The Pharma Innovation Journal 2021; SP-10(4): 358-365

www.ThePharmaJournal.com

The Pharma Innovation

ISSN (E): 2277- 7695
ISSN (P): 2349-8242
NAAS Rating: 5.23
TPI 2021; SP-10(4): 358-365

www.thepharmajournal.com
Received: 13-02-2021
Accepted: 21-03-2021

S Manikandan
PhD (AB), Institute of Agri Business Management, SKRAU, Bikaner, Rajasthan, India

Dr. Amita Sharma
Assistant Professor, Institute of Agri Business Management, SKRAU, Bikaner, Rajasthan, India

Corresponding Author:
S Manikandan
PhD (AB), Institute of Agri Business Management, SKRAU, Bikaner, Rajasthan, India

A study on big data analytics in malt barley production in Rajasthan

S Manikandan and Dr. Amita Sharma

Abstract
Agriculture plays significant role in almost every country economy in the world. It is identified that agriculture produces huge amount of data with wide velocity and wide range of variety in each second. Presenting the discussion on how big data analytics and how its tools use in malt barley production is the main aim of this paper.

Keywords: Big data analysis, agriculture, malt barley, big data tools and techniques

Introduction
Nowadays people want not only collect the data but also they need to understand them and need to identify the importance of the data set in order to take better decisions. Big data is considered as a large collection of dataset which is having high velocity, volume and varieties that make difficulty in processing and managing by using traditional techniques and tools (Elgendy and Elragal, 2014) [4]. It can be either structured, unstructured or semi-structured. An advanced analytic technique which can be used for analyzing big data to reveal unknown, hidden and useful patterns is identified big data analytics (Elgendy and Elragal, 2014) [4]. Hence big data plays a major part in decision making process.

Barley (*Hordeum vulgare* L.) is the most widely grown cereal crop over broad environmental conditions and is the world's fourth important cereal crop after wheat, maize and rice (USAID, 2014) [12]. Malt barley, due to its limited usage and high-quality requirements, is a unique niche product in international grain market representing only about 1.5% of the total world grain production (USAID, 2014) [12]. It is particularly interesting in the context of smallholder commercialization and food security since it has high value as both cash and food crop (Legesse *et al*., 2007) [6]. However, according to Mulatu and Grando (2011) [7], malt barley supplied to malt factory is produced by farmers having fragmented and small plots of land; as a result, the demand for malt barley are not being met. Asella malt factory's annual report, shows, only about 60% of the demand for malt barley is covered by domestic production while the rest is imported from foreign countries. Moreover, according to the report of Lemu Bilbilo district, from 55,245 hectares of land covered by cereal crops only 14% of the area is covered with malt barley (Weldeyohanis and Abddisa, 2017) [13].

As an outcome of rapid growth of countries, Information Technology has become major part in almost every field in a country like education, health services etc. Also agriculture plays a major part in each and every country economy. With the rapid development of the IT, it will become easier to collect data, store data and analyze data in order to derive useful information. IT can be used in agriculture as a direct tool for increasing the productivity as well as an indirect tool for encouraging farmers to take better decisions (Patel and Sayyed, 2014) [9]. Precision farming which is popular among developed countries largely uses IT to increase the productivity (Patel and Sayyed, 2014) [9]. It has been found that the world population will exceed the 9 billion in 2050 (Shekhar *et al*., 2017) [11]. Therefore, producing and delivering products in efficiently is highly important. Applying big data analytics in agriculture, gives highly advanced benefits like innovate solution for minimizing the usage of resources with having harvest of promising as in precision agriculture. Precision agriculture is mostly depended on analyzing the real time data which comes on large amount and from various sources such as soil, weather, air, equipment, availability etc. Smart farming which is mainly paid attention on how to use big data analytics in business processes to enhance their functions (Omo-Ojugo, 2017) [8]. Also the growth of data is quicker than the computational speed (Ch. Chandra Sekhar, 2018) [3].

Hence it is much needed to minimize the time consuming for analysis of data. Therefore by using the big data analytic tools and techniques, timely information can be obtained for famers to take better decisions on how to get the fruitful harvest and for the policy makers to take better decisions on import or export of products (Kumari *et al*., 2016) [5].

Methodology

This study includes the utilization of deductive methodology, which is viewed as more scientific in its methodology. Since deductive methodology implied beginning with ideas outlined from previous literature. This research structure aligned with this methodology.

- Identifying the factors involved in farming systems of Malt Barley in Rajasthan.
- To identify the factors involved in farming systems of Malt Barley in Rajasthan. Data's are collected using primary data collection approach.
- Identification of Big Data Applications in Malt Barley.
- To identify the factors involved in in farming systems of Malt Barley in Rajasthan. Data's are collected using secondary data collection approach.

Research design

Data collection procedure was done by analysing the primary and secondary sources. The methods involved in the primary data collection were through focus group and cognitive interview methods. Though it leftovers true that certain of info were composed via email and telephonic discussion, maximum of data gathering was conceded out by conducting a field trip.

The scope of this research is restricted to Malt barley in Rajasthan. The role and contribution of this research is evaluated using questionnaire. Data collected through questionnaire from the selected villages, selection of farmers will be done by ensuring minimum representation of main farming models in the study area (totally 405 farmers) using convenience sampling technique for the study.

Objectives of the study

To Study Big Data Applications in Malt Barley through Text Mining.

To explore applications of Big Data in farming models

Results and Discussion

Big Data Application in Malt Barley through Text Mining

Text mining in agriculture

Data mining is the process of extracting important and useful information from large data sets and is a relatively new inter-disciplinary concept involving data analysis and knowledge discovery from the databases (Richard Hogan, 2019) [10]. It uses a multi-facet approach, which includes statistical analysis, data visualization, neural networks, knowledge discovery, pattern recognition and database management (Ben Foley, 2009) [1]. Data mining techniques are utilized for future predictions of various crops which will help farmers to take most appropriate decision for their crops. Data mining techniques and machine learning techniques are utilized to investigate the effect of different parameters and make predictions of the crop production. India has highest production of many crops. The major crops cultivated in India are Rice, Wheat and Maize. On the basis of different cultivation season in India, crops can be divided into kharif, Rabi and zaid crops. Rice and Maize are Kharif crops, which is the summer or mansoon crop and cultivated in the months of January to December in all over India. Wheat is a Rabi crop, which is the spring harvest or winter crop and cultivated between the months of October to may in India. This work, analyse the malt barley production in the area of Rajasthan.

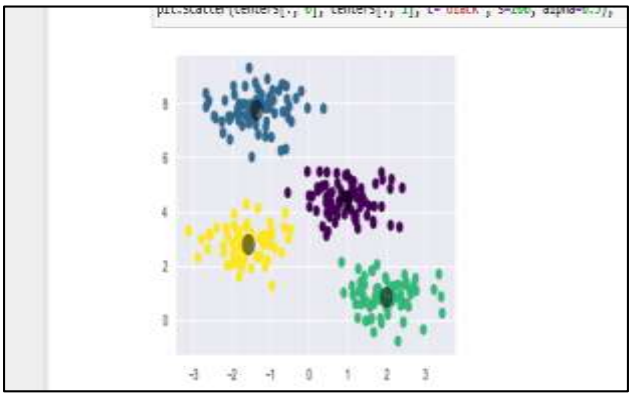
To predict rice crop production for Rajasthan state, India by selecting 33 districts of Rajasthan by considering parameters like precipitation, minimum temperature, average temperature, maximum temperature, reference crop evapotranspiration, area, production and yield for the Kharif season from 2010 to 2018.The algorithms used for crop production prediction clustering are RNN LSTM.

Clustering process

Clustering method is used to group text documents which has similar contents. It has partitions called clusters and each partition will have a number of documents with similar contents. Clustering makes sure that no document will be omitted from the search and it derives all the documents which has similar contents. K-means is the frequently used clustering technique. This technique also compares each cluster and finds how well the document are connected to each other. Companies use this technique to create a database with a similar document.

A sub-set of objects which are similar is called a cluster. High similarity occurs when the objects are in same clusters and the objects are dissimilar in other clusters Clustering is a data mining technique which maps the similar instance together, and dissimilar instance together, and dissimilar instance belong to diverse group based on data instance. The data instance is divided into subsets.

K-Means: K-means is the technique in which 'n' number of data objects are partitioned into 'k' number of groups according to some mean value with which the elements are compared. Each group has a member called as centroid which becomes the new mean value of that group. This mean value acts as the prototype of the cluster which helps in repeating the procedure until convergence is reached. This algorithm is very quick and can be run multiple times.



Source: Researcher's computation from secondary data

Fig 1: Cluster the top 4 barley production districts in Rajasthan

Malt Barley Growth percentage over 2010 to 2018

Production of Barley India for the last 8 years is presented in the below Figure 2. The trend shows a mixed result with the production Figures fluctuating around 1600 thousand MT since 2010. Also, good rainfall during 15-16 and 16-17 boosted production of wheat and hence overall production of barley is negative in last two years. The major producing states of Barley in India are Rajasthan.

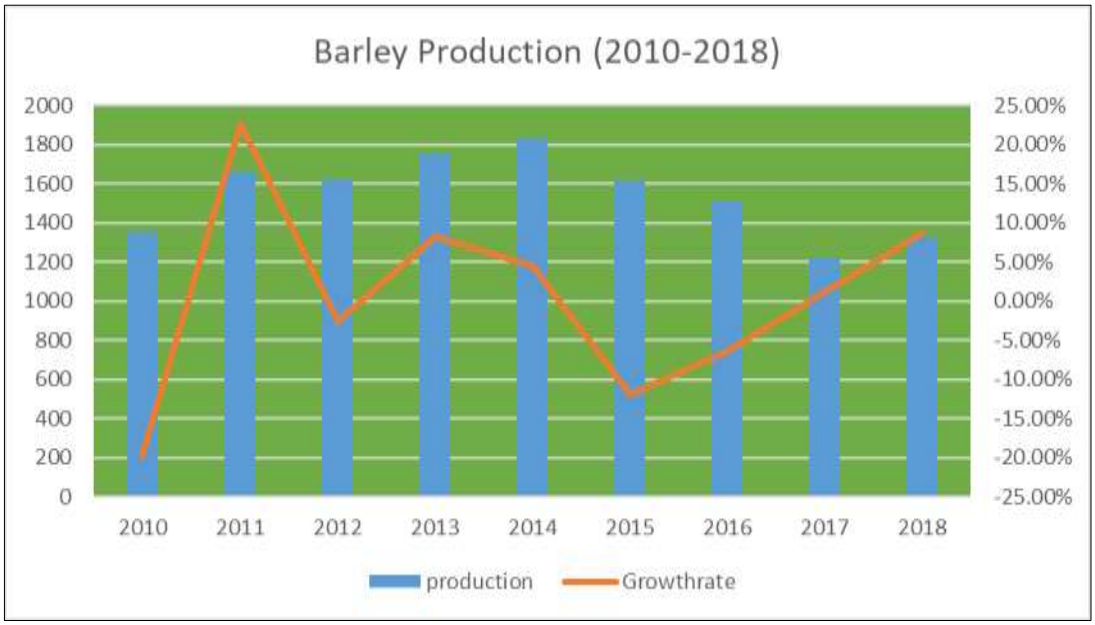


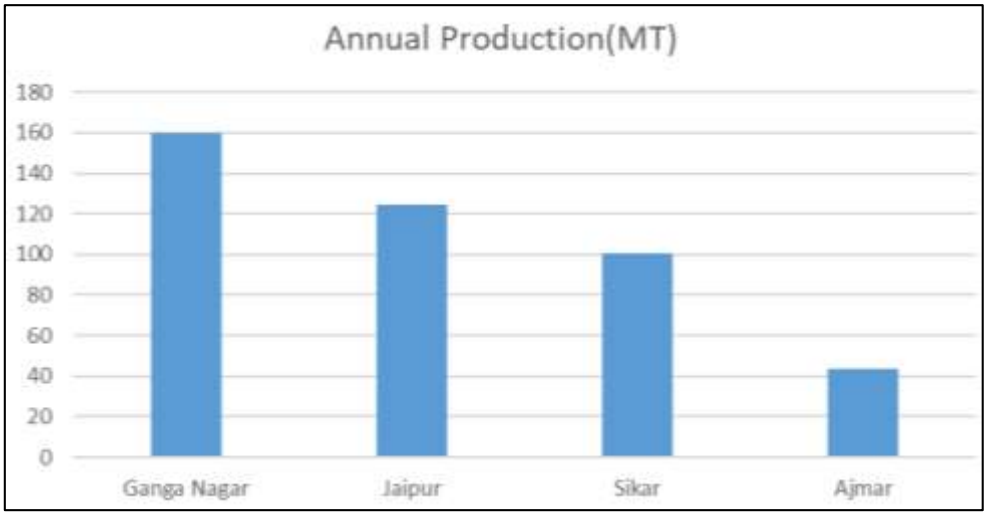
Fig 2: Malt Barley Production Rate over 2010 to 2018

Source: Researcher's computation from secondary data

A mixed trend of going up and down is shown in the trend. In 2010 the Barley production was rose to nearly 2000 thousand million tons and the rate of growth also showing an upward trend during this period. After that the rate of growth declined and by 2018 again the trends are growing.

Annual Production

The major Malt Barley producing districts in the Rajasthan are Ganga Nagar, Sikar, Jaipur, Ajmer. Ganga Nagar is the high barley producing district about 160,062 MT. The top 4 Barley producing districts of the Rajasthan are presented in Figure 3



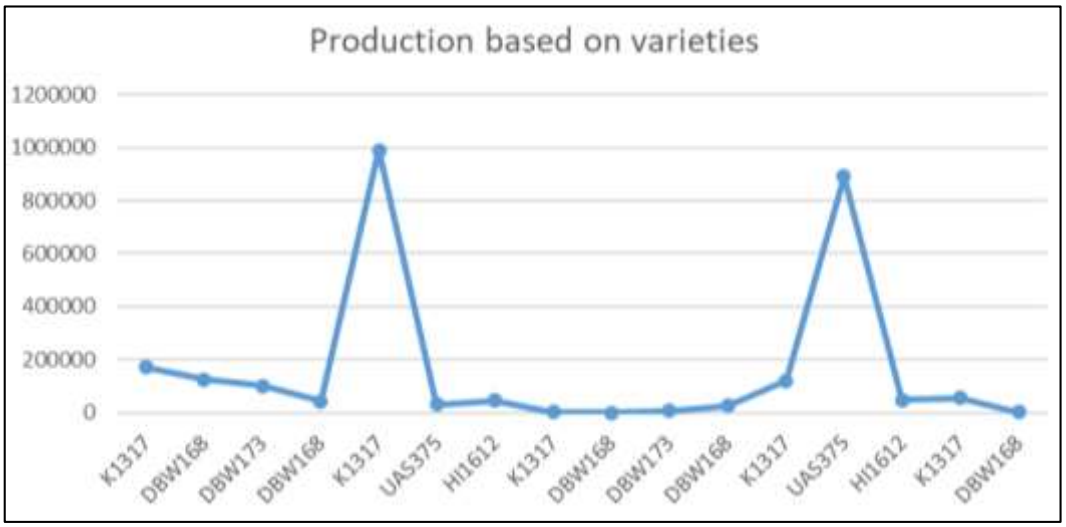
Source: Researcher's computation from secondary data

Fig 3: Annual Production of Barley (2018)

From the bar chart it is clear that Ganga Nagar is the fore runner in annual Barley production with 160 million tons and Ajmer with the least amount of production with 40 million tons annually. Jaipur contributes more than 120 million tons whereas Sikar produced 100 million tons of Barley every year. The production in these four places contributes together to the total barley production of our nation. Production varies in each place according to the geographical and climatic features of those areas. Farming techniques are also different in each of these places depending up on the techniques adopted by farmers of these places.

Annual Production based on varieties

The wheat varieties of K1317 and UAS375 are showing more level of produced outputs among all other varieties. K1317 had highest amount of production with 1000000 MT a year. UAS375 also has above 800000 MT annually. Whereas the malt barley production also more while using these wheat varieties followed by Barley.



Source: Researcher's computation from secondary data

Fig 4: Malt barley production Based on Variety

RNN-LSTM for Malt Barley Production Forecasting

Predicting Malt barley production, especially in Rajasthan state, as it is important in national economic programming. The production of Malt barley, in addition to the relationship with the genetics of the grower, adaptive terms, the effect of pests and pathology and weeds, quality of management and control during the growing season and so on. Therefore, it is not beyond the possibility of acquiring relationships or systems that can predict the highest precision with meteorological data. Today, there are many performance predictions models, most of which have generally been classified into two groups: a) statistical models, b) crop simulation models. Recently, the application of Artificial Intelligence (AI), such as Recurrent Neural Networks Algorithm has shown more efficiency in dissolving the problem. Applying them can make models easier and more accurate than complex natural systems with many inputs. In this research, an attempt has been made to develop a prediction model of Malt barley using RNN. If we design a network that correctly learns the relationships of effective previous year's production on Malt Barley, it can be used to estimate crop production in the long or short term and also with sufficient and useful data to obtain an RNN model for each area. In addition, the use of RNN may find the most effective factors in crop production. Therefore, some factors that your measures are difficult and cost effective can be ignored. In this case only the effect of climatic factors on barley has been applied. In computing and related fields, Recurrent neural networks are computational models that are able to learn to mechanize and recognize patterns. They are usually presented as systems of interconnected "neurons" that can calculate the values of the inputs by feeding information through the network. Like other machine learning methods, neural networks have been used to solve a wide variety of tasks that are difficult to solve through regular rule-based programming, including computer vision and voice recognition. The word network in the term "Recurrent neural network" refers to the inter-connections between neurons in the different layers of each system. This system contains three different layers they are the first layer has input neurons, which send data through synapses to the second layer of neurons, and then through more synapses to the third layer of the output neurons. More complex systems will have more layers of neurons with some increased layers of input neurons and output neurons. An RNN is typically defined by three types of parameters:

Dataset

In our dataset, we take year and month for malt barley production, the task is to predict the production. The data was available from January 2010 to December 2018, with 183 observations. The dataset includes the state_Name, Raj_district, Year, Season, Crop type, Area, Production, month-wise production on 2018, production in 33 district on Rajasthan. District names: Ajmer, Alwar, Banswara, Bhartpur, Bhilwara, Bikaner, Bundi, Chittorgarh, Dausa, Dholpur, Dungarpur, Sri Ganganagar, Hanmangarh, Jaipur, Jhalwar, Karauli, Kota, Rajsamand, Sawai Madhopur, Sirohi, Sikar, Udaipur.

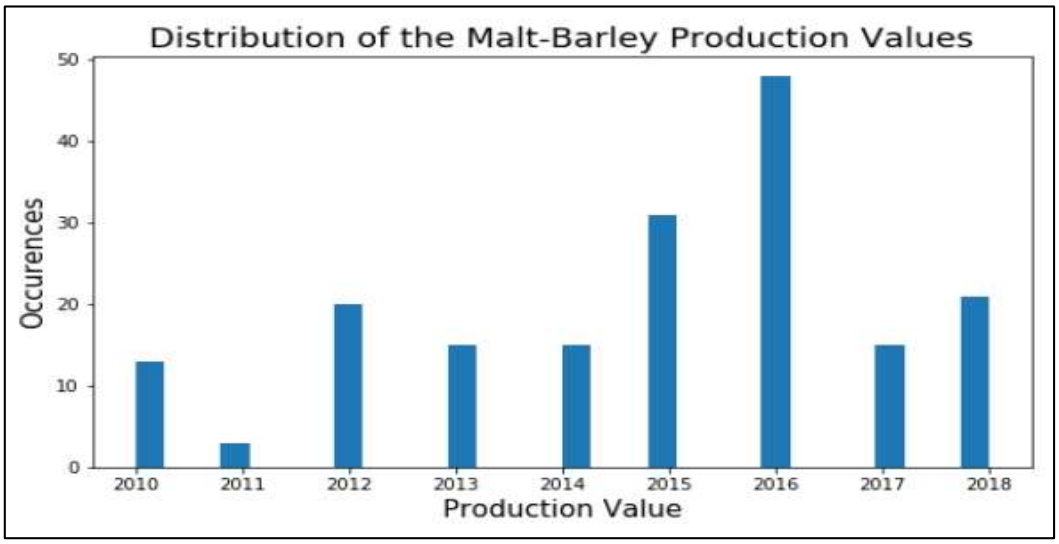
22]:

	Unnamed: 0	State_Name	ind_district	Crop_Year	Season	Crop	Area	Production	phosphorus	X1	X2	X3	X4
0	2	Rajasthan	ajmer	2012	kharif	malt-barley	2	2	1	64	29	155.712	573.769
1	3	Rajasthan	ajmer	2015	kharif	malt-barley	13	19	1	27	43	498.700	466.700
2	4	Rajasthan	ajmer	2016	kharif	malt-barley	15	24	1	19	27	442.600	498.700
3	5	Rajasthan	ajmer	2013	kharif	malt-barley	6	12	1	24	19	377.600	442.600
4	6	Rajasthan	ajmer	2018	kharif	malt-barley	4	7	1	12	24	492.300	377.600

23]: `X_finite = X_rj[np.isfinite(X_rj["X1"])]`

Source: Researcher's computation from Secondary Data

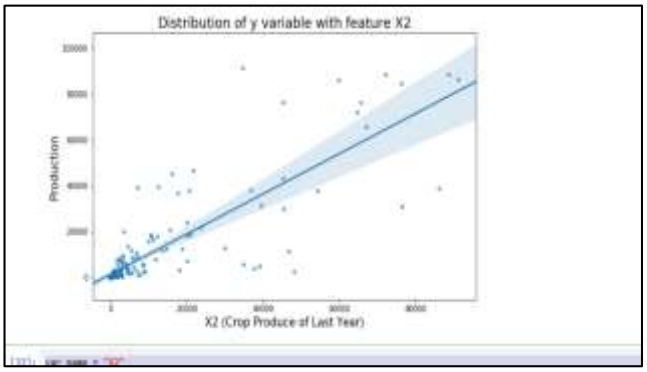
Fig 5: Snapshot of Dataset used for Malt Barley prediction



Source: Researcher's computation from Secondary Data

Fig 6: Year wise Malt-Barley production in the year 2010-2018

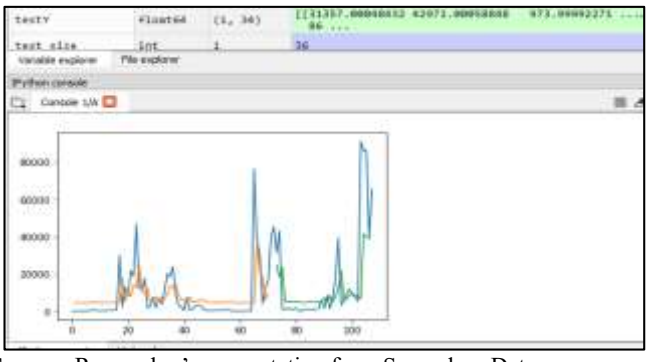
In the above snapshot, shows that overall malt production values in 33 districts. In the year 2016 Malt barley getting high production rate.



Source: Researcher's computation from Secondary Data

Fig 7: Production vs Crop produce of year 2018

The above snapshot shows the 2018 malt production values in 33 districts. Graph plot between overall production and last year's production (2018).



Source: Researcher's computation from Secondary Data

Fig 8: Prediction of Malt-barley production in 2019

In the Figure 8 green line shows predicted production value in the year 2019 for 33 Districts. It clearly shows Jaipur and Ganganagar may give better production rate.

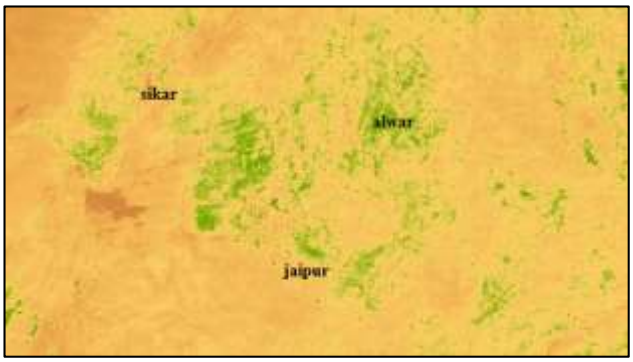
Vegetation Index-based Methods

It can be retrieved by establishing a regression equation for the field-measured FAPAR and vegetation indices such as the NDVI, which can be calculated from original images, reflectance images, and images after atmospheric correction. This method is convenient and flexible, but plant absorption changes with seasonal changes and is sensitive to vegetation type, growth stages, site environment, and other factors. Thus, the applications of this type of model are limited.

Studies have shown that the FAPAR and NDVI have a linear relationship under certain conditions. Myneni *et al.*, (1994) studied the relationship between the NDVI and FAPAR under different soil and atmospheric parameters using the transfer method. If the study is limited to the nearby sub satellite point, the effect of the atmospheric parameters and that of bidirectional reflectance can be ignored, and the influence of background can also be ignored if soil reflectance is medium. Therefore, Myneni *et al*., suggested that a linear relationship between the NDVI and FAPAR is tenable when the solar zenith angle is less than 60 degree, the observation angle of the nearby sub satellite point is less than 30 degree, the soil background reflectance is moderate (i.e., the approximate NDVI is 0.12), and the atmospheric optical thickness is less than 0.65 at 550 nm.

Roujean and Breon (1995) simulated the radiative transfer process within the canopy and surface reflection using the SAIL model, focusing on the relationship between the NDVI and FAPAR under different solar zenith angles, observation zenith angles. and relative azimuth angles. They found that the relationship will improve when the solar zenith angle or view zenith angle increases. The route of the light lengthens when it enters from the side, and the impact of the background reduces; however, it will also induce NDVI saturation for large LAI values.

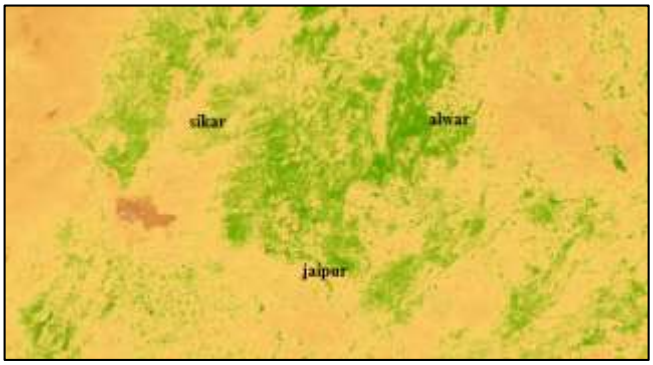
Malt barley Vegetation index in Rajasthan (2010 to 2018).



Source: Researcher's computation from Secondary Data

Fig 9: Crop Vegetation index of Sikar, Alwar and Jaipur from 01/01/2010 to 31/12/2010

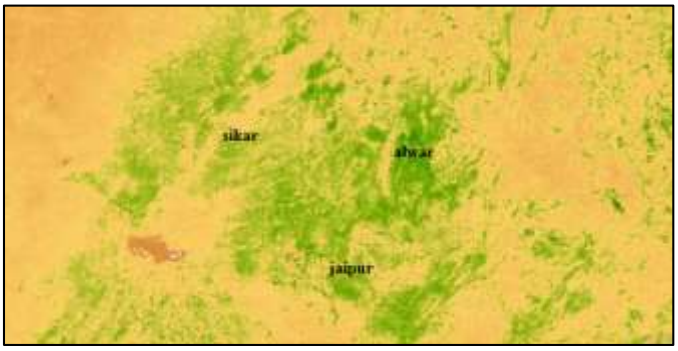
The crop index in the year 2010 at these 3 cities were presented in Figure 9 At Sikar and Jaipur, the crop index was very low as it is indicated by majority in orange colour. But Alwar has some greenish area which shows that it had some good crop index.



Source: Researcher's computation from Secondary Data

Fig 10: Crop Vegetation index of Sikar, Alwar and Jaipur from 01/01/2011 to 31/12/2011

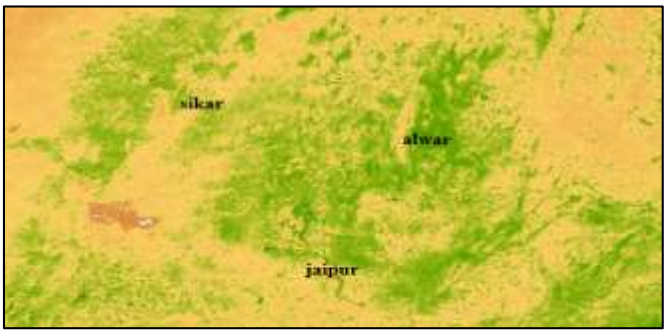
The crop index in the year 2011 at these 3 cities are presented in Figure 10, At Sikar and Jaipur, the crop index was little higher compared to 2010 as it is indicated by some greenish area. Alwar had some increased crop index when compared to the previous year as it had lot greener area.



Source: Researcher's computation from Secondary Data

Fig 11: Crop Vegetation index of Sikar, Alwar and Jaipur from 01/01/2012 to 31/12/2012

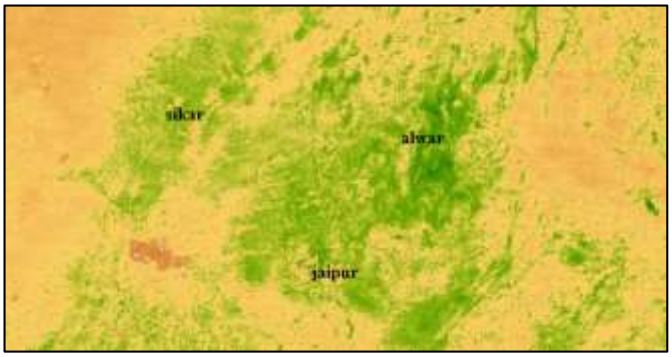
The crop index in the year 2012 at these 3 cities are shown in Figure 11, At Sikar and Jaipur, the crop index was higher compared to 2011 as it is indicated by lot more greenish area. Alwar had some increased crop index when compared to the previous year as it had huge greener area.



Source: Researcher's computation from Secondary Data

Fig 12: Crop Vegetation index of Sikar, Alwar and Jaipur from 01/01/2013 to 31/12/2013

The crop index in the year 2013 at these 3 cities are shown in Figure 4.12 At Sikar and Jaipur, the crop index was higher compared to 2012 as it is indicated by lot more greenish area. Alwar had some increased crop index when compared to the previous year as it had huge greener area.



Source: Researcher's computation from Secondary Data

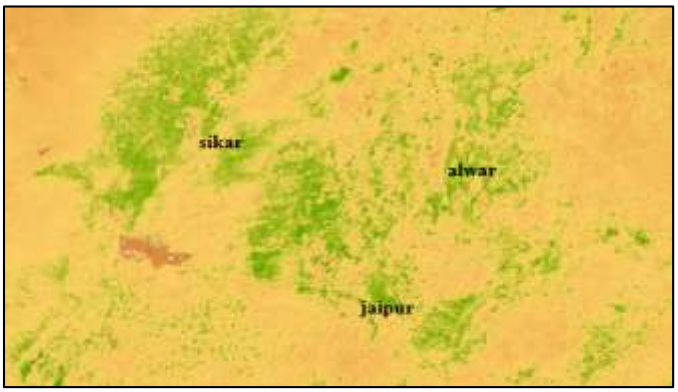
Fig 13: Crop Vegetation index of Sikar, Alwar and Jaipur from 01/01/2014 to 31/12/2014

The crop index in the year 2014 at these 3 cities are shown in Figure 13, At Sikar and Jaipur, the crop index was higher compared to 2013 as it is indicated by lot more greenish area. Alwar had some increased crop index when compared to the previous year as it had huge greener area.

Source: Researcher's computation from Secondary Data

Fig 14: Crop Vegetation index of Sikar, Alwar and Jaipur from 01/01/2015 to 31/12/2015

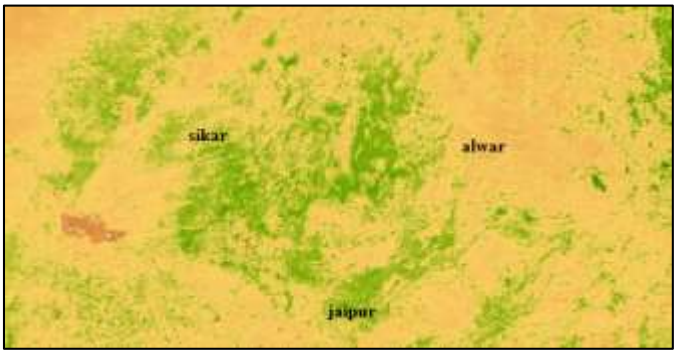
The crop index in the year 2015 at these 3 cities are shown in Figure 14, At Sikar and Jaipur, the crop index was higher compared to 2014 as it is indicated by lot more greenish area. Alwar had some increased crop index when compared to the previous year as it had huge greener area.



Source: Researcher's computation from Secondary Data

Fig 15: Crop Vegetation index of Sikar, Alwar and Jaipur from 01/01/2016 to 31/12/2016

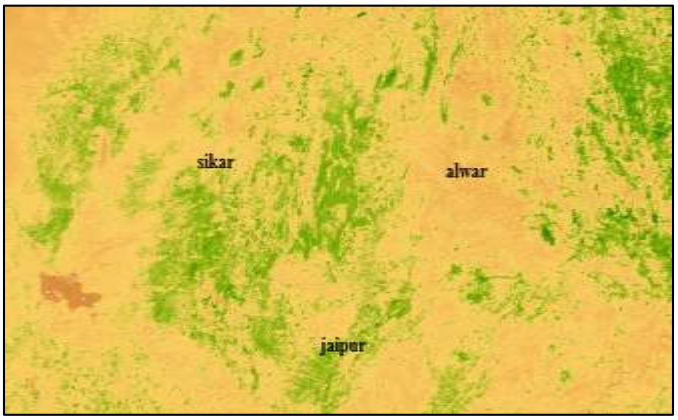
The crop index in the year 2016 at these 3 cities are shown in Figure 15, At Sikar and Jaipur, the crop index was reduced compared to 2015 as it is indicated by decrease more greenish area. Alwar had some decreased crop index when compared to the previous year as it had reduced greener area.



Source: Researcher's computation from Secondary Data

Fig 16: Crop Vegetation index of Sikar, Alwar and Jaipur from 01/01/2017 to 31/12/2017

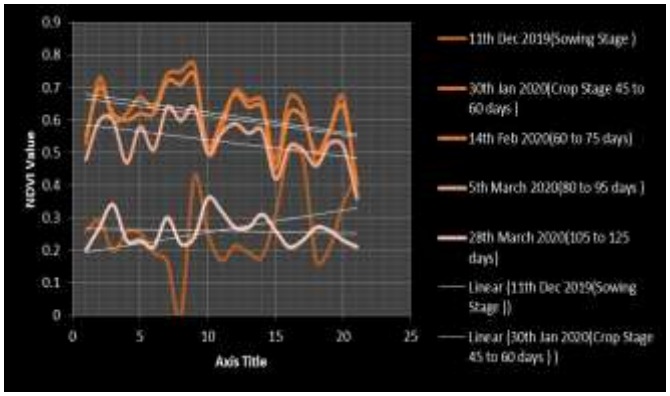
The crop index in the year 2017 at these 3 cities are shown in Figure 16, At Sikar and Jaipur, the crop index was up and down compared to 2016 as it is indicated by decrease more greenish area. Alwar had some decreased crop index when compared to the previous year as it had reduced greener area.



Source: Researcher's computation from Secondary Data

Fig 17: Crop Vegetation index of Sikar, Jaipur and Alwar from 01/01/2018 to 31/12/2018

The crop index in the year 2018 at these 3 cities are shown in Figure 17, At Sikar and Jaipur, the crop index was up and down compared to 2016 as it is indicated by decrease more greenish area. Alwar had drastic decrease in crop index when compared to the previous year as it had reduced greener area.
It is found that the crop index was seen with a pattern of increase in 2010-2015 and then decreased from 2015-2018, the main reasons behind this is lack of rainfall in those regions and fertility of the soil was also a factor due to which the crop index decreased showing decrease in the productivity.



Source: Researcher's computation from Secondary Data

Fig 18: NDVI Value of the Crop

From figure 18, it is evident that NDVI is at the lowest during the land preparation and seed sowing stage. During the crop stage of 45 to 60 days, the NDVI value is the highest at 0.77. A gradual decline is noted as the crop grows and matures reaching the harvesting stage as observed between 60 to 75 days and 80 to 95 days.
In order to determine "natural" increasing conditions in an area for a given time of year, NDVI values can be averaged over time. The health of vegetation in that position will then be defined relative to the norm by further study. NDVI can show where vegetation thrives and where it is under threat when studied over time, as well as changes in vegetation due to human activity such as deforestation, natural disasters such as wild fires, or changes in the phenological process of plants.

Conclusion

Currently, big data become more powerful with the development of the technology. Many industries have already used the analysis of the big data with the aid of technology. An agronomy-based framework for performing agricultural analytics has been explained. Also, with the recent development in the field of technology, platforms and tools. Integrating them with big data analytics and implementing them in the field of malt barley production has been introduced, explained and reviewed in detail. Finally, a few challenges have been discussed, with possible solutions to overcome it. As big data analytics is vastly adopted by many industries especially in agriculture, it has to face many challenges. As a result of these challenges, directions for the future are revealed for further improvement.

References

1. Ben Foley what is regression analysis and why should I use it? 2009 Retrieved at: https://www.surveygizmo.com/resources/blog/regression-analysis/ assesses as on 05.08.2018
2. Berhane L, Hailu G, Fekadu A. Barley production and research. Barley research in Ethiopia: Past work and

future prospects. Proceedings of the first Barley Research Review Workshop, 16–19 October 1993. IAR/ICARDA, Addis Ababa, Ethiopia 1996.
3. Ch. Chandra Sekhar J, Uday Kumar B, Kishor Kumar, Ch. Sekhar. "Effective use of Big Data Analytics in Crop planning to increase Agriculture Production in India," International Journal of Advanced Science and Technology 2018;113(9)
4. Elgendy N, Elragal A. "Big Data Analytics," in 14th Industrial Conference, ICDM 2014.
5. Kumari SV, Bargavi Dr. P, Subhashini U. "Role of Big Data Analytics in Agriculture," International Journal of Computational Science, Mathematics and Engineering, 2016;3(3).
6. Legesse G, Debebe S, Alemu T. Assessing the comparative advantage of malt barley production in Ethiopia. Application of a Policy Analysis Matrix. In African Crop Science Conference Proceedings 2007;8(1):1227-1230.
7. Mulatu B, Grando S. Barley Research and Development in Ethiopia. Proceeding of the 2nd National Barley Research and Development Review Workshop. 28–30 November 2006, HARC, Holetta, Ethiopia. ICARDA, Aleppo, Syria 2011.
8. Omo-Ojugo E. "Relevance of Big Data Analytics in Agriculture: Focus on Nigeria Agricultural Sector," International Journal of Science and Research 2017;7(11).
9. Patel S, Sayyed IU. "Impact of Information Technology in Agricultural Sector," International Journal of Food, Agriculture and Veterinary Sciences 2014;4(2):5.
10. Richard Hogan. how to calculate uncertainty 2019. https://www.isobudgets.com/calculate-test-uncertainty-ratio/, assessed as on 28.09.2019.
11. Shekhar S, Schnable P, LeBauer D, Baylis K, VanderWaa K. "Agriculture Big Data (AgBD) Challenges and Opportunities from Farm to Table: A Midwest Big Data Hub Community Whitepaper," Midwest Big Data Hub workshop 2017.
12. USAID (United States Agency for International Development), Agricultural Growth Program-Agribusiness and Market Development (AGP-AMDe). Addis Ababa, Ethiopia, 2014.
13. Weldeyohanis, Rijalu, Abddisa, Fikadu. Value Chain Analysis of Malt Barley (Hordeumvulgarel.): A Way Out for Agricultural Commercialization? The Case of Lemu Bilbilo District, Oromia Region, Ethiopia. Journal of Economics and Sustainable Development 2017;8:81-97.