Probabilistic Prediction Model Using Bayesian Inference in Climate Field: A Systematic Literature

Evi Ardiyani[1], Sri Nurdiati[2], Ardhasena Sopaheluwakan[3], Mohamad Khoirun Najib[4], Fadillah Rohimahastuti[5]

[1,2,4,5]Department of Mathematics, Mathematics and Natural Science Faculty, IPB University, Bogor 16680, Indonesia

[3]Center for Applied Climate Services, Agency for Meteorology Climatology and Geophysics, Jakarta 10720, Indonesia

ardiyanievi@apps.ipb.ac.id[1], nurdiati@apps.ipb.ac.id[2], ardhasena@bmkg.go.id[3], mkhoirun_najib@apps.ipb.ac.id[4], fadillahrohimahastuti@apps.ipb.ac.id[5]

	ABSTRACT
	Wildfires occur repeatedly every year and have a negative impact on natural ecosystems. Anticipation of wildfires is very necessary, therefore a prediction model is needed that can produce predictions with a good level of accuracy. One approach to develop probabilistic prediction models is Bayesian inference. The purpose of this research is to review the methods that can be used in developing probabilistic prediction models using the Bayesian approach. The methodology used is Systematic Literature Review (SLR) which can be used to provide a comprehensive review of Bayesian inference research in developing probabilistic prediction models. The research strategy used was the Boolean Technique applied to database sources including Scopus, IEEE Xplore, and ArXiv. The articles used have novelty and ease of explanation of Bayesian methods, especially predictions in the field of climate so that articles are selected based on inclusion and exclusion criteria. The results show that probabilistic models can provide more accurate results than deterministic models. The Bayesian Model Averaging (BMA) method is a widely used method because it is easy to implement and develop so that the prediction results can be more accurate. The development of probabilistic prediction models with a Bayesian approach has great potential to grow as seen from the development of the number of research publications over the past 5 years. The research position of probabilistic prediction models with Bayesian approaches in the field of climate is dominated by the research community in China with the main problems related to hydrology.

———————————— ◆ ————————————

A. INTRODUCTION

Wildfires are one of the natural disasters that occur annually. Wildfires have a negative impacts such as reducing the quality of natural ecosystems to public health problems around the fire site (Aflahah et al., 2019). In 2007, a forest fire in Greek town became the worst fire season on record, resulting in 84 deaths, 270,000 hectares burned and more than 3,000 homes destroyed in 110 villages (Ribeiro et al., 2019). In Indonesia, one of the largest wildfires occurred in 2015 with a burned area of 2.6 million (ha) (Endrawati, 2018) that caused dense haze covering 80% of Sumatra and Kalimantan (Santriwati et al., 2021). Forests have an important role in life, one of which is the producer of oxygen for human survival (Samsuri, Anita

Zaitunah, 2021). Forests in Indonesia are referred to as one of the lungs of the world that contributes oxygen for the sustainability of living things, especially the Kalimantan region. Kalimantan, which is also called the "heart of Borneo", is able to produce 40% of the oxygen on earth, so Kalimantan has been named the "Lungs of the World" since 2007 (Wahyuni & Suranto, 2021). Wildfires in Kalimantan are influenced by the local climate such as rainfall, humidity, temperature, and wind speed (Candra et al., 2017). The global climate also affects fires such as *El Niño Southern Oscillation* (ENSO) and *Indian Ocean Dipole* (IOD) (Najib et al., 2022; Nurdiati, Bukhari, et al., 2021). The occurrence of *El Niño,* namely the warm pase of the ENSO and the positive IOD either individually or together can significantly increase the scale of the wildfires phenomenon as happened in 2002, 2006, 2009, 2015, and 2019 (Fanin & Van Der Werf, 2017). ENSO and IOD phenomena are strongly influenced by global warming (Suhermat et al., 2021). As time goes by, global warming continues to increase which will increase the potential for wildfires in Kalimantan considering that Kalimantan has a close relationship with ENSO and IOD. Therefore, an early warning system for wildfires is needed. A deterministic prediction model has been built in 2021 by (Nurdiati et al., 2021). Deterministic models produce a single output with an error value. In recent decades, many researchers have developed probabilistic prediction models.

Probabilistic prediction models provide important information about the uncertainty of the error value, in other words, probabilistic prediction models pay attention to the error value (Zhang et al., 2022; Bazionis & Georgilakis, 2021; Z. Li et al., 2013;). There are several probabilistic prediction models, including probabilistic regression model, hidden Markov chain, and Bayesian. Danelljan et al. (2020) proposed a probabilistic regression formulation and applied it to visual tracking because visual tracking is still prone to errors and malfunctions with this method substantially improving performance. Nguyen-Le et al. (2020) combines the techniques of long-term memory and the hidden Markov model to predict fracture propagation problems in engineering. The primary benefit of the hidden Markov model is its capacity to learn with fewer data. Using a combination of long short-term memory and the hidden Markov model yields a more accurate solution than using only long short-term memory, particularly in the case of datasets that are deficient in information.

One of the best approach to building probabilistic models is the Bayesian Inference (Marin et al., 2005; Cowles, 2003). Bayesian inference-based probabilistic models have the basic concept of Bayes' theorem developed by Thomas Bayes in 1763 (Efron, 2013) in equation (1). The probability model has the main principle of updating information with new information to obtain new, more accurate knowledge (Natalia et al., 2020; Wang et al., 2002).

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \tag{1}$$

Bayesian inference is widely used to build a prediction model, especially in the field of biostatistics (Xie, 2022; Gaskins et al., 2021; Patrick, 2012), economic (Lekar et al., 2019; Gilboa et al., 2008; Cooper et al., 2004), and industry (Adedipe et al., 2020; Ghosh et al., 2020; Basuki et al., 2014). Bayesian models provide instruments for risk estimation and enable decision makers to integrate objective expert estimates with historical data. Using a Bayesian network

model, (Pendharkar et al., 2005) demonstrate how a confidence update procedure can be used to incorporate decision-making risk and compared the predictive performance of the Bayesian model to that of nonparametric neural network forecasting models and well-known regression tree forecasting models and demonstrate that the Bayesian model is a competitive model for forecasting software development effort. Deep Learning approaches in prediction models frequently produce deterministic estimates and do not account for the inherent uncertainty in model predictions. Abbaszadeh et al. (2022) presents a framework for probabilistic estimation that employs Bayesian Model Averaging (BMA). Taking into account model uncertainty, the proposed method generates more accurate and trustworthy soybean yield forecasts, as demonstrated by the results.

Climate change in the world greatly affects other natural phenomena that have the potential to cause natural disasters in some areas. Therefore, the importance of a prediction of climate change. Climate change affected the entire world, specifically India's fragile Himalayan mountain region, which has high significance due to being a climatic indicator, (Haq, 2022) developed a Deep Long Short-Term Memory (CDLSTM) model and optimized to forecast all Himalayan states' temperature and rainfall values. In the field of climate, the use of Bayesian inference to develop probabilistic prediction models is still small, especially in Indonesia. Therefore, this review article aims to study the application of the Bayesian approach in building probabilistic prediction models in the field of climate. For descriptive analysis and synthesis of articles, we use Systematic Literature Review (SLR). The findings of this study are expected to provide information on the current state of research and future research prospects related to the use of Bayesian inference in building probabilistic prediction models in climate analysis, especially in Indonesia.

B. METHODS

The methodological principles of this research were developed based on the Study Literature Review (SLR) by applying the Systematic Reviews and Meta-Analyses (PRISMA) methodology. PRISMA has a flowchart that can improve systematic reviews and meta-analyses (Moher et al., 2009). The PRISMA flowchart has four phases including identification, screening, eligibility, and inclusion. According to Albuquerque et al. (2021), There are 4 essential things in conducting SLR, including exploring the stages of literature assessment, identifying the Step process, developing research questions, and how to assess to overcome possible bias Exploring the stages of literature review and the step process can be done by Pickering & Byrne (2014) defining the topic, formulating research questions, identifying keywords, conducting electronic searches, assessing publications, data acquisition, data cleaning, testing publications, revising summary tables, method drafts, and writing revisions until the paper is finished. We used the general guidelines introduced by that focused on developing research questions and assessing articles to overcome possible bias in SLR. This study's descriptive analysis uses text mining to provide a comprehensive overview of the research topic. The detailed process can be seen in Figure 1 (Xiao & Watson, 2019).
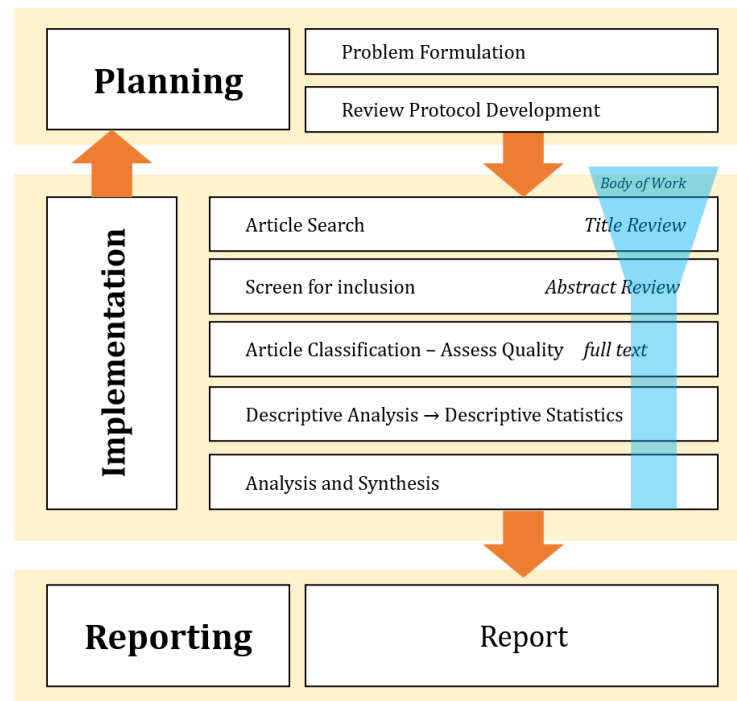
Figure 1. SLR Process

1. Research Strategy

The article search strategy used in this research uses the Boolean Technique. The Boolean technique is adapted to obtain essential search terms in the article collection. The search terms used are "Bayesian AND Probabilistic AND Prediction AND Climate." The database sources used in the article search process were sourced from Scopus, IEEE Xplore, and ArXiv. The selected articles used are based on the inclusion and exclusion criteria in Table 1.

Table 1. Inclusion and Exclusion Criteria

No.	Inclusion Criteria	Excision Criteria
1.	Must be articles and conference papers that published in the last 10 years (2012 – 2022).	Articles are written in another language.
2.	The article should include the development and implementation of the Bayesian inference for model prediction.	Bayesian implementation does not focus on prediction models.
3.	Must be written in English	Articles do not focus on climate.
4.	QA Value>80%, QA questions refer to Table 3.	

2. Research Questions

The purpose of this study was also to answer the research questions in Table 2. We used the general guidelines developed by Petticrew & Roberts (2008) which focuses on developing research questions and how to assess articles to avoid bias in SLR. Articles obtained in the search strategy are then screened based on the title and abstract in answering the research question, as shown in Table 2.

Table 2. Research Question

No.	Research Question
1.	What are the advantages and disadvantages of probabilistic prediction models?
2.	What are the advantages and disadvantages of Bayesian Inference in Probabilistic models?
3.	What methods are easy to implement and develop with Bayesian Inference?
4.	What are the countries that have developed the concept of Bayesian Inference in prediction models?

3. Quality Assurance

After screening, all articles were ready for quality testing. The article quality test was conducted based on the method developed by (Al-Emran et al., 2018), we used six testing criteria in Table 3. This method scores articles by following the rules: score 1 for articles with "yes" quality test questions, 0 for articles with "no" quality test questions, and 0.5 for papers with "partial" quality test questions. Three people conducted the assessment, and the scores obtained were averaged to get the final score in Table 3. Articles on this study used papers with an assessment percentage above 80% in this study, as shown in Table 3.

Table 3. Quality Assurance Question

No.	Quality Assurance
1	Are the research aims specified clearly?
2	Is the information presented clear and concise?
3	Does the study provide enough explanation of its methodology?
4	Do the findings of the study add to the understanding of Bayesian models?
5	Are the conclusions clearly identified?
6	Are the conclusions logical and concise with the flow of the paper?

C. RESULT AND DISCUSSION

Search results with keywords described in the research strategy section obtained 196 data, including articles, conference papers, conference reviews, reviews, and book chapters published in 2012 - 2022 with the big theme of Bayesian in climate. Following the inclusion and exclusion criteria, the papers that have been collected are filtered based on the type of paper so that the papers used are only types of articles and conference papers. Scientific articles filtered by type are sourced from 3 databases, namely Scopus, IEEE Xplore, and ArXiv, so it is necessary to remove duplicates. Articles sourced from the three databases duplicated a total of two articles. Articles were evaluated based on the title and abstract according to the inclusion and exclusion criteria in Table 1.

Furthermore, 52 articles that have passed the title and abstract screening are evaluated based on the research question in Table 2. Based on the title and abstract screening, 15 papers were assessed based on the quality assessment in Table 3 with the evaluation described in the previous section. Table 4 shows the results of the review of 16 articles. Three papers did not qualify with a percentage score below 80%, so only 13 papers will be used in this study. Details of the paper selection process are shown in Figure 2.
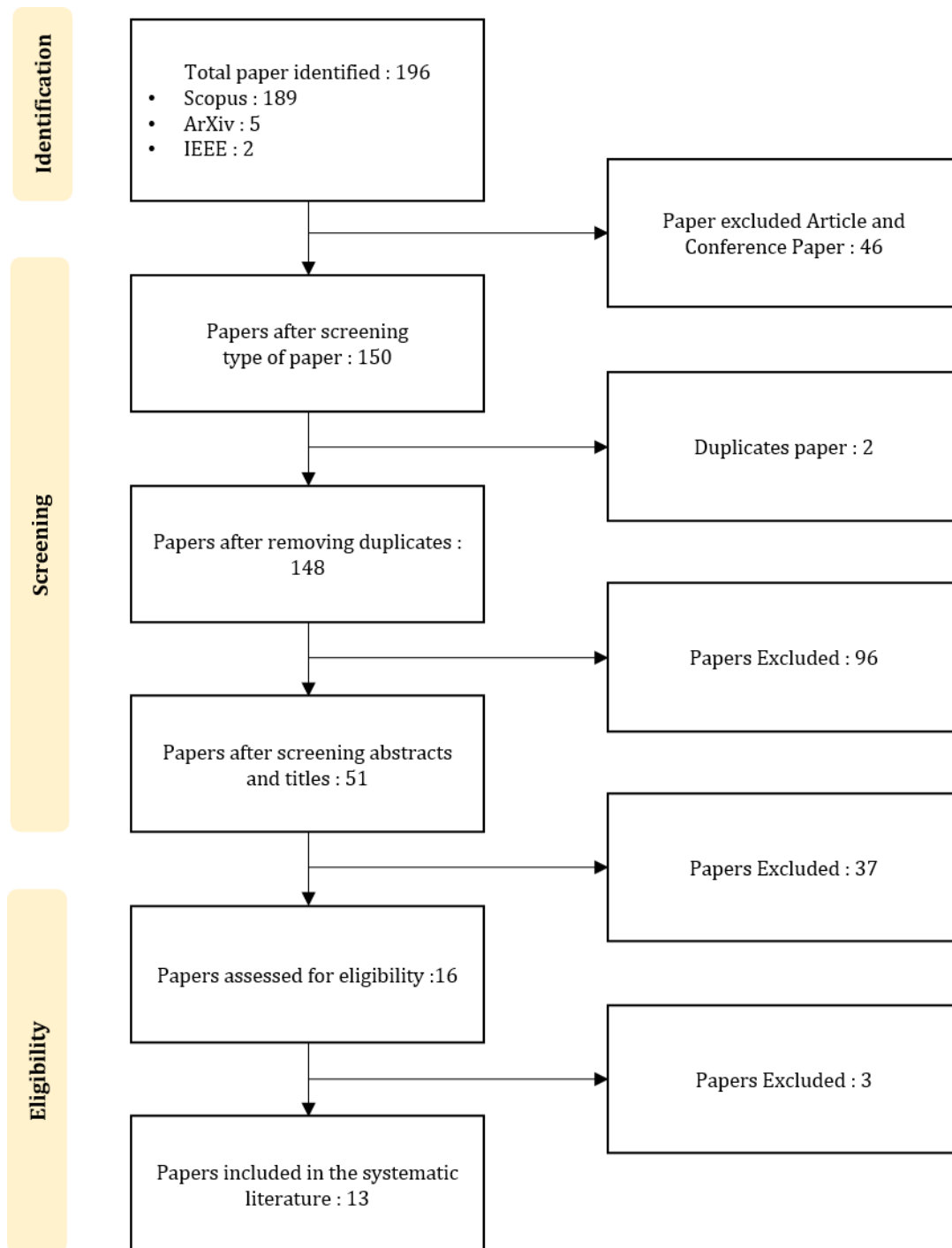
Figure 2. SLR diagram

1. Quality Assurance Results

Based on the questions in Table 4 and the ratings described in the previous section, the assessor gave ratings for the 16 articles used in this study. The maximum percentage score that can be obtained is 100%. Some articles scored below 80%, indicating that the articles needed to meet the quality assurance described in the previous section. This study does not use articles with a quality assurance value below 80%. The assessment results on 16 articles are shown in Table 4 with the information Q1-Q6 is Quality Assurance 1-6, A1-A3 is assessors 1-3, and S1-S16 is the source of articles 1-16, as shown in Table 4.

Table 4. Quality Assurance Results

		Q1			Q2			Q3			Q4			Q5			Q6			%
		A1	A2	A3	A1	A2	A3	A1	A2	A3	A1	A2	A3	A1	A2	A3	A1	A2	A3	
Source	S1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	97,2
	S2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	77,8
	S3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	97,2
	S4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	97,2
	S5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	100
	S6	1	1	1	1	1	1	1	1	0.5	1	1	1	1	1	1	1	1	1	91,7
	S7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	100
	S8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.5	1	90,3
	S9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.5	1	97,2
	S10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	94,4
	S11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	88,9
	S12	1	1	1	1	1	1	0	1	1	0	1	1	1	0	1	1	0.5	1	69,4
	S13	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0.5	1	72,2
	S14	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	91,7
	S15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	97,2
	S16	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	88,9

2. Descriptive Analysis

The position of Bayesian Inference research in probabilistic prediction models in the field of climate can be analyzed using diagrams and graphs. Figure 2 shows the number of citations of scientific articles. Figure 3 shows the classification results of related articles. Figure 4 shows the percentage of probabilistic prediction model development locations using Bayesian inference. Based on Figure 2, the number of articles cited each year is higher than those published each year. The highest publication was in 2013, and in the following year, no publication was related to it, indicating the need for more research on probabilistic prediction models using Bayesian inference, as shown in Figure 2.
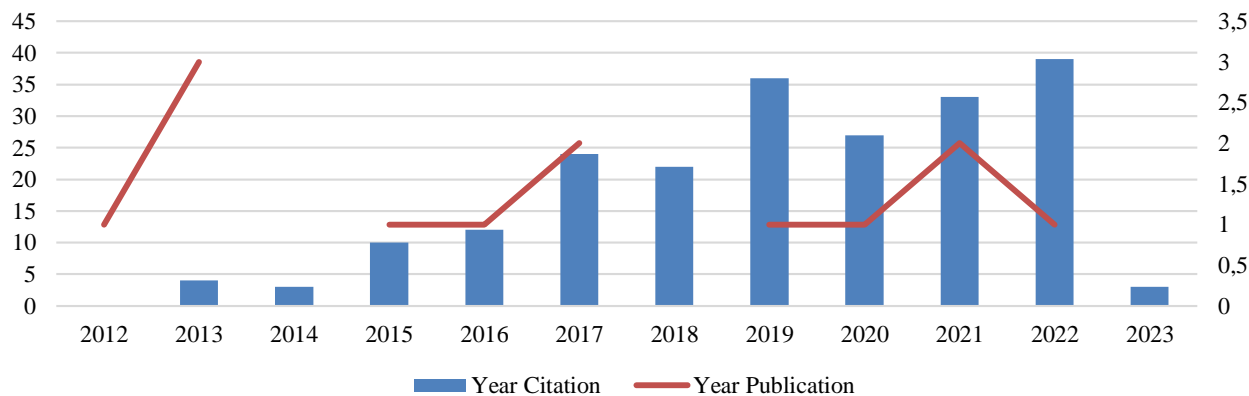

Figure 2. Articles citations and publications each year

Most publications (38.5%) used the Bayesian Model Average (BMA) approach in building probabilistic prediction models shown in Figure 3. Based on the results of these studies, BMA is one approach that is easy to implement and modify to produce good accuracy in probabilistic

prediction models. As a framework for model selection and combination, BMA may be used to a broad variety of issues. The determination of posterior model probabilities renders BMA a very simple model selection procedure. Under the Bayesian paradigm, such probabilities are clearly interpretable, the model space may be as large as necessary, and there is no need to keep track of the number of parameters or kind of penalty, as is needed by the information criterion approaches used in the statistical literature. The second and third widely used methods are Traditional Bayesian (TB) and Bayesian Network (BN). The selection of the prior model strongly influences both methods. This is different from the BMA method which uses the entire prior model as the posterior. In BN, there is no universally accepted method for building networks from data. This makes the design of Bayesian Networks a relatively large undertaking. A problem that arises is the fact that Bayesian networks can only exploit causal influences recognized by the person programming them, as shown in Figure 3.
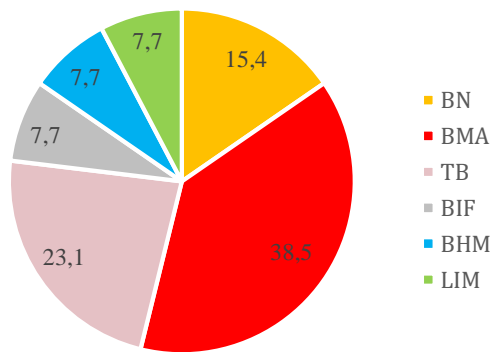

Figure 3. Sub-them Bayesian inference for each articles

Based on Figure 4 the countries that dominate the research on developing probabilistic prediction models using Bayesian inference are China, Germany, and Australia. Research in these three regions focuses on climate, with China in particular focusing on hydrology, as shown in Figure 4.
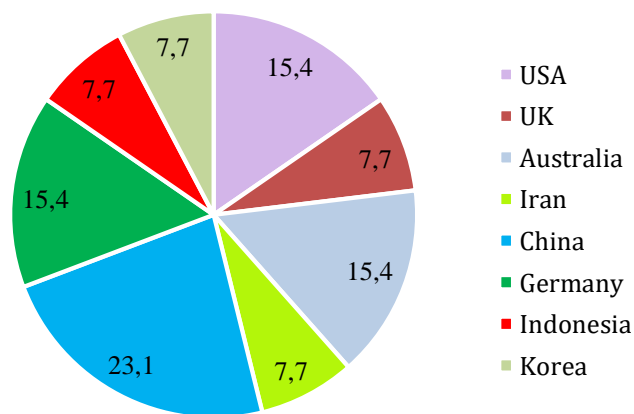

Figure 4. Percentage of articles by research location

3. Answer to Research Questions

Based on the objectives and findings summarized in Table 5, the drawback of the probabilistic prediction model is that it requires good prior knowledge because the prior knowledge strongly influences the computational results of the probabilistic model in the previous model. Therefore, its computation will require more development and computation time than the non-probabilistic model. Meanwhile, since the probabilistic model is an update of the previous model, the advantages of the probabilistic prediction model can provide more accurate prediction results. The accuracy of probabilistic models is also affected by the use of "uncertainty" in bias, model trends, and variability (Fan et al., 2017). Therefore, checking the quality of prediction accuracy can be done more efficiently. Bayesian inference in probabilistic models is one alternative to obtain prediction results. The Bayesian approach has the basic concept of updating the information in each model, such as the concept of probabilistic models, updating the information can significantly improve accuracy results. However, the computational time required is longer, and the proposed model tends to be "expensive" this is can be a disadvantage of Bayesian inference for probabilistic models. Although computationally expensive, Bayesian inference prediction models can provide good benefits considering the importance of the accuracy of prediction model results as a decision support system (Foster et al., 2020; Siegert et al., 2016).

The Bayesian approach has several methods, including Bayesian Model Averaging (BMA), Bayesian Network, Bayesian Hierarchical Model, and Bayesian Inferential Framework (Agustin & Adi, 2021; Alinezhad et al., 2021; Foster et al., 2020; Ji et al., 2019; Kapsch et al., 2012; Kim et al., 2017; Y. Li et al., 2022; Liang et al., 2013; Olson et al., 2016; Röpnack et al., 2013; Sharifahmadian & Latifi, 2013; Siegert et al., 2016). Based on the articles used in this study, 38% of studies used the BMA method. The BMA method is easy to implement and develop. The development of the BMA method and the provision of weights can increase the accuracy rate (Fan et al., 2017; Olson et al., 2016). The second order is traditional Bayesian which only uses one input, namely the error value or 'uncertainties' in the previous model. Next is the Bayesian network method for probabilistic prediction models. All of these methods utilize previous information, which aims to increase knowledge so that predictions can become more accurate. The three ways have characteristics whose usefulness can be adjusted according to the needs. The main problem in building a probabilistic prediction model with the Bayesian approach is the selection of several priors and models used as input for the next stage. The BMA and Bayesian Network methods are suitable for overcoming this problem. The BMA method is a method that is easy to implement and modify. There are three versions of the BMA method, which basically uses all priors or previous models in obtaining the posterior distribution. in the BMA method, the posterior distribution is assumed in equation 2 with g_k is Gaussian distribution and hyperparameter (mean $(a_k + b_k f_k)$ and variance σ_k^2). Parameter vector θ on model provides $(a_k, b_k), \sigma_k^2$ and w_k for each $k = 1, \ldots, K$. For clarity, the different schematics of the three versions are shown in Figure 3. For clarity, the different schematics of the three versions are shown in Figure 5.

$$p(y|f, \theta) = \sum_{k=1}^{K} w_k g_k(y|f_k) \tag{2}$$

Of the three versions of BMA, the BMA method was modified to improve the quality of the prior distribution, namely Leave One Out (LOO) BMA and Perturbed BMA. These modifications are inspired by the concept of leave one out of cross-validation and random forest. LOO BMA estimates the mean with the LOO method and then calculates the weight and variance with maximum likelihood. Whereas in perturbed BMA, the parameters are chosen randomly. Perturbed BMA generates multiple models by injecting randomness into the data, similar to a random forest.

Another method is the Bayesian Network (BN). Unlike the BMA method, which mainly uses all fairly similar prior distributions, the BN method is effective for variables and knowledge that vary from prior knowledge. These variables and knowledge are intended to be integrated into a framework. BN is suitable in such cases because it includes a probabilistic relationship between variables and allows for mutually independent variables, as shown in Figure 5.
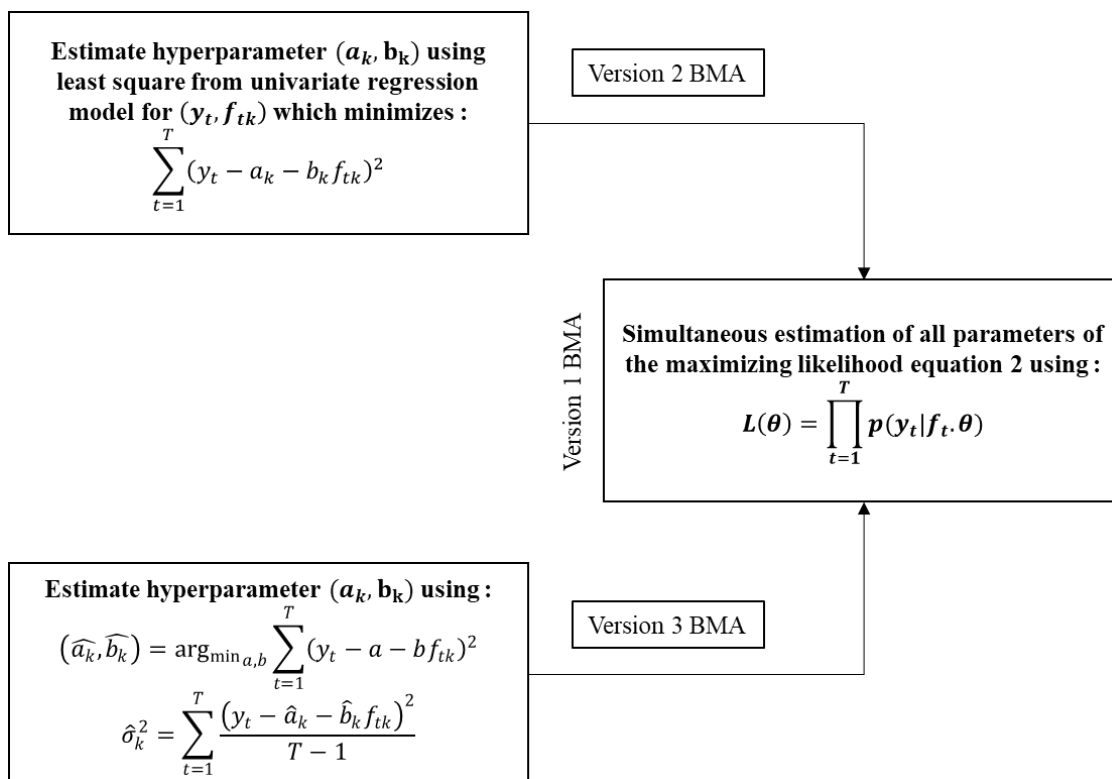

Figure 5. Three version of BMA

The countries that have developed the concept of Bayesian Inference in prediction models are China, Germany, Indonesia, Australia, the USA, the UK, Iran, and Korea. The majority of the research is in China, with the main problem in the field of hydrology. In other countries, studies focus on Sea Surface Temperature (SST), which is influenced by rainfall and global climate. While some studies are conducted globally, some focus on one region, such as East Nusa Tenggara, East Asia, and South Australia, because they have consistent drought patterns yearly. Consistent drought patterns are an advantage in making prediction models with good results.

Table 5. Summary of Articles

Source	Aims	Findings
Foster *et al.* 2020 (Foster et al., 2020)	Implementation, Calibration, and performance evaluation of Linear Inverse Model in SST anomaly forecasting using Bayesian inference and Probabilistic Scoring	• Bayesian inference provides better probabilistic and estimation capabilities when compared to traditional point estimation methods for LIM models. • The choice of prior distribution significantly impacts the estimation results, such as improving the ability of the model to capture the anomalous variability of SST. • Bayesian inference is computationally expensive.
Siegert *et al.* 2016 (Siegert et al., 2016)	Application of Bayesian Framework for verification and recalibration of ensemble forecast based on small hindcast dataset	• Bayesian framework can improve experience capabilities compared to more straightforward recalibration methods • Computational methods for Bayesian analysis tend to be "expensive." • Inferences depend on the assumption of appropriate parametric middles.
Fan *et al.* 2017 (Fan et al., 2017)	Using Bayesian inference, calculate model weights on climate change projection ensembles to create probabilistic projections.	• The method presented is more robust • Different priors used can lead to different conclusions • Bayesian inference considers "uncertainties" in model bias, trends, and internal variability, including errors in the observations used.
Alinezhad *et al.* 2021 (Alinezhad et al., 2021)	Hydro-climate projections with uncertainties in the Zayandeh-Rud river basin using Bayesian Model Averaging (BMA).	The BMA approach produces better estimates of meteorology, including temperature and rainfall, from the observed values during the base period.
Olson *et al.* 2016 (Olson et al., 2016)	Probabilistic projections using Regional Climate Models (RCMs) with Bayesian Model Averaging (BMA)	Weighted projections are well-calibrated and more precise than unweighted projections.
Liang *et al.* 2013 (Liang et al., 2013)	Forecasting on several combinations of hydrological models using Bayesian Model Averaging (BMA)	• BMA is always better than Simple Average Methods (SAM) • BMA approach is excellent and robust for integrating single models. • Confidence interval prediction with BMA can provide information to support flood control decisions.
Röpnack *et al.* 2013 (Röpnack et al., 2013)	Comparing ensembles of atmospheric predictions to enable relative measurements between different EPS systems	• Bayesian model suitable for comparing ensemble systems with each other • Based on the Bayesian model, the modeled temperature profile of COSMO-DE-EPS is more consistent than the observed profile of COSMO-SREPS.
Kim *et al.* 2014 (Kim et al., 2017)	New Bayesian Averaging Methods untuk probabilistic ensemble forecasting dengan menggabungkan deterministik ensemble forecast berdasarkan pendekatan regresi multivariat dengan Bayesian Model Averaging (BMA)	New Bayesian Averaging Methods for probabilistic ensemble forecasting by combining deterministic ensemble forecasting based on multivariate regression approach with Bayesian Model Averaging (BMA).
Kapsch *et al.* 2012 (Kapsch et al., 2012)	• Evaluate the performance of WTs that may cause damage by hail storms	• The proposed methods of LOO BMA and Perturbed BMA are better at predicting

	• Evaluate the capability of RCM simulation in reliably producing hail-related synoptic systems in WR distribution • Develop a probabilistic prediction model of the amount of hail damage	than the standard Bayesian Model Averaging approach. • The proposed model can be used as an early warning system to avoid structural damage to buildings and infrastructure, although proposed model must improve the model's accuracy.
Agustin *et al.* 2021 (Agustin & Adi, 2021)	Measuring disaster impact as a decision support system using Bayesian Network	The updated Bayesian approach can be used to guide maintenance strategies and operations of buildings and infrastructure.
Ji *et al.* 2019 (Ji et al., 2019)	Improving rainfall prediction capability in East Asia based on Ensemble Prediction System (EPS) outputs from ECMWF, NCEP, and UKMO from TIGGE dataset.	• Light rainfall events are more accurately predicted using the standard deterministic BMA with limited ability for moderate and heavy rainfall events. • The categorized deterministic BMA model is superior to the standard model, especially for moderate and heavy rainfall. • The categorized BMA deterministic model provides better-calibrated rainfall probabilities than the standard model. • The superiority of the BMA model to climatological forecasts only sometimes lasts for a more extended period.
Li *et al.* 2022 (Y. Li et al., 2022)	Developing Spatiotemporal Projection Based Bayesian Hierarchical Model (STP-BHM) for sub-seasonal rainfall forecasting.	• STP-BHM model can provide skillful and reliable probabilistic forecasts • Bayesian statistical model is more flexible and efficient for assessing multiple reliable sources • The STP-BHM model shows the functional predictive ability for both below-normal and above-normal events, and positive Brier ability is seen at all critical times • The STP-BHM model outperforms the NCEP S2S dynamic model when the lead time is more than 5s.
Sharifahmadian & Latifi 2013 (Sharifahmadian & Latifi, 2013)	Measuring the performance of the Water Environment Risk prediction model using Bayesian Network.	The proposed method effectively predicts Waster Environment Risk and improves water resource.

D. CONCLUSION AND SUGGESTIONS

In conclusion, this article aims to conduct a systematic review of probabilistic prediction models using Bayesian inference in the field of climate. The main goal of this review was to gain insight into probabilistic prediction models using Bayesian inference and to find the best approach to implement them. However, overall, probabilistic prediction models can improve accuracy. Methods used in building probabilistic prediction models in the climate field include Bayesian Inferential Framework, Bayesian Network, Bayesian Hierarchical Model, and Bayesian Model Average. BMA is easy to modify and implement, so the majority of papers, as much as 38%, use BMA as a method for building prediction models. The majority of countries that developed probabilistic prediction models using Bayesian is China, with its main problems related to hydrology.

REFERENCES

Abbaszadeh, P., Gavahi, K., Alipour, A., Deb, P., & Moradkhani, H. (2022). Bayesian Multi-modeling of Deep Neural Nets for Probabilistic Crop Yield Prediction. *Agricultural and Forest Meteorology*, *314*. https://doi.org/10.1016/j.agrformet.2021.108773

Adedipe, T., Shafiee, M., & Zio, E. (2020). Bayesian Network Modelling for the Wind Energy Industry: An Overview. *Reliability Engineering and System Safety*, *202*. https://doi.org/10.1016/j.ress.2020.107053

Aflahah, E., Hidayat, R., & Alfahmi, F. (2019). Pendugaan hotspot sebagai indikator kebakaran hutan di Kalimantan berdasarkan faktor iklim. *Pendugaan Hotspot Sebagai Indikator Kebakaran Hutan Di Kalimantan Berdasarkan Faktor Iklim*, *9*(2), 405–418. https://doi.org/10.29244/jpsl.9.2.405-418

Agustin, A. D., & Adi, T. J. W. (2021). Prediction Models of Infrastructure Resilience as a Decision Support System Based on Bayesian Network. *IOP Conference Series: Earth and Environmental Science*, *832*(1). https://doi.org/10.1088/1755-1315/832/1/012014

Al-Emran, M., Mezhuyev, V., Kamaludin, A., & Shaalan, K. (2018). The impact of knowledge management processes on information systems: A systematic review. *International Journal of Management*, *43*, 173–187.

Albuquerque, V., Dias, M. S., & Bacao, F. (2021). Machine learning approaches to bike-sharing systems: A systematic literature review. *ISPRS International Journal of Geo-Information*, *10*(2). https://doi.org/10.3390/ijgi10020062

Alinezhad, A., Gohari, A., Eslamian, S., & Saberi, Z. (2021). A probabilistic Bayesian framework to deal with the uncertainty in hydro-climate projection of Zayandeh-Rud River Basin. *Theoretical and Applied Climatology*, *144*(3–4), 847–860. https://doi.org/10.1007/s00704-021-03575-3

Basuki, M., Manfaat, D., Nugroho, S., & Dinariyana, A. A. B. (2014). Probabilistic risk assessment of the shipyard industry using the bayesian method. *International Journal of Technology*, *5*(1), 88–97. https://doi.org/10.14716/ijtech.v5i1.157

Bazionis, I. K., & Georgilakis, P. S. (2021). Review of Deterministic and Probabilistic Wind Power Forecasting: Models, Methods, and Future Research. *Electricity*, *2*(1), 13–47. https://doi.org/10.3390/electricity2010002

Candra, A., Thamrin, & Mubarak. (2017). Analisis Pengaruh Faktor Iklim Dan Kebakaran Hutan/Lahan Terhadap Konsentrasi Pm10 Di Kota Pekanbaru Selama Kurun Waktu Tahun 2011-2015. *Ilmu Lingkungan*, *11*(2), 209–227. https://doi.org/10.31258/jil.11.2.p.209-227

Cooper, N. J., Sutton, A. J., Abrams, K. R., Turner, D., & Wailoo, A. (2004). Comprehensive decision analytical modelling in economic evaluation: A Bayesian approach. *Health Economics*, *13*(3), 203–

226. https://doi.org/10.1002/hec.804

Cowles, M. K. (2003). Bayesian Statistical Modelling. *Journal of the American Statistical Association*, *98*(461), 256–257. https://doi.org/10.1198/jasa.2003.s262

Danelljan, M., Van Gool, L., & Timofte, R. (2020). Probabilistic regression for visual tracking. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 7181–7190. https://doi.org/10.1109/CVPR42600.2020.00721

Efron, B. (2013). Bayes' theorem in the 21st century. *Science*, *340*(6137), 1177–1178. https://doi.org/10.1126/science.1236536

Endrawati, E. (2018). Identifikasi Areal Bekas Kebakaran Hutan Dan Lahan Menggunakan Analisis Semi Otomatis Citra Satelit Landsat. *Seminar Nasional Geomatika*, *2*, 273. https://doi.org/10.24895/sng.2017.2-0.420

Fan, Y., Olson, R., & Evans, J. P. (2017). A Bayesian posterior predictive framework for weighting ensemble regional climate models. *Geoscientific Model Development*, *10*(6), 2321–2332. https://doi.org/10.5194/gmd-10-2321-2017

Fanin, T., & Van Der Werf, G. R. (2017). Precipitation-fire linkages in Indonesia (1997-2015). *Biogeosciences*, *14*(18), 3995–4008. https://doi.org/10.5194/bg-14-3995-2017

Foster, D., Comeau, D., & Urban, N. M. (2020). A Bayesian Approach to Regional Decadal Predictability: Sparse Parameter Estimation in High-Dimensional Linear Inverse Models of High-Latitude Sea Surface Temperature Variability. *Journal of Climate*, *33*(14), 6065–6081. https://doi.org/10.1175/JCLI-D-19-0769.1

Gaskins, J. T., Fuentes, C., & De La Cruz, R. (2021). A Bayesian nonparametric model for classification of longitudinal profiles. *Biostatistics*. https://doi.org/10.1093/biostatistics/kxab026

Ghosh, S., Pandita, P., Atkinson, S., Subber, W., Zhang, Y., Kumar, N. C., Chakrabarti, S., & Wang, L. (2020). Advances in Bayesian Probabilistic Modeling for Industrial Applications. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, *6*(3). https://doi.org/10.1115/1.4046747

Gilboa, I., Postlewaite, A. W., & Schmeidler, D. (2008). Probability and uncertainty in economic modeling. *Journal of Economic Perspectives*, *22*(3), 173–188. https://doi.org/10.1257/jep.22.3.173

Haq, M. A. (2022). CDLSTM: A novel model for climate change forecasting. *Computers, Materials and Continua*, *71*(2), 2363–2381. https://doi.org/10.32604/cmc.2022.023059

Ji, L., Zhi, X., Zhu, S., & Fraedrich, K. (2019). Probabilistic precipitation forecasting over East Asia using Bayesian model averaging. *Weather and Forecasting*, *34*(2), 377–392. https://doi.org/10.1175/WAF-D-18-0093.1

Kapsch, M. L., Kunz, M., Vitolo, R., & Economou, T. (2012). Long-term trends of hail-related weather types in an ensemble of regional climate models using a Bayesian approach. *Journal of Geophysical Research Atmospheres*, *117*(15). https://doi.org/10.1029/2011JD017185

Kim, Y., Kim, W., Ohn, I., & Kim, Y. O. (2017). Leave-one-out Bayesian model averaging for probabilistic ensemble forecasting. *Communications for Statistical Applications and Methods*, *24*(1), 67–80. https://doi.org/10.5351/CSAM.2017.24.1.067

Lekar, S., Shumeiko, D., Lagodiienko, V., & Valerii, N. (2019). CONSTRUCTION OF BAYESIAN NETWORKS IN PUBLIC ADMINISTRATION OF THE ECONOMY. *International Journal of Civil Engineering and Technology (IJCIET)*, *10*(04), 1379–1384.

Li, Y., Wu, Z., He, H., & Yin, H. (2022). Probabilistic subseasonal precipitation forecasts using preceding atmospheric intraseasonal signals in a Bayesian perspective. *Hydrology and Earth System Sciences*, *26*(19), 4975–4994. https://doi.org/10.5194/hess-26-4975-2022

Li, Z., Peck, K. K., Brennan, N. P., Jenabi, M., Hsu, M., Zhang, Z., Holodny, A. I., & Young, R. J. (2013). Diffusion tensor tractography of the arcuate fasciculus in patients with brain tumors: Comparison between deterministic and probabilistic models. *Journal of Biomedical Science and Engineering*, *06*(02), 192–200. https://doi.org/10.4236/jbise.2013.62023

Liang, Z., Wang, D., Guo, Y., Zhang, Y., & Dai, R. (2013). Application of Bayesian Model Averaging Approach to Multimodel Ensemble Hydrologic Forecasting. *Journal of Hydrologic Engineering*, *18*(11), 1426–1436. https://doi.org/10.1061/(asce)he.1943-5584.0000493

Marin, J. M., Mengersen, K., & Robert, C. P. (2005). Bayesian Modelling and Inference on Mixtures of Distributions. *Handbook of Statistics*, *25*, 459–507. https://doi.org/10.1016/S0169-

7161(05)25016-2

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Journal of Clinical Epidemiology*, *62*(10), 1006–1012. https://doi.org/10.1016/j.jclinepi.2009.06.005

Najib, M. K., Nurdiati, S., & Sopaheluwakan, A. (2022). Copula-based joint distribution analysis of the ENSO effect on the drought indicators over Borneo fire-prone areas. *Modeling Earth Systems and Environment*, *8*, 2817–2826. https://doi.org/10.1007/s40808-021-01267-5

Natalia, C., Surbakti, I. P., & Oktavia, C. W. (2020). Integrated ANP and TOPSIS Method for Supplier Performance Assessment. *Jurnal Teknik Industri*, *21*(1), 34–45. https://doi.org/10.22219/jtiumm.vol21.no1.34-45

Nguyen-Le, D. H., Tao, Q. B., Nguyen, V. H., Abdel-Wahab, M., & Nguyen-Xuan, H. (2020). A data-driven approach based on long short-term memory and hidden Markov model for crack propagation prediction. *Engineering Fracture Mechanics*, *235*. https://doi.org/10.1016/j.engfracmech.2020.107085

Nurdiati, S., Bukhari, F., Julianto, M. T., Najib, M. K., & Nazria, N. (2021). Heterogeneous Correlation Map Between Estimated ENSO And IOD From ERA5 And Hotspot In Indonesia. *Jambura Geoscience Review*, *3*(2), 65–72. https://doi.org/10.34312/jgeosrev.v3i2.10443

Nurdiati, S., Sopaheluwakan, A., Julianto, M. T., Septiawan, P., & Rohimahastuti, F. (2021). Modelling and analysis impact of El Nino and IOD to land and forest fire using polynomial and generalized logistic function: cases study in South Sumatra and Kalimantan, Indonesia. *Modeling Earth Systems and Environment*. https://doi.org/10.1007/s40808-021-01303-4

Olson, R., Fan, Y., & Evans, J. P. (2016). A simple method for Bayesian model averaging of regional climate model projections: Application to southeast Australian temperatures. *Geophysical Research Letters*, *43*(14), 7661–7669. https://doi.org/10.1002/2016GL069704

Patrick, B. (2012). Probability and measure. In *From Measures to Itô Integrals* (Anniversar). Wiley.

Pendharkar, P. C., Subramanian, G. H., & Rodger, J. A. (2005). A probabilistic model for predicting software development effort. *IEEE Transactions on Software Engineering*, *31*(7), 615–624. https://doi.org/10.1109/TSE.2005.75

Petticrew, M., & Roberts, H. (2008). Systematic Reviews in the Social Sciences: A Practical Guide. In *Systematic Reviews in the Social Sciences: A Practical Guide*. https://doi.org/10.1002/9780470754887

Pickering, C., & Byrne, J. (2014). The benefits of publishing systematic quantitative literature reviews for PhD candidates and other early-career researchers. *Higher Education Research and Development*, *33*(3), 534–548. https://doi.org/10.1080/07294360.2013.841651

Ribeiro, L. M., Viegas, D. X., Almeida, M., McGee, T. K., Pereira, M. G., Parente, J., Xanthopoulos, G., Leone, V., Delogu, G. M., & Hardin, H. (2019). Extreme wildfires and disasters around the world: Lessons to be learned. *Extreme Wildfire Events and Disasters: Root Causes and New Management Strategies*, 31–51. https://doi.org/10.1016/B978-0-12-815721-3.00002-3

Röpnack, A., Hense, A., Gebhardt, C., & Majewski, D. (2013). Bayesian model verification of NWP ensemble forecasts. *Monthly Weather Review*, *141*(1), 375–387. https://doi.org/10.1175/MWR-D-11-00350.1

Samsuri, Anita Zaitunah, O. R. (2021). Analisis Kebutuhan Ruang Terbuka Hijau: Pendekatan Kebutuhan Oksigen. *Jurnal Silva Tropika*, *5*(1), 305–320.

Santriwati, S., Halide, H., & Hasanuddin, H. (2021). Faktor Osean – Atmosfer untuk Memprediksi Titik Panas (Hostspot) di Wilayah Asia Tenggara Bagian Selatan. *Jurnal Geocelebes*, *5*(2), 116–130. https://doi.org/10.20956/geocelebes.v5i2.13454

Sharifahmadian, E., & Latifi, S. (2013). Water environment risk prediction using Bayesian network. *Conference Proceedings - IEEE SOUTHEASTCON*. https://doi.org/10.1109/SECON.2013.6567437

Siegert, S., Stephenson, D. B., Sansom, P. G., Scaife, A. A., Eade, R., & Arribas, A. (2016). A Bayesian framework for verification and recalibration of ensemble forecasts: How uncertain is NAO predictability? *Journal of Climate*, *29*(3), 995–1012. https://doi.org/10.1175/JCLI-D-15-0196.1

Suhermat, M., Dimyati, M., Supriatna, S., & Martono, M. (2021). Impact of Climate Change on Sea Surface Temperature and Chlorophyll-a Concentration in South Sukabumi Waters. *Jurnal Ilmu Lingkungan*, *19*(2), 393–398. https://doi.org/10.14710/jil.19.2.393-398

Wahyuni, H., & Suranto, S. (2021). Dampak Deforestasi Hutan Skala Besar terhadap Pemanasan Global di Indonesia. *JIIP: Jurnal Ilmiah Ilmu Pemerintahan*, *6*(1), 148–162. https://doi.org/10.14710/jiip.v6i1.10083

Wang, X., Bradlow, E. T., & Wainer, H. (2002). a General Bayesian Model for Testlets: Theory and Applications. In *ETS Research Report Series* (Vol. 2002, Issue 1). https://doi.org/10.1002/j.2333-8504.2002.tb01869.x

Xiao, Y., & Watson, M. (2019). Guidance on Conducting a Systematic Literature Review. *Journal of Planning Education and Research*, *39*(1), 93–112. https://doi.org/10.1177/0739456X17723971

Xie, L. (2022). The analysis and forecasting COVID-19 cases in the United States using Bayesian structural time series models. *Biostatistics and Epidemiology*, *6*(1), 1–15. https://doi.org/10.1080/24709360.2021.1948380

Zhang, Y. M., Wang, H., Bai, Y., Mao, J. X., & Xu, Y. C. (2022). Bayesian dynamic regression for reconstructing missing data in structural health monitoring. *Structural Health Monitoring*, *21*(5), 2097–2115. https://doi.org/10.1177/14759217211053779