

Application of Quality Management System in the Research Process: A Case Study for Plant Phenotyping Research

Alexander Galba , Eva Kánská , Vojtěch Mikeš , Jiří Vaněk , Jan Jarolímek 

Department of Information Technologies, Faculty of Economics and Management, Czech University of Life Science Prague, Czech Republic

Abstract

Phenomics research, driven by advancements in imaging and image processing, enables high-throughput measurements of plant traits, providing insights into growth, tissue development, and biochemical states. However, data accuracy is critical to reliable outcomes, especially in complex methods like 3D reconstruction and hyperspectral imaging. This study demonstrates the role of Quality Management Systems (QMS) in enhancing the research process in plant phenotyping. The study emphasizes the importance of a robust data quality assurance pipeline, focusing on error identification and improving data labeling processes through semi-automation. Root Cause Analysis (RCA) was employed to address discrepancies in annotated datasets and identify critical issues, such as misalignment in experimental protocols and operational errors, including the misplacement of irrigation hoses during data collection. Corrective actions, such as photo documentation and procedural revisions, significantly improved data quality. Additionally, algorithmic support streamlined the annotation process, increasing efficiency and data reliability. This integrated approach underscores the necessity of quality control in research, especially for geographically distributed teams working under variable conditions, and highlights the broader applicability of QMS in optimizing research outputs.

Keywords

Quality management system, data quality, plant phenotyping, research process, root cause analysis, data labeling process.

Galba, A., Kánská, E., Mikeš, V., Vaněk, J. and Jarolímek, J. (2024) "Application of Quality Management System in the Research Process: A Case Study for Plant Phenotyping Research", *AGRIS on-line Papers in Economics and Informatics*, Vol. 16, No. 4, pp. 79-86. ISSN 1804-1930. DOI 10.7160/aol.2024.160406.

Introduction

Phenomics research has benefited from the methods and devices allowing the high throughput measurements of plant traits at different levels, including growth-related traits, cells and tissue development, and biochemical and physiological states. Imaging and image processing developments enable capturing how these traits vary over time. Advancements in parallel and automated image acquisition allow for processing images of large plant populations under specific growth conditions. Image-processing capabilities enable 3D reconstruction of image data and automated quantification of biological features. (Simek et al., 2015) These advancements allow modeling at the systems level (Dhondt et al., 2013; Fiorani and Schurr, 2013; Kartal et al., 2021; Paulus, 2019; Pongpiyapaiboon et al., 2023; Sozzani et al., 2014; Vadez et al., 2015).

The goal of using digital methods is to achieve the best possible accuracy in identifying the

biological parameters of the observed plants. A prerequisite for such accuracy is a high-quality database. Errors in data are the basis for poor-quality outputs or erroneous research conclusions. To achieve the best possible research results, it is necessary to pay attention to the quality of the entire research process (Ronanki et al., 2022).

The application of Quality Management Systems (QMS) in plant phenotyping research plays a pivotal role in ensuring data accuracy and consistency, especially with the increasing complexity of hyperspectral imaging and sensor-based techniques. For instance, a quality assurance pipeline developed for hyperspectral imaging systems ensures that spatial and spectral quality parameters are accurately maintained, enabling reliable detection of plant diseases such as *Cercospora* Leaf Spot through convolutional neural network (CNN)-supported data analysis. This quality-assured approach is crucial for evaluating and refining imaging systems in real-time plant health monitoring (Detring

et al., 2024). Another critical aspect of QMS in plant phenotyping is data management, essential for handling the vast amounts of digital data generated by automated sensing systems. Ensuring that data-sharing practices align with FAIR (Findability, Accessibility, Interoperability, and Reusability) principles can enhance the reproducibility and efficiency of research processes. (Ugochukwu and Phillips, 2022) The integration of advanced phenotyping methods, such as 3D reconstruction and leaf surface estimation, further highlights the need for robust data collection and analysis quality control to optimize workflows and reduce processing times (Daiki and Noshita, 2024).

The application of Quality Management Systems in the research process ensures consistent standards, minimizes errors, and enhances the reliability of outcomes. In clinical and public health research, QMS is essential to maintain protocol integrity, prevent deviations, and uphold the credibility of research findings. By implementing systematic procedures, research teams can ensure rigorous data quality and operational efficiency (Isere and Omorogbe, 2024). Additionally, in biomedical laboratories, QMS helps address quality issues, such as result stability and replication crises, by optimizing research processes and increasing effectiveness and efficiency. Despite initial resistance due to resource allocation and bureaucracy concerns, QMS offers significant advantages in process control and reliability. (Brünschwitz and Kleymann-Hilmes, 2024) Furthermore, operations research methods integrated with QMS in engineering and industrial contexts enable continuous improvement through mathematical modeling and optimization techniques, ensuring product and service quality across all phases (Parker, 2024). The application of QMS in research laboratories, such as in the petroleum industry, improves client satisfaction, reduces failure rates, and fosters industry-academia collaboration, providing a robust framework for process optimization and innovation (Vianna et al., 2022).

This article's main objective is to apply the quality

management system in the research process to increase the quality of outputs.

Material and methods

Our research in the field of plant phenotyping is conducted by an international team located worldwide in various places and time zones. Researchers are located in the Czech Republic (The Czech University of Life Sciences), the Netherlands (Phenospex), India (ICRISAT - The International Crops Research Institute for the Semi-Arid Tropics), and Turkey (Cukurova University)

The research aims to estimate plant traits from 3D scans of plants acquired by the high-throughput phenotyping platform LeasyScan (built using Phenospex PlantEye F600 sensor). Subsequently, individual plant detection, organs, and other analyses are performed mainly using 3D computer vision methods. Obtaining 3D models outdoors brings efficiency to the entire research process. It does not require excessive manipulation with plants. Part of the process involves manually annotating datasets to train the artificial neural network models. The general scheme of the entire research is depicted in Figure 1.

The research commenced following the task assignment. Time sessions were set up for regular meetings where interim results were reviewed, schedules were discussed, and next steps were suggested. Control mechanisms were also part of the regular consultations. The method of selecting and checking a sample of data was chosen for control. The conclusions from the data control were regularly consulted, and measures were taken to correct the identified deficiencies.

The following procedure was chosen for the data flow: Data acquisition -> Data preprocessing -> Data Annotation. Data acquisition was performed in India using the LeasyScan platform. The source data was then stored and managed by Phenospex's Hortcontrol system, which is also available using BrAPI. Validated algorithms were used for



Source: Author

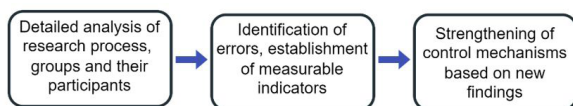
Figure 1: The general scheme of the entire research.

data preprocessing. The preprocessing consists of several steps, i.e., rotation, merging, voxelization, smoothing, and soil segmentation. The subsequent data annotation step was performed using the Segments.ai environment. All procedures were rechecked and verified (*Multi-Sensor Data Labeling Platform for Robotics and AV | Segments. Ai, n.d.; Phenospex - Smart Plant Analysis and Phenotyping Systems, n.d.*).

Over time, it became clear that the set rules needed to be revised to ensure the overall quality of the research process. Weaknesses in the quality of the data, rather than in the quality of the steps used, were identified. Regular online meetings of all groups and information sharing, including the control mechanisms, are still needed to eliminate the deficiencies.

The primary source of deficiencies was identified in the data annotation process for generating test and control datasets using artificial neural networks. The number of annotated plants for the ground-truth did not match the predictions. The first step to correcting this situation was to increase the qualification of persons designated for data labeling through regular training and more detailed checks. It turned out that this procedure only solved the consequences and not the source of the error. It was evident that it was necessary to revise the entire research process and find the source of the inaccuracies.

The control mechanisms, consisting of the selection of samples and their control, needed to be revised, and it was decided that it was necessary to find a way to check all the acquired data. At the same time, it was essential to check all steps to exclude the possibility of multiple sources of errors. To avoid further problems in the future, the methodological procedure expressed in Figure 2 Plan to improve the research process was chosen.



Source: Author

Figure 2: Plan to improve research process.

It was important for the whole process to choose a method to identify the bottleneck. During the discussion, it became clear that the methods used are mainly focused on the commercial area or processes in the public sector. Their application in research is only partially obvious. Most methods

focus on streamlining the process regarding time, costs, or risk analysis. Some methods are extensive, and their application would require longer (ISO). A common attribute of these methods is creating a detailed procedure diagram in various formats. This procedure is then supplemented with tools and procedures to detect deficiencies and thereby increase the quality of the entire process. The diagram expresses the sequence of causes and effects and the influence of the main actors. (Bali et al., 2021; *ISO 9001:2015 - Quality Management Systems - Requirements, n.d.*; Serrat, 2017; Shook, 2008; Starzyńska and Hamrol, 2013; Venkatasubramanian et al., 2003) The display allows a better understanding of the sequence of events and the identification of critical points. The methods then focus on these and increase quality by applying appropriate control mechanisms. It is a suitable tool to express the quality of the process in a measurable quantity. In practice, this may be time, the number of erroneous outputs, or a combination of similar values. Based on such a criterion, the success of the applied method can then be evaluated. An overview of the methods is shown in Figure 3 (Tari and Sabater, 2004a).

The assessment of which method is appropriate depends on the perspective of the participants and the definition of the target parameters. Finally, the point of view of the greatest possible generality in use and applicability for scientific research was chosen. From this point of view, the root cause analysis method was chosen. A Root Cause Analysis (RCA) is the process of identifying the most fundamental reason for the problem, which, if eliminated or corrected, would prevent the problem from reoccurring. Various techniques for analysis can be used for this method, and therefore, its application is suitable for multiple environments, including high-risk industries such as medicine or drug development (Andersen and Fagerhaug, 2000; Percarpio et al., 2008; Andersen, B. and Fagerhaug, T. (2006); Wilson et al., 1993; Yuniarto, 2012).

Creating a measurable parameter expressing the source data's quality and accuracy proved essential. In addition to the position in 3D space, the data from the device also contained information about the point's color. A method based on the height parameter of individual points (z-axis value) was chosen for the check. The points in each data sample were divided into 2 groups according to height. The division had the following assumptions:

The seven basic quality control tools	The seven management tools	Other tools	Techniques
Cause and effect diagram	Affinity diagram	Brainstorming	Benchmarking
Check sheet	Arrow diagram	Control plan	Departmental purpose
Control chart	Matrix diagram	Flow chart	analysis
Graphs	Matrix data analysis method	Force field analysis	Design of experiments
Histogram	Process decision	Questionnaire	Failure mode and effects analysis
Pareto diagram	Programme chart	Sampling	Fault tree analysis
Scatter diagram	Relations diagram		Poka yoke
	Systematic diagram		Problem solving methodology
			Quality costing
			Quality function deployment
			Quality improvement teams
			Statistical process control

Source: Tari and Sabater, 2004b

Figure 3: Overview of the methods.

- Group 1 may contain points belonging to the color of soil and flower pots
- Group 2 may contain points in a color belonging only to plants

Subsequently, the color in individual groups was checked algorithmically for all samples. It turned out that in Group 2, there were points whose color did not correspond to the points of plants. This finding was supplemented by the knowledge from the manual labeling of samples, where some points and shapes could not be labeled as parts of plants. The number of points not corresponding to plants in group 2 was determined as a measurable data quality criterion. This finding focused the research team's attention on preparing and implementing experiments when obtaining source data. In addition to creating a quantifiable indicator expressing the quality of the data, we focused on the entire research process.

Results and discussion

To apply RCA in our research, we created a process diagram. Our problem was the quality of the source data. After the analysis, we identified 4 areas that impact the source data. These were the technologies used, the settings of individual experiments, the execution of separate experiments, and the data processing methods. In these areas, any shortcomings could affect the quality of the data. We set up control mechanisms to verify the correct functioning in all areas. Some control mechanisms are determined by the type of area. For technical equipment, this meant performing a setting check

and calibration. To set up individual experiments, we conducted discussions with everyone involved in the settings. In the area of experiment execution, we conducted more detailed operator training. For processing methods, we prepared control outputs of the algorithms used to verify their functioning. However, we were unable to improve the quality of the data. After a detailed analysis of the experiment execution, it was found that the operators needed to adhere sufficiently to the established procedures. To rectify this situation, a requirement for photographic time documentation of each experiment was introduced. The documentation showed that the irrigation hoses (black color) were carelessly placed, causing their points to overlap with the source data (Figure 4).



Source: Author

Figure 4: Irrigation hoses.

After correcting errors in the execution of experiments and implementing photo documentation control, the source data quality increased. After corrective actions, we selected and reviewed a new data set of acquired data. It no longer showed anomalies in the 2nd group of points. The entire RCA tree diagram is shown in Figure 5.

Taking photo documentation and storing it for each experiment brought success by revealing the cause of the insufficient data quality. After discussing why the entire process of conducting experiments needed to be set up from the beginning, including photo documentation, it turned out that the reason was technical complexity. It was necessary to install a photo device for the outdoor scanner and connect it to the data storage. This measure seemed expensive and unnecessary.

Finding the error also resulted in another discovery:

unsatisfactory data labeling speed and labeling data quality. We discovered that trained operators still need help accurately labeling the designated class of objects. After an analysis, we changed the marking procedure so that the operators' duty was to mark only defined categories of plant organs (leaves, bay leaves, stems, and petiole) and to record the belonging of the parts to a specific plant in the form of recording identifiers in numerical terms. An algorithm was subsequently written to mark the whole plant, automatically generating this mark. Further streamlining of the labeling process was found in dropping the labeling of object detection points and replacing this process with an algorithm that produces this labeling automatically from labeled parts. These changes made it possible to include a larger group of operators in the labeling process, thereby speeding up the entire labeling process. The set procedure is shown in Figure 6.

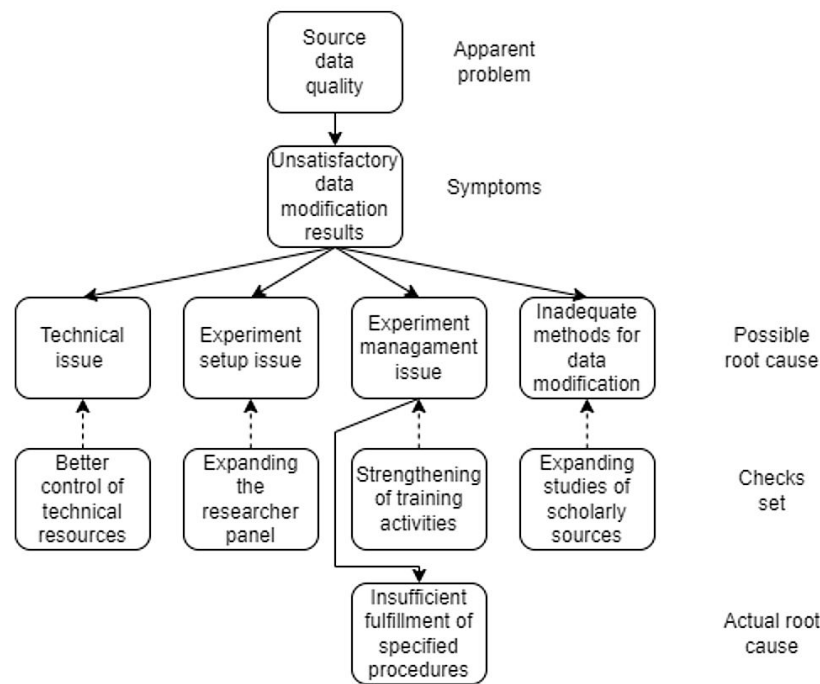


Figure 5: Root cause analysis tree diagram.

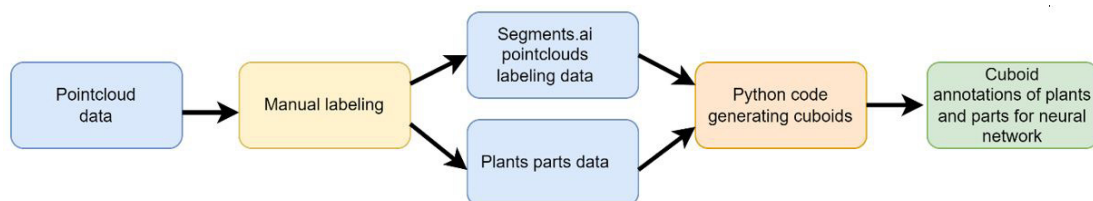
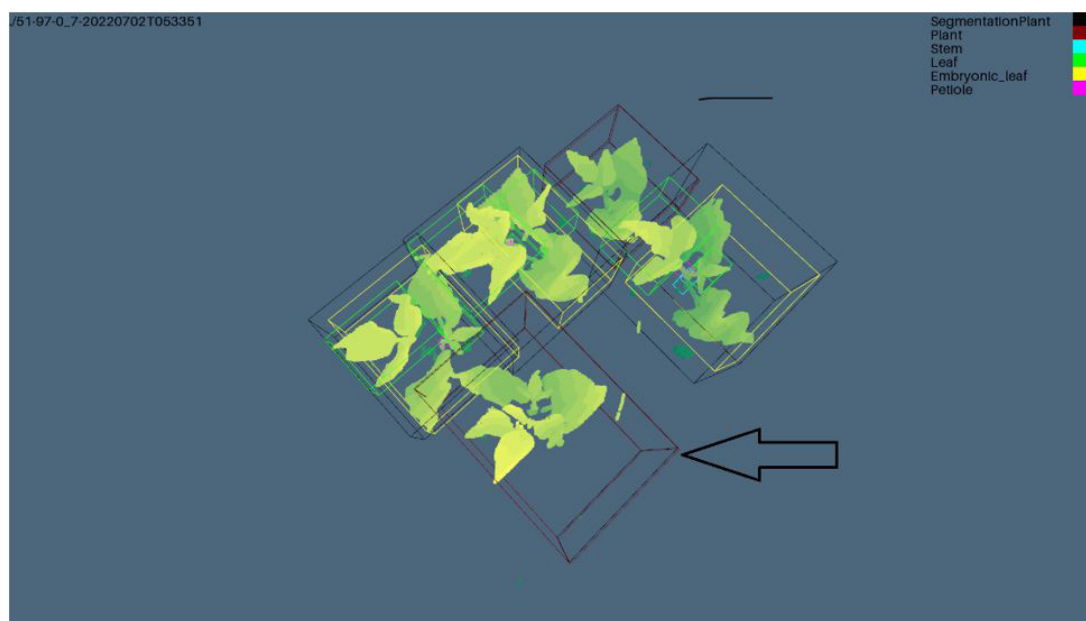


Figure 6: Semi-automated labeling process.



Source: Author

Figure 7: Visual control of labeling data.

Automatic cuboid generation also proved helpful in detecting incorrect labeling. The labeling process was supplemented with images of the generated cuboids, enabling more efficient control. Visual identification of suspicious samples was much more accessible. See Figure 7.

Conclusion

The mentioned research project is part of broader research in plant phenotyping using artificial intelligence. Our findings showed that using critical insight in the research process is very beneficial. The quality of research outputs directly depends on the data quality used. Our research takes place in an international team, geographically separated and in different time zones. The nature of the experiments performed does not allow their exact repetition. Cultivation of plants for experimental purposes in an outdoor environment cannot be repeated with identical conditions. All this makes it challenging to identify the sources of errors and inaccuracies. Creating a measurable parameter to express data quality was a guide to finding the source of errors. In the outdoor plant scanning environment, the acquired data is less rich than in the case of indoor scanning. Therefore, the highest possible quality of the source data was achieved important. The use of optimization methods for improving quality is mainly concentrated in the area of the process in the sphere of production, trade, and services.

Their use in a research environment is more challenging. The hierarchy of research teams is not strictly defined, and the results often depend on individual abilities, ideas, and creativity. The fact that we were not satisfied with the original quality of the data and implemented procedures to increase the quality of the entire process moved us to a higher level. At the same time, we increased the efficiency of part of the process (labeling), and this shortened the time frame and expanded the database for generating higher-quality research outputs. An important factor was also the fact that we introduced measures in source data quality in the following areas.

Acknowledgements

The results and knowledge included herein have been obtained owing to support from the following institutional grant. Internal grant agency of the Faculty of Economics and Management, Czech University of Life Sciences Prague, grant no. IGA 2023A0017.

This work was supported by the EC's Horizon Europe funding in the project CODECS grant agreement No. 101060179.

This work was supported by the EC's Digital Europe Programme in the project AGRITECH EU grant agreement No. 101123258.

Corresponding author:

Doc. Ing. Pavel Šimek, Ph.D.

Department of Information Technologies, Faculty of Economics and Management

Czech University of Life Sciences Prague,

Kamýčká 129, 165 00 Prague - Suchbátka, Czech Republic

E-mail: simek@pef.czu.cz

References

- [1] Andersen, B. and Fagerhaug, T. (2000) "Root cause analysis: simplified tools and techniques", pp. 155. [Online]. Available: https://books.google.com/books/about/Root_Cause_Analysis.html?hl=cs&id=i_EJAQAAMAAJ. [Accessed: Sept. 8, 2024]. ISBN 0873894669.
- [2] Andersen, B. and Fagerhaug, T. (2006) "Root Cause Analysis, Second Edition: Simplified Tools and Techniques", 240 p., Quality Oress. ISBN 9780873896924.
- [3] Bali, P., Kutikuppala, L. V. S., Avti, P. and Medhi, B. (2021) "Data Fraud and Essence of Data Verifiability, Quality Assurance Implementation in Research Labs", pp. 137-159. ISBN 978-981-16-3074-3. DOI 10.1007/978-981-16-3074-3_9.
- [4] Brünschwitz, S. and Kleymann-Hilmes, J. (2024) "Benefits and approaches of a quality management system in biomedical research laboratories", *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz*, Vol. 67, No. 1, pp. 99-106. E.ISSN ISSN 1437-1588. DOI 10.1007/S00103-023-03797-Y.
- [5] Daiki, S. and Noshita, K. (2024) "A comparative study of plant phenotyping workflows based on three-dimensional reconstruction from multi-view images", *BioRxiv*. DOI 10.1101/2024.03.21.586185.
- [6] Detring, J., Barreto, A., Mahlein, A.-K. and Paulus, S. (2024) "Quality Assurance of Hyperspectral Imaging Systems for Neural Network supported Plant Phenotyping". DOI 10.21203/RS.3.RS-4648326/V1.
- [7] Dhondt, S., Wuyts, N. and Inzé, D. (2013) "Cell to whole-plant phenotyping: the best is yet to come", *Trends in Plant Science*, Vol. 18, No. 8, pp. 428-439. ISSN 1360-1385. DOI 10.1016/J.TPLANTS.2013.04.008.
- [8] Fiorani, F. and Schurr, U. (2013) "Future Scenarios for Plant Phenotyping", *Annual Review of Plant Biology*, Vol. 64, pp. 267-291. ISSN 1543-5008. DOI 10.1146/annurev-arplant-050312-120137.
- [9] Isere, E. E. and Omorogbe, N. E. (2024) "Quality Management in Clinical and Public Health Research: A Panacea for Minimising and Eliminating Protocol Deviations in Research Operations", *Nigerian Medical Journal*, Vol. 65, No. 3, pp. 367-375. ISSN 0300-1652. DOI 10.60787/NMJ-V65I3-421.
- [10] ISO 9001:2015 - Quality management systems - Requirements. (n.d.) [Online]. Available: <https://webstore.ansi.org/standards/iso/iso90012015> [Nov. 15, 2023].
- [11] Kartal, S., Choudhary, S., Masner, J., Kholova, J., Stoces, M., Gattu, P., Schwartz, S. and Kissel, E. (2021) "Machine Learning-Based Plant Detection Algorithms to Automate Counting Tasks Using 3D Canopy Scans", *Sensors*, Vol. 21, No. 23, pp. 8022. ISSN 1424-8220. DOI 10.3390/S21238022.
- [12] Multi-sensor data labeling platform for robotics and AV | Segments.ai. (n.d.). [Online]. Available: <https://segments.ai/>. [April 28, 2024].
- [13] Parker, J. (2024) "The application of operation research in the quality management engineering", *Advances in Operation Research and Production Management*, Vol. 1, No. 1, pp. 25-32. ISSN 1687-9155. DOI 10.54254/3006-1210/DIRECT/2532.
- [14] Paulus, S. (2019) "Measuring crops in 3D: Using geometry for plant phenotyping", *Plant Methods*, Vol. 15, No. 1, pp. 1-13. ISSN 1746-4811. DOI 10.1186/S13007-019-0490-0/FIGURES/5.

- [15] Percarpio, K. B., Watts, B. V. and Weeks, W. B. (2008) "The effectiveness of root cause analysis: what does the literature tell us?", *Joint Commission Journal on Quality and Patient Safety*, Vol. 34, No. 7, pp. 391-398. ISSN 1553-7250 . DOI 10.1016/S1553-7250(08)34049-5.
- [16] Phenospex - Smart plant analysis and Phenotyping systems. (n.d.). [Online]. Available: <https://phenospex.com/> [Nov. 15, 2023].
- [17] Pongpiyapaiboon, S., Tanaka, H., Hashiguchi, M., Hashiguchi, T., Hayashi, A., Tanabata, T., Isobe, S. and Akashi, R. (2023) "Development of a digital phenotyping system using 3D model reconstruction for zoysiagrass", *The Plant Phenome Journal*, Vol. 6, No. 1, ISSN 2578-2703. DOI 10.1002/PPJ2.20076.
- [18] Ronanki, S., Pavlík, J., Masner, J., Jarolímek, J., Stočes, M., Subhash, D., Talwar, H. S., Tonapi, V. A., Srikanth, M., Baddam, R. and Kholová, J. (2022). An APSIM-powered framework for post-rainy sorghum-system design in India, *Field Crops Research*, Vol. 277, p. 108422. ISSN 0378-4290. DOI 10.1016/J.FCR.2021.108422.
- [19] Serrat, O. (2017) "The Five Whys Technique", In: *Knowledge Solutions*, pp. 307-310. Springer, Singapore. ISBN 978-981-10-0982-2. DOI 10.1007/978-981-10-0983-9_32.
- [20] Shook, J. (2008) "Managing to Learn: Using the A3 Management Process to Solve Problems, *Gain Agreement, Mentor and Lead*", pp. 10-11. ISBN 1934109207.
- [21] Šimek, P., Stočes, M., Vaněk, J., and Masner, J. (2015) "Mobile accessibility of information sources in the areas of agriculture, forestry, water management, food industry and rural development", *Agrarian Perspectives XXIV: Global Agribusiness and the Rural Economy*, pp. 440-446. ISBN 978-80-213-2581-4.
- [22] Sozzani, R., Busch, W., Spalding, E. P. and Benfey, P. N. (2014) "Advanced imaging techniques for the study of plant growth and development", *Trends in Plant Science*, Vol. 19, No. 5, pp. 304-310. ISSN 1878-4372. DOI 10.1016/J.TPLANTS.2013.12.003.
- [23] Starzyńska, B. and Hamrol, A. (2013) "Excellence toolbox: Decision support system for quality tools and techniques selection and application", *Total Quality Management and Business Excellence*, Vol. 24, No. 5-6, pp. 577-595. ISSN 0954-4127. DOI 10.1080/14783363.2012.669557.
- [24] Tarí, J. J. and Sabater, V. (2004a) "Quality tools and techniques: Are they necessary for quality management?", *International Journal of Production Economics*, Vol. 92, No. 3, pp. 267-280. ISSN 0925-5273. DOI 10.1016/J.IJPE.2003.10.018.
- [25] Ugochukwu, A. I. and Phillips, P. W. B. (2022) "Data sharing in plant phenotyping research: Perceptions, practices, enablers, barriers and implications for science policy on data management", *Plant Phenome Journal*, Vol. 5, No. 1. ISSN 2578-2703. DOI 10.1002/PPJ2.20056.
- [26] Vadez, V., Kholová, J., Hummel, G., Zhokhavets, U., Gupta, S. K. and Hash, C. T. (2015) "LeasyScan: a novel concept combining 3D imaging and lysimetry for high-throughput phenotyping of traits controlling plant water budget", *Journal of Experimental Botany*, Vol. 66, No.18, pp. 5581. ISSN 0022-0957. DOI 10.1093/JXB/ERV251.
- [27] Venkatasubramanian, V., Rengaswamy, R., Yin, K. and Kavuri, S. N. (2003) "A review of process fault detection and diagnosis: Part I: Quantitative model-based methods", *Computers and Chemical Engineering*, Vol. 27, No. 3, pp. 293-311. ISSN 0098-1354. DOI 10.1016/S0098-1354(02)00160-6.
- [28] Vianna, E. L. F., De Figueiredo, V. V., Da Silva, C. M. F., Bertolino, L. C. and Spinelli, L. (2022) "Impact of implementing quality control systems in laboratories associated with teaching and research institutions – The case study of the laboratory for macromolecules and colloids in the petroleum industry. *International Journal of Metrology and Quality Engineering*, Vol. 13, pp. 4. ISSN 2107-6847. DOI 10.1051/IJMQE/2022004.
- [29] Wilson, P. F. ., Dell, L. D. and Anderson, G. F. (1993) "*Root cause analysis: a tool for total quality management*", pp. 216. ASQC Quality Press, 1993. ISBN 9780873891639.
- [30] Yuniarto, H. A. (2012) "The Shortcomings of Existing Root Cause Analysis Tools", *Proceedings of the World Congress on Engineering 2012 Vol III*, WCE 2012, July 4 - 6, 2012, London, U.K. ISSN 2078-0958.