# Environmental Monitoring and Agricultural Insights: Analysis of Cotton Crop using PowerBI

Syeda Faiza Nasim<sup>a</sup>, Muhammad Khurram<sup>b</sup>, Aiza Kamran<sup>c</sup>, Syeda Alishba Fatima<sup>d</sup> and Asma Qaiser<sup>e</sup>

<sup>e</sup> Department of Computer Science, Iqra University, Karachi 76400, Pakistan Email:<u>asma.gaiser@iqra.edu.pk</u>

Received: 12/09/2023, Revised: 22/11/2023, Accepted: 15/12/2023

Abstract— To understand the variables affecting Pakistan's cotton output, our research paper study focuses on applying the PowerBI tool to analyze data related to the cotton crop. We gathered data from two fields in Rahimyar Khan and Shah Alam Shah, Matiari, Sindh, to analyze the soil moisture content, availability of fertilizer, and environmental factors to improve agricultural practices and increase crop yields. The dataset contains data from monitoring dates and factors like temperature, humidity, soil moisture content, and signal intensity. We forecast cotton crop output, improve planting schedules, and foresee possible issues like bug outbreaks using predictive analytics. The study offers practical suggestions for decision-making procedures about fertilizer application schedules, irrigation schedules, and the sustainability of cotton crops. Limitations include data quality and scalability challenges, and future research will focus on improving agricultural techniques for better cotton growing.

## Index Terms— Soil Moisture, Cotton Crop, Power BI, Agricultural Sustainability, Irrigation, 7 in 1 sensor.

## I. INTRODUCTION

E have selected the cotton crop for data collection and analysis in this study, due to its economic importance and widespread cultivation in Pakistan. Cotton is an important cash crop in Pakistan, with considerable economic and agricultural benefits [1-6]. By concentrating on cotton crop data, researchers can get insight into elements influencing cotton production such as environmental conditions, soil moisture levels, and fertilizer availability. This research can also provide useful information for improving agricultural operations, maximizing crop yields, and increasing cotton farming sustainability [7 - 12]We collect cotton crop data from two distinct fields in Pakistan. The data is gathered using a 7-in-1 sensor. These dataset values were taken in Rahimyar Khan and Shah Alam Shah, Matiari, Sindh with crop 2X in Pakistan, with coordinates at 28°15'6.44"N and 70°25'3.73"E with sowing a at 5 May 2023 covering 41.4 acres, 28°10'58.55"N, 70°18'28.66"E with sowing date 1 20thMay 2023 covering 52.6 acres, and 25°33'52.85"N, 68°26'41.39"E with sowing date 25 May 2023 covering 112 acres. The monitoring dates for the above-mentioned coordinates are May 8, May 15, and May 28, 2023. The dataset was further categorized into Data Added, Humidity, API Humidity, Temperature, API Temperature, Heat Index, Soil Moisture, Soil M Contain Moisture Raw Value, Soil Temp, Soil\_EC Raw, Signal Strength, and APIKey. SMSAck N\_For\_User P\_For\_User,K\_For User, Nitrogen, Phosphorus, Potassium, N\_From\_Admin, P\_From\_Admin, K From Admin, N\_From\_EC, P\_From\_EC, K From EC, Salinity, battery voltage, water requirement, and reset status [13-20].

## II. DATA DESCRIPTION

Device ID 205 has the largest dataset by both row count (2999) and data size (469 KB). This shows that Device ID 205 collects more data or at a faster rate than the other devices. Device ID 170 follows with a smaller dataset than Device ID 205, consisting of 1791 rows and 311.37 KB of data. Device 171 has the smallest dataset among the three, with 1618 rows and 239 KB of data [21-28].



<sup>&</sup>lt;sup>a</sup> Department of Computer Science & Information Technology, NED University of Engineering and Technology, Pakistan, Ph. +92-21-99261261, Email:<u>sfnasim@cloud.neduet.edu.pk/sfaizaadnan@gmail.com</u>

<sup>&</sup>lt;sup>b</sup> Department of Computer and Information Systems Engineering, NED University of Engineering and Technology, Pakistan, Ph. +92-21-99261261, Fax: +92-21-99261255, Email:<u>mkhurrum@neduet.edu.pk</u>.

<sup>&</sup>lt;sup>c,d</sup> Department of Computer Science & Information Technology, NED University of Engineering and Technology, Pakistan, Ph. +92-21 99261261,Email:<u>alishbafatima73@gmail.com,aizakamran709@gmail.com</u>,

# III. LITERATURE REVIEW

Examining how technological advances, analytics of data, and agriculture interact, our research study supports important ideas that highlight the value of using Power BI to conduct thorough agricultural data analysis. Below here Table 1 provides insights of the background study with its main target findings.

provides insights of the background study with its main target findings.					
Target Areas	Summary	Research Insights			
Integration of IoT Devices [9]	Enhances real- time data collection and monitoring in agriculture through IoT device integration within Power BI, providing timely insights into environmental conditions, crop health, and irrigation requirements.	Real-time monitoring insights enable rapid decision- making on environmental conditions, soil moisture levels, and fertiliser availability, critical for optimal crop management.	Data Visualization Techniques [10]	Utilizes effective data visualization techniques in Power BI to communicate complex agricultural data insights to stakeholders, facilitating informed decision- making through interactive visualizations, scatter plots, trend	Interactive visualizations help illustrate data linkages and trends comprehensively, aiding in decision- making processes [26, 30].
Machine Learning Algorithms [10]	Empowers stakeholders with predictive	Correlation study reveals relationships		lines, and correlation coefficients.	
	analytics capabilities by incorporating machine learning algorithms in Power BI, enabling the identification of patterns and trends in agricultural data for informed decision-making and resource optimization. [23- 25]	between critical variables like humidity, temperature, soil moisture, and nutrient levels across devices, aiding in data- driven decision- making.	Cross-Domain Data Integration [13]	Explores the benefits of integrating agricultural data with weather forecasts, market trends, and soil health assessments in Power BI to provide a holistic view of agricultural operations, enabling informed decision-making	Integration of data from various sources offers a comprehensive picture of agricultural operations, enhancing decision-making.
Sustainability Practices [11]	Promotes sustainable agricultural	Insights aid in improving agricultural		and optimization of farming practices. [2-5]	
	practices such as precision farming, resource optimization, and environmental impact assessment through data analysis using Power BI, contributing to long-term productivity and	maximizing crop yields, and increasing sustainability in cotton farming.[27-29]	Data Security and Privacy [12]	Addresses challenges related to data security, privacy, and regulatory compliance when handling sensitive agricultural data in Power BI, ensuring data integrity, confidentiality,	Implementation of robust data security measures safeguards agricultural data integrity and confidentiality, fostering trust among stakeholders [31- 35].

environmental

stewardship.

	and regulatory adherence.	
User Training and Adoption[15-17]	Empowers agricultural stakeholders with user training and adoption strategies to leverage Power BI effectively for data-driven decision-making, enhancing user proficiency and promoting the adoption of data analytics tools in farming practices [6-8].	Actionable recommendations provided for irrigation scheduling, fertilizer use, and cotton crop management methods, enhancing decision-making processes. [30-31]
Case Studies and Best Practices [12]	Showcases successful implementations of Power BI in agricultural data analysis through case studies and best practices, demonstrating tangible benefits and outcomes achieved in optimizing agricultural operations, increasing productivity, and promoting sustainability.	Insights from case studies guide decision-making, increase production, and promote sustainable agriculture techniques.

Table 1- Literature Findings

#### IV. MATH

We used data from Rahim Yar Khan (May 5, 2023, and May 10, 2023) and Shah Alam Shah (Matiari, Sindh, August 25, 2023) for our "Exploratory Data Analysis (EDA) Process." We created summary statistics and used KDE's pair charts and correlation matrices (heat map) to show the data correlations between each attribute. Applied visualisation and analytics using the pandas and Seaborn libraries on Microsoft Power BI desktop, as well as cleaning and manipulating the dataset in

Power Query Editor. To prepare the dataset for further research or data modelling, we took considerable effort to identify and remove outliers as illustrated here in Figure 1,2 and 3 [1]



Figure 1 - Device ID: 171



Figure 2 - Device ID: 170



A. RELATIONSHIP BETWEEN TEMPERATURE AND HUMIDITY

In Power BI Desktop, the scatter plot depicting the relationship between Temperature and Humidity suggests a discernible pattern as illustrated below here in Figure 4,5 and 6, revealing a moderate negative correlation with a correlation coefficient of -0.6642,-0.66 and -0.61 from device Id 170, 171 and 205 respectively. This indicates that as temperatures fluctuate, there is a tendency for humidity to decrease, providing valuable insights into the inverse association between these two variables [36-37].



Figure 4 - Device ID: 170







### B. Temperature vs API Humidity:

Implemented scatter plot in Power BI with temperature on X-axis and API humidity on y -axis. Since the correlation coefficient is negative: -0.046570113,-0.1for device Id 170 and171 respectively. It shows there is a weak negative correlation between them. The trend observed indicates that as temperature increases, API humidity tends to decrease slightly. Similarly for device ID 205 The value of correlation coefficient is 0.01, which shows there is a very weak positive linear relationship between temperature and API humidity. The correlation is closer to zero, suggesting that there is almost no linear association between the two variables in this case. As shown below here in Figure 7,8 and 9.



Figure 7- Device ID: 170



## C. Humidity vs API Humidity:

We generate another scatter plot in Power BI with humidity on x-axis and API humidity on Y-axis showing correlation coefficient. Device Id 170,171,205 depict positive correlation between humidity and API humidity with correlation coefficient of 0.1, 0.5, 0.64 respectively. The trend lines indicate that as humidity increases, API humidity tends to increase as well as shown below here in Figure 10,11 and 12.



Figure 10- Device ID: 170



Figure 11- Device ID: 171



Figure 12- Device ID: 205

## D. Temperature vs Heat Index:

Utilizing a Power BI scatter plot with temperature on X and heat index on Y-axis we observe a notable strong positive correlation coefficient 0.79386607, 0.7,0.89 between temperature and heat index of device Id 170,171 and 205 respectively. Enhance the visualization with trend lines for clearer insights. As trend suggests that as temperature increases, the heat index also tends to rise significantly.

#### E. API Temperature and Heat Index:

With a correlation coefficient of 0.701720004,0.6 and 0.84 indicating a strong positive correlation, of device Id 170,171 and 205 respectively. There is a clear relationship between API temperature and heat index. Employ a scatter plot in Power BI, placing API temperature against the heat index. The pattern observed shows an increase in API temperature is associated with a noticeable rise in the heat index. As shown in Figures 13,14 and 15.



Figure 13- Device ID: 170



Figure 14- Device ID: 171



#### F. HUMIDITY AND HEAT INDEX

Implemented Power BI scatter plot with humidity on X and heat index on Y o device Id 170,171 and 205 using correlation coefficient of -0.457539308, -0.3 and -0.37 respectively signifying a moderate negative correlation, there is a noteworthy relationship between humidity and heat index. The trend suggests that as humidity decreases, the heat index tends to increase moderately, shown in Figures 16,17 and 18.



## G. RELATIONSHIP BETWEEN TEMPERATURE AND API **TEMPERATURE**

With Power BI, we create a scatter plot for device ID 170,171 and 205 temperatures on X-axis and API temperature on Y Axis and vice versa to visualize the correlation coefficient of 0.91, 0.9.0.89 suggesting a strong positive relationship between the two variables. Observed Trend: The trend lines indicates that as the temperature increases, the API temperature also increases, and vice versa, as shown in Figures 19, 20 and 21.







Figure 21- Device ID: 205

## H. Trends of Soil Moisture on Nutrient levels

The plots visualize the relationship between soil moisture and nutrient levels for each nutrient. Three separate Scatter plots are created in Power BI Desktop for Nitrogen (N), Potassium (K) and for phosphorus (P). The scatter plot for nitrogen shows how nitrogen levels in the soil vary with changing soil moisture. Similarly, the scatter plot for potassium or phosphorus illustrates the relationship between soil moisture and the respective nutrient levels. Analyzing these plots helps in understanding how Soil moisture impacts the availability or concentration of each nutrient, aiding in agricultural or environmental assessments. As shown in Figures 22 and 23. Device ID 170:



Figure 22- Impact On Soil Moisture



Figure 23- Relationship between Nutrient values on Soil Moisture

Device ID 171:



Figure 24- Impact On Soil Moisture



Figure 25- Relationship between Nutrient values on Soil Moisture

#### Device ID 205:



Figure 26- Impact On Soil Moisture



Figure 27- Relationship between Nutrient values on Soil Moisture

#### V. RESULTS

Our Analysis of agricultural data using Power BI generated considerable insights across multiple dimensions. Correlation studies revealed correlations between critical variables such as humidity, heat index, temperature, API temperature, soil moisture, and nutrient levels across multiple devices. Trend analysis revealed trends that represent the impact of changing climatic elements on agricultural metrics, such as the positive correlation between temperature and API temperature, as well as the link between soil moisture and nutrient levels. Real-time monitoring insights enabled rapid decision-making on environmental conditions, soil moisture levels, and fertilizer availability, which is critical for optimal crop management. Interactive visualizations, such as scatter plots, trend lines, and correlation coefficients, helped to comprehensively illustrate data linkages and trends. Plotting a line chart in Power BI, a line graph plotted represents the trend over a one-year period. The x-axis represents time, likely labelled with months, covering the entire 2023 year and y-axis displays the levels of soil moisture and the respective nutrient (nitrogen, potassium, or phosphorus). The line on the chart depicts how the nutrient levels change with different devices over the course of the year in correlation with soil moisture. Observing the line provides insights into seasonal variations or patterns in nutrient levels based on the soil's moisture content as seen below in Figures 28, 29 and 30.



Figure 30- Device ID: 205

Comparing all above-mentioned data from multiple devices enabled the detection of fluctuations in agricultural factors, which aided in the optimization of farming operations and resource allocation. Furthermore, these predictive analytics using current data from different devices allowed for the forecast of cotton crop yields, the optimization of planting schedules, and the anticipation of possible concerns such as bug outbreaks, thereby increasing agricultural output.

#### VI. DATA AVAILABILITY:

Repository name: Zenodo

Data identification number: 10.5281/zenodo.8371502 Direct URL to data: https://doi.org/10.5281/zenodo.8371502 Repository name: figshare

Data identification number: 10.6084/m9.figshare.24186024.v1

#### VII. LIMITATIONS

The limitations of our research paper include data quality issues, limited data sources, a small sample size, reliance on assumptions, the need for technical skill, the potential impact of external factors, data security concerns, and scalability constraints.

#### VIII. FUTURE WORK:

These insights presented in this research paper gave actionable recommendations to aid decision-making processes about irrigation schedule, fertilizer use, and overall cotton crop management methods. Furthermore, the integration of data from various sources, including weather forecasts, soil sensors, and crop management systems, provided a comprehensive picture of agricultural operations and performance. Overall, these findings help to guide decision-making, increase production, and promote the use of sustainable agriculture techniques.

#### REFERENCES

- [1] McKinion, J.M., Jenkins, J.N., Akins, D., Turner, S.B., Willers, J.L., Jallas, E., & Whisler, F.D. (2001). Analysis of a precision agriculture approach to cotton production. Computers and Electronics in Agriculture, 32(2), 213-228. https://doi.org/10.1016/S0168-1699(01)00166-1
- Hodges, H.F., Whisler, F.D., Bridges, S.M., Reddy, K.R., McKinion, J.M. [2] (1997). Simulation in crop management: GOSSYM/COMAX. In: Peart, R.M., Curry, R.B. (Eds.), Agricultural Systems Modeling and Simulation. Marcel Decker, New York, pp. 235-281.
- [3] Jallas, E., Akins, D.C., Turner, S., Gourley, K.E., McKinion, J.M., Willers, J.L., Clouvel, P., Cretenet, M., Sequeira, R., Jenkins, J.N. (1999). Precision Farming, Myth or Reality: a Case Study in Mississippi Cotton Fields. Proceedings of Beltwide Cotton Production and Research Conferences, Orlando, FL, 3-7 January, 1999.
- McKinion, J.M., Jenkins, J.N., Akins, D., Turner, S.B., Willers, J.L., [4] Jallas, E., Whisler, F.D. (2001). Analysis of a precision agriculture approach to cotton production. Computers and Electronics in Agriculture, 32(2), 213-228.
- Reddy, K.R., McKinion, J.M. (1997). GOSSYM/COMAX. [5]
- Sequeira, R.A., Jallas, E. (1995). The simulation Model GOSSYM and its [6] decision support system, COMAX: its applications in American agriculture. Agric. Dev. 8, 25-34.
- Steiner, J.L., Williams, J.R., Jones, O.R. (1987). Evaluation of the EPIC [7] simulation model using a dryland wheat-sorghum-fallow crop rotation. Agron. J. 79, 732-738.
- Wilkerson, G.G., Jones, J.W., Boote, K.J., Ingram, K.T., Mishoe, J.W. [8] (1983). Modeling soybean growth for crop management. Trans. ASAE 26, 63-73.
- Williams, J.R., Jones, C.A., Kiniry, J.R., Spanel, D.A. (1989). The EPIC [9] crop growth model. Trans. ASAE 32 (2), 497-511.
- [10] Zhang, N., Taylor, R, Schrock, M., Starrenbogr, S. (1999). Application of a Field-Level Geographic Information System (FIS) in Decision Making for Precision Agriculture. ASAE Pap. 99-3046. ASAE, St. Joseph, MI.
- [11] T.Weng, Y.Xie, G.Chenetal., "Loadfrequencycontrolunder false data inject attacks based on multi-agent system method in multi-area power systems," International Journal of Distributed Sensor Networks, vol. 18, no. 4, pp. 155013292210904-155013292214618, 2022.
- [12] B.Tom'si'c, D.Markovic', V.Jankovic'etal., "Biodegradationofcellulosefib ersfunctionalized with cuo/cu2on an oparticles in combination with polycarboxylic acids," Cellulose, vol. 29, no. 1, pp. 287-302, 2022.
- [13] Z.Wu,X.Li,X.Liuetal.,"Membraneshellpermeabilityofrs198 microcapsules and their ability for growth promotingbioactivitycompoundreleasing,"RSCAdvances,vol.10,no.2, pp. 1159-1171, 2020. [4] A. Feng, J.Zhou, E.D. Vories, K.A. Sudduth, and

M.Zhang, "Yield estimation in cotton using uav-based multi-senso imagery,"BiosystemsEngineering,vol.193,no.3,pp.101–114, 2020.

- [14] S. Ale, N. Omani, S. K. Himanshu, J. P. Bordovsky, K. R. Thorp, and E. M. Barnes, "Determining optimum irrigation termination periods for cotton production in the Texas high plains," Transactions of the American Society of Agricultural and Biological Engineers, vol. 63, no. 1, pp. 105–115, 2020.
- [15] M. Fouzi, M. Thimma, M. Binsabt, A. A. Husain, and S. Aouabdi, "Stem cell growth and proliferation on rgd bioconjugated cotton fibers," Bio-Medical Materials and Engineering, vol. 32, no. 1, pp. 39–52, 2021.
- [16] M. Ku"çu"k and M. L. "Oveço" glu, "Fabrication of sio2–zno np/ zno nr hybrid coated cotton fabrics: the effect of zno nr growth time on structural and uv protection characteristics," Cellulose, vol. 27, no. 3, pp. 1773– 1793, 2020.
- [17] C.A.Moraes,L.W.deOliveira,E.J.deOliveira,D.F.Botelho, A. N. de Paula, and M. F. Pinto, "A probabilistic approach to assess the impact of wind power generation in transmission networkexpansionplanning,"ElectricalEngineering,vol.104, no. 2, pp. 1029–1040, 2021.
- [18] C.Caro-Ruiz,A.S.Al-Sumaiti,S.Rivera,andE.Mojica-Nava, "A mdpbased vulnerability analysis of power networks considering network topology and transmission capacity," IEEE Access, vol. 8, no. 1, pp. 2032–2041, 2020.
- [19] Y. Wang, J. Wang, L. Yao, and W. Y. Yin, "Emi analysis of multiscale transmission line network using a hybrid fdtd method," IEEE Transactions on Electromagnetic Compatibility, vol. 63, no. 4, pp. 1202– 1211, 2021.
- [20] W. Zhou, Y. Wang, and Z. Chen, "Impedance-decoupled modeling method of multiport transmission network in inverter-fed power plant," IEEE Transactions on Industry Applications, vol. 56, no. 1, pp. 611–621, 2020.
- [21] J.PonockoandJ.V.Milanovic, "Multi-objectivedemandside management at distribution network level in support of transmissionnetworkoperation,"IEEETransactionsonPower Systems, vol. 35, no. 3, pp. 1822–1833, 2020.
- [22] D. Bose, C. K. Chanda, and A. Chakrabarti, "Vulnerability assessment of a power transmission network employing complex network theory in a resilience framework," Microsystem Technologies, vol. 26, no. 8, pp. 2443–2451, 2020.
- [23] E. Herna'ndez-Rodr'iguez, L. H. Escalera-Va'zquez, E Mendoza, D. Garc'ia-'Avila, and M. Montoro Girona, "Reduced-impact logging maintain high moss diversity in temperate forests," Forests, vol. 12, no. 4, pp. 383–419, 2021.
- [24] C. L. Wiegand. The value of direct observations of crop canopies for indicating growing conditions and yield. The 18th International Symposium on Remote Sensing of Environment, 18, 1551 (Paris, 1984).

- [25] N. A. Quarmby, M. Milnes, T.L. Hindle and Silleos. The use of multitemporal NDVI measurements from AVHRR data for crop yield estimation and prediction. Inter. J. Remote Sens., 14, 199 (1993)
- [26] I. Ahmad, D. A. Awan, M. Bhatti, I. H. Akhtar, M. Ibrahim. Satellite Remote Sensing and GIS based Crop Forecasting System in Pakistan. Crop monitoring for improvement food security, 95 (FAO, 2014)
- [27] T. Sh. Beisenboyev, N.F. Bespalov, Salinization dynamics of irrigated soils and cotton productivity. Tashkent (1993)
- [28] Annual report of Ministry of Water Resources of the Republic of Uzbekistan (2018)
- [29] E3S Web of Conferences 227, 03001 (2021) GI 2021. <u>https://doi.org/10.1051/e3sconf/202122703001</u>
- [30] P. Kingra, D. Majumder. Application of Remote Sensing and GIS in Agriculture and Natural Resource Management Under Changing Climatic Conditions. Agri. Research J., 53(3), 295 (2016)
- [31] Z. J. Mamatkulov, E.Yu. Safarov, R.K. Oymatov, I.I. Abdurahmanov. Application of GIS and Remote Sensing in crop monitoring and yield forecasting in case of lowyielded farmlands. J. Problems Architecture. Const., Samarkand, Spatial Volume, 130 (2019)
- [32] C. Atzberger. Advances in remote sensing of agriculture: context description, existing operational monitoring systems and major information needs. J. Remote Sens, 5, 949 (2013)
- [33] T. Bernardes, M. A. Meriera, M. Adami, A. Giarolle and B. F. Rudorff. Monitoring biennial bearing effect on coffee yield using MODIS remote sensing imagery. Journal of Remote Sens, 4, 2492 (2012)
- [34] T. Sakamoto, M. Yokozawa, H. Toritani, M. Shibayama, N. Ishitsuka. A crop phenology detection method using time series MODIS data. Journal of Remote Sens Environ, 96, 366 (2005)
- [35] S. Macdonald, United States Department of Agriculture. Retrieved from www.usda.gov/oce/forum (2018) 12. Y. Huang, S. Thomson, Remote Sensing for Cotton Farming (2015)
- [36] N. Ramarao, Conformity Analysis of Cotton Crop using Remote Sensing and GIS, (Geospatial World, 2009) 14.
- [37] D. Atakhanov, Cotton yield forecasting in Tashkent province by using Remote Sensing techniques.https://www.wur.nl/en/activity/Cotton-yieldforecasting-in-Tashkentprovince-by-using-Remote-Sensingtechniques.htm (2013)