**Pak. J. Statist.**
**2024 Vol. 40(2), 199-224**

# PENALIZED QUANTILE REGRESSION AND EMPIRICAL MODE DECOMPOSITION FOR IMPROVING THE ACCURACY OF THE MODEL SELECTION

**Ali S.A. Ambark**[1,2] and **Mohd Tahir Ismail**[1§]

[1] School of Mathematical Sciences,
Universiti Sains Malaysia, Pulau Pinang, Malaysia
Email: ali.ambark@sebhau.edu.ly

[2] Department of Statistics, Faculty of Science,
Sebha University, Libya

[§] Corresponding author Email: m.tahir@usm.my

## ABSTRACT

Penalized quantile regression is a hybrid statistical technique used for selecting the significant and meaningful predictor variables that strongly affect the response variable, estimating conditional quantiles, and addressing the multicollinearity among the predictor variables. The empirical mode decomposition (EMD) technique is implemented to analyze nonlinear and nonstationary data into several components called the intrinsic mode functions (IMFs) and one residue. This paper aims to apply three proposed methods for penalized quantile regression, namely, EMD-QR-R, EMD-QR-L, and EMD-QR-ENET, to determine the decomposed components of predictor variables that exhibit the strongest effects on the response variable and address the multicollinearity between the decomposed components to improve the prediction accuracy. Simulation studies and real data applications illustrate the performance and implementation of the proposed method. Results show that the penalized quantile regression method with EMD can determine the decomposition components with the most significant impact on the response variable and improve prediction accuracy.

## KEYWORD

Quantile regression; Ridge Regularization; Lasso Regularization; Elastic-net Regularization; Empirical Mode Decomposition; Multicollinearity.

## 1. INTRODUCTION

Although the least-squares method has been widely used in statistical analysis and many other applications, including but not limited to finance, economics, environmental science, and society, with the best properties such as linearity, unbiasedness, and efficiency, it is not resistant to outliers and leverage points. Furthermore, least-squares predict only the impact of predictors mostly on the conditional mean of the response. On the other hand, quantile regression was proposed by (Koenker & Bassett, 1978) and has become an alternative to least squares in the presence of outliers, which estimates the conditional quantile of the response based on the predictors. In recent decades, with rapid development, quantile regression has become an appealing statistical tool in modern regression analysis.

Quantile regression studies the relationship between the response variable and the predictor variables at any quantile of the conditional distribution function providing more comprehensive visibility of the phenomena under study. The quantile regression has no distributional assumption about the error term in the model. It can give complete information about the relationship between the response variable and predictors on the whole conditional distribution (Koenker & Bassett, 1978). Moreover, quantile regression is robust against outliers and can handle heteroscedastic datasets (as opposed to linear regression). Quantile regression has become widely used in practical applications thanks to these properties. The quantile estimator can quantify the whole conditional distribution of the response variable conditional on predictors and provide an overall assessment of the predictor variables' effects at various quantiles of the response variable (Koenker, 2005).

Variable selection has become important in the model-building process. In many applications, the number of variables is enormous. Keeping useless variables in the model, on the other hand, is undesirable since it makes the model harder to comprehend and may impair its predictive performance. Many different penalties were proposed in order to achieve variable selection. For instance, Ridge (Hoerl & Kennard, 1970), Lasso (Tibshirani, 1996), SCAD (Fan & Li, 2001), elastic-net (Zou & Hastie, 2005), Variable selection has become important in the model-building process. In many applications, the number of variables is enormous. Keeping useless variables in the model, on the other hand, is undesirable since it makes the model harder to comprehend and may impair its predictive performance. Many different penalties were proposed to achieve variable selection. For instance, Ridge (Hoerl & Kennard, 1970), Lasso (Tibshirani, 1996), SCAD (Fan & Li, 2001), elastic-net (Zou & Hastie, 2005), adaptive Lasso (Zou, 2006), adaptive elastic-net (Zou & Zhang, 2009), and MCP (Zhang, 2010). Adaptive Lasso (Zou, 2006), adaptive elastic-net (Zou & Zhang, 2009), and MCP (Zhang, 2010).

Recently, the empirical mode decomposition (EMD) approach was presented by (Huang et al., 1998), which is an intuitive, direct, and adaptable method for decomposing nonlinear and nonstationary time series data. This approach is the first part of the Hilbert-Huang transform (HHT). In contrast to traditional approaches, EMD does not impose any a priori constraints on the data (such as stationarity or linearity) but instead lets the data speak for itself. Despite this approach being completely derived from empirical evidence and lacking a formal mathematical foundation, it may efficiently divide a data series into distinct components, each corresponding to a particular oscillation frequency. Since its inception, this technique has been used in a wide variety of fields, including economics (Huang et al., 2003), engineering (Yang et al., 2003), medicine (Yang et al., 2011), physics (Varadarajan et al., 2004), and environmental science (Huang et al., 1999). This technique supplies more accurate findings in many situations than traditional methodologies, uncovering novel patterns within the analyzed data sets. EMD decomposes the original time series data into several components called intrinsic mode functions (IMFs) and one residual. One can use these components as new predictors to predict the response variable and enhance regression analysis predictive performance.

This paper is organized as follows. Section 2 presents the literature review. Section 3 describes the proposed penalized quantile regression using Ridge, LASSO, and Elastic net penalties. In section 4, we compare the proposed methods by performing numerical experiments. Section 5 displays the proposed penalized quantile regression methods applied to the daily exchange rate, and the conclusions are presented in Section 6.

## 2. LITERATURE REVIEW

In the case of high-dimensional data, variable selection and parameter estimation are crucial in quantile regression in the case of high dimensional data. It generates a sparse model that is easy to interpret, efficient, and robust to outliers. Koenker (2004) was the first to use regularization in quantile regression, he used the L1-norm penalty to decrease the random effects in mixed-effect quantile regression models for longitudinal data. The L1-penalized quantile regression method was developed by (Belloni & Chernozhukov, 2011), (Ahmed & Ismail, 2014), (Wang et al., 2018), (Bonaccolto, 2019) and (Liu et al., 2020). The variable selection in quantile regression is implemented by (Amin et al., 2015), (Peng & Wang, 2015), (Shen et al., 2018), (Khan et al., 2019), (Wang et al., 2018) and (Hu et al., 2021) with SCAD and adaptive LASSO penalties. While (Yan & Song, 2019) with an adaptive elastic-net penalty. In (Ranganai & Mudhombo, 2021) study, Ridge, LASSO, and elastic net penalties are considered. In (Burgette et al., 2011) the Lasso and elastic-net penalties are used to identify important predictors.

In addition, (Mkhadri et al., 2017) suggested a coordinate descent algorithm for computing the penalized smooth quantile regression. (Zhong et al., 2016) used quantile regression to identify important covariates for a general class of ultrahigh dimensional single-index models. Hierarchical penalty in quantile regression has been suggested in the literature by (Kang et al., 2018). (Xu et al., 2017) introduced both the sampling method and the Lasso technique to quantile regression. (Ahn & Kim, 2018) proposed a penalized competing risks quantile regression model, including the group bridge and the adaptive group bridge. A new regularization method for performing simultaneous model selection in multiple quantiles regression was introduced by (Zou & Yuan, 2008). More recently, (Ciuperca, 2020) proposed quantile regression with an adaptive elastic net penalty for computing the adaptive elastic-net group quantile estimator of the regression parameters. (Shi & Wilke, 2020) used the (adaptive) group Ridge to identify the relevant variable in heterogeneous and high-dimensional data and proved that the estimates have an oracle property. In (Su & Wang, 2021) the elastic net is used to identify important predictors for quantile regression. (Yousif & Housain, 2021) proposed a new version of penalty functions for quantile regression called Atan to estimate parameters and variable selection.

Bayesian techniques for variable selection in quantile regression have gained popularity among many researchers. For example, (Alhamzawi et al., 2012) proposed a Bayesian adaptive Lasso in quantile regression. The Bayesian elastic net penalty approach was proposed by (Alshaybawee et al., 2017) for estimating and selecting variables in a single index quantile regression model. (Tang et al., 2020) studied the quantile regression with adaptive Lasso and Lasso penalty. (Tian et al., 2021) developed Bayesian bridge-randomized penalized quantile regression.

In recent years, many studies have combined the EMD algorithm and the penalized regularization regression method to determine the impact of decomposition components on the response variable; for example, (Shen et al., 2012) used a combination of ridge regression with EEMD, (Qin et al., 2016) and (Masselot et al., 2018) used the Lasso approach with EMD. In recent years, the authors (Al-Jawarneh et al., 2020) and (Al-Jawarneh & Ismail, 2021) combined EMD and elastic net penalty to select important predictor variables with significant effects on response variables.

Another interesting study by (Tweneboah et al., 2020) used the EEMD-based QQR. (Owusu et al., 2020) used both the ensemble empirical mode decomposition (EEMD) and the quantile-on-quantile regression (QQR) on spot and futures energy and precious metal prices in India. (Zhang et al., 2020) proposed a framework for probability density forecasting of wind speed based on QR and kernel density estimation (KDE). EMD is implemented to reduce the noise of raw wind speed series.

Time series processes are frequently used to represent variables of interest, which can lead to modelling and accuracy issues. Such time series are often nonstationary and nonlinear. Moreover, several predictors may suffer from multicollinearity problems. The existence of these issues in quantile regression may lead to an increase in the variability of parameter estimations, rendering the result less dependable. These also potentially affect prediction accuracy. In practice, it is often discovered that the data exhibits violation of the linear model assumptions or that the researchers are interested in modelling other values other than the mean of the response variable, for instance, the median and other quantiles(Chatterjee & Hadi, 2013) (Chatterjee & Hadi, 2013). It is well known that the quantile regression that was proposed by (Koenker & Bassett, 1978) required no assumptions to impose on the residual term, which is used as an alternative to least squares in the presence of assumptions. To address these issues, the penalized quantile regression based on the EMD method with Ridge (RQR), Lasso (LQR), and elastic-net (ENQR) penalties are proposed for variables selection and estimation of regression coefficients to improve the performances of the predictions further. The purpose of this study aims to determine the decomposition components of the original time-series predictors that display the most substantial effects on the response variable and address multicollinearity among the decomposition components at different quantiles. Therefore, the innovation and contributions of this study can be summarized as follows:

- The Ridge, LASSO and ENET penalties procedure to the QR framework are proposed for improving the prediction accuracy in model selection. EMD is implemented to decompose the nonstationary and nonlinear predictor signals into a finite set of IMF components and one residual component.
- The decomposition components are used as new predictor variables in the QR-R, QR-L and QR-ENET methods to select the decomposition components that have the most effect on the response variable and to address multicollinearity among the components.
- The simulation study and real dataset are conducted to evaluate the performance of the proposed methods.

## 3. METHODOLOGY

### 3.1 Empirical Mode Decomposition

Empirical Mode Decomposition (EMD) has been described by (Huang et al., 1998). The EMD is an adaptive signal analysis algorithm that decomposes a nonlinear and nonstationary signal into a finite number of functions called intrinsic mode functions (IMFs) and residual components through a sifting process. The IMFs cover local instantaneous frequencies; the low-level IMF includes high local frequencies, whereas the high-level IMF covers low local frequencies. Each IMF must fulfil the following two conditions:

i. The number of local extreme values (maxima and minima) and the number of zero-crossing must be equal or by differ at most one.

ii. The local mean must be zero, defined as the mean of the upper and lower envelopes.

A time series $x(t)$ is decomposed via EMD approach into some intrinsic mode function (IMFs) $C_k(t), k = 1, \dots., K$, and as well as a residual $r(t)$ representing the overall trend of the original series. Then to reconstruct the original data $x(t)$, the sum of all IMFs and final residual is calculated as:

$$x(t) = \sum_{k=1}^{K} C_k(t) + r(t) \tag{1}$$

The IMFs are produced using an iterative process known as “sifting”. The basic aim of the EMD technique is to remove riding waves and make wave profiles more symmetric. The following is a description of the sifting procedure:

1) Identify all local extrema, including minima and maxima of a time series signal $x(t)$.

2) Produce Envelope. Connect all local extrema with a cubic spline line to generate the upper envelope $x_{up}(t)$ and the lower envelope $x_{low}(t)$, respectively.

3) Calculate the mean of the upper and the lower envelopes $m(t) = \frac{x_{up}(t)+x_{low}(t)}{2}$

4) Subtract the mean $m(t)$ from the original time series $x(t)$ to obtain the component $h(t)$ as shown $h(t) = x(t) - m(t)$

5) Check whether series $h(t)$ is an IMF or not
   i. If not an IMF, substitute $h(t)$ for $x(t)$ and repeat the sifting process, which consists of step 1 to step 4 until $h(t)$ meets the conditions of IMF.
   ii. If $h(t)$ is an IMF according to the definition of IMF, then calculate the residual as: $r(t) = x(t) - h(t)$

6) Repeat steps 1–5 until the residual satisfies the stopping criterion of the standard deviation $SD_k$, which needs the normalized squared difference between two successive sifting operations to be small. The difference is defined as follows:

$$SD_k = \sum_{t=0}^{T} \frac{h_{k-1}(t) - h_k(t)^2}{h_{k-1}^2(t)} \tag{2}$$

To decompose the time series into a collection of IMFs and a residue via the EMD approach described above. We can further analyze the signal and compute the instantaneous frequency of each converted IMF by applying the Hilbert transform to each IMF. The entire process is called the Hilbert-Huang transform (HHT) (Huang et al., 1998).

### 3.2 Regularized Quantile Regression

Quantile regression was proposed by (Koenker & Bassett, 1978). Quantile regression can estimate the response in different quantiles of the data distribution. It is incredibly beneficial when the aim of the modelling is limited to a specific area of the data distribution, such as modelling extreme values on the top quantile of the data distribution.

The simple linear quantile regression model is given by,

$$y_i = \beta_0 + x_i^T \beta + \varepsilon_i \; for \; i = 1,2,3, \dots \dots, n \tag{3}$$

where $y_i$ are response variables, $x_i^T = (x_{i1}, x_{i2}, \dots, x_{ip})$ are the $p$ known predictors, $\beta_0$ is the intercept, indicates $\beta$ a $p \times 1$ the vector of yet estimated unknown regression coefficients (parameters) and $\varepsilon_i$ are random errors whose $\tau$th conditional quantile given Xi equals zero for $\tau \in (0,1)$. Then the $\tau th$ conditional quantile function may be estimated by solving

$$\hat{\beta}_\tau = arg_\beta min \sum_{i=1}^{n} \rho_\tau \left(y_i - \beta_0 - x_i^T \beta\right) \tag{4}$$

where $\rho_\tau$ the check function defined by

$$\rho_\tau(u) = \begin{cases} \tau u \; if \; u > 0 \\ (\tau - 1)u \; if \; u \leq 0 \end{cases}$$

#### 3.2.1 Ridge Penalty

(Hoerl & Kennard, 1970) introduced Ridge regression as one of the most popular alternative solutions to OLS. This method is used to improve the estimation of regression parameters in the case multicollinearity is present among the predictor variables.

The ridge estimate is given by

$$\beta(\tau R) = argmin \left\{ \sum_{i=1}^{n} \left( y_i - \beta_0 + \sum_{j=1}^{p} x_{ij} \beta_j \right) + \lambda \sum_{j=1}^{p} \beta_j^2 \right\} \tag{5}$$

where $\lambda$ is a positive ridge parameter in the range $0 < \lambda < 1$.

We consider QR penalized with the Ridge penalty (7) denoted by QR-R. The QR-R is given by the minimization problem of

The quantile regression with ridge regression uses the ridge coefficients to build the penalized quantile regression model. The solution of the ridge coefficient can be written as the following

$$\hat{\beta}_\tau = arg_\beta min \sum_{i=1}^{n} \rho_\tau \left( y_i - \beta_0 - \sum_{j=1}^{p} x_{ij} \beta_j \right) + \lambda \sum_{j=1}^{p} \beta_j^2 \tag{6}$$

### 3.2.2 LASSO Penalty

The LASSO (Least Absolute Shrinkage and Selection Operator) was proposed by (Tibshirani, 1996) which can effectively identify important predictor variables and estimate regression coefficients simultaneously.

LASSO imposes a slightly different penalty on the coefficient vector $j, j = 1,2, \ldots k - 1$ than Ridge. In the case of LASSO, the $\lambda$ parameter is multiplied by the $L_1$-norm of the vector $(\beta_1 \ldots \beta_{k-1})$ whereas Ridge uses the $L_2$-norm. The LASSO estimates are defined as

$$\beta(\text{LASSO}) = argmin\left\{\sum_{i=1}^{n}\left(y_i - \beta_0 + \sum_{j=1}^{p} x_{ij}\beta_j\right) s.t. \sum_{j=1}^{p}\left|\beta_j\right| \leq s\right\} \quad (7)$$

where $s \geq 0$ is a tuning parameter. The LASSO penalty is often called an $L_1$ penalty because of the first power in the penalty term.

The LASSO quantile regression estimates are defined as

$$\beta(\text{LASSO}) = argmin\left\{\sum_{i=1}^{n}\left(y_i - \beta_0 + \sum_{j=1}^{p} x_{ij}\beta_j\right) + \lambda \sum_{j=1}^{p}\left|\beta_j\right|\right\} \quad (8)$$

We first consider quantile regression with the LASSO penalty. The lasso regularized quantile regression (Li & Zhu, 2008) is given by

$$\hat{\beta}_\tau = arg_\beta min \sum_{i=1}^{n} \rho_\tau\,(y_i - \beta_0 - x_i^T\beta) + \lambda \sum_{j=1}^{p}\left|\beta_j\right|,$$

where $\lambda$ is the penalty parameter (regularizer) that controls the amount of shrinkage.

### 3.2.3 Elastic Net Penalty

The elastic net penalty proposed by (Zou & Hastie, 2005) combines the ridge and LASSO penalties. Elastic Net coefficient estimates are obtained by using an Elastic Net penalty to reduce the regression loss function.

$$\sum_{j=1}^{p}\left[\alpha\left|\beta_j\right| + (1-\alpha)\beta_j^2\right] \leq k \quad (9)$$

when the value of $\alpha = 0$ then we get to ridge regression, but if $\alpha = 1$, then we get LASSO regression and the Elastic Net penalty for $0 \leq \alpha \leq 1$. Regularized quantile regression based on elastic-net regularized is defined as:

$$\hat{\beta}_\tau = argmin\left\{\sum_{i=1}^{n} \rho_\tau(y_i - \beta_0 - x_i^T\beta) + \lambda \sum_{j=1}^{p}\left[\alpha\left|\beta_j\right| + (1-\alpha)\beta_j^2\right]\right\} \quad (10)$$

### 3.3 Multicollinearity

Multicollinearity is a common problem in multiple regression analysis. It occurs when two or more predictor variables in a regression model are highly correlated. Such an interrelationship between explanatory variables obscures their link with the explained variable. Furthermore, the reliability of the regression analysis is decreased and reduced in the presence of multicollinearity by the low quality of the resultant estimates. The variance inflation factor (VIF) test will be used to measure the correlation value between predictor variables. When variance inflation factor (VIF) values are less than 10, there is no multicollinearity (Januaviani et al. 2019; Tamura et al., 2017).

$$VIF_j = \frac{1}{1 - R_j^2}. \tag{11}$$

### 3.4 Proposed Methods

In this section, based on the principles of Regularized Quantile regression and EMD introduced in previous sections, QR-R, QR-L and QR-ENET methods based on the EMD method are proposed to explain the effects of the decomposition components via EMD on the response variable and addressing multicollinearity as shown in Figure 1. The proposed method can be summarized as the following steps:

#### 3.4.1 Ridge Quantile Regression Based on EMD (EMD-QR-R)

**Step 1:** The EMD method was used to decompose the original time series data $x(t)$ into several components named $IMF_k$ $(k = 1,2, \ldots . K)$, as well as a residual component r.

$$x(t) = \sum_{k=1}^{K} C_k(t) + r(t) \tag{12}$$

**Step 2:** All the decomposed components obtained from the predictor variable $x(t)$ are utilized to select the subset of components that exhibited the most impact (Masselot et al., 2018):

$$y(t) = \sum_{j=1}^{p} \left[ \sum_{k=1}^{K} C_{jk}\beta_{jk} + r_{jk}(t)\beta_{jk} \right] + \varepsilon(t) \tag{13}$$

**Step 3:** The check whether there is a multicollinearity problem between the decomposition components by utilizing Variance Inflation Factor (VIF).

$$\mathrm{VIF_j} = \frac{1}{1 - \mathrm{R_j^2}} \tag{14}$$

**Step 4:** The optimal tuning parameter value is determined, which gives the minimum MSE based on the 10-fold CV method.

$$\lambda_R = \underset{s=1,2,\ldots.S}{argmin}\left(MSE_{\lambda_s}\right) \tag{15}$$

$$MSE_{\lambda_s} = \frac{1}{10} \sum_{d=1}^{10} RSS_{\lambda_s}^{d}$$

**Step 5:** Apply the Ridge penalized quantile regression method:

$$\hat{\beta}_\tau = arg_\beta min \sum_{i=1}^{n} \rho_\tau \left( y_i - \beta_0 - \sum_{j=1}^{p} \left( \sum_{k=1}^{K} C_{jk}(t)\, \beta_{jk} - r_j(t)\beta_{jk+1} \right) \right) + \lambda \sum_{j=1}^{p} \beta_j^2 \tag{16}$$

QR with the Ridge (L2) penalty has been proposed as a solution to the multicollinearity.

### 3.4.2 LASSO Quantile Regression Based on EMD (EMD-QR-L)

The first three steps are the same as in the EMD-QR-R method, and the remaining step are as follows:

**Step 4:** The optimal tuning parameter value is determined, which gives the minimum MSE based on the 10-fold CV method.

$$\lambda_L = \underset{s=1,2,\ldots.S}{argmin}\left(MSE_{\lambda_s}\right) \tag{17}$$

$$MSE_{\lambda_s} = \frac{1}{10}\sum_{d=1}^{10} RSS_{\lambda_s}^{d}$$

**Step 5:** Apply the LASS0 penalized quantile regression method:

$$\hat{\beta}_\tau = arg_\beta min \sum_{i=1}^{n} \rho_\tau \left( y_i - \beta_0 - \sum_{j=1}^{p} \left( \sum_{k=1}^{K} C_{jk}(t)\, \beta_{jk} - r_j(t)\beta_{jk+1} \right) \right) + n\lambda \sum_{j=1}^{p} \left|\beta_j\right| \tag{18}$$

QR with the LASSO (L1) penalty has been proposed as a solution to variable selection.

### 3.4.3 Elastic Net Quantile Regression Based on EMD (EMD-QR-ENET)

The first three steps are the same as in the EMD-QR-R method, and the remaining step are as follows:

**Step 4:** The optimal tuning parameter value is determined, which gives the minimum MSE based on the 10-fold CV method.

$$\lambda_{Enet} = \underset{s=1,2,\ldots.S}{argmin}\left(MSE_{\lambda_s}\right) \tag{19}$$

$$MSE_{\lambda_s,\alpha} = \frac{1}{10}\sum_{d=1}^{10} RSS_{\lambda_s,\alpha}^{d}$$

**Step 5:** Apply the Elastic Net penalized quantile regression method:

$$\hat{\beta}_{\tau} = \arg min \left[ \sum_{i=1}^{n} \rho_{\tau} \left( y_i - \beta_0 - \left( \sum_{j=1}^{p} \left( \sum_{k=1}^{K} C_{jk}(t)\, \beta_{jk} - r_j(t)\beta_{jk+1} \right) \right) \right) + \lambda \sum_{j=1}^{p} \left[ \alpha \left| \beta_j \right| + (1-\alpha)\beta_j^2 \right] \right] \tag{20}$$

**Step 6:** In this study, we will use the simulation and a real-world example to compare the performances of the prediction methods proposed with those of traditional methods, namely, QR-R, QR-L, and QR-ENET.

To compare the accuracy performances between the proposed methods and the existing methods, namely, R-QR, L-QR, and ENET-QR. Evaluation metrics are used as follows: the residual sum of squares (RSS), the prediction error (PE) the mean squared error (MSE), and the mean absolute error (MAE).To compare accuracy performances between the proposed methods with the existing methods, namely, R-QR, L-QR and ENET-QR. Evaluation metrics are used following: the residual sum of squares (RSS), prediction error (PE), the mean squared error (MSE) and Mean Absolute Error (MAE).

$$PE = \frac{1}{n} \sum_{i=1}^{n} \rho_{\tau}(y_i - \hat{y}_i) \tag{21}$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{22}$$

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{23}$$

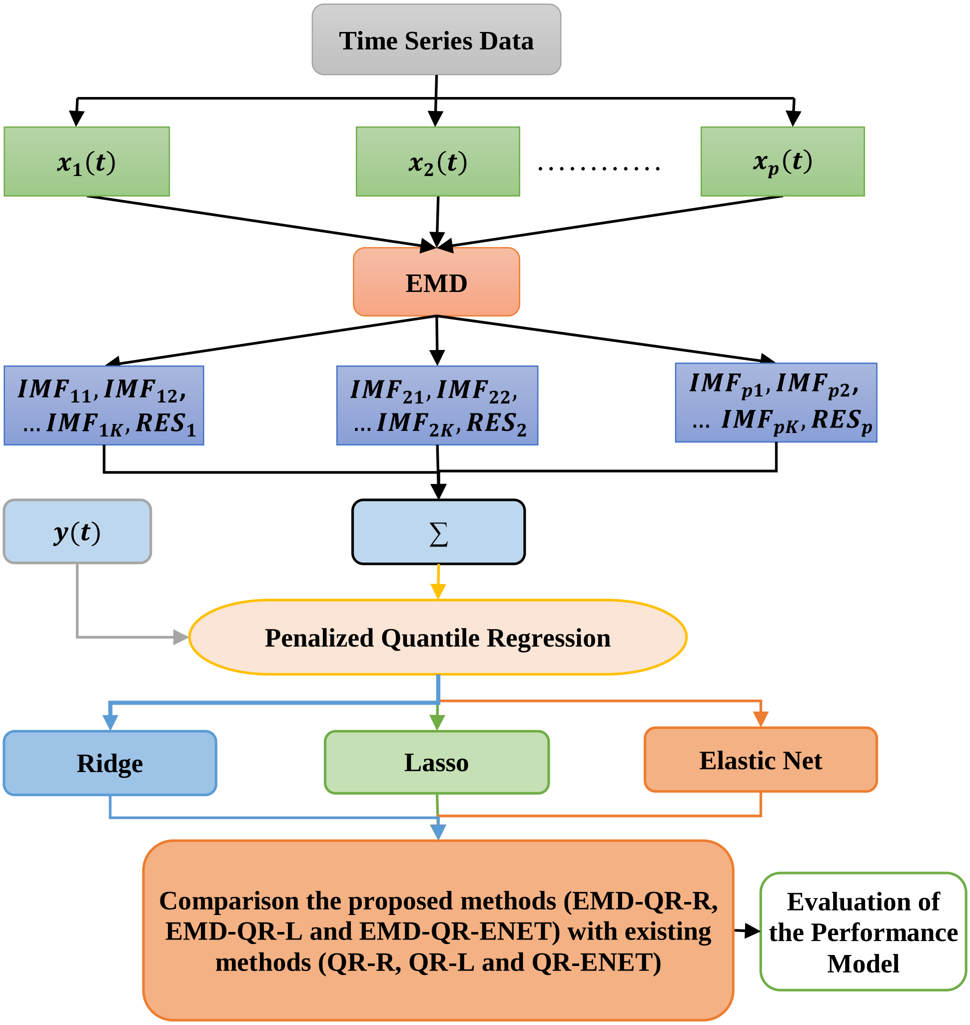


**Figure 1: Flowchart of the Proposed**

## 4. NUMERICAL EXPERIMENTS

In this section, we evaluate the finite sample property of our proposed methods and compare it with other methods through simulation studies and empirical analysis of a real-world example to demonstrate that the performance of our proposal methods represent an effective variable-selection procedure and enhances prediction accuracy.

### 4.1 Simulation Study

In this section, we perform a simulation study to investigate the finite-sample performance of penalized quantile regression under the Ridge, the LASSO, and the ENET penalty functions. The predictor variables $(x_1, x_2, x_3)$ and response variables $(y(t)$ were simulated from actual signals selected from the work of (Al-Jawarneh et al., 2021; Qin et al., 2016).

Here, two different scenarios are considered where we consider that types of random errors are $N{\sim}(0{,}1)$ and $X^2{\sim}(2)$, for $\varepsilon$. Throughout the simulations we generated, we assume the sample size is 250, and the time-domain between ($0 \le t \le 9$), We repeated each experiment 500 times at three different quantiles, namely $\tau = (0.25, 0.5, 0.75)$.

$$y(t) = \sin(\pi t) + \sin(2\pi t) + \cos(6\pi t) + \cos(13\pi t) + \varepsilon$$

$$x_1(t) = \sin(2\pi t) + \cos(\pi t) + \sin(5\pi t) + \sin(9\pi t) + \varepsilon$$

$$x_2(t) = 0.2t + \sin(\pi t) + \cos(6\pi t) + \cos(9\pi t) + \varepsilon$$

$$x_3(t) = \sin(\pi t) + \sin(8\pi t) + \cos(7\pi t) + \cos(13\pi t) + \varepsilon$$

### 4.2 Results and Discussion

We compared the suggested modeling process's prediction accuracy to three traditional approaches utilized in the literature: (1) ridge penalized quantile regression between the response variable $y$, and all the decomposed components of predictor variables ($x_j$); (2) lasso penalized quantile regression between the response variable y and all the decomposed components of predictor variables ($x_j$); (3) elastic net penalized quantile regression between the response variable y and all the decomposed components of predictor variables ($x_j$).

**Table 1**
**Variance Inflation Factors**

| $C_{1,1}$ | $C_{1,2}$ | $C_{1,3}$ | $C_{1,4}$ | $C_{1,5}$ | $C_{1,6}$ | $C_{1,7}$ | $R_1$ | $C_{2,1}$ | $C_{2,2}$ | $C_{2,3}$ | $C_{2,4}$ | $C_{2,5}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1.049 | 1.064 | 1.082 | 1.162 | 2.099 | 2.234 | 63.610 | 14.844 | 1.064 | 1.086 | 1.159 | 1.463 | 3.427 |
| $C_{2,6}$ | $C_{2,7}$ | $C_{2,8}$ | $R_2$ | $C_{3,1}$ | $C_{3,2}$ | $C_{3,3}$ | $C_{3,4}$ | $C_{3,5}$ | $C_{3,6}$ | $C_{3,7}$ | $C_{3,8}$ | $R_3$ |
| 14.833 | 55.882 | 103.026 | 50.144 | 1.088 | 1.099 | 1.039 | 1.295 | 2.004 | 3.677 | 44.889 | 178.595 | 23.972 |

Table 1 shows the multicollinearity test among the number of decomposition components extracted by the EMD algorithm. We decomposed $x_1(t)$ and $x_3(t)$ to five IMF components and a residual component, $x_2(t)$ to seven IMF components and a residual component, Table 1 presents the VIF values for the decomposition components. The result indicates that some of these values are greater than 10, indicating a high multicollinearity problem exists between the decomposition components.

Table 2 shows the optimal values of the tuning parameter $\lambda$ that were chosen using a 10-fold CV (cross-validation). In addition, Tables 3 and 4 summarize the simulation results of the prediction performance relying on lambda values that were chosen in Table 2 in terms of Bias, MASE, MAE, and MSE for the cases of $\varepsilon \sim N(0{,}1)$ and $\varepsilon \sim \chi 2(2)$, respectively, to compare the results of the QR-R, QR-L, and QR-ENET models based on EMD with the QR-R, QR-L, and QR-ENET models without EMD with the quantile indices $\tau = 0.25$, 0.50 and 0.75. The superiority of the proposed methods is also obvious in the prediction accuracy at different quantiles, especially when $\theta = 0.5$ in terms of the averages of Bias, MASE, MAE, and MSE.

**Table 2**
**Optimal Values of λ in Proposed Methods via 10-fold Cross Validation**

| Quantile | Method | $\varepsilon \sim N(0,1)$ | | $\varepsilon \sim X^2(2)$ | |
|---|---|---|---|---|---|
| | | $\lambda_{min}$ | $\lambda_{1se}$ | $\lambda_{min}$ | $\lambda_{1se}$ |
| 0.25 | EMD-QR-R | 0.09226464 | 0.35545490 | 0.14819489 | 0.47563898 |
| | EMD-QR-L | 0.01626342 | 0.04393155 | 0.02121867 | 0.05878540 |
| | EMD-QR-ENET | 0.03769530 | 0.09440782 | 0.04687355 | 0.05053135 |
| 0.50 | EMD-QR-R | 0.12347958 | 0.38727731 | 0.14174847 | 0.49723369 |
| | EMD-QR-L | 0.01779694 | 0.05085077 | 0.02922003 | 0.05282554 |
| | EMD-QR-ENET | 0.06001606 | 0.04571545 | 0.04699365 | 0.26338103 |
| 0.75 | EMD-QR-R | 0.10659933 | 0.20269875 | 0.14094516 | 0.86923302 |
| | EMD-QR-L | 0.01941768 | 0.04319107 | 0.01188580 | 0.02924384 |
| | EMD-QR-ENET | 0.02572568 | 0.03944290 | 0.03715861 | 0.02063793 |

Table 3 shows the results of the simulation experiment for the case of $\varepsilon \sim N(0,1)$ We find that the EMD-QR-R, EMD-QR-L and EMD-QR-ENET are superior to the other models in terms of estimation, and predictive ability. In addition, the difference between $\tau = 0.25$, $\tau = 0.5$, and $\tau = 0.75$ is relatively large. Such as, the results of the Bias values show that the smallest value is achieved using the EMD-QR-ENET method, the value changes from 0.7248 when $\tau = 0.25$ to 0.5067 by decreasing 30.09%, then changes to 0.7682 by increasing 5.65% when $\tau = 0.50$. Furthermore, in terms of the averages of PE and MAE, the smallest value of PE is 0.0061751 achieved by EMD-QR-ENET when $\tau = 0.5$, while the biggest value is 0.4779517 achieved by QR-R with $\tau = 0.75$. For MAE the smallest value is 0.5654065 achieved by EMD-QR-ENET with $\tau = 0.5$, while the biggest value is achieved by QR-R with $\tau = 0.25$.

**Table 3**
**Simulation Results for the Case of $\varepsilon \sim N(0,1)$ with different τ**

| τ | Method | MSE | Bias | MAE | PE | Optimal λ |
|---|---|---|---|---|---|---|
| 0.25 | QR-R | 1.4940271 | 0.50001 | 1.0004058 | 0.1767725 | 0.35674556 |
| | QR-L | 1.4937402 | 0.48301 | 1.0000753 | 0.1743084 | 0.05255001 |
| | QR-ENET | 1.4930349 | 0.48509 | 0.9999021 | 0.1743408 | 0.10947919 |
| | EMD-QR-R | 0.7431496 | 0.21558 | 0.6774484 | 0.1154234 | 0.10423921 |
| | EMD-QR-L | 0.7131301 | 0.20349 | 0.6665723 | 0.1121351 | 0.01672008 |
| | EMD-QR-ENET | 0.7152909 | 0.20519 | 0.6669050 | 0.1131898 | 0.03624545 |
| 0.50 | QR-R | 0.9784170 | 0.00489 | 0.7970440 | 0.0298182 | 0.343735237 |
| | QR-L | 0.9798689 | 0.00509 | 0.7978155 | 0.0303973 | 0.005068826 |
| | QR-ENET | 0.9795799 | 0.00506 | 0.7975363 | 0.0303197 | 0.010560053 |
| | EMD-QR-R | 0.5242612 | 0.00151 | 0.5787048 | 0.0071754 | 0.128049813 |
| | EMD-QR-L | 0.5008937 | 0.00154 | 0.5658945 | 0.0063031 | 0.023141965 |
| | EMD-QR-ENET | 0.5001667 | 0.00153 | 0.5654065 | 0.0061751 | 0.050104191 |
| 0.75 | QR-R | 1.3741551 | 0.4040219 | 0.9299064 | 0.4779517 | 0.65621619 |
| | QR-L | 1.3636166 | 0.3931201 | 0.9279268 | 0.4719328 | 0.02796614 |
| | QR-ENET | 1.3640482 | 0.3929475 | 0.9280903 | 0.4721575 | 0.04714104 |
| | EMD-QR-R | 0.7584165 | 0.2110771 | 0.7044213 | 0.3443763 | 0.10674075 |
| | EMD-QR-L | 0.7053932 | 0.1896935 | 0.6810476 | 0.3243694 | 0.01627920 |
| | EMD-QR-ENET | 0.7148415 | 0.1955307 | 0.6850912 | 0.3305906 | 0.03892253 |

Table 4 presents the simulation results for the proposed methods and the other considered methods for the data generated from $\varepsilon \sim X^2(2)$. It can be observed that the performance of the proposed methods achieves the smallest MSAE, MAE, and PR compared with the QR-R, QR-L, and QR-ENET methods, which shows that the proposed model has the ability to achieve more prediction accuracy. From the results in Figures 4 and 5, it is evident that the reported test errors (MSE and MAE) of the proposed methods are the lowest among the considered regularized methods in almost both simulation scenarios.

**Table 4**
**Simulation Results for the Case of $\boldsymbol{\varepsilon \sim X^2(2)}$ with Different $\boldsymbol{\tau}$**

| τ | Method | MSE | Bias | MAE | PE | Optimal λ |
|---|---|---|---|---|---|---|
| **0.25** | QR-R | 1.504268 | 0.508294 | 1.0106202 | 0.1784608 | 0.329169 |
| | QR-L | 1.496153 | 0.493564 | 1.0085365 | 0.1760503 | 0.042175 |
| | QR-ENET | 1.494747 | 0.493854 | 1.0080611 | 0.1760542 | 0.085246 |
| | EMD-QR-R | 1.014441 | 0.288913 | 0.8098327 | 0.1342447 | 0.153227 |
| | EMD-QR-L | 1.006221 | 0.282986 | 0.8049735 | 0.1332406 | 0.020666 |
| | EMD-QR-ENET | 1.004010 | 0.28633 | 0.8049404 | 0.1340565 | 0.056242 |
| **0.50** | QR-R | 0.9910238 | 0.00571 | 0.7996400 | -0.03482165 | 0.385199 |
| | QR-L | 0.9916905 | 0.005454 | 0.7997428 | -0.03356119 | 0.039013 |
| | QR-ENET | 0.9914121 | 0.005596 | 0.7997226 | -0.03390704 | 0.097459 |
| | EMD-QR-R | 0.7266575 | 0.002003 | 0.6889818 | -0.00820746 | 0.190364 |
| | EMD-QR-L | 0.7053599 | 0.001654 | 0.6787889 | -0.00459781 | 0.024484 |
| | EMD-QR-ENET | 0.7049594 | 0.001850 | 0.6784066 | -0.00526482 | 0.064675 |
| **0.75** | QR-R | 1.3778080 | 0.398878 | 0.9312214 | -0.4744644 | 0.518371 |
| | QR-L | 1.3702163 | 0.389343 | 0.9297602 | -0.4682408 | 0.117851 |
| | QR-ENET | 1.3701583 | 0.389523 | 0.9295945 | -0.4678270 | 0.245523 |
| | EMD-QR-R | 1.0166191 | 0.289679 | 0.8110795 | -0.4024530 | 0.152722 |
| | EMD-QR-L | 0.9909606 | 0.278224 | 0.8029164 | -0.3942394 | 0.017734 |
| | EMD-QR-ENET | 0.9938541 | 0.280586 | 0.8030988 | -0.3960574 | 0.054957 |

From Figures 2 and 3, we see that the EMD-QR-L and EMD-QR-ENET methods can select the true significant decomposition components and parameter estimation simultaneously in the real problem, except for the EMD-QR-R method that selects all decomposition components in the final model at different quantiles. That is, the EMD-QR-R method cannot do decomposition component selection. The result from Figure 2 indicates that the EMD-QR-L and EMD-QR-ENET methods select approximately 16 and 18 non-zero coefficients, respectively, at all three levels of quantile (i.e. 0.25, 0.50, and 0.75). To sum up, the results of this study reveal that the proposed methods achieve better performance than traditional methods. The proposed methods have the lowest test error as compared to all competitors at all levels of quantiles because of their robust nature.

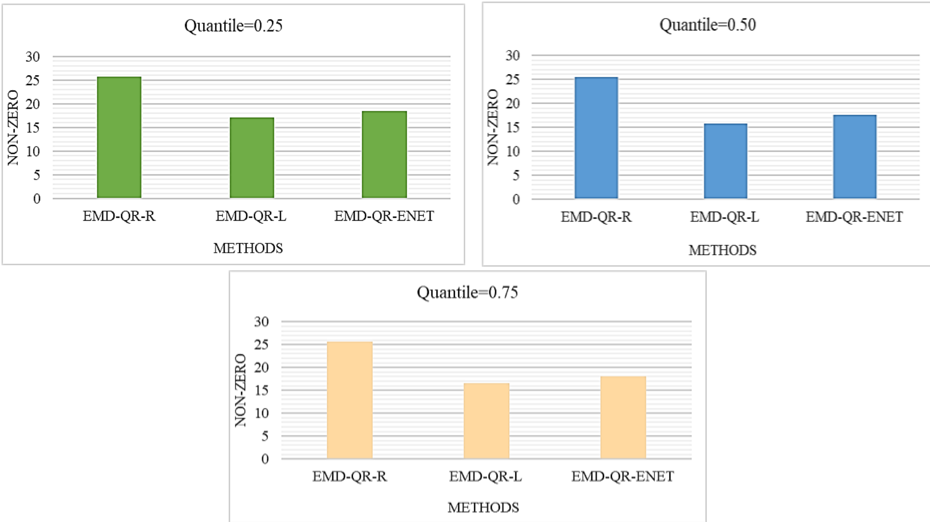


**Figure 2: The Number of Non-Zero Coefficient with $\boldsymbol{\varepsilon \sim N(0,1)}$ with different $\boldsymbol{\tau}$**

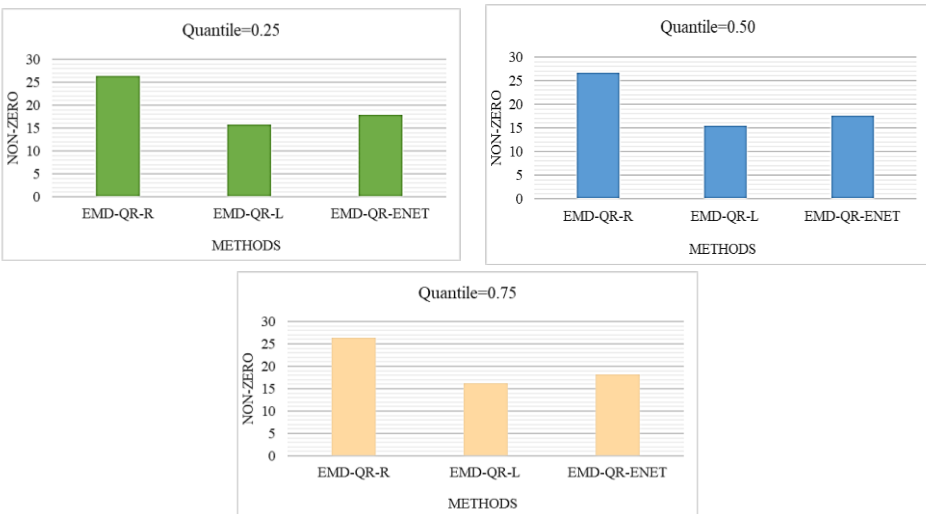


**Figure 3: The Number of Non-Zero Coefficient with $\boldsymbol{\varepsilon \sim X^2(2)}$ with different $\boldsymbol{\tau}$**

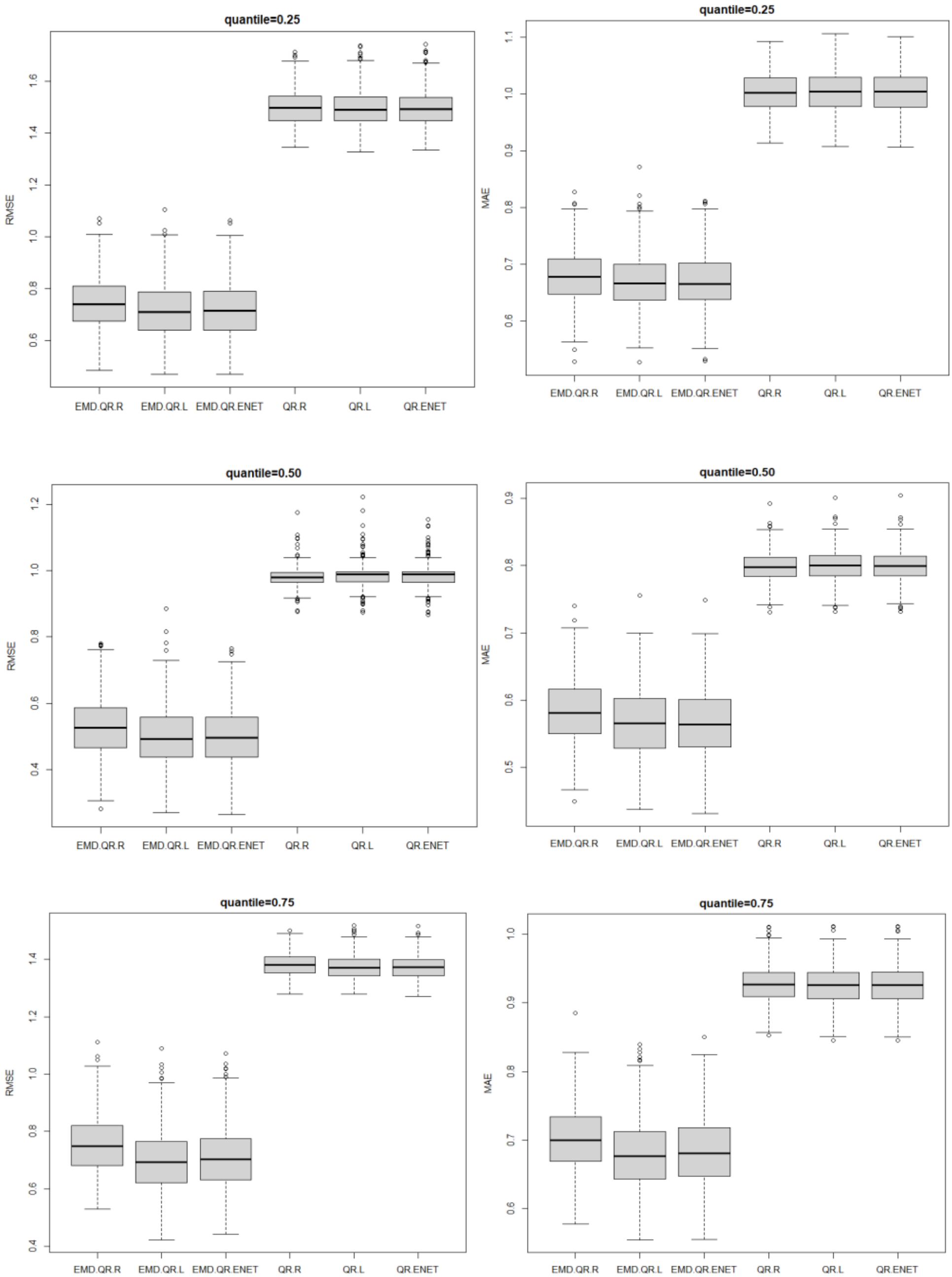


**Figure 4: Boxplots for MSE and MAE for the case of $\varepsilon \sim N(0, 1)$ with different τ**

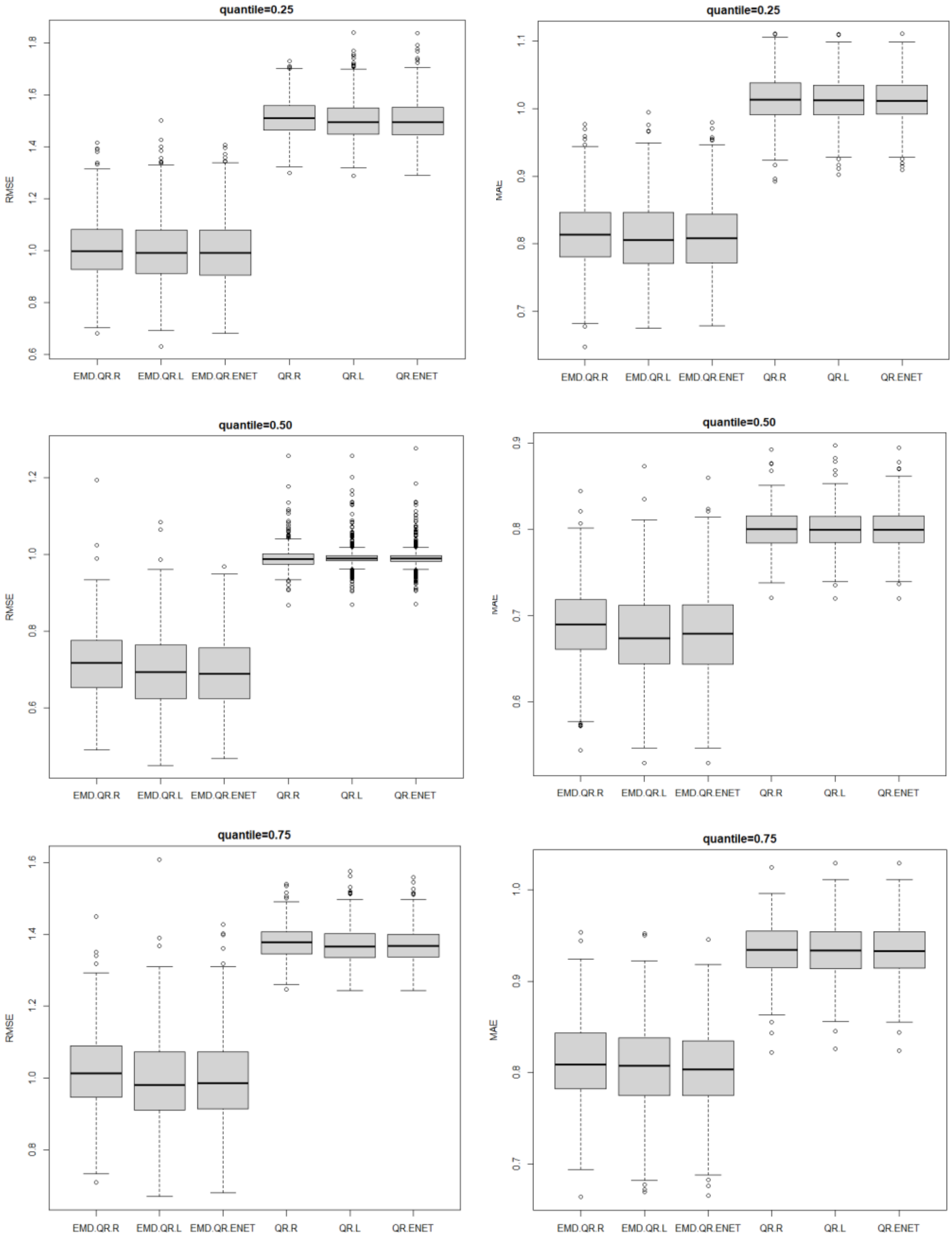


**Figure 5. Boxplots for MSE and MAE for the case of $\varepsilon \sim X^2(2)$ with different $\tau$**

## 5. APPLICATION

In this section, we present an empirical analysis of a real-world example based on the four countries' daily exchange rates against the US dollar (USD) to evaluate the performance of the proposed methods.

### 5.1 A Real Data Example

We use the data set used in (Al-Jawarneh et al., 2021) to illustrate the application of the proposed methods, the data set contains the close daily exchange rates of four countries against the USD. Those selected countries were Taiwan (TAW/USD) as the response variable, Malaysia (MYR/USD), Japan (JAP/USD), and China (CHN/USD) as predictor variables from Mar. 27, 2015 to Oct. 25, 2019 and consisted of 1,196 observations. These data can also be obtained from the Wall Street Journal database https://www.wsj.com.

### 5.2 Results and Discussion for Daily Exchange Rates

Figure 6 shows the plot of the daily close exchange rates for the predictor variables MALAYSIA $(x1(t))$, JAPAN $(x2(t))$, and CHINA $(x3(t))$. When the TAIWAN is the response variable $y(t)$, the signals show both long-term trends and short-term fluctuations in overtime, as can be seen in (Figure 6) which shows that the signals are nonstationary and nonlinear. Therefore, it is better to apply penalized quantile regression based on EMD methods, namely, EMD-QR-R, EMD-QR-L and EMD-QR-ENET compared with penalized quantile regression methods. To assess the prediction ability of the selected model and parameters estimation, 70% of the observations are used as training data and the remaining 30% as test data. Table 5 presents the descriptive statistics of the variables in the model to characterize the nature of the variables considered. The following is an analytic presentation of these measurements for each model variable.

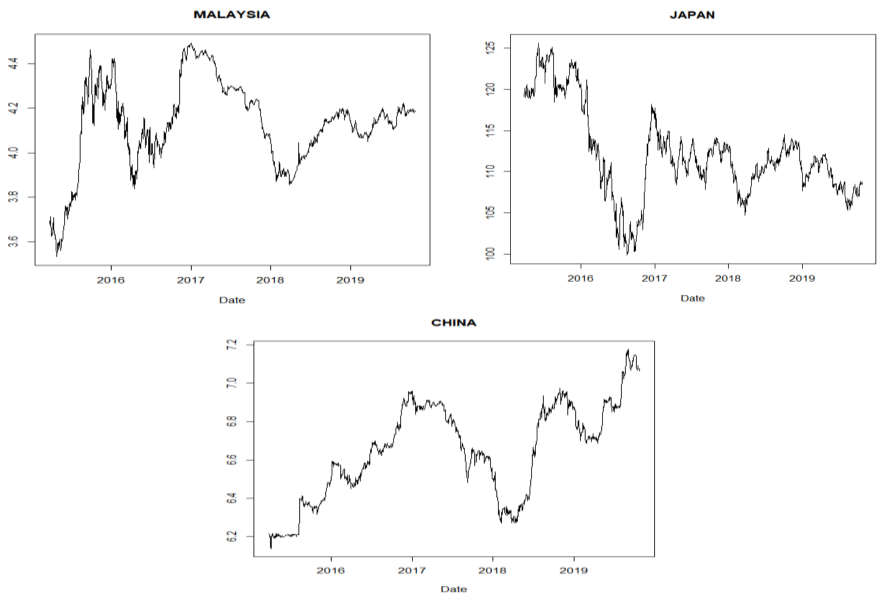


**Figure 6: The Daily Close Exchange Rates are Plotted Over Time**

**Table 5**
**Descriptive Statistics**

| Variables | Mean | Standard Deviation | Skewness | Kurtosis |
|---|---|---|---|---|
| **TAW/USD (y)** | 29.92712 | 1.0213564 | 0.3347141 | 2.549761 |
| **MYR/USD (x1)** | 4.170099 | 0.1939055 | -0.5551624 | 3.439429 |
| **JAP/USD(x2)** | 109.4827 | 5.4782274 | 0.4567056 | 2.908487 |
| **CHN/USD (x3)** | 6.742283 | 0.2476597 | -0.1838290 | 2.086378 |

Figure 7 shows the decomposition components via the sifting for the EMD algorithm of the original predictors. The MALAYSIA $x1(t)$ is decomposed into eight IMF components and a residual component, JAPAN $x2(t)$ is decomposed into eight MF components and a residual component, while CHINA $x3(t)$ is decomposed into eight IMF components and a residual component.

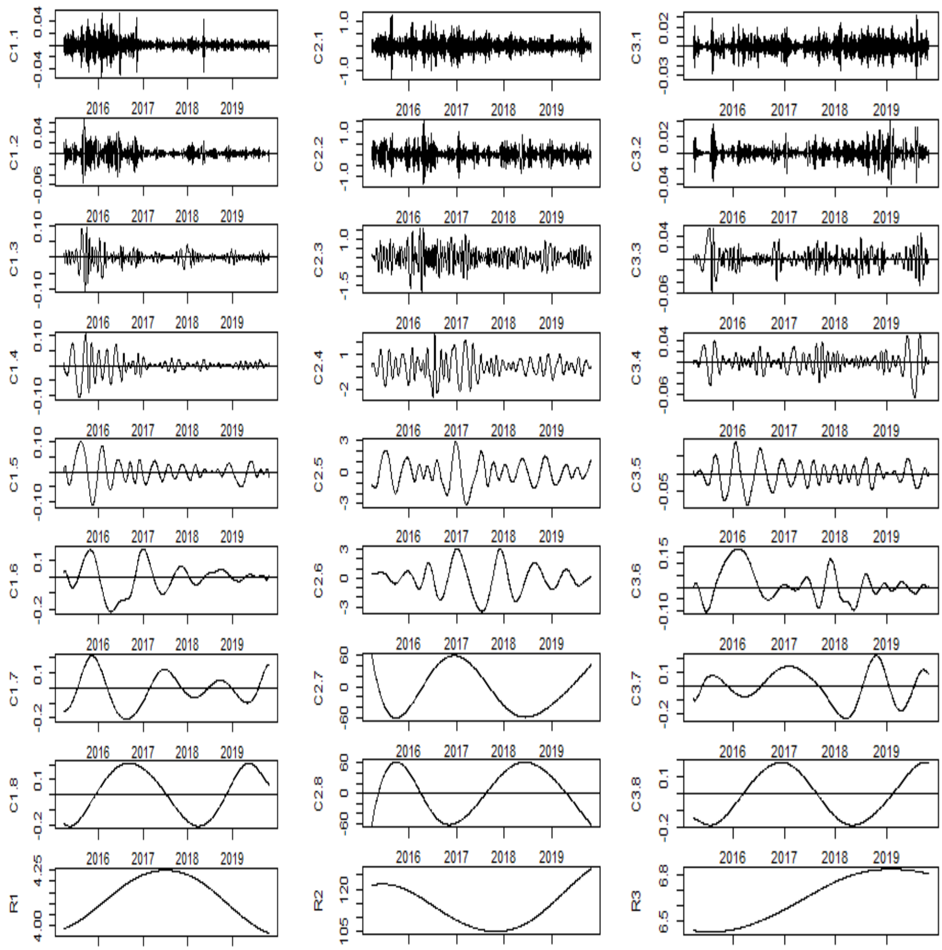


**Figure 7: EMD Decomposition of the Daily Close Exchange Rates**

Table 6 illustrates the Variance Inflation Factor (VIF), which measures the inflation of parameter estimates for all the decomposition components of the model's MYR, JAP, and CHN variables, which will be used to identify the factors causing the problem of multicollinearity. These indicators were calculated for the regression parameters of all the model's decomposition components. The multicollinearity of the decomposition components was calculated as in Table 6.

Table 6 shows that the VIF values for some of the decomposition components ($C_{1,8}$, $C_{2,5}$, $C_{2,6}$, $C_{2,8}$, $R_2$, $C_{3,8}$ and $R_3$) are greater than 10, indicating a high multicollinearity problem exists between the decomposition components of the MYR, JAP and CHN variables.

**Table 6**
**Variance Inflation Factors (VIF)**

| $C_{1,1}$ | $C_{1,2}$ | $C_{1,3}$ | $C_{1,4}$ | $C_{1,5}$ | $C_{1,6}$ | $C_{1,7}$ | $C_{1,8}$ | $R_1$ | $C_{2,1}$ | $C_{2,2}$ | $C_{2,3}$ | $C_{2,4}$ | $C_{2,5}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1.05 | 1.04 | 1.12 | 1.10 | 2.19 | 3.05 | 4.89 | 14.95 | 4.76 | 1.27 | 1.24 | 1.12 | 1.08 | 13.36 |
| $C_{2,6}$ | $C_{2,7}$ | $C_{2,8}$ | $R_2$ | $C_{3,1}$ | $C_{3,2}$ | $C_{3,3}$ | $C_{3,4}$ | $C_{3,5}$ | $C_{3,6}$ | $C_{3,7}$ | $C_{3,8}$ | $R_3$ | |
| 15.86 | 5.52 | 28.65 | 20.58 | 1.67 | 1.36 | 1.34 | 1.19 | 1.38 | 2.05 | 8.11 | 20.89 | 90.56 | |

**Table 7**
**Optimal Values of λ in Proposed Methods via 10-Fold Cross Validation**

| $\tau$ | EMD-QR-R | | EMD-QR-L | | EMD-QR-ENET | |
|---|---|---|---|---|---|---|
| | $\lambda_{min}$ | $\lambda_{1se}$ | $\lambda_{min}$ | $\lambda_{1se}$ | $\lambda_{min}$ | $\lambda_{1se}$ |
| 0.25 | 0.168877 | 0.179413 | 0.016888 | 0.021513 | 0.038381 | 0.042029 |
| 0.50 | 0.179993 | 0.191223 | 0.017999 | 0.020939 | 0.040908 | 0.043459 |
| 0.75 | 0.147613 | 0.166607 | 0.014761 | 0.016164 | 0.033549 | 0.040228 |

Table 7 presents the optimal values of the tuning parameter λ achieved via 10-fold cross-validation in proposed methods. Table 8 reports the comparison results of three proposed methods with traditional methods. Performance is evaluated via MSE, MAE, and PE. A lower Bias, MSE, MAE, and PE indicates a better performance. Due to the limited space, we only report the result of three levels of quantiles, i.e. $\tau = (0.25, 0.5, 0.75)$.

By comparison, we can see that the proposed methods achieve the smallest value of MSE, MAE, and PE among all of the other methods. For example, the smallest value of MSE of the QR-R, QR-L, QR-ENET, EMD-QR-R, EMD-QR-L and EMD-QR-ENET respectively are 1.0567778, 0.9185961, 1.0031082, 0.6618645, 0.5557124 and 0.4897215 with $\tau = 0.5$, while the biggest value are 2.1089951, 2.1664858, 2.1515059, 0.8186193, 0.5945223 and 0.5456317 with $\tau = 0.25$. The result illustrates the significant superiority of the proposed methods in prediction accuracy. We also provided a bar chart (see Figure 8) that shows the MSE results for each method in our study. The chart indicates that penalized quantile regression with EMD achieves remarkably less prediction error than penalized quantile regression without EMD.

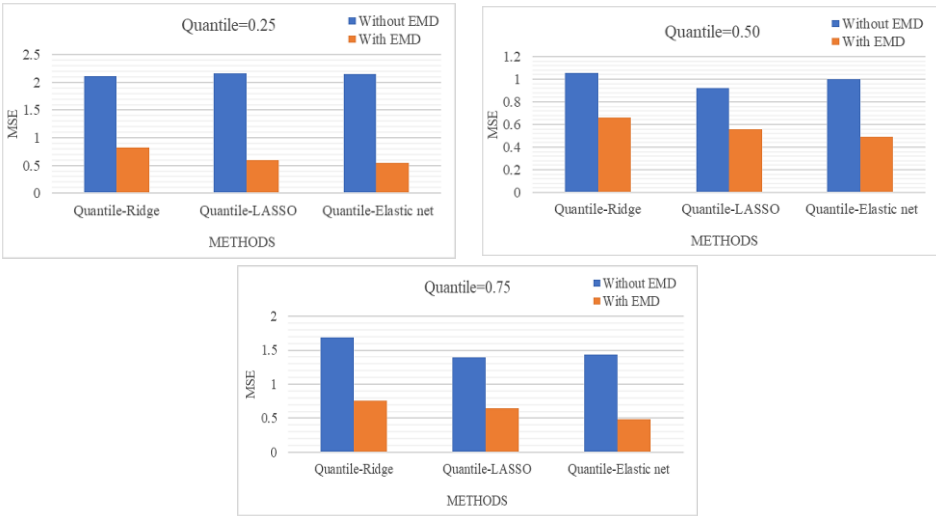


**Figure 8: MASE Results of the Methods for Real Data with/without using EMD**

**Table 8**
**Performance Comparison of Methods using the Daily Close Exchange Rates**

| τ | Method | MSE | Bias | MAE | PE |
|---|---|---|---|---|---|
| 0.25 | QR-R | 2.1089951 | 0.8062739 | 1.0599337 | 0.20156848 |
| | QR-L | 2.1664858 | 0.8025703 | 1.0733641 | 0.20064259 |
| | QR-ENET | 2.1515059 | 0.8031484 | 1.0698911 | 0.20065749 |
| | EMD-QR-R | 0.8186193 | 0.3855621 | 0.6930179 | 0.09639053 |
| | EMD-QR-L | 0.5945223 | 0.1379728 | 0.6223372 | 0.03449320 |
| | EMD-QR-ENET | 0.5456317 | 0.1456226 | 0.5823144 | 0.03640565 |
| 0.50 | QR-R | 1.0567778 | 0.0758499 | 0.7794459 | 0.03792494 |
| | QR-L | 0.9185961 | -0.0154339 | 0.7180875 | 0.01258439 |
| | QR-ENET | 1.0031082 | 0.0562682 | 0.7550317 | 0.01359601 |
| | EMD-QR-R | 0.6618645 | -0.0583486 | 0.5939205 | 0.02917431 |
| | EMD-QR-L | 0.5557124 | 0.005509380 | 0.5943449 | 0.00275469 |
| | EMD-QR-ENET | 0.4897215 | 0.006413746 | 0.5375484 | 0.0032068 |
| 0.75 | QR-R | 1.6858104 | -0.7991616 | 1.1211380 | 0.5993712 |
| | QR-L | 1.3983563 | -0.7248874 | 1.0182253 | 0.5436656 |
| | QR-ENET | 1.4338077 | -0.7367931 | 1.0310333 | 0.5525948 |
| | EMD-QR-R | 0.7532370 | -0.2651068 | 0.6555788 | 0.1988301 |
| | EMD-QR-L | 0.6427820 | -0.1359814 | 0.6336818 | 0.1019861 |
| | EMD-QR-ENET | 0.4817885 | -0.1546586 | 0.5412115 | 0.1159940 |

The estimated coefficients of the decomposition components of the selected model from EMD-QR-R, EMD-QR-L and EMD-QR-ENET are presented in Table 9. We find that all proposed methods can reduce the number of decomposition components at all three levels of quantile (i.e. 0.25, 0.50 and 0.75), except for EMD-QR-R method. This method does not have the ability to do select the decomposition components, and thus, all the decomposition components are entered into the final model. While the numbers of non-zero coefficients of EMD-QR-L and EMD-QR-ENET methods are different at each level of quantiles, such as, the EMD-QR-L method selects nineteen decomposition components when $\tau = 0.5$, whereas the EMD-QR-L method selects eighteen decomposition components at the same level. Thus, the EMD-QR-L and EMD-QR-ENET methods are more accurate in selecting non-zero coefficients than other regression methods.

**Table 9**
**Coefficient Estimation for the Decomposition Components for Proposed Methods**

| Variables | EMD-QR-R | | | EMD-QR-L | | | EMD-QR-ENET | | |
|---|---|---|---|---|---|---|---|---|---|
| | **0.25** | **0.50** | **0.75** | **0.25** | **0.50** | **0.75** | **0.25** | **0.50** | **0.75** |
| $C_{1,1}$ | 0.0179 | 0.0220 | 0.0153 | 0.0276 | 0.0201 | 0.0183 | 0.0256 | 0.0196 | 0.0151 |
| $C_{1,2}$ | 0.0170 | 0.0271 | 0.0253 | 0.0331 | 0.0233 | 0.0301 | 0.0318 | 0.0245 | 0.0282 |
| $C_{1,3}$ | 0.0074 | 0.0121 | 0.0411 | 0.0366 | 0.0373 | 0.0435 | 0.0330 | 0.0412 | 0.0405 |
| $C_{1,4}$ | -0.0234 | 0.0295 | 0.0417 | 0.0116 | 0.0126 | 0.0209 | 0.0128 | 0.0095 | 0.0223 |
| $C_{1,5}$ | 0.0401 | -0.0027 | 0.0290 | 0.1043 | 0.1092 | 0.1123 | 0.0971 | 0.0967 | 0.0709 |
| $C_{1,6}$ | 0.0572 | 0.0205 | 0.0054 | 0.1756 | 0.1581 | 0.1365 | 0.1752 | 0.1505 | 0.1019 |
| $C_{1,7}$ | -0.0164 | 0.0285 | 0.0360 | 0 | 0.0315 | 0.1635 | 0 | 0.0209 | 0.0632 |
| $C_{1,8}$ | 0.2151 | 0.2787 | 0.2555 | 0.7424 | 0.7426 | 0.7780 | 0.6800 | 0.6850 | 0.6369 |
| $R_1$ | -0.0849 | -0.1262 | -0.1347 | 0 | 0 | 0 | 0 | 0 | 0 |
| $C_{2,1}$ | 0.0065 | 0.0107 | 0.0036 | 0.0006 | 0 | 0 | 0.0024 | 0 | 0 |
| $C_{2,2}$ | -0.0033 | -0.0038 | 0.0031 | 0.0002 | 0 | 0.0060 | 0 | 0 | 0.0039 |
| $C_{2,3}$ | 0.0157 | 0.0087 | -0.0109 | 0 | 0 | 0 | 0.0015 | 0 | -0.0034 |
| $C_{2,4}$ | 0.0361 | 0.0244 | 0.0147 | 0.0017 | 0.0087 | 0.0244 | 0.0060 | 0.0111 | 0.0281 |
| $C_{2,5}$ | 0.1024 | 0.0539 | 0.0053 | 0.0728 | 0.0602 | 0.0100 | 0.0766 | 0.0500 | 0.0134 |
| $C_{2,6}$ | 0.0300 | 0.0379 | 0.0625 | -0.0097 | -0.0249 | 0.0148 | -0.0167 | -0.0321 | 0.0077 |
| $C_{2,7}$ | -0.0089 | -0.0643 | -0.0550 | -0.1974 | -0.1823 | -0.1080 | -0.2035 | -0.1977 | -0.1897 |
| $C_{2,8}$ | -0.0248 | 0.0201 | 0.0146 | 0 | 0 | 0 | 0 | 0 | 0 |
| $R_2$ | 0.1652 | 0.2359 | 0.2344 | 0.6469 | 0.6440 | 0.6438 | 0.4348 | 0.4842 | 0.4625 |
| $C_{3,1}$ | 0.0054 | 0.0066 | 0.0094 | 0.0005 | 0.0036 | 0.0061 | 0 | 0.0046 | 0.0060 |
| $C_{3,2}$ | 0.0069 | 0.0079 | 0.0022 | 0.0041 | 0 | 0 | 0 | 0 | 0 |
| $C_{3,3}$ | -0.0070 | 0.0006 | 0.0225 | -0.0040 | 0 | 0.0072 | -0.0022 | 0 | 0.0096 |
| $C_{3,4}$ | -0.0078 | 0.0436 | 0.0290 | 0.0262 | 0.0239 | 0.0209 | 0.0230 | 0.0161 | 0.0332 |
| $C_{3,5}$ | 0.0129 | 0.0288 | 0.0751 | 0.0143 | 0.0363 | 0.0377 | 0.0131 | 0.0456 | 0.0629 |
| $C_{3,6}$ | 0.1330 | 0.1834 | 0.2317 | 0.2297 | 0.2060 | 0.1028 | 0.2242 | 0.2096 | 0.1753 |
| $C_{3,7}$ | 0.0544 | 0.0696 | 0.0437 | -0.0742 | -0.0441 | -0.1168 | -0.0962 | -0.0429 | -0.0084 |
| $C_{3,8}$ | 0.0876 | 0.0989 | 0.0875 | 0 | 0 | 0 | 0 | 0 | 0 |
| $R_3$ | -0.1914 | -0.2601 | -0.2546 | 0 | 0 | 0 | -0.2249 | -0.1697 | -0.1347 |

## 6. CONCLUSION

In this paper, we propose a novel penalized quantile regression based on EMD for nonstationary and nonlinear predictor variables. It performs variable selection and robust parameter estimation simultaneously. In the proposed models, the selection of the decomposition components via the EMD method has been improved, which greatly affects the response variable. This paper aims to enhance the prediction accuracy of the model selection by selecting significant decomposition components in the final model, suggesting robust parameter estimation, and dealing with high multicollinearity among the decomposition components.

The results from numerical simulations and real applications for the daily close exchange rates illustrate that the proposed methods perform better than other quantile regression models at different quantiles. Penalized quantile regression methods with EMD perform better than those without EMD counterparts. The results also showed the superior performance of the proposed methods in estimation and variable selection when multicollinearity is present or absent and building a final model free from multicollinearity and are also resistant to outliers/or heavy-tailed distributions. Thus, overall, penalized quantile regression based on EMD has higher accuracy and is superior to penalized quantile regression methods.

## REFERENCES

1. Ahmed, N.S. and Ismail, E.A.R. (2014). Selection of Variables in Quantile Regression (Linear Lasso- Goal Programming). *Asian Journal of Applied Sciences*, 2(06), 853-866.
2. Ahn, K.W. and Kim, S. (2018). Variable selection with group structure in competing risks quantile regression. Statistics in Medicine, 37(9), 1577-1586.
3. Alhamzawi, R., Yu, K. and Benoit, D.F. (2012). Bayesian adaptive Lasso quantile regression. *Statistical Modelling*, 12(3), 279-297.
4. Al-Jawarneh, A.S., Ismail, M.T. and Awajan, A.M. (2021). Elastic Net Regression and Empirical Mode Decomposition for Enhancing the Accuracy of the Model Selection. *International Journal of Mathematical, Engineering and Management Sciences*, 6(2), 564-583.
5. Al-Jawarneh, A.S., Ismail, M.T., Awajan, A.M. and Alsayed, A.R.M. (2020). Improving accuracy models using elastic net regression approach based on empirical mode decomposition. *Communications in Statistics: Simulation and Computation*, 51(7), 4006-4025.
6. Al-Jawarneh, A.S. and Ismail, Mohd. T. (2021). Elastic-Net Regression based on Empirical Mode Decomposition for Multivariate Predictors. *Pertanika Journal of Science and Technology*, 29(1), 199-215.
7. Alshaybawee, T., Midi, H. and Alhamzawi, R. (2017). Bayesian elastic net single index quantile regression. *Journal of Applied Statistics*, 44(5), 853-871.
8. Amin, M., Song, L., Thorlie, M.A. and Wang, X. (2015). SCAD-penalized quantile regression for high-dimensional data analysis and variable selection. *Statistica Neerlandica*, 69(3), 212-235.
9. Belloni, A. and Chernozhukov, V. (2011). Ell;1-penalized ruantile regression in high-himensional sparse models. *Annals of Statistics*, 39(1), 82-130.

10. Bonaccolto, G. (2019). Critical Decisions for Asset Allocation via Penalized Quantile Regression. *SSRN Electronic Journal*, 2005, 1-20.
11. Burgette, L.F., Reiter, J.P. and Miranda, M.L. (2011). Exploratory quantile regression with many covariates: An application to adverse birth outcomes. *Epidemiology*, 22(6), 859-866.
12. Chatterjee, S. and Hadi, A.S. (2013). Regression analysis by example. In shed by John Wiley & Sons, Inc., Hoboken, New (Vol. 40, Issue 12). https://doi.org/10.1080/02664763.2013.817041
13. Ciuperca, G. (2020). Adaptive elastic-net selection in a quantile model with diverging number of variable groups. *Statistics*, 54(5), 1147-1170.
14. Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, 96(456), 1348-1360.
15. Hoerl, A.E. and Kennard, R.W. (1970). Ridge Regression: Applications to Nonorthogonal Problems. Technometrics, 12(1), 69-82.
16. Hu, J., Chen, Y., Zhang, W. and Guo, X. (2021). Penalized high-dimensional M-quantile regression: From L1 to Lp optimization. *Canadian Journal of Statistics*, 49(3), 875-905.
17. Huang, N.E., Shen, Z. and Long, S.R. (1999). A new view of nonlinear water waves: The Hilbert Spectrum 1. In *Annu. Rev. Fluid Mech* (Vol. 31).
18. Huang, N.E., Wu, M.L., Qu, W., Long, S.R. and Shen, S.S. (2003). Applications of Hilbert-Huang transform to non-stationary financial time series analysis. *Applied stochastic models in business and industry*, 19(3), 245-268.
19. Huang, Shen, Z., Long, S.R., Wu, M.C., Snin, H.H., Zheng, Q., Yen, N.C., Tung, C.C. and Liu, H.H. (1998). The empirical mode decomposition and the Hubert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 454(1971), 903-995.
20. Kang, J., Shin, S.J., Park, J. and Bang, S. (2018). Hierarchically penalized quantile regression with multiple responses. *Journal of the Korean Statistical Society*, 47(4), 471-481.
21. Khan, D.M., Yaqoob, A., Iqbal, N., Wahid, A., Khalil, U., Khan, M., Rahman, M.A. A., Mustafa, M.S. and Khan, Z. (2019). Variable Selection via SCAD-Penalized Quantile Regression for High-Dimensional Count Data. *IEEE Access*, 7, 153205-153216.
22. Koenker, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, 91(1), 74-89.
23. Koenker, R. (2005). Quantile regression. In *Nature Methods* (Vol. 16, Issue 6). https://doi.org/10.1038/s41592-019-0406-y
24. Koenker, R. and Bassett, G. (1978). Regression Quantiles. *Econometrica*, 46(1), 33. https://doi.org/10.2307/1913643
25. Li, Y. and Zhu, J. (2008). L1-norm quantile regression. *Journal of Computational and Graphical Statistics*, 17(1), 163-185.
26. Liu, Y., Zeng, P. and Lin, L. (2020). Generalized ℓ1-penalized quantile regression with linear constraints. *Computational Statistics and Data Analysis*, 142, 106819.

27. Januaviani, T.M.A., Gusriani, N., Joebaedi, K., Supian, S. and Subiyanto, S. (2019). The best model of LASSO with the LARS (least angle regression and shrinkage) algorithm using Mallow's Cp. *World Scientific News*, (116), 245-252.
28. Masselot, P., Chebana, F., Bélanger, D., St-Hilaire, A., Abdous, B., Gosselin, P. and Ouarda, T.B. (2018). EMD-regression for modelling multi-scale relationships, and application to weather-related cardiovascular mortality. *Science of The Total Environment*, 612, 1018-1029.
29. Mkhadri, A., Ouhourane, M. and Oualkacha, K. (2017). A coordinate descent algorithm for computing penalized smooth quantile regression. *Statistics and Computing*, 27(4), 865-883.
30. Owusu, P., Kumar, A. and Padhan, H. (2020). Analysis of EEMD-based quantile-in-quantile approach on spot- futures prices of energy and precious metals in India. *Resources Policy*, 68(May), 101731.
31. Peng, B. and Wang, L. (2015). An iterative coordinate descent algorithm for high-dimensional nonconvex penalized quantile regression. *Journal of Computational and Graphical Statistics*, 24(3), 676-694.
32. Qin, L., Ma, S., Lin, J.C. and Shia, B.C. (2016). Lasso Regression Based on Empirical Mode Decomposition. *Communications in Statistics: Simulation and Computation*, 45(4), 1281-1294.
33. Ranganai, E. and Mudhombo, I. (2021). Variable selection and regularization in quantile regression via minimum covariance determinant based weights. *Entropy*, 23(1), 1-21.
34. Shen, Y., Liang, H.Y. and Fan, G.L. (2018). Penalized empirical likelihood for quantile regression with missing covariates and auxiliary information. *Communications in Statistics-Theory and Methods*, 47(8), 2001-2021.
35. Shen, Z., Feng, N. and Shen, Y. (2012). Ridge regression model-based ensemble empirical mode decomposition for ultrasound clutter rejection. *Advances in Adaptive Data Analysis*, 4(01n02), 1250013.
36. Shi, S. and Wilke, R.A. (2020). Variable selection with group structure : exiting employment at retirement age — a competing risks quantile regression analysis. In *Empirical Economics* (Issue 0123456789). Springer Berlin Heidelberg.
37. Su, M. and Wang, W. (2021). Elastic net penalized quantile regression model. *Journal of Computational and Applied Mathematics*, 392, 113462.
38. Tamura, R., Kobayashi, K., Takano, Y., Miyashiro, R., Nakata, K. and Matsui, T. (2017). Best subset selection for eliminating multicollinearity. *Journal of the Operations Research Society of Japan*, 60(3), 321-336.
39. Tang, Q., Zhang, H. and Gong, S. (2019). Bayesian regularized quantile regression analysis based on asymmetric laplace distribution. *Journal of Applied Mathematics and Physics*, 8(1), 70-84.
40. Tian, Y.Z., Tang, M.L., Chan, W.S. and Tian, M.Z. (2021). Bayesian bridge-randomized penalized quantile regression for ordinal longitudinal data, with application to firm's bond ratings. *Computational Statistics*, 36(2), 1289-1319.
41. Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288.

42. Tweneboah, G., Owusu Junior, P. and Kumah, S.P. (2020). Modelling the asymmetric linkages between spot gold prices and African stocks. *Research in International Business and Finance*, 54, 101246.
43. Varadarajan, N. and Nagarajaiah, S. (2004). Wind response control of building with variable stiffness tuned mass damper using empirical mode decomposition/Hilbert transform. *Journal of Engineering Mechanics*, 130(4), 451-458.
44. Wang, L., Van Keilegom, I. and Maidman, A. (2018). Wild residual bootstrap inference for penalized quantile regression with heteroscedastic errors. *Biometrika*, 105(4), 859-872.
45. Xu, Q., Cai, C., Jiang, C., Sun, F. and Huang, X. (2017). Sampling Lasso quantile regression for large-scale data. *Communications in Statistics: Simulation and Computation*, 47(1), 92-114.
46. Yan, A. and Song, F. (2019). Adaptive elastic net-penalized quantile regression for variable selection. *Communications in Statistics - Theory and Methods*, 48(20), 5106-5120.
47. Yang, A.C., Fuh, J.L., Huang, N.E., Shia, B.C., Peng, C.K. and Wang, S.J. (2011). Temporal associations between weather and headache: Analysis by empirical mode decomposition. *PLoS ONE*, 6(1), 1-6.
48. Yang, J.N., Lei, Y., Pan, S. and Huang, N. (2003). System identification of linear structures based on Hilbert-Huang spectral analysis. Part 2: Complex modes. *Earthquake Engineering & Structural Dynamics*, 32(10), 1533-1554.
49. Yousif, A.H. and Housain, W.J. (2021). Atan regularized in quantile regression for high dimensional data. *Journal of Physics: Conference Series* (Vol. 1818, No. 1, p. 012098).
50. Zhang, C.H. (2010). Nearly unbiased variable selection under minimax concave penalty. *Annals of Statistics*, 38(2), 894-942.
51. Zhang, L., Xie, L., Han, Q., Wang, Z. and Huang, C. (2020). Probability density forecasting of wind speed based on quantile regression and kernel density estimation. *Energies*, 13(22), 1-24.
52. Zhong, W., Zhu, L., Li, R. and Cui, H. (2016). Regularized quantile regression and robust feature screening for single index models. *Statistica Sinica*, 26(1), 69-95.
53. Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American Statistical Association*, 101(476), 1418-1429.
54. Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 67(2), 301-320.
55. Zou, H. and Yuan, M. (2008). Regularized simultaneous model selection in multiple quantiles regression. *Computational Statistics & Data Analysis*, 52(12), 5296-5304.
56. Zou, H. and Zhang, H.H. (2009). On the adaptive elastic-net with a diverging number of parameters. *Annals of Statistics*, 37(4), 1733-1751.