Recent Trends of Big Data in Precision Agriculture: a Review

A I Sourav*, A W R Emanuel

Magister Informatika, Universitas Atma Jaya Yogyakarta, Yogyakarta, Indonesia, 55281.
*e-mail: 195303060@students.uajy.ac.id

Abstract. Recent developments in the field of technology have led to a renewed interest in the field of smart agriculture. The current smart agricultural system produces and depends on large amounts of data, yet, it is hard to process the vast amounts of data using traditional data analysis systems. Big Data technologies have attracted much attention among researchers due to their potential to handle large amounts of data. Thus, for numerous possibilities and powerful data processing capabilities, Big Data continues to become a hot topic of research in the agricultural field. This is indicated by the existing several pilot projects as well as various investigations that have been performed in the Big Data area. Therefore, this study attempts to review the recent trends of Big Data technology in the field of precision agriculture. The research focuses on the latest applications of Big Data technology in multi-disciplinary agriculture domains. A total of 25 recently published studies about Big Data in agric-areas were selected for the review process after proper screening. The selected papers were categorized according to their fields of study, purposes, methods applied, as well as their contributions in the agri-field. The findings indicated that Big Data application areas are expanding, and the impacts are significant in precision agriculture.

1. Introduction

The world’s population is increasing persistently. According to the United Nations, the current world population is 7.7 billion and it is estimated to grow around 8.5 billion by the year 2030, and 9.7 billion in 2050 [1]. With this massive population growth, one of the utmost challenges is food production. Therefore, agriculture is the only way to produce adequate food that will be able to feed this enormous world population. However, the traditional methods of farming are inadequate for providing enough food. Due to the continuously expanding population, the resources required for agriculture such as; land, freshwater, and energy continue to get limited day by day and barriers like climate change effects and rapid urbanization cause a reduction in food production. It is, therefore, emergent to introduce new methods and techniques in the field of agriculture. The technological revolution of the 21st century has changed the traditional concepts in almost every primary sector of human life, one of which is agriculture. The idea of agriculture has been altered a lot in the last few decades. New technologies are introduced to enhance agricultural production. Applying the latest technical concepts like cloud computing [2], Internet of Things (IoT) [2], Big Data [3], data mining [4] and artificial intelligence helps in shaping and advancing agricultural activities. More so, the use of these technologies helps farmers to make smart decisions and actions to enhance farming processes.

Recently, the Internet of Things (IoT) and wirelessly connected devices have opened a door for several possibilities to solve complex problems of agriculture. By deploying sensors and machinery
across agricultural farms, farmers can keep track of valuable data like temperature, humidity, soil conditions, and the water level in the soil [4]. The amount of combined data from multiple farms is huge. Therefore, sensors can be used for data collection as the first stage of the tasks carried out at the farms. However, executing the immense amounts of data to meaningful information requires a systematic process. At this stage, the concept of Big Data analysis in farming activities becomes a necessity. Big Data is not only applied in farming processes but also the entire supply chain. The processed data using Big Data applications gives farmers a precise insight into the real-time condition of the farm, which helps them to make smart decisions [5]. Big Data in the agricultural sectors and precision agriculture is comparatively a new concept. It is still at an early development stage. Although the number of researches in the Big Data field is emerging, quality scientific publications are still limited to the agriculture topic and the future possibilities of Big Data in this sector are not yet well determined [5].

Thus, this review aims to go through the recent research activities conducted in the last five years to explore the current trends of Big Data in the agricultural sector. From the research perspective, the study addressed two research questions; (1) What are the current roles of Big Data in precision agriculture? (2) What are the shared contributions of Big Data in the agricultural sector? The review answers these two research questions based on extracting the recent trends, application fields, tools method applied, as well as the contributions of using Big Data technology in agri-area. The outcome of this study indicates that Big Data has gained much attention and popularity among the researchers to be applied in agricultural sectors. It is expected that the findings of this study will help future researchers to conduct further researches in the related field.

2. Literature Review
It is challenging to define Big Data with a fixed definition. According to Furht and Villanustre [6], Big Data is a complex set of data set, which is difficult to process using traditional data management and processing technologies within a tolerable elapsed time. The authors defined Big Data according to three main characteristics; i.e., volume, velocity, and variety. Volume refers to the size of data, velocity refers to the various processes of transferring the data including time and latency of data handling, and variety refers to the different formats of data such as text, image, video, audio, digital documents, and emails. Data can be structured, unstructured, or semi-structured [7]. Xindong Wu, Xingquan Zhu et al. [8] suggested three key characteristics of Big Data; i.e., (1) heterogeneous and massive diverse data source, (2) autonomous with decentralized control, (3) composite and expanding, all the in data and knowledge association. In situations where traditional data techniques are slow to respond, Big Data technology can perform with better efficiency [7]. Due to its compelling advantages, Big Data is used in various industries such as smart grid, e-health, Internet of Things (IoT), public utilities, transportation and logistics, governance and political services, etc [7].

The emergence of diverse technologies is shifting traditional farming practices. The term ‘Precision Agriculture’ was introduced after the rapid evolution of technology in the last few decades. Lindblom et al. [9] and Li et al. [10] defined precision agriculture as a facet of site-specific crop management, which represents a new kind of farming management concept rather than the traditional concepts that include observation, monitoring, and precise measurement of factors in food production. Big Data assists both farmers, policymakers, and people involved in the food production and supply chain to manage decision making, administration planning, and better processing [11]. Many related efforts include Big Data platforms and applications focusing on various problems in the agricultural sector such as; supply chain management [12], distribution channel [13], productivity improvement [14], and monitoring environmental factors [15].

3. Methodology
The bibliographic analysis of this study includes three steps i.e.; (1) gathering relevant studies, (2) filtering the works that are directly related to the topic of the study, (3) detailed investigation of related articles. Figure 1 illustrates the process of the literature screening process. In the first step, a search for
relevant works were performed through renowned scientific databases, mainly IEEE Explore, ScienceDirect, and Google Scholar – a web scientific indexing service. The search query used specific keywords of: “Big Data” AND [“Precision Agriculture” OR “Smart Farming” OR “Agriculture”].

To limit the publication range, only publications from 2016 to 2020 were considered for the review. For better-quality information, the authors filtered the collected papers according to SCImago Journal Ranking (SJR indicator). The publications from journals bearing a rank range of Q1 to Q4 were considered as good sources of information. The publications without any journal ranking were not considered.

Figure 1: Literature screening process

After the filtering process, the authors considered only 25 papers. The chosen studies met the authors’ requirements and showed the actual relation of Big Data in the field of precision agriculture. Finally, in the analysis phase, the authors analyzed the chosen papers considering their purposes, methods, platforms, tools used, their fields of application, and impacts achieved to determine the research questions mentioned before.

4. Result and Discussion

To answer the research questions, the authors run the review process in two phases. In the first phase, the selected papers were arranged based on different domains, namely; the purpose of the study, method or platform used, tools used, and the agri-areas of the papers.

Table 1. Arrangement of recent studies according to purpose, method/platform/tools used, and application fields

| No. | Purpose | Method, Platform, Tools | Agri-area | Ref. |
|-----|---------|-------------------------|-----------|-----|
| 1   | To generate a better decision-making report. | Hadoop framework | Distribution channel | [13] |
| 2   | To examine environmental changes and analyze cultivation techniques | Internet of Things (IoT), 3D cluster correlation, farmers’ behavior analysis | Environment monitoring | [15] |
| 3   | To increase IoT in the agricultural industry and reduce human interaction. | IoT, Cloud computing and Big Data | IoT monitoring system | [16] |
| 4   | To introduce a fault analysis system for agricultural machinery. | IoT and Big Data platform | Fault analysis of agricultural machinery | [17] |
| 5   | To predict the garlic price trends. | Support vector machine (SVM) model and ARIMA-SVM model based on Big Data | Price prediction | [18] |
| No. | Purpose                                                                 | Method, Platform, Tools                                                                 | Agri-area                                                                 | Ref. |
|-----|-------------------------------------------------------------------------|-----------------------------------------------------------------------------------------|----------------------------------------------------------------------------|------|
| 6   | To establish an environment-production model and build a system of crop growth environment perception and intelligent decision. | Zigbee protocol, wireless sensor node network                                          | Environment perception and intelligent decision                            | [19] |
| 7   | To predict rainfall.                                                    | Regression model and neural network models                                              | Weather forecasting                                                       | [20] |
| 8   | To develop a soil analysis system and understand the problems related to agricultural lands. | Hadoop framework, 3d reconstruction, statistical analysis                              | Soil analysis                                                             | [21] |
| 9   | To propose and implement Big Data analytics architecture for an Agro advisory system. | NoSQL database, multiple linear regression                                              | Agro-advisory system, crop yield prediction                               | [22] |
| 10  | To build a Big Data architecture based on the user’s needs.             | Profile-based architecture                                                              | Smart decision making                                                     | [23] |
| 11  | To identify crop diseases based on previously available data.           | Hadoop and Hive tools                                                                   | Crop disease identification                                               | [24] |
| 12  | To allow smart management of water and fertilizer in agriculture.       | Internet of Things (IoT)                                                                | Smart irrigation and fertilizer management                                 | [25] |
| 13  | To estimate the environmental impact of agriculture.                    | Geospatial analysis using an online software platform                                   | Environmental impact of agriculture                                        | [11] |
| 14  | To establish a framework to support farmers.                            | Web-based application                                                                  | Sustainable agricultural development                                       | [26] |
| 15  | To establish a database and allow traceability.                         | ICT platform                                                                            | Sugarcane harvesting operations                                           | [14] |
| 16  | To establish a smart decision system.                                   | Frame designation and design process                                                    | The agricultural intelligent decision system                              | [27] |
| 17  | To improve sustainability management in supply chain design for the valorization of agriculture wastes. | Life-cycle assessment (LCA) method                                                      | Supply chain management                                                   | [12] |
| 18  | To maximize crop production.                                           | Data mining techniques (PAM, CLARA, and DBSCAN)                                         | Crop production and environmental factors analysis.                       | [28] |
| 19  | To increase crop production and control of agricultural cost.           | IoT and cloud-based Big Data analysis                                                   | Fertilizer requirements, crop analysis, market and stock requirements for the crop | [29] |
| 20  | To reduce disaster risks in agriculture                                 | Geospatial (Big) data processing through farm machinery telemetry, wireless             | Disaster risk management                                                  | [30] |
Table 1 illustrates the recent studies on Big Data applications in the agricultural sector. It also gives us insights into our first research question. According to Table 1 classification; recent Big Data studies cover a vast area of different applications in the agricultural sector. They serve diverse purposes such as generating valuable information from a large amount of data, forecasting, increasing production rate, reducing production cost and food loss, reducing disaster risks, introducing sustainability in the food supply chain, providing a smart decision, automating the farming process and reduce human interaction, fault analysis, fertilizer and water management, crop disease identification, etc. These purposes are difficult to solve by using traditional methods because of their complexity. Where Big Data technologies can handle these issues with better efficiency, faster responsiveness, and wide scalability. The methods, tools, and platforms for applying Big Data in the agricultural sector are also emerging rapidly. Finally, in the different domains of agri-area, Big Data is a high performing technology. Agricultural domains like precision farming, food security, supply chain, environment monitoring, agricultural zone, and land identification, fertilizing, disease identification and prevention, soil analysis, intelligent decision systems are using Big Data as a major supporting technology to enhance yield production. We predict that in the future, Big Data analytics will play some dominating role in the agricultural sector.

Table 2 gives insights about the second research question. It illustrates the contributions or achieved impacts of Big Data applications in the agricultural sector.

Table 2. Impact of Big Data analysis on the agricultural field

| References | Impact achieved |
|------------|-----------------|
| [11], [13], [14], [19], [20], [23], [28] | Intelligent decision-making support |
| [14], [20], [28], [29] | Better management and operation guidelines |
| [17] | Improves crop production |
| [18] | Demonstrates the practical significance of Big Data analysis technology in agricultural production. |
| [27] | Short-term garlic price prediction. |
| [11], [13], [14], [19], [20], [23], [28] | Improves the performance of the agricultural intelligent decision systems. |
Table 2 shows that recent studies found many exciting and powerful impacts of Big Data applications in different agri-areas. One of the most common contributions of using Big Data in agri-area is that it helps farmers and the stakeholders related to agri-activities to make an intelligent decision. It makes the decision-making process accurate and faster. In addition to that, it increases crop yield production rate and provides stakeholders with better managerial guidelines. With Big Data analytics, weather predictions like rainfall and storm forecasting become more accurate. The artificial neural network, along with Big Data technology can identify crop disease based on the previously stored massive amount of data, which reduces disaster and pollution risks in crop production. Big Data analytics is now becoming a mounting solution for market analysis and predicting the price of specific goods. It helps to make complex supply chain processes easier to handle. The applications of Big Data in the Agri sector also helps farmers and stakeholders to control the production cost. It also becomes a convenient solution for smart and automatic recommended systems. The findings observed in this study thus show that the current applications of Big Data have a significant impact on the agricultural field.

5. Conclusion
The traditional methods of farming are not adequate to produce food for the future world’s population. New technologies and techniques continue to appear to increase the production of crops and utilize resources precisely. With the technological advancement, the agricultural data has entered the era of Big Data. This article reviews recent studies on Big Data analysis in agriculture and concludes that these growing technologies will play a lead role in agricultural development in near future. 25 articles are selected mostly from an application and contribution perspective of Big Data in various agri-areas. This review shows how Big Data application areas in the agricultural sector are continuously emerging. The rapidly expanding application areas and noticeable contributions indicate that the future dependency of Big Data in the agricultural sector will increase. Advancements in technology will also make it easier and available for the mass application of Big Data applications in the future. This study, however, a preliminary exploratory study to picture the current trends and applications of Big Data in agricultural fields. This paper can be a knowledge base for future explorers in this area. It is expected that this study will help future researchers as a reference and will provide knowledge about the recent applications and contributions of Big Data in the agricultural area. As an extension of this work, the researchers want to explore more effective applications of Big Data in agri-areas in the future.

Acknowledgments
We would like to thank Universitas Atma Jaya Yogyakarta, department of Magister Informatika, International Affairs office, Kemitraan Negara Berkembang (KNB) scholarship program, and the government of Indonesia for supporting this work. We also thank our anonymous reviewers; whose fair and thorough feedback significantly improved this paper.

References
[1] United Nations, *World Population Prospects 2019*. Department of Economic and Social Affairs.
Population Division (2019).

[2] N. Ahmed, D. De, and I. Hussain, “Internet of Things (IoT) for Smart Precision Agriculture and Farming in Rural Areas,” IEEE Internet Things J., vol. 5, no. 6, pp. 4890–4899, 2018, doi: 10.1109/JIOT.2018.2879579.

[3] X. Pham and M. Stack, “How data analytics is transforming agriculture,” Bus. Horiz., vol. 61, no. 1, pp. 125–133, 2018, doi: 10.1016/j.bushor.2017.09.011.

[4] J. Muangprathub, N. Boonnam, S. Kajornkasirat, N. Lekbangpong, A. Wanichsombat, and P. Nillaor, “IoT and agriculture data analysis for smart farm,” Comput. Electron. Agric., vol. 156, no. January 2019, pp. 467–474, 2019, doi: 10.1016/j.compag.2018.12.011.

[5] S. Wolfert, L. Ge, C. Verdouw, and M. J. Bogaardt, “Big Data in Smart Farming – A review,” Agric. Syst., vol. 153, pp. 69–80, 2017, doi: 10.1016/j.agsy.2017.01.023.

[6] A. Oussous, F. Z. Benjelloun, A. Ait Lahcen, and S. Belfkih, “Big Data technologies: A survey,” J. King Saud Univ. - Comput. Inf. Sci., vol. 30, no. 4, pp. 431–448, 2018, doi: 10.1016/j.jksuci.2017.06.001.

[7] A. Kamilaris, A. Assumpcio, A. B. Blasi, M. Torrellas, and F. X. Prenafeta-Boldú, “Estimating the Environmental Impact of Agriculture by Means of Geospatial and Big Data Analysis: The Case of Catalonia,” in From Science to Society, 2018, pp. 39–48, doi: 10.1007/978-3-319-65687-8_4.

[8] J.-P. Belaud, N. Prioux, C. Vialle, and C. Sablayrolles, “Big data for agri-food 4.0: Application to sustainability management for by-products supply chain,” Comput. Ind., vol. 111, pp. 41–50, 2019, doi: https://doi.org/10.1016/j.compind.2019.06.006.

[9] T. A. Nielsen and K. L. Sørensen, “A comparative review of the state and advancement of Site-Specific Crop Management in the UK and China,” Land Use Policy, vol. 70, pp. 399–415, 2018, doi: 10.1016/j.landusepol.2017.08.040.

[10] J. Lindblom, C. Lundström, M. Ljung, and A. Jonsson, “Promoting sustainable intensification in precision agriculture: review of decision support systems development and strategies,” Precis. Agric., vol. 18, pp. 309–331, 2017, doi: 10.1007/s11119-016-9491-4.

[11] F. Tseng, H. Cho, and H. Wu, “Applying Big Data for Intelligent Agriculture-Based Crop Selection Analysis,” IEEE Access, vol. 7, pp. 116965–116974, 2019, doi: 10.1109/ACCESS.2019.2935564.

[12] S. Liu, L. Guo, H. Webb, X. Ya, and X. Chang, “Internet of Things Monitoring System of Modern Eco-Agriculture Based on Cloud Computing,” IEEE Access, vol. 7, pp. 37050–37058, 2019, doi: 10.1109/ACCESS.2019.2903720.

[13] D. Li, Y. Zheng, and W. Zhao, “Fault Analysis System for Agricultural Machinery Based on Big Data,” IEEE Access, vol. 7, pp. 99136–99151, 2019, doi: 10.1109/ACCESS.2019.2928973.

[14] B. Wang et al., “Research on hybrid model of garlic short-term price forecasting based on big data,” Comput. Mater. Contin., vol. 57, no. 2, pp. 283–296, 2018, doi: 10.32604/cmc.2018.03791.
[19] F. Hao, X. Luo, and C. Mu, “Research on Key Technologies of Intelligent Agriculture Based on Agricultural Big Data,” in 2016 International Conference on Smart City and Systems Engineering (ICSCSE), 2016, pp. 598–601, doi: 10.1109/ICSCSE.2016.0161.

[20] M. R. Bendre, R. C. Thool, and V. R. Thool, “Big Data in Precision Agriculture Through ICT: Rainfall Prediction Using Neural Network Approach,” in Proceedings of the International Congress on Information and Communication Technology, 2016, pp. 165–175, doi: 10.1007/978-981-10-0767-5_19.

[21] G. M. Alves and P. E. Cruvinel, “Big Data Environment for Agricultural Soil Analysis from CT Digital Images,” in 2016 IEEE Tenth International Conference on Semantic Computing (ICSC), 2016, pp. 429–431, doi: 10.1109/ICSC.2016.80.

[22] P. Shah, D. Hiremath, and S. Chaudhary, “Big Data Analytics Architecture for Agro Advisory System,” in 2016 IEEE 23rd International Conference on High Performance Computing Workshops (HiPCW), 2016, pp. 43–49, doi: 10.1109/HIPCW.2016.015.

[23] S. Lamrhari, H. Elghazi, T. Sadiki, and A. El Faker, “A profile-based Big data architecture for agricultural context,” in 2016 International Conference on Electrical and Information Technologies (ICEITT), 2016, pp. 22–27, doi: 10.1109/ICEITT.2016.7519585.

[24] R. Kaur, R. Garg, and H. Aggarwal, “Big data analytics framework to identify crop disease and recommendation a solution,” in 2016 International Conference on Inventive Computation Technologies (ICICT), 2016, vol. 2, pp. 1–5, doi: 10.1109/INVENTIVE.2016.7824791.

[25] P. Zhang, Q. Zhang, F. Liu, J. Li, N. Cao, and C. Song, “The Construction of the Integration of Water and Fertilizer Smart Water Saving Irrigation System Based on Big Data,” in 2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC), 2017, vol. 2, pp. 392–397, doi: 10.1109/CSE-EUC.2017.258.

[26] S. Mekruksavanich and T. Cheosuwan, “Visual Big Data Analytics for Sustainable Agricultural Development,” in 2018 International Joint Symposium on Artificial Intelligence and Natural Language Processing, iSAI-NLP 2018 - Proceedings, 2018, pp. 1–5, doi: 10.1109/iSAI-NLP.2018.8692910.

[27] J. Zhao and J. Guo, “Big data analysis technology application in agricultural intelligence decision system,” in 2018 IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), 2018, pp. 209–212, doi: 10.1109/ICCCBDA.2018.8386513.

[28] J. Majumdar, S. Naraseeyappa, and S. Ankalaki, “Analysis of agriculture data using data mining techniques: application of big data,” J. Big Data, vol. 4, no. 1, p. 20, 2017, doi: 10.1186/s40537-017-0077-4.

[29] S. Rajeswari, K. Suthendran, and K. Rajakumar, “A smart agricultural model by integrating IoT, mobile and cloud-based big data analytics,” in 2017 International Conference on Intelligent Computing and Control (I2C2), 2017, pp. 1–5, doi: 10.1109/I2C2.2017.8321902.

[30] T. Rzeznik et al., “Disaster risk reduction in agriculture through geospatial (Big) data processing,” ISPRS Int. J. Geo-Information, vol. 6, no. 8, p. 238, 2017, doi: 10.3390/ijgi6080238.

[31] K. Charvat et al., “Advanced Visualisation of Big Data for Agriculture as Part of Databio Development,” in IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, 2018, pp. 415–418, doi: 10.1109/IGARSS.2018.8517556.

[32] L. J. Young, M. Hyman, and B. R. Rater, “Exploring a Big Data Approach to Building a List Frame for Urban Agriculture: A Pilot Study in the City of Baltimore,” J. Off. Stat., vol. 34, no. 2, pp. 323–340, 2018, doi: 10.2478/jos-2018-0015.

[33] J. Granda-Cantuña, C. Molina-Colcha, S. Hidalgo-Lupera, and C. Valarezo-Varela, “Design and Implementation of a Wireless Sensor Network for Precision Agriculture Operating in API Mode,” in 2018 International Conference on eDemocracy eGovernment (ICDEG), 2018, pp. 144–149, doi: 10.1109/ICDEG.2018.8372346.

[34] S. Parvin et al., “Smart Food Security System Using IoT and Big Data Analytics,” in 16th
International Conference on Information Technology-New Generations (ITNG 2019), 2019, pp. 253–258, doi: 10.1007/978-3-030-14070-0_35.

[35] K. Perakis, F. Lampathaki, K. Nikas, Y. Georgiou, O. Marko, and J. Maselyne, “CYBELE – Fostering precision agriculture & livestock farming through secure access to large-scale HPC enabled virtual industrial experimentation environments fostering scalable big data analytics,” *Comput. Networks*, vol. 168, p. 107035, 2020, doi: https://doi.org/10.1016/j.comnet.2019.107035.