Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.18576/amis/180307

Integration of GSTARIMA Model with Heteroskedastic Error and Kriging for Climate Forecasting: A Systematic Review

Putri Monika[1], *Budi Nurani Ruchjana*[2,*], *Atje Setiawan Abdullah*[3] *and Rahmat Budiarto*[4]

[1]Doctoral Program of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Sumedang 45363, Indonesia
[2]Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Sumedang 45363, Indonesia
[3]Department of Computer Science, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Sumedang 45363, Indonesia
[4]College of Computer Science and Information Technology, Al-Baha University, Alaqiq 65779-7738, Saudi Arabia

Received: 12 Feb. 2024, Revised: 2 Mar. 2024, Accepted: 16 Mar. 2024
Published online: 1 May 2024

Abstract: This paper discusses the systematic literature review (SLR) for the integration of the Generalized Space-Time Autoregressive Integrated Moving Average (GSTARIMA) model with heteroscedastic error and the Kriging method for climate forecasting. The GSTARIMA model is one of the Spatio-Temporal Models with powerful forecasting capabilities. GSTARIMA model with Autoregressive Conditional Heteroscedasticity (ARCH) model to overcome the non-constant error variance and Kriging method for forecasting at unobserved locations. The modelling framework and procedures follow the data analytics life cycle methodology to handle climate big data. This paper aims to show the gap analysis in the research of the GSTARIMA model for climate modelling. The SLR method includes three stages: collecting papers from the database, filtering and selection process using the PRISMA method, and conducting a gap analysis for future work. This research inspires researchers to contribute to improving and refining the model, making it a more potent and valuable tool in climate forecasting.

Keywords: GSTARIMA, heteroscedastic error, Kriging method, climate, data analytics life cycle

1 Introduction

The climate is a statistical description of the average variability of the relevant quantities over months to years, referred to as average weather [1]. In addition, it includes several interrelated elements, such as temperature, rainfall, humidity, atmospheric conditions, and wind patterns [2]. Climate change is a pressing global issue of paramount importance that demands comprehensive research. The Intergovernmental Panel on Climate Change (IPCC) was founded by scientists worldwide to research the concept. The sixth assessment report of the IPCC explains that climate change affects ecosystem conditions, human activities, the global water cycle, infrastructure, health, and others [3].

The handling of climate change is in the world's spotlight, included in the pillars of Sustainable Development Goals (SDGs) [4,5]. Meanwhile, climate management is the 13th goal of the SDGs, with the mission statement, "Take urgent action to combat climate change and its impacts by regulating emissions and promoting developments in renewable energy? [6]. Climate change is a worldwide concern that impacts every country on the planet. On the other hand, climate change's impact differs depending on area and country. Some countries are more vulnerable than others to the consequences of climate change. For example, low-lying island nations such as the Maldives and Tuvalu risk being submerged by rising sea levels [7]. Droughts and desertification are being caused by climate change in Africa, leading to food shortages and displacement [8]. Climate change creates more frequent and extreme wildfires, hurricanes, and floods in the United States [9, 10,11]. Climate change generates heatwaves, droughts, and flooding throughout Europe [12]. Indonesia is vulnerable to the effects of climate change, including catastrophic occurrences such as floods, droughts, and

* Corresponding author e-mail: budi.nurani@unpad.ac.id

storms, as well as long-term changes caused by sea-level rise. Rising sea levels and extreme weather events such as floods, droughts, and hurricanes are already wreaking havoc on Indonesia's coastal communities, infrastructure, and ecosystems [13]. For the three future eras of the 2020s, 2050s, and 2080s, the average annual temperature in Afghanistan is predicted to rise by 1.8 C, 3.5 C, and 4.8 C, respectively [14]. As a result, major weather events have occurred, including four years of flooding that have submerged half of the country. Turkey has seen frequent and intense heatwaves, droughts, and wildfires [15]. Water scarcity is also harming the country's agricultural industry and economy. Climate change is very detrimental regarding materials, infrastructure, and people's lives. Therefore, it is essential to forecast future climate conditions to take preventive, mitigation, and adaptation actions [16,17].

Spatio-Temporal modelling for climate forecasting has made significant progress in recent years, as evidenced by the diverse and innovative approaches presented in the summarized articles. These approaches mainly focus on harnessing the power of mathematics-statistical modelling, deep learning, and data assimilation techniques to understand and predict climate-related phenomena [18,19,20]. The Spatial-temporal representation combines location and time to define a phenomenon and understand the relationship between spatial and temporal changes. The Space-Time Autoregressive (STAR) Model with homogenous characteristics between locations, with the same parameters for each location, is the most often used Spatio-Temporal model based on the Box Jenkins method [21]. The Generalized Space-Time Autoregressive (GSTAR) model is created with diverse inter-location characteristics. The parameters of the GSTAR model differ depending on location [22]. The GSTARIMA model is a development model for analyzing nonstationary data and model errors following the Moving Average Model [23].

The GSTAR-ARCH model is a statistical model used to assess climate data with heteroscedastic error variance, indicating that the error variance is not constant. This model extends the Spatio-Temporal Model by accounting for heteroscedasticity caused by auto-regressive prior knowledge in stationary data. Furthermore, the GSTARI-ARCH model has been designed to handle non-stationary data. A GSTAR model that considers the heteroscedasticity of errors is also developed by involving exogenous variables called the GSTARI-X-ARCH model. This model forecasts climate through rainfall with exogenous variables in relative humidity. Another Spatio-Temporal model that considers heteroscedastic errors is the STARMA-GARCH Model, which forecasts temperature [24].

Previous research used the Kriging Method to predict phenomena at unobserved locations. Kriging method is used for interpolation and forecasting Temperature in Mosul and Baghdad City [25]. Kriging method, land-use regression (LUR), and LightGBM (light gradient boosting machine) methods were combined to predict PM2.5 concentrations [26]. In Spatio-Temporal modelling, the GSTAR Model is integrated with the Kriging method to forecast rainfall at unobserved locations in West Java [27].

Spatio-Temporal modeling in climate forecasting requires using big data analytics [28]. A big data approach is needed, such as knowledge discovery in databases (KDD) data mining, and data analytics lifecycle. These approaches make it possible to find patterns and forecast the future by extracting meaningful information from enormous datasets. The volume, velocity, and variety of data and the necessity for efficient data processing and analysis tools are some hurdles in using big data analytics for climate forecasting [29]. Analyzing climate forecasts requires collecting, cleaning, and managing data from various sources, including historical weather data, sensor data, satellite data, and more.

This study summarises previous research on Spatio-Temporal forecasting models with heteroskedastic errors and the Kriging method applied to climate data. This research attempts to cover several areas, such as Spatio-Temporal models for stationary and non-stationary data, methods for parameter estimation in the models, forecasting at unsampled locations, and the potential to integrate Spatio-Temporal models with Heteroskedastic errors and Kriging for climate forecasting. Ultimately, this review contributes to a broader understanding of integrated Spatio-Temporal Models with Heteroskedastic errors and the Kriging Method for climate and highlights avenues for further research and innovation in this critical area. To facilitate the analysis process, we formulate the following research questions (RQs):

RQ1 How to integrate the GSTARIMA model with heteroskedastic errors using the Kriging method?
RQ2 How to forecast climate phenomena using integrating GSTARIMA and Kriging models through a data analysis life cycle approach?

RQs were examined and explored by reviewing previous results and searching literature on databases. The results were filtered and selected using the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) method. Furthermore, relevant articles are presented in a state-of-the-art manner to identify research gaps. A bibliometric method was also used to show the linkage of keywords for each article. The review stage was performed to analyze search results and discuss new research. Potential new research was provided to be studied and developed on the GSTARIMA Model and its application.

2 Materials and Methods

2.1 Theoretical Background

2.1.1 The GSTARIMA Model

In 1980, Pfeifer and Deutch introduced the STAR Model, assuming each location has the same characteristics [30, 31]. The STAR model was developed into a GSTAR model because the assumptions in the STAR model do not match the reality in the field, where there is a diversity of characteristics in each location. The GSTAR model introduced by Ruchjana assumes that the characteristics of each location are heterogeneous. The GSTAR(p, λ_k) model has a time order of p and a spatial order of λ_k expressed in matrix form through equation (1) [22]:

$$\mathbf{z}(t) = \sum_{k=1}^{p} \sum_{l=0}^{\lambda_k} \left[\boldsymbol{\Phi}_{kl} \mathbf{W}^{(l)} \mathbf{z}(t-k) \right] + \mathbf{e}(t), \tag{1}$$

where
$Z(t)$: the value of the observation at time t,
$Z(t-1)$: the value of the observation at time $t-1$,
ϕ : a parameter that indicates the influence of the value of $Z(t-1)$ on the value of $Z(t)$,
$e(t)$: the value of error.

The GSTARMA model expands the GSTAR model by adding MA error elements. The GSTARMA model is applied to stationary data [32]. The GSTARMA model developed on nonstationary data is called the GSTARIMA model. Min et al. [23] first introduced the GSTARIMA model with application to urban traffic network modeling and short-term traffic flow forecasting. The GSTARIMA model $(p_{\lambda_k}, d, q_{v_k})$with d being the differencing order is expressed in Equation (2):

$$\begin{aligned} \mathbf{y}(t) = & \sum_{k=1}^{p} \sum_{l=0}^{\lambda_k} \left[\boldsymbol{\Phi}_{kl} \mathbf{W}^{(l)} \mathbf{y}(t-k) \right] \\ & - \sum_{k=1}^{q} \sum_{l=0}^{v_k} \left[\boldsymbol{\Theta}_{kl} \mathbf{W}^{(l)} \mathbf{e}(t-k) \right] + \mathbf{e}(t), \end{aligned} \tag{2}$$

where

$$\begin{aligned} \mathbf{y}(t) &= \mathbf{z}(t) - \mathbf{z}(t-1), \mathbf{y}(t-1) \\ &= \mathbf{z}(t-1) - \mathbf{z}(t-2), \ldots, \mathbf{y}(t-k) \\ &= \mathbf{z}(t-k) - \mathbf{z}(t-k-1), \end{aligned} \tag{3}$$

$\mathbf{z}(t)$: a vector of variables of size $(N \times 1)$ at time t,
$\mathbf{z}(t-k)$: vector of variables of size $(N \times 1)$ at time $(t-k)$,
λ_k : spatial order in the k-th autoregressive,
v_k : spatial order of the k-th moving average,
$\boldsymbol{\Phi}_{kl}$: autoregressive and space-time parameters at time order k and spatial order l of size $(N \times N)$ in the form of diagonal matrix $\left(\Phi_{kl}^{(1)}, \Phi_{kl}^{(2)}, \Phi_{kl}^{(3)}, \ldots, \Phi_{kl}^{(N)}\right)$,
$\boldsymbol{\Theta}_{kl}$: MA parameters at time order k and spatial order l of size $(N \times N)$ in the form of diagonal matrix $\left(\Theta_{kl}^{(1)}, \Theta_{kl}^{(2)}, \Theta_{kl}^{(3)}, \ldots, \Theta_{kl}^{(N)}\right)$,
$\mathbf{W}^{(l)}$: weight matrix of size $(N \times N)$ at spatial order $l, l = 0, 1, 2, \ldots, \lambda_k$ containing $w_{ii} = 0$ and $\sum_{i \neq j} w_{ij} = 1$,
$\mathbf{e}(t)$: error vector of size $(N \times 1)$ at time t, assuming $\mathbf{e}(t)^{iid} N\left(\mathbf{0}, \sigma^2 \mathbf{I}\right)$.

2.1.2 Autoregressive Conditional Heteroscedasticity (ARCH) and Generalized-ARCH (GARCH) Model

Although the GSTARIMA model assumes constant error variance, applying climate data often shows non-constant error variance. To overcome this, the GSTARIMA model is integrated with the Autoregressive Conditional Heteroscedasticity (ARCH) Model. This time series model detects variance heteroscedasticity using historical data [33]. Describing the ARCH(p) model, researchers use the following expression [33]:

$$h_t = \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \mathbf{e}_{t-i}^2 \,; i = 1, 2, 3, \ldots, p. \tag{4}$$

In Equation (4), the variables represented include:
h_t : the conditional variance at time t,
α_0 : the intercept or constant error,
$\alpha_1, \alpha_2, \ldots, \alpha_p$: ARCH model parameters,
$\alpha_0 > 0$ and $\alpha_i \geq 0$.

Bollerslev (1986) developed the GARCH model, which is a development of the ARCH model. Equation (5) presents the GARCH (p,q) Model equation.

$$h_t = \sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i e_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i}, \tag{5}$$

where $p \geq 0$, $q > 0$, $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, \ldots, q$, $\beta_i \geq 0$, $i = 1, \ldots, p$.

2.1.3 Kriging Method

The Kriging method is a geostatistical interpolation technique used to predict variable values at unobserved locations based on variable values observed at other locations. This method assumes that variable values have a spatial structure related to the distance and direction between observation locations. A semivariogram is required to calculate the Kriging Method. An experimental semivariogram is calculated based on measurement data collected from the field or observations at a particular location. The formula for calculating the experimental semivariogram is as follows [34]:

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(k)} [Z(x_i+h) - Z(x_i)]^2, \tag{6}$$

where

$\hat{\gamma}(h)$	:	semivariogram value at distance h,
$Z(x_i)$	:	observation value at location x_i,
$Z(x_i+h)$	:	observation value at location x_i+h,
$N(h)$	:	many pairs of data that have distance h,
h	:	distance between 2 locations.

Theoretical semivariograms can be divided into Spherical, Gaussian, and Exponential Models. The Spherical Model is a model that assumes that spatial dependence has a certain maximum distance or radius. This Model is used if the spatial dependence decreases with distance and reaches a threshold value at a specific radius, after which the semivariogram value becomes constant. The semivariogram function of the Spherical Model can be expressed as [34]:

$$\gamma(h) = \begin{cases} c\left[\left(\frac{3h}{2a}\right) - \left(\frac{h}{2a}\right)^3\right], & h \le a \\ c, & h > a \end{cases}. \tag{7}$$

The Exponential Model is a model that assumes that spatial dependence decreases exponentially with distance between locations. The semivariogram function of the Exponential Model can be expressed as [34]:

$$\gamma(h) = \begin{cases} c\left[1 - exp\left(\frac{-h}{a}\right)\right], & h \le a \\ c, & h > a \end{cases}. \tag{8}$$

The Gaussian Model is a model that assumes that spatial dependence has a symmetric pattern and decreases exponentially with distance between locations. The semivariogram function of the Gaussian Model can be expressed as [34]:

$$\gamma(h) = \begin{cases} c\left[1 - exp\left(\frac{-h}{a}\right)^2\right], & h \le a \\ c, & h > a \end{cases}, \tag{9}$$

where

h	:	distance between sample locations,
c	:	sill value,
a	:	range.

The semivariogram also provides the weights used in interpolation. The Kriging method aims to determine the value of the Kriging weight θ_i, which minimizes the estimator's variance so that a BLUE (Best Linear Unbiased Estimator) estimator is obtained. The Kriging estimator $\hat{Z}(x_0)$ can be written as follows [34]:

$$\hat{Z}(x_0) - \xi(x_0) = \sum_{i=1}^{n} \theta_i [Z(x_i) - \xi(x_i)], \tag{10}$$

where

$\hat{Z}(x_0)$	:	Kriging estimator at unobserved location x,
x_i	:	the ith data location adjacent to the unsampled location x,
$\xi(x_0)$	:	expectation value of $Z(x_0)$,
$\xi(x_i)$	:	expectation value of $Z(x_i)$,
n	:	many sample data used for estimation,
θ_i	:	weight value at location i.

2.1.4 Data Analytics Life Cycle

Climate data has the Big Data criteria of volume, variety, and velocity. Big Data could be more efficient when analyzed using traditional methods. The Data Analytics Life Cycle methodology is designed to handle Big Data problems and data science projects. The Data Analytics Life Cycle consists of six phases, including [35]:

- Discovery → At this stage, researchers must study, search, and investigate facts, identify problems, and develop context and understanding of the data sources needed to support research.
- Data Preparation → Next, data is cleaned to identify missing values or noisy data. The results of data cleaning are transformed by aggregating daily data into monthly or according to the needs of the analysis. In this case, pre-processing data is obtained and ready for processing and analysis.
- Model Planning → At this stage, the model planning that will be used for analysis is carried out.
- Model Building → Researchers divide the results of data preparation into in-sample data (training) and out-sample data (testing) to do forecasting.
- Communicate Results → Researchers present forecasting results using visualizations in the form of time series plots, choropleth maps, diagrams, and others.
- Operationalize → The final stage is operationalized, and researchers provide final reports, recommendations, scripts, and technical documents. In addition, researchers can apply the model to the appropriate environment.

2.2 Collected Article

The PRISMA method is a widely used guide and methodological framework for conducting and presenting systematic reviews and meta-analyses [25]. The method provides the results of a systematic review, including completeness and clarity in reporting. The PRISMA method is supported by flowcharts in selecting articles [26,27].

The first stage in the PRISMA method is a literature search. Meanwhile, a literature search using keywords was carried out in this research in four databases, namely Google Scholar, Dimensions, Science Direct, and Scopus. The keywords entered in the database consist of four codes connected with ”OR” and ”AND”. The criteria selected in the collection of articles include:

1.The publication type selected is article research and conference paper.
2.Written in English
3.The range of article publications is 2000-2023.
4.The title, abstract, or keywords contain the words presented in Table 1.

Table 1: Keywords used for literature search.

Codes	Keywords
A	(”Spatio Temporal” OR ”GSTAR” OR ”GSTARIMA” OR ”Generalized Space-Time Autoregressive”)
B	(”Heteroscedasticity” OR ”ARCH” OR ”GARCH” OR ”Seemingly Unrelated Regression” OR ”SUR” OR ”Kriging Method”)
C	(”Data Analytics Life Cycle” OR ”Data Mining” OR ”Big Data Approach” OR ”Climate Change” OR ”Extreme Rainfall” OR ”Weather” OR ”Temperature”)
D	A AND B AND C

The keywords provided in Table 1 are input into the database, followed by the enter key to initiate a search. After displaying the search results, criteria 1-3 pertain to the publication type, language selection, and publication year range are configured to filter articles under the specified parameters. Subsequently, eligible articles are downloaded in .bib, .csv, and .ris formats. The number of article findings in each database is recorded for utilization as reference material in the subsequent stage.

The second stage involves the selection of articles, which is carried out through a manual process to ensure relevance. Specifically, the criteria for selecting relevant articles explore the GSTAR model and its application. The articles included at this stage comprise the ones obtained from the initial database search and those found manually through citation searching. The stages in article selection are explained as follows [39,40,41]:

(a)Duplicate selection aims to remove duplicate articles found. Duplication can be found in databases or literature sources with almost the same or similar structure. The duplication selection stage can be conducted with special software such as Jabref and Mendeley reference managers to compare titles, abstracts, and content.
(b)The relevance of the title and abstract is selected by assessing and ensuring that it matches the topic criteria. Titles and abstracts of selected articles are read in their entirety, and irrelevant ones are excluded at this stage.
(c)The full selection aims to determine whether the discussion and content in the article are relevant to the topic. All articles are accessed and read manually to ensure their appropriateness. Articles that fail to meet the established criteria or do not pertain to the subject matter under investigation are hereby excluded from the subsequent phases of the process

The final stage in the PRISMA method is the articles review, explaining, and answering the RQs presented in Section 1.

3 Results and Discussion

3.1 Results of Literature Search and Dataset Analysis

The results of the literature search are presented in Table 2, where code A produces 213,557 articles, code B produces 1,121,262, code C produces 7,525,693, code D is searched by combining code A, B, and C the ”AND” connector to produce 286 articles.

Table 2: Keyword search results in the database.

Codes	Scopus	Dimensions	EBSCO-Host	Total
A	101,483	69,050	34,024	213,557
B	339,122	515,898	266,242	1,121,262
C	1,381,753	4,046,170	2,097,770	7,525,693
D	77	71	138	286

The manual selection stage of the article is carried out as follows:

(a)At the initial stage, duplicate selection is conducted to identify 161 articles as duplicates and removed from the study.
(b)The selection stage is based on the relevance of the title and abstract, where 35 articles are selected as relevant and considered for further research.
(c)In the full paper accessibility selection stage, 60 articles can be accessed and downloaded for further selection.
(d)In the full paper relevance selection stage, the entire contents of the 18 articles are read and analyzed to determine their relevance. Relevant papers were also added using another method, citation search, resulting in 32 relevant articles. So, a total of 48 review articles were obtained that were relevant to the topic discussed.

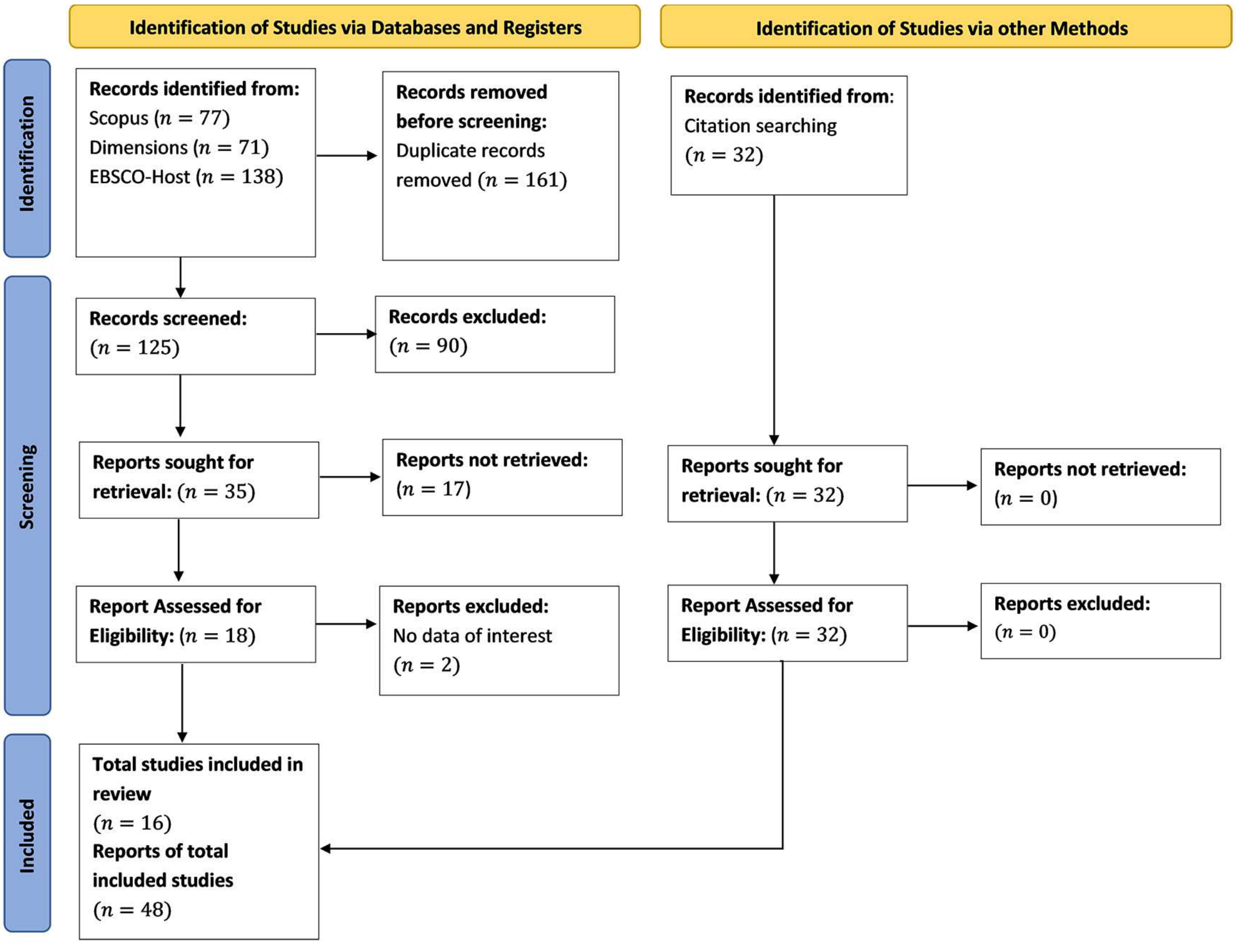


Fig. 1: PRISMA Diagram for Relevant Article Selection on Spatio Temporal Model.

These stages are presented visually in the PRISMA diagram in Figure 1 with three stages: identification, screening, and inclusion. Identification includes the duplication selection stage in stage (a). Screening consists of stages (b) and (c) for selecting title-abstract and full paper. Finally, inclusion explains the number of research articles relevant to the topic.

3.2 Bibliometric Analysis

The next stage describes the selected articles in bibliometric mapping used as a visualization method to analyze the pattern of relationships between scientific articles [42,43,44]. This paper uses bibliometric maps to visualize scientific networks involving keywords in 48 articles. The visualization results are in circles and clusters distinguished by different colors. The circles on the bibliometric map represent the number of related publications by keyword. A circle with a large size indicates several keywords with similar relationships between scientific articles. Clusters in a bibliometric map show connected circles and represent scientific articles with similarities in context, such as topics [45]. Furthermore, bibliometric mapping keyword analysis is obtained using VOSviewer to understand the structure, patterns, and relationships between scientific articles [46]. VOSviewer analyzes keywords that frequently appear in articles and identifies the relevant ones. The results of the bibliometric mapping for keyword analysis with VOSviewer are presented in Figure 2.

Figure 2 was created using the VosViewer software, and 48 relevant articles are saved in .ris format. Article files are inputted into VOSviewer, a mapping selected for co-occurrence words. The bibliometric mapping in Figure 2 shows that the co-occurrence of keywords consists of five clusters. These clusters indicate the link between "Spatio-Temporal Models" and "Climate." Forecasting climate is done chiefly with Spatio-Temporal Models and Time Series Models. In Figure 2, it can be seen that there are clusters that show climate variables that are often used by researchers, such as rainfall, Pacific Decadal Oscillation (PDO), atmospheric pressure, etc.

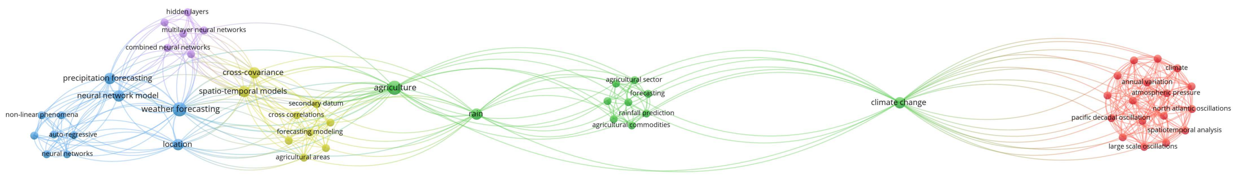


Fig. 2: Bibliometric mapping of keywords contained in 48 relevant articles on Spatio Temporal Model.

As revealed by an analysis of 48 relevant articles, the state-of-the-art in this field highlights significant progress in several key topics shown in Table 3. First, ”GSTARIMA models” are emerging as a prominent approach to analyzing Spatio-Temporal data. This cutting-edge model combines the capabilities of time series analysis and spatial relationships, enabling a comprehensive understanding of complex interactions. Secondly, the exploration of ”Heteroscedastic Error” in this study is in terms of overcoming the non-constant variance of errors in the GSTARIMA Model. By addressing these heteroscedastic errors, researchers aim to improve the accuracy and reliability of their forecasts, ultimately leading to more robust modeling results. In addition, ”Kriging,” a geostatistical interpolation technique, plays an essential role in spatial analysis. This method incorporates the estimation of unknown values based on observed values in the vicinity, incorporating spatial correlation. Collectively, these advances show the evolving research landscape in spatial-temporal analysis, featuring the integration of cutting-edge methodologies such as the GSTARIMA Model, the consideration of heteroscedastic errors, and the application of techniques such as Kriging to unravel complex spatial patterns and relationships.

Table 3 provides a comprehensive overview of the research developments related to the GSTARIMA/Spatio-Temporal model. Several studies have been conducted on Spatial-temporal modeling while considering heteroscedastic errors. Kumar et al. [24] used a STARMA-GARCH model to forecast monthly temperatures, resulting in minimal Mean Absolute Percentage Error (MAPE) values in their predictions. Similarly, Monika et al. [28] used the GSTARI-X-ARCH model to forecast rainfall influenced by humidity, showing favorable forecast accuracy. In a different context, Akbar et al. [75] introduced the GSTARMAX model to forecast air pollutants in Surabaya, achieving low Root Mean Square Error (RMSE) values. Furthermore, several articles have integrated the spatial-temporal and kriging models. Dhaher et al. [25] applied the Spatio-Temporal-Kriging model for temperature interpolation and prediction in Baghdad and Mosul cities. Dai et al. [26] used four methods, including LUR, LightGBM, ML, and Kriging, to forecast PM2.5 concentrations, which resulted in satisfactory R2 accuracy. Pramoedyo et al. [90] adopted the GSTARX-SUR-Kriging model to forecast cocoa plant diseases affected by rainfall with reasonably accurate forecast results. However, Abdullah et al. [27] used the GSTAR-Kriging model to forecast rainfall in unobserved locations and produced fairly reliable predictions. Shu-qin et al. [89] explored two different approaches, namely GWR and Kriging methods.

3.3 GAP Analysis

Conducting a GAP analysis based on relevant articles illustrates the evolving research landscape in Spatio-Temporal modeling, heteroscedastic errors, and Kriging methodologies for forecasting climate and environmental data. These articles collectively represent vital insights and areas that need further exploration. The research that has been evaluated demonstrates a high propensity to use GSTARIMA models’ capacity to forecast climate-related variables like temperature, precipitation, and air pollutants [27,53,65,73,74,75,86]. A common thread is the evaluation of model performance metrics, especially MAPE, RMSE, R2, and MSE. However, the gap lies in comprehensively exploring complex parameter configurations in the GSTARIMA framework, especially in dynamic Spatio-Temporal systems. In addition, progress still needs to be made in validating these models using more sophisticated techniques, especially in handling larger, higher-dimensional data sets.

It is clear that heteroskedastic errors are critical to climate prediction, and special attention has been paid to using ARCH and GARCH models to address this issue [24,28,77]. Researchers concentrate on achieving higher prediction accuracy, indicated by lower RMSE and MSE values. However, there are differences in research in dealing with complex and non-linear forms of heteroscedasticity, which can arise from complex climate datasets. Identifying more flexible methods to handle this complexity could be an interesting subject of investigation. By incorporating Kriging into a Spatio-Temporal model, discernible trends can be identified, especially in interpolation and forecasting climate variables [25,26,27,66,81]. This study relies heavily on RMSE and MAE as tools to assess prediction

Table 3: State-of-the-art from 48 relevant articles.

No	Reference	Models	Dataset	Application	Model Assumptions				Model Performance Analysis			
					MA Component	**Exogenous Variable**	**Hetero. Error**	**Kriging Method**	MAPE	RMSE	MSE	Accuracy
1	Dhaher et al. [25]	Kriging, Spatio-Temporal	Temperature Data in Mosul and Baghdad city	Interpolate and Forecasting Temperatures	-	-	-	✓	-	A) Mosul = 0.16; B) Baghdad= 1.05; C) A+B=0.61	-	-
2	Puica et al. [47]	ST, Kriging	Historical forecasts and actual wind speed observations	Predicting wind speeds in Southern California	-	-	-	✓	Temporal Model = 5.7%-18.8%	-	-	-
3	Garcıa et al. [48]	ST model with a Generalized Extreme Value (GEV)	The temperature data used come from a simulation of the climate using the community WRF model	Bayesian hierarchical spatio-temporal model with a Generalized Extreme Value (GEV) parametrization of the extreme data to analyze the dataset.	-	-	-	-	-	-	-	-
4	Huguenin et al. [49]	ST	The mean total monthly precipitation, the El-Ni˜no Southern Oscillation (ENSO), the Atlantic Multidecadal Oscillation(AMO), and the Caribbean Low-Level Jet (CLLJ)	Examining the meteorological factors that affect extreme precipitation events in a Costa Rican basin	-	-	-	-	-	-	-	-
5	Zhang et al. [66]	ST, deep learning	The global SST, Sea Surface Temperature record is from 1998, ENSO events	ENSO Forecasting	-	-	-	-	-	-	MSE value from 0.019 to 0.026; ED-PredRNN (FP32/Mixed)= 0.0273/0.0330	-
6	Raman et al. [51]	ST, Kriging The collected subsurface water samples, in situ field analysis was performed for	water temperature (°C), dissolved oxygen (mg/l), pH, total alkalinity (mg/l), and sp. conductivity (μS/cm).	River water quality-Ganga	-	-	-	✓	-	OK (Gaussian) Alkalinity= 5.62; Sp. Cond= 15.28	-	-
7	Fung et al. [52]	ST, Kriging, GWR, IDW	The daily rainfall data for Peninsular Malaysia from 1988 to 2017	Spatio Temporal analysis of rainfall in Malaysia	-	-	-	✓	-	IDW=91.2;OK(exp)= 100.9; OK(lin)= 107.7; GWR=112.8; MGWR=95.9	-	$R_{I}DW^{2}$=0.540; $R_{O}KE^{2}$=0.362; $R_{O}KL^{2}$=0.189; $R_{G}WR^{2}$=0.480; $R_{M}GWR^{2}$=0.574
8	Dai et al. [26]	LUR, LightGBM, ML, Kriging	PM2.5 site monitoring data (http://106.37.208.233:20035/)	Spatio-Temporal Characteristics of PM2.5 Concentrations	-	-	-	✓	-	-	-	R2= 0.976 (average for 2016–2021)
9	Kumar et al. [24]	STARMA, GARCH	Temperature Data (https://power.larc.nasa.gov/data-accessviewer/)	Forecasting Monthly Temperature	✓	-	✓	-	MAPE for Max. Temperature 2-4% and MAPE for Temperature Range 10-12%	-	-	-
10	Monika et al. [28]	GSTARI-X-ARCH	Climate Data (https://power.larc.nasa.gov/data-accessviewer/)	Forecasting Climate in West Java	-	✓	✓	-	MAPE In-Sample= 20%, MAPE Outsample= 19%	-	-	-
11	Mukhaiyar et al. [53]	GSTAR	The average daily wind speed from NOAA	Predict the occurrence of Hurricane Katrina	-	-	✓	-	MAPE= 6.86	-	MSE=0.86	MAD=0.70
12	Sofi et al. [54]	ST	The maximum temperature from NASA POWER database	Temperature forecasting	-	-	-	-	-	ConvLSTM= 3.298	ConvLSTM =986.86	-
13	Xiang et al. [55]	ST Transformer U-Net (ST-UNet)	Dataset of hourly temperature forecasts from 2013 to 2017	Temperature forecasting	-	-	-	-	-	ST-UNet= 0.8763	-	-

Table 3: State-of-the-art from 48 relevant articles (continued).

No	Reference	Models	Dataset	Application	Model Assumptions				Model Performance Analysis			
					MA Component	Exogenous Variable	Hetero. Error	Kriging Method	MAPE	RMSE	MSE	Accuracy
14	Zhao et al. [56]	Coupled forecast system model version 2 (CFSv2) with the multilinear regression model	SST, the daily precipitation data	The study analyzes the spatial-temporal distribution of precipitation in the Huai river basin and explores the relationship between precipitation and global SST.	-	-	-	✓	-	Information fusion with CFSv2 statistical downscaling model=72.31	-	-
15	Anshuka et al. [57]	LSTM deep learning model	The NOAA Climate Prediction Centre Global Unified Gauge-Based Analysis of Daily Precipitation data and the global analyses of monthly Kaplan Sea Surface temperature anomalies data	Spatio-temporal hydrological extreme forecasting in the South Pacific region using a long short-term memory deep learning model.	-	-	-	-	-	RMSE Value from 0.295 to 0.357	The lowest MSE (0.2)	Average = 0.75
16	Hou et al. [58]	Multi-In and Multi-Out (MIMO) model	Optimum Interpolation SST (OI-SST) data from the National Oceanic and Atmospheric Administration (NOAA) and select the region of Ni˜no 3.4 as the target region for prediction.	Predict sea surface temperature (SST) at different temporal scales	-	-	-	-	For Monthly: MIMO-122= 3.55%; MIMO-911= 2.63%; MIMO= 2.24%; MIMO-433= 2.37%; MIMO-244= 3.59%	For Monthly: MIMO-122= 1.17; MIMO-911= 0.90; MIMO= 0.76; MIMO-433= 0.82; MIMO-244= 1.19	For Monthly: MIMO-122= 1.37; MIMO-911= 0.81; MIMO= 0.58; MIMO-433= 0.66; MIMO-244= 1.41	-
17	Yu et al. [59]	ST graph attention network (STGAT)	Dataset of hourly air temperature observations from 2013 to 2017 in Beijing, China	Air temperature forecasting	-	-	-	-	-	Zhe= 1.2968; Min= 1.3149; Yue= 1.3988	-	-
18	Chen et al. [60]	ST, Kriging	The annual average rainfall	Rainfall forecasting	-	-	-	✓	-	Low RMSE value	-	-
19	Thorson et al. [61]	ST, Bayesian hierarchical framework	Dataset of fish community data from 1977 to 2018 in the Gulf of Alaska	Spatio-temporal ecosystem model for forecasting community reassembly under changing climate conditions	-	-	-	-	-	-	-	-
20	Kong et al. [62]	Deep spatio-temporal forecasting model (Deep-STF)	Dataset of hourly temperature observations from 2013 to 2017 in China	Multi-site weather prediction post-processing	-	-	-	-	-	Deep-STF= 2.41	-	-
21	Kuo et al. [63]	Kriging	The sensors and the weather stations (http://e-service.cwb.gov.tw)	Comparing Kriging Estimators	-	-	-	✓	-	RMSE¡3	-	MAE¡3
22	Ghorbani et al. [64]	ST, Kriging	The rainfall data and mean air temperature data for the period 2009-2019	Rainfall and mean air temperature forecasting	-	✓	-	✓	-	Rainfall RMSE value from 5 to 21, Temp. RMSE value from 1 to 3	-	-
23	Iriany et al. [65]	GSTAR, SUR, NN	Precipitation data	Comparison GSTAR-SUR-NN for precipitation forecasting	-	-	✓	-	-	RMSE=5.8684	-	MAD=3.8917

Table 3: State-of-the-art from 48 relevant articles (continued).

No	Reference	Models	Dataset	Application	Model Assumptions				Model Performance Analysis			
					MA Component	Exogenous Variable	Hetero. Error	Kriging Method	MAPE	RMSE	MSE	Accuracy
24	Zhang et al. [66]	ST, Kriging	Data for three fixed locations from APDRC (Asia-Pacific Data Research Center)	-	-	-	-	✓	-	-	MSE=0.744	MAE=0.751
25	Cui et al. [67]	ST, Kriging	The daily rainfall data from 1971 to 2010	Rainfall forecasting	-	-	-	✓	-	-	-	$R^2 = 0.77$
26	Takafuji et al. [68]	ST	Dataset of weather patterns from 2008 to 2017 in Patagonia	Spatiotemporal forecasting model for weather patterns in Patagonia	✓	-	-	✓	-	RMSE range 1-15	-	-
27	Li et al. [69]	ST, Kriging	The wind speed, air temperature, and Rainfall data.	Temperature prediction	-	✓	-	✓	-	Space-time prediction=0.640; Pure time forecasting=1.713	-	-
28	Amato et al. [70]	ST -Empirical Orthogonal Functions (EOFs)	Case study of air temperature prediction in a complex Alpine region of Europe.	Forecasting air temperature	-	-	-	✓	-	-	-	-
29	Su et al. [71]	ML, Kriging	NFI datasets	Estimating aboveground biomass	-	-	-	✓	-	RF=52.08%; RFOK=52.05%; RFCK=51.60%	-	RF=24.56; RFOK=23.47; RFCK=22.14
30	Nowak et al. [72]	ST, Kriging	Monthly rainfall data recorded across a network of weather stations in the MDB	Rainfall forecasting	-	-	-	✓	-	The minimum RMSE of 0.9515	-	-
31	Iriany et al. [73]	GSTAR, SUR, NN	Precipitation Data in Malang	Precipitation Forecasting	-	-	✓	-	-	General= 5.3131	-	R^2=0.6177
32	Sulistyono et al. [74]	GSTAR, SUR	Rainfall Data	Rainfall forecasting in agricultural areas	-	-	✓	-	-	Training=5.779; Testing=10.433	-	-
33	Akbar et al. [75]	GSTARMAX	Air Pollutant Data	Forecasting Air Pollutant in Surabaya	✓	✓	✓	-	-	A smaller RMSE Value	-	-
34	Sjahid et al. [76]	GSTARMA	The concentration of PM10 pollutants	Prediction of PM10 pollutant in Surabaya	✓	-	-	-	-	-	-	-
35	Hølleland et al. [77]	ST-GARCH	Dataset of sea surface temperature anomalies	-	✓	-	✓	-	-	-	-	-
36	Xiao et al. [78]	ST-Deep Learning	The daily Optimum Interpolation Sea Surface Temperature produced by the National Oceanic and Atmospheric Administration (NOAA).	Sea surface temperature (SST) field prediction using time-series satellite data	-	-	-	-	Linear SVR = 3.292%; LSTM (1-feature) = 3.193%; LSTM (n-features) = 3.523%; ST-DL = 2.879%	Linear SVR = 0.8660C; LSTM (1-feature) = 0.8500C; LSTM (n-features) = 0.8800C; ST-DL = 0.7590C	-	-
37	Mukhopadhyay et al. [79]	ST, Bayesian Hierarchical Model	Monthly rainfall data	Rainfall forecasting	-	-	-	-	-	RMSE values from 41.10 to 54.78	-	-
38	Mashford et al. [80]	ST, Bayesian Hierarchical Autoregressive Model	The rainfall data from the Ovens catchment	Rainfall forecasting	-	-	-	✓	-	RMSE values from 1 to 11	-	-
39	Venetsanou et al. [81]	ST-Kriging	Precipitation and temperature dataset	Prediction precipitation and temperature	-	-	-	✓	-	-	Prec. MPI=25.7 and 0.; Prec.HadGEM2=30.3 and 304.8; Temp. MPI=8.9 and 2.5; Temp. HadGEM2=6.6 and 14.7	-
40	Abdullah et al. [27]	GSTAR-Kriging	Rainfall Data	Predicting Rainfall Data at Unobserved Locations in West Java	-	-	-	✓	Model I=8.97%; Model II=12.51%; Model III=7.72%	-	-	-

Table 4: State-of-the-art from 48 relevant articles (continued).

No	Reference	Models	Dataset	Application	Model Assumptions				Model Performance Analysis			
					MA Component	Exogenous Variable	Hetero. Error	Kriging Method	MAPE	RMSE	MSE	Accuracy
41	Wang et al. [82]	ST, Kriging	Annual mean temperature (temperatures averaged over a whole year) for 62 years (1950–2011)	Estimation of areal values of near-land-surface temperatures	-	-	-	✓	-	-	-	-
42	You et al. [83]	ST, Bayesian Kriging	Data are collected at the six locations for the two scenarios: cooling and non-cooling on punch in the injection process.	Prediction in squeeze casting	-	-	-	✓	-	-	-	Higher accuracy
43	Martınez et al. [84]	ST, Kriging	Monthly Precipitation of a Hydrogeological Zone in Meta (Colombia)	Precipitation forecasting	✓	-	-	✓	-	-	-	Cross-validation analysis
44	Borrelli et al. [85]	ST, Kriging	The average density of the meteorological stations used to interpolate the grid-based map of rainfall erosivity	Rainfall forecasting	-	-	-	✓	-	-	-	R_{CV}^2 = 0.777 for the cross-validation, R^2= 0.779 for the fitting
45	Nisak [86]	GSTARIMA-SUR	Rain Fall Data in Malang Southern Region Districts	Forecasting rainfall	✓	-	✓	-	-	Tangkilsari=5.263	-	R2=0.6481
46	Chang et al. [87]	ST	Dataset of precipitation observations from 1979 to 2012	A spatiotemporal model to analyze the changes in precipitation patterns over time and space	-	-	-	-	-	-	-	-
47	Carvalho et al. [88]	ST	Dataset of daily rainfall observations from 1961 to 2010 in Brazil	Estimate daily rainfall data	-	-	-	✓	-	-	MSEmod= 19.77; MSEkrig=26.15; MSEcokrig=24.04	-
48	Shu-qin et al. [89]	GWR, Kriging	Climate and Socio-economic variable	Variability of Soil Organic Matter influenced by climate and socio-economic	-	-	-	✓	-	-	-	-

accuracy. However, areas still need to be addressed when creating an adaptive Kriging method that can capture the temporal and spatial changes present in complex climate data. Due to these limitations, there is a possibility for more complex techniques that are adapted to changing trends and non-stationary data.

3.4 The Framework for model integration for climate forecast

3.4.1 The Integration of GSTARIMA Model with Heteroskedastic Error and Kriging Method for Forecasting

After reviewing previous researchers and performing gap analysis, a conceptual integration model of GSTARIMA with heteroskedastic error and the Kriging method is made to answer RQ1. The GSTARIMA model is processed following the Box-Jenkins method, which includes identification, parameter estimation, and diagnostic checking. The initial identification of the GSTARIMA model is to determine the stationarity of the data. If the stationary test results show that the data is not stationary, then a differencing process is carried out until stationary data is obtained. Next, check the order of the model univariately with the ARIMA Model. The model order is received from the results of the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots. Models with the same order are selected for further multivariate and Spatio-Temporal modeling. Regarding Spatial-temporal modeling, a weight matrix shows the diversity of locations. The order of the Spatio-Temporal Model is obtained based on the calculation of the Space-Time Autocorrelation Function (STACF) and Partial Space-Time Autocorrelation Function (STPACF).

Furthermore, parameter estimation for the GSTARI Model is carried out using the Ordinary Least Square (OLS) method. The error generated by the GSTARI Model is re-modeled to obtain the GSTIMA Model using the Maximum Likelihood (MLE) method. The GSTARI Model and GSTIMA Model are combined to produce the GSTARIMA Model. On the other hand, if exogenous variables influence the response variable, it becomes the GSTARIMA-X Model. Furthermore, predictions are made on the testing data for the GSTARIMA Model. The last stage of the model diagnostic check to determine the model error is white noise and homoscedasticity.

The GSTARIMA Model errors with heteroscedasticity errors are re-estimated following the ARCH/GARCH Model to overcome the non-constant variance of the errors. GSTARIMA Model errors are divided into mean equations and variance equations. The mean equation of the GSTARIMA Model error is estimated using the MLE method, and the variance equation is estimated using the GLS method. Integrating the GSTARIMA Model with the ARCH/GARCH Model can minimize the model error. This model is only able to forecast at locations that have observed values.

Regarding climate data, some areas do not have observation stations. The GSTARIMA and ARCH/GARCH models are then integrated with the Kriging method. The Kriging method is proven to forecast phenomena at unobserved locations. Estimated parameters in the GSTARIMA-ARCH Model are input to obtain parameters at unobserved locations. Furthermore, experimental and theoretical semivariogram calculations are carried out to obtain Kriging weights from unobserved locations. The estimated parameters for the unobserved locations are simulated to get the data at the unsampled locations. Finally, the data at unsampled locations are forecasted with the GSTARI-MA-ARCH Model. The integration of the GSTARIMA Model, ARCH/GARCH Model, and Kriging Method can forecast the phenomenon at unobserved locations in the future.

3.4.2 Data Analytics Life Cycle for Climate Forecasting

The conceptual integrated model of GSTARIMA, ARCH/GARCH, and the Kriging Method is then used to forecast climate that meets the criteria of Big Data. Regarding answering RQ2, the modeling flow follows the data analytics life cycle methodology in Figure 3. The initial stage begins with discovery, problem identification, determination of data sources to be processed, and hypotheses that are proven using theorems and mathematical formulas. The next step involves data preparation inputting climate data into the process. Raw climate data is taken at a daily interval and cleaned to eliminate missing value data. Daily data is transformed by aggregating daily data into monthly data. In model planning, mathematical model integration is carried out. At this stage, the theorem that answers the research hypothesis is created. The GSTARIMA model is developed following the Box-Jenkins method. Integrating the GSTARIMA model, ARCH/GARCH and Kriging method requires complex mathematical reasoning, especially in estimating model parameters. The integrated model is used in the model building stage, with the data preparation results inputted. Climate data is divided into training data and testing data. The forecasting results are interpreted by the model obtained. Furthermore, visualization is carried out at the communication results stage, and recommendations are obtained. The last step is to operationalize the results of discoveries in Model development with theorems on mathematical modeling and dissemination.

4 Conclusions

In conclusion, a systematic literature review was conducted to develop the Integration GSTARIMA model with heteroscedastic error and the Kriging method for climate forecasting. A comprehensive search and analysis of the literature was performed to provide a clear understanding of the latest research. This research uses PRISMA and bibliometric methods to analyze the development of this topic. In this paper, the study's results in integrating the GSTARIMA Model with the ARCH/GARCH Model can overcome the problem of non-constant error variance. The GSTARIMA and ARCH models provide an overview of multivariate modeling affected by time, location, and non-stationary data. On the other hand, the GSTARIMA/Spatio-Temporal Model can only forecast at the observed location. Through the integration of the GSTARIMA Model with the Kriging Method, it has been discovered that the prediction of Spatio-Temporal phenomena becomes feasible for unobserved locations in the future. The development of the GSTARIMA Model, ARCH/GARCH, and Kriging Method allows the discovery of theorems in mathematical modeling. The application of this model to climate data uses the data analytics life cycle methodology for more detailed processing and more accurate information.

Acknowledgement

The authors thank the Rector Universitas Padjadjaran, who gave support for the dissemination of research for students and lecturers through the Padjadjaran Excellence Fastrack Scholarship(BUPP) and and Academic Leadership Grant (ALG) with contract number 1549/UN6.3.1/PT.00/2023. The authors also thank the reviewers, Prof. Eddy Hermawan, Prof. Diah Chaerani, and Prof. Sukono, for their valuable review and discussion. This paper is also supported by RISE_SMA project year 2019-2024.

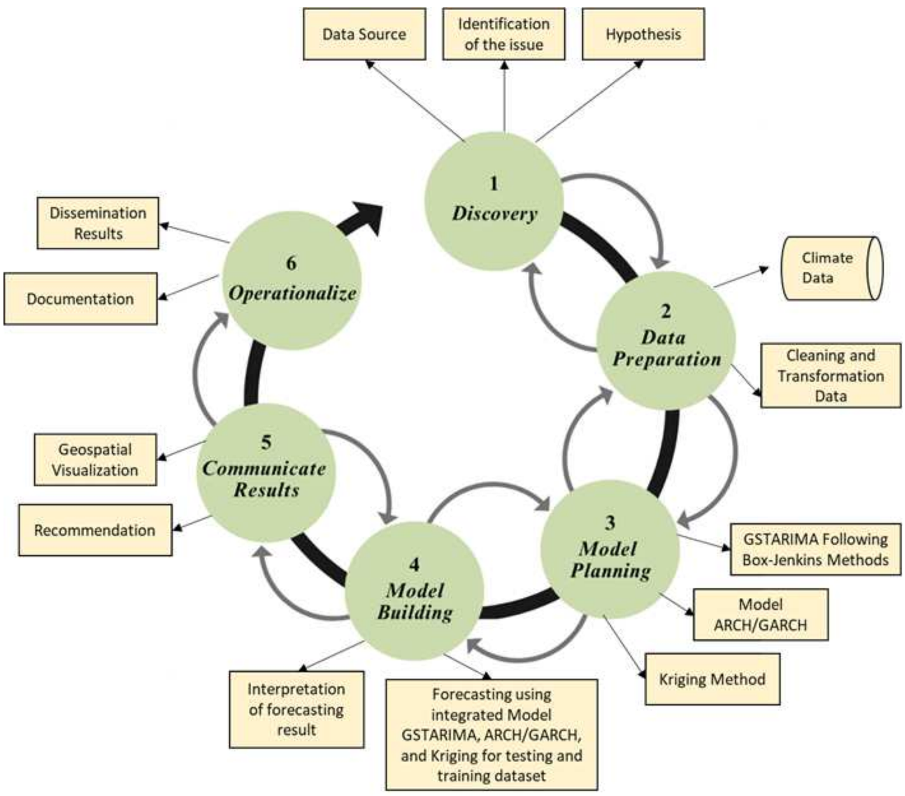


Fig. 3: Data Analytics Life Cycle for Integrated GSTARIMA, ARCH, and Kriging.

References

[1] Werndl, C., On Defining Climate and Climate Change, In Proceedings of the British Journal for the Philosophy of Science, Oxford University Press **67**, 337–364 (2016).

[2] Nunes, L.J.R., Analysis of the Temporal Evolution of Climate Variables Such as Air Temperature and Precipitation at a Local Level: Impacts on the Definition of Strategies for Adaptation to Climate Change, Climate **10**, 154 (2022). doi:10.3390/cli10100154

[3] Pörtner, H.O., Roberts, D.C., Adams, H., Adler, C., Aldunce, P., Ali, E., Ibrahim, Z.Z., Climate Change 2022: Impacts, Adaptation and Vulnerability, IPCC: Switzerland (2022).

[4] Alfarizi, M., Yuniarty Literature Review of Climate Change and Indonesia's SDGs Strategic Issues in a Multidisciplinary Perspective, In Proceedings of the IOP Conference Series: Earth and Environmental Science, Institute of Physics **1105** (2022).

[5] Kelman, I., Linking Disaster Risk Reduction, Climate Change, and the Sustainable Development Goals, Disaster Prev Manag **26**, 254–258 (2017). doi:10.1108/DPM-02-2017-0043

[6] Nerini, F., Francesco, Sovacool, B., Hughes, N., Cozzi, L., Cosgrave, E., Howells, M., Tavoni, M., Tomei, J., Zerriffi, H., Milligan, B., Connecting Climate Action with Other Sustainable Development Goals, Nat. Sustain. **2**, 674-680 (2019).

[7] Kelman, I., Orlowska, J., Upadhyay, H., Stojanov, R., Webersik, C., Simonelli, A.C., Procházka, D.; Němec, D., Does Climate Change Influence People's Migration Decisions in Maldives?, Clim Change **153**, 285–299 (2019). doi:10.1007/s10584-019-02376-y

[8] Khalifa, M., Eltahir, E.A.B., Assessment of Global Sorghum Production, Tolerance, and Climate Risk, Front. Sustain. Food. Syst. **7**, 1184373 (2023). doi:10.3389/fsufs.2023.1184373

[9] Coop, J.D., Parks, S.A., Stevens-Rumann, C.S., Ritter, S.M., Hoffman, C.M., Extreme Fire Spread Events and Area Burned under Recent and Future Climate in the Western USA, Global Ecology and Biogeography **31**, 1949–1959 (2022). doi:10.1111/geb.13496

[10] Ayyad, M., Hajj, M.R., Marsooli, R., Climate Change Impact on Hurricane Storm Surge Hazards in New York/New Jersey Coastlines Using Machine-Learning, NPJ Clim. Atmos. Sci. **6**, 88 (2023). doi:10.1038/s41612-023-00420-4

[11] Hennighausen, H., Liao, Y., Nolte, C., Pollack, A., Flood Insurance Reforms, Housing Market Dynamics, and Adaptation to Climate Risks, J. Hous. Econ., 101953 (2023). doi:10.1016/j.jhe.2023.101953

[12] Klingelhöfer, D., Braun, M., Brüggmann, D., Groneberg, D.A., Heatwaves: Does Global Research Reflect the Growing Threat in the Light of Climate Change?, Global Health **19**, 56 (2023). doi:10.1186/s12992-023-00955-4

[13] Han, W., Zhang, L., Meehl, G.A., Kido, S., Tozuka, T., Li, Y., McPhaden, M.J., Hu, A., Cazenave, A., Rosenbloom, N., et al., Sea Level Extremes and Compounding Marine Heatwaves in Coastal Indonesia, Nat. Commun. **13**, 6410 (2022). doi:10.1038/s41467-022-34003-3

[14] Sidiqi, M., Kasiviswanathan, K.S., Scheytt, T., Devaraj, S., Assessment of Meteorological Drought under the Climate Change in the Kabul River Basin, Afghanistan, Atmosphere (Basel) **14**, 570 (2023). doi:10.3390/atmos14030570

[15] Aksu, H,. Nonstationary Analysis of the Extreme Temperatures in Turkey, Dynamics of Atmospheres and Oceans **95**, 101238 (2021). doi:10.1016/j.dynatmoce.2021.101238

[16] Ray Biswas, R., Rahman, A., Adaptation to Climate Change: A Study on Regional Climate Change Adaptation Policy and Practice Framework, J. Environ. Manage. **336**, 117666 (2023). doi:10.1016/j.jenvman.2023.117666

[17] Hong, T., Malik, J., Krelling, A., O'Brien, W., Sun, K., Lamberts, R., Wei, M., Ten Questions Concerning Thermal Resilience of Buildings and Occupants for Climate Adaptation, Build. Environ. **244**, 110806 (2023). doi:10.1016/j.buildenv.2023.110806

[18] Munandar, D., Ruchjana, B.N., Abdullah, A.S., Pardede, H.F., Literature Review on Integrating Generalized Space-Time Autoregressive Integrated Moving Average (GSTARIMA) and Deep Neural Networks in Machine Learning for Climate Forecasting, Mathematics **11**, 2975 (2023).

[19] Yuan, T., Zhu, J., Wang, W., Lu, J., Wang, X., Li, X., Ren, K., A Space-Time Partial Differential Equation Based Physics-Guided Neural Network for Sea Surface Temperature Prediction, Remote Sens (Basel) **15**, 3498 (2023). doi:10.3390/rs15143498

[20] Li, N., Chen, X., Qiu, J., Li, W., Zhao, B., Spatio-Temporal Characteristics and Trend Prediction of Extreme Precipitation—Taking the Dongjiang River Basin as an Example, Water (Basel) **15**, 2171 (2023). doi:10.3390/w15122171

[21] Pfeifer, P.E., Deutsch, S.J., A Three-Stage Iterative Procedure for Space-Time Modeling, Technometrics **22**, 35-47 (1980).

[22] Borovkova, S.A., Lopuhaä, H.P., Nurani, B., Generalized STAR Model with Experimental Weights, In Proceedings of the Proceedings of the 17th International Workshop on Statistical Modelling, 139–147 (2002).

[23] Min, X., Hu, J., Zhang, Z., Urban Traffic Network Modeling and Short-Term Traffic Flow Forecasting Based on GSTARIMA Model, 13th International IEEE Annual Conference on Intelligent Transportation Systems, 1535–1540 (2010).

[24] Kumar, R.R., Sarkar, K.A., Dhakre, D.S., Bhattacharya, D., A Hybrid Space–Time Modelling Approach for Forecasting Monthly Temperature, Environmental Modeling & Assessment **28**, 317–330 (2022). doi:10.1007/s10666-022-09861-2

[25] Dhaher, G., Shexo, A., Using Kriging Technique to Interpolate and Forecasting Temperatures Spatio-Temporal Data, European Journal of Pure and Applied Mathematics **16**, 373–385 (2023). doi:10.29020/nybg.ejpam.v16i1.4613

[26] Dai, H., Huang, G., Wang, J., Zeng, H., Zhou, F., Spatio-Temporal Characteristics of PM2.5 Concentrations in China Based on Multiple Sources of Data and LUR-GBM during 2016–2021, Int. J. Environ. Res. Public Health **19**, 6292 (2022). doi:10.3390/ijerph19106292

[27] Abdullah, A.S., Matoha, S., Lubis, D.A., Falah, A.N., Jaya, I.G.N.M., Hermawan, E., Ruchjana, B.N., Implementation of Generalized Space Time Autoregressive (GSTAR)-Kriging Model for Predicting Rainfall Data at Unobserved Locations in West Java, Applied Mathematics and Information Sciences **12**, 607–615 (2018). doi:10.18576/amis/120316

[28] Monika, P., Ruchjana, B.N., Abdullah, A.S., GSTARI-X-ARCH Model with Data Mining Approach for Forecasting Climate in West Java, Computation **10**, 204 (2022). doi:10.3390/computation10120204

[29] Dagaeva, M., Garaeva, A., Anikin, I., Makhmutova, A., Minnikhanov, R., Big Spatio-Temporal Data Mining for Emergency Management Information Systems, In Proceedings of the IET Intelligent Transport Systems, Institution of Engineering and Technology **13**, 1649–1657 (2019).

[30] Pfeifer, P.E., Deutsch, S.J., A STARIMA Model-Building Procedure with Application to Description and Regional Forecasting, Transactions of the Institute of British Geographers **5**, 330-349 (1980). doi:10.2307/621846

[31] Pfeifer, P.E., Deutsch, S.J., A Three-Stage Iterative Procedure for Space-Time Modeling, Technometrics **22**, 35–47 (1980).

[32] Di Giacinto, V., A Generalized Space-Time ARMA Model with an Application to Regional Unemployment Analysis in Italy, Int. Reg. Sci. Rev. **29**, 159–198 (2006). doi:10.1177/0160017605279457

[33] Engle, R.F., Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, Econometrica **50**, 987 (1982). doi:10.2307/1912773

[34] Montero, J.-M., Fernandez-Aviles, G., Mateu, J., Spatial and Spatio-Temporal Geostatistical Modeling and Kriging, First Edition, John Wiley & Sons Ltd, (2015).

[35] Dietrich, D., Heller, B., Yang, B., Data Science & Big Data Analytics, John Wiley & Sons Inc., (2015).

[36] Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., et al., The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews, International journal of surgery **88**, 105906 (2021).

[37] Osayande, I., Ogunyemi, O., Gwacham-Anisiobi, U., Olaniran, A., Yaya, S., Banke-Thomas, A., Prevalence, Indications, and Complications of Caesarean Section in Health Facilities across Nigeria: A Systematic Review and Meta-Analysis, Reprod Health **20**, 1-21 (2023).

[38] Saboyá Acosta, L.P., Urbina-Cardona, J.N., Current State of Knowledge of Páramo Amphibians in Colombia: Spatio Temporal Trends and Information Gaps to Be Strengthened for Effective Conservation, Trop. Conserv. Sci. **16**, 19400829231169984 (2023).

[39] Sukono, Juahir, H., Ibrahim, R.A., Saputra, M.P.A., Hidayat, Y., Prihanto, I.G., Application of Compound Poisson Process in Pricing Catastrophe Bonds: A Systematic Literature Review, Mathematics **10**, 2668 (2022).

[40] Ibrahim, R.A., Sukono, Napitupulu, H., Ibrahim, R.I., How to Price Catastrophe Bonds for Sustainable Earthquake

Funding? A Systematic Review of the Pricing Framework, Sustainability **15**, 7705 (2023).

[41] Firdaniza, F., Ruchjana, B.N., Chaerani, D., Radianti, J., Information Diffusion Model in Twitter: A Systematic Literature Review, Information **13**, 13 (2022). doi:10.3390/info13010013.

[42] Gil, M., Wróbel, K., Montewka, J., Goerlandt, F., A Bibliometric Analysis and Systematic Review of Shipboard Decision Support Systems for Accident Prevention, Saf. Sci. **128**, 104717 (2020).

[43] Li, J., Jiang, Y., The Research Trend of Big Data in Education and the Impact of Teacher Psychology on Educational Development During COVID-19: A Systematic Review and Future Perspective, Front. Psychol. **12**, 753388 (2021).

[44] Bravo-Toledo, L., Barreto-Pio, C., López-Herrera, J., Milla-Figueroa, C., Pilco-Nuñez, A., Virú-Vásquez, P., Global Research Trends in Emergy and Wastewater Treatment: A Bibliometric Analysis, Environmental Research, Engineering and Management **79**, 16-36 (2023). doi:10.5755/j01.erem.79.1.30824

[45] Tian, H., Chen, J., A Bibliometric Analysis on Global EHealth, Digit. Health **8**, 20552076221091352 (2022). doi:10.1177/20552076221091352

[46] Mohamed, B., Marzouk, M., Bibliometric Analysis and Visualisation of Heritage Buildings Preservation, Herit. Sci. **11**, 1-20 (2023).

[47] Puica, M., Benth, F.E., A Spatio-Temporal Model for Predicting Wind Speeds in Southern California, Commun Stat Case Stud Data Anal Appl **9**, 321–349 (2023). doi:10.1080/23737484.2023.2217137

[48] Garcıa, J.A., Acero, F.J., Portero, J., A Bayesian Hierarchical Spatio-Temporal Model for Extreme Temperatures in Extremadura (Spain) Simulated by a Regional Climate Model, Clim. Dyn. **61**, 1489–1503 (2023). doi:10.1007/s00382-022-06638-x

[49] Huguenin, C.N., Serafin, K.A., Waylen, P.R., A Spatio-Temporal Analysis of the Role of Climatic Drivers Influencing Extreme Precipitation Events in a Costa Rican Basin, Weather Clim Extrem **42**, 100602 (2023) doi:10.1016/j.wace.2023.100602

[50] Zhang, X., Wang, Y., Wei, L., Jiang, J., Lin, P., Liu, H., Spatiotemporal Networks for ENSO Forecasting with LICOM3 and Remote Sensing Data, Eng. Appl. Artif. Intell. **125**, 106641 (2023). doi:10.1016/j.engappai.2023.106641

[51] Raman, R.K., Bhor, M., Manna, R.K., Samanta, S., Das, B.K., Statistical and Geostatistical Modelling Approach for Spatio-Temporal Assessment of River Water Quality: A Case Study from Lower Stretch of River Ganga, Environ Dev Sustain **25**, 9963-9989 (2023). doi:10.1007/s10668-022-02472-7

[52] Fung, K.F., Chew, K.S., Huang, Y.F., Ahmed, A.N., Teo, F.Y., Ng, J.L., Elshafie, A., Evaluation of Spatial Interpolation Methods and Spatiotemporal Modeling of Rainfall Distribution in Peninsular Malaysia, Ain Shams Engineering Journal **13**, 101571 (2022). doi:10.1016/j.asej.2021.09.001

[53] Mukhaiyar, U., Ramadhani, S., The Generalized STAR Modeling with Heteroscedastic Effects, CAUCHY: Jurnal Matematika Murni dan Aplikasi **7**, 158-172 (2022). doi:10.18860/ca.v7i2.13097

[54] Sofi, S.S., Oseledets, I., A Case Study of Spatiotemporal Forecasting Techniques for Weather Forecasting, ArXiv preprint arXiv:2209, 14782 (2022).

[55] Xiang, L., Guan, J., Xiang, J., Zhang, L., Zhang, F., Spatiotemporal Model Based on Transformer for Bias Correction and Temporal Downscaling of Forecasts, Front. Environ. Sci. **10**, 1039764 (2022). doi:10.3389/fenvs.2022.1039764

[56] Zhao, J., Xu, J., Wang, G., Jin, J., Hu, X., Guo, Y., Li, X., Spatial–Temporal Distribution and Forecasting Model of Precipitation Using Dynamic-Statistical Information Fusion, Journal of Water and Climate Change **13**, 1425–1447 (2022). doi:10.2166/wcc.2022.375

[57] Anshuka, A., Chandra, R., Buzacott, A.J.V., Sanderson, D., van Ogtrop, F.F., Spatio Temporal Hydrological Extreme Forecasting Framework Using LSTM Deep Learning Model, Stochastic Environmental Research and Risk Assessment **36**, 3467–3485 (2022), doi:10.1007/s00477-022-02204-3

[58] Hou, S., Li, W., Liu, T., Zhou, S., Guan, J., Qin, R., Wang, Z., MIMO: A Unified Spatio-Temporal Model for Multi-Scale Sea Surface Temperature Prediction, Remote Sens (Basel) **14**, 2371 (2022). doi:10.3390/rs14102371

[59] Yu, X., Shi, S., Xu, L. A Spatial–Temporal Graph Attention Network Approach for Air Temperature Forecasting, Appl. Soft. Comput. **113**, 107888 (2021). doi:10.1016/j.asoc.2021.107888

[60] Chen, H., Sheng, S., Xu, C.Y., Li, Z., Zhang, W., Wang, S., Guo, S., A Spatiotemporal Estimation Method for Hourly Rainfall Based on F-SVD in the Recommender System, Environmental Modelling and Software **144**, 105148 (2021). doi:10.1016/j.envsoft.2021.105148

[61] Thorson, J.T., Arimitsu, M.L., Barnett, L.A.K., Cheng, W., Eisner, L.B., Haynie, A.C., Hermann, A.J., Holsman, K., Kimmel, D.G., Lomas, M.W., et al., Forecasting Community Reassembly Using Climate-Linked Spatio-Temporal Ecosystem Models, Ecography **44**, 612-625 (2021). doi:10.1111/ecog.05471

[62] Kong, W., Li, H., Yu, C., Xia, J., Kang, Y., Zhang, P., A Deep Spatio-Temporal ForecastingModel for Multi-SiteWeather Prediction Post-Processing, Commun. Comput. Phys. **31**, 131-153 (2021). doi:10.4208/cicp.OA-2020-0158

[63] Kuo, P.-F., Huang, T.-E., Putra, I.G.B., Comparing Kriging Estimators Using Weather Station Data and Local Greenhouse Sensors, Sensors **21**, 1853 (2021). doi:10.3390/s21051853

[64] Ghorbani, M.A., Mahmoud Alilou, S., Javidan, S., Naganna, S.R., Assessment of Spatio-Temporal Variability of Rainfall and Mean Air Temperature over Ardabil Province, Iran, SN. Appl. Sci. **3**, 1-10 (2021). doi:10.1007/s42452-021-04698-y.

[65] Iriany, A., Rosyida, D., Arifin, A., A Comparison of GSTAR-SUR Models and a Hybrid GSTAR-SUR/Neural Network Model on Residuals of Precipitation Forecasting, Commun. Stat. Simul. Comput. **50**, 2782-2792 (2021). doi:10.1080/03610918.2019.1615625

[66] Zhang, R., Yang, S., Wang, Y., Wang, S., Gao, Z., Luo, C., Three-Dimensional Regional Oceanic Element Field Reconstruction with Multiple Underwater Gliders in the Northern South China Sea, Applied Ocean Research **105**, 102405 (2020). doi:10.1016/j.apor.2020.102405

[67] Cui, Y., Pan, C., Liu, C., Luo, M., Guo, Y. Spatiotemporal Variation and Tendency Analysis on Rainfall Erosivity in the Loess Plateau of China, Hydrology Research **51**, 1048-1062 (2020). doi:10.2166/nh.2020.030

[68] Takafuji, E.H. de M., da Rocha, M.M., Manzione, R.L., Spatiotemporal Forecast with Local Temporal Drift Applied to Weather Patterns in Patagonia, SN. Appl. Sci. **2**, 1001 (2020). doi:10.1007/s42452-020-2814-0

[69] Li, S., Griffith, D.A., Shu, H., Temperature Prediction Based on a Space–Time Regression-Kriging Model, J. Appl. Stat. **47**, 1168-1190 (2020). doi:10.1080/02664763.2019.1671962

[70] Amato, F., Guignard, F., Robert, S., Kanevski, M., A Novel Framework for Spatio-Temporal Prediction of Environmental Data Using Deep Learning, Sci. Rep. **10**, 22243 (2020). doi:10.1038/s41598-020-79148-7

[71] Su, H., Shen, W., Wang, J., Ali, A., Li, M., Machine Learning and Geostatistical Approaches for Estimating Aboveground Biomass in Chinese Subtropical Forests, For. Ecosyst. **7**, 64 (2020). doi:10.1186/s40663-020-00276-7

[72] Nowak, G., Welsh, A.H., Improved Prediction for a Spatio-Temporal Model, Environ Ecol Stat **27**, 631-648 (2020), doi:10.1007/s10651-020-00447-3

[73] Iriany, A., Rosyida, D., Sulistyono, A.D., Ruchjana, B.N., Precipitation Forecasting Using Neural Network Model Approach, In Proceedings of the IOP Conference Series: Earth and Environmental Science, Institute of Physics Publishing **458**, 012020 (2020).

[74] Sulistyono, A.D., Hartawati, Iriany, A., Suryawardhani, N.W., Iriany, A., Rainfall Forecasting in Agricultural Areas Using GSTAR-SUR Model, In Proceedings of the IOP Conference Series: Earth and Environmental Science, Institute of Physics Publishing **458**, 012041 (2020).

[75] Akbar, M.S., Setiawan, Suhartono, Ruchjana, B.N., Prastyo, D.D., Muhaimin, A., Setyowati, E., A Generalized Space-Time Autoregressive Moving Average (GSTARMA) Model for Forecasting Air Pollutant in Surabaya, J. Phys. Conf. Ser. **1490**, 12022 (2020). doi:10.1088/1742-6596/1490/1/012022

[76] Sjahid, M., Akbar, Setiawan, Suhartono, Ruchjana, B.N., Prastyo, D.D., Prediction of PM10 Pollutant in Surabaya Using Generalized Space-Time Autoregressive Moving Average, Investigacion Operacional **41**, 990-998 (2020).

[77] Hølleland, S., Karlsen, H.A., A Stationary Spatio-Temporal GARCH Model, J. Time Ser. Anal. **41**, 177-209 (2019). doi:10.1111/jtsa.12498

[78] Xiao, C., Chen, N., Hu, C., Wang, K., Xu, Z., Cai, Y., Xu, L., Chen, Z., Gong, J., A Spatiotemporal Deep Learning Model for Sea Surface Temperature Field Prediction Using Time-Series Satellite Data, Environmental Modelling and Software **120**, 104502 (2019). doi:10.1016/j.envsoft.2019.104502

[79] Mukhopadhyay, S., Ogutu, J.O., Bartzke, G., Dublin, H.T., Piepho, H.P., Modelling Spatio-Temporal Variation in Sparse Rainfall Data Using a Hierarchical Bayesian Regression Model, J. Agric. Biol. Environ. Stat. **24**, 369-393 (2019). doi:10.1007/s13253-019-00357-3

[80] Mashford, J., Song, Y., Wang, Q.J., Robertson, D., A Bayesian Hierarchical Spatio-Temporal Rainfall Model, J. Appl. Stat. **46**, 217–229 (2019). doi:10.1080/02664763.2018.1473347

[81] Venetsanou, P., Anagnostopoulou, C., Loukas, A., Lazoglou, G., Voudouris, P., Minimizing the Uncertainties of RCMs Climate Data by Using Spatio-Temporal Geostatistical Modeling, Earth Sci. Inform. **12**, 183-196 (2018). doi:10.1007/s12145-018-0361-7

[82] Wang, H., Pardo-Igúzquiza, E., Dowd, P.A., Yang, Y., Optimal Estimation of Areal Values of Near-Land-Surface Temperatures for Testing Global and Local Spatio-Temporal Trends, Comput. Geosci. **106**, 109-117 (2017), doi:10.1016/j.cageo.2017.06.002

[83] You, D., Jiang, X., Cheng, X., Wang, X., Bayesian Kriging Modeling for Spatiotemporal Prediction in Squeeze Casting, International Journal of Advanced Manufacturing Technology **89**, 355-369 (2017). doi:10.1007/s00170-016-9078-2

[84] Martınez, W.A., Melo, C.E., Melo, O.O., Median Polish Kriging for Space–Time Analysis of Precipitation, Spat. Stat. **19**, 1-20 (2017). doi:10.1016/j.spasta.2016.10.003

[85] Borrelli, P., Diodato, N., Panagos, P., Rainfall Erosivity in Italy: A National Scale Spatio-Temporal Assessment, Int. J. Digit Earth **9**, 835-850 (2016), doi:10.1080/17538947.2016.1148203

[86] Nisak, S.C., Seemingly Unrelated Regression Approach for GSTARIMA Model to Forecast Rain Fall Data in Malang Southern Region Districts, CAUCHY **4**, 57-64 (2016). doi:10.18860/ca.v4i2.3488

[87] Chang, W., Stein, M.L., Wang, J., Kotamarthi, V.R., Moyer, E.J., Changes in Spatiotemporal Precipitation Patterns in Changing Climate Conditions, J. Clim. **29**, 8355-8376 (2016). doi:10.1175/JCLI-D-15-0844.s1

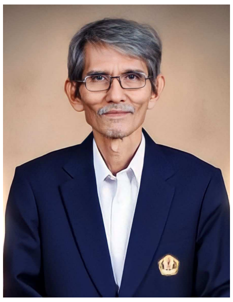
[88] De Carvalho, J.R.P., Nakai, A.M., Monteiro, J.E.B.A., Spatio-Temporal Modeling of Data Imputation for Daily Rainfall Series in Homogeneous Zones, Revista Brasileira de Meteorologia **31**, 196-201 (2016), doi:10.1590/0102-778631220150025

[89] Shi, S.Q., Cao, Q.W., Yao, Y.M., Tang, H.J., Peng, Y., Wu, W.B., Xu., H.Z., Jia, L., Li, Z.G., Influence of Climate and Socio-Economic Factors on the Spatio-Temporal Variability of Soil Organic Matter: A Case Study of Central Heilongjiang Province, China, J. Integr. Agric. **13**, 1486-1500 (2014). doi:10.1016/s2095-3119(14)60815-7

[90] Pramoedyo, H., Ashari, A., Fadliana, A., Forecasting and Mapping Coffee Borer Beetle Attacks Using GSTAR-SUR Kriging and GSTARX-SUR Kriging Models, ComTech: Computer, Mathematics and Engineering Applications **11**, 65-73 (2020). doi:10.21512/comtech.v11i2.6389

Putri Monika, received B.Sc. degree in Mathematics,Faculty Mathematics and Natural Sciences, Universitas Padjadjaran, Indonesia in 2021, Magister in Mathematics,Faculty Mathematics and Natural Sciences, Universitas Padjadjaran, Indonesia in 2022. Currently, she is a student in doctor program of Mathematics, Faculty Mathematics and Natural Sciences, Universitas Padjadjaran, Indonesia. Her research interest include Spatio Temporal Modelling, Time Series Analysis, Stochastics Processes, and Big Data Analytics.

Budi Nurani Ruchjana, received B.Sc. degree in Mathematics from Universitas Padjadjaran, Indonesia in 1987, Magister in Applied Statistics from Institut Pertanian Bogor, Indonesia in 1992, and Doctor of Mathematics and Natural Sciences from Institut Teknologi Bandung, Indonesia, in 2002. Currently, she is a full professor at Department of Mathematics, Universitas Padjadjaran, Indonesia. Her research interest include Spatio Temporal Modeling, Stochastics Processes, Time Series Analysis, Spatial Analysis, Geostatistics and Ethnomathematics.



Atje Setiawan Abdullah, received B.Sc. degree in Mathematics from Universitas Padjadjaran, Indonesia in 1985, Magister in Management and Industrial Technology from Institut Teknologi Bandung, Indonesia in 1989, Magister and Doctor of Computer Science from Universitas Gadjah Mada, Indonesia in 2004 and 2009, respectively. Currently, he is a full professor at Department of Computer Science, Universitas Padjadjaran, Indonesia. His research interest include Spatial Data Mining, Management and Information Systems, Decision Support System, and Ethno-informatics.

Rahmat Budiarto, received B.Sc. degree in Mathematics from Bandung Institute of Technology, Indonesia in 1986, M.Eng. and Dr.Eng. in Computer Science from Nagoya Institute of Technology, Japan in 1995 and 1998, respectively. Currently, he is a full professor at Dept. of Computer Science, Albaha University, Saudi Arabia. His research interests include intelligent systems, brain modeling, IPv6, network security, Wireless sensor networks, and MANETs.