A Systematic Review of Quantile Regression in Varying Coefficient Models for Longitudinal Data

B. Tantular, BN. Ruchjana, Y. Andriyana, and A. Verhasselt

***Abstract*— Varying coefficient models have some regression coefficients allowed to vary as smooth functions of other variables. Applying varying coefficient models for longitudinal data determined the effect of different covariates between the time variable called the time-varying coefficient model. Several researchers used this model with a different approach. Quantile regression is a technique that uses P-splines as an estimation procedure. This research conducts a systematic literature review of the peer-reviewed papers on varying coefficient models with quantile objective function inspired by Hastie and Tibshirani. Furthermore, it shows a comprehensive bibliometric analysis involving a co-authorships network of the productive authors as well as a bibliometric map with the clustered term. The varying coefficient model and quantile regression are used to exposes a thematic analysis. Finally, the varying coefficient model, which includes time and spatial effects, is an interesting topic for further research.**

***Index Terms*—longitudinal data, p-splines, quantile regression, systematic literature review, varying coefficient models**

I. Introduction

VARYING coefficient models, introduced by Hastie and Tibshirani in 1993, showed that some regression coefficients vary as smooth functions of other variables [1]. The varying coefficient models have been used by researchers in various data structures, including longitudinal structures. Therefore, the variables effect among the time variables (t), called time-varying coefficient models, is determined using this model for longitudinal data. Several researchers have used this model, including [2] which applied the two-step estimation method, [3] used the expansion of basis function and variable selection, and [4] combined the P-splines method with non-negative garrote variable selection as the proposed variable selection technique.

A general form of a time-varying coefficient model for longitudinal data is

$$Y(t_{ij}) = \beta_0 + \sum_{p=1}^{P} \beta_p(t_{ij}) X^{(p)}(t_{ij}) + \varepsilon(t_{ij}) \tag{1}$$

where $i = 1, 2, 3, \ldots, n$ and $j = 1, 2, 3, \ldots, N_i$, $Y(t)$ is the response variable, $X^{(p)}(t)$ is the p-th covariate and $\beta_k(t)$ is the k-th parameter at time t where t_{ij} is the change in coefficient $X^{(p)}$ through the unspecified function $\beta_k(.)$. In this case, n is the number of subjects measured, while N_i is the number of repeated measurements of subject i. N_i indicates that this model can be used in a balance or unbalanced design. Also, this model assumes that the model's error $\varepsilon(t_{ij})$ is independent of $X(t)$.

Mean regression is used in various models due to the ease of parameter estimation. However, it requires many assumptions which limit its use. This method should not be implemented when the data contains any outlier or leverage. Koenker and Basset [5] introduced quantile regression, which is a generalization of median regression. This method is robust to the outlier and provides more information was obtained about the distribution of responses. Furthermore, it is used to determine the effect of the response distribution's covariates of location, scale, and shape. Several researchers used different estimation procedures of quantile regression in varying coefficient models, such as [6] apply a two-step estimation procedure and [7] used the basis function approach, and [8] who proposed P-splines quantile objective functions. Andriyana and Gijbels [9] developed an extended model of P-splines quantile regression in varying coefficient models for simple heteroscedastic error. The model proposed by [10] includes methods to avoid crossing the conditional quantile estimators.

Median regression can be generalized to quantile regression, which describes the response and explanatory variables' relationships at the centralized size such as mean or median, as well as various response variable quantiles [11]. According to [8], the unconditional quantile is applied to that of univariate data, while the conditional quantile is the quantile of the regression models.

Let Y is a response random variable, and $X^{(1)}, \ldots, X^{(p)}$ are explanatory variables contained in model (1). Therefore conditional quantile regression of model (1) is

$$q_\tau(Y(t) \mid X(t), t) = \mathbf{X}^T(t)\boldsymbol{\beta}^\tau(t) \tag{2}$$

where $\boldsymbol{\beta}^\tau = (\beta_0^\tau, \beta_1, \ldots, \beta_p)^T$ and $\beta_0^\tau = \beta_0 + F_e^{-1}(\tau)$. In this case, $F_\varepsilon^{-1}(\tau) = \inf\{u : F_\varepsilon(u) \geq \tau\}$ for $0 \leq \tau_1 \leq \tau_2 \leq 1$ applies $F_\varepsilon^{-1}(\tau_1) \leq F_\varepsilon^{-1}(\tau_2)$ where τ-*th* is a quantile value of

Manuscript received February 05, 2022; revised October 03, 2022. This work was supported in part by Universitas Padjadjaran through research grant No: 5884/UN6.D/LT/2019

*B. Tantular is a PhD student in Department of Mathematics, Universitas Padjadjaran, West Java, 45363, Indonesia, e-mail: (bertho@unpad.ac.id).

BN. Ruchjana is a Professor of Statistics, Universitas Padjadjaran, West Java, 45363, Indonesia, e-mail: (budi.nurani@unpad.ac.id).

Y. Andriyana is a lecturer in Department of Statistics Universitas Padjadjaran, West Java, 45363, Indonesia, e-mail: (y.andriyana@unpad.ac.id).

A. Verhasselt is a Professor of Statistics, Hasselt University, Martelarenlaan 42, BE3500 Hasselt, Belgium, e-mail: (anneleen.verhasselt@uhasselt.be).

error ε. This error term is assumed independent to **X**. In varying coefficient models, $\beta_k(t_{ij})$ has a function

$$\beta_k(t_{ij}) \approx \sum_{l=1}^{m_k} \alpha_{kl} B_{kl}(t_{ij}; v_k) \tag{3}$$

where k = 1, ..., p. This $\beta_k(t_{ij})$ is a normalized B-splines approximation with degree v_k. $B_{kl}(t_{ij};v_k)$ is a function of basis B-splines with degree v_k, at $l = 1, \ldots, m_k$, where $m_k = u_k + v_k$ and $u_k + 1$ is an equidistance knots for *k-th* component.

Coefficients $\alpha = (\alpha_0^T, \ldots, \alpha_p^T)^T$, where $\alpha_k^T = (\alpha_{k1}, \ldots, \alpha_{kmk})^T$, are unknown so it must be estimated by minimizing the following objective function

$$S(\alpha) = \sum_{i=1}^{N} \frac{1}{N_i} \sum_{j=1}^{N_i} \rho_\tau \left(Y_{ij} - \sum_{k=0}^{P} \sum_{l=1}^{m_k} \alpha_{kl} B_{kl}(t_{ij}; v_k) X_{ij}^{(k)} \right) + \sum_{k=0}^{P} \sum_{l=d_k+1}^{m_k} \lambda_k \left| \Delta_k^d \alpha_{kl} \right| \tag{4}$$

where $\lambda_k > 0$ is k-th smoothing parameter that controls the penalty and goodness-of-fit. Δ_k^d is *d-th* order differencing operator. In addition, a penalty was added in B-splines to avoid overfitting, allowing the maintenance of many knots while limiting the effect. According to [12], the combination of B-splines and penalties approach is called P-splines.

Quantile objective function (4) contain $\rho_\tau(.)$ called *the pinball loss function* or *check-function.* This function is non-differentiable, hence cannot be minimized by ordinary methods [5]. The linear programming problem is used to minimize or maximize the function with constraints. Also, both equality and inequality constraints are used in linear programming. Interior point methods can be used to minimizing quantile objective function (4) on linear programming [13], [14].

This research focuses on scientific articles about varying coefficient models associated with quantile regression and P-splines estimation procedure on longitudinal data. A systematic literature review method of all peer-reviewed papers that contain these topics was presented. Additionally, the results data analysis and discussion were provided. This literature review will be useful in selecting new topics for further research.

II. METHODS

The method used is a systematic literature review containing several steps. The first step is topic formulation, which involves defining keywords relevant to the topic. Furthermore, the second step is a study design, where the database sources are determined. The next step is sampling and data collection, including searching, saving, and merging procedures. The following step is data analysis, where the right tools to analyze the data are determined. Then, the last step is reporting. However, four conditions must be met in this method, including depth and rigor, replicability, usability, and helpful format [15].

As previously stated, varying coefficient models are interesting to analyze, especially on longitudinal data. The basis of the regression model and its estimation procedure are keywords associated to this model. In this context, the specified keywords are varying coefficient models with additional longitudinal data, quantile regression, and P-splines.

There are several sources of scientific article databases, a few can be obtained for free, including Google scholar, Dimensions, and Science Direct. Google scholar is a great database but contains articles and books, proceedings, thesis, or dissertations. However, Dimensions has a filtering method for some conditions despite having an extensive database, while Science direct database only contains articles.

The three database sources use a similar search method with a different layout. Furthermore, searching methods should use quotation marks, the conjunction OR for a keyword with similar meaning, and the conjunction AND to add another keyword. Parentheses are added for a group of keywords. Searching is done sequentially by adding a word or group of keywords and filtering by publication year and type. The full data or title and abstract options can be chooses in search setting. An alternative to obtaining raw citations database was Publish or Perish, the software that retrieves and analyzes academic citations using data sources such as Google Scholar [16]. Generally, this citations database contains information such as title, authors, publication date, keywords, publisher and issue number.

A searching technique was carried out using the keywords mentioned above of English-written international peer-reviewed journals [16]. Furthermore, code A was used for the group of keywords “longitudinal data” AND ("varying coefficient model" OR "varying coefficient models"), code B for ("quantile regression" OR "quantile regressions"), and code C for ("p-splines" OR "p-spline" OR "penalized spline" OR "penalized splines"). Subsequently, code D for adding C into B, and code E for adding D into A. Table 1 showed the number of papers found from a sequential search.

There are several options to save the search result database, but only Refman and BibTex formats exist in all sources. The Bibtex format is used due to the ease to merge by text editor or reference manager software. Meanwhile, JabRef is the reference manager software used to merge and manage the reference database. The merged database may contain similar data or irrelevant topics. Therefore, the duplicate articles are removed or the contents of the articles are merged, and the non-articles database should be removed. The final database is saved in comma-separated value (CSV) format that matches Dimensions, which can be used in the analysis software.

The data analysis procedure consists of a summary of publications, citation analysis, top of journal and publisher, authorship and co-authorship, and term analysis of title and abstract. Additionally, the open-source R statistical software was used with the bibliometrix package [17].

This package can perform a comprehensive bibliometric analysis of the article database based on sources, authors, and terms. In addition, trends can be visualized and analyzed in the form of bibliometric mapping using VOSViewer software [18]. The results of the analysis were presented in tables and graphs or map, and the details were explained in the subsequent sections.

TABLE I
NUMBER OF PAPES IN DATABASE SEARCHING

Code	Keywords	Source			Total
		Google Scholar	Dimensions	Science Direct	
A	"longitudinal data" AND ("varying coefficient model" OR "varying coefficient models")	1,850	1,269	113	3,232
B	("quantile regression" OR "quantile regressions")	17,000	35,768	6,934	59,702
C	("P-splines" OR "P-spline" OR "penalized spline" OR "penalized splines")	11,400	6,093	1,151	18,644
D	A AND B	428	351	28	807
E	**D AND C**	**94**	**71**	**2**	**167**

III. RESULT AND ANALYSIS

A. *Systematic Literature Review Process*

The systematic literature review process is divided into four stages. The first is the identification stage, where specific keywords applied in the search is defined, to determine the topic formulation. Then the second is the searching stage, which includes the initial and advance search. The initial A and B keywords search using Google Scholar, Dimensions, and Science Direct generates 1,516 documents. However, the addition of keywords C in advance produces 167 documents. The next stage is filtering, which involves the removal of non-articles documents, and the merging of documents with similar records. After filtering, the documents in the database reduce to 86. The last stage is the selection which is divided into two parts, including selection by reading the title and abstract and selection by reading the full-text content. This is carried out manually to obtain a list of final papers relevant to the topics. Fig. 1 shows the detail of this process and the number of documents from each.

B. *Bibliometric Analysis*

Bibliometric analysis was conducted using a filtered database with 86 potential papers. The data was analyzed using biblioshiny as a web interface of the bibliometrix R package. This is an easy-to-use web application that analyzes bibliometric data in several formats (Scopus, Web of Science, PubMed, Dimensions). Therefore, the Dimensions format was used, containing information such as title, authors, publication date, keywords, publisher, volume, issue number, and citation.

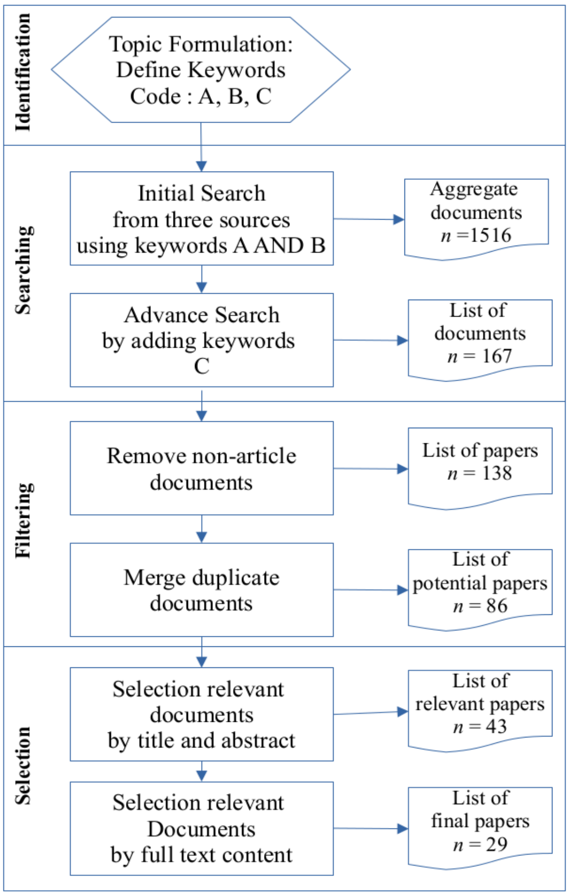

Fig. 1. Systematic literature review process.

Summary of publications

This describes a summary of the research on quantile regression in varying coefficient models for longitudinal data. The result shows 86 papers from 49 publishers, with an average of 7.82 per year and 2.76 average citations per paper. Also, there are 176 authors, where 4 are single authors.

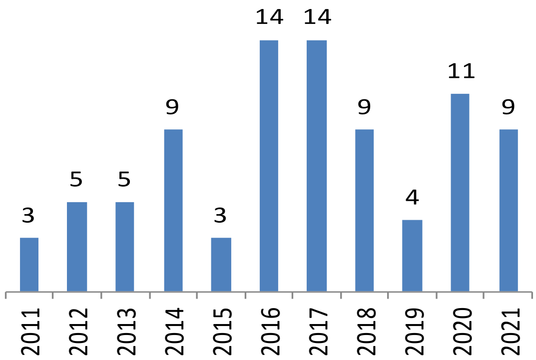

Fig. 2. The number of papers on varying coefficient models, quantile regression and P-splines in three database published between 2011 until 2021

Fig. 2 shows the number of papers in the last ten years (2011-2021, 2021 until June). There has been an increase in the number of papers published until 2017. Subsequently, a decrease was occurred in 2018, then increase again in the last years.

Several journals published articles related to varying coefficient models, quantile regression, and P-splines in longitudinal data. There are 49 journals in this database in total, and Table 2 showed the top 15 journals on these topics.

TABLE 2
TOP 15 JOURNAL PUBLISHING THE MOST PAPERS ON VARYING COEFFICIENT MODEL, QUANTILE REGRESSION AND P-SPLINES IN LONGITUDINAL DATA

Journal	Papers
Computational Statistics & Data Analysis	6
Journal of Multivariate Analysis	5
Annals of The Institute of Statistical Mathematics	4
Asta AdvancesiIn Statistical Analysis	4
Electronic Journal of Statistics	4
Statistical Papers	4
Journal of Nonparametric Statistics	3
Journal of Systems Science and Complexity	3
Journal of The American Statistical Association	3
Statistical Modelling	3
Test	3
Bernoulli	2
Journal of Applied Statistics	2
Journal of Physics Conference Series	2
Journal of Statistical Planning And Inference	2

The database presented the most cited papers with relevant topics on varying coefficient models, quantile regression, and P-splines. Furthermore, the first paper on varying coefficient model by [1] has been cited more than 2282 and 889 times on Google Scholar and Dimensions, respectively Table 3 depicts the most cited paper title and abstract containing keywords varying coefficient model and quantile regression in the last ten years.

TABLE 3
THE MOST CITED PAPERS THAT TITLE AND ABSTRACT CONTAIN KEYWORDS VARYING COEFFICIENT MODEL AND QUANTILE REGRESSION

Journal	Authors	Cited
Variable selection of varying coefficient models in quantile regression	Noh, Hohsuk; Chung, Kwanghun; Van Keilegom, Ingrid	26
P-splines quantile regression estimation in varying coefficient models	Andriyana, Y.; Gijbels, I.; Verhasselt, A.	16
Quantile regression in varying-coefficient models: non-crossing quantile curves and heteroscedasticity	Andriyana, Y.; Gijbels, I.; Verhasselt, A.	9
Quantile regression in heteroscedastic varying coefficient models	Andriyana, Y.; Gijbels, I.	5
Identification and estimation in quantile varying-coefficient models with unknown link function	Yue, Lili; Li, Gaorong; Lian, Heng	3

Most productive authors

This section describes network authorships and bibliometric mapping on varying coefficient models, quantile regression, and P-splines. Furthermore, the R package bibliometrix was used to analyze authorships. About 176 authors have written papers on these topics in total, where 4 were written by single authors. Fig. 3 illustrated the number of authors per paper in full detail.

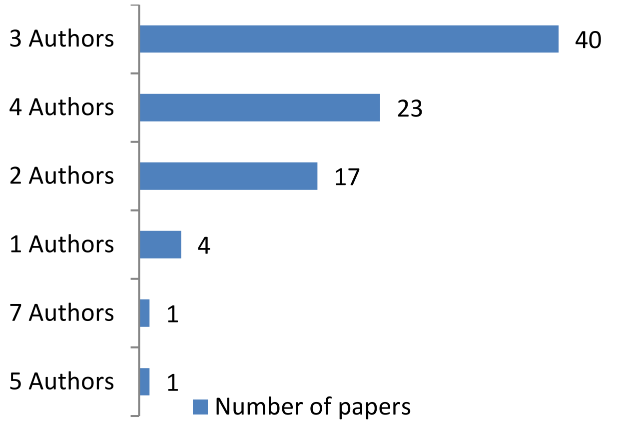

Fig. 3. Number of authors per papers on varying coefficient models, quantile regression and P-splines in three databases published between 2011 until 2021

Some authors have written more than one paper on these topics. Table 4 shows the most productive authors who publish several papers on varying coefficient models, quantile regression, and P-splines in longitudinal data. The second and third column depicts the number of papers and the impact factor of authors, respectively,

TABLE 4
THE MOST PRODUCTIVE AUTHORS WHO WRITE PAPERS ON VARYING COEFFICIENT MODEL, QUANTILE REGRESSION AND P-SPLINES IN LONGITUDINAL DATA

Authors	Number of Papers	H-Index
Lian H	10	3
Zhao W	7	2
Gijbels I	6	2
Andriyana Y	5	1
Verhasselt A	5	2
Zhang J	4	2
Kneib T	4	2
Li G	3	2
Ma S	3	2
Greven S	3	1

These top ten authors produced more than 61% of papers from all papers in this database in last decade.

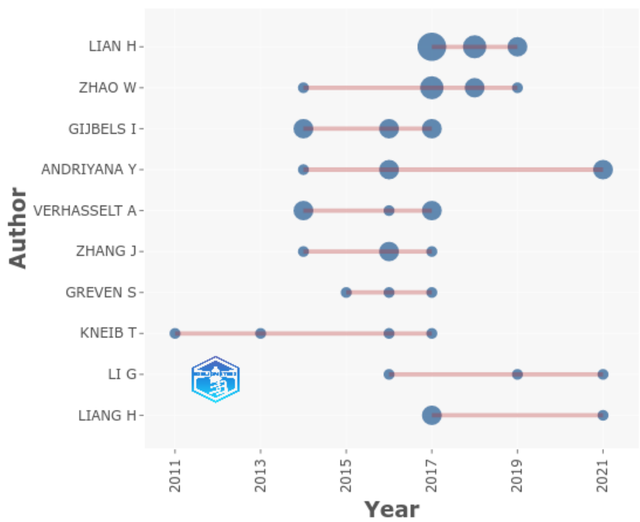

Fig. 4. Authors production over time on varying coefficient models, quantile regression and P-splines in three databases published between 2011 until 2021

Fig. 4 demonstrated production papers of the top ten productive authors working with varying coefficient models, quantile regression, and P-splines in longitudinal data. The blue circle represents the number of papers, while the larger circle indicates more papers. Meanwhile, the darker circle shows the significant total citation index per year.

Network of co-authorships

Co-authorships network on the varying coefficient model, quantile regression, and P-splines was created by obtaining bibliographic data from the main authors related to their co-authors as shown in Fig. 5. There are 8 clusters of collaboration authors, indicated by different colors with connected lines. This is due to specific keywords in the search process hence every cluster has similar themes in its papers.

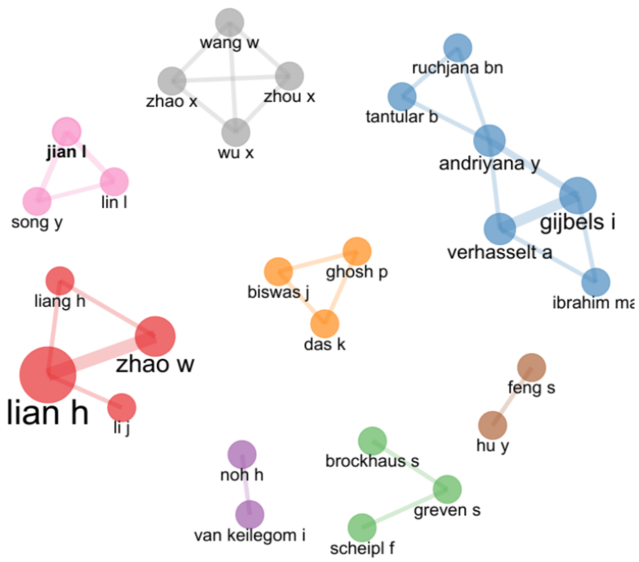

Fig. 5. Collaborative network between the 26 authors working on varying coefficient models, quantile regression and P-splines in three databases published between 2011 until 2021

A cluster map is to build a co-authorships network using Bibliometric coupling of abstracts terms (Fig. 6). This map shows a co-authorships network with clusters of abstract terms that are similar (Fig. 7).

Fig. 6 demonstrated that the colors of the circle indicate 4 clusters. The clusters contain three prominent abstract terms related to the co-authorships network as shown in Fig. 7.

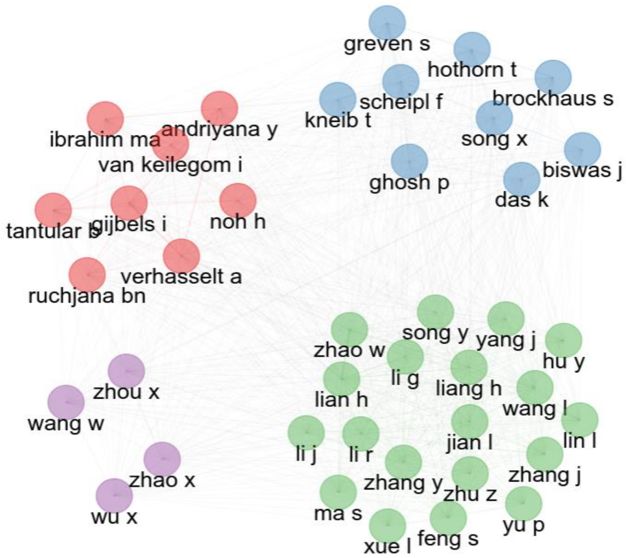

Fig. 7. Co-authorships network with clusters by authors on varying coefficient models, quantile regression and P-splines

Furthermore, networks with the same colors are linked to a cluster of author coupling, and a connected thin line indicates the network between authors.

For example, the red cluster has an abstract containing varying coefficient models, longitudinal data structure, and quantile objective function. These abstract terms are observed in the papers of 8 authors in the red cluster.

C. *Research theme mapping*

Topics within the quantile regression and varying coefficient models are highlighted in this section. The VOSviewer software creates a network of co-occurrence terms by the title and abstract fields, and terms with similar meaning are grouped using VOSviewer thesaurus. The co-occurrences terms in this network are indicated by the distance between two terms. Also, the VosViewer network creates several clusters, shown in different colors, where each is represented by one color.

Theme mapping and clustering

Fig. 8 showed that the co-occurrence network of the terms used in quantile regression and varying-coefficient models literature. The analysis results show that there are 5 clusters with at least 5 members, indicated by the colors red, green, blue, yellow, and purple.

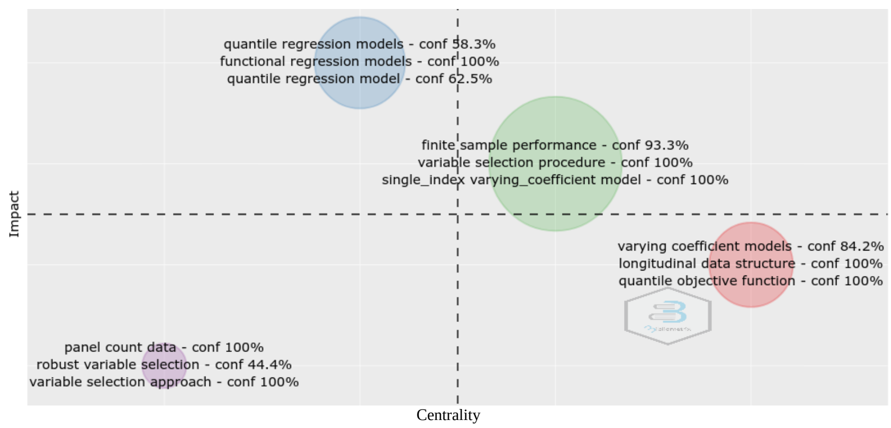

Fig. 6. Clusters by authors coupling on varying coefficient models, quantile regression and P-splines

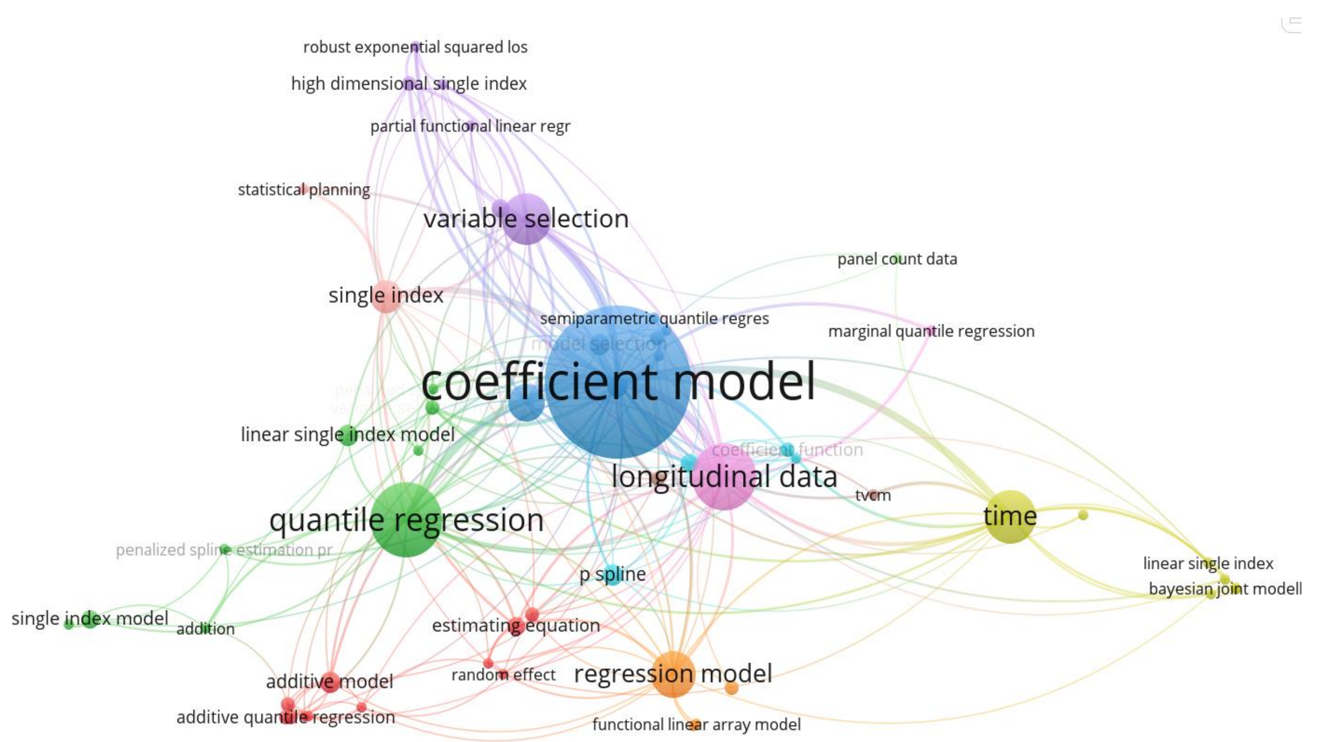

Fig. 8. Co-occurance network of terms on varying coefficient models, quantile regression and P-splines

TABLE 5
CLUSTER MEMBERSHIP OF TITLE AND ABSTRACT TERMS RELATED TO VARYING COEFFICIENT MODEL, QUANTILE REGRESSION AND P-SPLINES IN LONGITUDINAL DATA

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
additive model	addition	coefficient model	bayesian joint model	high dimensional single index
additive quantile regression	linear single index model	linear model	linear single index model	modal regression
estimating equation	M-estimator	model selection	method-ological review	partial functional linear regression model
knot	penalized spline estimation	regression coefficient	parameter estimation	robust
penalize spline	quantile regression	semi - parametric quantile regression	quantile single index model	robust exponential square loss
random effect	semi - parametric estimation	smooth function	statistical inference	variable selection
simulation	single index model	varying coefficient model	time	
smoothing parameter	Variable selection method			
smoothness				
spline smoothing				
unknown link function				

Table 5 described the keywords of the 5 clusters that have at least 5 members.

According to Table 5 the opportunity to do research relevant to this topic is widely opened based on the keywords from each cluster, it can be seen that. The associations of a selected keyword with others are obtained by hover the cursor over the word selected. Suppose the quantile regression keywords are selected from cluster 4. Subsequently, the keywords are directly related to the quantile regression. Fig. 9 illustrated the result of these network terms.

Also, keywords associated with quantile regression are selected to observe the association with other words. The term coefficient model has a similar meaning with varying coefficient models. Table 5 showed that the coefficient model and varying coefficient model are in the same cluster. In this case, due to the selection of coefficient model term in cluster 3, the keywords directly related to the (varying) coefficient model are as shown in Fig. 10.

The R package bibliometrix software can analyze the conceptual structure of themes, for instance, thematic map and evolution. Furthermore, the associated keywords (three words terms) of quantile regression and the varying coefficient model were produced using the thematic map as shown in Fig. 11, divided into four quadrants. The coloring circles clusters represent the size of the circles, which states the number of papers. Every circle shows only a maximum of three associated keywords, though it contains more than three.

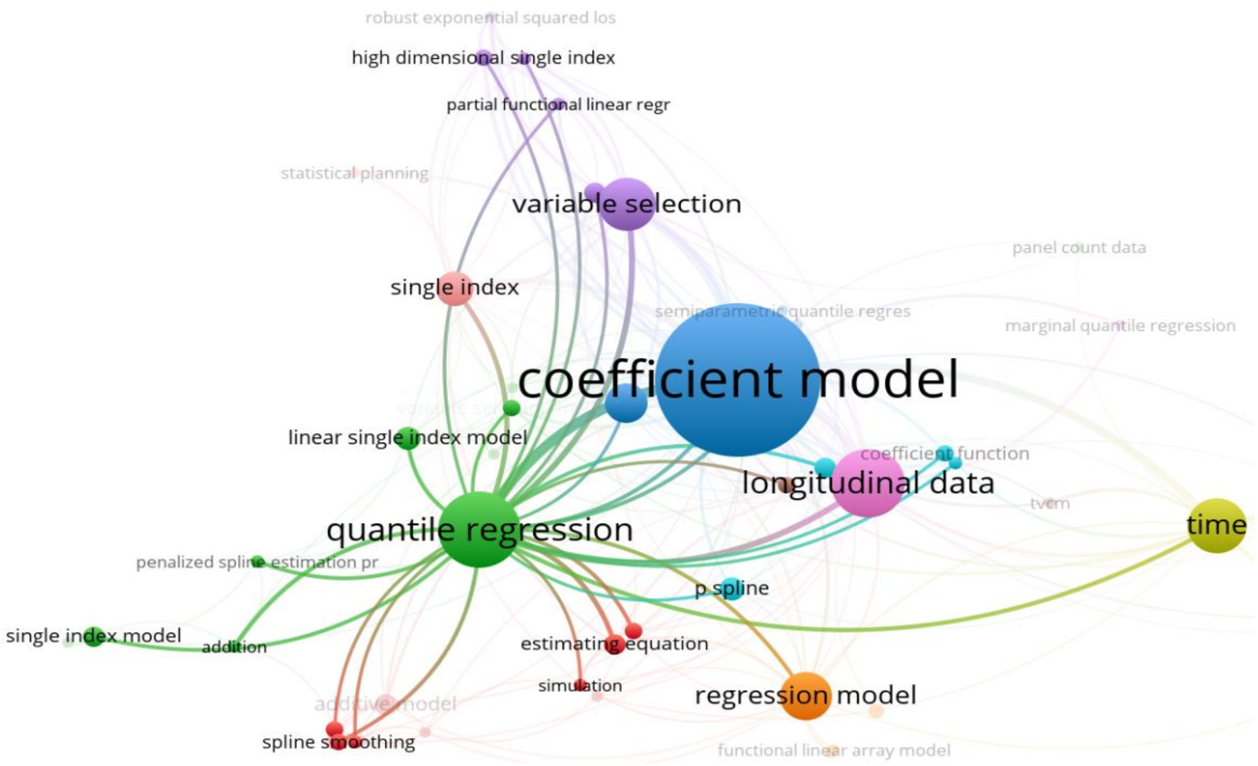

Fig. 9. Quantile regression title and abstract network terms

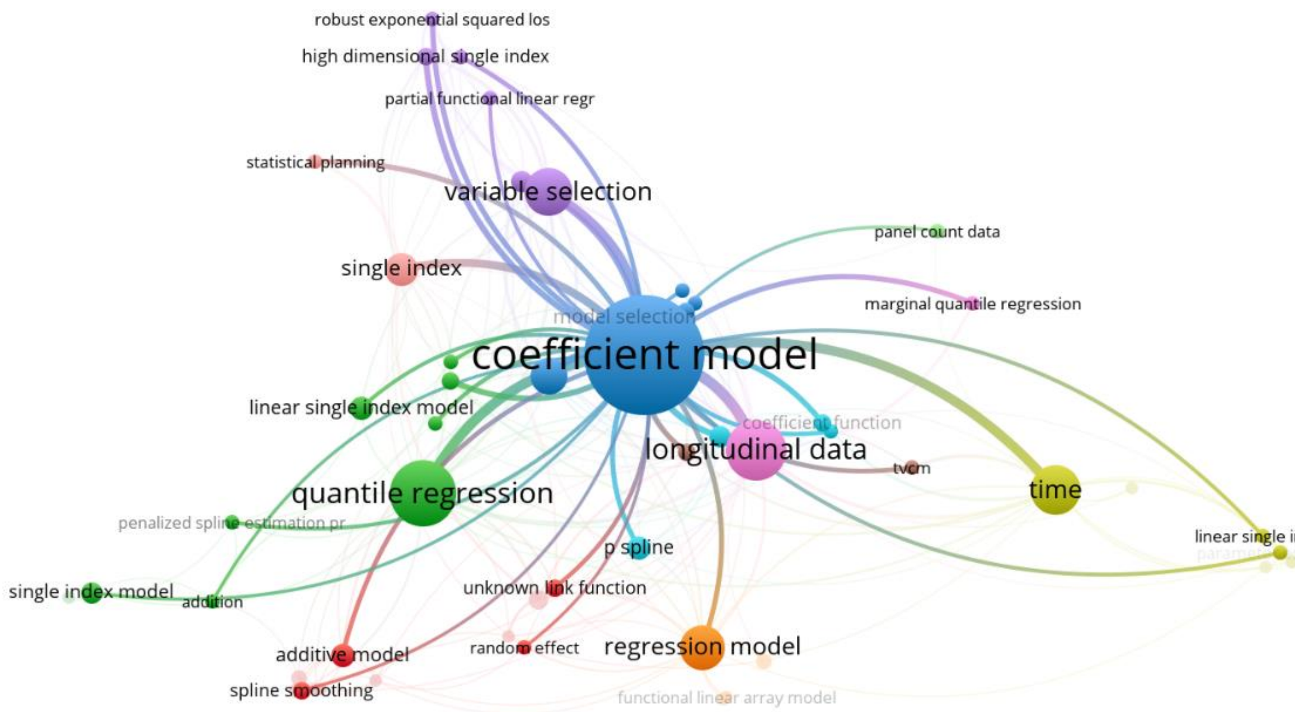

Fig. 10. (Varying) coefficient model title and abstract network terms

Theme evolution

History and continuity are the critical factors in analyzing the theme. The development of the theme links some keywords with others, which may expand or sharpen it. Fig. 12 showed the thematic evolution of important keywords. This analysis was divided into three periods 2011-2014, 2015-2017, and 2018-2021. Furthermore, the initial keywords in the first period include variable selection method, chain monte carlo, finite sample performance, unknown link function, and additive quantile regression. The keyword varying coefficient model develops a variable selection method, unknown link function, and additive quantile regression during the second period. This expands to a partially linear single index, quantile regression model, and panel count data. The variable selection procedure extended to chain monte carlo and finite sample performance though they were related during the third period.

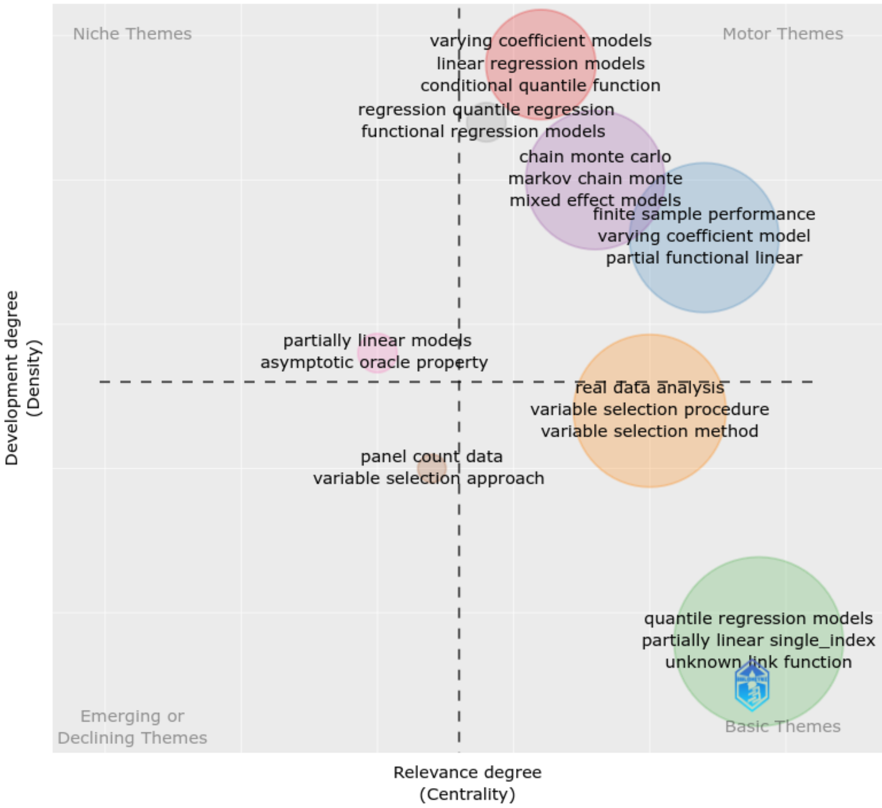

Fig. 11. Themes quadrant of varying coefficient models and quantile regression

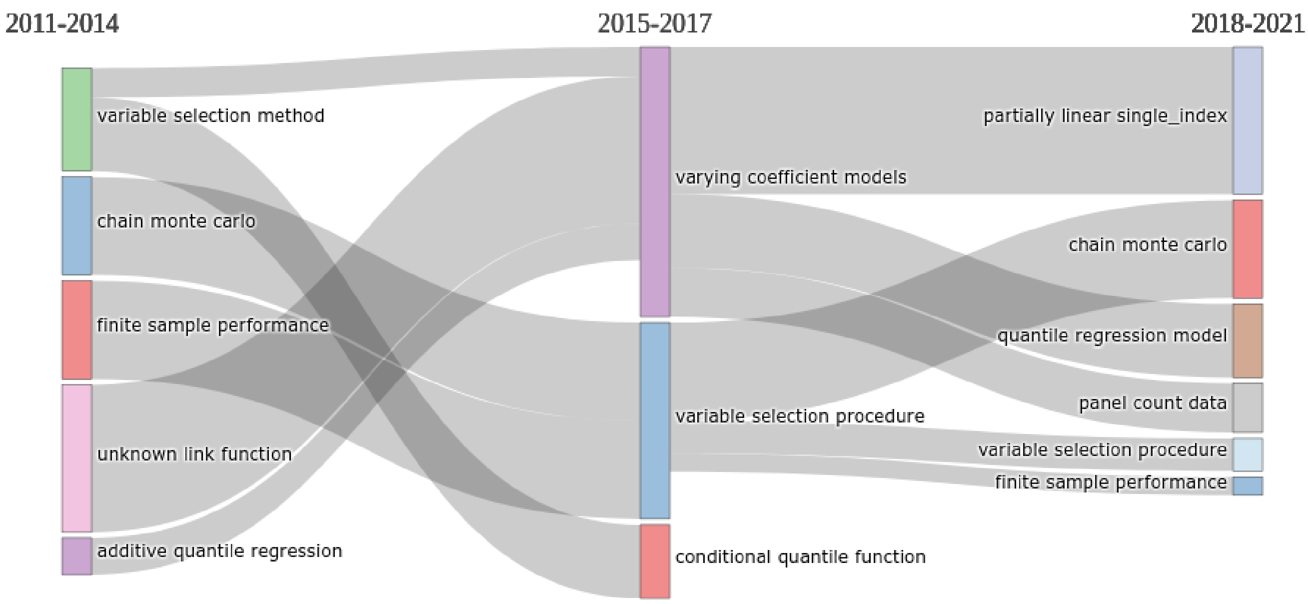

Fig. 12. Quantile regression title and abstract network terms

IV. Discussion

The results of the bibliometric analysis showed the most cited papers, that the title and abstract contain keywords varying coefficient model and quantile regression, dominated by Andriyana, Gijbels and Verhasselt (Table 3). They are the top 5 of the most productive authors in varying coefficient model, quantile regression and P-splines in longitudinal data (Table 4). They are also in the same cluster in the collaborative network between authors (Fig 5) that working on varying coefficient model, quantile regression and P-splines. Besides, they have working together in the same topics which the abstract term contain varying coefficient model, quantile regression and P-splines (Fig 6). Furthermore, they are connected in co-authorships networks in the same cluster (Fig 7).

The relationship between topics is illustrated by Fig. 8, 9, and 10 that both the quantile regression and (varying) coefficient model are related to the longitudinal data and time. In many situations, longitudinal data has a structure that can be measured in time and space. However, there are no keywords related to space or spatial based on the analysis. Some researchers use regression models and predictions with spatial effects, such as [19] and [20], while others use spatial modeling related to varying coefficient such as [21]. This is a possible future research topic. The discussion on varying coefficient models that include time and spatial effects is an interesting topic to study for the future agenda.

Table 6 shows the result of including the keywords space-time or spatiotemporal in the searching process. This result demonstrated that 6 documents contained specific keywords of the varying coefficient model, quantile regression, P-splines, space-time, and longitudinal data. On examination of these documents, the keywords were only mentioned in the documents but were not the main words. The searching method is expanded to include the keywords varying coefficient model and space-time or spatiotemporal. This method searches for titles and abstracts containing these keywords in Dimensions and Google scholar, resulting 7 and 2 documents, respectively. Furthermore, there are two papers specifically using space-time or spatiotemporal coupled with a varying coefficient model. The first paper by [22] discussed the varying coefficient model with both space and time effects. Also, the spline and kernel smoothing estimation methods were used in the context of mean regression to describe the time-varying coefficient result in the plot and the spatial-varying coefficient in the map.

TABLE 6
NUMBER OF PAPERS IN DATABASE SEARCHING WITH ADDING KEYWORDS "SPACE-TIME" OR "SPATIO TEMPORAL"

Code	Keywords	Source			Total
		Google Scholar	Dimensions	Science Direct	
A	"longitudinal data" AND ("varying coefficient model" OR "varying coefficient models")	760	353	10	**1123**
B	("quantile regression" OR "quantile regressions")	3510	2076	306	**5892**
C	("P-splines" OR "P-spline" OR "penalized spline" OR "penalized splines")	1360	772	107	**2239**
D	A AND B	144	72	7	**223**
E	**D AND C**	0	6	0	**6**

The second paper by [23] discussed varying coefficient models for spatiotemporal clustered data. Meanwhile, more papers with unspecific keywords were found in Google scholar. They are still related to spatiotemporal and varying coefficient models, such as [24], [25], and [26]. They had different primary focus on the themes. These papers were in the context of mean regression, and not quantile regression. Furthermore, there is an opportunity to develop this model in the context research of quantile regression.

V. CONCLUSION

The analysis uses sources from Google Scholar, Dimensions, and Science Direct database. There are no restrictions on the type of paper published through these sources. Unlike papers published in Scopus or Web of Science, those obtained from these databases include non-indexed journals and books. Additionally, for more accurate and objective result on the bibliometric analysis, Scopus or the Web of Science database sources are preferable.

This paper summarizes of the work on quantile regression in varying coefficient models. It shows a bibliographic mapping of the topics, the network of co-authorships, and the cluster of authors with similar specific topics. This bibliometric analysis reveals that there are several topics on quantile regression and varying coefficient models to be studied further. The results showed the most cited papers, that the title and abstract contain these keywords dominated by Andriyana, Gijbels and Verhasselt as the top 5 of the most productive authors. They are in the same cluster of the collaborative between authors and working together in this topic.

Based on the discussion, there are only a few articles that include space-time related to varying coefficient model in the context mean regressions, so the space-time varying coefficient model related to quantile regression as well as its estimation procedure is open for research development.

ACKNOWLEDGMENT

The authors are grateful for suggestions and inputs from the Rector of Universitas Padjadjaran.

REFERENCES

[1] T. Hastie and R. Tibshirani, "Varying-coefficient models," *J. R. Stat. Soc. Ser. B*, vol. 55, no. 4, pp757-779, 1993.
[2] J. Fan and W. Zhang, "Statistical methods with varying coefficient models," *Stat. Interface*, vol. 1, no. 1, pp179-195, 2008.
[3] L. Wang, H. Li, and J. Z. Huang, "Variable selection in nonparametric varying-coefficient models for analysis of repeated measurements" *J. Am. Stat. Assoc.*, vol. 103, no. 484, pp1556-1569, 2008.
[4] A. Antoniadis, I. Gijbels, and A. Verhasselt, "Variable Selection in Varying-Coefficient Models Using P-Splines," *J. Comput. Graph. Stat.*, vol. 21, no. 3, pp638-661, 2012.
[5] R. Koenker and G. Bassett, "Regression Quantiles," *Econometrica: Journal of The Econometric Society*, vol. 46, no. 1, p33-50, 1978.
[6] Y. Mu and Y. Wei, "A dynamic quantile regression transformation model for longitudinal data," *Stat. Sin.*, vol. 19, pp1137-1153, 2009.
[7] H. J. Wang, Z. Zhu, and J. Zhou, "Quantile regression in partially linear varying coefficient models," *Ann. Stat.*, vol. 37, no. 6B, pp3841-3866, 2009.
[8] Y. Andriyana, I. Gijbels, and A. Verhasselt, "P-splines quantile regression estimation in varying coefficient models," *Test*, vol. 23, no. 1, pp153-194, 2014.
[9] Y. Andriyana and I. Gijbels, "Quantile regression in heteroscedastic varying coefficient models," *AStA Adv. Stat. Anal.*, vol. 101, no. 2, pp151-176, 2017.
[10] Y. Andriyana, I. Gijbels, and A. Verhasselt, "Quantile regression in varying-coefficient models: non-crossing quantile curves and heteroscedasticity," *Stat. Pap.*, vol. 59, no. 4, pp1589-1621, 2018.
[11] R. Koenker, *Quantile Regression*. Cambridge University Press, UK, 2005.
[12] P. H. C. Eilers and B. D. Marx, "Flexible smoothing with b-splines and penalties," *Stat. Sci.*, vol. 11, no. 2, pp89-121, 1996.
[13] R. Koenker, P. Ng, and S. Portnoy, "Quantile smoothing splines," *Biometrika*, vol. 81, no. 4, pp673-680, 1994.
[14] R. Koenker and P. Ng, "A Frisch-Newton algorithm for sparse

quantile regression," *Acta Math. Appl. Sin.*, vol. 21, no. 2, pp225-236, 2005.
[15] R. W. Palmatier, M. B. Houston, and J. Hulland, "*Review articles: Purpose, process, and structure*." Springer, 2018.
[16] C. Oppenheim, *The publish or perish book, vol. 29, no. 2*. Tarma Software Research Pty Limited Melbourne, Australia, 2011.
[17] M. Aria and C. Cuccurullo, "bibliometrix: An R-tool for comprehensive science mapping analysis," *J. Informetr.*, vol. 11, no. 4, pp959-975, 2017.
[18] N. J. Van Eck and L. Waltman, "VOSviewer manual," *Leiden: Univeristeit Leiden*, vol. 1, no. 1, pp1-53, 2013.
[19] I. G. N. M. Jaya, Y. Andriyana, and B. Tantular, "Spatial Prediction of Malaria Risk with Application to Bandung City, Indonesia" *IAENG International Journal of Applied Mathematics*, vol. 51, no. 1, pp199-206, 2021.
[20] Yuliana Susanti, Sri Sulistijowati H., Hasih Pratiwi, and Twenty Liana, "Paddy Availability Modeling in Indonesia Using Spatial Regression." *IAENG International Journal of Applied Mathematics*, vol. 45, no. 4, pp398-403, 2015.
[21] I. G. N. M. Jaya, A. S. Abdullah, E. Hermawan, B. N. Ruchjana, "Bayesian Spatial Modeling and Mapping of Dengue Fever: A Case Study of Dengue Fever in the City of Bandung Indonesia" *Int. J. Appl. Math. Stat.* vol. 54, no 3, pp94-103, 2016.
[22] N. Serban, "A space-time varying coefficient model: The equity of service accessibility," *Ann. Appl. Stat.*, pp2024-2051, 2011.
[23] J. Lee, M. E. Kamenetsky, R. E. Gangnon, and J. Zhu, "Clustered spatio-temporal varying coefficient regression model," *Stat. Med.*, vol. 40, no. 2, pp 465-480, 2021.
[24] Z. Lu, D. J. Steinskog, D. Tjøstheim, and Q. Yao, "Adaptively varying-coefficient spatiotemporal models," *J. R. Stat. Soc. Ser. B (statistical Methodol.*, vol. 71, no. 4, pp859-880, 2009.
[25] L. E. Nieto-Barajas, "Bayesian regression with spatiotemporal varying coefficients," *Biometrical J.*, vol. 62, no. 5, pp1245-1263, 2020.
[26] C. Song, X. Shi, and J. Wang, "Spatiotemporally Varying Coefficients (STVC) model: A Bayesian local regression to detect spatial and temporal nonstationarity in variables relationships," *Ann. GIS*, vol. 26, no. 3, pp277-291, 2020.

Bertho Tantular received the B.S. degree from Department of Statistics Universitas Padjadjaran in 1998 and Master of Science from Institut Pertanian Bogor in 2009. He starts Doctor of Mathematics at Faculty of Mathematics and Natural Science, Universitas Padjadjaran in 2019. He is active in Longitudinal Data and Multilevel Analysis research.