3 | 0.0000 | 0.0020 | 0.0001 | 0.9999 |
| M4 | 0.0000 | 0.62698 | 0.0000 | 1.0000 |
| M5 | 0.0000 | 0.87009 | 0.0000 | 1.0000 |
| M6 | 0.0000 | 0.46844 | 0.0000 | 1.0000 |
| M7 | 1.0000 | 0.43347 | 2.3072 | 0.03684 |

**Table 5:** Prediction intervals: Linear quantile regression averaging.

| Date | Actual FDI | $Q_{0.25}$ | $Q_{0.5}$ | $Q_{0.975}$ |
|---|---|---|---|---|
| 2011 | 4139289123 | 4119050821 | 4139289108 | 4139289123 |
| 2012 | 4626029122 | 4598699833 | 4626029126 | 4626029127 |
| 2013 | 8232518816 | 8232518815 | 8232518815 | 8232518815 |
| 2014 | 5791659020 | 5791659020 | 5823624444 | 5823624444 |
| 2016 | 1521139945 | 1509122141 | 1521139942 | 1521139945 |

#### 3.2.4 Forecast evaluation for all models

RMSE, MAE and MAPE are used to evaluate the forecasting accuracy of best prediction models. The model with smallest RMSE, MAE and MAPE is the best model. It is clear from Table 6 that the three best models are M6, M7 and M8, since their RMSE, MAE and MAPE are less compared to those of all other models. Model M7 is the best of them all. Figure 4 shows the time series plots between the actual FDI and four different types of kernels. For all these four kernels, we can see that their forecast (dotted line) doesn't follow the actual FDI (solid line) very well, since most forecasts deviate from

**Table 6:** Models forecast evaluation.

| Models | RMSE | MAE | MAPE |
|---|---|---|---|
| M3 | 990412816 | 710514555 | 20.34699 |
| M4 | 872954158 | 390396977 | 4.742133 |
| M5 | 935633917 | 460601035 | 6.130039 |
| M6 | 548418213 | 245260093 | 2.979163 |
| M7 | 14295367 | 6393175 | 0.1103855 |
| M8 | 14295370 | 6393088 | 0.1103845 |

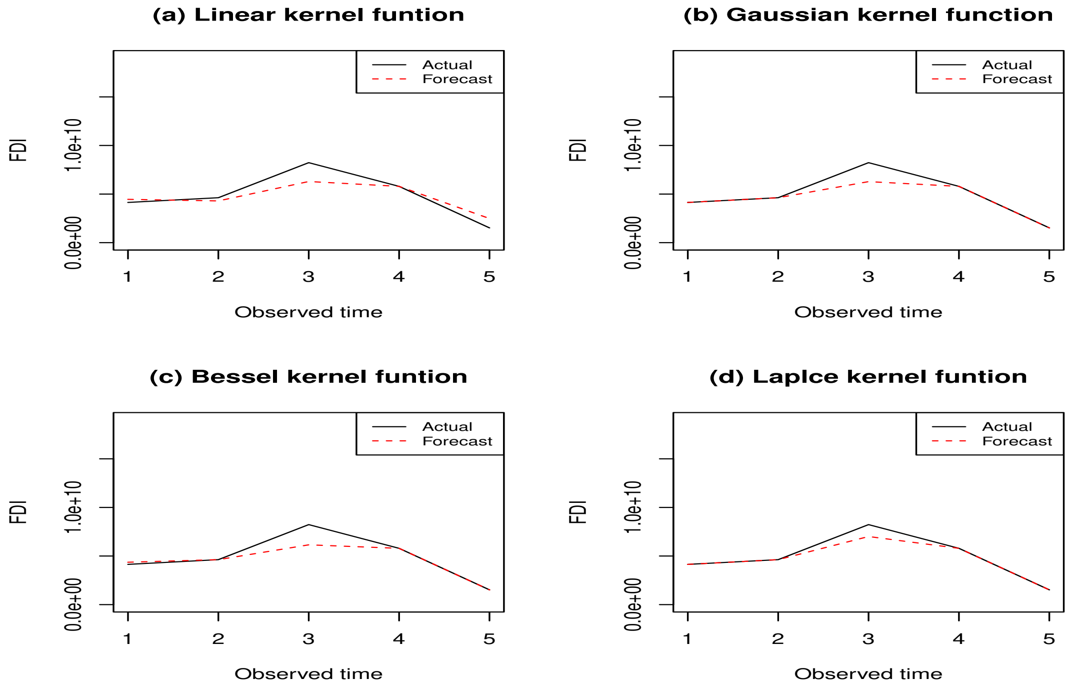


**Fig. 4:** Time series plot (kernel functions forecast).

the actual FDI. On the graph of the linear kernel function, we can see that the forecast goes below and above the actual FDI, meaning that this model has more under and over predictions. The Bessel, Gaussian and Laplace kernels have more under predictions. Figure 5 top panel and Figure 5 bottom panel are showing the time series plots of the actual FDI and the forecast of the models ANOVA radial basis kernel function and linear quantile regression averaging. The graphs show that the forecast for these two models follow the actual FDI remarkably well, as there are no visible deviations.

### 3.2.5 Performance comparison of best models

Prediction intervals

This section presents the comparison of the performance of the models, for the best three chosen models (M6, M7 and M8). Tables 7, 8 and 9 are showing the prediction intervals width (PIW) and residuals for models M6, M7 and M8. All the prediction intervals were constructed under the 95% confidence level. The PIWs were constructed by taking the difference between the upper bound of 0.975 and lower bound of 0.25 and the residuals where obtained from the difference between the actual FDI and its forecast ($\tau = 0.5$).

**Table 7:** Prediction intervals widths: M6.

| Date | Actual FDI | $Q_{0.25}$ | $Q_{0.5}$ | $Q_{0.975}$ | PIW | Residuals |
|---|---|---|---|---|---|---|
| 2011 | 4,139,289,123 | 4,139,289,119 | 4,139,289,124 | 7,658,408,711 | 3,519,119,592 | -1.313 |
| 2012 | 4,626,029,122 | 4,615,340,098 | 4,626,029,126 | 7,965,387,527 | 3,350,047,429 | -3.599 |
| 2013 | 8,232,518,816 | 5,246,630,120 | 7,253,495,221 | 8,232,518,860 | 2,985,888,740 | 9,790,23,595 |
| 2014 | 5,791,659,020 | 4,862,831,701 | 5,791,659,021 | 8,135,666,559 | 3,272,834,858 | -0.900 |
| 2015 | 1,521,139,945 | 1,840,239,977 | 1,521,139,968 | 7,652,734,546 | 5,812,494,569 | -22.695 |

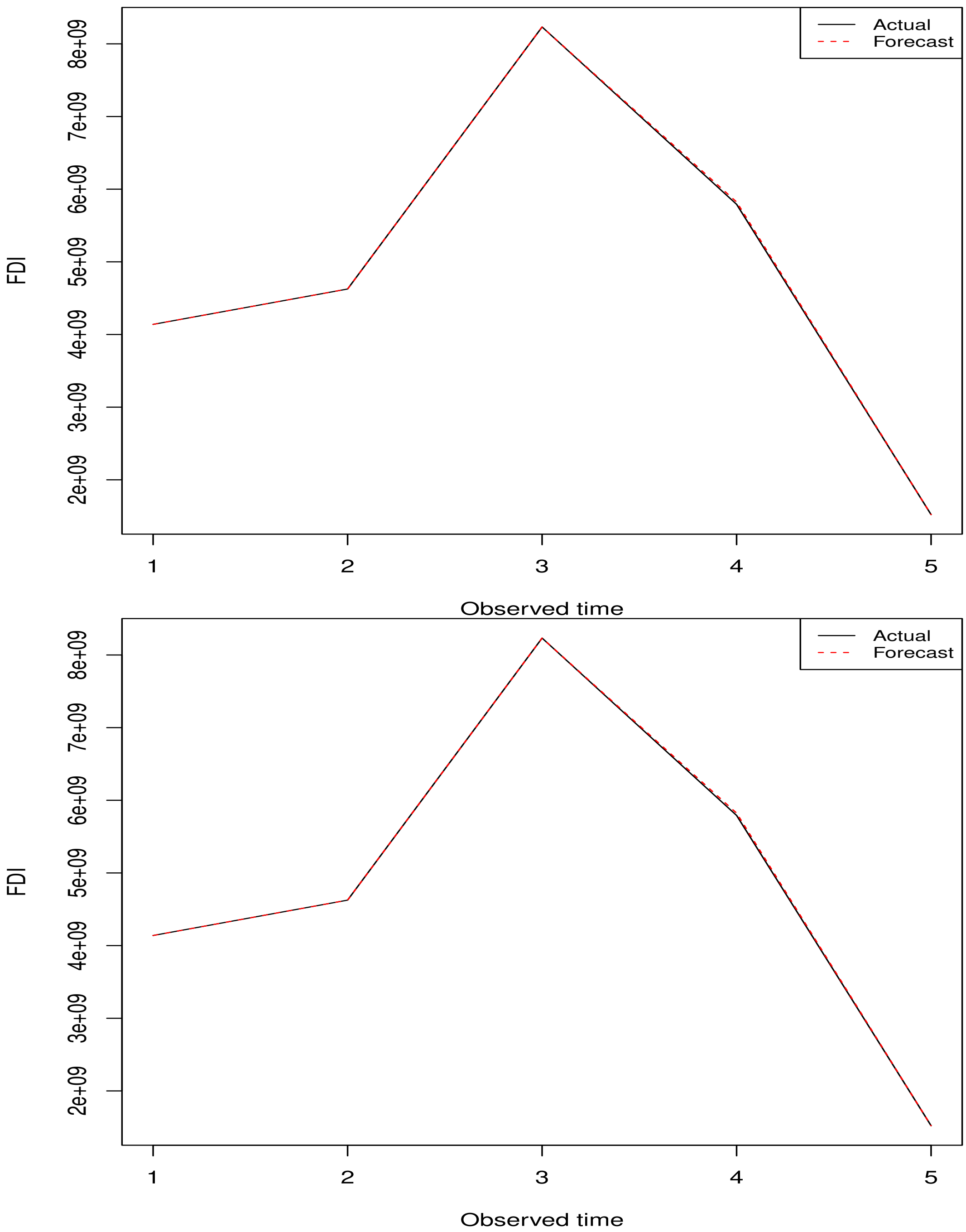


**Fig. 5:** Top: Time series plot of FDI with forecasts from model M7 (ANOVA radial basis kernel). Bottom: Time series plot of FDI with forecasts from model M8 (LQRA).

The accuracy check of the prediction intervals was done using PICP, PINAW, PINAD, over prediction count and under prediction count. It can be seen from Table 10 that the prediction interval coverage probability (PICP) of model M8 is 100%, meaning that for M8 the prediction interval is covering all the target values. For model M8 all the target values are covered completely or we have 100% coverage. Models M6 and M7 are both having the PICP of 80%, meaning that the models M6 and M7 have 80% coverage. The PICPs of M6 and M7 are not valid since they are less than 95%. This means that for the model M8, the constructed PIs cover the target values with high probability. The PINAW values of the models M6 and M7 are wider than the PINAW value of model M8. The PINAD values of the models M6 and M7 are

**Table 8:** Prediction intervals width: M7.

| Date | Actual FDI | $Q_{0.25}$ | $Q_{0.5}$ | $Q_{0.975}$ | PIW | Residuals |
|---|---|---|---|---|---|---|
| 2011 | 4,139,289,123 | 4,139,289,043 | 4,139,289,107 | 6,359,577,977 | 2,220,288,934 | 15.6874 |
| 2012 | 4,626,029,122 | 4,626,029,122 | 4,626,029,127 | 6,486,153,500 | 1,860,124,378 | -4.5999 |
| 2013 | 8,232,518,816 | 7,545,980,351 | 8,232,518,379 | 8,232,518,826 | 6,865,384,75 | 436.6208 |
| 2014 | 5,791,659,020 | 5,791,659,022 | 5,823,624,433 | 7,577,278,212 | 1,785,619,190 | -319654.22 |
| 2015 | 1,521,139,945 | 1,521,139,941 | 1,521,139,942 | 3,594,033,866 | 2,072,893,925 | 3.3053 |

**Table 9:** Prediction intervals widths: M8.

| Date | Actual FDI | $Q_{0.25}$ | $Q_{0.5}$ | $Q_{0.975}$ | PIW | Residuals |
|---|---|---|---|---|---|---|
| 2011 | 4,139,289,123 | 4,119,050,821 | 4,139,289,108 | 4,139,289,123 | 20,238,302 | 14.7 |
| 2012 | 4,626,029,122 | 4,598,699,833 | 4,626,029,126 | 4,626,029,127 | 27,329,294 | -3.6 |
| 2013 | 8,232,518,816 | 8,232,518,815 | 8,232,518,815 | 8,232,518,815 | 0 | 0.6 |
| 2014 | 5,791,659,020 | 5,791,659,020 | 5,823,624,444 | 5,823,624,444 | 31,965,424 | -31,965,419 |
| 2016 | 1,521,139,945 | 1,509,122,141 | 1521139942 | 1521139945 | 12,017,804 | 3.3 |

**Table 10:** Comparisons of M6, M7, and M8.

| | PICP(% ) | PINAW (%) | PINAD (%) | Over prediction | Under prediction |
|---|---|---|---|---|---|
| M6 | 80 | 56.4 | $9.509\times^{-3}$ | 0 | 1 |
| M7 | 80 | 25.7 | $5.662\times^{-11}$ | 0 | 1 |
| M8 | 100 | 0.27 | 0 | 0 | 0 |

also wide, compared to the value of model M8. Model M8 is considered to be the best model, compared to models M6 and M7. Since model M8 has 100% coverage and narrow PINAW and PINAD. Model M8 was also confirmed to be the superior model by the under and over predictions. Model M8 has zero under and above predictions. Figure 6 top panel and Figure 6 bottom panel are showing the density plots and box and whisker plots for prediction intervals widths for the Models M6, M7 and M8, respectively. The density plot of model M8 seems to be the best compared to the density plot of models M6 and M7 since the density of model M8 is narrow. This also confirms that model M8 is the best.

Forecast error distribution

Table 11 gives summary statistics of the residuals of the best three models. Model M8 has the lowest mean forecast error compared to the models M6 and M7 and also has the lowest standard deviation error. This also shows that model M8 is still the best model. All these three models have the same kurtosis and skewness values, except that, the skewness of model M6 is positive.

**Table 11:** Summary statistics for the residuals of forecast of M6, M7 and M8.

| | Mean | Median | Minimum | Maximum | Kurtosis | Skewness | Standard deviation |
|---|---|---|---|---|---|---|---|
| M6 | 195804713 | -1 | -23 | 979023595 | -0.92 | 1.073313 | 437832665 |
| M7 | -6392992 | 3 | -31965413 | 437 | -0.92 | -1.073313 | 14295418 |
| M8 | -6393081 | 1 | -31965420 | 15 | -0.92 | -1.073313 | 14295372 |

Empirical results on statistical tests of equal forecast accuracy and test of conditional predictive ability

We present the results from the DM and GW tests. In all cases the null hypothesis is: $H_0$ : forecasts are equally accurate.

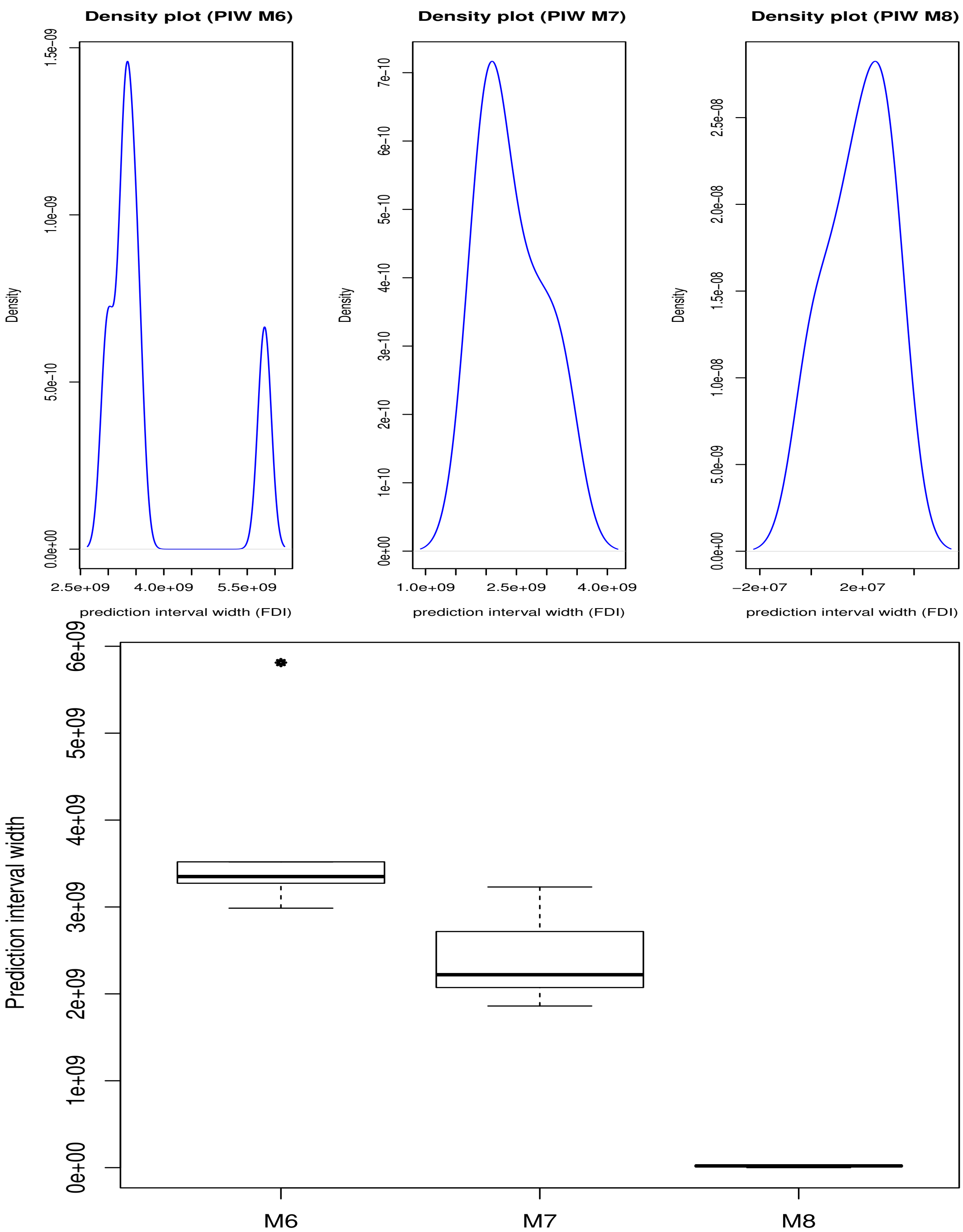


**Fig. 6:** Top: Density plot for PIW (M6, M8 and M7). Bottom: Box and whisker plot of PIW (M6, M8 and M7).

The three best models considered are M6 (using the Laplace radial basis kernel), M7 (using the ANOVA radial basis kernel) and M8 (based on linear quantile regression averaging (LQRA)). The null hypotheses are:
$H_0$ forecasts from M6 and M7 are equally accurate.
$H_0$ forecasts from M6 and M8 are equally accurate.
$H_0$ forecasts from M7 and M8 are equally accurate.

Based on the results from Table 12 the DM tests show that in all cases the forecasts are equally accurate. For the GW test we shall use the following to denote dominance, $M_i < M_j, \forall i \neq j$ to mean that $M_j$ dominates $M_i$. From Table 12,

**Table 12:** Model comparisons.

| DM test results | | | |
|---|---|---|---|
| Null hypothesis | Test statistic | p-value | Result |
| M6 = M7 | 1.4587 | 0.2184 | Equally accurate |
| M6 = M8 | 1.4587 | 0.2184 | Equally accurate |
| M7 = M8 | 1.4809 | 0.2128 | Equally accurate |
| GW test results | | | |
| Null hypothesis | Test statistic | p-value | Result |
| M6 = M7 | 2.0000 | 0.2184 | Sign of mean loss is (+). M7 dominates M6 |
| M6 = M8 | 2.0000 | 0.2184 | Sign of mean loss is (+). M8 dominates M6 |
| M7 = M8 | 2.0001 | 0.2128 | Sign of mean loss is (-). M7 dominates M8 |

$M6 < M7$, $M6 < M8$, $M8 < M7$ implies that $M6 < M8 < M7$. Therefore M7 has the highest predictive power since it dominates the other two models.

## 4 Discussion

To avoid unnecessary predictors and overfitting, variable selection was done using Lasso. All the predictor variables were selected to be included in the kernel regression models. This meant that the GDP, GFCF, LP, CPI, PI, Exp and noltrend could be regarded as the major determinants of FDI in South Africa. The noltrend, GDP, GFCF and Exp were found to have more influence on the inflow of FDI compared to the other variables

Results from the nonparametric local linear kernel quantile regression models showed that all the actual FDI values were found to be inside the prediction intervals (developed from the models Linear (M3), Gaussian (M4), Bessel (M5), Laplace radial basis (M6) and ANOVA radial basis (M7)), even though some of the prediction intervals are wider than others. Even though the prediction intervals for models Linear (M3), Gaussian (M4), Bessel (M5) and Laplace radial basis (M6) cover all the actual values, the times series plots showed that the forecasted values do not follow the actual FDI very well. However the forecast values of model M7 did follow the actual FDI very well, this made ANOVA radial basis to be the best model for forecasting FDI in South Africa. Based on the accuracy measures, the ANOVA radial basis kernel model was also confirmed to be the best model, since it had the lowest values of MAPE, MAE and RMSE compared to Linear, Gaussian, Bessel and Laplace radial basis models. Forecasts from the ANOVA radial basis kernel model were found to follow the actual FDI values remarkably well and all the forecast values were inside the prediction intervals. The results are consistent with those of [36], who established that in multidimensional regression problems, the ANOVA radial basis kernel performs very well. The forecasts from the Linear, Gaussian, Bessel, Laplace radial basis and ANOVA radial basis models were then combined using linear quantile regression averaging resulting in M8.

Forecast evaluation was done for all the models that were built in this study. Based on the accuracy measures, the best three models were Laplace radial basis, ANOVA radial basis and linear quantile regression averaging, since their RMSE, MAE and MAPE were less compared to those of all other models. A comparative analysis of the best three chosen models was also done. An analysis of the prediction intervals was done using PICP, PINAW, PINAD, over-prediction count and under prediction count. The ANOVA radial basis model was considered to be the best fitting model, compared to models Laplace radial basis and ANOVA radial basis, since it had 100% coverage probability and narrow PINAW and PINAD. The linear quantile regression averaging model was also confirmed to be the superior model for forecasting FDI as it had no under and over predictions. The forecast error distribution also confirmed that the ANOVA radial basis model was the best model for forecasting the FDI. However, based on the conditional predictive ability of two competing forecasting methods using the GW tests, the ANOVA radial basis kernel model was found to have the highest predictive ability. On using the DM tests the null hypothesis that the forecasts of two competing are equally accurate was not rejected for all cases at the 5% level of significance. Based on all these forecast evaluation metrics as well as the statistical tests for predictive accuracy, the ANOVA radial basis kernel is therefore considered to be the best model. Having the model that can accurately forecast the FDI inflows can help South Africa to increase its economic growth.

## 5 Conclusion

In this study, we presented and analysed the kernel quantile regression models in forecasting FDI in South Africa for over 20 years (years 1996 to 2015). Results from this study showed that combining forecasts using quantile regression averaging results in more accurate forecasts compared to using individual models. These results are consistent with those of [37], [38], among others. However, the ANOVA radial basis kernel was considered to be the best fitting model based on the forecast accuracy measures used and the predictive ability tests carried out. Forecasts of FDI inflows are important to decision-makers in government who have to ensure economic growth in South Africa. According to the findings, the FDI inflows to SA seemed to fluctuate a lot between the years 1996 to 2015.

## Author Contributions

The results of this study were obtained from a submitted masters dissertation at the University of Venda. Conceptualization, N.N. and C.S.; methodology, N.N., C.S. and A.B; software, N.N. and C.S.; validation, N.N., C.S. and A.B.; formal analysis N.N. and C.S.; original draft preparation, N.N.; writing, review and editing, N.N., C.S. and A.B.; supervision, C.S. and A.B. All authors have read and agreed to the submitted version of the manuscript.

## Funding

This research received no external funding.

## Acknowledgments

The authors are grateful to numerous people for their helpful comments on this article.

## Conflict of interest

The authors declare that they have no conflict of interest.

## References

[1] T.A. Pugel, Foreign investment in the us food-marketing system: Discussion. *American Journal of Agricultural Economics*, **65(2)**, 423-425, (1983).
[2] J.R. Markusen, J.R. and A.J. Venables, Foreign direct investment as a catalyst for industrial development, *European Economic Review*, **43(2)**, 335-356, (1999).
[3] N. Bermeo. Does electoral democracy boost economic equality? *Journal of Democracy*, **20(4)**, 21-35, (2009).
[4] K. Brooks. *Improving opportunities for the rural poor in South Africa*. World Bank, unpublished Working Paper, September (2000), 30p.
[5] G. Agrawal, G. and M.A. Khan. Impact of FDI on GDP: A comparative study of china and India, *International Journal of Business and Management*, **6(10)**, 71-79, (2011).
[6] H. Görg and D. Greenaway. Much ado about nothing? do domestic firms really benefit from foreign direct investment? *The World Bank Research Observer*, **19(2)**, 171-197, (2004).
[7] L. Resmini. The determinants of foreign direct investment in the ceecs: new evidence from sectoral patterns, *Economics of transition*, **8(3)**, 665-689, (2000).
[8] A.A. Bevan and S. Estrin. *The determinants of foreign direct investment in transition economies*, Technical Report 342, William Davidson Institute at the University of Michigan (2000).
[9] F. Schneider and B.S. Frey. Economic and political determinants of foreign direct investment, *World development*, **13(2)**, 161-175, (1985).
[10] A. Klimek. Greenfield foreign direct investment versus cross-border mergers and acquisitions: the evidence of multinational firms from emerging countries, *Eastern European Economics*, **49(6)**, 60-73, (2011).
[11] A. Biswas. Forecasting Net Foreign Direct Investment Inflows in India: Box-Jenkins ARIMA Model, *International Journal of Management & Business Studies*, **5(3)**, 49-58, (2015).

[12] A.S. Henry, E.E. Elijah, A.A. Gwani and J. Simon. Time Series ARIMA Model for Predicting Nigeria Net Foreign Direct Investment, *World Scientific News*, **128(2)**, 348-362, (2019).
[13] R. Koenker and K. Hallock. Quantile regression: An introduction, *Journal of Economic Perspectives*, **15(4)**, 43-56, (2001).
[14] Z. Chunying. *A quantile regression analysis on the relations between foreign direct investment and technological innovation in china*, In Information Technology, Computer Engineering and Management Sciences (ICM), 2011 International Conference **4**, 38-41, (2011).
[15] W. Jyun-Yi and H. Chih-Chiang. Does foreign direct investment promote economic growth? evidence from a threshold regression analysis, *Economics Bulletin*, **15(12)**, 1-10, (2008).
[16] J. Paniagua, E. Figueiredo and J. Sapena. Quantile regression for the FDI gravity equation, *Journal of Business Research*, **68(7)**, 1512-1518, (2015).
[17] S. Girma and H. Görg. *Foreign direct investment, spillovers and absorptive capacity: Evidence from quantile regressions*, Technical Report 1248, Kiel Institute for the World Economy (2005).
[18] M.L. Huang and C. Nguyen. A nonparametric approach for quantile regression, *Journal of Statistical Distributions and Applications*, **5(3)**, 1-14, (2018). https://doi.org/10.1186/s40488-018-0084-9
[19] A. Karatzoglou, A. Smola and K. Hornik. *Kernlab*, R package, version 0.9-26, (2018), Available online at https://cran.r-project.org/web/packages/kernlab/index.html. (Accessed on 2 May 2018).
[20] R. Koenker and G. Bassett Jr. Regression quantiles, *Econometrica*, **46(1)**, 33-50, (1978).
[21] D.P. McMillen. Linear and nonparametric quantile regression: In Quantile Regression for Spatial Data, *Springer*, 13-27, (2013).
[22] Y. Li, Y. Liu and J. Zhu. Quantile regression in reproducing kernel Hilbert spaces, *Journal of American Statistical Association*, **102(477)**, 255-268, (2007).
[23] S. Bang, S-H. Eo, M. Jhun and H.J. Cho. Composite kernel quantile regression, *Communications in Statistics - Simulation and Computation*, **46(3)**, 2228-2240, (2017).
[24] R. Tibshirani. Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society. Series B (Methodological)*, **58(1)**, 267-288, (1996).
[25] J. Friedman, T. Hastie, N. Simon and R. Tibshirani. *Lasso and elastic-net regularized generalized linear models: glmnet*, R package, version 2.0-10, (2017). Available online at https://cran.r-project.org/web/packages/glmnet/glmnet.pdf. (Accessed on 7 March 2018).
[26] A. Khosravi, S. Nahavandi and D. Creighton. Construction of optimal prediction intervals for load forecasting problems, *IEEE Transactions on Power Systems*, **25(3)**, 1496-1503, (2010).
[27] H. Quan, D. Srinivasan and A. Khosravi. Short-term load and wind power forecasting using neural network-based prediction intervals, *IEEE transactions on neural networks and learning systems*, **25(2)**, 303-315, (2014).
[28] P.F. Christoffersen. Evaluating interval forecasts, *International economic review*, **39(4)**, 841-862, (1998).
[29] F.X. Diebold and R. Mariano. Comparing predictive accuracy, *J. Bus. Econ. Statist.*, **13(1)**, 253-265, (1995).
[30] A. Tarassow and S. Schreiber. *FEP–the forecast evaluation package for gretl*, version 2.41, (2020). Available online at http://ricardo.ecn.wfu.edu/gretl/cgi-bin/current_fnfiles/unzipped/FEP.pdf. (Accessed on 1 March 2018).
[31] U. Triacca. *Comparing Predictive Accuracy of Two Forecasts*, (2018). [Online]. Available: http://www.phdeconomics.sssup.it/documents/ Lesson19.pdf
[32] J. Lago, G. Marcjaszd, B. De Schuttera and R. Weron. Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark, *Applied Energy*, **293**, 1-21, (2021).
[33] L.C. Osabutey and C. Okoro. Political risk and foreign direct investment in Africa: The case of the Nigerian telecommunications industry, *Thunderbird International Business Review*, **57(6)**, 417-429, (2015).
[34] I. Takeuchi, Q.V. Le, T.D. Sears and A.J. Smola. Nonparametric quantile estimation, *Journal of Machine Learning Research*, **7**, 1231-1264, (2006).
[35] L. Thomas, J. Leape, M. Hanouch and R. Rumney. *Foreign direct investment in South Africa: The initial impact of the trade. Development, and Cooperation Agreement between South Africa and the European Union*, CREFSA, London School of Economics (2005).
[36] T. Hofmann, B. Schölkopf and A. J. Smola. Kernel methods in machine learning, *Ann. Statist.*, **36(3)**, 1171-1220, (2008).
[37] P. Mpfumali, C. Sigauke, A. Bere and S. Mulaudzi. Day ahead hourly global horizontal irradiance forecasting: An application to South African data, *Energies*, **12(18)**, 1-28, (2019).
[38] K. Maciejowska, J. Nowotarski and R. Weron. Probabilistic forecasting of electricity spot prices using Factor Quantile Regression Averaging, *Int. J. Forecast.*, **32**, 957-965, (2016).