E-ISSN: 2987-1069
P-ISSN: 2987-1018
Siber Journal of Advanced Multidisciplinary (SJAM)
Siber Nusantara Research
+62 812 1046 7572 https://research.e-siber.org research.siber.info@gmail.com Jalan Marina Indah Raya No. 1 Pantai Indah Kapuk, Kamal Muara, Penjaringan, Jakarta Utara, Indonesia.

DOI: https://doi.org/10.38035/sjam.v2i2
Received: June 22nd, 2024, **Revised:** July 03rd, 2024, **Publish:** July 17th, 2024

Utilization of Big Data Analytics in Understanding Climate Change Patterns: Latest Trends and Findings on Climate Patterns in Indonesia

Abu Sopian[1], Dedi Setiadi[2], Novita Fitriani[3]
[1]Universitas Mohammad Husni Thamrin Jakarta, Jakarta Timur, Indonesia, abu.sopian355@gmail.com
[2]Universitas Mohammad Husni Thamrin Jakarta, Jakarta Timur, Indonesia, ddstd2304@gmail.com
[3]Universitas Mohammad Husni Thamrin Jakarta, Jakarta Timur, Indonesia, justnovii12@gmail.com

Corresponding Author: abu.sopian355@gmail.com[1]

Abstract: Climate change has become a global concern. In Indonesia, a deep understanding of climate change patterns is increasingly important in planning effective mitigation and adaptation measures. In this study, we explore the utilization of big data analytics to understand climate change patterns and identify trends findings related to climate patterns in Indonesia based on variables such as temperature, humidity, rainfall, and wind speed. Data used were collected from satellite observations, weather stations from 2010 to 2020. Purpose this study aims to explore the utilization of big data analytics in understanding climate change patterns in Indonesia, with a focus on identifying trends and recent findings in climate patterns. Methodology–We applied big data analytics techniques such as descriptive statistical analysis, spatial regression modeling, and cluster analysis to identify climate change patterns and recent trends in Indonesia. Findings–Through big data analysis, we successfully identified significant climate change patterns in Indonesia. Practical implications the findings of this study can provide a better understanding of climate change dynamics in Indonesia, serving as a basis for decision-making in natural resource management, disaster risk mitigation, and climate change adaptation strategies at both regional and national levels and can serve as a reference for researchers.

Keyword: Big Data Analytics, Climate Change, Climate Patterns.

INTRODUCTION

Indonesia, with its unique geographic and climatic diversity, faces distinctive challenges in addressing the impacts of climate change. (Stanley, 2024) A comprehensive understanding of climate change patterns at both regional and national levels is crucial for making informed decisions to safeguard communities and the environment. The utilization of big data analytics presents an effective tool for understanding and anticipating climate change in Indonesia, particularly in planning effective mitigation and adaptation measures. This study delves into the application of big data analytics to unravel climate change patterns and identify recent trends and findings pertaining to climate patterns in Indonesia. The analysis encompasses a wide range of variables, including minimum temperature (°C), maximum temperature (°C), average temperature (°C), average humidity (%), rainfall (mm), sunshine duration (hours), maximum wind speed (m/s), direction of maximum wind speed (°), average wind speed (m/s), and predominant wind direction (°). Data encompassing the period from 2010 to 2020 was gathered from diverse sources, including satellite observations, weather stations, and global climate models.

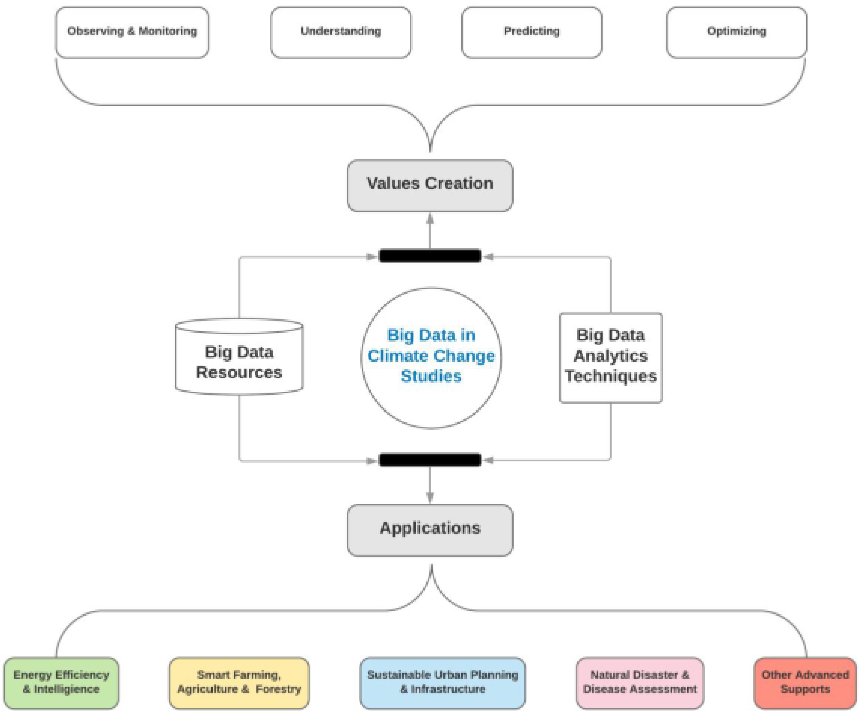
The issue addressed in this study focuses on exploring the utilization of big data analytics in understanding climate change patterns in Indonesia? This research aims to investigate the utilization of big data analytics in comprehending climate change patterns in Indonesia, with a specific emphasis on identifying trends and recent findings in climate patterns. This study provides a novel contribution to understanding the role of big data analytics in comprehending climate change patterns in Indonesia. The findings and methodologies proposed herein can serve as valuable references for researchers and practitioners interested in climate analysis, disaster risk mitigation, and climate change adaptation strategies in this region.

Literature Review

Big Data represents a vast collection of information originating from diverse sources and in various formats, generated rapidly. (Wang, Yang, Wang, Sherratt, & Zhang, 2020) The process of data analysis involves scrutinizing and transforming raw data into understandable information, while data science encompasses a range of analyses, programming tools, algorithms, predictive statistical analyses, and machine learning aimed at identifying and extracting patterns from raw data. Essentially, Big Data focuses on systematically analyzing, extracting, or handling data from datasets that are too large or complex to be efficiently processed by traditional data processing software. This necessitates scaling up through multiple nodes to process efficiently. In essence, Big Data can be explained by the 5Vs: volume (amount of data), velocity (speed of incoming data), variety (diversity of data formats), veracity (data accuracy), and value (benefits of data). The purpose of utilizing Big Data is to provide tools for data management and analysis in the face of continuously increasing data volumes. (Sebestyén, Czvetkó, & Abonyi, 2021)

Climate change has emerged as one of the most pressing global challenges that demand immediate attention. (Seddon, Chausson, Berry, Girardin, Smith, & Turner, 2020) Addressing this multifaceted issue requires a holistic and integrated approach that draws upon knowledge from diverse disciplines. This comprehensive understanding serves as the foundation for designing effective solutions that account for the complexities, uncertainties, and interconnectedness of the various systems involved. (Sebestyén, Czvetkó, & Abonyi, 2021)

According to a 2019 study by Hossein Hassani, Xu Huang, and Emmanuel Silva titled "Big Data and Climate Change," which reviewed over 100 research papers, energy efficiency is currently the most prominent area where big data analysis is being applied in climate change research. Smart agriculture and natural disaster assessment follow closely behind. While there's extensive research on big data's role in smart cities, its application for climate change sustainability is still in its early stages. Disease management also hasn't received significant exploration yet. (Hassani, Huang, & Silva, 2019) In the following section, we'll delve into the framework of big data utilization within climate change studies:



Source: Research Results (Hassani, Huang, & Silva, 2019)
Figure 1. The framework of big data utilization within climate change studies

Based on the research conducted by Jie Lu, Anjin Liu, Yiliao Song, and Guangquan Zhang in their paper titled "Data-driven decision support under concept drift in streamed big data" in 2019, this position paper discusses the foundational framework and applicable techniques in streaming big data and concept drift deviation for D3M (Data-driven Decision Making). This study initially establishes the technical framework for real-time D3M based on the concept of drift and details the characteristics of high-volume streaming data. Methodologies and key approaches for detecting concept drift and supporting D3M are highlighted and presented. Finally, further research directions, methods, and related procedures for utilizing streaming data to support decision-making in concept drift environments are identified. We hope the observations in this paper will support researchers and professionals in understanding the fundamentals and research directions of D3M in the stream big data environment. (Lu, Liu, Song, & Zhang, 2019)

Based on the research conducted by Bogdan Bochenek and Zbigniew Ustrnul in their paper titled "Machine Learning in Weather Prediction and Climate Analyses - Applications and Perspectives" in 2022, an analysis has been carried out on 500 most relevant scientific articles published since 2018, regarding machine learning methods in the field of climate and numerical weather prediction using the Google Scholar search engine. The most commonly interesting topics in the abstracts have been identified, and some of them have been examined in detail: in numerical weather prediction research - photovoltaic and wind energy, atmospheric physics and processes; in climate research - parameterization, extreme events, and climate change. With the created database, it is also possible to extract the most commonly examined meteorological fields (wind, rainfall, temperature, pressure, and radiation), methods (Deep Learning, Random Forest, Artificial Neural Networks, Support Vector Machines, and XGBoost), and countries (China, US, Australia, India, and Germany) in this topic. By critically reviewing the literature, the authors attempt to predict the direction of future research in this field, with the main conclusion being that machine learning methods will be a major feature in future weather forecasting. (Bochenek & Ustrnul, 2022)

Based on the research conducted by James H. Faghmous and Vipin Kumar in their paper titled "A Big Data Guide to Understanding Climate Change" in 2014, this article introduces data science readers to the challenges and opportunities in mining large climate datasets, with an emphasis on the difference between climate data mining and traditional big data approaches. We focus on the data, methods, and challenges of implementation that must be overcome for big data to fulfill its promise in climate science applications. More importantly, we highlight research indicating that relying solely on traditional big data techniques will yield questionable findings, and instead, we propose a data science paradigm guided by theory that utilizes scientific principles to constrain big data techniques and the process of interpreting results to extract accurate insights from large climate datasets. (Faghmous & Kumar, 2014)

Based on the literature review above, it can be concluded that research on the Utilization of Big Data Analytics in Understanding Climate Change Patterns remains an intriguing topic for further investigation. (Bibri, 2019) Considering the research on Big Data Analytics indicates that a data-driven approach in exploring the utilization of big data analytics in understanding climate change patterns in Indonesia, with a focus on identifying trends and recent findings in climate patterns. By leveraging big data analysis, we can identify more effective and efficient solutions to address the impacts of climate change. However, challenges persist in terms of integrating complex data and reducing uncertainty in future climate predictions. Cross-disciplinary collaboration and sustainable investment in data infrastructure are needed to optimize adaptation strategies in the future. By leveraging big data analytics to unravel climate change patterns and identify recent trends and findings pertaining to climate patterns in Indonesia based on the aforementioned variables, we aim to gain a comprehensive understanding of local climate change patterns. This knowledge will be instrumental in devising effective mitigation and adaptation strategies.

METHOD

In this study, we utilized climate data from various sources including satellite observations, weather stations, and global climate models. We applied big data analytics techniques such as descriptive statistical analysis and cluster analysis to identify climate patterns and trends in Indonesia. The research process was structured within a research framework comprising identification, problem formulation, data collection, data analysis, and conclusion stages.

In the identification stage, we identified the research topic related to exploring big data for climate change. Subsequently, in the problem formulation stage, based on the identification, we formulated the research problem addressed in this study: how to explore the utilization of big data analytics in understanding climate change patterns in Indonesia? The next stage of the research was data collection, where we gathered data through literature review from research journals, while large climate data sources were obtained from open data repositories such as Kaggle. The collected data included climate data in Indonesia with variables such as minimum temperature (°C), maximum temperature (°C), average temperature (°C), average humidity (%), rainfall (mm), sunshine duration (hours), maximum wind speed (m/s), direction of maximum wind speed (°), average wind speed (m/s), and predominant wind direction (°) from 2010 to 2020.

Subsequently, the data analysis stage involved using big data analysis techniques for data processing and manipulation. This encompassed various actions such as data cleaning, manipulating data structures, converting data formats, and extracting relevant climate patterns and trends from the available data. The final stage was the conclusion stage, where we summarized the overall findings to address the research problem and reinforce the discovered results in this study.

RESULTS AND DISCUSSION

Based on the research findings obtained through big data analysis, we have successfully identified significant climate patterns and trends in Indonesia. These include monthly average rainfall, annual average rainfall for the years 2010-2020, regional annual average temperature changes for the years 2010-2020, Average Temperature & Rainfall, Top 10 Average Rainfall Provinces, Top 10 High Temperature Provinces, and Monthly Sunshine Duration (hours). The data used are sourced from Kaggle as secondary data. Below is the climate data in Indonesia from 2010 to 2020, comprising 589,265 data records:

	date	Tn	Tx	Tavg	RH_avg	RR	ss	ff_x	ddd_x	ff_avg	ddd_car	station_id
0	01-01-2010	21.4	30.2	27.1	82.0	9.0	0.5	7.0	90.0	5.0	E	96001
1	02-01-2010	21.0	29.6	25.7	95.0	24.0	0.2	6.0	90.0	4.0	E	96001
2	03-01-2010	20.2	26.8	24.5	98.0	63.0	0.0	5.0	90.0	4.0	E	96001
3	04-01-2010	21.0	29.2	25.8	90.0	0.0	0.1	4.0	225.0	3.0	SW	96001
4	05-01-2010	21.2	30.0	26.7	90.0	2.0	0.4	NaN	NaN	NaN	NaN	96001
...	...	...	...	...	...	...	...	...	...	...	...	...
589260	27-12-2020	25.2	31.2	29.2	74.0	0.0	1.4	4.0	280.0	2.0	C	97980
589261	28-12-2020	25.3	31.6	28.1	78.0	NaN	3.0	12.0	260.0	2.0	C	97980
589262	29-12-2020	24.6	32.3	28.4	81.0	NaN	6.5	5.0	260.0	2.0	SW	97980
589263	30-12-2020	25.2	32.6	28.4	80.0	0.0	2.4	7.0	260.0	2.0	C	97980
589264	31-12-2020	24.3	32.0	26.7	86.0	26.6	5.8	7.0	350.0	2.0	C	97980

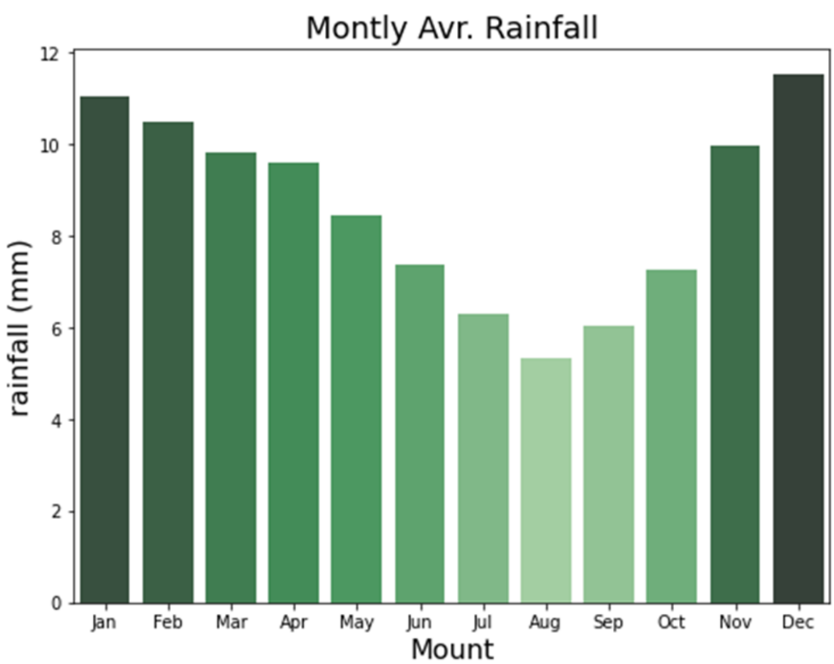
589265 rows × 12 columns

Source: Research Results

Figure 2. The climate data in Indonesia from 2010 to 2020, comprising 589,265 data records

1) Average Rain Mountly

Based on the results of big data analysis, the highest monthly average rainfall occurs in December and January, while the lowest occurs in August and September.

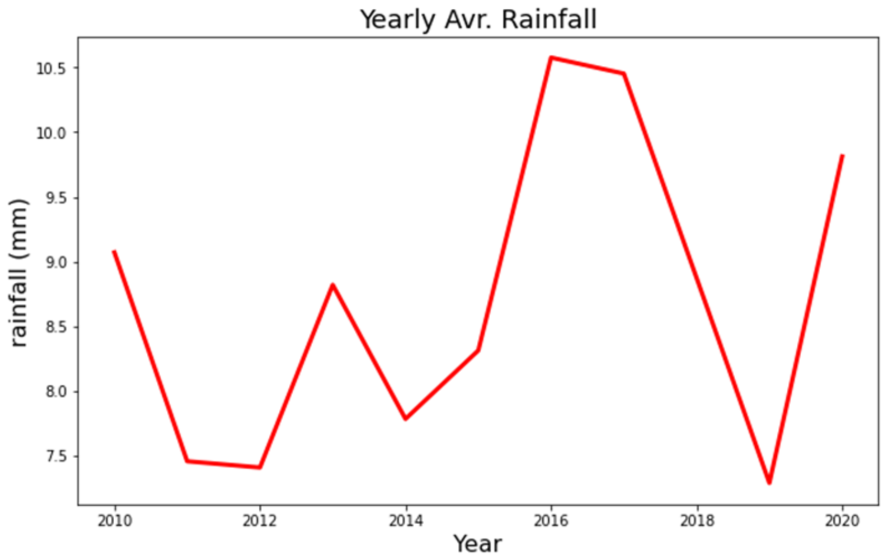


Source: Research Results

Figure 3. The highest monthly average rainfall occurs in December and January

2) Average Rain Yearly 2010-2020

Based on the results of big data analysis, the annual average rainfall from 2010 to 2020 shows that the highest average rainfall occurred in the year 2016.

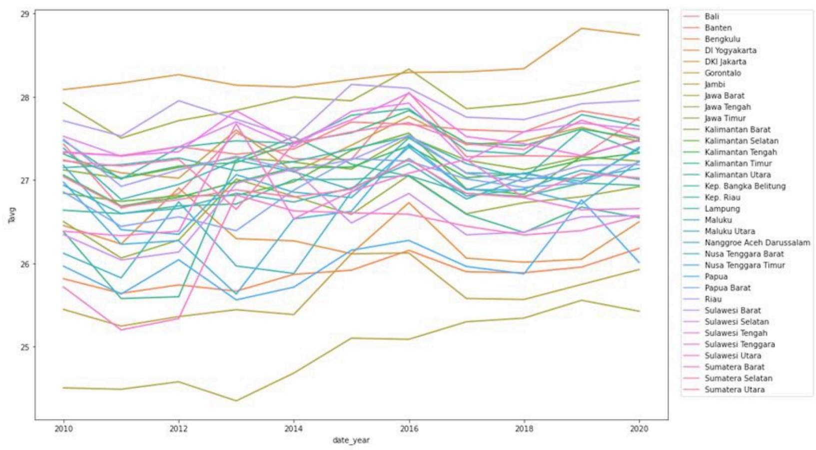


Source: Research Results

Figure 4. The annual average rainfall from 2010 to 2020

3) Regional annual average temperature change 2010-2020

Meanwhile, the regional annual average temperature changes from 2010 to 2020 indicate an increase in temperature change in the year 2019, with the highest province being Bali Province.

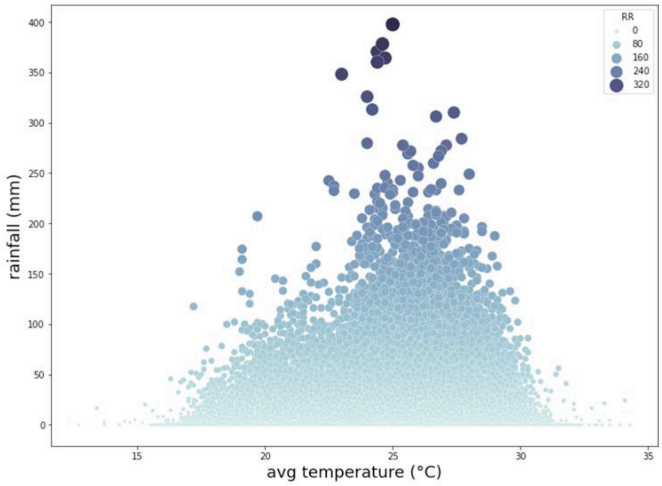


Source: Research Results

Figure 5. The regional annual average temperature changes from 2010 to 2020

4) Avg Temperature & Rainfall

The average temperature is 28°C, and the highest rainfall is 320mm.



Source: Research Results

Figure 6. The average temperature is 28°C, and the highest rainfall is 320mm

5) Top 10 Avg. Most Rainly Province

Top 10 provinces with the highest rainfall average. West Sumatra Province is the highest.

Table 1. Top 10 Provinces with the Highest Rainfall Average

Province_Name	
Sumatera Barat	15.065.508
Sulawesi Selatan	11.214.450
Papua	11.190.000
Kalimantan Barat	10.816.452
Bengkulu	10.682.361
Kalimantan Tengah	10.295.150
Kalimantan Utara	9.707.385
Jawa Barat	9.680.245
Sumatera Selatan	9.418.426
Sumatera Utara	9.310.504

6) Top 10 Avg. High Temp Province

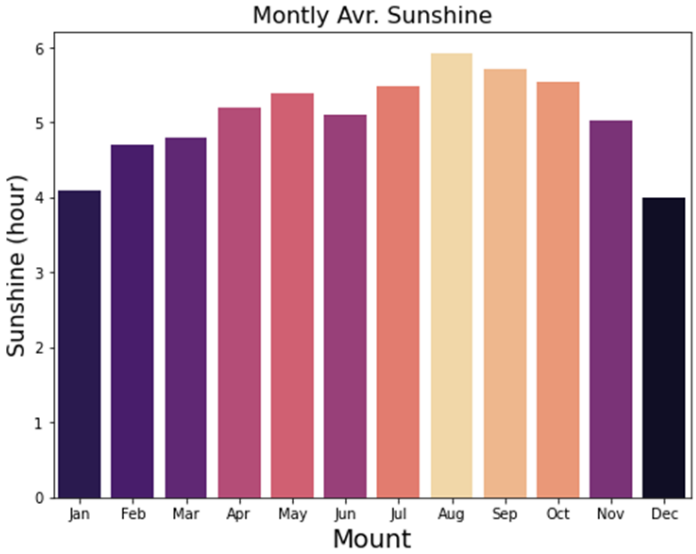
Top 10 Provinces with High Temperatures. The province with the highest temperature is DKI Jakarta.

Table 2. Top 10 Provinces with High Temperatures

Province_Name	
DKI Jakarta	28.285858
Jawa Tengah	27.930590
Sulawesi Barat	27.805781
Sulawesi Tengah	27.536687
Banten	27.522975
Kep. Riau	27.499683
Sulawesi Tenggara	27.491315
Kalimantan Timur	27.413652
Sumatera Selatan	27.360074
Gorontalo	27.349468

7) Monthly Avg. Duration of Sunshine (hour)

The highest monthly average sunshine duration is in August.



Source: Research Results

Figure 7. The highest monthly average sunshine duration is in August

CONCLUSION

Based on the research findings, it can be concluded that, through big data analysis, the utilization of big data analytics can be explored to understand climate change patterns in Indonesia by examining significant climate patterns and trends, including monthly average rainfall, annual average rainfall for the years 2010-2020, regional annual average temperature changes for the years 2010-2020, Average Temperature & Rainfall, Top 10 Average Rainfall Provinces, Top 10 High Temperature Provinces, and Monthly Sunshine Duration (hours). The results of this research can further be used as a foundational contribution for predictive research in decision-making regarding natural resource management, disaster risk mitigation, and climate change adaptation strategies at regional and national levels.

Limitations in the study include the availability of quality large-scale data and the need for data normalization for effective data utilization. Furthermore, this study did not predict climate change by observing climate patterns and trends in the future, but only explored Indonesian climate big data to observe patterns and trends from 2010 to 2020. Therefore, future research could delve into deep learning topics for predicting climate patterns and trends in the future.

REFERENCE

Bibri, S. E. (2019). On the sustainability of smart and smarter cities in the era of big data: an interdisciplinary and transdisciplinary literature review. *Journal of Big Data*, 2-64.

Bochenek, B., & Ustrnul, Z. (2022, January). Machine Learning in Weather Prediction and Climate Analyses-Applications and Perspectives. *Atmosphere, XIII*, 1-16.

Faghmous, J. H., & Kumar, V. (2014, September). A big data guide to understanding climate change. *Mary Ann Liebert, inc, II*(3), 155-163.

Hassani, H., Huang, X., & Silva, E. (2019, February). Big Data and Climate Change. *Big Data and Cognitive Computing MDPI, III*, 1-17.

Lu, J., Liu, A., Song, Y., & Zhang, G. (2019, October). Data-driven decision support under concept drift in streamed big data. *Complex & Intelligent Systems, VI*, 157-163.

Sebestyén, V., Czvetkó, T., & Abonyi, J. (2021, March). The Applicability of Big Data in Climate Change Research: The Importance of System of Systems Thinking. *Frontiers in Environmental Science, IX*, 1-26.

Seddon, N., Chausson, A., Berry, P., Girardin, C. A., Smith, A., & Turner, B. (2020). Understanding the value and limits of nature-based solutions to climate change and other global challenges. *Philosophical Transactions of the Royal Society, 375*(1794), 1-12.

Stanley, T. (2024). Carbon 'known not grown': Reforesting Scotland, advanced measurement technologies, and a new frontier of mitigation deterrence. *Environmental Science and Policy, 151*, 1-11.

Wang, J., Yang, Y., Wang, T., Sherratt, S. R., & Zhang, J. (2020). Big Data Service Architecture: A Survey. *Journal of Internet Technology, 21*(2), 393-405.