Application of Big Data and Business Analytics

This page intentionally left blank

Application of Big Data and Business Analytics

EDITORS

DR SNEHA KUMARI
Assistant Professor, Vaikunth Mehta National Institute of Cooperative Management, Maharashtra, India

DR K. K. TRIPATHY
I.E.S. Director, Vaikunth Mehta National Institute of Cooperative Management, Maharashtra, India

DR VIDYA KUMBHAR
Deputy Director, Symbiosis Institute of Geoinformatics, Symbiosis International (Deemed University), Maharashtra, India

emerald
PUBLISHING

United Kingdom – North America – Japan – India – Malaysia – China

Emerald Publishing Limited
Howard House, Wagon Lane, Bingley BD16 1WA, UK

First edition 2021

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-80043-885-9 (Print)
ISBN: 978-1-80043-884-2 (Online)
ISBN: 978-1-80043-886-6 (Epub)

ISOQAR certified Management System, awarded to Emerald for adherence to Environmental standard ISO 14001:2004.

ISOQAR REGISTERED
Certificate Number 1985
ISO 14001

INVESTOR IN PEOPLE

Acknowledgments

The editors of the book would like to express their gratitude to each and everyone who directly or indirectly supported us during this project. We would like to thank all for their contribution in the intellectual capital of the book. An extended thanks to all the authors for their hard work, diligence and dedication in serving the readers. Their contribution is appreciable.

Words fail to express our heartiest thanks and profound sense of gratitude to Dr Rameshwar Dubey for his commendable guidance, motivation and support. Our sincere thanks to Dr Urvashi Rathod and Dr Yogesh Patil for providing the support and motivation to carry out this project work. We would also like to thank the reviewers Dr Manisha Dinesh Bhosle, Dr Nisha Bharti and others for their constant support and help during this work.

We would like to express our gratitude to the Editorial Advisory Board members who helped us with their timely reviews and the team of Emerald who supported us constantly at every phase of the book.

We would like to thank the Almighty for guiding us with mature thoughts and decision-making for the success of the book. The support of family remains unparalleled in any success and the same goes here as well. We would like to thank our parents, family members and friends who helped us stay motivated during the entire course of the book. Everyone has been a great source of inspiration, motivation and support for this work.

This page intentionally left blank

Contents

About the Editors

Dr Sneha Kumari is a PhD as a full-time Junior Research Fellow from the Symbiosis International (Deemed) University. She has completed her Undergraduate in Agriculture at the Indian Council of Agriculture Research Fellowship and her Masters in Agribusiness Management under Indian Council of Agriculture Research. She has also completed her PGDBA in Human Resource Management and Marketing Management. She is an Assistant Professor at Vaikunth Mehta National Institute of Cooperative Management – a National Institute of Ministry of Agriculture and Farmers Welfare, Government of India. She has a rich experience as Statistical Officer, Researcher and Assistant Professor. She is also associated with different institutes for various educational and research related project assignments. She has published several research papers in the area of big data, agriculture, sustainability in ABDC and Scopus journals and has attended several national and international conferences.

Dr K. K. Tripathy is a PhD from Department of Management Studies of Indian Institute of Technology Delhi on "Micro-finance Management and its Impact on Rural Livelihoods"; Masters in Economics and Bachelor of Law (LLB). He is an Officer of Indian Economic Service (IES), Government of India. He joined the IES in 1999. Prior to joining the IES, he also worked as an Executive Magistrate in the Government of Odisha after joining the Odisha Administrative Service in 1998. He is presently the Director of the Vaikunth Mehta National Institute of Cooperative Management – a National Institute of Ministry of Agriculture and Farmers Welfare, Government of India. During his 21 years of public service, he has served in the Ministry of Human Resource Development (MHRD), Ministry of Agriculture and Farmers Welfare, Ministry of Rural Development, Ministry of Food Processing Industries, Planning Commission and United Nations Development Programme. He has published around 70 research articles and has attended several international and national conferences. His areas of research are rural development, microfinance and agriculture.

Dr Vidya Kumbhar is PhD in Computer studies from Symbiosis International (Deemed University). Her area of specialization includes data science and analytics, software development, big data handling, spatial data mining, spatial decision support system and agro-informatics. She is also associated with different institutes for various educational-related assignments. She has worked on software development projects in National Informatics Centre, Pune and currently working on different research projects funded by government and private organizations. She has published various research papers in National and International Journals and conferences. She is also acting as a reviewer for national and international journals indexed in Scopus and web of science.

Preface

According to Kitchin (2014) Big data is defined as huge volume of structured and unstructured data. Boyd and Crawford (2012) have defined big data's cultural, technological and scholarly phenomenon while Fan, Han, & Liu (2014) have defined big data as the ocean of information.

5 Vs of Big Data

While the term "big data" is relatively new, the act of gathering and storing large amounts of data is characterized by 5 Vs (Jeble, Kumari, & Patil, 2016)

Volume: Organizations collect data from a variety of sources, including business transactions, social media and information from sensor or machine-to-machine data. In the past, storing it would've been a problem – but new technologies (such as Hadoop) have eased the burden.
Velocity: Data streams in at an unprecedented speed and must be dealt with in a timely manner. Radio Frequency Identification (RFID) tags, sensors and smart metering are driving the need to deal with torrents of data in near-real time.
Variety: Data comes in all types of formats – from structured, numeric data in traditional databases to unstructured text documents, email, video, audio, stock ticker data and financial transactions.

At Statistical Analysis System (SAS), two additional dimensions is considered when it comes to big data:

Variability: In addition to the increasing velocities and varieties of data, data flows can be highly inconsistent with periodic peaks. Is something trending in social media? Daily, seasonal and event-triggered peak data loads can be challenging to manage. Even more so with unstructured data.
Veracity: Today's data comes from multiple sources, which makes it difficult to link, match, cleanse and transform data across systems. However, it's necessary to connect and correlate relationships, hierarchies and multiple data linkages or data can quickly spiral out of control.

Big Data and Business Analytics is conceived to provide a platform for academicians and practitioners to identify and explore the solutions to various problems in society, environment and industry using advance analytic tools.

Business analytics is the process of converting data into insights (Xavier, Srinivasan, & Thamizhvanan, 2011). It is "the extensive use of data, statistical and quantitative analysis, explanatory and predictive models, and fact-based management to drive decisions and actions." With the increase in the availability of data, analytics has now become a major element in both the top line and the bottom line of any organization. However, the rate of absorption of analytics in decision-making is slow. This is due to the fact that there are several ambiguities in the definition and scope of analytics (Jeble, Kumari, Venkatesh, & Singh, 2019). An effective use of analytics must grow with time and experience in most individuals. There is much more in analytics besides descriptive data collection and reporting. By 2025 there will be an increasing need for more data analytics to be involved in business. Effective performance management analytics is an integration of IT-based solutions, management accounting applications and analytical methods. Therefore descriptive, predictive and prescriptive analytics are essential for any business.

Use of analytics in better decision-making has evolved since past. In late 1960s technology-based analytics had been the base of the decision support system (Jeble, Kumari, & Patil, 2018). Later in 1987 scanner panel data was used to analyze decision-making in retail shops which was further followed by OLAP, a software analytical tool. In 1990 Enterprise resource planning system became the prime use for analytics in company. This led to the evolution of analytics using internet, e-commerce, mobiles, sensors and software analytical tools. With time Big Data Predictive Analytics has been used in decision-making in different streams. Analytics in big data has been useful in improving the visibility and coordination (Dubey et al., 2018). Businesses can have better decision-making capability with the use and better understanding of data analytics (Agrawal, 2014).

Why is Big Data Important?

The importance of big data doesn't revolve around the availability of data, but the purpose of data. One can take data from any source and analyze it to find answers that enable (1) cost reductions, (2) time reductions, (3) new product development and optimized offerings, and (4) smart decision-making. When big data is combined with high-powered analytics, managerial decisions can be performed such as:

- Determining root causes of failures, issues and defects in near-real time.
- Generating sustainable solutions for any stream.
- Recalculating entire risk portfolios in minutes.
- Detecting fraudulent behavior before it affects any organization.

Application of Big Data and Business Analytics

Big data has been in use by government institutes for forecasting weather patterns, discovering seismic activities that predicts earthquakes and preparing descriptive reports. It has been in use by the Economists to stimulate economic growth.

Big data mining is the patterns in the data that is normally not looked by the users. This unlooked data also leads to several important informations which can make decision-making smooth.

Big data is used in health insurance for predicting customer dissatisfaction through speech to text data from call center recordings (Devenport & Dyche, 2013). Several retail banks have focused on exploiting big data at times of financial crisis for doing a better job.

Linked In has used big data and data scientists to develop product features and product offerings. This has helped the consumers as well as the companies to make decisions about the product.

Google have constantly developed new products and services that have big data algorithms for search. Most of the companies are masters in developing standard reports and multidimensional reports through big data analytics. In many companies, big data is directly focused on products, services and customers. Senior managers have used predictive analytics as the next step in data analytics.

Big data not only allows knowledge discovery efforts but also they need to promote them. The sooner the business executives understand the value of knowledge discovery the better competitor they become. This can lead to high level innovations and high rewards.

Big Data and Business Analytics for Decision-Making

In the current era, world is challenged with demanding customers, high competition, short product lifecycles, rising costs of labor and materials, unemployment and unsustainability. Globalization is making it even more challenging as blurring boundaries among countries create level playing field for selling products and services across the globe. Firms need to make operational, tactical and strategic decisions based on available information. In addition to traditional decision support systems, big data provides additional tools to arrive at decisions.

Big data can provide valuable competitive intelligence (Jeong et al., 2016), help in dramatic cost reductions, substantial improvements and development of sustainable goals for the world.

Objectives

Data Science & Business Analytics will bring together researchers, engineers and practitioners and encompass wide and diverse topics of application in almost every field. It will also invite the participation of scholars, analysts and data scientists to present their ideas, concepts and proof of works indicating application of Big Data and Business Analytics.

Target Audience

The primary target audience of this book includes researchers, academicians and data scientist from a variety of disciplines interested in analyzing and application of big data analytics. A secondary target audience consists of data analysts, students and scholars pursuing advanced study in big data.

Organization of the Book

The book is organized into eight chapters. A brief description of each of the chapters follows:

Chapter 1 authored by Sneha Kumari, Vidya Kumbhar and K. K. Tripathy analyses and compares the big data on soil parameters of a district with the standards paving a way for mapping the crops with suitability of soil health The major component of agriculture production includes the type of seed, soil, climatic conditions, irrigation pattern, fertilizer, weed control and technology used. Soil is one of the prime elements in modern times for agriculture. Soil is one of the primary and important factors of crop production. The available soil nutrient status and external applications of fertilizers decides the growth of crop productivity. This chapter focuses on the application of soil data on soil health management for sustaining agriculture understanding the causal relationship between soil health parameters, cropping pattern and crop productivity.

Chapter 2 authored by Nilisha Prashant Itankar, Yogesh Patil, Prakash Rao and Viraja Bhat reviews the big data on heavy metals playing a crucial role in the economic development of a nation. Industries utilizing heavy metals, consequently, emanate large volume of metal containing liquid effluents. Since metals are non-renewable and finite resources, their judicious and sustainable use is the key. Hazardous metal laden water poses threat to human health and ecology. Apart from metals, these industrial effluents also consist of toxic chemical. Conventional physical–chemical approach technologies are not efficient enough, consume energy and are thus not cost effective. It is known that biomaterials (i.e., microorganisms, plants & agricultural biomass) have the ability to bind metals, in some cases selectively, from aqueous medium known as "metal biosorption."

Chapter 3 authored by Girish Joshi and Bindya Kohli designs a conceptual framework with clear land bank ownership records in dematerialized form, so that it can be easily maintained and traded like commodities. Secondary data published in the form of previous research work in Geographical Information System and business analytics has been used to analyze the cases on land bank dematerialization. This chapter discusses on practices like digitization of land records, conceptual dematerialization, use of analytical dashboards for analysis of data and associated benefits. This change can help India to transform its land management process and will help to explore commercial utilization of agricultural land and urban land plots for planned development.

Chapter 4 authored by Tihana Škrinjarić presents the potentials of including online search volume data in modeling the demand series of consumer products. Forecasting future demand for products of a company represents one of the important parts of planning and conducting business in general. Thus, the purpose of this chapter is twofold. The first purpose is to give a critical overview of the existing research on the topic of forecasting and now casting demand and consumption. The other purpose is to fill the gap in the literature by empirically comparing several approaches of modeling and forecasting demand and consumption on real data.

Chapter 5 authored by Vishita Rajesh Khanna analyses and compares variety of tastes in the food industries. For sustaining in such environment companies create their unique selling point and big data helps them to analyze market situation for such purpose. Companies combine big data with technologies like machine learning and artificial intelligence to get faster and more personalized experiences. This can be an opportunity for the food industries to reduce food loss and gain better returns on investment by going for a digital transformation.

Chapter 6 authored by Ritambhara Singh reviews to develop a strong research base for the academia and the industry to understand the importance of data analytics in International trade. This chapter focuses on the case of cotton trade from India and explores different methodologies developed by the World Bank and International Trade Center to analyze the big data available on export and import. Through big data analysis, this chapter finds out the export performance, market demand, export potential and attractive markets for Indian cotton. This chapter also explores the trade competitiveness of Indian cotton over the years. The data through appropriate analysis can address some simple yet complicated questions in trade like what export potential the commodity holds, if the commodity is competitive or not in international market, what are new markets to look up to, and other similar questions. In other words, this information could make huge difference in decision-making of traders and policymakers directly, and farmers indirectly.

Chapter 7 authored by Raj Krishna deals with the big data of Aadhaar Project of the Central Government of India, its features, its impact upon the welfare schemes of government. This chapter paves the way for the Aadhaar scheme.

Chapter 8 authored by Dr Irem Ucal Sari, Duygu Sergi and Burcu Ozkan establishes and presents customer segmentation and RFM analysis first, then a real case application of RFM analysis on customer segmentation for a fuel company. At the end of the application part, possible strategies for the company are generated.

Dr Sneha Kumari
Assistant Professor, Vaikunth Mehta National Institute of Cooperative Management, Pune, India

Dr K. K. Tripathy
I.E.S. Director, Vaikunth Mehta National Institute of Cooperative Management, Pune, India

Dr Vidya Kumbhar
Assistant Professor, Symbiosis Institute of Geoinformatics, Symbiosis International Deemed University, Pune, India

References

Agrawal, D. (2014). Analytics based decision making. *Journal of Indian Business Research*, *6*(4), 332–340.

Boyd, D., & Crawford, K. (2012) Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, *15*(5), 662–679.

Devenport, T. H., & Dyché, J. (2013). Big data in companies. *International Institute for Analytics*. Retrieved from https://www. sas. com/content/dam/SAS/en_us/doc/whitepaper2/bigdatabigcompanies-106461. Accessed on March 12, 2020.

Dubey, R., Luo, Z., Gunasekaran, A., Akter, S., Hazen, B. T., & Douglas, M. A. (2018). Big data and predictive analytics in humanitarian supply chains: Enabling visibility and coordination in the presence of swift trust. *International Journal of Logistics Management*, *29*(2), 485–512.

Fan, J., Han, F., & Liu, H. (2014). Challenges of big data analysis. *National Science Review*, *1*(2), 293–314.

Jeble, S., Kumari, S., & Patil, Y. (2016). Role of big data and predictive analytics. *International Journal of Automation and Logistics*, *2*(4), 307–331.

Jeble, S., Kumari, S., & Patil, Y. (2018). Role of big data in decision making. *Operations and Supply Chain Management – An International Journal*, *11*(1), 36–44.

Jeble, S., Kumari, S., Venkatesh, V. G., & Singh, M. (2019). Influence of big data and predictive analytics and social capital on performance of humanitarian supply chain. *Benchmarking: An International Journal*, *27*(2), 606–633.

Jeong, Y. S., Kim, B., Choi, H. J., & Lee, J. S. (2016, January). Issues about combination of temporal information and spatial information. 2016 International Conference on Big Data and Smart Computing (BigComp) (pp. 462–465). Retrieved from https://ieeexplore.ieee.org/abstract/document/7425971. Accessed on June 12, 2020.

Kitchin, R. (2014). Big data, new epistemologies and paradigm shifts. *Big Data & Society*, *1*(1). doi:10.1177/2053951714528481

Xavier, M. J., Srinivasan, A., & Thamizhvanan, A. (2011). Use of analytics in Indian enterprises: An exploratory study. *Journal of Indian Business Research*, *3*(3), 168–179.

Foreword

World has generated data in every field giving rise to big data. There is plenty of data everywhere and data storage is becoming critical at present. The importance of big data doesn't revolve around the availability of data, but the purpose of data. Researchers, academicians, policy makers and practitioners are consistently driving ways to find out the application of big data. With so much of data, it is time to understand the big data and how analytics can help in better decision making and manage things. Application of Big Data and Business Analytics at present needs to be explored among academicians, practitioners, policy makers and researchers. The book will have academic and managerial implications to manage the decision-making process. This book explores a number of perspectives on how big data and business analytics can help in better decision making. The book can be an asset for the readers at the present time. The authors from different countries and universities have made a contribution in organizing their research ideas into research chapter meeting the scope of the book. The chapters in the book have been selected carefully, providing a fine balance between trends in big data analytics and it's application in different streams. The chapters have diverse themes in application of big data analytics.

The chapters selected have been classified into Themes like data visualization, multiple aspects of data analytics, predictive analytics, application of data analytics in industry, agriculture and service sector followed by the challenges in digital technologies. The chapter on Customer Segmentation Using RFM Analysis: Real Case Application on a Fuel Company establishes and presents a real case application of RFM analysis on customer segmentation for a Fuel company with possible strategies for the company are generated. The chapter on Applications of Big Data Analytics: A Boon for the Food Industry analyses big data with technologies like machine learning and artificial intelligence to get faster and more personalized experiences generating an opportunity for the food industries to reduce food loss and gain better returns on investment by going for a digital transformation. The chapter on Big Data for Sustainable Rural Development with special reference to MGNREGA addresses the application of big data related to rural employment with special reference to the world's largest public works and wage employment generating poverty alleviation programme – Mahatma Gandhi National Rural Employment Guarantee Act (MGNREGA) on Rural Development. The authors have done an appreciable work in presenting the applications of big data and business analytics in different managerial decision making. The application of big data can be seen in descriptive, predictive and prescriptive analytics.

All the chapters are topical. The chapters are well-balanced covering the application of big data and business analytics by academicians, researchers, industrial experts, policy makers and practitioners. This book will bring together researchers, engineers and practitioners and encompass wide and diverse topics of application in almost every field. It will also invite the participation of scholars, analysts and data scientists to analyze the application of Big Data and Business Analytics by the contributors from different countries. The book paves a way for the readers to understand how big data can be efficiently utilized in better managerial applications. Dr. Sneha Kumari, Dr. K. K. Tripathy; Vaikunth Mehta National Institute of Cooperative Management and Dr. Vidya Kumbhar, Symbiosis International (Deemed University) have done a commendable job as book editors in making the application of big data analytics research available for a wide audience.

Dr. Lt. Col. Anupama Munshi
Veteran, Faculty and Researcher, India

Dr. Lt. Col. Anupama Munshi (retd) is a doctorate in Management with 17 years of experience in Indian Army in handling big data of entire gamut of Human Resource Development functions , Human Resource Management and Industrial Relations. She has dealt with the application of big data in decision making in the Army. She is an expert in application of big data in designing & implementing training programs to enhance efficiency & motivation levels. She has also applied the big data for imparting teaching and training to officers of Indian Army as well as officers of other armies in subjects like Quality control, Logistics & Supply chain management, Transport management, Tendering and procurement for Defense supplies. She is the First Lady Officer of Indian Army to command an Independent Food Inspection Unit for providing logistics support to a specialized brigade and was first officer of Army Service Corps to be awarded General Officer Commanding in Chief's Commendation card for outstanding service. She has also worked as a Professor in Symbiosis International Deemed University, Pune and is the visiting faculty for many renowned Management Institutes in Pune. She is consistently involved in researches of application of big data in the Indian Army.

Editorial Advisory Board

This page intentionally left blank

Chapter 1

Data Analytics for Soil Health Management and Their Crop Mapping in Satara District

Sneha Kumari, Vidya Kumbhar and K. K. Tripathy

Abstract

The major component of agriculture production includes the type of seed, soil, climatic conditions, irrigation pattern, fertilizer, weed control, and technology used. Soil is one of the prime elements in modern times for agriculture. Soil is also one of the primary and important factors for crop production. The available soil nutrient status and external applications of fertilizers decide the growth of crop productivity (Annoymous, 2017). The upcoming research question that needs to be addressed is What is the application of soil data on soil health management for sustaining agriculture? Driven by the need, the aim of the present study is (a) to explore the soil parameters of a district, (b) compare the values with the standards, and (c) pave a way for mapping the crops with suitability of soil health. This study will not only be beneficial for the district to take appropriate steps to improve the soil health but also would help in understanding the causal relationship among soil health parameters, cropping pattern, and crop productivity.

Keywords: Soil; crop mapping; micronutrients; macronutrients; data analytics; fertilizers; pH

Background

Indian agriculture planning has always stressed on increasing production and productivity of agriculture crops to meet not only the ever increasing demands of the consumers but also meeting growing livelihood requirements of millions

Application of Big Data and Business Analytics, 1–23

doi:10.1108/978-1-80043-884-220211002

of small and marginal farmers of the country. Agriculture production is highly influenced by the appropriate usage of agri inputs. The major components of agriculture production include the type of seed, soil, area suitability, climatic conditions, irrigation pattern, fertilizer, weed control, and technology. With time, an increased use of fertilizers and pesticides have resulted in crop productivity growth. With an excess use of fertilizers, the country has maintained its position in the top in fruits, vegetables, and food grains. An increase in the production has questioned the detrimental impact on soil in India. The soil has been deteriorated to the extreme and the production is somewhere having an adverse impact on human health.

The spatial variation of soil chemical and physical properties influences soil and crop management efficiency (Bhatti, Mulla, & Frazier, 1991). There is a need to focus on soil properties to increase the crop productivity. Soil is the core component of agri inputs which further drives the use of fertilizers, agrochemicals, and water. The soil nutrient balance has to be maintained to build the gap between potential and actual yield (Tittonell, Shepherd, Vanlauwe, & Giller, 2008). In India, more focus is laid on soil nutrient balance practices for specific cropping systems in irrigated areas as these areas are economically more important for national food security (Rego, Rao, Seeling, Pardhasaradhi, & Rao, 2002; Rego et al., 2003). Immense scope of work exists to ensure degree of nutrient balance in less endowed dry land or semi-arid cropping systems. Soil quality has a high degree of influence on the cropping system. Indian farmers struggle to grow their crops because of soil nutrient deficiency in their farm land. Lack of soil health and the existing ineffective soil health management system at the grassroot level drive the farmers to apply more and more; this is done by the use of fertilizers, agrochemicals, and bio controlling agents. Despite several efforts made by farmers, there has been a decline in the production and productivity. This is due to insufficient mapping of the crops with the soil type and characteristics. Therefore, it is essential to map the crop with the soil dimensions and find out underlying factors of low productivity and possible ways toward effective soil health management.

Introduction

The natural resources like soil, climate, and water profoundly influence the sustainability of cropping patterns. Severe soil degradation and conversion of fertile land for civilian and infrastructure development have compelled researchers, academicians, and policy makers to map the suitability of soil for different agriculture uses. The importance of soil and its improvement has been receiving greater impetus. This is often followed by land evaluation. According to FAO (1976), land evaluation is defined as the process of land assessment involving surveys and studies used for specific purpose. Soil suitability studies are beneficial for agriculture sector as it provides information on the choice of crops to be grown for maximizing crop production. The suitability of land is a balance between soil restoration and degradation. For effective planning and utilization, soil site studies have become the necessity at present. Soil suitability models also provide guidelines to the policy makers for mapping the most suitable crops with the soil. Although Big Data have been generated by the agriculture universities and research institutes, yet it has not been correlated with soil suitability for sustaining crop production. The upcoming research question that needs to be addressed is what is the application of soil data on soil health management for sustaining agriculture. Driven by the need, the present study envisages (a) to explore the soil parameters, (b) to compare the values with the prescribed standards, and (c) to pave a way for mapping the crops with suitability of soil health. Through this study, an attempt has been made to map the suitability of crop with the present soil conditions.

Literature Review

Soil, climate, nutrient status, and management ability influence the crop yield (Fowler, 1999). Sustainable agriculture depends on good soil health to overcome the low productivity of crops (Kumari & Patil, 2017). It is essential to overcome the adverse effects of soil before sowing of seeds. Crop wise land or area suitability is a prerequisite for increasing the crop productivity. Crop production also depends on precipitation, temperature, elevation, soil, fertilizers used, and land types used. Soil is one of the primary and important factors of crop production. It is a consolidation of minerals and organic materials that help in the growth of crop (Annoymous, 2017). The available soil nutrient status and external applications of fertilizers decide the growth of crop productivity (Annoymous, 2017). Thus, land suitability and soil parameters are very important to increase the crop yield. Soil nutrient maps are designed using geographic information system (GIS)-based technique to provide information on site-specific fertilizer application to increase the crop productivity in dairy farms of Ireland and China (Fu, Tunney, & Zhang, 2010). The spatial distribution features of soil nutrient contents and their response to land use for red soil in the area of the Zhuxi small watershed were identified to improve the crop productivity (Li, Chen, Chen, & Chen, 2011). The statistical methods were also applied to study the variability of soil properties such as soil pH, soil organic carbon (OC), total nitrogen (N), phosphorus (P), and potassium (K) in the MaiNegus Catchment of Tigray, Northern Ethiopia (Tesfahunegn, Tamene, & Vlek, 2011). The GIS-based multicriteria overlay analysis was used to study the soil suitability for the cotton crop. The detailed analysis of morphological and physico-chemical properties concluded that not only climatic parameters but also the position of the landscape is important for the crop suitability analysis (Walke, Reddy, Maji, & Thayalan, 2012). The multivariate geo-statistical approach along with PCA was applied to study the spatial relationship between soil and crop variables in durum wheat field at the farm of the University of Tuscia, Viterbo, Central Italy. The study was useful for offering decision support for the development of site-specific crop management strategies (Casa & Castrignanò, 2008). The personal digital assistant (PDA)-based soil fertility management information system using embedded GIS was designed for Jianshui country that helped for the recommendation for N, P, and K fertilizers maintaining the soil nutrient balance (Xiaolin, Linnan, Lin, Wengfeng, &

Limin, 2012). The GIS-based multicriteria analysis method is used for the decision-making process of land suitability analysis for the crop to enhance the crop productivity (Elsheikh et al., 2013; Walke et al., 2012; Wan et al., 2009). The geo-statistical approach is applied for more effective integration of soil knowledge of classifications, covariates with statistical spatial prediction of soil properties, and crop productivity improvement (Lark, 2012). The multivariate geo-statistical analysis of spectral data was used to decide soil-specific crop suitability over space and time (Casa & Castrignanò, 2008). It was also applied to produce nutrient maps and to provide information on site-specific fertilizer application to increase the crop productivity (Fu et al., 2010; Tesfahunegn et al., 2011). The statistical and geo-statistical analysis techniques were applied to identify the spatial variability of soil total nitrogen (STN) and soil total phosphorus (STP) and to identify the relationship among precipitation, temperature, STN, and STP densities (Bi, Li, Liu, Guo, & Jun, 2009; Liu, Shao, & Wang, 2013). The study of chemical properties such as soil pH, OC, and soil physical properties such as elevation along with rainfall and climate parameters was used to improve the wheat crop yield (Sarkar et al., 2014). The literature confirms that the soil nutrient variation plays an important role for crop production.

Importance of Soil Parameters for Crop Yield

Soil tests identify the percentage of the nutrients, micronutrients, and biophysical properties of the soil. It is a practical way to determine lime and fertilizer needed for specific crop to increase the crop yield (Peters, Laboski, & Bundy, 2007). Soil test levels are positively related to the number of soil samples collected per field. The accuracy level of soil sampling determines the fertility index of the corresponding field (Dinkins & Jones, 2007). Knowledge about spatial variance of soil fertility is of significant importance for implementing effective crop management to increase the crop yield (Gitas, Douros, Minakou, Silleos, & Karydas, 2009; Hedoin, Ellsworth, & Hornbaker, 2007; Karydas, Gitas, Koutsogiannaki, Lydakis-Simantiris, & Silleos, 2009). Soil chemical properties, fertility, and soil physical properties are important for natural growth of the crop (Wang, Sun, Jia, Li, & Xu, 2008).

Soil Chemical Properties

Soil chemical properties help to maintain nutrient status of the soil. The pH level between 7.5 and 8.57 shows that the soil is less alkaline and good for most of the crops (Brady, Weil, & Weil, 2008). Based on the soil pH value, the soil may be classified as saline, sodic, and acidic soils. The soil salinity is indicated by soil electrical conductivity (EC) (Brady et al., 2008). It has been found that saline and sodic soils are not good for the crop growth. The soil OC is the best source of energy and helps for growth as well as increment in nutrient availability through mineralization (Körschens et al., 2013). Soil calcium is a basic element of soil which helps to create the cell structure of the crop and increases the water penetration (Horneck, Ellsworth, Hopkins, Sullivan, & Stevens, 2007; Kirkby & Pilbeam, 1984).

Soil Fertility Parameters

The soil productivity depends on the soil fertility parameters. Soil is composed of micro- and macronutrients. Macronutrients comprise nitrogen, phosphorous, calcium, sulfur, magnesium, and potassium, while micronutrients are manganese, molybdenum, boron, iron, chlorine, zinc, and copper. These nutrients are essential for crop growth. Soil nitrogen helps foliage grow strong and affects the plant's leaf development. It also gives plants the green color because of its assistance with chlorophyll production. Soil phosphorous assists with the growth of roots and flowers, helping the plants to survive harsh climates and environmental stressors. Soil potassium strengthens plants, helps contribute to early growth, and assists the plants in retaining water. It also protects the plants from contracting diseases and insects. Soil magnesium helps in the green color development of the plant. Soil sulfur aids in the production of amino acids, proteins, enzymes, and vitamins. Soil calcium aids in the growth and development of cell walls. Soil zinc is helpful for the plant growth. It helps in the growth of roots as well as carbohydrates and chlorophyll formation (Alloway, 2008; Mousavi, 2011; Mousavi, Galavi, & Rezaei, 2013). The soil magnesium is an important and essential element for different crop functions. It helps to improve the nutrient capacity of the crop along with photosynthesis. It is one of the important elements of chlorophyll (Gerendás & Führs, 2013; Huber & Jones, 2013). Soil iron is also an important element for crop production. It helps to improve the metabolism and the respiration of the crop. It is an important chlorophyll element, which aids in nitrogen fixation in soil (Shenker & Chen, 2005; Zuo & Zhang, 2011). Soil copper is another important soil fertility parameter which helps in protein and carbohydrates metabolism and improves photosynthesis and respiration of the crop (Gupta, 1997; Rehm & Schmitt, 1997). Along with copper contents, potassium and phosphorous in the soil also play an important role in photosynthesis (Dzotsi et al., 2010; Wang, Zheng, Shen, & Guo, 2013). This helps to improve the drought resistance of soil (Kumari, Patil, & Rao, 2020), maintaining the soil moisture contents, and helps to improve the crop yield.

Soil Physical Properties

Soil physical properties play an important role in improving moisture retention and water holding capacity. The soil becomes coarse and allows the water to flow through soil particles and clay, while silt holds the water contents for a long time (Selassie, Ayalew, Elias, &Getahun, 2014).

Research Objective and Scope

Every year, the soil testing laboratory in each district compiles a series of data on physical and chemical properties of the soil. The soil data on pH, nitrogen, phosphorus, and potassium (NPK), micronutrients, texture, density, and moisture have generated a Big Data on soil health. The upcoming research question that needs to be addressed is what is the application of soil data on soil health management for sustaining agriculture. Driven by the need, the aim of the present study is to (a) explore the soil parameters like pH, EC, OC, phosphorous, micronutrients, macronutrients, density, and texture of a district; (b) compare the values with the standards; (c) compare the data of soil from 2015 to 2016 and 2018 to 2019 available from soil health survey; and (d) pave a way for mapping the crops with suitability of soil health. The effective utilization of soil parameters such as soil nutrients, soil micronutrients, soil pH, and soil moisture improves the crop growth, quality of crop, and crop yield. The soil suitability decision for major crops in semi-arid and arid areas will improve the agricultural performance of the district. The scope of the study is limited to Man taluka of Satara district.

Research Methodology

The soil data on soil pH, EC, OC, phosphorous (P), potassium (K), copper (Cu), iron (Fe), magnesium (Mg), and zinc (Zn) have been collected from the soil testing laboratories of Satara district of the Indian State of Maharashtra. The other parameters of soil, such as moisture percentage, $CaCO_3$, coarse sand percent, silt percent, clay percent, fine sand percent, apparent density, pore space percent, volume percent, specific density, texture, calcium, magnesium, and sodium percent in the soil have also been collected from the agriculture soil testing laboratories. About 33,000 samples of soil data have been obtained from Man taluka[1] of Satara district in the year 2018–2019 and 7,000 samples in the year 2015–2016. Data of soil samples collection have been taken from the Man (Dahiwadi) taluka of Satara district. The procedure for formulating soil site criteria involved rating of soil parameters from 1 to 5, where 1 indicates very low and 5 indicates very high. For micronutrients, soil rating has been converted into 1 and 2, indicating low- and high-content values. The parameters of the soil have been mapped with the standard soil health requirements. Based on the soil site database, an attempt has been made to develop suitability for major crops grown in Man Taluka.

Result and Discussion

After obtaining data from soil health card[2] under the Department of Agriculture Cooperation and Farmers' Welfare, Man taluka was chosen for further study. The soil data of the year 2015–2016 and 2018–2019 were collected from the soil data survey conducted by the Ministry of Agriculture and Farmers' Welfare. Table 1 shows the statistics of micro- and macronutrients of soil of the 2 years. It has been observed from the soil health survey that around 7,600 soil samples were taken for the year 2015–2016 and 33,000 soil samples for the year 2018–2019. The study shows that the soil of Man taluka is deficient in micronutrients like boron and macronutrients like sulfur and nitrogen.

---

[1]Maan taluka is situated on Man river in Satara district of Maharashtra in India.
[2]Soil Health Card scheme aims at promoting soil test based and balanced use of fertilizers to enable farmers to increase crop yield with low input cost.

Table 1. Descriptive Statistics of Soil Health of Man Taluka of Satara District.

	Year[a]	Number of Soil Samples	Mean	SD	SEM
OC	1	7,688	1.15	0.564	0.006
	2	33,790	2.80	0.754	0.004
N	1	2,715	1.19	0.531	0.010
	2	33,733	2.52	0.880	0.005
P	1	7,691	1.28	0.983	0.011
	2	33,788	3.67	1.187	0.006
K	1	7,689	1.39	0.857	0.010
	2	33,787	3.83	0.800	0.004
S	1	7,690	0.01	0.116	0.001
	2	33,786	0.13	0.334	0.002
Zn	1	7,684	0.63	0.482	0.005
	2	33,784	0.47	0.499	0.003
Fe	1	7,686	0.68	0.468	0.005
	2	33,779	0.18	0.382	0.002
Cu	1	7,686	0.99	0.105	0.001
	2	33,763	0.98	0.143	0.001
Mn	1	7,685	0.91	0.292	0.003
	2	33,598	0.45	0.498	0.003
B	1	7,690	0.04	0.188	0.002
	2	33,784	0.48	0.500	0.003

Note: OC, organic carbon; N, nitrogen; P, phosphorous; K, potassium; S, sulfur; Zn, zinc; Fe, iron; Cu, copper; Mn, manganese; B, boron.
For OC, N, P, and K, 5-point Likert scale has been used where 1 = very low and 5 = very high to describe the degrees of nutrient content in the soil samples.
For S, Zn, Fe, Cu, Mn, and B, the soil data reflect as deficient or sufficient expressed as 0 and 1.
[a]Year 1 indicates 2015–2016 and year 2 is 2018–2019.

Table 1 indicates that the average content of nutrients has deteriorated from 2015–2016 to 2018–2019 for the micronutrients like manganese and copper. While for the rest of the nutrients, there is a significant increase in the mean.

Further comparative study has been conducted for the two years using the independent sample *t*-test. The below-mentioned hypothesis has been set.

> *H0*. There is no significant difference in the soil properties for the year 2015–2016 and 2018–2019.

> *H1*. There is significant difference in the soil properties for the year 2015–2016 and 2018–2019.

The result shown in Table 2 demonstrates that there is a significant difference in the soil properties for the periods under reference as the level of significance is found to be less than 0.05. The soil properties have changed with time.

After conducting the *t*-test, the study has undergone descriptive statistics of each micro- and macronutrients reported by the soil survey data. The significance level (equal variance not assumed) shows that there is a significant difference in the nutrients over the years.

Comparative Study of Soil Properties for the Years 2015–2016 and 2018–2019

OC. Fig. 1 shows that the OC has improved in 2018–2019 over 2015–2016. The OC was very low in 2015–2016, which has improved over the last three years.

One to five scale has been used to study this soil property where 1 is for very low, 2 is for low, 3 is medium, 4 is high, and 5 is very high. The present OC of Man taluka is medium. The OC of the soil has improved significantly as compared to 2015–2016.

Nitrogen. Fig. 2 shows that the nitrogen content has also improved significantly in the current year as compared with the past. However, the soil is still not self-sufficient in nitrogen content. Nitrogen status of soil has been improved significantly.

Phosphorous. Fig. 3 shows that the soil has significantly become rich in phosphorous as compared with the past years.

Potassium. Fig. 4 depicts that the potassium level has also increased significantly in the soil.

Sulfur. Fig. 5 shows that the soil is deficient in sulfur content and not much improvement has been seen by the soil in the past three years. It has been observed that the sulfur content of the soil has deteriorated compared to the past three years. Sulfur is an important macronutrient that affects crop yield. Change in atmospheric conditions and agriculture practices has led to deterioration in sulfur of the soil. Sulfur deficiency can be treated with gypsum, intake of fertilizers containing sulfur intake, manures, and maintaining pH level of the soil.

Zinc. Fig. 6 shows that the soil is not self-sufficient in the presence of zinc. Zinc is involved in fruit, seed, and root development. Zinc availability to crops declines with an increase in the pH of the soil, low soil temperatures, and low organic matter.

Iron. Fig. 7 shows that the soil has deteriorated in iron content.

Copper. Soil is self-sufficient in copper as demonstrated in Fig. 8.

Manganese. Fig. 8 depicts that soil has shown improvement in the manganese content.

Boron. Fig. 10 shows that the soil has significantly improved the boron content.