BAREKENG: Journal of Mathematics and Its Applications
September 2026 Volume 20 Issue 3 Page 1793-1806
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol20iss3pp1793-1806

COMPARATIVE STUDY OF FUNCTIONAL LINEAR REGRESSION AND AN INTERACTION MODEL FOR MODELING RAINFALL IN INDONESIA

Khusnia Nurul Khikmah[1], **A'yunin Sofro**[2*]

[1]*Mathematics Department, Faculty of Mathematics and Natural Sciences, Universitas Negeri Surabaya*
[2]*Actuarial Sciences Department, Faculty of Mathematics and Natural Sciences, Universitas Negeri Surabaya*
Jln. Komplek Universitas Negeri Surabaya, C8 Building, Ketintang, Gayungan, Surabaya, Jawa Timur, 60231, Indonesia

*Corresponding author's e-mail: * **ayuninsofro@unesa.ac.id***

Article Info

Article History:
Received: 4th May 2025
Revised: 29th September 2025
Accepted: 16th March 2026
Available online: 8th April 2026

Keywords:
Functional data;
Linear functional effects;
Rainfall;
Regression;

ABSTRACT

Regression is a widely used statistical modeling method. Its applications have evolved, including functional data analysis, which offers flexibility by modeling data over specific time intervals with varying case-specific observations. Functional regression has several categories, one of which is functional prediction regression. This study uses functional prediction regression to model comprehensive climate data from the World Bank Climate Change Knowledge Portal, specifically Indonesia's average annual temperature and rainfall from 1951 to 2016. An advantage of functional prediction regression is its ability to model based on the dataset. We compare two models: the first with a functional linear effect and the second with a linear interaction combination effect. The models are estimated using boosting and the Generalized Additive Model for Location, Scale, and Shape (GAMLSS). Results show that the functional linear model better fits Indonesia's rainfall data, yielding a smaller Akaike's Information Criterion by 170.351.

How to cite this article:

K. N. Khikmah and A. Sofro., "COMPARATIVE STUDY OF FUNCTIONAL LINEAR REGRESSION AND AN INTERACTION MODEL FOR MODELING RAINFALL IN INDONESIA", *BAREKENG: J. Math. & App.,* vol. 20, no. 3, pp. 1793-1806, Sep, 2026.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

Climate change is a global issue driven by shifts in weather patterns. One of the most critical dimensions of weather variation in Indonesia is rainfall. The climate is defined as the result of long-term atmospheric and weather processes [1]. So, climate change is a change in the atmosphere and weather processes over a long period. Based on several studies, climate change has an impact on many aspects, some of which are the spread of the COVID-19 disease [2], [3], which affects the ability of the agricultural system [4], and impacts on forests in the northeastern United States and Canada [5]. In Indonesia, climate change significantly alters temperature and rainfall patterns over time, with cascading impacts on agriculture [6] established the fundamental patterns of climate change impacts across Indonesia's diverse ecosystems, providing critical baseline data for subsequent studies. These climatic shifts particularly affect agricultural systems, as demonstrated by [7], [8]. Water resources, and natural disasters like floods [9] showing through global modeling that increased precipitation intensity in tropical climates like Indonesia's directly elevates flood risks by 12-24% per degree of warming. While rainfall dynamics have been extensively studied in recent decades, existing models often fail to fully capture the complex temporal dependencies in climate data. A gap of this study addressed through functional data analysis. Mainly, the functional data approach remains underutilized in rainfall modeling despite its flexibility to handle time-varying patterns [10]. Our research contributes by systematically comparing functional prediction regression methods (boosting vs. GAMLSS) and demonstrating the performance of the functional linear model and functional linear interaction effects modeling accuracy for Indonesia's rainfall data, and offering a more robust analytical framework for climate researchers.

Functional data analysis is a flexible method for exploring data in certain time intervals with data that can vary between cases and may have different distances [11], [12]. Climate and weather case data are one of the best applications of functional data analysis [13], [14], [15]. Some functional data analysis developments involve spatial [16], joint modeling using principal components [17], and using regression [18]. Research in functional data using the functional prediction regression model, in theory, can be seen at [19] and in medicine application, see [20], [21]. Research on functional data with functional prediction regression has been carried out in many previous studies. Some studies are based on data electroencephalography (EEG) and electromyography (EMG). Functional regression includes several modeling approaches, such as functional prediction regression, an advanced method in functional data analysis. Functional prediction regression can test models based on the overall data; see [20], [22]. Another advantage of functional prediction regression is a model that can model and thus predict based on data settings.

Functional prediction regression can be modeled through boosting and generalized additive models for location scale and shape. Boosting is an algorithm developed for statistical models so that the model can be analyzed [23], [24] and can be implemented through FDboost by using the gradient boosting algorithm in estimating the model. Gradient boosting is a new algorithm in estimating functional data that chooses advantages for high-dimensional data and an algorithm that selects base learners to update the most suitable additional predictors so that they are more effective in modeling [25]. Another method is generalized additive models. The approach for location scale and shape is a model that can improve functional regression, which has advantages in analyzing high-dimensional data and can estimate several parameters simultaneously, making it more complex and efficient. Functional prediction regression through generalized additive models for location scale and shape (GAMLSS). However, this study's application was not without limitations. Estimating several parameters simultaneously demands substantial computational power and careful model adjustment, posing a significant barrier in research involving large datasets that are the imitations encountered in this study.

The advantage of the GAMLSS method is that it enhances functional regression by efficiently analyzing high-dimensional data. Given these advantages, this study applies functional prediction regression via GAMLSS and boosting to model rainfall patterns in Indonesia—a region where rainfall dynamics have been extensively studied [26], [27], [28]. Specifically, the main objective is to directly compare the performance of these two approaches in modeling rainfall, focusing on Indonesia's average annual temperature and rainfall as key variables.

2. RESEARCH AND METHODS

2.1 Functional Data Analysis

Functional data is multivariate data in which dimensions apply sequences, where ε_{ij} is measurement of noise assumed i.i.d. with mean 0 and variance σ^2, X_{ij} is observed discrete value of the i-th function at time point t_j, t is continuous-time and the function $x_i(t)$ is functional data, then [29].

$$X_{ij} = x_i(t) + \varepsilon_{ij}. \tag{1}$$

Functional data analysis is based on the idea that data due to time series from discrete observations create functional data representing all observation functions to provide information [30]. Briefly, multivariate time series data can also be analyzed using the functional data analysis approach.

Meanwhile, functional data analysis is an analysis and theory of data that varies in a time or location continuum related to functions, images, and forms in a functional set [31]. Functional data analysis has a simple dataset from where $n = 1,2,\ldots,N$, N curves are observed at the same interval, for example, $[x, y]$ and for all t points that are not known the value of the curve is $t \in [x, y]$, where j= $1,2,\ldots,J_n$ and $t_{j,n} \in [x, y]$ then.

$$X_n(t_j, n) \in \mathcal{R}. \tag{2}$$

The base expansion used is the B-spline base.

$$X_t = \sum_{k=1}^{m+l-1} c_k Bk(t, \tau), \tag{3}$$

where c is the vector coefficient. In functional data analysis, discrete observations are converted into continuous curves using smoothing techniques. One such method is the use of B-spline basis expansion, that have advantages in terms of the structure of the polynomial, allowing for local adaptation to sudden changes in the data trend, smoothness constraints controlled by knot placement, and spline order prevents overfitting, and computational efficiency enables scalable analysis of high-dimensional functional data. One common approach uses B-spline basis expansion, which builds a flexible and smooth representation of the underlying functional trend in the data, where the object observed at a certain point $(t_{j,n})$ and at each point t has a value of $x_n(t)$ with $n = 1,2,3,\ldots,N$ then:

$$X_n(t): t \in [x, y]. \tag{4}$$

Base expansion in general in functional data analysis where c is the vector coefficient and $\phi(t)$ as the base system [29].

$$X(t) = \phi(t)c = \sum_{k}^{K} c_j \phi_k(t). \tag{5}$$

The application of several methods, namely in health, environment [30], and multivariate analysis, is the basis of functional data analysis [11]. Functional data has advantages. Within the limitations of the observations, it can produce a refined functional representation and goals that function well with effectiveness in solving functional problems [32]. One of the functional data analysis methods is used in weather and climate problems [33]. With technological developments, one of the functional data tools is boosting [18]. Because functional data analysis is related to functions, images, and forms in a functional set, several applications of functional data are used in modeling [34]. Functional data modeling is a topic that is part of functional data analysis that is developing, one of which is through regression. The functional regression Model is categorized into three models, namely Functional prediction regression, commonly called scalar-on-function regression (scalar response and functional covariates); functional response regression, commonly called function-on-scalar regression (scalar covariates and functional responses); and Function-on-function regression (functional covariates and functional responses) [20]. This study focuses specifically on the functional prediction regression model.

2.2 Functional Prediction Regression

One form of modeling in functional data is regression, which is used in functional prediction regression, where functional linear regression is a multivariate regression for functional data. Functional prediction regression can also be called scalar-on-function regression. Functional linear regression Model with $\{(X_1,Y_1),(X_2,Y_2),\ldots,(X_n,Y_n)\}$ data for X_i is a predictor function and $i=1,2,\ldots,N, X=\{X_i\to R,\ i=1,2,\ldots,n\}$, $a\in\mathcal{R}$ is the intercept, $s\in S=[S_1,S_2]$ with $S_1< S_2$ and $S_1,S_2\in\mathcal{R}$, ϵ is error, and b is the functional regression coefficient then [34], [35].

$$Y_i = a + \int_S b(t)X_i(t)dt + \varepsilon_i. \tag{6}$$

Functional regression $X_i(t)=\sum_{k=1}^{K_X} X_{ik}^*\phi_k(t)$ where K_X is design matrix for the x-th effect and $(X_i^*)=[X_{i1}^*,X_{i2}^*,\ldots,X_{iK_X}^*]'$ and $\phi(t)=[\phi_1(t),\phi_2(t),\ldots,\phi_{K_X}(t)]'$ and $b(t)=\sum_{k=1}^{K_b} b_k^*\psi_k(t)$ where K_b is design matrix for the b-th effect and $b^*=[b_1^*,b_2^*,\ldots,b_{iK_b}^*]'$ and $\psi(t)=[\psi_1(t),\psi_2(t),\ldots,\psi_{K_b}(t)]'$, for more explanation see [18].

$$g\{E(Y_i)\} = a + \int_S b(t)X_i(t)dt, \tag{7}$$

$$\int_S b(t)X_i(t)dt = \mathbf{X}_\mathbf{i}^*\mathbf{J}_{\boldsymbol{\phi},\boldsymbol{\psi}}\mathbf{b}^*. \tag{8}$$

In Eq. (6) $\varepsilon_i\in\mathcal{R}$ is a random error that is independent of X_i and the variance is finite [20], [34], [36].

$$\int_S^T E(X^2) < \infty. \tag{9}$$

Suppose the additional predictor $h(x)$ for scalar response is $g\{E(Y_i)\}$, then [18], [20].

$$g\{E(Y_i)\} = a + \int_S b(t)X_i(t)dt = h(x) = \sum_{j=1}^{J} h_j(x), \tag{10}$$

where x is the functional covariate vector, $g\{E(Y_i)\}$ is the functional representation of the response result, and $h_j(x)$ is the base function representation of the partial effects. The additional predictor $h(x)$ for functional prediction regression using FDboost, because the desired response is scalar, is not crucial for time. It has several potential base-learners for functional covariates and scalar responses for additional predictors.

Table 1. Potential Additional Predictors of $h(x)$ for Functional Prediction Regression

Model	$\boldsymbol{h(x)}$
Model 1	$a + \int_S b(t)X_i(t)dt$
Model 2	$a + zf(z_i) + \int_S b(t)X_i(t)dt + z\int_S b(t)X_i(t)dt$

Modeling with the functional prediction regression can be modeled through boosting. Boosting is an algorithm developed for statistical models to analyze the model [37]. The way to select base-learners to update additional predictors that function as an iteration enhancer of the model to become more complex where the number of boosting iterations controls the complexity of the model [38].

Functional prediction regression can be implemented via boosting through FDboost by using an algorithm to estimate the model. One of the estimations in modeling with the functional prediction regression can use the gradient boosting algorithm. The gradient-boosting algorithm selects advantages for high-dimensional data and selects base learners to update the most suitable additional predictor model. In general, the estimation using the gradient boosting algorithm for the functional regression model has an algorithm that clearly explain in [18], [39]. Functional prediction regression through boosting can be explored from plot-modelled data. Such as knowing the smooth effects, the results of the model's prediction, and the residuals. The first step of the boosting method is to set an effect function, fitting the model by using gradient boosting, to decide the best model by considering the m_{stop}. The second way to analyze functional data is by using generalized additive models for location scale and shape. Andreas Fuest, Andreas Mayr, and Sonja Greven developed additive models for location scale and shape.

The approach has advantages in analyzing high-dimensional data and can estimate several parameters simultaneously. The first step of using generalized additive models for location scale and shape is to set an effect function, then fit the model, and find the best model by considering Akaike's information criteria value. Finally, Akaike's information criteria value will be used in this paper to measure the model's performance. The best model selection is based on Akaike's information criteria, which is the smallest value of the model.

2.3 Aikake's Information Criterion (AIC)

Akaike Information Criterion (AIC) is an information criterion that shows a relative measure of the suitability of a statistical model. The AIC value indicates a measure of the model's fit with the data modeling [40]. AIC were first introduced in 1973 by Akaike. The value of AIC minimizes information loss in model selection and provides information used to determine the best model. The best model obtained is obtained based on the smallest AIC value. In general, AIC where k is the parameter of the model, and L is the maximum value of the likelihood function used to estimate the model is formulated as [41], [42].

$$AIC = 2k - 2ln(L). \tag{11}$$

3. RESULTS AND DISCUSSION

The data is about average annual temperature and rainfall from 1951 until 2016 in Indonesia or for 66 years from the World Bank Climate Change Knowledge Portal website https://climateknowladgeportal.worldbank.org/country/indonesia. The data display in functional data is given in Fig. 1. To find an explanation of functional data, one of which is a plot with a broader description, see [43]. The data plot, functional data in Indonesia from 1951 to 2016, was formed starting from the data contained in the 66×2 matrix with intervals, $[0,1]$ in Fig. 1.

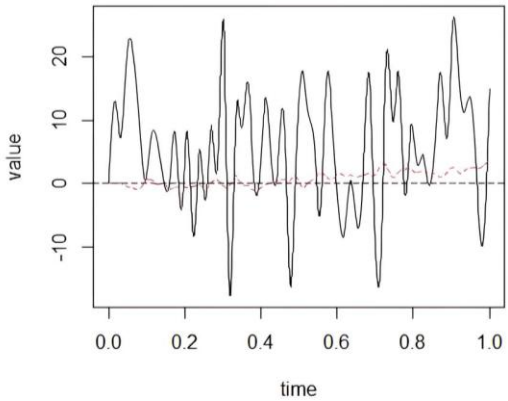


Figure 1. Functional Data Plots of Average Annual Temperature and Annual Average Rainfall In Indonesia From 1951 To 2016

The graph in Fig. 1 shows a time series graph of Indonesia's average annual temperature in black and annual rainfall in red. The data are collected from 1951 to 2016. Temperature values range from -10 to 20, and rainfall values range from 0.0 to 1.0, and have been plotted against time, showing their dynamic patterns over 65 years. The temperature graph shows significant fluctuations. This fluctuation is shown by the period's sharp increase, with peaks approaching 20 °C and drops below 0 °C, indicating significant interannual variability, which may be related to climate phenomena. It shows a clear long-term warming trend and considerable interannual variability. This overall increase is punctuated by sharper warming and occasional cooler periods. Notably, the clear peaks and troughs align with known large-scale climate phenomena; for example, the black graph, whose peaks typically bring warmer and drier conditions to the region.

Meanwhile, the rainfall graph shows smoother oscillations but with wet phases by the values approaching 1.0 and dry phases by the values approaching 0.0, reflecting seasonal climate cycles. This functional graph provides a holistic view of Indonesia's climate trends, emphasizing the interconnection between temperature and rainfall dynamics over time by inverting the relationship observed in some years where temperature decreases coincide with rainfall. It shows an apparent deviation from the average pattern. Fluctuations show an inverse relationship with temperature during extreme years, particularly during strong events, where minimum rainfall coincides with maximum temperature. This visual inverse correlation is a key characteristic of the region's climate and underscores the interconnected nature of these two variables. The relevance of this functional plot to the subsequent analysis is twofold. First, it justifies using functional data methods, as complex nonlinear patterns are more appropriately modeled as smooth curves than discrete data points. Second, the observed covariation between the temperature and precipitation curves provides a strong rationale for the subsequent analysis stage to quantitatively assess the nature and strength of their dynamic relationship over time. By examining these functional forms, we can move beyond simple correlation to understand how the curve shapes of these climate variables relate to each other.

The data temperature and rainfall have 66 observations, $f(z_i)$ is the smooth effect. $\alpha_k(t_k, z_i)$ is interaction effects on functional covariates, namely the use of rainfall data (independent variable), and x_{ki}, $k = 1$, namely in the form of functional variables. For example, in Table 1 $g\{E(Y_i)\} = \sum_{j=1}^{J} h_j(x)$. The model 1, namely:

$$g\{E(Y_1)\} = \sum_{j=1}^{J} h_j(x) = \int_S b(t)X_i(t)dt,$$

and model 2, namely:

$$g\{E(Y_1)\} = \sum_{j=1}^{J} h_j(x) = zf(z_i) + \int_S b(t)X_i(t)dt + z\int_S b(t)X_i(t)dt.$$

The model $g\{E(Y_i)\}$ was standardized with a range $[-1,1]$. The above model is a model with response scalar and functional covariate, which has base-learners. In this study, the scalar response was chosen for the time it was neglected, and the signal function was chosen for the covariate function. Base learners for functional covariates have several predictors for scalar-on-function regression or functional prediction regression, which can be seen where additional predictor functions for this study were selected with the help of FDboost. The results of generating functional covariates are presented in Fig. 2.

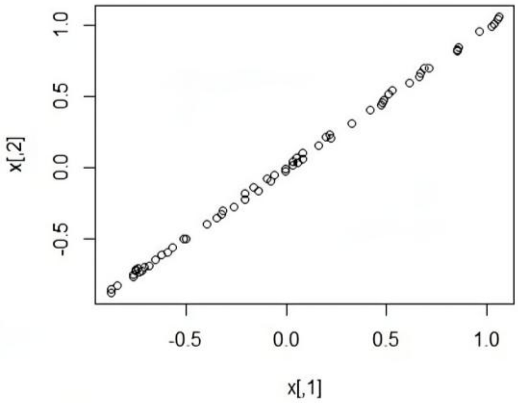


Figure 2. The Plot of the Functional Results of the Covariates in the Range $[-1,1]$

The additional predictor function model 1 used has a linear functional effect around the intercept, and the model is calculated using the boosting aid. The results of the linear functional effects around the intercept of the additional predictor functions of model 1 are given in Fig. 3, which is the base-learners of the rainfall data that are made as functional covariates that show the estimation results of the linear functional effects of the additional predictor functions to the equation $\int_S b(t)X_i(t)dt$.

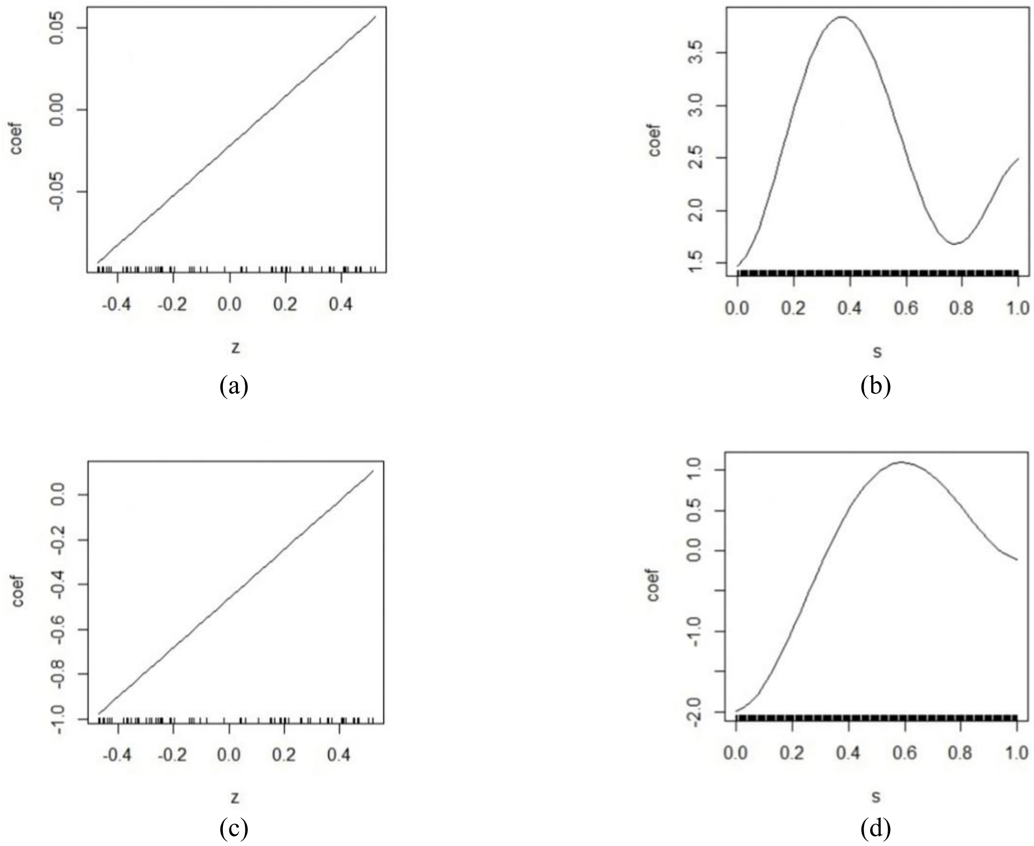
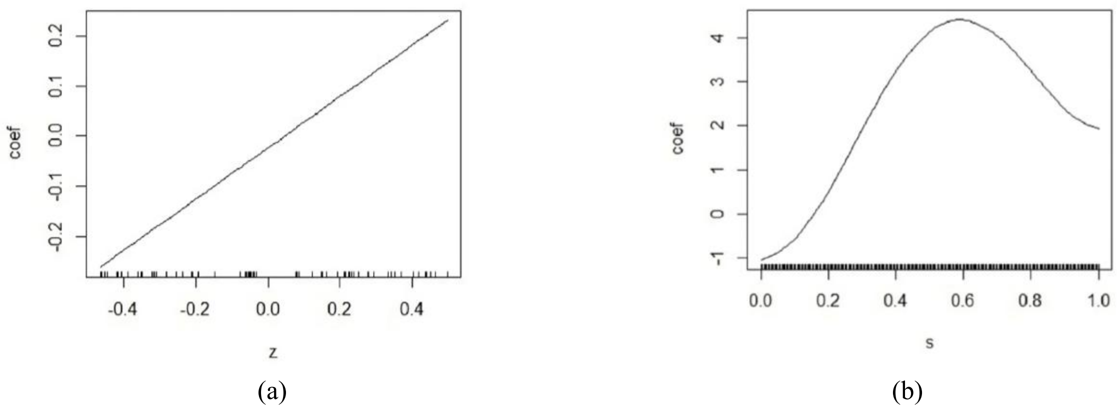
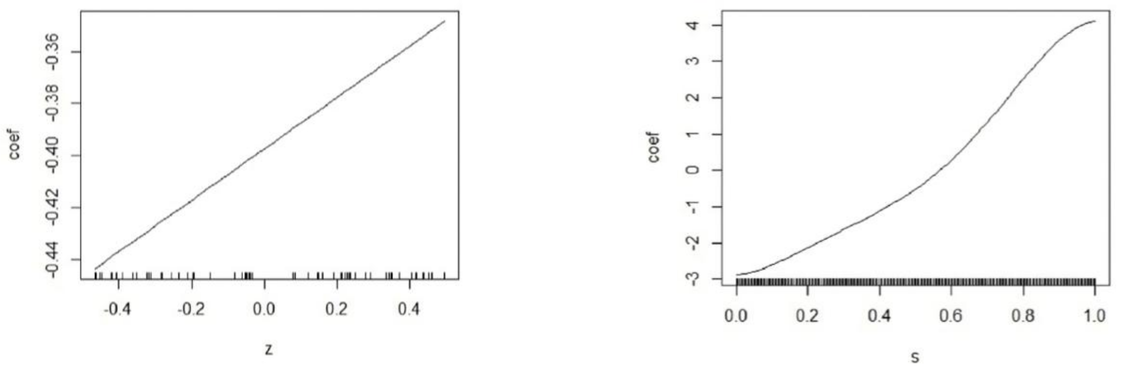


Figure 3. (a) Covariate-Specific of Function Effect of Temperature Term z_1, (b) Smooth Effect of Model Estimation Coefficient in Model 1 with the Optimal Number of Boosting in Long-Term Rainfall Pattern $\beta_1(s)$, (c) Covariate-Specific of Function Effect of Temperature Extreme Precipitation Threshold z_2, (d) Smooth Effect of Model Estimation Coefficient in Model 1 with the Optimal Number of Boosting in Seasonal Rainfall Pattern $\beta_2(s)$

Like the additional predictor function model 1, model 2 affected the intercept. Calculating the model using boosting assistance via FDboost, namely the combined effect of linear interactions. The results of the combined effect of linear interactions around the intercept of the additional predictor functions of model 2 are given in Fig. 4, namely the base-learners of rainfall data which are made as functional covariates which show the estimation results of the linear functional effects of the additional predictors on the equation $\int_s b(t)X_i(t)dt + z\int_s b(t)X_i(t)dt$.

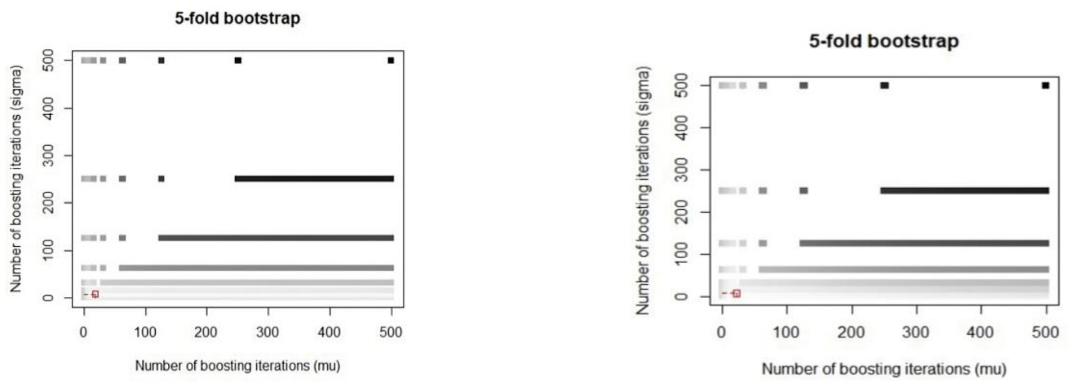




(c) (d)

Figure 4. (a) Covariate-Specific of Function Effect of Temperature Interaction Term z_1, (b) Smooth Effect of Model Estimation Coefficient in Model 2 with the Optimal Number of Boosting in Long-Term Rainfall Pattern $\beta_1(s)$, (c) Covariate-Specific of Function Effect of Temperature Extreme Precipitation Threshold z_2, (d) Smooth Effect of Model Estimation Coefficient in Model 2 with the Optimal Number of Boosting in Seasonal Rainfall Pattern $\beta_2(s)$

In modeling data through regression with the scalar response, the model complexity is influenced and controlled by boosting iterations and requires an estimation coefficient. One of the coefficients estimate functions is to obtain a measure of the uncertainty brought through bootstrapping. The bootstrap result has a zero-bias interval, but the variability of the estimation coefficient is still there. This interval measures all the uncertainty generated by the model selection as the uncertainty of the actual coefficient. In Fig. 5, the estimated coefficients generated by model 1 and model 2 are plotted together with the estimates calculated on the 5-fold bootstrap, which can be seen in the gray graph. In contrast, the black graph represents the mean function over the bootstrap sample.



(a) (b)

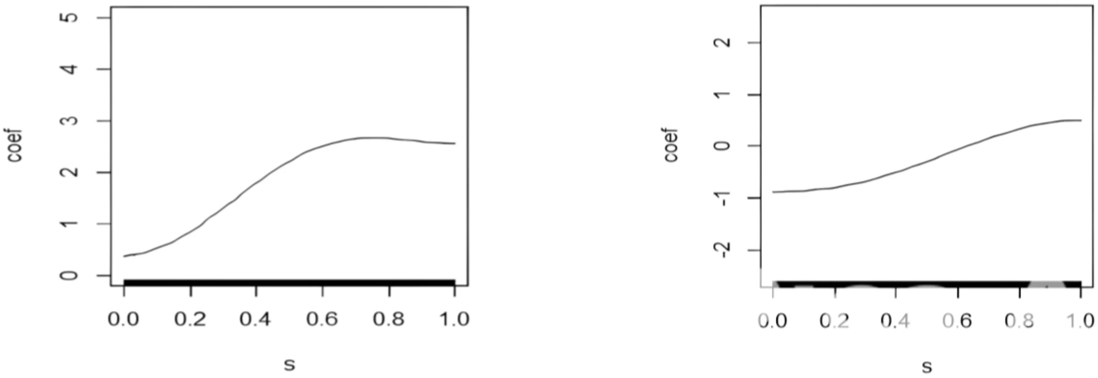
Figure 5. Plot the Bootstrap for Rainfall Values with Quantiles of 95% (a) Model 1, (b) Model 2

From Fig. 5, the dashed red line shows that the boosting number of the optimal iterations and the optimal number of iterations were estimated by bootstrap with 5-folds. The number of boosting iterations with parameters where the regression spline adjustment (knots) used is 16. In the functional prediction regression boosting in Eq. (6), $b(t)$ is smooth effects that have residuals (ε_i). Such a model relies on functional covariates and includes smooth residues. Thus, to adjust the effect on the model with additional predictors, namely model 1 and model 2, where $b(t)$ is the basis for all effects to achieve the equality of estimates.

Table 2. The m_{stop} Value Table Results from the Model with Functional Prediction Regression using Boosting

Model	m_{stop}	
	μ	σ
Model 1	18	8
Model 2	22	8

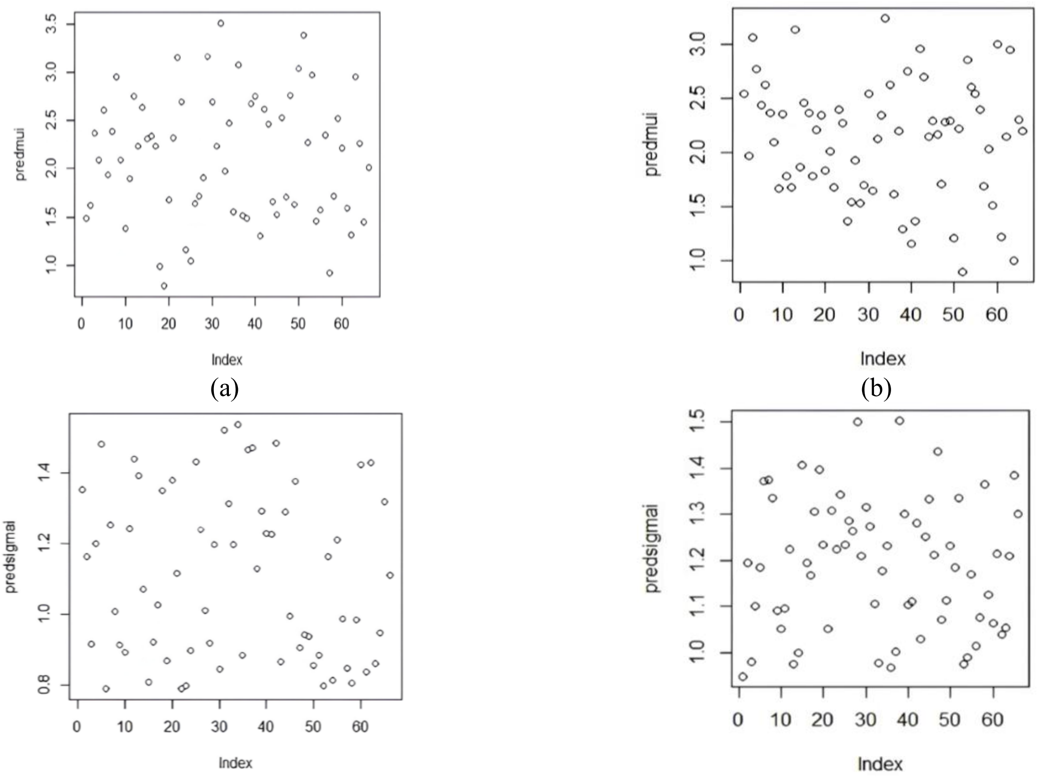
Bootstrap retention iteration for 5-fold bootstrap for base-learners where model 1 m_{stop} boosting iteration with bootstrap for mu parameter is $m = 18$, and sigma is $m = 8$ whereas for model 2 m_{stop} the boosting iteration with bootstrap for mu parameter is $m = 22$ and sigma is $m = 8$. So, the worse the model, the higher the m_{stop} bootstrap it has because the optimization in the boosting iteration is getting longer. The result of v or the number of steps in the boosting iteration is 0.1. The model made has smooth effects or linear functional effects $f(z_i)$, namely the effect of the interaction between functional covariates and other scalar variables. The effect of the interaction results on the model symbolized by $f(z_i)$ is given in Fig. 6, which is the estimation result of the bootstrap coefficient for rainfall made as a functional covariate from model 1 and model 2, namely the linear functional effect and the combined effect of linear interactions.



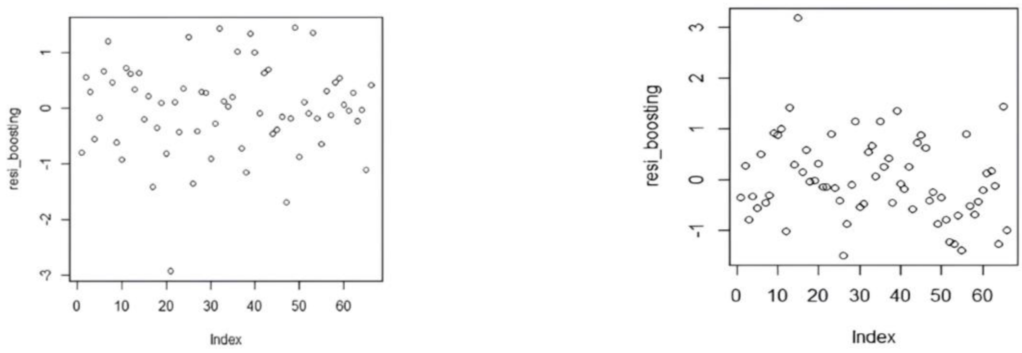
(a) (b)

Figure 6. The Plot of the Combined Effect of Linear Interactions $f(z_i)$ (a) Model 1, (b) Model 2

The prediction results of the model using boosting with two models where each model has parameters, namely mu and sigma used, the results obtained with the QQ-plot and have residuals given in Fig. 7.



(a) (b) (c) (d)



(e) (f)

Figure 7. QQ-Plot (a) μ Plot of Model 1, (b) μ Plot of Model 2, (c) σ Plot of Model 1, (d) σ Plot of Model 2, (e) Residual Plot of Model 1, and (f) Residual Plot of Model 2

Besides being able to be modeled through boosting, functional prediction regression or scalar-on-function regression can also be modeled through generalized additive models for location scale and shape with the help of R. Like modeling via boosting, modeling through generalized additive models for location scale and shape also provides results regarding functional covariates and estimates with different representations with boosting. In functional prediction regression through generalized additive models for location scale and shape, such as via boosting using the additional predictor $h(x)$ in Table 1 as model 1, which allows linear functional effects and model 2 effects from a combination of linear interactions. Modeling functional prediction regression through generalized additive models for location scale and shape is obtained.

Table 3. Table of model results with functional prediction regression using GAMLSS

Model	Parameter		Estimate	Error	t-value	p-value
Model 1	μ	a	2.0997	0.1015	20.892	2.000×10^{-16}
		z	0.2447	0.321	0.787	0.434
	σ	a	-0.27429	0.087	-3.151	0.003
		z	1.14361	0.339	3.365	0.001
Model 2	μ	a	2.13939	0.090	23.707	2.000×10^{-16}
		z	0.62578	0.326	1.921	0.059
	σ	a	-0.18345	0.087	1.921	0.039
		z	0.19184	0.289	0.664	0.509

From the Table 3 of model results, it is obtained for model 1, namely:

$$2.0997 + \int b_k(t_k)x_{ki}(t_k)dt_k + 0.62578\int \alpha_k(t_k, x_i)x_{ki}(t_k)dt_k, \qquad (12)$$

and for model 2 based on 16 can be written as

$$2.13939 + 0.62578f(z_i) + \int b_k(t_k)x_{ki}(t_k)dt_k + 0.19184\int \alpha_k(t_k, x_i)x_{ki}(t_k)dt_k. \qquad (13)$$

From the model in Eq. (12) functional prediction regression through generalized additive models for location scale and shape, the value of the degrees of freedom in modeling is 9.628938, and the residual degrees of freedom is 56.37106. Meanwhile, in the Eq. (13) functional prediction regression model through the generalized additive models for location scale and shape, the value of the degrees of freedom in modeling is 9.551611, and the residual degrees of freedom are 54.44839. The analysis results provide important practical insights into rainfall modeling, where the effect of the temperature covariate on rainfall is non-linear and changes smoothly over its range of values. Conversely, Model 2's assumption of simple linear interaction may be too rigid, potentially overlooking critical conditions such as threshold effects or diminishing returns. In this case, functional effects can model how the relationship between temperature and rainfall intensity changes differently across temperature regimes. Therefore, Model 1 not only better models the data but also provides a more physically plausible framework for understanding and predicting rainfall dynamics.

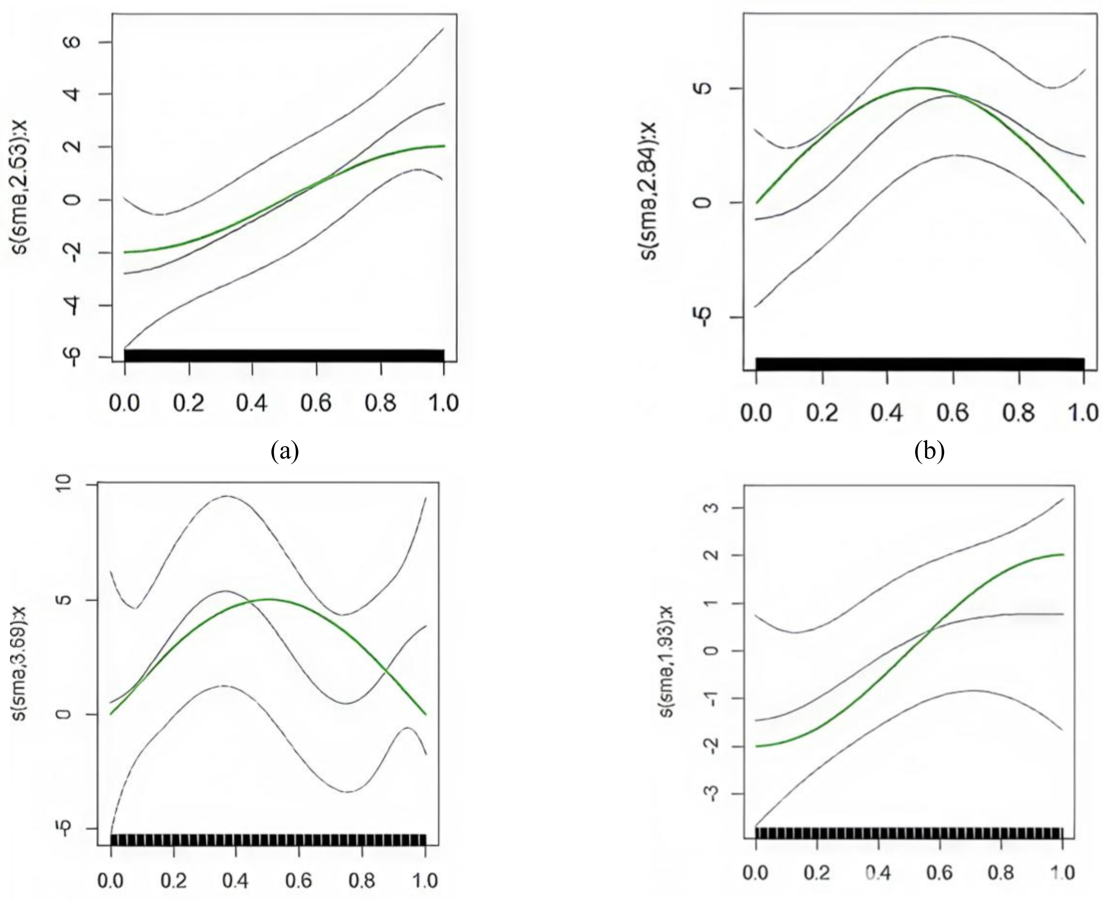


Figure 8. Evaluation Points and Linear Functional Effects (a) Model 1 for μ Parameter, (b) Model 1 for the σ Parameter, (c) Model 2 for μ Parameter, (d) Model 2 for the σ Parameter

The following are the results of the Akaike Information Criterion (AIC) values.

Table 4. Akaike's Information Criterion (AIC) Value Results

Model	AIC
Model 1	170.351
Model 2	182.187

This Akaike's Information Criterion (AIC) value shows the accuracy of testing climate change data with the functional prediction regression model. The result provides information from the two models that the best model for rainfall modeling is model 1 because it has a smaller Akaike's Information Criterion (AIC) value. Statistically, the lower AIC of Model 1 indicates a better balance between goodness-of-fit and model complexity, suggesting that additive functional effects capture rainfall patterns more parsimoniously than interaction terms, which may overfit the data. This superiority implies that Indonesia's rainfall variability is environmentally driven by simpler, temporally consistent climate mechanisms rather than complex interactions between covariates. The plot of the modeling results containing the evaluation points and effects on functional covariates for model 1 with additional predictors producing linear functional effects. Then, for model 2 with additional predictors producing the combined effect of linear interactions, it is given in Fig. 8 for the left plot is the model evaluation points for the mu parameter. The plot to the right is the model evaluation points for the sigma parameter. The green line represents the effect of each model's estimated coefficients.

The accuracy obtained by this study is based on Table 4, where the best model based on the smallest AIC is the functional prediction regression model with linear functional effects. The results found in this study align with previous research on [44], which shows that functional prediction regression with linear functional effects has a measure of model goodness based on AIC values that are better than functional prediction regression models with combined effects of linear interactions. Other research on [45], [46] shows how the performance of functional prediction regression or scalar-on-function regression methods with linear functional effects is suitable for modeling and solving various problems.

In general, this study boasts its research results in the modeling section. It provides important insights into analyzing functional data with functional linear effects and combined effects of linear interactions based on their goodness of fit through AIC values, especially in the case of rainfall. The comparative analysis of the two proposed relationships shows that the model can capture the functional linear interactions of the data well. In addition, the regulation mechanism inherent in the functional model itself has the advantages previously described in the introduction section, and the comprehensive evaluation proposed in this study by highlighting the combined effect in the model provides empirical evidence for the performance of functional prediction regression or scalar-on-function regression models in modeling rainfall problems.

4. CONCLUSION

The study results show the effectiveness of functional prediction regression in modeling Indonesia's annual rainfall and temperature from 1951 to 2016. By comparing the results of boosting and GAMLSS approaches, this study found that model 1 (functional linear effects) outperformed model 2 (linear interaction effects), evidenced by lower bootstrap iterations of μ=18 and σ=8 with the goodness of modeling by a smaller AIC value of 170.351 and 182.187. Crucially, the superior performance of model 1 implies that the relationship between rainfall and temperature in Indonesia is best captured by a smooth, dynamic function rather than a simple linear interaction. This suggests that the influence of temperature on rainfall varies in a complex, non-constant manner throughout the year. This finding provides a more nuanced understanding of Indonesia's climate system, which is essential for improving the accuracy of regional climate projections. These results highlight the utility of functional data analysis for climate modeling, with potential extensions to methods like mixed GAM computation vehicle with functional principal component analysis, MGCV-FPCA, or function-on-function regression for future research.

Author Contributions

Khusnia Nurul Khikmah: Data curation, Formal Analysis, Methodology, Project Administration, Resources, Software, Visualization, Writing - Original Draft, Writing - Review and Editing. A'yunin Sofro: Conceptualization, Funding Acquisition, Investigation, Supervision, Validation, Writing - Review and Editing. All authors had approved the final version and discussed the results and contributed to the final manuscript.

Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Acknowledgment

We are also profoundly grateful for providing the necessary resources and fostering a conducive research environment within the Mathematics Department at Universitas Negeri Surabaya.

Declarations

The authors declare no conflicts of interest.

Declaration of Generative AI and AI-assisted Technologies

Generative AI tools were used solely for language refinement (grammar, spelling, and clarity). The scientific content, analysis, interpretation, and conclusions were developed entirely by the authors. The authors reviewed and approved all final text.

REFERENCES

[1] G. Bayon, I. N. Bindeman, A. Trinquier, G. J. Retallack, and A. Bekker, "LONG-TERM EVOLUTION OF TERRESTRIAL WEATHERING AND ITS LINK TO EARTH'S OXYGENATION," *Earth Planet Sci Lett*, vol. 584, p. 117490, 2022. doi: https://doi.org/10.1016/j.epsl.2022.117490

[2] G. F. Ficetola and D. Rubolini, "CLIMATE AFFECTS GLOBAL PATTERNS OF COVID-19 EARLY OUTBREAK DYNAMICS," 2020.

[3] P. Shi *et al.*, "THE IMPACT OF TEMPERATURE AND ABSOLUTE HUMIDITY ON THE CORONAVIRUS DISEASE 2019 (COVID-19) OUTBREAK-EVIDENCE FROM CHINA," *MedRxiv*, pp. 2020–2023, 2020. doi: https://doi.org/10.1101/2020.03.22.20038919

[4] M. Habib-ur-Rahman *et al.*, "IMPACT OF CLIMATE CHANGE ON AGRICULTURAL PRODUCTION; ISSUES, CHALLENGES, AND OPPORTUNITIES IN ASIA," *Front Plant Sci*, vol. 13, p. 925548, 2022. doi: https://doi.org/10.3389/fpls.2022.925548

[5] L. Rustad *et al.*, "CHANGING CLIMATE, CHANGING FORESTS: THE IMPACTS OF CLIMATE CHANGE ON FORESTS OF THE NORTHEASTERN UNITED STATES AND EASTERN CANADA," *Gen. Tech. Rep. NRS-99. Newtown Square, PA: US Department of Agriculture, Forest Service, Northern Research Station. 48 p.*, vol. 99, pp. 1–48, 2012. doi: https://doi.org/10.2737/NRS-GTR-99

[6] B. H. Prabowo, "THE THE ECONOMIC CASCADING EFFECT ON FUTURE CLIMATE CHANGE AND AGRICULTURE ECONOMIC IN INDONESIA: ENGLISH," *Tamansiswa Management Journal International*, vol. 3, no. 1, pp. 4–7, 2021. doi: https://doi.org/10.54204/TMJI/Vol312021002

[7] O. I. O. M. F. Ilunga, "THE RAINFALL FACTOR OF CLIMATE CHANGE EFFECTS ON THE AGRICULTURAL ENVIRONMENT: A REVIEW," 2017.

[8] J. A. Lim *et al* "MITIGATING THE REPERCUSSIONS OF CLIMATE CHANGE ON DISEASES AFFECTING IMPORTANT CROP COMMODITIES IN SOUTHEAST ASIA, FOR FOOD SECURITY AND ENVIRONMENTAL SUSTAINABILITY—A REVIEW," *Front Sustain Food Syst*, vol. 6, p. 1030540, 2023. doi: https://doi.org/10.3389/fsufs.2022.1030540

[9] H. Tabari, "CLIMATE CHANGE IMPACT ON FLOOD AND EXTREME PRECIPITATION INCREASES WITH WATER AVAILABILITY," *Sci Rep*, vol. 10, no. 1, p. 13768, 2020. doi: https://doi.org/10.1038/s41598-020-70816-2

[10] P. A. White and A. E. Gelfand, "MULTIVARIATE FUNCTIONAL DATA MODELING WITH TIME-VARYING CLUSTERING," *Test*, vol. 30, no. 3, pp. 586–602, 2021. doi: https://doi.org/10.1007/s11749-020-00733-z

[11] M. Deb and T. K. Chakrabarty, "FUNCTIONAL TIME SERIES ANALYSIS OF AGE-SPECIFIC FERTILITY RATES: VISUALIZING THE CHANGE IN THE AGE-PATTERN OF FERTILITY IN INDIA," *International Journal of Advanced Statistics and Probability, 4 (2)*, pp. 79–89, 2016. doi: https://doi.org/10.14419/ijasp.v4i2.6182

[12] J.-L. Wang, J.-M. Chiou, and H.-G. Müller, "FUNCTIONAL DATA ANALYSIS," *Annu Rev Stat Appl*, vol. 3, no. 1, pp. 257–295, 2016. doi: https://doi.org/10.1146/annurev-statistics-041715-033624

[13] M. Pedersen, *FUNCTIONAL ANALYSIS IN APPLIED MATHEMATICS AND ENGINEERING*. CRC press, 2018. doi: https://doi.org/10.1201/9780203755426

[14] I. Shah, F. Jan, and S. Ali, "FUNCTIONAL DATA APPROACH FOR SHORT-TERM ELECTRICITY DEMAND FORECASTING," *Math Probl Eng*, vol. 2022, no. 1, p. 6709779, 2022. doi: https://doi.org/10.1155/2022/6709779

[15] P. Kokoszka and M. Reimherr, *INTRODUCTION TO FUNCTIONAL DATA ANALYSIS*. Chapman and Hall/CRC, 2017. doi: https://doi.org/10.1201/9781315117416

[16] I. Martínez-Hernández and M. G. Genton, "RECENT DEVELOPMENTS IN COMPLEX AND SPATIALLY CORRELATED FUNCTIONAL DATA," 2020. doi: https://doi.org/10.1214/20-BJPS466

[17] C. Li, L. Xiao, and S. Luo, "JOINT MODEL FOR SURVIVAL AND MULTIVARIATE SPARSE FUNCTIONAL DATA WITH APPLICATION TO A STUDY OF ALZHEIMER'S DISEASE," *Biometrics*, vol. 78, no. 2, pp. 435–447, 2022. doi: https://doi.org/10.1111/biom.13427

[18] S. Brockhaus, M. Melcher, F. Leisch, and S. Greven, "BOOSTING FLEXIBLE FUNCTIONAL REGRESSION MODELS WITH A HIGH NUMBER OF FUNCTIONAL HISTORICAL EFFECTS," *Stat Comput*, vol. 27, pp. 913–926, 2017. doi: https://doi.org/10.1007/s11222-016-9662-1

[19] X. Zhang, J.-M. Chiou, and Y. Ma, "FUNCTIONAL PREDICTION THROUGH AVERAGING ESTIMATED FUNCTIONAL LINEAR REGRESSION MODELS," *Biometrika*, vol. 105, no. 4, pp. 945–962, 2018. doi: https://doi.org/10.1093/biomet/asy041

[20] J. S. Morris, "FUNCTIONAL REGRESSION," *Annu Rev Stat Appl*, vol. 2, no. 1, pp. 321–359, 2015. doi: https://doi.org/10.1146/annurev-statistics-010814-020413

[21] S. Greven and F. Scheipl, "A GENERAL FRAMEWORK FOR FUNCTIONAL REGRESSION MODELLING," *Stat Modelling*, vol. 17, no. 1–2, pp. 1–35, 2017. doi: https://doi.org/10.1177/1471082X16681317

[22] S. Brockhaus, A. Fuest, A. Mayr, and S. Greven, "SIGNAL REGRESSION MODELS FOR LOCATION, SCALE AND SHAPE WITH AN APPLICATION TO STOCK RETURNS," *J R Stat Soc Ser C Appl Stat*, vol. 67, no. 3, pp. 665–686, 2018. doi: https://doi.org/10.1111/rssc.12252

[23] S. Brockhaus, M. Melcher, F. Leisch, and S. Greven, "BOOSTING FLEXIBLE FUNCTIONAL REGRESSION MODELS WITH A HIGH NUMBER OF FUNCTIONAL HISTORICAL EFFECTS," *Stat Comput*, vol. 27, no. 4, pp. 913–926, 2017. doi: https://doi.org/10.1007/s11222-016-9662-1

[24] S. Greven and F. Scheipl, "A GENERAL FRAMEWORK FOR FUNCTIONAL REGRESSION MODELLING," *Stat Modelling*, vol. 17, no. 1–2, pp. 1–35, 2017. doi: https://doi.org/10.1177/1471082X16681317

[25] L. Lusa, "GRADIENT BOOSTING FOR HIGH-DIMENSIONAL PREDICTION OF RARE EVENTS," *Comput Stat Data Anal*, vol. 113, pp. 19–37, 2017. doi: https://doi.org/10.1016/j.csda.2016.07.016

[26] S. Arjasakusuma, Y. Yamaguchi, Y. Hirano, and X. Zhou, "ENSO-AND RAINFALL-SENSITIVE VEGETATION REGIONS IN INDONESIA AS IDENTIFIED FROM MULTI-SENSOR REMOTE SENSING DATA," *ISPRS Int J Geoinf*, vol. 7, no. 3, 2018. doi: https://doi.org/10.3390/ijgi7030103

[27] A. F. Athiyyah *et al.*, "MOLECULAR EPIDEMIOLOGY AND CLINICAL FEATURES OF ROTAVIRUS INFECTION AMONG PEDIATRIC PATIENTS IN EAST JAVA, INDONESIA DURING 2015–2018: DYNAMIC CHANGES IN ROTAVIRUS GENOTYPES FROM EQUINE-LIKE G3 TO TYPICAL HUMAN G1/G3," *Front Microbiol*, vol. 10, p. 940, 2019. doi: https://doi.org/10.3389/fmicb.2019.00940

[28] Y. Fan *et al.*, "RECONCILING CANOPY INTERCEPTION PARAMETERIZATION AND RAINFALL FORCING FREQUENCY IN THE COMMUNITY LAND MODEL FOR SIMULATING EVAPOTRANSPIRATION OF RAINFORESTS AND OIL PALM PLANTATIONS IN INDONESIA," *J Adv Model Earth Syst*, vol. 11, no. 3, pp. 732–751, 2019. doi: https://doi.org/10.1029/2018MS001490

[29] P. Hall, H. G. Müller, and J. L. Wang, "PROPERTIES OF PRINCIPAL COMPONENT METHODS FOR FUNCTIONAL AND LONGITUDINAL DATA ANALYSIS," 2006. doi: https://doi.org/10.1214/009053606000000272

[30] S. Ullah and C. F. Finch, "APPLICATIONS OF FUNCTIONAL DATA ANALYSIS: A SYSTEMATIC REVIEW," *BMC Med Res Methodol*, vol. 13, pp. 1–12, 2013. doi: https://doi.org/10.1186/1471-2288-13-43

[31] J. O. Ramsay and B. W. Silverman, *APPLIED FUNCTIONAL DATA ANALYSIS: METHODS AND CASE STUDIES*. Springer, 2002. doi: https://doi.org/10.1007/b98886

[32] J. O. Ramsay and C. Dalzell, "SOME TOOLS FOR FUNCTIONAL DATA ANALYSIS," *J R Stat Soc Series B Stat Methodol*, vol. 53, no. 3, pp. 539–561, 1991. doi: https://doi.org/10.1111/j.2517-6161.1991.tb01844.x

[33] H. L. Shang and R. J. Hyndman, "NONPARAMETRIC TIME SERIES FORECASTING WITH DYNAMIC UPDATING," *Math Comput Simul*, vol. 81, no. 7, pp. 1310–1324, 2011. doi: https://doi.org/10.1016/j.matcom.2010.04.027

[34] T. T. Cai and P. Hall, "PREDICTION IN FUNCTIONAL LINEAR REGRESSION," 2006. doiL: https://doi.org/10.1214/009053606000000830

[35] J. D. Tucker, J. R. Lewis, and A. Srivastava, "ELASTIC FUNCTIONAL PRINCIPAL COMPONENT REGRESSION," *Statistical Analysis and Data Mining: The ASA Data Science Journal*, vol. 12, no. 2, pp. 101–115, 2019. doi: https://doi.org/10.1002/sam.11399

[36] F. Pfisterer, L. Beggel, X. Sun, F. Scheipl, and B. Bischl, "BENCHMARKING TIME SERIES CLASSIFICATION--FUNCTIONAL DATA VS MACHINE LEARNING APPROACHES," *arXiv preprint arXiv:1911.07511*, 2019.

[37] A. Mayr, H. Binder, O. Gefeller, and M. Schmid, "THE EVOLUTION OF BOOSTING ALGORITHMS," *Methods Inf Med*, vol. 53, no. 06, pp. 419–427, 2014.doi: https://doi.org/10.3414/ME13-01-0122

[38] P. Bühlmann and B. Yu, "BOOSTING WITH THE L 2 LOSS: REGRESSION AND CLASSIFICATION," *J Am Stat Assoc*, vol. 98, no. 462, pp. 324–339, 2003.doi: https://doi.org/10.1198/016214503000125

[39] S. Brockhaus, D. Rügamer, and A. Stöcker, "FDBOOST: BOOSTING FUNCTIONAL REGRESSION MODELS," *R package version 0.2-0,* https://CRAN. R-project. org/package=FDboost, 2016.

[40] S. Ray and B. Bhattacharyya, "AVAILABILITY IN DIFFERENT SOURCE OF IRRIGATION IN INDIA: A STATISTICAL APPROACH," *Ecosystem*, p. 109, 2015.

[41] S. Portet, "A PRIMER ON MODEL SELECTION USING THE AKAIKE INFORMATION CRITERION," *Infect Dis Model*, vol. 5, pp. 111–128, 2020. doi: https://doi.org/10.1016/j.idm.2019.12.010

[42] J. E. Cavanaugh and A. A. Neath, "THE AKAIKE INFORMATION CRITERION: BACKGROUND, DERIVATION, PROPERTIES, APPLICATION, INTERPRETATION, AND REFINEMENTS," *Wiley Interdiscip Rev Comput Stat*, vol. 11, no. 3, p. e1460, 2019. doi: https://doi.org/10.1002/wics.1460

[43] C. M. Crainiceanu, J. Goldsmith, A. Leroux, and E. Cui, *FUNCTIONAL DATA ANALYSIS WITH R*. Chapman and Hall/CRC, 2024. doi: https://doi.org/10.1201/9781003278726

[44] P. T. Reiss, J. Goldsmith, H. L. Shang, and R. T. Ogden, "METHODS FOR SCALAR-ON-FUNCTION REGRESSION," *International Statistical Review*, vol. 85, no. 2, pp. 228–249, 2017. doi: https://doi.org/10.1111/insr.12163

[45] M. W. McLean, G. Hooker, and D. Ruppert, "RESTRICTED LIKELIHOOD RATIO TESTS FOR LINEARITY IN SCALAR-ON-FUNCTION REGRESSION," *Stat Comput*, vol. 25, pp. 997–1008, 2015. doi: https://doi.org/10.1007/s11222-014-9473-1

[46] A. Bauer, F. Scheipl, H. Küchenhoff, and A.-A. Gabriel, "AN INTRODUCTION TO SEMIPARAMETRIC FUNCTION-ON-SCALAR REGRESSION," *Stat Modelling*, vol. 18, no. 3–4, pp. 346–364, 2018. doi: https://doi.org/10.1007/s11222-014-9473-1