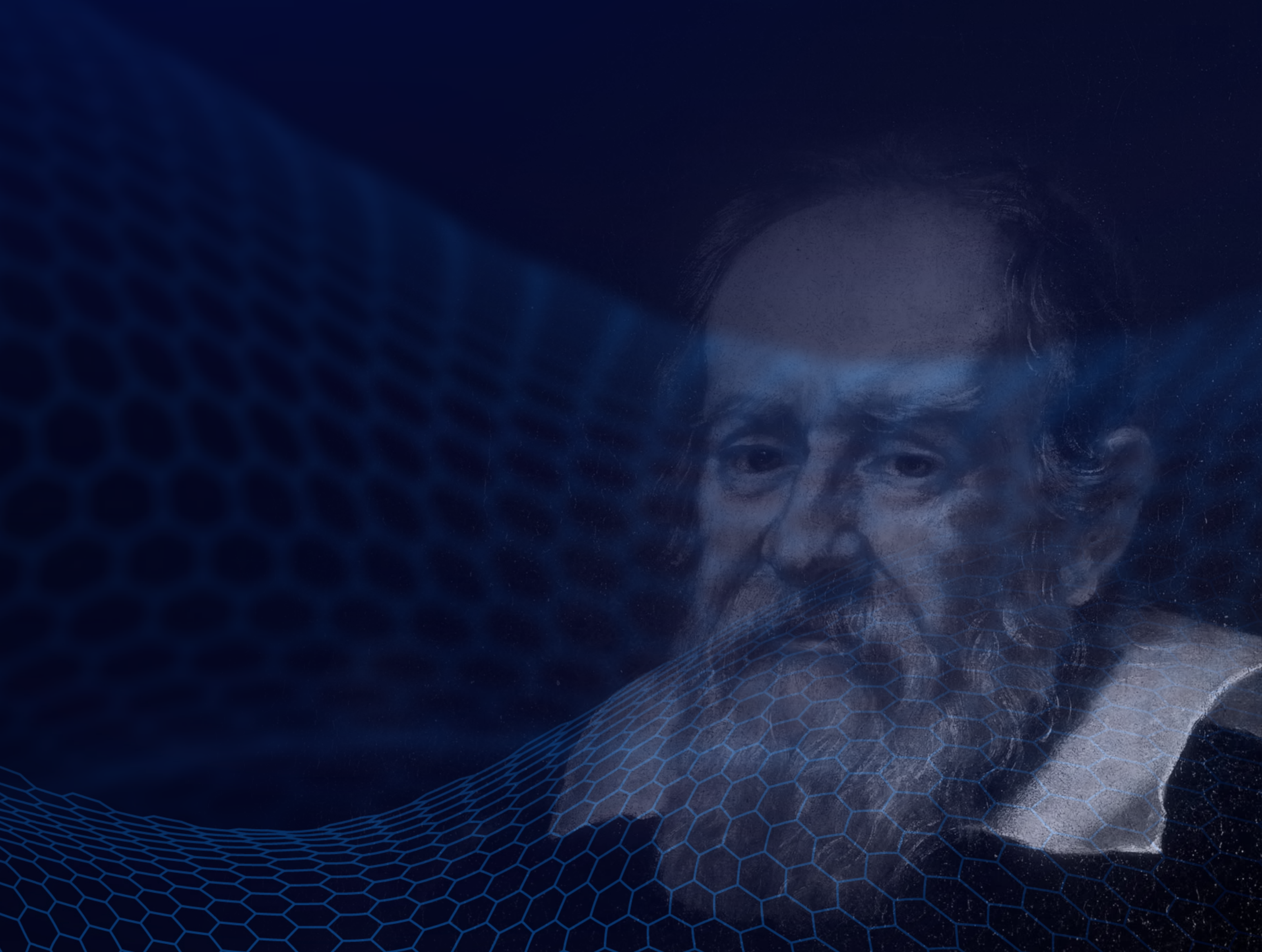
Journal of Engineering and Technology for Industrial Applications

ITEGAM JETIA

ISSN 2447-0228

June 2021

Volume 07 / No 29



Editor-in-Chief: J. C. Leite

www.itegam-jetia.org

ITEGAM - JETIA ISSN 2447-0228 Online
Journal of Engineering and Technology for Industrial Applications (JETIA)

O **ITEGAM-JETIA: Journal of Engineering and Technology for Industrial Applications** is a publication of the Galileo Institute of Technology and Education of the Amazon (ITEGAM), located in the city of Manaus since 2008. JETIA publishes original scientific articles covering all aspects of engineering. Our goal is the dissemination of research original, useful and relevant presenting new knowledge on theoretical or practical aspects of methodologies and methods used in engineering or leading to improvements in professional practice. All the conclusions presented in the articles It should be state-of-the-art and supported by current rigorous analysis and balanced assessment. Public magazine scientific and technological research articles, review articles and case studies.

JETIA will address topics from the following areas of knowledge: Mechanical Engineering, Civil Engineering, Materials and Mineralogy, Geosciences, Environment, Information and Decision Systems, Processes and Energy, Electrical and Automation, Mechatronics, Biotechnology and other Engineering related areas.

Publication Information:
ITEGAM-JETIA (ISSN 2447-0228), (online) is published by Galileo Institute of Technology and Education of the Amazon on a every two months (February, April, June, August, October and December).

Contact information:

Web page: www.itegam-jetia.org
Email: editor@itegam-jetia.org
Galileo Institute of Technology and Education of the Amazon (ITEGAM).

Joaquim Nabuco Avenue, No. 1950. Center. Manaus, Amazonas. Brazil.
Zip Code: 69020-031. Phone: (92) 3584-6145.

ITEGAM - JETIA ISSN 2447-0228 Online
Journal of Engineering and Technology for Industrial Applications (JETIA)
INSTITUTO DE TECNOLOGIA ITEGAM

ITEGAM-JETIA. v.7, n.29. June of 2021. Manaus - Amazonas, Brazil. ISSN 2447-0228 (ONLINE)
https://www.itegam-jetia.org

SUMMARY

Journal of Engineering and Technology for Industrial Applications

ISSN ONLINE: 2447-0228

ITEGAM-JETIA

Manaus, v.7 n.29, p. 04-12. May/Jun, 2021

DOI: https://doi.org/10.5935/jetia.v7i29.754

RESEARCH ARTICLE OPEN ACCESS

THE INFLUENCE OF PERFORMANCE EXPECTANCY ON E-PROCUREMENT ADOPTION MODEL IN DEVELOPING COUNTRIES: TANZANIANS PERCEPTION

Deus Nichodemus Shatta*[1] and France Shayo[2]

[1] Department of Business and Entrepreneurship Studies, National Institute of Transport, Dar Es Salaam Tanzania.
[2] Department of Marketing and Entrepreneurship, The Open University of Tanzania, Dar Es Salaam Tanzania.

[1] http://orcid.org/0000-0003-1852-6070, [2] http://orcid.org/0000-0002-3585-292X

Email: *deusshatta@gmail.com, france.shayo@out.ac.tz

ARTICLE INFO

Article History
Received: May 15th, 2021
Accepted: June 21th, 2021
Published: June 30th, 2021

Keywords:
Performance Expectancy,
E-Procurement,
Adoption.

ABSTRACT

This study aimed at filling the knowledge gap by examining the influence of performance expectancy on e-procurement adoption model in developing countries, Tanzania in particular. The mediation effects of relative advantage and attitude were considered in this study. Positivism philosophy and cross-sectional survey research design were adopted. In addition, stratified sampling technique was adopted and a sample of 157 respondents was used. Questionnaire and documentary review were used for data collection. Partial Least Squares Structural Equation Modeling with the help of SmartPLS 3 software was used to analyze the collected data. The findings reveal that performance expectancy has direct influence on e-procurement adoption model. Likewise, performance expectancy has indirect influence on e-procurement adoption model through relative advantage and attitude. Basing on these findings, it is concluded that performance expectancy has both direct and indirect influences on e-procurement adoption model. Therefore, it is recommended that the Government of Tanzania and other Governments based in developing countries to pay attention to the performance expectancy on the way to e-procurement adoption systems.

I. INTRODUCTION

I.1 OVERVIEW AND PROBLEM SETTING

Behavioural intention to use technology is said to be influenced by performance expectancy because it provides benefits to consumers in performing certain activities [1]. However, many studies concerning e-procurement adoption systems have not taken into consideration the performance expectancy as a critical success factor which influences new technology adoption [2-7]. The study by [4] aimed at establishing the factors that influence e-procurement adoption technology in the organization. The findings of this study revealed that e-procurement adoption is influenced by three critical success factors namely internal needs, improved customers relationship and reduction of labour costs. In this study by [4] was recommended that allocation of financial resources, top management support and end-user buy-in can influence e-procurement adoption in any organization. The study by [5] used a cross-sectional research design to establish the factors that influence e-procurement adoption system. A questionnaire was used to collect data from 46 respondents. The collected data were analyzed by using descriptive statistics with the help of IBM SPSS statistics 21. The findings of this study revealed the following factors which influence e-procurement adoption; employees and management commitment to success of adoption, reliability of information technology and supplier's performance and monitoring the performance of e-procurement system. Based on these findings, this study recommended to the organization to address the factors which were identified towards success of e-procurement adoption. [6] conducted a study to examine the role of e-procurement strategy on the performance of an organization. A cross-sectional research design and descriptive statistics analysis with the help of IBM SPSS statistics 21 were adopted. In addition, inferential statistics with the help of multiple regression and correlation analysis was used to analyze the collected data. The customer service level. The findings revealed that the performance of the organization is affected significantly by the service level strategy, procurement cost reduction strategy, inventory optimization strategy and buyer/supplier collaboration strategy, audit and compliance strategy. The study by [6] recommended that the organization needs a good e-procurement system to use in order to reduce the effort and time required for procurement of goods, works and services. The new procurement system was anticipated to eliminate the traditional paper chain from ordering to payment during procurement process. The study by [8] investigated the factors influencing e-procurement adoption. Questionnaire survey was used to collect data from 213 organizations. Descriptive statistics was used to analyze data. The findings revealed that the eliminating geographical barriers, benefits of e-procurement, effective communication among buyer and suppler were most important factors influencing e-procurement adoption. In connection to that, [9] carried out the study whose aim was to understand advantages of e-procurement. A questionnaire was used to collect data from 155 public officials already were using e-procurement and data were processed using IBM SPSS statistics 21 software. The findings revealed that e-procurement adoption results into reduction of corruption, improves monitoring and accountability [9].

Furthermore, [10] aimed to understand the concept of e-procurement, its evolution and adoption in the market economy as well as higher education. The findings revealed that successful implementation of e-procurement would require the following critical success factors: top management support; user acceptance of e-procurement systems; employees and management commitment to success of adoption; reliability of information technology and supplier performance; monitoring the performance of e-procurement systems; Other critical success factors identified were: training of staff in procurement practices; risk perception and continuous measurement of the key benefits, best practices and actual selection of the system [10].

Studies by [8, 9] and [10] would have been more interesting if the authors had explained the indirect influences of performance expectancy towards e-procurement adoption. The study by [11] aimed at investigating the level of e-Perolehan system success and the factors that contribute to this success, as perceived by the Government users. The structural equation modeling was used to analyze the data. System compatibility, users' attitude and mimetic pressures were found to be the main significant critical success factors of the e-Perolehan system. The study by [12] examined the factors influencing e-procurement adoption in an organization. Simple random sampling technique was used to sample 97 participants. A questionnaire was used for data collection and the collected data were analyzed by using descriptive statistics. The findings revealed that training of both end users and procurement staff was an important factor in e-procurement adoption. It was recommended that management of financial institutions need to train both end users and procurement staff in order to keep the dynamism of e-procurement [12]. However, some few studies regarding information technology adoption models hypothesized performance expectancy to effect new technology adoption [13], [7], [14]. For example, [7] examined factors influencing adoption of mobile payment services. The study by [7] used survey method to collect data and the collected data were analyzed by using Partial Least Squares Structural Equation Modeling (PLS-SEM). The findings revealed that the consumer' behavioural intention to use mobile payment services is influenced significantly by perceived trust, social influence and effort expectancy. However, performance expectancy and facilitating condition were found to influence insignificantly the consumer' behavioural intention to use mobile payment services [7]. Likewise, the study by [13] revealed that the performance expectancy had insignificant influence towards green e-business adoption. Although the studies by [7] and [13] considered the performance expectancy as an important factor in new technology adoption but both of them did not base on e-procurement adoption. This study was carried out to fill this gap by examining the influence of performance expectancy on e-procurement adoption model in developing countries, Tanzania in particular. In addition to that the Government of Tanzania has been undertaking reforms in the public procurement process with the aim of supporting e-procurement system which is more efficient and transparent in terms of its performance [15]. However, it was noted that regardless of the benefits and the efficient of procurement performance associated with adoption of e-procurement system, 326 (30.9 per cent) of the trained suppliers for piloting Tanzania National electronic Procurement System (TANePS) adoption in the country were reluctant to register in the system [16, 17]. This situation may disturb the expected benefits and efficient procurement performance of the system.

I.2 OBJECTIVE OF THE STUDY

The objective of this study was to examine the influence of performance expectancy on e-procurement adoption through relative advantage and attitude as mediator constructs.

I.3 THEORICAL FRAMEWORK

In reviewing the theoretical literature, it was noted that at this time, researchers combine UTAUT or TOE model with external theories or models or other variables in order to address weaknesses of the theory or model [13], [7], [18]. It has been argued that it is essential to use at least two theories to attain a better understanding of multidimensional new information technology adoption due to the limitations that one theory or model may have [18]. Currently, there exists no collectively accepted theory to explain new technologies adoption and this situation has left researchers in a "theoretical confusion" because scholars have been developing their own theories or extending the existing ones to cater for their research problems [13], [7], [19-22]. For example, the study by [13] developed a model by using all constructs from the original UTAUT, however, one construct was added to furnish the research problem. Likewise, the study by [13] developed a model by using constructs from UTAUT, TOE and Institutional Theory to outfit for research problem.

I.3.1 Theoretical Gap

The existing theory adopted in the current study (UTAUT) and the model employed (TOE) are clarifying inadequately the combined perspectives (individual' and organizational' perspectives). The current study has developed an integrative model comprehensively explains the combined perspectives by using direct and indirect relationships of the constructs from TOE model and UTAUT of which the original UTAUT and TOE are not explaining clearly. The constructs relative advantage and attitude from TOE model and the construct performance expectancy from UTAUT were integrated and thought to influence directly and indirectly the Tanzania National e-Procurement System (TANePS) adoption. The direct and indirect relationships of these constructs are not comprehended clearly in the original and modified TOE models and UTAUTs. Prior studies with regard to new technology adoption have not applied the complete UTAUT model as found in origin and most of these studies employed only a subset of the model and that moderators were typically dropped or were replaced by other moderators [21], [1], [13], [7], [22]. Figure 1 and 2 presents modified UTAUTs of the new technology acceptance and use with an indication for the role of attitude and new moderator which was not suggested by the origin UTAUT.

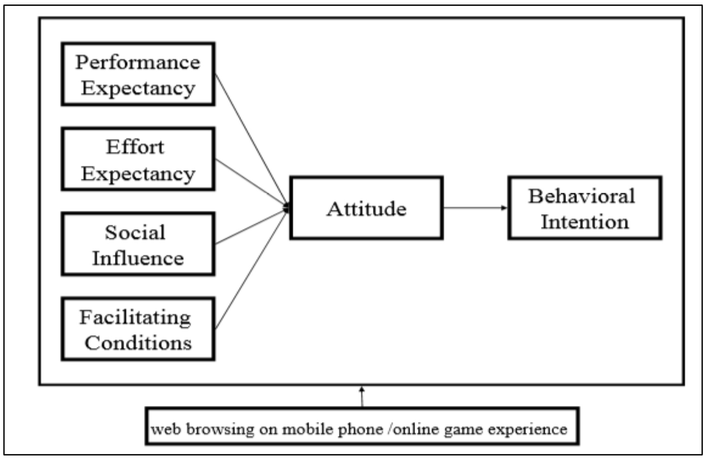


Figure 1: Modified UTAUT.
Source: [21].

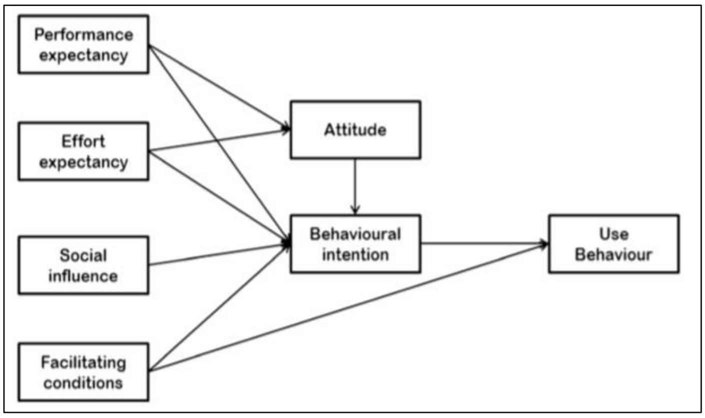


Figure 2: Modified UTAUT.
Source: [22].

Previous studies have also not applied all constructs from TOE model as found in [23]. Most prior studies had been employing only some of the constructs from each element of TOE model towards new technologies adoption [11], [24-27]. For example, the study by [11] adopted the TOE model to examine the level of e-Perolehan system achievement and the influences that subsidize to this achievement, as professed by the Government employers but the study included only attitude of users, learning of organization, expectation of suppliers, compatibility of the system and imitative pressure. The study by [25] also adopted the TOE model to explore the influence of these factors on the innovation decision of firms in the context of a developing country. Figure 3 and 4 presents modified models by [25] and by [26] which ignored some constructs from the original TOE model.

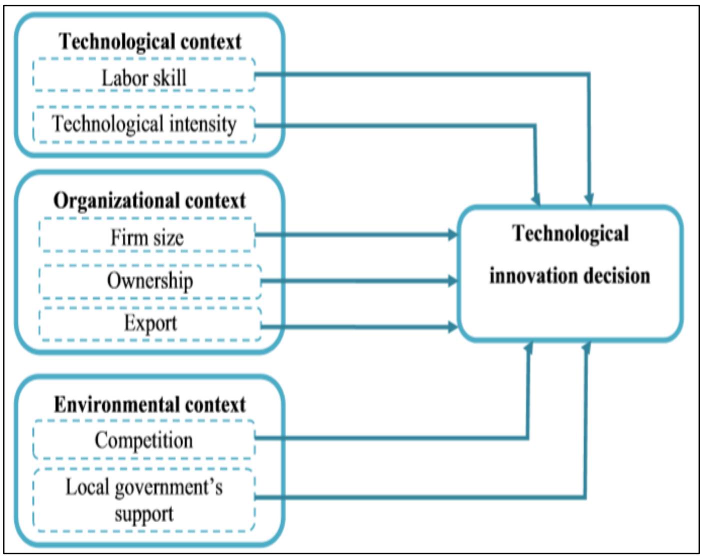


Figure 3: Modified TOE Model.
Source: [25].

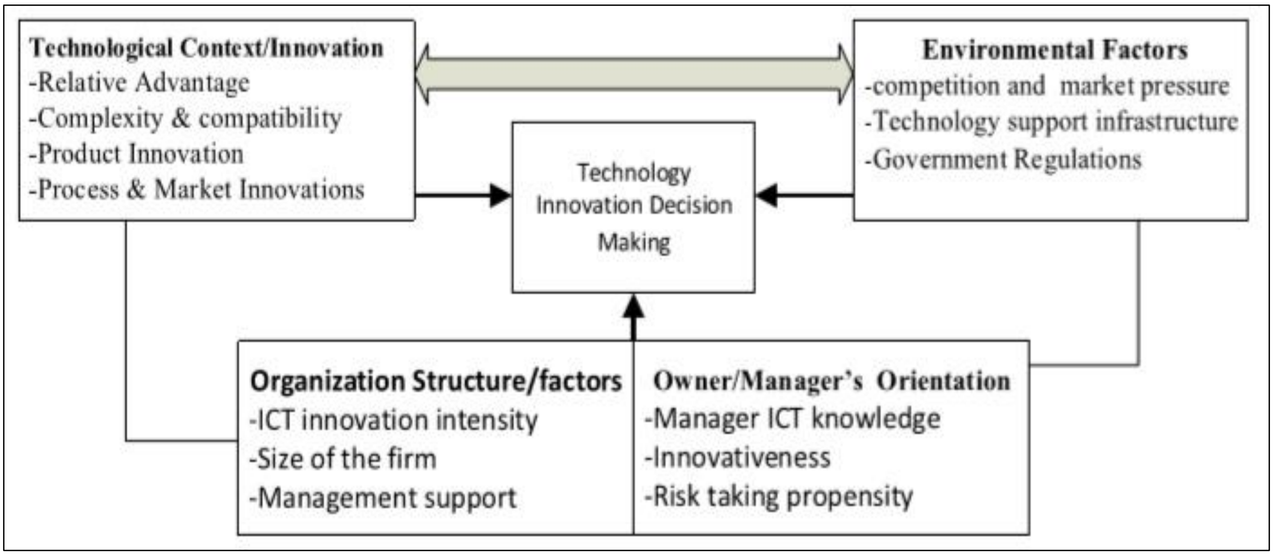


Figure 4: Modified TOE Model.
Source: [26].

A possible reason why prior studies may not have utilized all constructs from the origin UTAUT and TOE model is because some of the constructs might not be useful in their research problems towards new technologies adoption and use context. Basing on this fact, the current study has adopted TOE model and UTAUT in order to suit its research problem and objectives with regard to TANePS adoption [13], [18], [7], [28], [14], [18], [29].

I.4 MODEL CONSTRUCTS AND HYPOTHESES

This study explains three endogenous constructs and one exogenous construct. Endogenous constructs involve relative advantage, attitude and TANePS adoption. The exogenous construct is performance expectancy.

I.4.1 Performance Expectancy (PE)

[1] defined performance expectancy as a degree to which using technology will provide direct and indirect benefits to consumers in performing certain activities, and is theorized to influence behavioural intention to use technology. On the other hand, the origin UTAUT theory by [19] suggests that performance expectancy is thought to influence the new technology adoption. This argument of theory was supported by [24] who asserted that the easier it is to use a technology, the greater the expected benefits from the technology with regard to performance enhancement. The underlying principle under performance expectancy is that individuals make decisions rationally and systematically on the basis of the system available.

In this study, performance expectancy was postulated to positively, directly and indirectly effect TANePS adoption basing on the fact from the theory and the underlying principle.

H_1: Performance Expectancy has positive and direct effect on TANePS adoption.

H_2: Performance Expectancy has positive and indirect effect on TANePS adoption through attitude.

H_3: Performance Expectancy has positive and indirect effect on TANePS adoption through relative advantage.

I.4.2 Relative Advantage (RA)

Relative advantage is the extent to which the innovation is viewed by users to be better than the existing idea; i.e., perceived cost and benefits [30]. However, perceived benefits include 'direct and indirect benefits' like reduction in transaction errors and transaction costs, improved data accuracy and faster tendering process [31]. In this study, relative advantage was postulated to positively effect TANePS adoption.

H_4: Relative advantage has positive and direct effect on TANePS adoption.

I.4.3 Attitude (AT)

User's attitude in this context of study means having positive or negative mindset of buyers or suppliers towards e-procurement adoption in public sector [11] which includes like or dislike of the system; willing to learn or not willing to learn about the application of the system; committed or not committed to use the system in public procurement process; frequently interact or not frequently interact with the system in public procurement process. In this study attitude was postulated to positively effect TANePS adoption in public sector.

H_5: Attitude has positive and direct effect on TANePS adoption.

I.4.4 Conceptual Model of the Study

For the purpose of validation of these assertions a number of direct and indirect effects of the constructs on TANePS adoption were conceptualized as depicted in Figure 5.

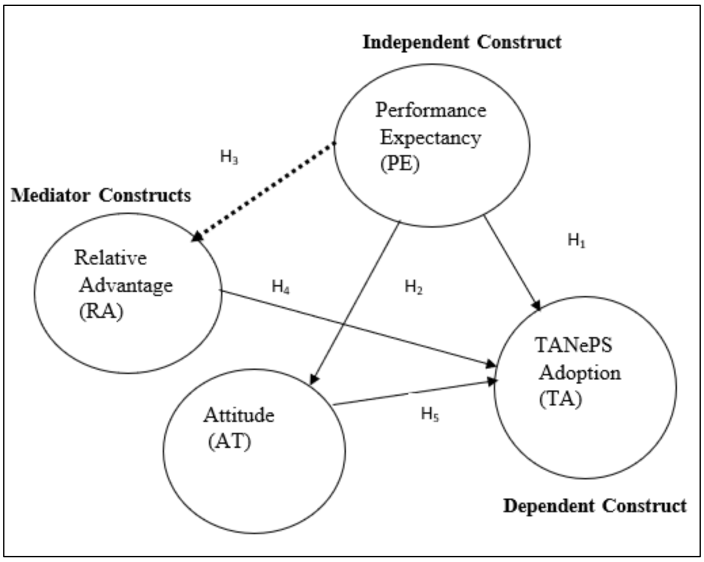


Figure 5: Conceptual Model of the Study.
Source: Authors, (2021).

Key:

——▶ Relationships Exist in Literature.

........▶ Theoretical Gap (Does not Exist in Literature).

II. METHODS

Positivism philosophy and cross-sectional survey research design were adopted. In addition, stratified sampling technique was adopted and a sample of 157 respondents was used. Questionnaire and documentary review were used for data collection. Based on research model of this study, the rule of thumb proposed by [32, 33] for applying PLS-SEM and SmartPLS 3 software in data analysis was used to justify the minimum number of sample size required. Forty (40) respondents were regarded to be the minimum number of sample size for testing the hypotheses of research model. This is due to the fact that the exogenous construct (performance expectancy) had a maximum of four indicators. Per rule of thumb proposed by [33], four indicators, times ten was equivalent to forty (40). Therefore, a sample of 157 respondents was sufficient for data analysis in this study.

The respondents' general information was operationally measured by using non-parametric measurements for sex and education level. The responses were based on nominal scale attitude questions. However, respondents' age and experience were operationally measured by using parametric measurements of which complete age and experience were filled in the questionnaire. The dependent variable, mediator variables and the independent variable were operationally measured by using non parametric measurements. TANePS adoption had four indicators (service performance, efficiency, transparency and information quality), performance expectancy had four indicators (simplification of procurement process, simplification of payment process, removal of chances of corruption, reduction of procurement cycle time), relative advantage had four indicators (reduction in transaction errors, transaction costs, improved data accuracy, faster tendering process) and attitude had four indicators (acceptance of the system, intention to learn, plan to use, intention to take advantage). The responses were based on a five Likert scale (ordinal) questions.

Questionnaires with closed ended questions were assigned numbers to enable the process of quantitative data analysis to be more accurate and simpler. The collected quantitative data were analyzed by using descriptive statistics with the help of IBM SPSS

Statistics software 21 Version and by using Partial Least Squares Structural Equation Modeling (PLS-SEM) with the help of SmartPLS 3 software. Based on the complexity of the research model and the sample size of this study, PLS-SEM was the best non-parametric multivariate method that could allow data analysis by incorporating both (direct and indirect) relationships between the constructs and between the indicators and unobservable variables of the research model.

III. FINDINGS

III.1 EDUCATION LEVEL OF RESPONDENTS

The findings concerning education level against type of respondents revealed that majority of the respondents from procuring entities had higher levels of education than suppliers. Procurement experts from procuring entities who responded in this study were as follows: eight per cent were holding diploma in procurement, 50 per cent were holding bachelor degree in procurement and 42 per cent were holding post graduate degrees related to procurement. For suppliers who responded in this study the finding revealed that 66.7 per cent were holding diplomas in different fields, certificates in various fields, certificates of secondary and primary education, 26.3 per cent were holding bachelor degree in different fields and 7.0 per cent of the respondents were holding post graduate degrees in different fields. These findings implied that the information and data provided by the respondents for this study were actual and comprehensive. Table 1 shows the education level of respondents against type of respondents.

Table 1: Education Level of the Respondents.

		Type of Respondent		Total
		Procurement Expert	**Suppliers**	
Education Level of Respondents	Primary Educ.	0	4	4
	Secondary Educ.	0	15	15
	Certificate Level	0	5	5
	Diploma Level	8	14	22
	Bachelor Degree	50	15	65
	Master's Degree	41	4	45
	PhD Degree	1	0	1
Total		**100**	**57**	**157**

Source: Authors, (2021).

III.2 MEAN AND STANDARD DEVIATION OF THE RESPONDENT'S EXPERIENCE

Apart from education of the respondents, the study also revealed that about 10 years was a mean of the respondents with regard to their experience which had good implications for the data and information collected for this study. However, there was significant dispersion of experience in years from one respondent to another due to the fact that the standard deviation from the mean was about 8 years close to its mean. Figure 6 shows the mean and standard deviation of the respondents' experience.

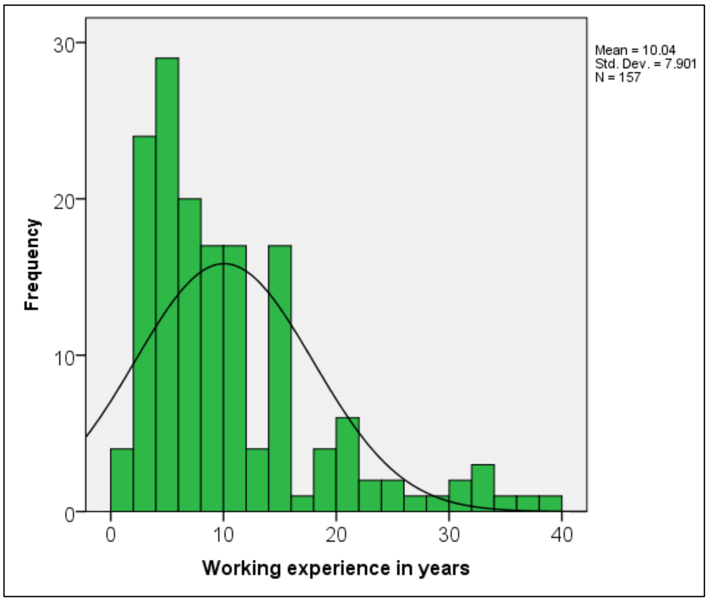


Figure 6: The Mean and Standard Deviation of the Respondents' Experience.
Source: Authors, (2021).

III.3 MEAN AND STANDARD DEVIATION OF THE RESPONDENTS' AGE

The study also revealed that about 40 years was the mean age of the respondents. In addition, there was insignificant dispersion of years from one respondent to another due to the fact that the standard deviation from the mean age was only about 10 years. This finding implied that the ages of the respondents were not scattered and therefore, the information and data provided for this study was very comprehensive and genuine. Figure 7 shows the mean and standard deviation of the respondents' ages of the study.

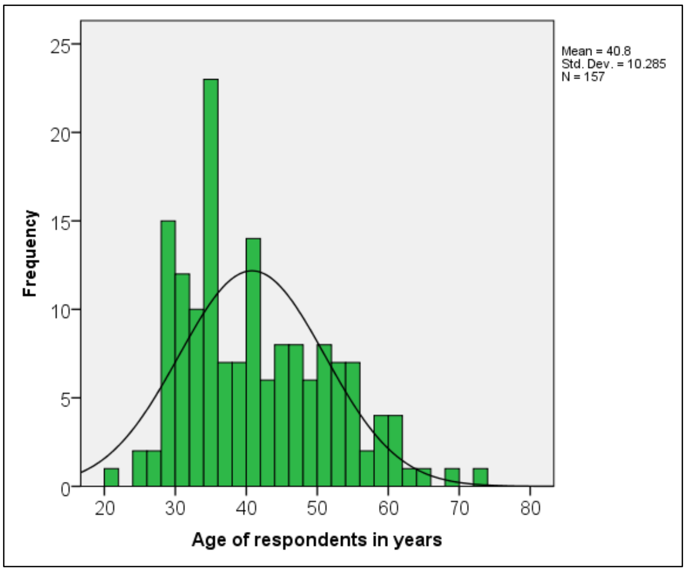


Figure 7: Mean and Standard Deviation of the Respondents' Ages.
Source: Authors, (2021).

III.4 RELEVANCE OF PATH COEFFICIENTS AND STATISTICAL SIGNIFICANCE OF THE HYPOTHESIZED RELATIONSHIPS

After performing PLS algorithm, the report from SmartPLS 3 for relevance of the path coefficients of the research model showed that the path coefficients of the hypothesized relationships of performance expectancy and TANePS adoption were positive which meant that an increase in one standard deviation of the performance expectancy results into increase of the rate of TANePS adoption. In addition, the direct and indirect relationships of the performance expectancy were supported. Figure 8 and 9 show the relevance of path coefficients and statistical significance of the hypothesized relationships.

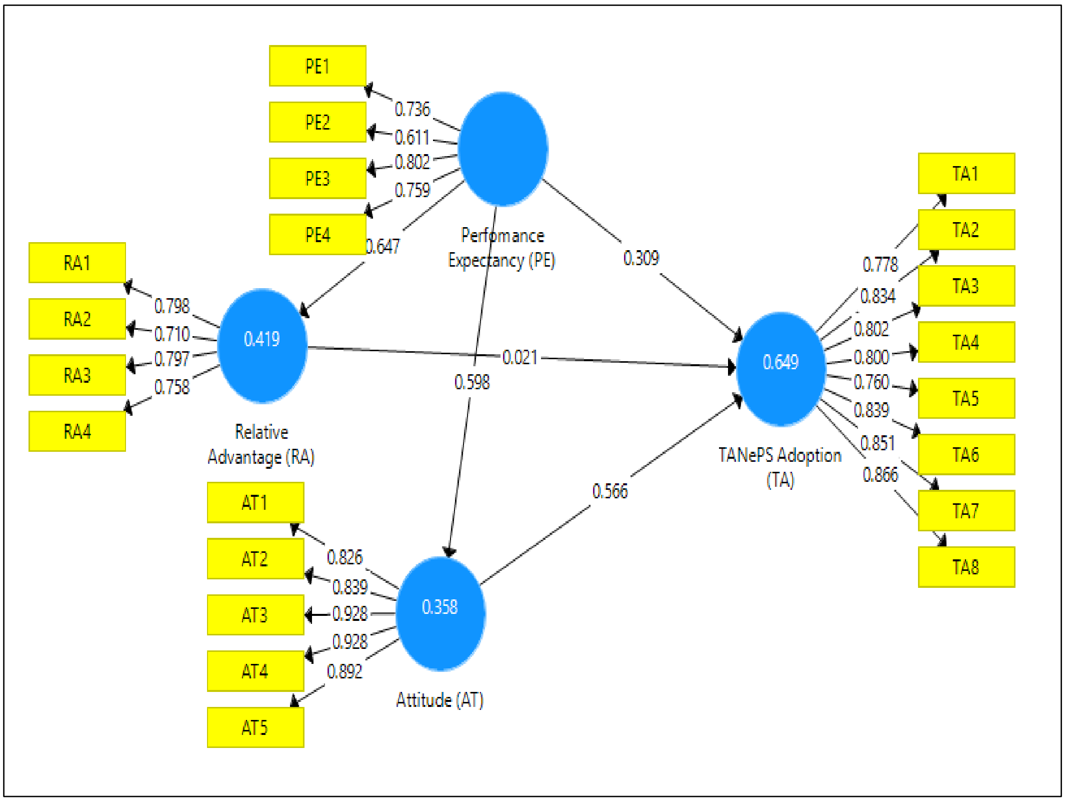


Figure 8: Relevance of the Path Coefficients.
Source: Authors, (2021).

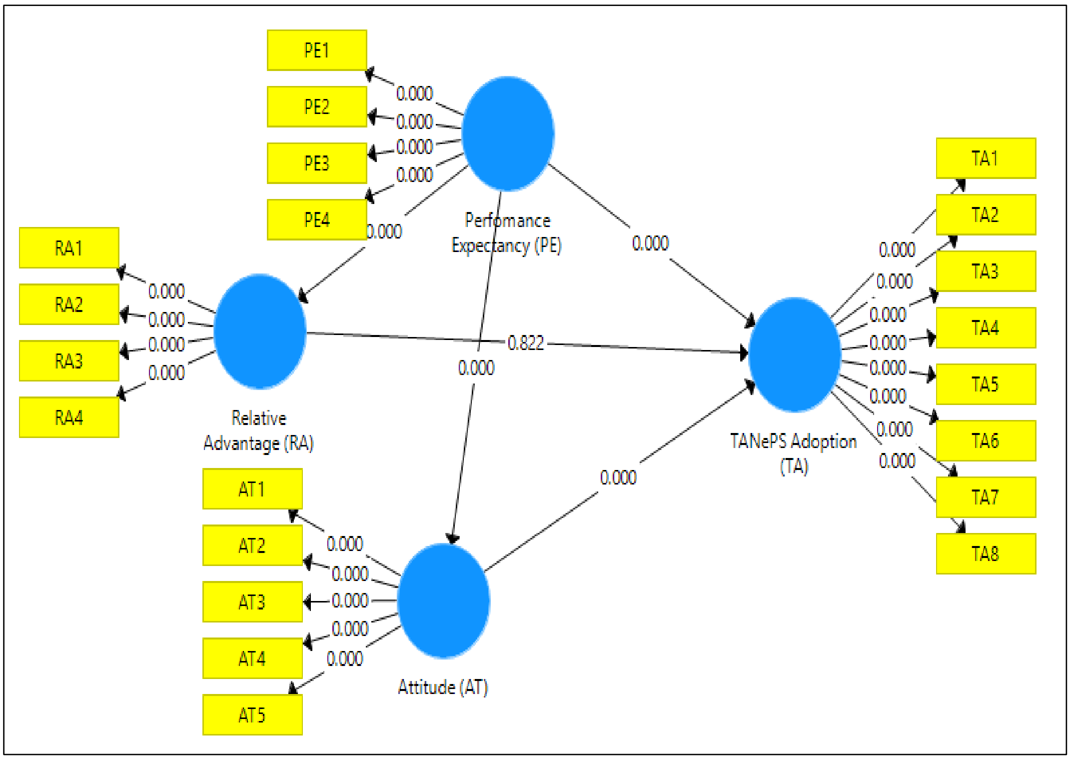


Figure 9: Statistical Significance of the Hypotheses.
Source: Authors, (2021).

This study adopted mediators (relative advantage and attitude) that accounted for part of the relationships between predictor (performance expectancy) and the outcome (TANePS adoption) and the steps for testing mediating effects were adopted as described by [33]. The direct effect does not have to be significant for full mediation, if the indirect effect is insignificant then no mediation, the situation of partial mediation occurs when both the direct and indirect effects are significant while full mediation occurs when the direct effect is not significant and the indirect effect is significant. The hypotheses tested from the conceptual model based on direct and indirect effects of performance expectancy towards TANePS adoption and are summarized in Table 2.

Table 2: Findings of Direct and Indirect Hypotheses.

Hypothesis	Path	Influence	P-value	Remark
H_{1a}	PE->TA H_1	Direct	0.000	Supported
H_{1b}	PE->RA->TA H_3, H_4	Indirect	0.000	Supported
H_{1c}	PE->AT->TA H_2, H_5	Indirect	0.000	Supported

Source: Authors, (2021).

IV. DISCUSSION OF FINDINGS

IV.1 THE INFLUENCE OF PERFORMANCE EXPECTANCY

In this study, performance expectancy was proposed to positively, directly and indirectly influence TANePS adoption. The results of both direct and indirect influences showed positive path coefficients which implied that an increase in one standard deviation of the performance expectancy converted into an increment of the rate of TANePS adoption and both relationships were found statistically significant (p-values were < 0.05). These results implied that the relationships exist in real life. Likewise, these results are parallel to the previous studies of [21] and [22] which indicated that the direct and indirect influences of performance expectancy to new technologies adoption through attitude was statistically significant. However, the findings of the current study contrast from some prior studies. For example, [7] revealed that performance expectancy (PE) was found statistically insignificant towards Behavioural Intention to utilize Mobile. The study by [7] reasoned that performance expectancy being immaterial could propose that buyers don't just depend on performance expectancy so as to embrace and utilize portable payment directions. However, this factor additionally doesn't influence buyers' choices according to the study by [7]. In addition, the study by [13] characterized performance expectancy as the degree to which an individual accepts that utilizing a specific framework will upgrade their activity execution. A similar report set that the higher the performance expectancy among the small and medium travel industry undertakings, the higher the effect on expectation to embrace and utilize green e-business. These results indicated the immaterial influence of performance expectancy on green e-business reception because of the way that the p-value was more prominent than 0.05, while the path coefficient was under 0.2, suggesting that, it was less significant for conversation and accordingly, it was disposed of.

V. CONCLUSION

V.1 THEORETICAL CONTRIBUTION

Despite the fact that these results of the current study are consistent with the origin UTAUT but the origin UTAUT considered only individual perspectives as criticized by [22]. In this way, the current study has added new information with respect to the impact of performance expectancy on new innovation (TANePS) adoption because it has included two points of view; public and private divisions of which the current theories and prior studies are lighting up incapably. The determinant of performance expectancy from UTAUT to influence positively, directly, and indirectly the reception of new innovation (TANePS) in the public to cooperating with relative advantage and attitude from TOE is a substantial miss in the existing theories and models. In addition, this study argues that performance expectancy is one of the powerful constructs which TOE model did not consider in technological factors as one of the determinants for adopting new technologies. This argument is valid and has been justified by this study. Taking into account the relationships between performance expectancy and relative advantage and attitude for TANePS adoption, the findings revealed that performance expectancy had both direct and indirect influences towards TANePS adoption. Since performance expectancy influences new technology adoption both directly and indirectly, then, this study suggests that it should be added in technological factors list in TOE model.

V.2 EMPIRICAL CONTRIBUTION

Taking into consideration the focus of the previous empirical studies conducted in developing countries and the focus of this study, it is therefore concluded that the empirical gap which was existing has been filled. This is due to the fact that, this study has included both the perception of buyers (procurement experts from procuring entities) and suppliers a substance that the existing empirical literature was explaining insufficiently. Furthermore, the integrated conceptual model with constructs influencing directly and indirectly the adoption of new technology (TANePS) for buyers' and suppliers' perspectives has been well comprehended in this study of which the existing empirical studies are explaining the indirect relationships inadequately. The indirect relationships between performance expectancy from UTAUT and relative advantage and attitude from TOE model have been sufficiently comprehended in this study, a substance that the existing empirical literature was missing. In this study, positivism philosophy and deductive-forecasting approach were adopted. In connection to that, PLS SEM with the help of SmartPLS 3 was adopted in data analysis of which the existing empirical literature related to e-procurement adoption have used these approach and method insufficiently.

V.3 PRACTICAL IMPLICATION

The results from the tested hypotheses in the conceptual model has practical implications in terms of applicability of TANePS in the public sector. With reference to the relationship between relative advantage and TANePS adoption. For relative advantage being statistically insignificant in direct relationship with TANePS adoption could suggest that buyers and suppliers during TANePS adoption, expansion and reform do not only rely on relative advantage of the system in order to adopt and use the new system. Hence, relative advantage of the system does not influence buyers' and suppliers' decisions to adopt and use TANePS. However, performance expectancy and attitude influence buyers' and suppliers' decisions to adopt and use TANePS because the direct and indirect relationships with TANePS adoption were found statistically significant.

V.4 RECOMMENDATION

Policymakers need to restructure the ICT and public procurement policies to be in line with e-procurement adoption using the final model of this study. The final integrated model of this study has implications in public procurement policy with regard to the factors influencing TANePS adoption, expansion and reform of the system in the public sector. This study has identified performance expectancy as a predictor factor while relative advantage and attitude are mediators towards TANePS adoption as an outcome. The public procurement policy makers should take into consideration these factors during expansion and reform of TANePS.

VI. LIMITATIONS OF THE STUDY AND AREAS FOR FUTURE RESEARCH

This study was limited to three constructs which explained only 64.9% of the variance in users' intention to use TANePS which implies more constructs are required to be included. It is therefore recommended that further research may increase the number of constructs to improve the variance in users' intention to use TANePS. Additionally, in this study, attention has been paid to the application of reflective model with constructs inducing the indicators. It is therefore recommended that further research would be necessary to apply formative model with specified constructs which are induced by the indicators and have received considerable attention in the recent past.

VII. AUTHOR'S CONTRIBUTION

Conceptualization: Deus Nichodemus Shatta.
Methodology: Deus Nichodemus Shatta.
Investigation: Deus Nichodemus Shatta.
Discussion of results: Deus Nichodemus Shatta.
Writing – Original Draft: Deus Nichodemus Shatta.
Writing – Review and Editing: Deus Nichodemus Shatta.
Resources: Deus Nichodemus Shatta.
Supervision: France Shayo.
Approval of the final text: France Shayo.

VIII. ACKNOWLEDGEMENT

The authors acknowledge the Tanzania Commission for Science and Technology (COSTECH) for its facilitation and co-ordination of the workshop on scientific paper writing which was held at Bagamoyo for the purpose of promoting research development activities in Tanzania.

VIII. REFERENCES

[1] V. Venkatesh, J. Y. L., & Xu, X. Thong. "Consumer Acceptance and Use of Information Technology: Extending the Unified Theory" MIS Quarterly, vol. 36, 1, pp. 157–178, 2012.

[2] J. Iles. "How is e-procurement related to the success of U.S. cities' sustainable purchasing policies?" Arizona State University, ASU School of Sustainability, 2017.

[3] F. B. Latif. "The Adoption of Electronic Procurement in Turkey" Advances in Economics and Business, vol. 2,2, 85-91, 2014.

[4] L. M. Mgidlana. "Factors affecting the adoption of e-procurement technologies from the supplier perspective" (Master's Research Project) Gordon Institute of Business Science, University of Pretoria, 2013.

[5] J. M. Mose, J. M. Njihia, & P. O. Magutu. "The critical success factors and challenges in e-procurement adoption among large scale manufacturing firms in Nairobi, Kenya" European Scientific Journal, vol. 9, 13, pp.1857- 7431, 2013.

[6] N. I. Shale. "Role of e-procurement strategy on the performance of state corporations in Kenya" (PhD Thesis) School of Entrepreneurship and Procurement Management in the College of Human Resource Development, Jomo Kenyatta University of Agriculture and Technology, 2014.

[7] E. A. Taluka. "Understanding factors influencing consumer behavioral intention to use mobile payment services in rural Tanzania: A Case of Pwani Region" (Master's Dissertation) University of Dar es Salaam, 2016.

[8] E. O., Ibem, E.B., Aduwo, P., Tunji-Olayeni, E. K. Ayo-Vaughan, & O. U. Uwakonye. Factors influencing e-procurement adoption in the Nigerian building industry. Journal of Construction Economics and Building, 16 (4), 54-67, 2016.

[9] D. Malekia. E-Procurement as anti-corruption tool in public procurement in Tanzania. European Journal of Business and Management, 7 (14), 2018.

[10] J. Watuleke. E-procurement: evolution and adoption. a review of literature, IJRDO-Journal of Educational Research, 2456-2947, 2017.

[11] E. S. Kassim, & H. Hussin. "A success model for the Malaysian Government e-procurement system: the buyer perspective" International Journal of Electronic Government Research (IJEGR), vol.9,1, 2013.

[12] K. O. Ombat. Relationship between e-procurement systems and performance of procurement function in commercial banks in Kenya. International Journal of Economics, Commerce and Management, United Kingdom, 3 (12), 2015.

[13] J. J. Masele. "Adoption of green e-business applications for sustainable tourism development in developing countries the case of Tanzania" (PhD Thesis), Department of Business Informatics I, VLBA-University of Oldenburg, 2014.

[14] T. Oliveira, & M. F. Martins. Literature review of information technology adoption models at firm level, The Electronic Journal Information Systems Evaluation, 14 (1). Retrieved from www.ejise.com, 2011.

[15] United Republic of Tanzania National Information and Communications Technology Policy, Government Printers, Dar es Salaam, 2016.

[16] United Republic of Tanzania Public Procurement Regulatory Authority, Annual performance evaluation report for 2018/2019 financial year, 2019.

[17] United Republic of Tanzania Public Procurement Regulatory Authority, Annual performance evaluation report for 2017/2018 financial year, 2018.

[18] A. Kiwanuka. Acceptance process: the missing link between UTAUT and Diffusion of Innovation Theory. American Journal of Information Systems, 3, (2), 40-44, doi: 10.12691/ajis-3-2-3. 2015.

[19] V. Venkatesh, M. G., Morris, G. B. Davis, & F. D. Davis. "User acceptance of information technology: Toward a Unified View", MIS Quarterly, vol. 27, 3, pp. 425-478, 2003.

[20] D. N. Shatta, F. A. Shayo, & J. N. Layaa. Influence and linkage of buyers'-suppliers' attitude towards e-procurement adoption in developing countries: Tanzania context. International Journal of Applied Research in Management and Economics, https://doi.org/10.33422/ijarme.v3i1.282, 2020b.

[21] L. S. Chen, C. J., L. Y., Kuan, & H. Huang Applicability of the UTAUT model in playing online game through mobile phones: moderating effects of user experience, IEEE Int'l Technology Management Conference, 2011.

[22] Y. K. Dwivedi, N. P., Rana, A., Jeyaraj, M., Clement, & M. D. Williams Re-examining the Unified Theory of Acceptance and Use of Technology (UTAUT): towards a revised theoretical model, DOI, o10.1007/s10796-017-9774-y., 2017.

[23] L. Tornatzky, & M. Fleischer, The Process of Technology Innovation, Lexington, MA, Lexington Books, 1990.

[24] M. Aboelmaged. The drivers of sustainable manufacturing practices in Egyptian SMEs and their impact on competitive capabilities: A PLS-SEM model. Journal of Cleaner Production, 175, 207-221, doi: ohttps://doi.org/10.1016/j.jclepro.2017.12.053, 2018.

[25] T. T. Hue. The determinants of innovation in Vietnamese manufacturing firms: an empirical analysis using a technology–organization–environment framework. Eurasian Business Review, 9(3), 247-267, doi:10.1007/s40821-019-00125-w., 2019.

[26] S. M. Chege, & D. Wang. The influence of technology innovation on SME performance through environmental sustainability practices in Kenya. Journal of Technology in Society, 60, 101210, doi: https://doi.org/10.1016/j.techsoc.2019.101210, 2020.

[27] Y. Zhang, J. Sun, Z. Yang, & Y. Wang. Critical success factors of green innovation: Technology, organization and environment readiness. Journal of Cleaner Production, doi: https:// doi.org/10.1016/j.jclepro.2020.121701., 2020.

[28] D. N. Shatta, F. A. Shayo, & J. N. Layaa. Determinants of e-procurement adoption model for green procurement in developing countries: experience from Tanzania. International Academic Journal of Procurement and Supply Chain Management, 3(2), 1-18. Retrieved from http://www.iajournals.org/articles/iajpscm_v3_i2_1_18.pdf., 2020a.

[29] B. Hwang, C. Huang, & C. Wu. A TOE approach to establish a green supply chain adoption decision model in the semiconductor industry, Journal of Sustainability, doi:10.3390/su8020168., 2016.

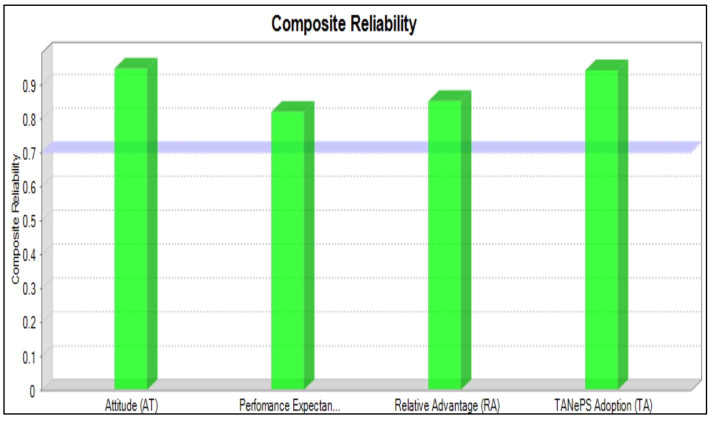
[30] E. M. Rogers. Diffusion of Innovations (5th ed.). Free Press, New York, 2003.

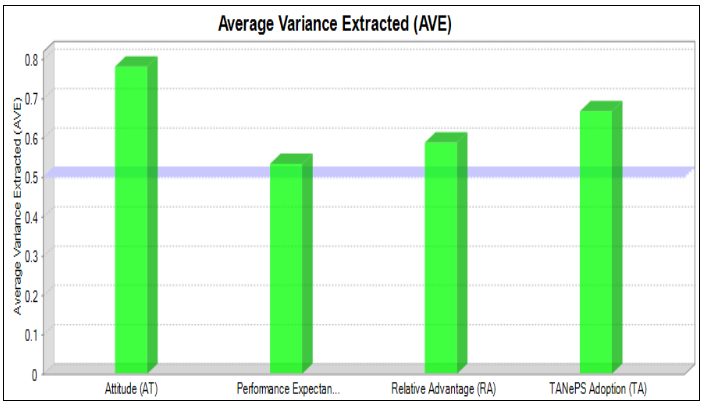
[31] M. Suleiman. Adoption of e-procurement and value addition: Tanzanian context. European Journal of Business and Management, 7(14), 2015.

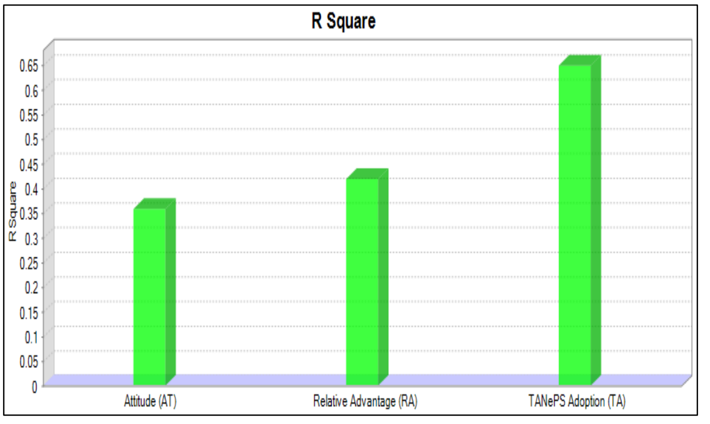
[32] J. F. Hair, C. M. Ringle, & M. Sarstedt. “PLS-SEM: Indeed, a silver bullet” The Journal of Marketing Theory and Practice, vol.19, 2, pp. 139–152, 2014.

[33] J.F. Hair, J.J., Risher, M. Sarstedt, & C.M. Ringle, (2018). “When to use and how to report the results of PLS-SEM” European Business Review, Permanent link to this document: available at: https://doi.org/10.1108/EBR-11-2018-0203 (May, 20,2020).

IX. APPENDIX A: KEY FINDINGS







Journal of Engineering and Technology for Industrial Applications

ISSN ONLINE: 2447-0228

ITEGAM-JETIA

Manaus, v.7 n.29, p. 13-24. May/Jun, 2021

DOI: https://doi.org/10.5935/jetia.v7i29.755

RESEARCH ARTICLE OPEN ACCESS

MINE SUBSIDENCE PREDICTION USING GENE EXPRESSION PROGRAMMING BASED ON MULTIVARIABLE SYMBOLIC REGRESSION

Hadi Rasouli[1], Kourosh Shahriar*[2] and Sayyed Hasan Madani[3]

[1, 2, 3] Department of Mining and Metallurgy Engineering, Amirkabir University of Technology, Tehran, Iran.

[1] http://orcid.org/0000-0002-1422-6225, [2] http://orcid.org/0000-0002-8561-6984, [3] http://orcid.org/0000-0002-6447-3646

Email: hadi.rasouli@aut.ac.ir, *k.shahriar@aut.ac.ir, hmadani@aut.ac.ir

ARTICLE INFO

Article History

Received: May 27th, 2021

Accepted: June 18th, 2021

Published: June 30th, 2021

Keywords:

Longwall mining,

Maximum subsidence,

Empirical model,

Gene expression programming.

ABSTRACT

Accurate prediction of surface subsidence becomes a significant challenge for active industrial companies in coal mining fields due to the importance of the economic impacts of longwall mining-induced subsidence. This article explores a new variant of genetic programming, namely gene expression programming (GEP). The GEP-based method is utilized to present a new mathematical formula for subsidence prediction in longwall coal mining. The derived model includes both geometrical and geological variables. The data set consists of field measurements obtained through 37 longwall panels of Ulan coal mine, NSW, Australia. The GEP-based model concluded satisfactory subsidence prediction outcomes compared to other empirical methods such as NCB, DMR, ACARP, and IPM. The predictive ability of the GEP-based models, which captures the complex nonlinear effects of the critical factors on the magnitude of subsidence, resulted in a statistically significant improvement in predictive capacity compared to the aforementioned empirical methods. The sensitivity analysis results indicated that Panel width and cover thickness with 31% and 23% were the most influential parameters in the proposed model. Also, the extracted seam thickness, thick layer location, and thick layer thickness had 19%, 16%, and 11% impact on the GEP proposed model, respectively.

I. INTRODUCTION

Longwall coal mining is the most common underground extraction method globally because of its relatively low cost, high safety, and efficiency in mining seams from depths. Longwall coal mining subsidence can affect groundwater resource and causes changes in permeability, porosity, and groundwater levels. Longwall-induced subsidence also has impacts on surface waters and associated ecosystems. Subsidence impacts can be divided into impacts on ecology, hydrology, geomorphology, and topography. Due to the importance of the mentioned effects, accurate prediction of surface subsidence due to longwall coal mining becomes a significant challenge in mining engineering. This importance includes environmental, economic, and social aspects.

Surface subsidence prediction methods, in general, include four categories: Empirical, numerical, hybrid, and physical methods. Empirical methods are based on the back analysis of field measurements and are the most common subsidence prediction methods. Some examples of these methods are the National coal Board (NCB) method, Australian Department Mineral Resource (DMR) method, Australian Coal Research Program (ACARP), influence profile method and Incremental Profile Method (IPM) [1-5].

Numerical methods use various mathematical functions to study ground movements in and around the longwall panels. Some examples of these methods include: cutting cantilever beam, key strata, and Voussoir beam structure methods [6-8].

Hybrid methods involve various mixtures of back-analysis of field observation data and using numerical and intelligent techniques. The fourth category is physical methods, which provide visual means but have little predictive value. In this study, Symbolic regression (SR)-based Gene Expression Programming (GEP) is used for subsidence prediction in longwall coal mines [9-10]. SR is a cluster of regression analysis methods to find and build the best model for accuracy and simplicity. This method requires many input and output data points to find an accurate regression. Similar to the genetic algorithm, this method includes both simple, linear chromosomes of fixed length. Like the parse trees of Genetic Programming (GP), GEP uses ramified structures of different sizes and shapes. The final output of the GEP model is a series of linear chromosomes of fixed length. In GEP, the genotype and phenotype are separated, and the model can benefit from all the evolutionary advantages. The main objective of this study is to develop a new mathematical model for predicting the maximum subsidence of longwall panels using the GEP method.

Among the various geometrical and geological factors that influence mine subsidence, five parameters include: panel width (W), extracted seam thickness (T), Cover thickness (H), the thickness of overburden thick layer (t), and the location of the massive unit above workings (Y) are used in the proposed model. Figure 1 shows the schematic of the longwall mining method [11].

The following sections of the paper are organized as follows: Section 2 gives brief literature surveys on the subsidence prediction methods. The most applicable empirical methods in Australia are discussed in more detail. In section 3, the method of research is explained. Discussion and results are provided in section 4. Finally, the conclusion is given in section 5.

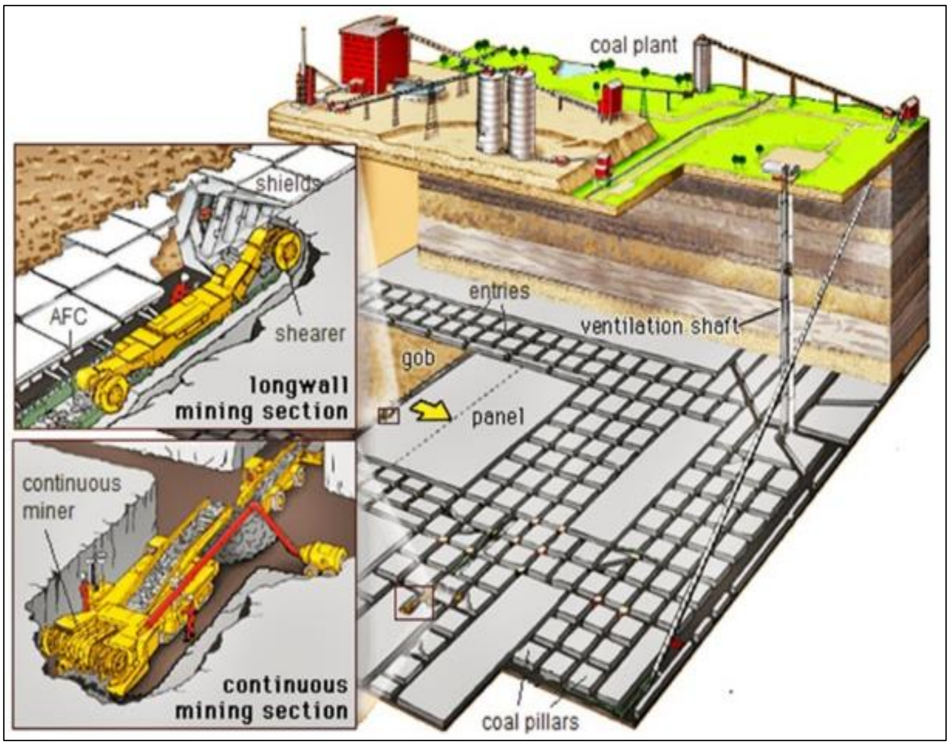


Figure 1: Schematic of longwall coal mining.
Source: [11].

II. LITURATURE SURVEY

The accuracy of mine subsidence prediction methods should never be taken for granted. The magnitude of subsidence depends on the input parameters of the specific site conditions. Empirical methods are based on the actual field measurements. These approaches predict subsidence based on parameter relationships developed from field monitoring and experience [1-4, 12-14]. The most widely used methods for predicting longwall mining-induced subsidence in Australia are described in detail.

National Coal Board in the UK proposed a subsidence prediction method during the 1960s [1]. NCB method results are based on the UK geology and do not predict subsidence magnitude accurately for other countries. During the 80s and 90s, the NCB method had been used in Australia. DMR, ACARP, and IPM methods have now replaced them. The NCB method gave good predictions when used in British mining conditions, but it provided much higher values than measured data in Australia. The difference in calculated subsidence magnitude is because the rock mechanics, geological conditions, and overlying strata of the extracted coal seam in British coalfields are different from those in Australia. The strata rocks in Britain are generally less strong and competent. Therefore, for a given seam thickness, the calculated maximum subsidence by the NCB method is greater than it would typically be for the Australian mining conditions. Figure 2 shows the NCB curves for subsidence prediction in different w/h ratios for caving and solid stowing cases.

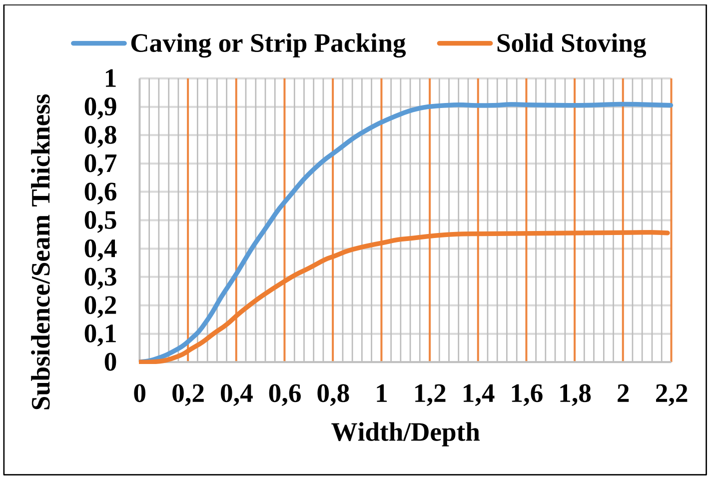


Figure 2: NCB subsidence prediction method.
Source: [1].

As a result of extensive research, NEW South Wales Department of Mineral Resources (DMR) introduced a new subsidence prediction model for Australian conditions [2]. DMR model is a modified form of the NCB method for subsidence prediction. The graphical charts of the DMR proposed model were presented in three handbooks for major coalfields of NSW. This method is also applicable for greenfield sites and where a worst-case scenario prediction is required. The data inputs to the DMR method are limited to the panel geometric variables include panel width, cover depth, and seam thickness. Like other empirical models, the DMR model is only suitable for subsidence prediction when geometrical variables are within the ranges that the model has been developed. In model development, The database from over fifty Newcastle and southern were used to model construction. As Figure 3 shows, this model is applicable when the panel (W/H) ratio is 0.2 to 2.0 and covers depth ranges between 70 m and 350 m. In recent years, the DMR method has mostly been superseded by the incremental profile method.

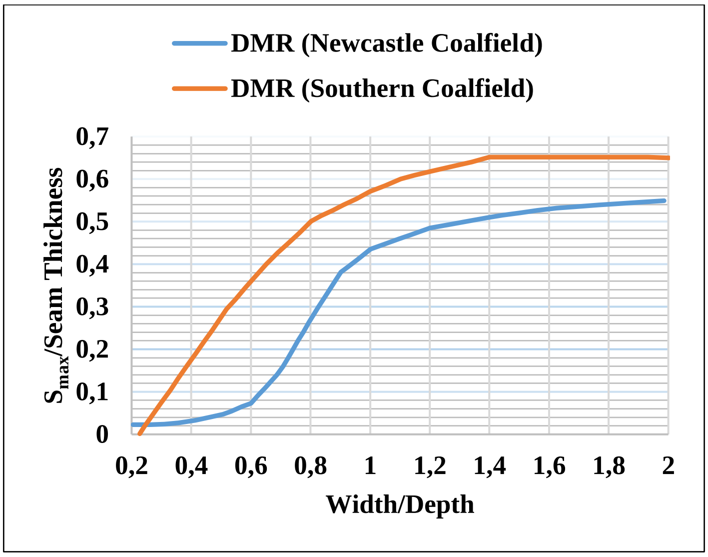


Figure 3: DMR subsidence prediction method.
Source: [2].

In 2003, the Australian Coal Research Program introduced a new model for subsidence prediction in Newcastle Coalfields [3]. ACARP model was developed based on the LAMODEL program and provided a reliable subsidence prediction model using both geometrical and geological information of longwall panels (See Figure 6).

The main focus was on the behavior of massive sandstone and conglomerate strata above the extracted coal seam. The massive geological units are classified into high, moderate, and low SRP. In the next step, according to obtained SRP factor and the thickness of the massive unit (Figure 4), maximum subsidence magnitude can be calculated from Figure 7. Upper and Lower bound prediction lines of this method for depths between 50m to 150 m is presented in Figure 7. For others depth diagrams are presented in [3]. Geometrically, the subsidence above a series of longwalls is strongly influenced by the panel width, the cover thickness, and the extraction height. Regarding geology, massive strata units above longwall panels result in reduced subsidence compared to longwall panels with similar geometry but thinner strata units.

$$S_{max} = \sqrt{12(1-\nu 2)/t)\ (\gamma H/E)\ (W2/4)} \qquad (1)$$

Where are:

ν – Poisson ratio (dimensionless),

t – Overburden thickness (m),

γ – Unit weight (N/m^3),

H – Cover thickness (m),

E –Young Modulus (N/m^2),

W– Panel width (m).

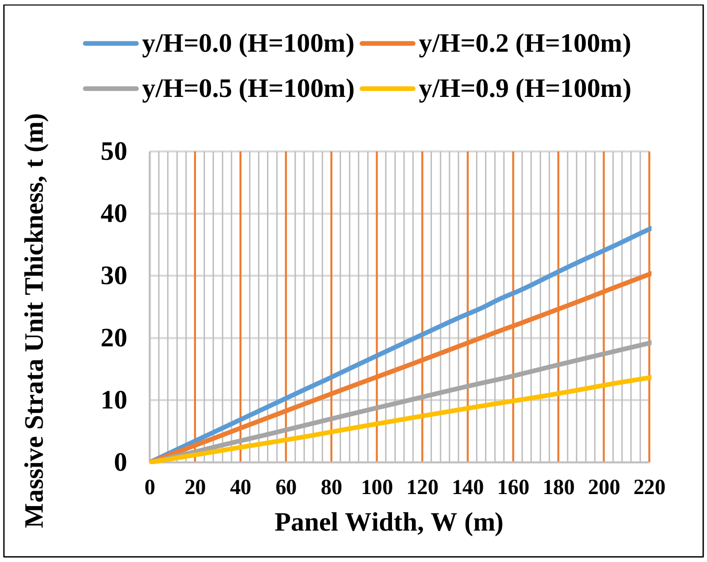


Figure 4: ACARP empirical model for predicting SRP above panels with cover thicknesses between 50 and 150m.
Source: [3].

IPM method was proposed to predict subsidence in the Newcastle Coalfield [4]. To predict subsidence with IPM requires panel width (W) and the thickness of cover (H). Figure 5 shows the details of subsidence prediction by the IPM method.

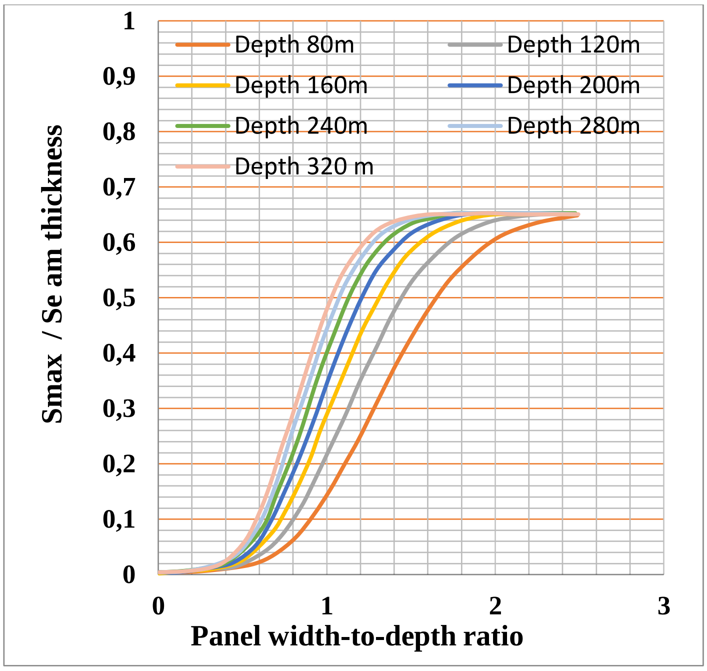


Figure 5: Subsidence profiles obtained using the incremental profile method.
Source: [4].

Empirical methods require extensive field measurement data to develop relationships applied to the mine subsidence prediction. However, when there are few case histories, an alternative approach is needed.

Numerical modeling methods simulate the geological and geotechnical conditions of the mining site to predict the

impact of various mining scenarios. Numerical modeling methods are often used when the mining method, strata conditions, and coal seam thickness differ from situations used in previously presented empirical methods. A conceptual model of the geotechnical conditions and the proposed mining methods must be developed before using numerical modeling. Computer modeling to simulate subsidence has been undertaken with various levels of success over the past years [15-17].

Developments in machine learning fields have created several new computer-aided data mining and hybrid approaches applicable for prediction problems. Artificial Neural Networks (ANN) have extensively been used to develop the nonlinear relationships between input parameters in mining and other geotechnical engineering systems [18-21]. A genetic algorithm (GA) is a robust stochastic approach for predicting various civil and mining problems. In contrast with ANNs and GA, the application of GP and its different variants, such as GEP in mining engineering, is entirely new and original. Various studies have also shown that GP, Linear Genetic programming (LGP), Multi-Expression Programming (MEP), and Gene Expression Programming (GEP) have advantages over ANNs in dealing with prediction problems [22-23].

The GEP and GP-based methods have useful applications where other standard modeling methods are complicated or detailed information for model construction don't exist [24-32].

III. METHODOLOGY

Genetic programming (GP) is one of the machine learning techniques that searches a program space. Gene Expression Programming (GEP) is an advanced form of GP technique. Ferreira presented gene expression programming [9, 10]. Traditional regression analyses, in some cases, have significant uncertainties. The regression analysis has a considerable impediment relating to complex processes. Besides that, the application of regression methods in model construction is the normality assumption of residuals. The capability of classical regression methods is also limited for the formulation of complexity, idealization of material behavior, and excessive empirical Parameters. The Gene ecpresion programming approach overcomes the constraints of various subsidence prediction methods that were previously presented. Contrary to many other soft computing tools, GEP provides prediction equations that can readily be used for subsidence prediction in longwall coal mines. The constitutive models derived using these methods can be incorporated into the different models. They may also be used to quickly check on other empirical models such as NCB, DMR, ACARP, and IPM. Similar to the genetic algorithm, this method includes both the simple, linear chromosomes of fixed length.

Like the parse trees of GP uses ramified structures of different sizes and shapes. The final output of the GEP model is a series of linear chromosomes of fixed length. In GEP, the genotype and phenotype are separated, and the model can benefit from all the evolutionary advantages. Ferreira book includes a basic algorithm of GEP and its implementation details. GEP method consists of five essential components: function set, terminal set, fitness function, control parameters, and termination points. GEP uses a fixed length of character strings to represent solutions to the problems, expressed as parse trees of different sizes and shapes. These trees are called GEP expression trees (ETs). One advantage of the GEP technique is that genetic diversity is highly simplified as genetic operators work at the chromosome level. Another advantage of GEP is its unique, multi-genic nature which allows the evolution of more complex programs composed of several subprograms. The fundamental of the GEP is schematically represented in Figure 8. The algorithm uses the following steps until a termination condition is achieved:

(1) Randomly generation of the fixed-length chromosome of each individual for the initial population;

(2) Chromosome expression as ET and fitness evaluation of individuals;

(3) Selection of the best individuals according to the fitness function;

(4) Repeating the previous stages to define several generations or until an acceptable solution is found. Standard fitness functions that are used in GEP model evaluation are as follows:

III.1 NUMBER OF HITS

When the precision is chosen for the evaluated models, the number of hits fitness function favors other fitness functions in evaluating the goodness of constructed model. The error can be either absolute or relative. The fitness F (ij) of an individual program i for fitness case j is evaluated by Equation 2 [9]:

$$\text{If } E_{(\text{ij})} \leq p, \quad \text{then } F_{(\text{ij})} = 1; \quad \text{else } f_{(\text{ij})} = 0 \qquad (2)$$

Where are:

E (ij) – is the error of an individual program i for fitness case j,

p –is the precision,

F (ij) – is the fitness of an individual program i for case j.

Equation 3 and Equation 4 show the error of an individual program in the cases in which errors are absolute and relatives, respectively [8]:

$$E_{(ij)} = \left|P_{(ij)} - T_j\right| \qquad (3)$$

$$E_{(ij)} = \left|\frac{P_{(ij)} - T_j}{T_j} \cdot 100\right| \qquad (4)$$

Where are:

$P_{(ij)}$ – Predicted value by the program i for the case j,

T_j– Target value for the case j.

$f_{\max} = n$, can occur when n is the number of fitness cases.

III.2 MEAN SQUARE ERROR (MSE)

The mean squared error expresses how close a regression line is to a set of points. It does this by taking the distances from the points to the regression line (these distances are the errors) and squaring them. The squaring is necessary to remove any negative signs. E (ij) is the error of an individual program i for fitness case j. Equation 5 and Equation 6 show the error of an individual program in instances in which errors are absolute and relatives:

$$E_i = \frac{1}{n}\sum_{j=1}^{n}(P_{(ij)} - T_j)^2 \qquad (5)$$

Where are:
n – Number of fitness cases,
E_i – Absolute error for the program i,
$P_{(ij)}$ – Predicted value by the program i for the case j,
T_j– Target value for the case j.

$$E_i = \frac{1}{n}\sum_{j=1}^{n}\left(\frac{P_{(ij)} - T_j}{T_j}\right)^2 \qquad (6)$$

Where are:
E_i– is the relative error for the program i.
Other parameters are previously stated in Equation 5.
Thus, for a perfect fitness, $P_{(ij)} = T_j$, and $E_{ij} = 0$.
To evaluate the fitness f_i of the program i, the following.

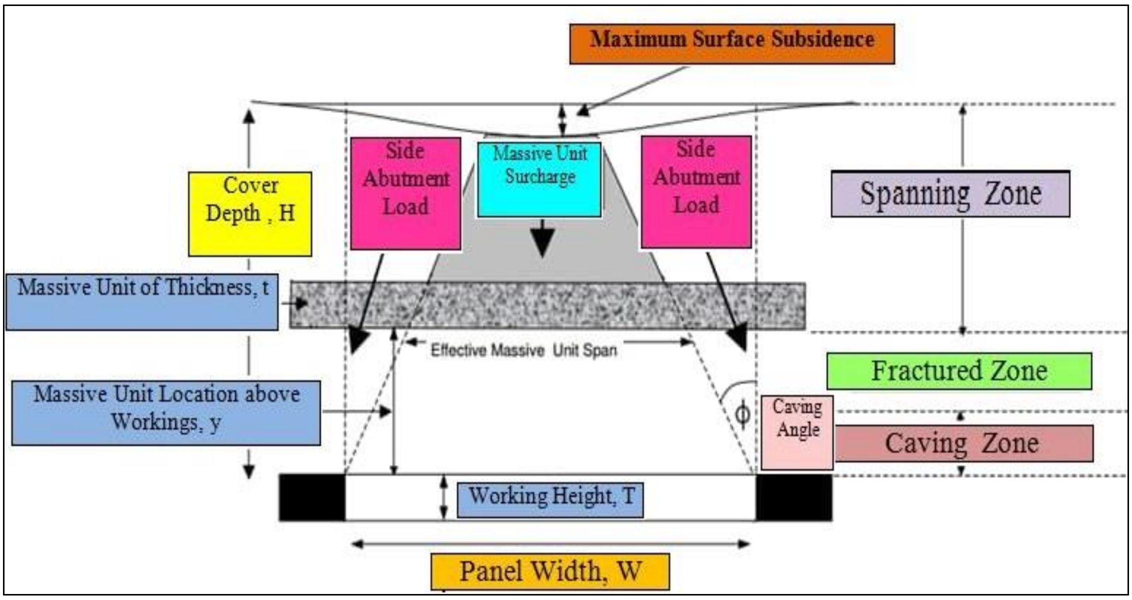


Figure 6: Key parameters of ACARP subsidence prediction method.
Source [3].

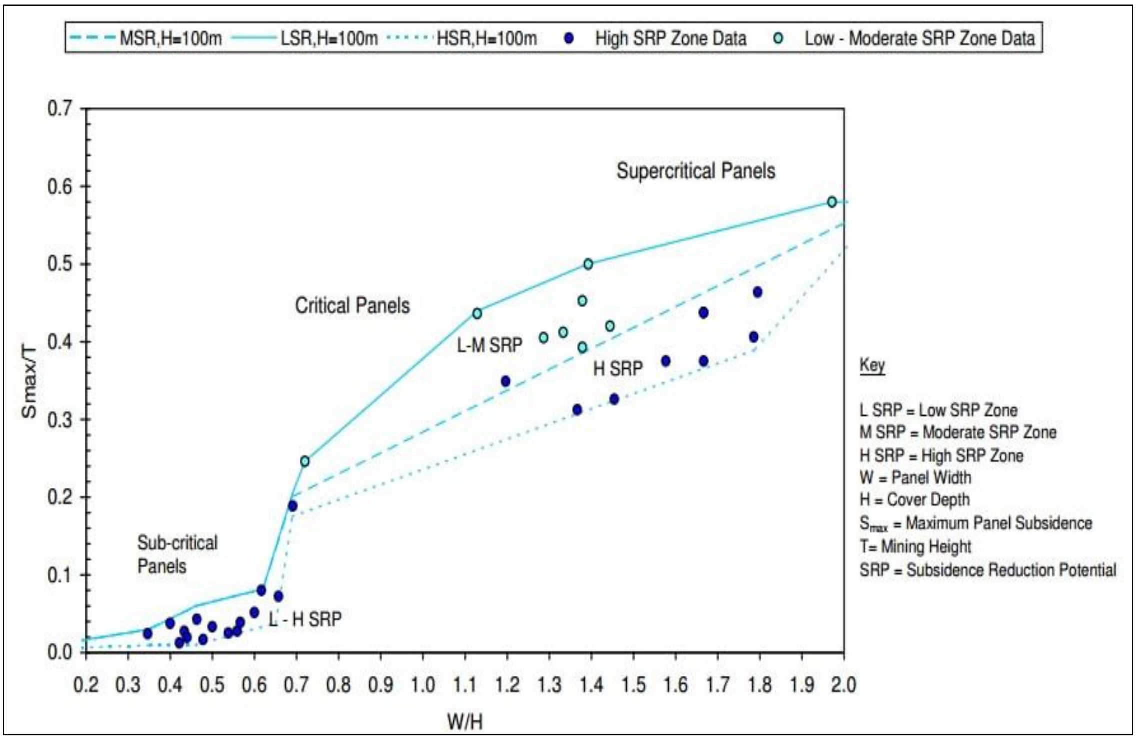


Figure 7: Empirical model for predicting subsidence above panels (cover thicknesses 50 m to 150 m and low to high SRP zones).
Source [3].

Equation can be used [9]:

$$f_i = 1000 \cdot \frac{1}{1 + E_i} \qquad (7)$$

Where are:
f_i–Total fitness of the program i,
E_i– Total error of the program i.
f_i ranging from 0 to 1000, with 1000 corresponding to the ideal conditions.

III.3 R-SQUARE

R-square is the square of the Pearson product-moment correlation coefficient, which can be calculated as Equation 8 [9]:

$$R_i = \frac{n\sum_{j=1}^{n}(T_j P_{(ij)}) - (\sum_{j=1}^{n} T_j)(\sum_{j=1}^{n} P_{(ij)})}{\sqrt{\left[n\sum_{j=1}^{n} T_j^2 - (\sum_{j=1}^{n} T_i)^2\right]\left[n\sum_{j=1}^{n} P_{(ij)}^2 - (\sum_{j=1}^{n} P_{(ij)})^2\right]}} \quad (8)$$

Where are:
R_i– Pearson correlation coefficient.
Other parameters are stated in Equation 5.

The fitness of an individual program is a function of the correlation coefficient and is defined by the Equation 9 [9]:

$$f_i = 1000 \cdot R_i^2 \quad (9)$$

f_i ranges from 0 to 1000, with 1000 corresponding to the ideal fitness.
d. Precision and Selection Range.

The fitness f_i of program i is expressed by Equation 10 and Equation 11 for absolute and relative errors [9]:

$$f_i = \sum_{j=1}^{n}(R - |P_{(ij)} - T_j|) \quad (10)$$

$$f_i = \sum_{j=1}^{n}(R - |\frac{P_{(ij)} - T_j}{T_j} \cdot 100|) \quad (11)$$

Where are:
R – selection range
$P_{(ij)}$, and T_j are the parameters previously defined.
Thus, for a perfect fit, $P_{(ij)} = T_j$ for all cases *and* $f_i = R$.

In the second step, the terminals and functions are determined. The third step involves selecting chromosomal structures such as head and tail size and the number of genes. and finally, in the fourth step, the linking function is defined, and genetic operators are determined. Detailed descriptions of stages 2 to 4 are mentioned in [9].

IV. RESULTS AND DISCUSSION

Data from Ulan mines have been used in the modeling process [33-38]. Ulan coal mines are located in New South Wales, Australia, and are very similar in geological and geotechnical characteristics. The collected empirical data include 37 longwall panels and 119 measured subsidence from them. Most of the model data, about 79% of them were related to old Ulan mine. This section describes Ulan coal mines characteristics. Ulan Coal Mine Complex (UCMC) is situated in the central west of New South Wales. It is located near the village of Ulan, approximately 38 kilometers north-northeast of Mudgee and 19 kilometers northeast of Gulgong. Coal mining started in the Ulan area in the 1920s, consisting of Old Ulan, No.3 underground mines, Ulan West areas, and open-cut mining. Figure 9 shows the locality plan of the complex [38].

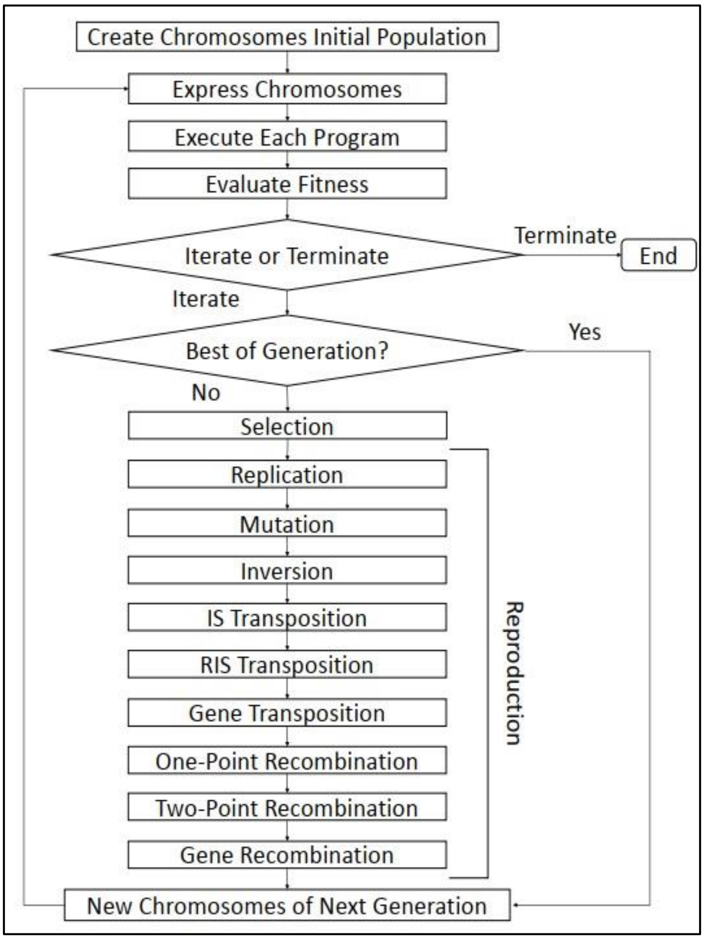


Figure 8: The flowchart of a gene expression algorithm.
Source: [9].

Figure 9: locality of the Ulan coal mines complex.
Source: [38].

The coal seams in the region range in thickness from 0.4 to 1o meters, and the Ulan coal seam is the thickest among them. Except for the Ulan, other seams are uneconomical due to high ash content. Massive Triassic sandstone, siltstone, and Narrabeen conglomerate overlie the Permian Coal Measures. Figure 10 shows Ulan coal mine longwalls and their positions [33].

The research aims to develop a new mathematical model for predicting the maximum subsidence of longwall panels using the GEP method. The GEP model uses the five influencing input parameters as Equation 12:

$$\text{Smax} = \text{F}(\text{W}, \text{T}, \text{H}, \text{t}, \text{Y}) \qquad (12)$$

Where are:
W– Panel width (m),
T– Extracted seam thickness (m),
H – Cover thickness (m),
t – Thickness of overburden thick layer (m),
Y– Massive unit location above workings (m).

The fitness function for model development is the mean square error (MSE). The mean squared error Ei of an individual program i is evaluated using Equation 5 and Equation 6. For assessing the fitness fi of individual program i, Equation 6 is used. Various parameters used in the GEP model are shown in Table 1.

Table 1: Experiment parameters of the GEP model.

Parameters	Values
Fitness Function	Equation 6
Population Size	30
Number of Generation	1000
Head Length	5
Number of Genes	3
Chromosome Length	33
Function Set	+, -, /, *, tan, inv
Terminal Set	c0, …, c3
Link Function	+
Mutation Rate	0.004
Inversion Rate	0.01
IS Transposition Rate	0.01
RIS Transposition Rate	0.01
One-Point Recombination	0.3
Two-Point Recombination	0.3
Gene Recombination	0.1
Random Numbers	[-100,100]

Source: Authors, (2021).

Ulan longwall panels have all been carefully scheduled according to the timetable provided in Table 2.

Table 2: Extraction timetable of Ulan longwall panels.

Long wall	Start	Finish	Longwall	Start	Finish
LW1	07-12-86	30-11-87	W4	01-09-13	02-05-14
LW2	20-12-87	15-10-88	LW28	03-07-14	05-03-15
LW3	30-11-88	31-08-89	W5	06-05-15	01-03-16
LW4	05-12-89	15-09-90	LW29	02-05-16	01-11-16
LW5	15-10-90	04-01-92	W6	02-01-17	01-10-17
A	15-05-92	30-08-92	LW30	02-12-17	04-04-18
B	05-10-92	28-02-93	W7	05-06-18	07-04-19
LW6	15-03-93	30-07-93	LW31	08-06-19	01-02-20
LW7	07-10-93	30-05-94	W8	09-04-20	07-11-20
LW8	15-06-94	15-02-95	LW32	08-01-21	09-07-21
LW9	22-03-95	26-10-95	W9	09-09-21	10-03-22
LW10	05-12-95	23-08-96	W10	11-05-22	08-10-22
LW11	25-10-96	26-11-97	LW33	09-12-22	08-02-23
LW12	23-10-97	01-07-98	W11	11-04-23	11-05-23
LW13	29-07-98	21-04-99	Ulan West		
LW14	21-07-99	01-04-00	UW1	02-01-12	01-01-14
LW15	31-05-00	22-02-01	UW2	01-03-14	01-03-15
LW16	20-03-01	08-10-01	UW3	01-05-15	01-05-16
LW17	06-11-01	21-07-02	UW4	01-07-16	01-09-17
LW18a	26-07-02	23-02-03	UW5	01-11-17	01-11-18
LW19	11-04-03	03-11-03	UW6	10-01-19	12-05-20
LW20a	10-12-03	10-12-03	UW7	13-07-20	10-10-21
LW21	27-10-04	27-10-04	UW8	10-12-21	12-06-23
LW22	23-09-05	23-09-05	UW9	11-08-23	10-01-25
LW23	25-10-06	11-09-07	UW10	12-03-25	12-04-26
LW24	05-11-07	20-03-08	UW11	12-06-26	10-08-28
W1	26-05-08	12-02-09	North 1		
LW25	01-04-09	01-11-09	LW-C	01-04-14	08-09-14
W2	01-01-10	01-11-10	LW-D	01-04-11	08-09-11
LW26	01-11-11	01-0911	LW-E	01-04-12	08-09-12
W3	01-11-11	03-08-12	LW-F	01-04-13	08-09-13
LW27	01-11-12	01-07-13	LW-G	01-04-15	08-09-15

Source: Authors, (2021).

For example, Table 3 shows a part of the measured subsidence data of old Ulan coal mine panels.

Table 3: Measured subsidence of old Ulan coal mine panels.

LW Panel	Measured Subsidence (m)	LW Panel	Measured Subsidence (m)
A	1.2	11C	1.4
B	0.93	11X	1.4
1	1.5	12D	1.3
5	1.0	13D	1.3
6	0.13	14D	1.1
7	1.0	15D	0.96
8	1.0	16D	1.1
9	1.2	17D	1.2
10	1.3	18E	1.1
11	1.4	19E	1.2

Source: [39].

Table 4 expresses the Performance results of the proposed GEP model in its different stages. Table 6 suggests the final GEP model equations and sub-trees in details.

Table 4: Proposed GEP model performance in different stages.

Model	**Experiment vs Prediction**		
Stage	R	MAD	RMSE
Learning	0.904	0.191	0.202
Test	0.892	0.184	0.197
Validation	0.881	0.181	0.192

Source: Authors, (2021).

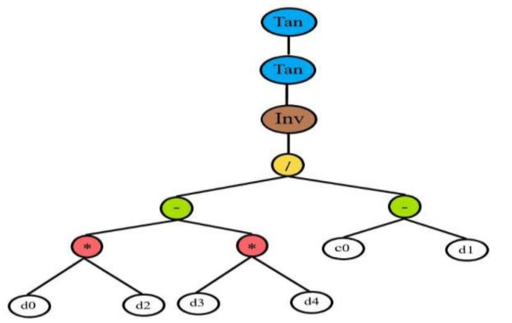
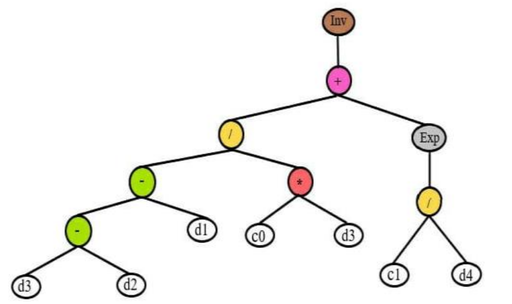
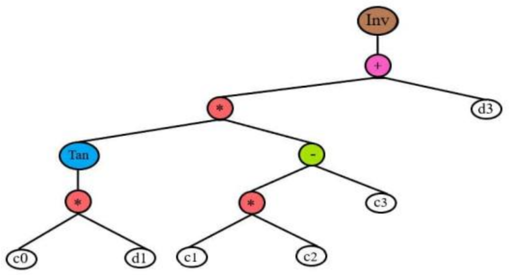
Table 5 and Figure 10 shows the importance of each parameter in the proposed GEP model. Panel width parameter with 31%, cover thickness with 23%, extracted seam thickness with 19%, thick layer location with 16%, and thick layer thickness with 11% impact the proposed model.

Table 5: Sensivity analysis results of the proposed GEP model.

Parameter	**Unit**	**Symbol**	**Importance in the GEP model (%)**
Panel width (W)	m	d0	31%
Cover thickness (H)	m	d1	23%
Extracted seam thickness (T)	m	d2	19%
Thickness of overburden thick layer (t)	m	d4	11%
Massive unit location above workings (Y)	m	d3	16%

Source: Authors, (2021).

Table 6: Sub-equations and sub-trees of the proposed GEP model results.

Equations	**Expression Trees (Ets)**
$\text{Tan}\left[\tan\left(\frac{c9 - d1}{(d0 * d2) - (d3 * d4)}\right)\right]$ (13) Where are: d0 – is the Panel width (m), d1 – is the cover thickness (m), d2 – is the thickness of extracted seam (m), d3 – is the thickness of overburden thick layer (m), d4 – is the massive unit location above workings (m), c0 = 11.2418.	
$\left[(d3 - d2) - \frac{d1}{c6 * d3}\right] + e^{\frac{c5}{d4}}$ (14) Where are: d1 – is the cover thickness (m), d2 – is the thickness of extracted seam (m), d3 – is the thickness of overburden thick layer (m), d4 – is the massive unit location above workings (m), c0 = -1.9457 , c1 = -8.8704.	
$\text{Tan}(c0 * d1)] * [(c2 * c4) - c1] + d3]$ (16) Where are: d1 – is the cover thickness (m), d3 – is the thickness of overburden thick layer (m), c0 = 144.932 , c1= 1.6201, c2 = 9.0104 , c3 = -3.7042.	

Source: Authors, (2021).

Table 7: Summary of the most applicable empirical methods for prediction of mine subsidence in Australia.

Method	Use in Australia	Main advantages	Main disadvantages
NCB	Less commonly used today and is used more as a rule of thumb to compare with other methods	Fast, low cost, requires fewer input parameters	The accuracy of the method when used in Australia is low
DMR	Widely used before 90s. It is used as an auxiliary method along with other subsidence prediction methods	Fast, low cost, requires fewer input parameters	Doesn't take into account the presence of the overlying thick strata
ACARP	Relatively common method	Takes into account the presence of the overlying thick strata	Its application limited to Newcastle, New South Wales
IPM	The most common subsidence prediction method in Australia	It can be implemented quickly and cheaply relative to numerical modeling methods	It requires more input data than NCB and DMR methods

Source: Authors, (2021).

Table 8: Subsidence prediction results in the most applicable empirical methods for prediction of mine subsidence in Australia.

Mine	W	H	T	t	Y	NCB	DMR	ACARP	IPM	GEP	Measured Subsidence
Ulan West	261	75	3.2	10	15	2.88	1.68	1.72	1.68	1.72	1.91
Ulan West	261	75	3.2	15	20	2.88	1.68	1.72	1.68	1.91	1.82
Ulan West	315	140	3.2	20	65	2.88	1.68	1.22	1.42	1.48	1.36
Ulan West	315	215	3.2	35	30	2.88	1.6	1.47	1.42	1.13	0.97
Old Ulan	216	160	2.54	15	13	2.27	1.19	1.17	1.29	1.55	1.37
Old Ulan	216	162	2.52	18	19	2.27	1.19	1.23	1.27	1.49	1.36
Old Ulan	216	167	2.54	20	23	2.29	1.17	1.27	1.25	1.44	1.35
Old Ulan	198	169	2.54	30	55	2.23	1.14	1.07	1.14	0.84	0.64

Source: Authors, (2021).

For validation purposes, the performance of the proposed GEP model and other maximum subsidence prediction methods reviewed in this research are compared with measured subsidence. The DMR, ACARP, and IPM methods had been proposed based on the database from NSW. Ulan coal mine complex has the same geological and geotechnical similarities with those of mentioned methods. Although the NCB method is proposed for UK mines, it has been used as a rule of thumb in comparing results of other empirical methods.

Table 7 provides a summary of the main empirical and proposed GEP models. Validation results of the proposed GEP model vs. other methods are provided in Table 8.

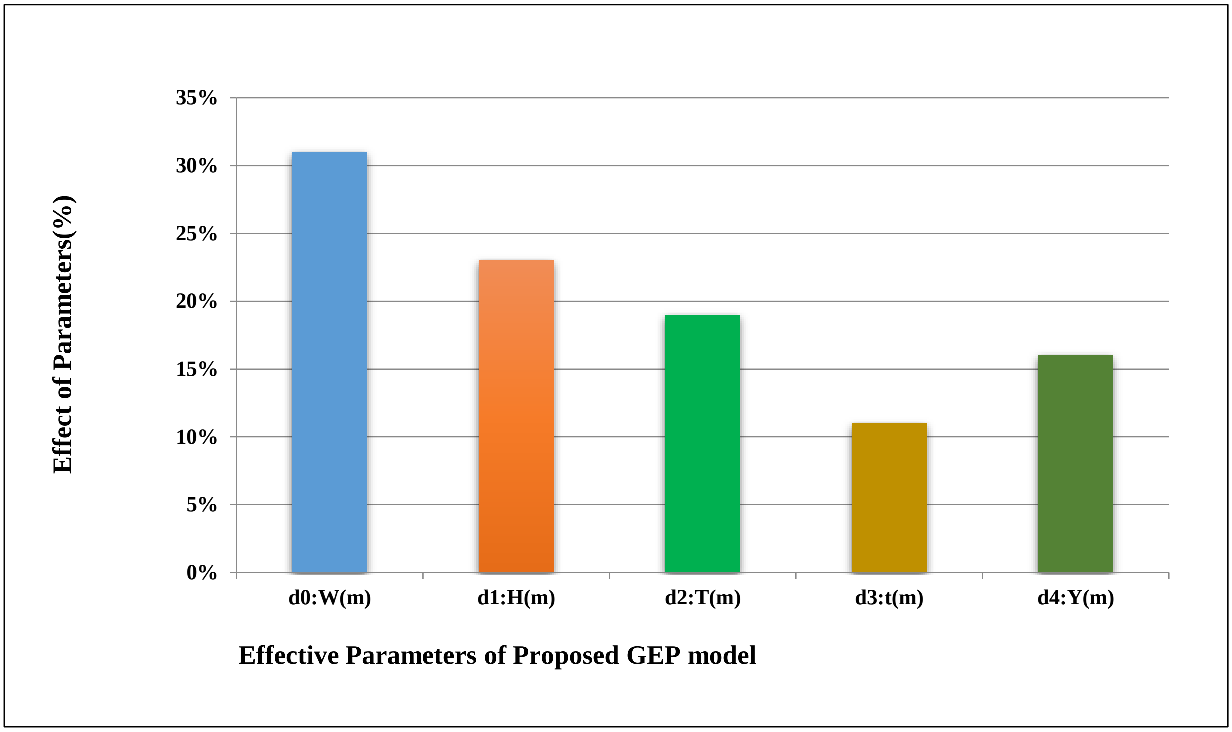


Figure 10: Sensitivity analysis results of the proposed GEP model.
Source: Authors, (2021).

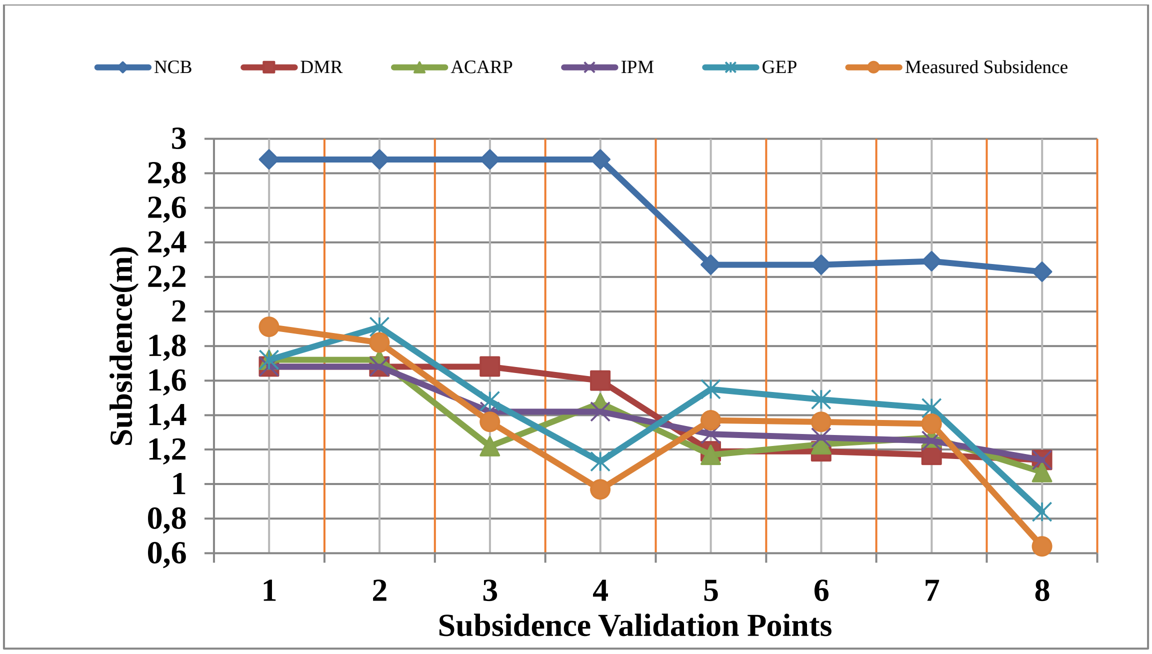


Figure 11: Prediction methods results vs. measured data.
Source: Authors, (2021).

As shown in Figure 11, the NCB method is based on the UK longwall mines. comparing the results of NCB method with the measured data is unsatisfactory. Compared to the previous methods presented with the measured data, the IPM has a minor error and higher accuracy. The IPM method has been proposed by Mine Subsidence Engineers Consultant into proprietary software and is available at a cost. Compared with IPM, the proposed GEP method has higher accuracy and requires less time and cost. Figure 12 and Table 9 suggest the statistics between measured subsidence and subsidence resulted from various prediction methods.

Table 9: Statistics results of the proposed GEP model vs. other methods.

Prediction method	Experiment vs Prediction			
	R	R^2	MAD	RMSE
NCB	0.46	0.21	1.26	1.28
DMR	0.72	0.52	0.26	0.34
ACARP	0.74	0.54	0.22	0.27
IPM	0.88	0.76	0.19	0.24
GEP	0.88	0.77	0.18	0.19

Source: Authors, (2021).

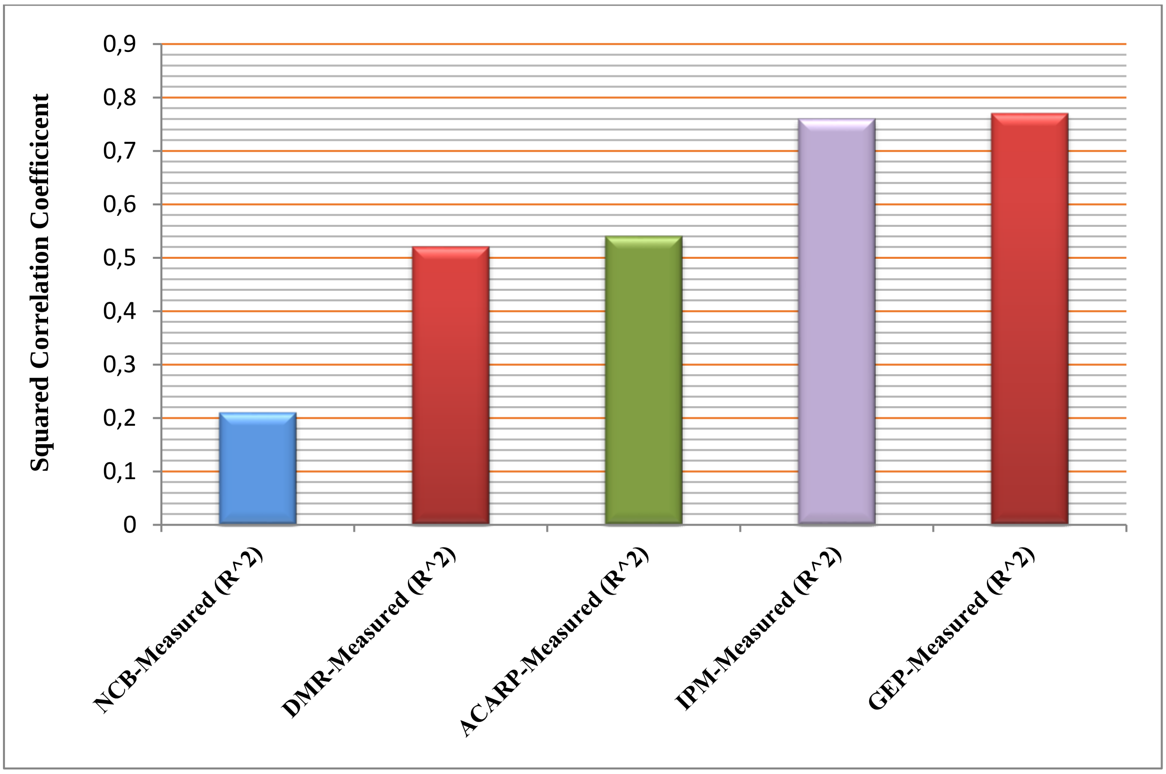


Figure 12: Squared correlation coefficients of measured subsidence and other prediction methods.
Source: Authors, (2021).

V. CONCLUSION

The present research introduced a new gene expression programming model based on multivariable symbolic regression for subsidence prediction in Ulan longwall mines. GEP can use linear, nonlinear functions and constant numbers without prior information about the final model. The proposed model was carried out through measured data as training inputs. Finally, a new mathematical formula for subsidence prediction was proposed. The maximum vertical subsidence was modeled in various affecting parameters (W, H, T, t, y). The proposed model was constructed using measured subsidence data from Ulan coal mines. By comparing the results of the proposed GEP model with NCB, DMR, ACARP, and IPM methods, it was observed that the GEP-based model was accurate enough and had the potential to be used in mines with similar geological and geotechnical conditions. The sensitivity analysis results indicated that Panel width with a 31% effect was the most influential parameter in the proposed model. Also, the impact of panel depth, extracted seam thickness, location, and thickness of the thick overburden layer were 23%, 19%, 16%, and 11%, respectively. The proposed model is expected to help predict subsidence where geological and geotechnical conditions are similar to that of the Ulan coal mines. Future research can include some new parameters individually or mixed to present new optimized subsidence prediction models.

VI. AUTHOR'S CONTRIBUTION

Conceptualization: Hadi Rasouli and Kourosh Shahriar.
Methodology: Hadi Rasouli, Kourosh Shahriar, and Sayyed Hasan Madani.
Investigation: Hadi Rasouli and Sayyed Hasan Madani.
Discussion of results: Hadi Rasouli and Kourosh Shahriar.
Writing – Original Draft: Hadi Rasouli and Sayyed Hasan Madani.
Writing – Review and Editing: Kourosh Shahriar.
Resources: Kourosh Shahriar and Sayyed Hasan Madani.
Supervision: Koroush Shahriar and Sayyed Hasan Madani.
Approval of the final text: Hadi Rasouli, Kourosh Shahriar, and Sayyed Hasan Madani.

VII. REFERENCES

[1] National Coal Board, "Subsidence engineers handbook", 2nd edition, Mining Department, London, UK, pp. 04-10, 1975.

[2] L. Holla, "Surface Subsidence Prediction in the Newcastle Coalfield", Department of Mineral Resources, NSW, Australia, pp. 13-23, 1987.

[3] MAN-001/1, "Summary of the Modified ACARP 2003 Empirical Model", Ditton Geotechnical Services, NSW, Australia, https://majorprojects.planningportal.nsw.gov.au, pp. 1-152, 2013.

[4] MSEC 309, "The Prediction of Subsidence Parameters of Proposed Austar Longwalls A6 to A17", Mine Subsidence Engineering Consultants, NSW, Australia, https://majorprojects.planningportal.nsw.gov.au, pp. 21-27, 2008.

[5] M. Hood, R.T. Evy, and L.R. Riddle, "Empirical methods of subsidence prediction, a case study from Illinois", International Journal of Rock Mechanics, 20 (4), pp. 153-170, 1983.

[6] C. He, J. Xu, "Subsidence prediction of overburden strata and surface based on the Voussoir beam structure", Advance in Civil Engineering, 1, pp. 1-14, 2018.

[7] Z. Li, J. Xu, "The effects of the rotational speed of voussoir beam structures formed by key strata on the ground pressure of stopes", International Journal of Rock Mechanics, 108, pp. 67-79,2018.

[8] C. Liu, "Voussoir beam model for lower strong roof strata movement in longwall mining", J. Rock Mech. Geotech. Eng. 9 (6), pp. 1171–1176, 2018.

[9] C. Ferreira, "Gene expression programming, Mathematical modelling by an artificial intelligence", vol.21, Springer, United Kingdom, pp. 1-492, 2006.

[10] C. Ferreira, "Gene expression programming: a new adaptive algorithm for solving problems", Complex Systems, 13(2), pp. 87–129, 2001.

[11] https://www.britannica.com/technology/coal-mining, Last accessed 03/20/2017.

[12] J. Yue, W. Cheng, and L. Fan, "The Study on mathematical model of urban land subsidence based on statistical analysis", International Conference on Management and Service Science, Wuhan, China, pp. 20–22, September 2009.

[13] G. Yu, W. Mi, D. Wang, L. Gao, and S. Lu, "Research on the relationship between the surface dynamic subsidence and overburden separated strata of coal mine and its model", Procedia Eng,, 191, pp. 196–205,2017.

[14] B. Wang, J. Xu, and D. Xuan, "Time function model of dynamic surface subsidence assessment of grout-injected overburden of a coal mine", Int. J. Rock Mech. Min. Sci, 104, pp. 1–8,2018.

[15] H. Li, J. Zha, G. Guo, "A new dynamic prediction method for surface subsidence based on numerical model parameter sensitivity", J. Clean Prod, 233, pp. 1418–1424, 2019.

[16] P. Tzampoglou, C. Loupasakis, " Numerical simulation of the factors causing land subsidence due to overexploitation of the aquifer in the Amyntaio open coal mine", Greece, Hydro Research, 1, pp. 8 –24, 2019.

[17] A. Zingano, A. Weiss, " Subsidence over room and pillar retreat mining in a low coal seam", Int. J. Min. Sci. Technol, 29, pp. 51–57, 2019.

[18] Y. Cha, W. Choi, and O. Buyukozturk, "Deep learning-based crack damage detection using convolutional neural networks", Comput. Aided Civ. Infrastruct. Eng, 32, pp. 361–378, 2017.

[19] M.E. Yetkin, M.K. Ozfirat, "Selection of thick coal seam mining method using analytic hierarchy process", JETIA, vol. 5, no. 20, pp. 06-11, Dec. 2019.

[20] M. Juszczyk, A. Le'sniak, "Modelling construction site cost index based on neural network ensembles", Symmetry, 11, pp. 411, 2019.

[21] V. Singh, S. Bano, A.K. Yadav, and S. Ahmad, "Feasibility of artificial neural network in civil engineering", Int. J. Trend Sci. Res. Dev, 3, pp. 724–728, 2019.

[22] B.S. Narendra, P.V. Sivapullaiah, S. Suresh, and S.N. Omkar, "Prediction of unconfined compressive strength of soft grounds using computational intelligence techniques: a comparative study", Computers and Geotechnics, Vol. 33, pp. 196-208, 2006.

[23] C. Kayadelen, O. Naydın, M. Fener, A. Demir, and A. Ozvan,"Modeling of the angle of shearing resistance of soils using soft computing systems", Expert Systems with Applications, Vol. 36, pp. 18- 26, 2009.

[24] M.N. Amar, "Prediction of hydrate formation temperature using gene expression programming", Journal of Natural Gas Science and Engineering, 87, pp. 24-36, 2021.

[25] I. Azim, J. Yang, M.F. Iqbal, M.F., Z. Mahmood, F. Wang, and Q.F. Liu, "Prediction model for compressive arch action capacity of RC frame structures under column removal scenario using gene expression programming", Structures, 25, pp. 212-228, 2020.

[26] A. Fathi, M. Mazari, and M. Saghafi, "Multivariate global sensitivity analysis of rocking responses of shallow foundations under controlled rocking", Eighth International Conference on case histories in geotechnical engineering, Geo- Congress, GSP 307, pp. 490-498, 2019.

[27] T. Hong, K. Jong, and Ch. Koo, "An optimized gene expression programming model for forecasting the national CO2 emissions in 2030 using the metaheuristic algorithms", Applied Energy, 228, pp. 208-220, 2018.

[28] A.I. Lawal, S. Kwon, O.S. Hammed, and M.A. Idris, "Blast-induced ground vibration prediction in granite quarries: An application of gene expression

programming, ANFIS, and sine cosine algorithm optimized ANN", International Journal of Mining Science and Technology, 31, pp. 265-277, 2021.

[29] H. Majidfar, B. Jahangiri, W.G. Buttlar, and A.H. Alavi, "New machine learning-based prediction models for fracture energy of asphalt mixtures", Measurement 135, pp. 438–451, 2019.

[30] S.M. Mousavi, P. Aminian, A.H. Gandomi, A.H. Alavi, and H. Bolandi, "A new predictive model for compressive strength of HPC using gene expression programming", Advances in Engineering Software, 45, pp. 105-114, 2012.

[31] S. Moyano, O. Reyes, H. Fardun, and S. Ventura, "Performing multi-target regression via gene expression programming-based ensemble models", Neurocomputing, 432, pp. 275-287, 2021.

[32] Y.Z. Murad, R. Hunifat, and W. Al-Bodour, (2020): Interior Reinforced Concrete Beam-to-Column Joints Subjected to Cyclic Loading: Shear Strength Prediction using Gene Expression Programming. Case Studies in Construction Materials, 13, pp. 1-9, 2020.

[33] ULA 3367, "Subsidence assessment of Ulan coal mine", Strata Control Technology, NSW, Australia, https://www.glencore.com.au/operations-and-projects/coal/current-operations/ulan-coal/reporting-documents, pp. 1-119, 2009.

[34] ULA 3370, "Subsidence assessment, North 1 underground mining area", Strata Control Technology, NSW, Australia, https://www.glencore.com.au/operations-and-projects/coal/current-operations/ulan-coal/reporting-documents, pp. 3-19, 2009.

[35] ULA, "Analysis of subsidence results from longwall W2, comparison with predictions", Strata Control Technology, NSW, Australia, https://www.glencore.com.au/operations-and-projects/coal/current-operations/ulan-coal/reporting-documents, pp. 1-92, 2009.

[36] ULA 4510, "2015 annual review of subsidence monitoring at Ulan west mine", Strata Control Technology, NSW, Australia, https://www.glencore.com.au/operations-and-projects/coal/current-operations/ulan-coal/reporting-documents, pp. 1-33, 2009.

[37] ULN SD PLN 0103, "Ulan West Extraction Plan for longwalls LW1 to LW4", Ulan Coal GLENCORE, NSW, Australia, https://www.glencore.com.au/operations-and-projects/coal/current-operations/ulan-coal/reporting-documents, pp. 11-29, 2015.

[38] https://www.glencore.com.au/operations-and-projects/coal/current-operations/ulan-coal, NSW, Australia, Last accessed 07/14/2019.

[39] WHT/1, Mine subsidence impact assessment for the proposed longwall panels LWs 1 to 14, No.4 underground area, Moolarben coal project, NSW, Australia, http://www.moolarbencoal.com.au/icms-docs/259060-response-to-submissions-appendices-8-9.pdf, pp. 30-47, September 2006.

Journal of Engineering and Technology for Industrial Applications

ISSN ONLINE: 2447-0228

ITEGAM-JETIA

Manaus, v.7 n.29, p. 25-35. May/Jun, 2021

DOI: https://doi.org/ 10.5935/jetia.v7i29.759

RESEARCH ARTICLE OPEN ACCESS

APPLICATION OF STATISTICAL PROCESS CONTROL (SPC) IN THE VOLUME VARIATION OF A VOLATILE LIQUID FUEL

Everaldo de Queiroz Lima*[1], Cecília Lenzi[2], Maurício Moyses Machado[3] and Fátima Geisa Mendes Teixeira[4]

[1, 2, 4] Chemical Engineering, Lutheran University Center of Manaus – CEULM/ULBRA, Manaus, Amazonas, Brazil.
[3] University of Uberaba – UNIUBE. Uberaba, Minas Gerais, Brazil.

[1] http://orcid.org/0000-0002-6887-4594, [2] http://orcid.org/0000-0003-3123-0610, [3] http://orcid.org/0000-0002-8509-7117, [4] http://orcid.org/0000-0002-1106-980X

Email: *everaldolima.am@gmail.com, cecilia.lenzi@yahoo.com.br, mauricioepga@gmail.com, fatimageisa2015@gmail.com

ARTICLE INFO

Article History
Received: June 09th, 2021
Accepted: June 18th, 2021
Published: June 30th, 2021

Keywords:
Control chart,
Fuel,
Statistical Process Control,
Volume.

ABSTRACT

Statistical Process Control makes it possible to monitor the quality characteristics of interest, ensuring their maintenance within pre-established limits and indicating when to adopt corrective and improvement actions. The problem addressed in this work The large variation in the net mass of a volatile fuel derived from petroleum (gasoline A) stored in atmospheric tanks at a Fuel Distributor in Manaus-AM. Data on the variation in the volume of gasoline A were studied in the process of transferring between tanks. The data for the study were obtained from the Distributor's Net Cargo Tonnage Certificates (CNCT) and Quality Certificates (QC) and tabulated in an electronic spreadsheet using the Microsoft Excel software from Windows, in which the graphs of the volume variation values, specification limits (LSL and USL) and process control limits (LCL and UCL) were also plotted. All results point to an objective way of demonstrating the high degree of safety required in this process, the studied process control parameters that form the variation of the ambient volume and volume at 20 °C of gasoline A, remained within the limits of control and as for the capacity indexes it was concluded that the process produced high capacity indexes.

I. INTRODUCTION

Statistical Process Control (SPC) is a statistical technique applied to production that allows the systematic reduction of the variability in the characteristics of the quality of interest, contributing to the improvement of the intrinsic quality, productivity, reliability and cost of what is being produced [1]. The production, storage and transport of fuels in terms of volumetric quantity is a growing concern, the volume of fuels can be compromised from production at an oil refinery or biofuel plant to the final consumer tank. [2, 3]. The amount of fuel in storage and distribution tanks can undergo several changes, causing the volume to differ from the amount of fuel that initially leaves what was produced [2, 4]. The regulatory body for the activities that integrate the oil, natural gas and biofuels industries in Brazil, linked to the Ministry of Mines and Energy (MME), is the National Agency of Petroleum, Natural Gas and Biofuels (ANP), the federal agency that executes the national policy for the sector with a focus on ensuring fuel supply, product quality and protecting consumer interests [5].

II. THEORETICAL REFERENCE

II.1 STATISTICAL PROCESS CONTROL

The SPC is an inspection system by sampling, operating throughout the process, with the objective of verifying the presence of special causes, that is, causes that are not natural to the process and that can harm the quality of the manufactured product. Once the special causes are identified, we can act on them, continuously improving the production processes and, therefore, the quality of the final product. [6, 7]. The SPC makes it possible to monitor the quality characteristics of interest, ensuring their maintenance within pre-established limits and indicating when to adopt corrective and improvement actions. It allows the systematic reduction of variability in quality characteristics, in an effort to improve the intrinsic quality, productivity and reliability of what is being produced or supplied. Shewhart control charts stand out among the SPC tools for operational simplicity and effectiveness in detecting problems [8]. The SPC provides an x-ray of the process, identifying its variability and enabling the control of this variability over time through continuous data collection, analysis and blocking of possible special causes that are making the system unstable [7, 9]. The main objective of the SPC is to enable an effective quality control, carried out by the operator himself in real time. This increases the operator's commitment to the quality of what is being produced and frees management for improvement tasks [9]. The SPC makes it possible to monitor the characteristics of interest, ensuring that they will remain within pre-established limits and indicating when corrective and improvement actions should be taken. It is important to emphasize the importance of detecting defects as early as possible, to avoid adding raw material and labor to a defective product [7].

II.2 CONTROL GRAPHICS

The SPC is operationalized through control charts, which are used to monitor the performance of a process from the definition of an acceptable control range. The control chart is used to analyze trends and patterns that happen over time. Its main purpose is to monitor a process, checking if it is under statistical control indicating its range of variation [6]. There are two types of control charts: for variables and for attributes. The control graphs for attributes refer to the quality characteristics that classify items in conforming and non-conforming, while the control graphs for variables are based on the measurement of quality characteristics on a continuous scale, as shown in Figure 1 [6, 7, 9].

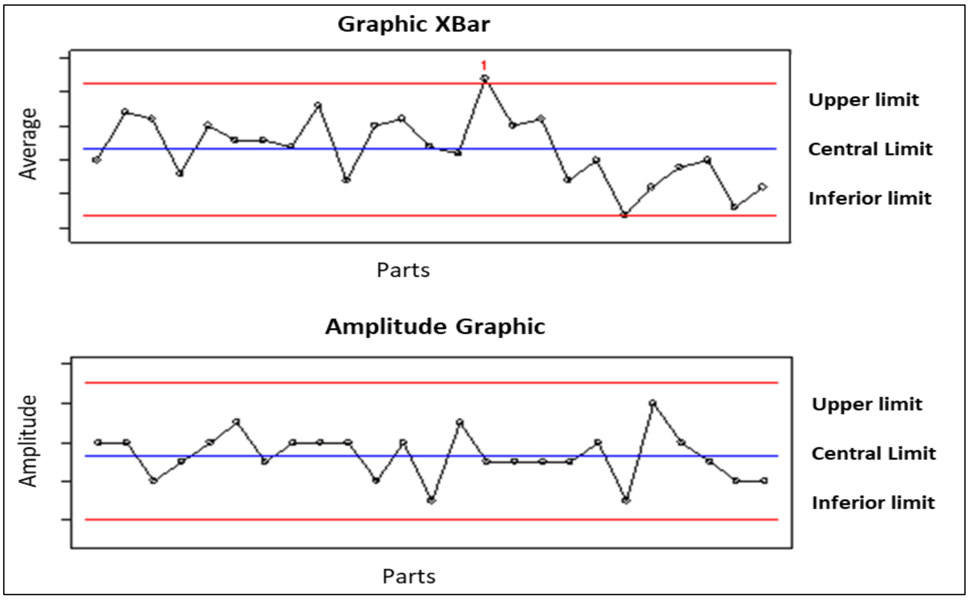


Figure 1: Control Chart Template.
Source: [17].

Figure 1 shows a control chart that is composed of a Central Line (CL) that represents the average value or central limit of the quality characteristic corresponding to the situation of the process under control and a pair of control limits: one of them located below the central line called Lower Control Limit (LCL) and another located above the central line, called Upper Control Limit (UCL) [10, 11].

II.3 CONTROL LIMITS

The control of the process is usually done through the control graph for individual values or gráfico graph. The variability of the process can be monitored either through the standard deviation control graph, called the S graph, or by the graph for the moving amplitude, called the R graph. To better exemplify the function of each graph, the x graph monitors the variability between samples and the S or R graph monitors the variability within the sample. The points plotted on the control graphs are joined by straight lines sequentially and are interpreted according to horizontal lines, called UCL, ML and LCL given by Equations 1, 2 and 3 for the mean graph and 5, 6, and 7 for the graph of the moving range [7, 12].

Control limits for the graph of individual measures.

$$UCL_x = \bar{x} + E_2.Rm \quad (1)$$

$$LM_x = \bar{x} \quad (2)$$

$$LCL_x = \bar{x} - E_2.Rm \quad (3)$$

Where:

$$E_2 = \frac{3}{d_2.} \quad (4)$$

Where $\bar{x}$ is the average of the graph points for each variable, $Rm = |x_i - x_{i-1}|$, that is, the difference between two subsequent values for each variable, and the parameters E_2 and d_2 are tabulated [7, 12].

II.4 VOLATILITY

Volatility is a quantity that is related to the easiness of the substance to pass from the liquid to the vapor or gaseous state. This facility depends on the referential; therefore, volatility is always relative: it takes into account two substances, one of which is the reference substance [10, 13].

The relative volatility between a substance A and a substance B is defined as follows:

$$\alpha_{AB} = \frac{y_{Ae}/x_{Ae}}{y_{Be}/x_{Be}} \quad (5)$$

Where: α_{AB} is the relative volatility between A and B. y_{Ae} and y_{Be} are molar fractions of A and B, respectively, in the vapor phase in equilibrium with the liquid phase. x_{Ae} and x_{Be} they are molar fractions of A and B, respectively, in the vapor phase in equilibrium with the liquid phase.

Relative volatility below one ($\alpha_{AB} < 1$) indicates that B is more volatile than A; otherwise, if the relative volatility is greater than one ($\alpha_{AB} > 1$), A is more volatile than B. If the liquid phase is an ideal mixture, Raoult's law can be accepted as valid:

$$p_A = P_A^{sat}.x_A \quad (6)$$

$$p_B = P_B^{sat}.x_B \quad (7)$$

Where: p_A and p_B, are partial pressures of A and B, respectively. P_A^{sat} and P_B^{sat} are vapor pressures of A and B, respectively. x_A and x_B are molar fractions of A and B, respectively, in the liquid phase.

If the vapor phase is an ideal gas, Dalton's Law applies:

$$y_A = \frac{p_A}{P} \quad (8)$$

$$y_B = \frac{p_B}{P} \quad (9)$$

Where: y_A and y_B are the molar fractions of A and B, respectively, in the vapor phase. P the total system pressure.

Replacing the eq. (6) in eq. (8) and eq. (7) in eq. (9), we will have:

$$y_A = \frac{P_A^{sat}.x_A}{P} \quad (10)$$

$$y_B = \frac{P_B^{sat}.x_B}{P} \quad (11)$$

Replacing the eq. (10) and eq. (11) in eq. (5), we will have:

$$\alpha_{AB} = \frac{P_A^{sat}}{P_B^{sat}} \quad (12)$$

That is, in cases of totally ideal liquid-vapor equilibrium, the relative volatility between two substances in a mixture is a simple relation of their vapor pressures [10, 13].

III. MATERIALS AND METHODS

III.1 COMPANY DESCRIPTION – PLACE OF STUDY

The work was carried out at a company in the oil sector in the city of Manaus, which has been operating since 2000 and has been consolidating its position as one of the largest distributors of petroleum fuels and biofuels in Brazil. As defined by the ANP, the company under study carries out the activity of distribution of liquid fuels that is of public utility and comprises the acquisition, storage, mixing, transportation, commercialization and quality control of fuels [14].

It recently underwent a process to expand its storage capacity, which went from 15.000 m^3 to 75.000 m^3, according to Regulatory Standard n° 20 (NR-20) of the Ministry of Labor is classified as a class III hazardous facility and carries out storage, transfer, handling and handling of combustible and flammable liquids [15].

III.2 STAGES OF WORK

The work was carried out on the changes in the volume of gasoline A in the process of transfers between tanks, as described in Figure 2.

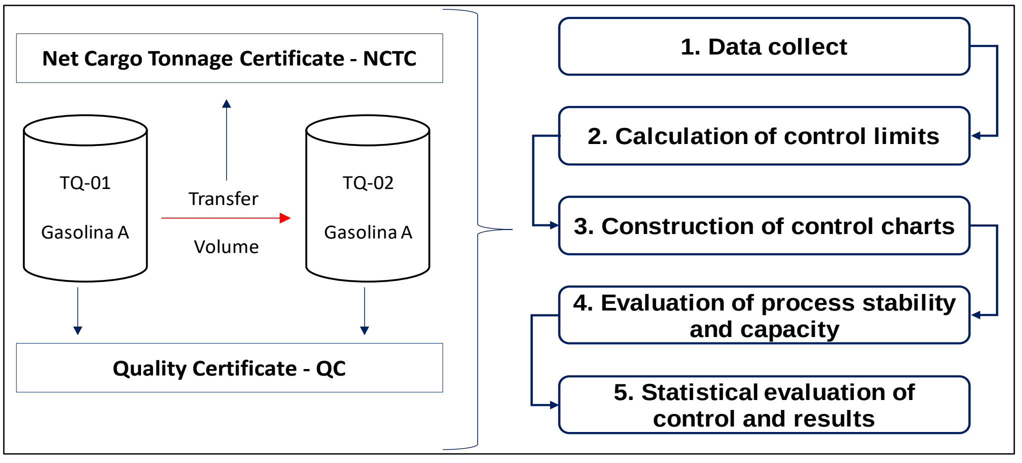


Figure 2: Fuel transfer flow and methodology flowchart.
Source: [7] and [12].

III.2.1 DATA COLLECT

Gasoline A volume data were collected in the transfer between tanks (TQ) processes. As shown in figure 2, data on volume at room temperature and volume at 20 °C from TQ-01 (shipping tank) and TQ-02 (receiving tank) were collected over a period of two months. These data were obtained from Certificates of Net Cargo Tonnage (CNCT), which is the official document for certification and quantification of the volume of fuel handled in oil terminals. The data of temperature, density and specific gravity at 20 °C were obtained from the Gasoline A Quality Certificate (QC), referring to the batch of work in the sending tank and in the receiving tank.

III.2.2 CALCULATION OF CONTROL LIMITS

The control of the process was carried out through the control chart of individual values and by control charts for mobile amplitudes. The points plotted on the control graphs were joined by straight lines sequentially that were interpreted in terms of horizontal lines, called UCL, ML and LCL that took place according to the methodology of [7] and [12].

III.2.3 CONSTRUCTION OF CONTROL GRAPHICS

The volume data at room temperature and the volume converted at 20 ° C of the transfers were organized in a spreadsheet and plotted the graphs using Microsoft Excel software from Windows. They checked how the process behaves and made an average ($\bar{x}$) and mobile range (Rm) of the data obtained to determine the UCL and the LCL to verify the expected standards of the process and define the limits according to the methodology proposed by [7], that defines the SCP for continuous and batch processes.

III.2.4 EVALUATION OF PROCESS STABILITY AND CAPACITY

The relationship between the levels of variability or the stability of the process in relation to the specification requirements, were made through the analysis of the capacity of the process. The verification of the ability of the process to safely meet the specifications was demonstrated by calculating the parameter C_{pk}, defined by Equations 13, 14 and 15. This parameter represents the ratio between the specification tolerance and the total dispersion of the process, according to the methodology of [7] and [12].

$$C_{pk} = \min[C_{pi}, C_{ps}] \quad (13)$$

$$C_{pi} = \frac{\bar{x}-LIE}{3.s} \quad (14)$$

$$C_{ps} = \frac{LSE-\bar{x}}{3.s} \quad (15)$$

Where LSL is the Lower Specification Limit and USL is the Upper Specification Limit, the s Where LSL is the Lower Specification Limit and USL is the Upper Specification Limit, the □ is the average of the samples.

The index C_{pk} evaluates the distance from the process average to the specification limits, taking the one that is less, that is, it is defined as being the lowest value between C_{pi} and C_{ps}, therefore, more critical in terms of the chances of producing out-of-specification items. If $C_{pk} \leq 1$ (unable process), $1 \leq C_{pk} \leq 1{,}33$ (acceptable process) e $C_{pk} \geq 1{,}33$ (capable process).

III.2.5 STATISTICAL EVALUATION OF CONTROL AND RESULTS

Data that exceeded the limit lines were verified and other Quality/Process engineering tools were used, such as Brainstorm and Ishikawa Diagram, which are discussions of ideas and a graphical form used as an analysis to represent influencing factors (causes) on a given problem. (effect) to find the root cause and define the process capacity indices [16].

IV. RESULTS AND DISCUSSIONS

The data of 33 transfers (samples) of Gasoline A from tank 01 to tank 02 in the months of August and September 2020 were studied. Observations of process control related to the parameters of volume at room temperature and volume at 20 °C were used. The information collected from the NCTC and product QC are described in Table 1.

Table 1: Data collected from CACL and QC for transfers of Gasoline A (33 samples - August and September 2020).

Nº.	Dispatcher Tank (TQ 01)						Receiving Tank (TQ 02)					
	Tank temp. (ºC)	**Dens. read (g/cm³)**	**Sample temp. (ºC)**	**Esp. Mass at 20 ºC (kg/m³)**	**Ambient volume (L)**	**Volume at 20 ºC (L)**	**Tank temp. (ºC)**	**Dens. read (g/cm³)**	**Sample temp. (ºC)**	**Esp. Mass at 20 ºC (kg/m³)**	**Ambient volume (L)**	**Volume at 20 ºC (L)**
1	31,5	0,712	29,5	719,9	1.241.029	1.224.236	30,5	0,717	30,0	725,2	1.243.243,0	1.228.130
2	31,0	0,718	30,0	726,2	1.250.090	1.234.216	30,0	0,717	29,0	724,4	1.254.595,0	1.240.037
3	29,0	0,721	28,5	728,0	1.070.774	1.059.712	28,0	0,718	27,5	724,2	1.073.099,0	1.063.134
4	30,0	0,715	28,0	721,6	1.935.762	1.013.110	29,0	0,720	28,0	726,6	1.931.089	1.911.055
5	30,0	0,710	28,5	717,1	997.157	985.329	29,0	0,708	27,5	714,3	995.546	984.831
6	29,0	0,706	28,5	713,1	964.288	953.874	28,0	0,704	26,5	709,5	964.287	954.930
7	29,5	0,707	28,0	713,7	1.026.635	1.014.950	28,5	0,712	28,0	718,7	1.028.494	1.018.176
8	28,0	0,705	27,5	711,3	330.876	327.683	27,0	0,702	26,0	707,1	330.652	327.825
9	28,0	0,714	26,0	719,0	912.786	904.177	27,0	0,710	26,5	715,4	914.522	906.896
10	28,5	0,712	27,5	720,2	776.243	768.446	27,5	0,717	26,5	724,4	776.845	770.046
11	29,0	0,717	28,5	724,0	1.406.613	1.391.908	28,0	0,720	26,5	725,3	1.409.735	1.396.690
12	29,0	0,717	27,0	722,8	1.032.362	1.021.530	28,0	0,722	27,0	727,7	1.033.674	1.024.177
13	28,0	0,716	26,5	721,4	219.162	217.110	27,0	0,715	26,5	720,4	218.192	216.399
14	28,5	0,709	28,0	715,7	908.299	899.106	27,5	0,706	26,5	711,5	908.623	900.407
15	27,5	0,711	26,0	716,0	832.622	825.195	26,5	0,716	25,0	720,1	833.066	826.706
16	29,5	0,703	29,0	710,6	1.575.481	1.557.381	28,5	0,701	27,5	707,4	1.579.472	1.563.082
17	31,5	0,705	29,5	713,0	512.228	505.152	30,5	0,703	29,5	711,0	513.398	506.886
18	31,0	0,710	30,0	718,3	1.153.299	1.138.303	30,0	0,715	28,5	722,0	1.155.210	1.141.709
19	29,0	0,713	28,5	720,0	1.200.225	1.187.528	28,0	0,710	27,0	715,8	1.203.303	1.191.847
20	30,0	0,719	28,0	725,6	1.098.599	1.085.895	29,0	0,715	28,0	721,6	1.098.678	1.087.110
21	30,0	0,720	28,5	727,0	414.281	409.510	29,0	0,725	27,5	731,1	412.711	408.487
22	29,0	0,721	28,5	728,0	913.011	902.528	28,0	0,724	27,0	729,7	914.090	905.741
23	29,5	0,720	28,0	726,6	1.218.226	1.206.291	28,5	0,725	27,5	731,1	1.220.256	1.208.463
24	28,0	0,717	27,5	723,2	1.017.680	1.008.202	27,0	0,716	25,5	720,6	1.017.412	1.009.057
25	28,0	0,710	26,0	715,0	1.059.505	1.049.393	27,0	0,707	28,5	714,1	1.062.680	1.053.784
26	28,5	0,707	27,5	713,3	476.090	471.237	27,5	0,712	27,0	717,8	476.378	472.152
27	29,0	0,706	28,5	713,1	454.002	449.099	28,0	0,704	27,0	709,9	455.905	451.487
28	29,0	0,707	27,0	712,9	928.658	918.621	28,0	0,705	27,5	711,3	927.966	919.011
29	28,0	0,705	26,5	710,5	561.877	556.441	27,0	0,710	26,0	715,0	562.394	557.698
30	28,5	0,714	28,0	720,6	1.769.183	1.751.539	27,5	0,711	26,5	716,4	1.773.054	1.757.258
31	27,5	0,712	26,0	717,0	1.112.712	1.102.816	26,5	0,708	25,5	712,6	1.112.855	1.104.166
32	29,5	0,711	29,0	718,5	1.039.391	1.027.729	28,5	0,716	27,5	722,2	1.042.358	1.032.011
33	28,0	0,717	26,0	722,0	1.028.154	1.018.543	27,0	0,720	26,0	724,9	1.026.245	1.017.927

Source: Authors, (2020).

IV.1 ASSESSMENT OF AMBIENT VOLUME VARIATION

Figure 3 shows the graph of the values collected from the variation in the ambient volume of Gasoline A for 33 transfers, it is observed that the process remains within the specification limits, which are established NBR 13787 and ANP Resolution N° 23/2004 [18, 19], which are -0,6 and +0,6% of volume variation for LSL and USL, respectively.

Figure 4 shows the control chart of the individual values for the Ambient Volume parameter for the 33 lots. It is observed that the process remains within the control limits, which were calculated for the process, which are -0,34 and +0,52% for the LCL and UCL, respectively, with a mean (ML) around +0,09% for the control chart of individual values (*x*).

The values of specification limit and control limit for individual values, evaluating the parameter of ambient volume were organized in a spreadsheet and are shown in Table 2.

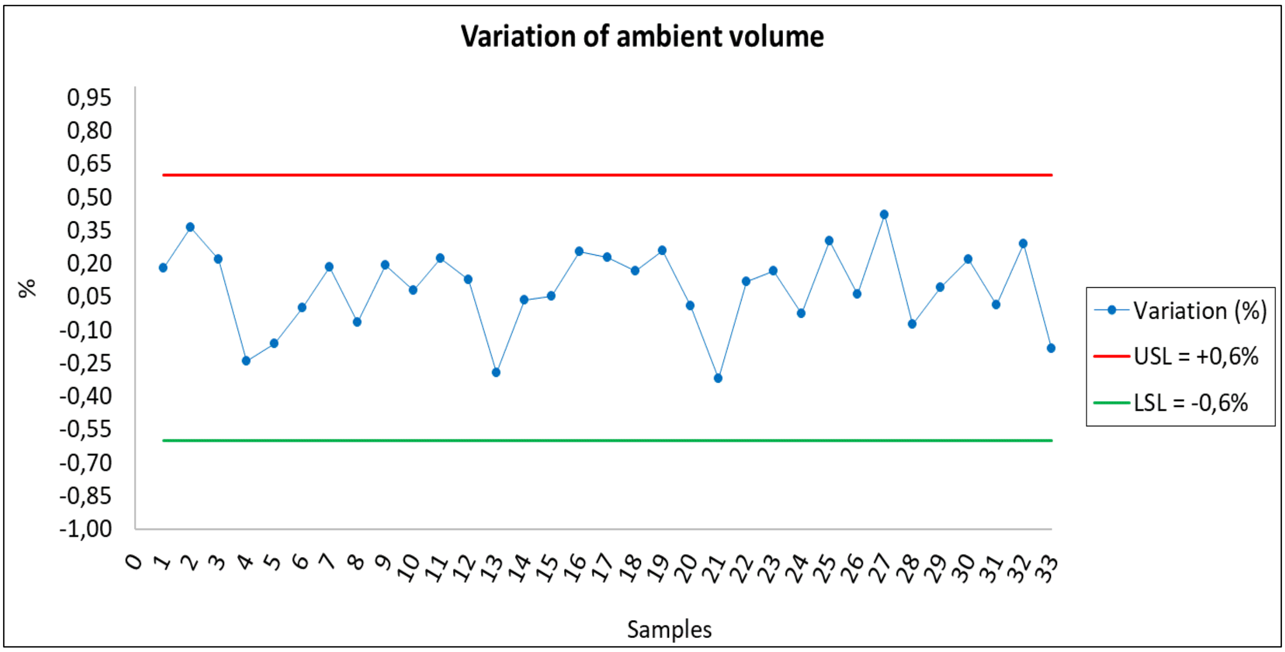


Figure 3: Graph of values and specification limits for Ambient Volume.
Source: Authors, (2020).

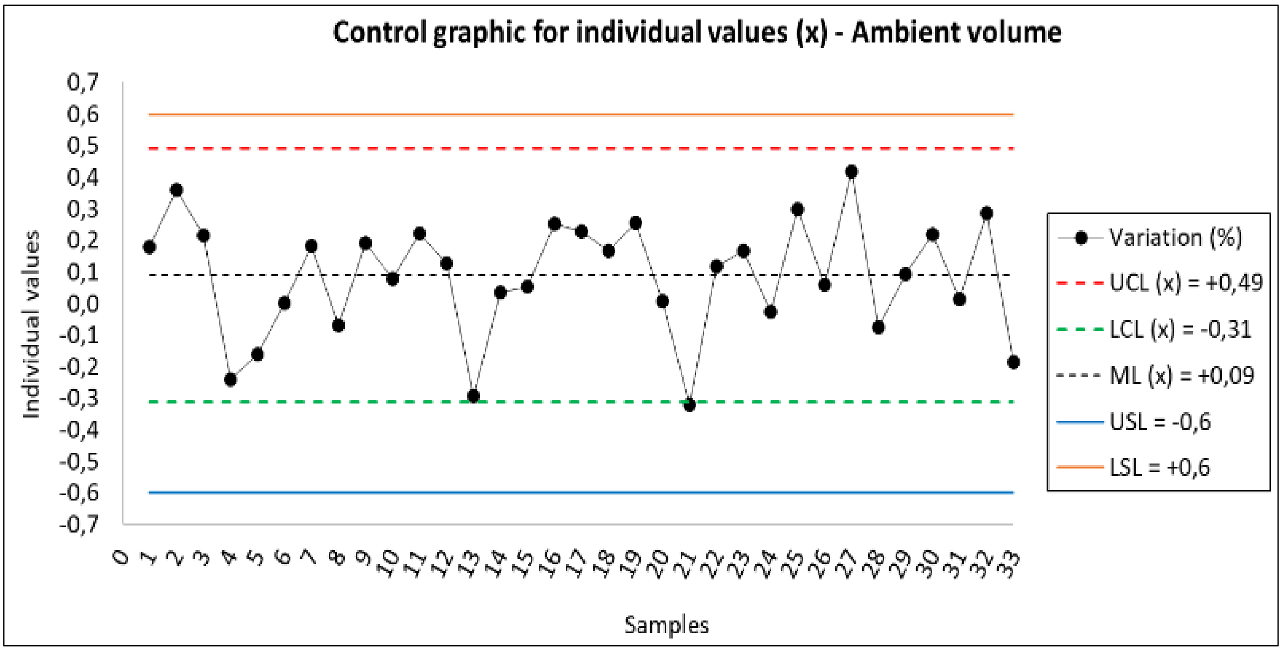


Figure 4: Control chart of individual values (*x*) - Ambient volume.
Source: Authors, (2020).

Table 2: Volume variation (%) in transfers between tanks (33 samples - August and September 2020).

Nº.	Ambient Volume (L)				Volume at 20 ºC (L)			
	Dispatched (V_{TQ1})	Received (V_{TQ2})	Difference (V_{TQ2} - V_{TQ1})	Variation ambiente volume (%)	Dispatched (V_{TQ1})	Received (V_{TQ2})	Difference (V_{TQ2} - V_{TQ1})	Variation volume at 20 ºC (%)
1	1.241.029	1.243.243	2.214	0,18	1.224.236	1.228.130	3.894	0,32
2	1.250.090	1.254.595	4.505	0,36	1.234.216	1.240.037	5.821	0,47
3	1.070.774	1.073.099	2.325	0,22	1.059.712	1.063.134	3.422	0,32
4	1.935.762	1.931.089	-4.673	-0,24	1.913.110	1.911.055	-2.055	-0,11
5	997.157	995.546	-1.611	-0,16	985.329	984.831	-498	-0,05
6	964.288	964.287	-1	0,00	953.874	954.930	1.056	0,11
7	1.026.635	1.028.494	1.859	0,18	1.014.950	1.018.176	3.226	0,32
8	330.876	330.652	-224	-0,07	327.683	327.825	142	0,04
9	912.786	914.522	1.736	0,19	904.177	906.896	2.719	0,30
10	776.243	776.845	602	0,08	768.446	770.046	1.600	0,21
11	1.406.613	1.409.735	3.122	0,22	1.391.908	1.396.690	4.782	0,34
12	1.032.362	1.033.674	1.312	0,13	1.021.530	1.024.177	2.647	0,26
13	219.162	218.192	-970	-0,29	217.110	216.399	-711	-0,17
14	908.299	908.623	324	0,04	899.106	900.407	1.301	0,14
15	832.622	833.066	444	0,05	825.195	826.706	1.511	0,18
16	1.575.481	1.579.472	3.991	0,25	1.557.381	1.563.082	5.701	0,37
17	512.228	513.398	1.170	0,23	505.152	506.886	1.734	0,34
18	1.153.299	1.155.210	1.911	0,17	1.138.303	1.141.709	3.406	0,30
19	1.200.225	1.203.303	3.078	0,26	1.187.528	1.191.847	4.319	0,36
20	1.098.599	1.098.678	79	0,01	1.085.895	1.087.110	1.215	0,11
21	414.281	412.711	-1.570	-0,32	409.510	408.487	-1.023	-0,19
22	913.011	914.090	1.079	0,12	902.528	905.741	3.213	0,36
23	1.218.226	1.220.256	2.030	0,17	1.206.291	1.208.463	2.172	0,18
24	1.017.680	1.017.412	-268	-0,03	1.008.202	1.009.057	855	0,08
25	1.059.505	1.062.680	3.175	0,30	1.049.393	1.053.784	4.391	0,42
26	476.090	476.378	288	0,06	471.237	472.152	915	0,19
27	454.002	455.905	1.903	0,42	449.099	451.487	2.388	0,53
28	928.658	927.966	-692	-0,07	918.621	919.011	390	0,04
29	561.877	562.394	517	0,09	556.441	557.698	1.257	0,23
30	1.769.183	1.773.054	3.871	0,22	1.751.539	1.757.258	5.719	0,33
31	1.112.712	1.112.855	143	0,01	1.102.816	1.104.166	1.350	0,12
32	1.039.391	1.042.358	2.967	0,29	1.027.729	1.032.011	4.282	0,42
33	1.028.154	1.026.245	-1.909	-0,19	1.018.543	1.017.927	-616	-0,06

Source: Authors, (2020).

IV.2 ASSESSMENT OF VOLUME VARIATION AT 20 °C

Figure 5 shows the graph of the values collected from the volume variation at 20 °C of Gasoline A for 33 transfers, it is observed that the process remains within the specification limits, which are established by NBR 13787 and resolution ANP n° 23/2004 [18, 19], which are -0,6 and +0,6% of volume variation for LSL and USL, respectively.

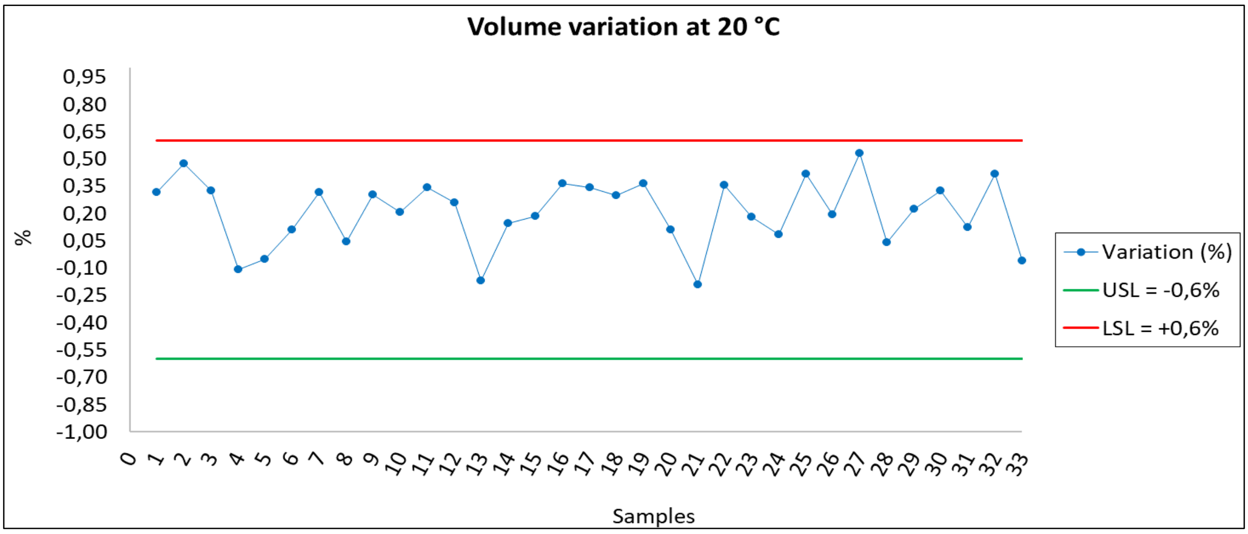


Figure 5: Graph of values and specification limits for Volume at 20 °C.
Source: Author (2020).

Figure 6 shows the control chart of the individual values for the Volume parameter at 20 °C for the 33 lots. It is observed that the process remains within the control limits, which were calculated for the process, which are -0,19 and +0,61% for LCL and UCL, respectively, with an average (LM) around +0,21% for the control chart of individual values (x).

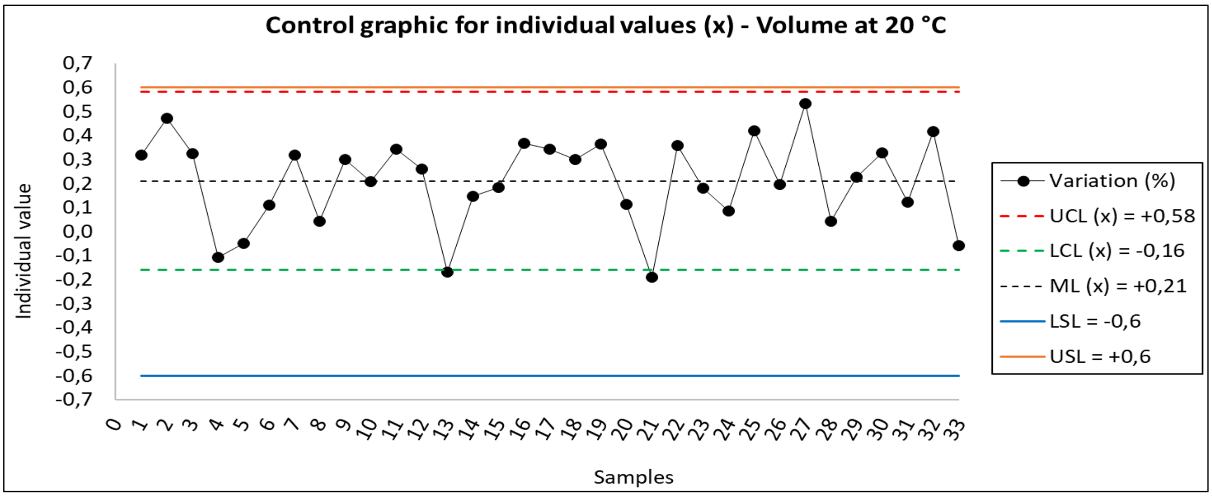


Figura 6: Control chart of individual values (*x*) -Volume at 20 °C.
Source: Authors, (2020).

The values of specification limit and control limit for individual values, evaluating the volume parameter at 20 °C, were organized in a spreadsheet and are shown in Table 2.

IV.3 PROCESS CAPABILITY ASSESSMENT

After proving the stability of the process, the process capacity index was calculated as described in Table 3.

Table 3: Control limits and capacity study.

Parameter	Ambient volume	Volume at 20 °C
Especification limits	-0,6 a + 0,6%	-0,6 a + 0,6%
Control limits (x)	-0,31 a + 0,49%	-0,16 a + 0,58%
Standard deviation	0,184	0,184
Cpi	1,245	1,463
Cps	0,931	0,712
Cpk = mín[Cpi,Cps]	**0,931**	**0,712**
Conclusion	Inability process	Inability process

Source: Authors, (2020).

For the Ambient Volume variable, the value of C_{pk} = 0,931, considering the standard deviation calculated for the 0,184. Having resulted $C_{pk} \leq 1{,}33$ the process is qualified as Incapable.

For the variable Volume at 20 °C, the value of C_{pk} = 0,712, considering the standard deviation calculated for the 0,184. Having resulted $C_{pk} \leq 1{,}33$ the process is qualified as Incapable.

IV.4 ACTION PLAN

Since, according to the result of the Process Capacity Index (C_{pk}), the process was considered Incapable, in this condition a brainstorm (sharing of ideas) was carried out in which everyone involved in the fuel transfer process (liquid bulk operator, operations supervisor and research student) exposed the ideas for surveying the causes that resulted in the undesired effect and from this survey, a Cause-Effect Diagram or Ishikawa Diagram was built, as shown in Figure 7.

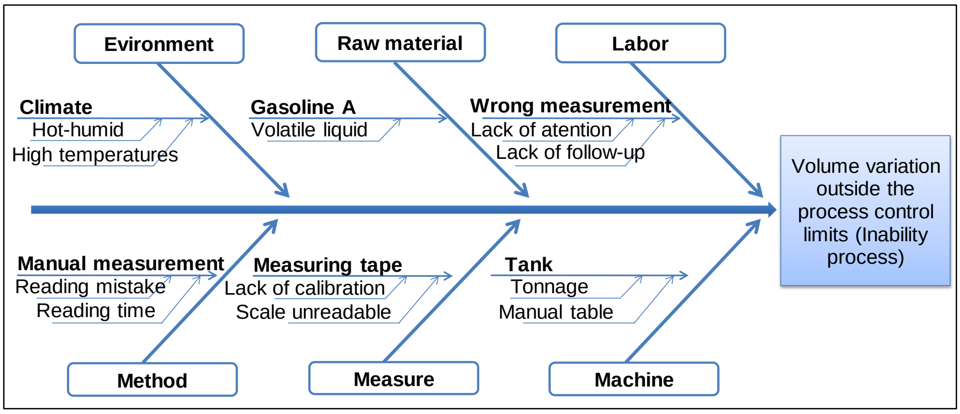


Figure 7: Ishikawa Diagram (cause-efectt) for volume variation.
Source: Authors, (2020).

From the Cause-Effect Diagram, a check-list was elaborated in which the variables and causes that led to the unwanted effect were listed. For each cause discussed for the process, it was verified whether or not the company met the requirement to inhibit the effect, and by elimination, work on the root cause.

As described in the check-list, the root cause of the unwanted effect for the process was the labor variable, which is caused by the wrong measurement that originated from the lack of attention during measurements and the lack of monitoring of a third party regarding the process carried out.

Once the root cause was verified, the data of 15 more transfers of Gasoline A from tank 01 to tank 02 were studied in October 2020 to review the fuel transfer process. The activity was monitored by an operations supervisor and the student researcher and carried out by the liquid bulk operator.

Observations of process control related to the parameters of volume at room temperature and volume at 20 °C were used. The information collected from the product quality certificates received and ANP specifications are described in Table 4 and Table 5**.**

Table 4: Data collected from CACL and QC for transfers of Gasoline A (15 samples - October 2020 - process review).

Nº.	Dispatcher Tank (TQ 01)						Receiving Tank (TQ 02)					
	Tank temp. (ºC)	**Dens. read (g/cm³)**	**Sample temp. (ºC)**	**Esp. Mass at 20 ºC (kg/m³)**	**Ambient volume (L)**	**Volume at 20 ºC (L)**	**Tank temp. (ºC)**	**Dens. read (g/cm³)**	**Sample temp. (ºC)**	**Esp. Mass at 20 ºC (kg/m³)**	**Ambient volume (L)**	**Volume at 20 ºC (L)**
1	30,0	0,719	28,0	725,6	1.098.599	1.085.895	29,0	0,715	28,0	721,6	1.098.678	1.087.110
2	30,0	0,710	28,5	717,1	997.157	985.329	29,0	0,708	27,5	714,3	995.546	984.831
3	29,0	0,721	28,5	728,0	913.011	902.528	28,0	0,724	27,0	729,7	914.090	905.741
4	27,5	0,711	26,0	716,0	832.622	825.195	26,5	0,716	25,0	720,1	833.066	826.706
5	28,0	0,717	27,5	723,2	1.017.680	1.008.202	27,0	0,716	25,5	720,6	1.017.412	1.009.057
6	28,0	0,705	26,5	710,5	561.877	556.441	27,0	0,710	26,0	715,0	562.394	557.698
7	28,5	0,709	28,0	715,7	908.299	899.106	27,5	0,706	26,5	711,5	908.623	900.407
8	28,5	0,707	27,5	713,3	476.090	471.237	27,5	0,712	27,0	717,8	476.378	472.152
9	29,0	0,717	27,0	722,8	1.032.362	1.021.530	28,0	0,722	27,0	727,7	1.033.674	1.024.177
10	29,0	0,707	27,0	712,9	928.658	918.621	28,0	0,705	27,5	711,3	927.966	919.011
11	29,0	0,706	28,5	713,1	964.288	953.874	28,0	0,704	26,5	709,5	964.287	954.930
12	28,0	0,705	27,5	711,3	330.876	327.683	27,0	0,702	26,0	707,1	330.652	327.825
13	27,5	0,712	26,0	717,0	1.112.712	1.102.816	26,5	0,708	25,5	712,6	1.112.855	1.104.166
14	28,5	0,712	27,5	720,2	776.243	768.446	27,5	0,717	26,5	724,4	776.845	770.046
15	29,5	0,720	28,0	726,6	1.218.226	1.206.291	28,5	0,725	27,5	731,1	1.220.256	1.208.463

Source: Authors, (2020).

Table 5: Volume variation (%) in transfers between tanks (15 samples - October 2020 - process review).

Nº.	Ambient Volume (L)				Volume at 20 ºC (L)			
	Dispatched (V_{TQ1})	**Received (V_{TQ2})**	**Difference ($V_{TQ2} - V_{TQ1}$)**	**Variation ambiente volume (%)**	**Dispatched (V_{TQ1})**	**Received (V_{TQ2})**	**Difference ($V_{TQ2} - V_{TQ1}$)**	**Variation volume at 20 ºC (%)**
1	1.098.599	1.098.678	79	0,01	1.085.895	1.087.110	1.215	0,11
2	997.157	995.546	-1.611	-0,16	985.329	984.831	-498	-0,05
3	913.011	914.090	1.079	0,12,	902.528	905.741	3.213	0,36
4	832.622	833.066	444	0,05	825.195	826.706	1.511	0,18
5	1.017.680	1.017.412	-268	-0,03	1.008.202	1.009.057	855	0,08
6	561.877	562.394	517	0,09	556.441	557.698	1.257	0,23
7	908.299	908.623	324	0,04	899.106	900.407	1.301	0,14
8	476.090	476.378	288	0,06	471.237	472.152	915	0,19
9	1.032.362	1.033.974	1.312	0,13	1.021.530	1.024.177	2.647	0,26
10	928.658	927.966	-692	-0,07	918.621	919.011	390	0,04
11	964.288	964.287	-1	0,00	953.874	954.930	1.056	0,11
12	330.876	330.652	-224	-0,07	327.683	327.825	142	0,04
13	1.112.712	1.112.855	143	0,01	1.102.816	1.104.166	1.350	0,12
14	776.243	776.845	602	0,08	768.446	770.046	1.600	0,21
15	1.218.226	1.220.256	2.030	0,17	1.206.291	1.208.463	2.172	0,18

Source: Authors, (2020).

IV.4.1 Assessment of Ambient Volume Variation – Process Review

Figure 8 shows the graph of the values collected from the variation in the ambient volume of Gasoline A for 15 transfers, it is observed that the process remains within the specification limits, which are established by NBR 13787 and ANP resolution n° 23/2004 [18, 19], which are -0,6 and +0,6% of the volume variation for the USL and LSL, respectively.

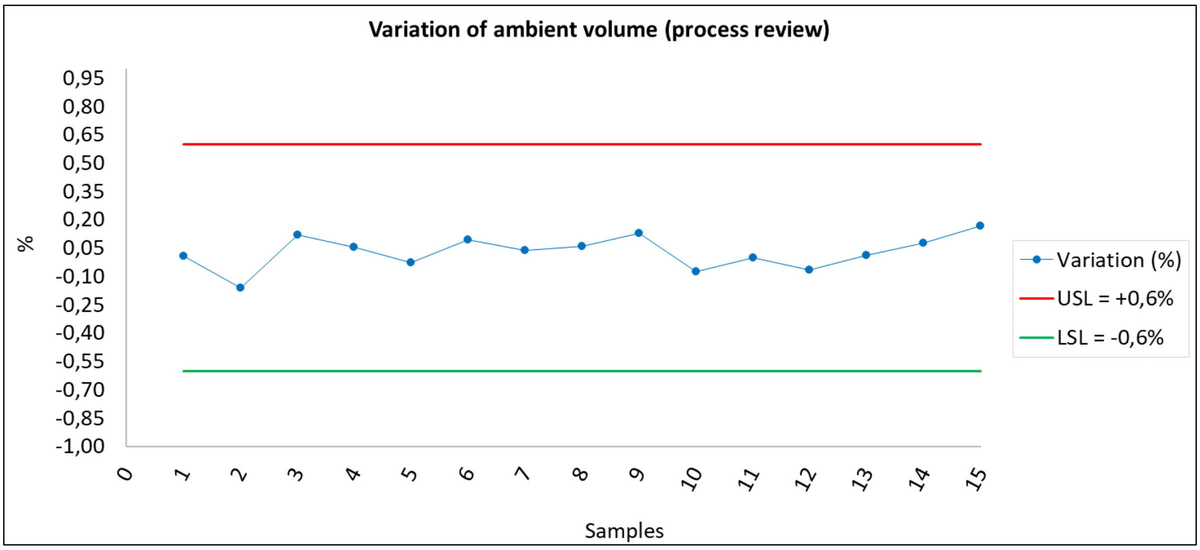


Figure 8: Graph of values and specification limits for Ambient Volume - process review.
Source: Author (2020).

Figure 9 shows the control chart of the individual values for the parameter of the Ambient Volume for the 15 samples. It is observed that the process remains within the control limits, which were calculated for the process, which are -0,26 and +0,32% for the LCL and UCL, respectively, with an average (ML) around +0,03% for the control chart of individual values (x).

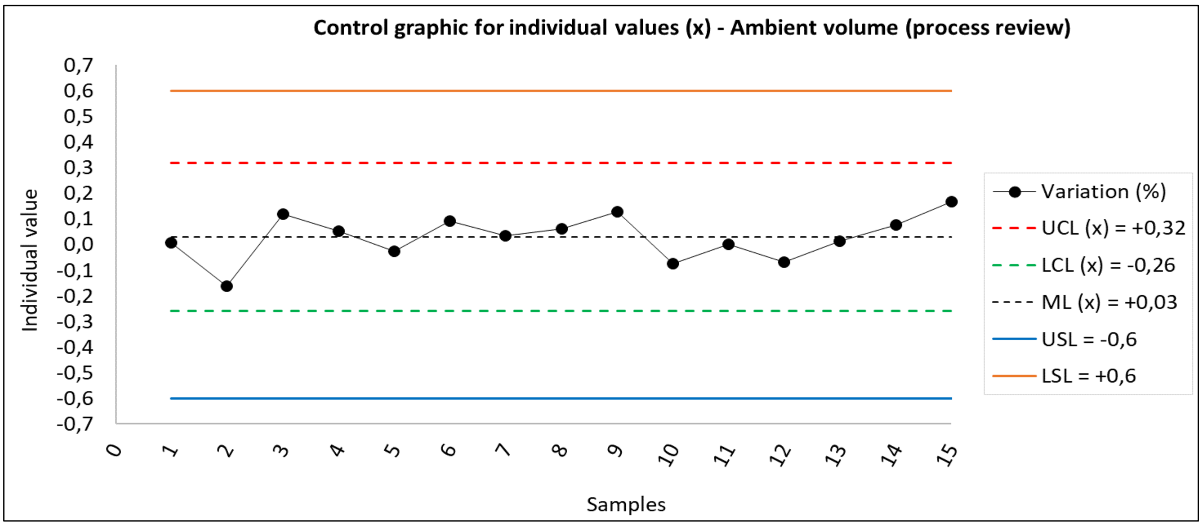


Figure 9: Control chart of individual values (x) - Ambient volume (process review).
Source: Authors, (2020).

The values of specification limit and control limit for individual values, evaluating the parameter of ambient volume for review of the process were organized in a spreadsheet and are presented in Table 5.

IV.4.2 Assessment of Volume Variation at 20 °C – Process Review

Figure 10 shows the graph of the values collected from the volume variation at 20 °C of Gasoline A for 15 transfers, it is observed that the process remains within the specification limits, which are established by NBR 13787 and resolution ANP n° 23/2004 [18, 19], which are -0,6 and +0,6% of volume variation for LSL and USL, respectively.

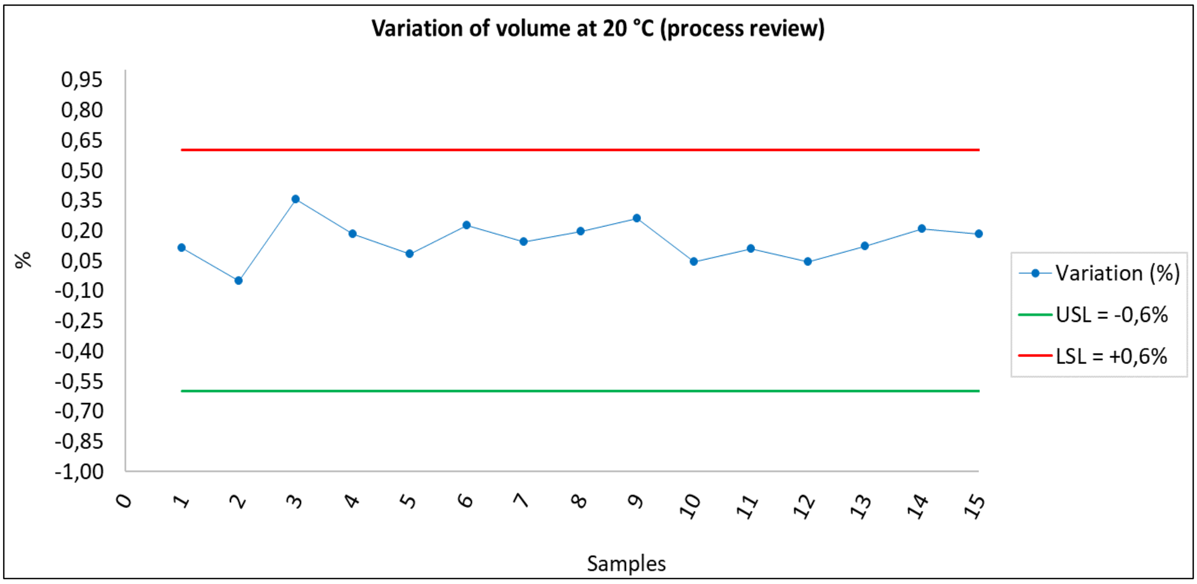


Figure 10: Graph of values and specification limits for Volume at 20 °C - process review.
Source: Authors, (2020).

Figure 11 shows the control chart of the individual values for the Volume parameter at 20 °C for the 15 lots. In both figures, it is observed that the process remains within the control limits, which were calculated for the process, which are -0,12 and +0,42% for LCL and UCL, respectively, with an average (ML) around + 0,15% for the control chart of individual values (x).

The specification limit and control limit values for the individual values, evaluating the volume parameter at 20 °C, were organized in a spreadsheet and are presented in Table 5.

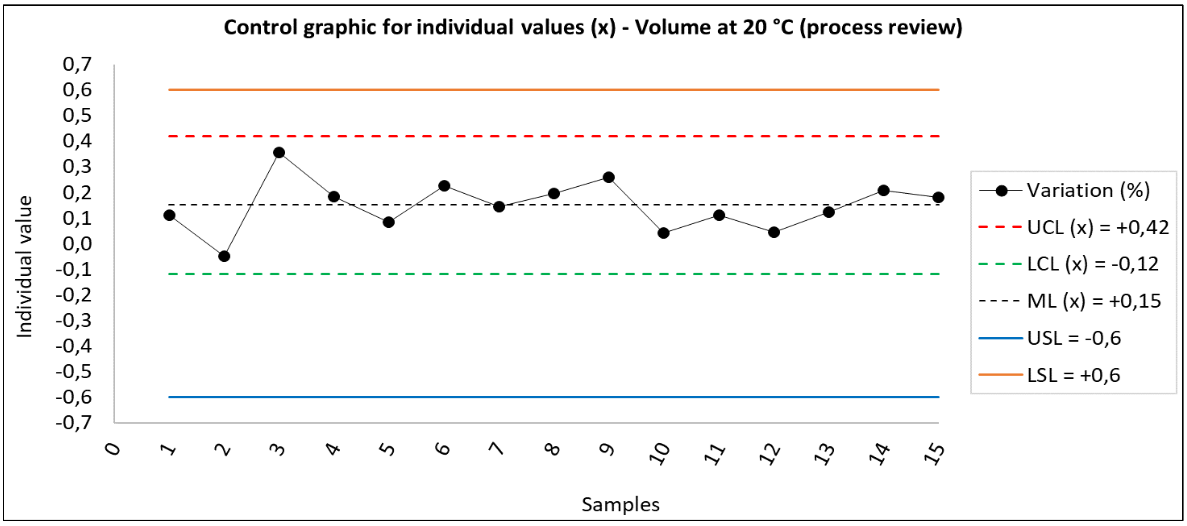


Figura 11: Control chart of individual values (x) - Volume at 20 °C - process review.
Source: Authors, (2020).

IV.4.3 Process Capability Assessment – Process Review

After proving the stability of the process, the process capacity index was calculated as described in table 6.

Table 6: Control limits and capacity study (process review).

Parameter	Ambient volume	Volume at 20 °C
Especification limits	-0,6 a + 0,6%	-0,6 a + 0,6%
Control limits (x)	-0,34 a + 0,52%	-0,2 a + 0,6%
Standard deviation	0,087	0,099
Cpi	2,412	2,505
Cps	2,196	1,515
Cpk = mín[Cpi,Cps]	**2,196**	**1,515**
Conclusion	Capable process	Capable process

Source: Authors, (2020).

For the Ambient Volume variable, the value of C_{pk} = 2,191, considering the standard deviation calculated for the 0,087. Having resulted $C_{pk} \geq 1,33$ the process is qualified as Capable.

For the variable Volume at 20 °C, the value of C_{pk} = 1,515, considering the standard deviation calculated for the 0,099. Having resulted $C_{pk} \geq 1,33$ the process is qualified as Capable.

V. CONCLUSIONS

In this work, Statistical Process Control (SPC) was used to demonstrate the performance of changes in the volume of gasoline A during the transfer process between tanks at a fuel distributor in the city of Manaus / AM and to establish a study for the use of this tool control.

An analysis of the graphs described in Figures 8, 9, 10 and 11 and Table 6, show that the process is capable of meeting all control specifications from a review of the process that was followed up in August and September of 2020. All the process control parameters studied, ambient volume and volume at 20 °C, remained within the control limits and as for the capacity indexes, it was concluded that the process produced high capacity indexes.

The SCP used here enabled the use of information accumulated in historical receipt data that had not been used for other purposes, allowing the knowledge of the levels of variation produced by the process, which can be a starting point for implementing a process of continuous improvement.

The use of SCP is also a systematic of continuous process validation, since it can be managed continuously. The use of SCP enabled the unveiling of unnoticed instabilities by simply comparing the targets to specifications and an opportunity to trigger continuous improvement actions.

VI. AUTHOR'S CONTRIBUTION

Conceptualization: Everaldo de Queiroz Lima and Fátima Geisa Mendes Teixeira.
Methodology: Everaldo de Queiroz Lima and Fátima Geisa Mendes Teixeira.
Investigation: Everaldo de Queiroz Lima.
Discussion of results: Everaldo de Queiroz Lima and Cecília Lenzi.
Writing – Original Draft: Everaldo de Queiroz Lima.
Writing – Review and Editing: Everaldo de Queiroz Lima and Cecília Lenzi.
Resources: Everaldo de Queiroz Lima.
Supervision: Cecília Lenzi and Maurício Moyses Machado.
Approval of the final text: Everaldo de Queiroz Lima and Cecília Lenzi.

VII. ACKNOWLEDGMENTS

MSc. Cecília Lenzi for guidance, to the Lutheran University Center of Manaus (CEULM-ULBRA) and Coordination of the Graduation in Chemical Engineering for the availability of the course, structure and management.

VIII. REFERENCES

[1] Costa, A. F. B.; Epprecht, E.; Carpinetti, L. C. R. "Controle Estatístico de Qualidade". Editora Atlas, 1° ed., São Paulo, 2004.

[2] Royalfic. "Fatores responsáveis por perdas de combustíveis". Site Royal FIC, 2018. Available: https://www.royalfic.com.br/conheca-7-fatores-responsaveis-por-perdas-de-combustivel-nos-postos/ [Accessed July 16, 2020].

[3] Diniz, A. F.; Schoendorfer, C. L.; Siqueira, L. C. G. "Estimated emission of volatile organic compounds from a fueling station and its impact on the environment". Anais, XIV Safety, Health and Environment Wolrd Congress, Cubatão, 2014.

[4] BRASILPOSTOS. "Perdas de combustíveis que geram prejuízos". Site da Brasil Postos, 2018. Available: https://www.brasilpostos.com.br/noticias/equipamentos/medidor-volumetrico-de-combustiveis/8-tipos-de-perdas-de-combustiveis-que-geram-prejuizo-na-revenda/ [Accessed July 16, 2020].

[5] BRASIL. "Institucional – Atribuições da ANP". Agência Nacional de Petróleo, Gás Natural e Biocombustíveis, 2020. Available: http://www.anp.gov.br/institucional [Accessed July 17, 2020].

[6] Lima, A. A. N.; Lima, J. R.; Silva, J. L.; Alencar, J. R. B.; Soares, J. L.; Lima, L. G.; Rolim, P. J. "Application of statistical process control in the pharmaceutical industry". Revista Ciências Farmacêuticas Básica Aplicada, v. 27, n.3, p. 177-187, 2006.

[7] Ramos, A. W. "CEP para processos contínuos e em bateladas". Fundação Vanzolini, ed. Edgar Blucher, 1º ed., São Paulo, 2012.

[8] Henning, E., Walter, O. M. C. F., Souza, N. S., Samohyl, R. W. "A study for the application of statistical process control charts in water quality indicators". Rev. Sistema e Gestão, pag. 2-13, vol. 09, n01, 2014.

[9] Ribeiro, J. L. D. E Caten, C. S. T. "Controle Estatístico do Processo". FEENG – UFRGS, Porto Alegre, 2012.

[10] Maciel, J. C. S. L. "Study of the liquid-vapor balance of the water + ethanol + ionic system aiming at the separation of anhydrous alcohol". Dissertação, 128 p. Engenharia Química, Universidade Federal do Rio Grande do Norte, Natal, 2012.

[11] Louzada, F.; Diniz, C.; Ferreira, P.; Ferreira, E. "Controle Estatístico de Processo: Uma abordagem prática para engenharia". Editora LTC, 1° ed., Rio de Janeiro, 2013.

[12] Montgomery, D. C. "Introdução ao controle estatístico da qualidade". 4° ed., John Wiley & Sons, Inc., 2001.

[13] Fernandes, F. A. N.; Pizzo, A. M.; Júnior, D. M. "Termodinâmica Química". Editora UFC, 1° ed., Fortaleza, 2006.

[14] BRASIL. "Resolução ANP n. 58/2014 – Regulamentação de atividade de distribuição de combustíveis no Brasil". Agência Nacional de Petróleo, Gás Natural e Biocombustíveis, DOU 20/10/2014. Available: http://legislacao.anp.gov.br/?path=legislacao-anp/resol-anp/2014/outubro&item=ranp-58-2014 [Acessed June 15, 2020].

[15] BRASIL. "Norma Regulamentadora n° 20 – Saúde e Segurança no Trabalho com Inflamáveis e Combustíveis. Ministério do Trabalho, DOU 06/07/1978". Available: https://enit.trabalho.gov.br/portal/images/Arquivos_SST/SST_NR/NR-20-atualizada-2019.pdf [Acessed June 7, 2020].

[16] Follmer, E. "Implementation of statistical process control in one step of the wire galvanizing process: case study in the acid pickling process". Trabalho de Conclusão de Curso, 29 p., Engenharia de Produção, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2013.

[17] PORTAL ACTION. "Controle Estatístico de Processo, 2020". Available: http://www.portalaction.com.br/controle-estatistico-do-processo [Acessed May 12, 2020].

[18] BRASIL. "NBR 13787 – Armazenamento de Combustíveis Líquidos e Inflamáveis". 2° ed. Rio de Janeiro, 2013. Available: https://www.docsity.com/pt/nbr-13787-controle-de-estoque-dos-sistemas-de-armazenamento-subterraneo-de-combustiveis-s/4782172/ [Acessed October 15, 2020].

[19] BRASIL. "Resolução ANP n. 23/2004 – Regulamentação de Estoques de Combustíveis". Agência Nacional de Petróleo, Gás Natural e Biocombustíveis, DOU 20/10/2014. Available: http://legislacao.anp.gov.br/?path=legislacao-anp/resol-anp/2004/novembro&item=ranp-23--2004 [Acessed October 15, 2020].

Journal of Engineering and Technology for Industrial Applications

ISSN ONLINE: 2447-0228

ITEGAM-JETIA

Manaus, v.7 n.29, p. 36-46. May/Jun, 2021

DOI: https://doi.org/10.5935/jetia.v7i29.750

RESEARCH ARTICLE OPEN ACCESS

A SYMBOLIC ATTRIBUTE-BASED ACCESS CONTROL MODEL FOR DATA SECURITY IN THE CLOUD

Iyabo Felicia Oyeyinka*[1], Sunday Idowu[2] and Afolashade Kuyoro[3]

[1] Centre for Information Technology Management, Yaba College of Technology, Yaba, Lagos State, Nigeria.
[2, 3] Computer Science Department, School of Computing and Engineering Sciences, Babcock University, Ilishan-Remo, Ogun State, Nigeria.

[1] http://orcid.org/0000-0001-5826-0575, [2] http://orcid.org/0000-0002-0013-0265, [3] http://orcid.org/0000-0001-7235-7744

E-mail: *oyeyinka0467@pg.babcock.edu.ng *(kemi.oyeyinka@yabatech.edu.ng), idowus@babcock.edu.ng, kuyoros@babcock.edu.ng

ARTICLE INFO

Article History
Received: May 05th, 2021
Accepted: June 21th, 2021
Published: June 30th, 2021

Keywords:
Access control,
Authentication,
Cloud Computing,
Data Security,
Encryption.

ABSTRACT

There have been several attempts made in literature to develop access control techniques to stem data security problems. Many of these techniques had been found to have one deficiency or other. Hence, this study developed a Symbolic Attribute-Based Access Control (SABAC) system for data security in the cloud service environment. SABAC system was implemented by developing Hash-tag Symbol Authentication (HSA) algorithm using the Message Digest-5 encryption. SABAC utilizes a 3-Tier continuous authentication method by combining the use of username and password, HSA code, and real-time image monitoring and verification. HSA code is generated by combining 5-tuple user attributes and the string generated from the user's image using Obfuscation Technique. The concatenated string is converted to hexadecimal which serves as input to MD5 to produces a unique HSA code. SABAC was evaluated using three major security metrics of confidentiality, integrity, and avail-ability. The result of security metrics tests showed a confidence level of 99.993%, integrity threshold of 99.998%, and availability throughput of 150 users/second. This implies that SABAC is highly efficient for cloud data security. It shows that hackers would find it impossible to match any fake identity with valid HSA in the database. The study concluded that SABAC could be used for access control in a cloud environment for assuring data security. It was recommended that the SABAC system should be adopted by Cloud Solution Providers and Security Specialists.

I. INTRODUCTION

In the conventional form of computing, the cost of setting up computer services is very high, including infrastructure maintenance, technical training, software licensing, and upgrading costs. The way information communication technology systems are managed has changed with improved Technology. This has resulted in a dramatic reduction in the cost of computing services and operation. Furthermore, while the cost of running IT services and many other related IT operational activities are reduced, the fear of data theft, impersonation, and eavesdropping, as well as other forms of cybercrime, remains a common drawback [1]. Security of data and resource has continuously been a major challenge of computing services. Cybercrime is a serious threat to computing technology, people, organizations, and nations, according to a survey published by [2], and the situation is compounded by data loss. According to [3] in the Verizon Business Data Bridge Investigation report, cyber vulnerability and attacks accounted for 43% of breaches on the web in 2020. It was shown in the report that out of these breaches, hacking accounted for 45%, social attacks 16% and human error caused 22% of breaches. This is more than double the proportion in 2019.

Therefore, a successful solution to the problems of computing vulnerability needs to keep data out of the reach of unauthorized access. One way of doing this is by monitoring people's rights by using a secure method of access control through a reliable system of authentication. In access control, it is a requirement to check the users' rights before granting permission. At any given time, access is limited or granted to the approved user and the implementation of access management techniques must be capable of detecting any form of impersonation or infringement. Many access control techniques has been proposed, these include: Mandatory Access Control (MAC), Discretionary Access Control (DAC), Role-Based Access Control (RBAC), Relationship-Based Access Control (ReBAC), Attribute-Based Access Control (ABAC) etc. ABAC is effective and efficient for big companies or big data [4]. All the existing Access Control techniques, out of which RBAC and ABAC are the most popular, have shortcomings fitting into the data security problems in the cloud service environment. RBAC has the problem of role explosion which happens when an organization's roles outnumber its actual users, and role changes may cause the authorization associated with each role to be modified or removed. There is a limitation to the volume of data this tool can handle hence it can only be used for a private cloud. ABAC is policy-based but neutral in that it can be used in many access control strategies to communicate various types of policies and the decision-making capability is very dynamic, which is a very significant benefit. ABAC uses the attributes of the user and resource for decision making. However, user attributes are also at risk of being compromised by intruders due to the multiple rules involved in permitting the user. It is not specifically defined what attribute should be, the effect of altering attribute is not known. Besides, the policy's implementation of ABAC is complex and difficult due to the fact that it is rule-based. It requires a lot of policies for each of the constraints to be implemented, such as object, subject and other underlined entities. In addition, there was no method to classify and monitor users' attributes as they travel through the system. ABAC leveraged on authorization more than authentication [5] hence we proposed a new system that is authentication-based.

In this paper a Symbolic Attribute-Based Access Control (SABAC) model is presented. It aims at protecting users' attributes so that their data is not exposed to intruders. To accomplish this, a symbolic code is generated by concatenating the image of the user with other attributes to produce a unique hashtag code for each users. In SABAC technology, the picture is a very important attribute that has been captured and encrypted to compute the hashtag using MD-5. The model is novel because the encrypted image is converted to code that hides the user's identity from an unauthorized person.

The remaining of this paper is organized as follows: Section 2 describes the background of the studies and related works, section 3 presents the design of SABAC and section 4 concludes the paper.

II. BACKGROUND AND RELATED WORKS

II.1 BACKGROUND

The overview of the previous works on access control indicated that ABAC is reliable and more efficient than others in literature but its shortcoming makes it not perfect for use in the cloud. So far, any appropriate formal specification for ABAC and its operational functionality has not been recorded. ABAC is characterized by NIST as an access management system where approvals are granted or refused on the basis of attributes assigned to the subjects, objects, environmental, and a collection of policy specifying the attributes [6]. The ABAC model was strongly acknowledged to be in development process and still have varying viewpoints on it [7].

ABAC is made up of Constraint, Permission, Object, Subjects and Users. Each of these has an attribute attached; Object Attribute (OA), Subject Attribute (SU), User Attribute (UA) [8]. The terms attributes are applied differently by organizations depending on what the system is meant to do. In some cases the ABAC model consists of entities called subject; that can be referred to as device, program, mechanism, user etc. Likewise, the object can be server, network, archive, and so on. Whereas, the context attributes are the environmental factors and constraint are the policies. The entities depicted as user will generally represent a person, a system representation of user assigned a selected session or time-frame could be also be a subject. An object could be a resource to be accessed by a user. The context may be when the access is made, location, the device actions, etc. Precise user characteristics include names, date of birth, rental day, address, telephone number, occupational title, identification number, etc. The addresses of the Internet Protocol, Global Positioning System are the values associated with location. The user needs to be assigned attributes to identify permission to access the services.

ABAC permits a lot of flexibility and individualization in access control management by combining the user attribute to data for decision making on the resources. This quality is more reliable to exploit in cloud service environment than using users' role or hierarchy as in RBAC. ABAC, provides a lot of management for the protection of data than the RBAC. It's built towards handling large information like what it's obtainable in the cloud computing [9]. The ABAC model works through a combination of authentication and access control authorization, which seems too cumbersome and poorly fixed. Authorization governs the approval of users to access resources dynamically with a number of rules guiding authorization. User will solely be permissible to access objects that its attributes are confirmed to align with the configured policies. This should be verified to meet the requirements of authorization after approval. Additionally, constraint specification is a sensitive aspect of achieving viable and robust access policy in a company that uses ABAC and it causes implementation difficulty as a result of several attributes to be specified for the entities concerned. That suggests there are several policies to be specified in ABAC unlike in RBAC that make use of the singular role of the user in the organization.

Any security model's behaviour dictates its simplicity. [10] claimed that ABAC's policy requirements are too many and that implementations is also become tasking. Also, ABAC is difficult to evaluate because of different characteristics and policies. If a system is not easy enough to be used by anybody without strict guidelines, it may be labelled as weak or not successful. ABAC focuses mainly on what the user is doing with the resource, which is authorization and mellows down on authentication since the system is more on policy [11].

II.2 RELATED WORKS

In their article Cross Bread Role-based Access Control for Extended Security at Azure in Cloud Computing. The proposed Advanced RBAC Architecture is a form of ontological definition which keeps a records of backup data. The log and user limit per position is sent to the server on the cloud. There is no backup in the previous framework that can lead to data security risks in the event of data failure. These new features have the advantage of improving cloud storage protection and avoiding data loss. In terms of the number of roles per user, the architecture addressed in this work has a constraint [12].

In "A New Semantic Role-Based Access Control Model for Cloud Computing", stated that the present role-based access control is user-centric, if indeed the user has no permission, the demanded access function would be denied. Meanwhile the semantic model implemented functions and responsibilities of the request of the user. The model contains the broker, the user, and the knowledge-based layers respectively. The user layer does the job of dispatching the request to the broker layer. It is necessary to note that user's requests cannot be rejected under this system, rather they should be re-organized to find suitable roles and positions that suit the user's purpose. This will be forwarded to the authorization agent, who will submit the assignments of the responsibilities and functions to the knowledge-based layer. The database consist of SPARQL, which is responsible for running a query for the user's discovered roles and functions once it is applicable to the user's permission request. This layer also consists of a graph of ontology that maps the permissions, functions, and features. The graph showed the direct relationship at the request level between roles and functions as opposed to the relationship at the permission level between roles and functions. Although this model is very flexible and versatile for cloud services, but searching for a matching position and feature that suits the request of the user may be time-consuming [13].

Study on ABAC and the implementation of IaaS models was carried out by [5]. This was based on the current model and the shortcomings of the role explosion in RBAC. The inconveniences of RBAC necessitated the development of the ABAC model. It is made up three phases; harmonization of the DAC, MAC, and RBAC features for ABAC, Second, in ABAC, a new model was developed to manage the dynamic characteristics of RBAC and its extension. Finally, to handle user attributes, an administrative model was developed. The generalized user-role assignment model is called the administrative model (GURA). This model was poor because if membership laws for permission are not met, an automated modification for an automated features selection should be provided, but such a requirement was not obtainable in the current ABAC.

The [14] invented and developed Cipher-text Policy-Based Encryption (CP-ABE) to manage secure access in his work, Toward Successful Access Control using Attributes and Pseudo Roles. Two-step access controls were developed and validated through Bilayer Access Control (BLAC) approach. In BLAC, level one confirms if user has made requests to access BLAC pseudo-roles. The level two tested the rules for further access restrictions within the related BLAC policies. BLAC, therefore, makes good use of attributes while retaining the benefits of RBAC. One problem noted in this study is that the user's right to privacy can be exposed because several characteristics are used in the process of requesting for access. Also, due to the repetitive nature of the policies, user's privacy can be exposed to attack. In BLAC, the attributes of the required resources and the requester are forwarded to the access Decision agent on the remote server. BLAC model was still prone to attack even with the policies been stored remotely.

The [15], concentrating on infrastructure-as-a-service, coordinated in the 4th yearly report for state of cloud adoption 2015. The survey asked 930 IT specialists about the use of cloud computing and associated systems. Participants were drawn from project manager to managers and experts who have represented organizations across various sectors. The report revealed the need for wider knowledge and proper education on cloud technology. And consumers should be more enlightened about the fear of losing their data on account of company folding-up to the cloud owners.

In his Ph.D. thesis, Cryptographic Enforcement of Attribute-Based Authentication, [16]. There was an enumeration of the distinction between the various ABA systems. The investigator also tracked the numerous authentication methods in literature that were listed as traceable and untraceable ABA. These elements represent new fields of information security for researchers.

In the scheme of [17] four stages of its authentication procedures used stationary probability. These are: setup, registration phase, phase of authentication, and phase of changing passwords. Because of the normal username and password used, this work has restrictions that could be prone to modification by an unauthorized user.

In “The Design of Hybrid Cloud Migration Techniques”, reviewed some recent benefit of cloud computing. Organizations have been motivated by enhancement and cost effectiveness to migrate their data, software, and other material into the cloud. The issues concerning the rate at which cloud was adopted was discussed. An effective way for information to be transferred to the cloud were suggested. Improved protection and access management methods, ease of use of migration software, as well as improved monitoring tools were proposed as areas of further study [18].

In their paper, “Authentication Techniques in Cloud Computing: A Review”, [19] presented the different techniques of authentication. It says that if user data is not secure, the cloud has lost its importance. The protection of data in the cloud is very significant. Username and password was among the different techniques listed but turned out to be the least effective of the techniques discussed and is expected to become increasingly irrelevant since it cannot cope with current security challenges on the cloud.

The [20], proposed the Advanced Encryption Standard (AES) Data Encryption versus Decryption Algorithm. AES is among the best methods used for encrypting and decrypting data in the field of cryptography. He said there was no documentation as to where this algorithm could be breached by hackers. It operate in the following major three sizes such as 128, 192, and 256 bits. The block consist of 128-bit cipher size block each.

This research, “A User-Centric Access Management System for Cloud Computing” by [21] showed the concerns users have on the protection of their migrated data and the thought of losing it. It introduced an access management strategy that invoke PKI for user to secure their products such that the security challenges can be addressed. The PKI techniques work for small data, if you use it with big data, it can make your computation problems worse. Secondly, the certificate authority that converts it into a digital certificate by digitally signing the public key will issue a false certificate by fooling users into submitting their data to incorrect certificate.

Between 1989 and 1992, Ron Rivest modeled many types of MD. MD2, MD4, and MD5 are the families of message digests. In 1990, the MD4 algorithm was developed for its computational velocity in software processing. The message is divided into 512-bit consist blocks and produces a digest length of 128 bits. It is also vulnerable to collision attacks, but it generates more resistance than MD2 [22]. Rivest created an updated version called Message Digest 5 (MD5) in 1992 to replace the earlier MD4 because of security problems [23]. It divides the message into 512-bit blocks and, like MD4, generates a 128-bit length digest. If there is an unintended corruption, it is mainly used to verify data integrity. Collision attacks and pre-image attacks are also vulnerable [24].

In ABAC, the following issues were identified which necessitated the design of SABAC;

- Attributes are easy to impersonate,
- Attributes lack standard,
- ABAC is policy-based and it is complex to implement
- and lack of monitoring mechanism for attributes’ identification.

Therefore, there is need to build an access control model that is more stable and user-friendly.

III. DESIGN OF SABAC

There are two main aspects of access control: authentication and authorization. This study concentrated on authentication. An authentication tool named Symbolic Attribute-Based Access Control (SABAC) was developed using The Message Digest-5 (MD5) encryption method to produce Hashtag Symbolic Authentication (HSA); a multi-factor, image-based access control mechanism to enhance the ABAC model. Confidentiality, integrity, and availability are the security metrics used for assessing SABAC performance. SABAC is a tripartite model that incorporates the following principles;

(i) Integrity: the approved users are granted access by combining user attributes and the HSA feature.

(ii) Confidentiality: user data was encrypted and decrypted to ensure confidentiality

(iii) Availability: This was accomplished by maintaining a time the data need achieved.

The MD5 encryption algorithm was used to build the HSA framework. This produces a code that is unique for every user. Each facial image used in the authentication process has the unique code attached to it. The SABAC model was developed and improved with the incorporation of Message Digest 5 (MD5) encryption algorithm to produce HSA key. The following principles was used to generate the key; what you have, what you know, and the extension of what you have. The attributes of the users; last name, first name, and picture are what you know, while what you have is the password. These entities have been joined together using the MD5 encryption to produce the HSA key to make an extension of what you have.

Several stages of adding and padding bits substantially made up MD5 algorithm. Depending on the message length, it takes many rounds of iteration to totally transmit message. The longer the message is, the more rounds for each block of message to be processed. In certain other cases, the bit length may be shorter than the expected norm to calculate the MD5, padding may be required for such bit length. To build the standard bit, a 64-bit representation must be inserted. The MD5 storage and process is stored in a buffer [25]. The steps for carrying out the MD5 algorithm are the following: Append padding bit, append representation of 64 bits, Initialize the MD buffer, process every block, and output. Due to the use of three-way authentication, with the symbolic strategy, an attacker cannot readily break into the cloud resources. The HSA code is far too lengthy to assume, and even if the names were guessed, the image feature could not be fabricated. None of the existing systems in literature used an image to produce a key for authentication systems as a major requirement. SABAC is a multi-faceted enhanced technology that incorporates the uniqueness of the picture of the user to ensure that the confidentiality, integrity, and availability of the cloud service are maintained during data access.

The advent of cloud computing has led to several security issues and a great deal of research in the field of cloud security is needed. In the cloud computing service environment, the SABAC model is a solution to these cloud security problems and can prevent many cyber-crime attacks and other security challenges. With SABAC, threats such as eavesdropping, denial of service attack, insider attack, outsider attack, passwords forgery, impersonation etc. are prevented. Refer to Figure 1 for the flow diagram of the SABAC conceptual model. SABAC's conceptual framework is a process that continues without a break. HSA is an important component for tracking authorized users in all phases of SABAC authentication system. HSA acts as a booster which needs to show often while the user access resources. This indicates that the user is authenticated. The SABAC model was developed for cloud systems and is deployed using Infrastructure as a Service (IaaS).

Please note that in the SABAC authentication process, the user image is converted to create the HSA key. A series of iteration processes are repeated when the HSA key is created. Each time a user attempts to log into the system, the symbolic key is compared to the content of the hash table for that user. Figure1 describes the five phases of SABAC.

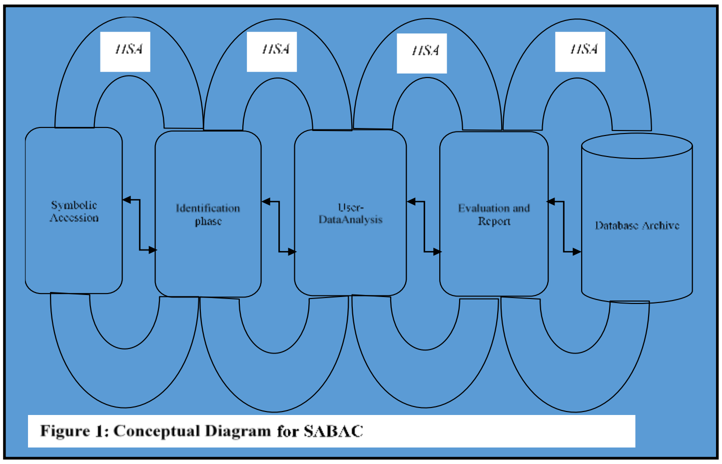


Figure 1: Conceptual Model of SABAC.
Source: Authors, (2020).

The phases used in this model are the symbolic accession process, the identification phase, the analysis of the user's data phase, the report evaluation phase, and the data archive phase. These five stages have an HSA that applies to it, and the transition to the fifth stage includes archiving the users' information.

The symbolic accession process, which involves user registration and identity capture, is the first phase. It is presumed that while some users are new, some may have registered with the owner of the resource. Two processes are involved in the Symbolic Accession phase; new user registration and signing in of existing users. In new user registration, the device prompts the user to supply his/her last name, first name, username and password. To generate the HSA key. The data entered by the user are acknowledged and saved in the repository. To complete the registration process, for new users, the system undertakes the following steps:

i. The user is instructed to take a photograph. The device's camera is turned on, and a picture of the user is taken.

ii. The module generates an image-id for the snapped image. The HSA code is generated as follow:
- The name of the user such as first-name, other-names, and last-name are saved as 64-bits string.
- The user's image is translated to string of 64-bit.
- Every attribute that have been collected are concatenated and converted to hexadecimal values.
- To produce a hash-tag key an MD5 encryption is applied.

iii. The hash-tag key genereated is store in database for reference and comparison next time the user intends to use the system.

The second process is the confirmation of the existing user profile. The Login procedure for registered individual to get entrance to the system includes the following:

i. The system prompts user to enter username and password.
ii. User's picture is automatically snapped by the system camera.
iii. The data stored in the archive is retrieved and compared with the new ones; the latest picture must be identical to what is in the archive.
iv. A new image-id is created, and the process is compared to the current image-id; the image's similarity and confidence threshold are calculated and saved in the database. If all of the information is identical to that in the database, access is granted; otherwise, access is denied.

Identification phase is the second phase. During this phase, the system serves as a trusted authority. The identification process identifies all the entities that produced the symbolic elements. The trusted authority will do a thorough check of the attribute combined to formed the HSA, and then connect them to the HSA key. The Symbolic portion is transformed for visual confirmation and readability. Since the HSA key is linked to the user image stored in the database, adding it to this stage improves the user's traceability. This distinguishes the work because the symbolic accession process ensures that the user attributes are captured correctly and encrypted. The identification phase ensures access is given to the authorized user by ensuring simple verification process. For example, a session can shut down unexpectedly, thereby allowing an attacker to login through the last login session that was not properly turned off. Identification phase is meant to detect such an unauthorized attempt. SABAC uses the HSA code generated and image classification process to ensure that the confidence level is within 95-100 percent and similarity thresholds of the system is 1 for the verified users. However, if the values are below 95 percent and 0 for both measures, the system classified the user as not confirmed.

The next phase is the users' data analysis. This is a process which includes tracking of user, validating user action, retrieval of hidden/encrypted data, confirming the transaction date and time stamps, etc. The HSA key is used for tracing who handles a specific activity on the cloud resource. This helps to ensure that the system confidentiality (unauthorized person not allowed into the system), integrity (to ensure data is valid and accurate) and availability (time and date stamp testing to verify that the system is fast enough during authentication and the resource are readily available for use).

The report evaluation phase assesses the process that takes place in the platform. It includes clarifying whether the described and evaluated components are indeed important to the kind of data held by the user. Also, apart from maintain confidentiality, integrity and availability, the HSA key is used for the justification of user's record that was retrieved and analyzed. Information must be:

i.

Identical with the retrieved original.

ii.

The encoded value encrypted is always the same.

iii. The retrieved information is independently checked and analyzed.
iv. If an intruder makes an attempt, a warning message may be sent to an approved person.

Password changes are an important part of this operation, and they should be done on a regular basis. The username, old password, and new password are all provided, with the old password being overwritten. HSA and New image-id are also generated and stored. This approach ensures that the SABAC system's credibility is maintained. The efficacy of the service's accessibility, the evaluation of the user HSA, password modification, unauthorized attempt, timestamp, percentage of trust level, and similarity are all available for reporting. The device also shows the user status; 0 to indicated unauthorized user and 1 to indicate authorized user.

Database Archive is the last phase. The administrator is responsible for the protection and storage of database of the user's data, as well as the organization. For further processing, the HSA and other attributes given are stored in the database archive. The image is used as one of the major prerequisite attribute needed for accessing the database; this is controlled and enforced by encryption of the user attributes and HSA information collected. The database is useful for traceability in the case of any user dispute or if hackers or attackers attempt to unlawfully break into the system. If there is no backup, it means that if a disaster occurs, the company's data is lost. This is one of the shortcomings in some of the existing cloud platforms. SABAC supports the use of a secure backup system that can be used during repairs, downtime, or system failures. To ensure the continuous availability of both service and resources, this is a primary requirement for using this security framework. Figure 2 displays the SABAC Database Archive.

Security metrics (Integrity, Confidentiality, and Availability) were used to assess SABAC model to measure its efficiency. SABAC model performance is enhanced by the HSA algorithm. The integrity, confidentiality, and availability of the data on the IaaS platform are guaranteed by this new access control model. HSA is a symbolically generated code that connects user permission with their attributes, facial appearance inclusive. An encrypted code is generated using the MD5 hashing algorithm using the specified attributes. Each individual has a distinct appearance that defines who they are by nature; this characteristic is converted to a 64-bit hexadecimal value. The names of the user, image, username and password were combined together using message digest to generate one-way encryption code for authentication purposes. The MD5 uses 128-bit memory with minimum of 512-bit message length for generating the hash code. The message length can be smaller than the minimum value required, it is padded by adding a 64-bit value to it.

The Symbolic code is used to create a connection between the picture that has already been captured and the face that was under monitoring. HSA is used to find out who should have access as contained in the access monitoring list and what roles and resources they have access to. For instance, in the SABAC cloud security system, the user can possess individual permission right to do any of the following on the service platform depend on their subscription; right to change, remove, update, and view data.

The HSA uses the MD-5 encryption algorithm to generate the hash-tag code. HSA is a distinctive code that offers a user access to cloud platform services. SABAC is differentiated from other access control models in literature by its symbolic feature. The symbol differentiates one user from another. It's seemingly difficult to alter the user's facial image if other details can be manipulated. Shown in Figure 3 is the HSA process flowchart diagram. Also, the algorithm for HSA is also shown in Figure 4.

There are two stages in the HSA process flow mechanism. These are login and request to use cloud services. In the first stage, new users are signed up by creating a personal profile. The attributes are concatenated and converted to hexadecimal which serves as input into MD5 encryption algorithm to generate an HSA key. Existing users log-in by providing their username and password, and the device conducts confidentiality authentication and information validation by face recapturing. The second stage is the request for user accessibility to cloud services. The process includes: the request to access the database, the system's integrity authentication check to assess the user's restriction level, the process of face validation, the verification of availability authentication, and the face validation.

The SABAC algorithm is as follows:

```
Begin
    If 'Existing User'
    Input username
    Accept password
    Input HSA
            If 'username' and 'password' is valid
                    Capture face
                    Verify face
            Else
                    Exit
            End if
            If face is valid
                    Generate HSA
                    Retrieve user's attributes
                    Compare user supplied HSA with computed HSA
            Else
                    Exit
            End if
            If HSA is correct
                    Confirm captured face
                    Grant Access
                    Calculate Confidentiality
                    Calculate Integrity
                    Calculate Availability
            Else
                    Exit
            Endif
            While User Is logged in;
                    Capture face at interval
                    Verify face
                    If face is 'not valid'
                            Exit
                    Else
                            Continue
                    Endif
            End While
    Else
            Create username and password
            Capture user facial image
            Convert image to string
            Concatenate (username + password + image)
            Convert code to hexadecimal
            Apply MD5 encryption
            Store attribute in database
            Exit
Endif
```

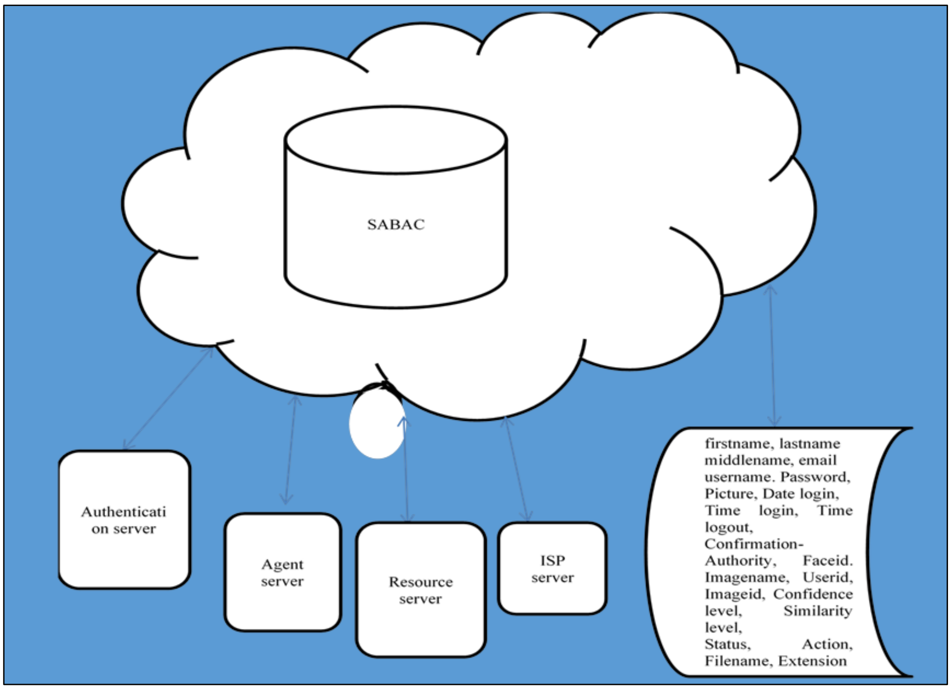


Figure 2: SABAC Database Archive.
Source: Authors, (2021).

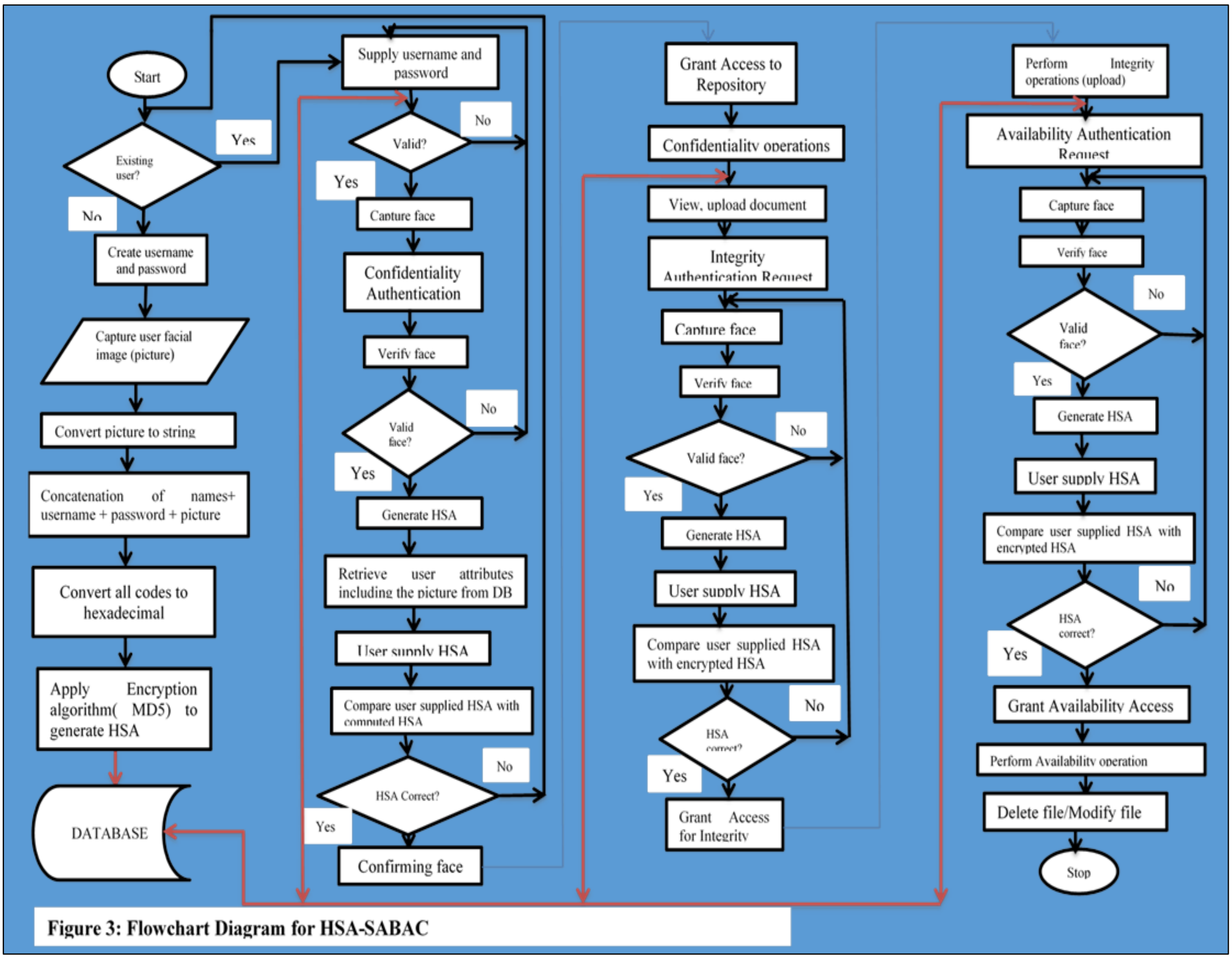


Figure 3: SABAC Model Flowchart
Source: Authors, (2021).

IV. RESULTS AND DISCUSSIONS

SABAC performance is evaluated using three security metrics: confidentiality, integrity, and availability.

i. Availability

This is the number of users authenticated per unit of time which is also called throughput. It is determined in this study by adding the time required to validate a user and the time required to produce HSA. The number of users divided by the time it takes to authenticate each user is the general formula for throughput (availability). The formula for throughput can be deduced as follows:

Let P represent the availability performance, T total time taken to validate user, and U represent the number of users.
The throughput is implemented thus;

Throughput

$$P = \frac{U}{\sum_{i=1}^{m} T_i} \tag{1}$$

where i=1…m

One way to determine the efficiency of SABAC system is to look at its throughput. In measuring throughput, the length of time the system spent in authenticating each user was recorded. This gives an insight to how fast the system attends to users request and how many users it can handle per time.

Table 1, shows the image validation time and authentication time for each group of users. It is worth noting that image validation time is a subset of the authentication time for each user and it determines significantly the time SABAC spends on authenticating such user.

It is observed from Figures 4, that the image validation time and authentication time is fairly stable while the number of users increased. This makes SABAC to be stable and always available to users as the load increases in the cloud. In figures 5, it is observed that SABAC becomes rugged and maintains an acceptable speed of authentication even as the workload increases.

In Figure 5, the throughput of SABAC was presented. The system become stable as the load increases, for instance, when the number of users is 40 and above, the throughput become higher as the number of users increase. It is observed that the throughput increased by 15,000% to 1,828 users per second between 10 users and 100 users. This implies that the system become more efficient as higher number of users enter the system.

In the cloud security system, the system should disallow unauthorised person from accessing the resource. The system's precision is determined by the following possibilities: false positive, true negative, false negative, and true positive. At all levels of access to cloud services, the system should not allow a false positive.

Table 1: Availability

Number of Users	Image Validation Time (ms)	Authentication time (ms)	Total Time (ms)
10	1.28	4.19	5.47
20	1.29	4.33	5.62
30	1.34	4.41	5.75
40	1.49	4.52	6.01
50	1.55	4.59	6.14
60	1.61	4.7	6.31
70	1.64	4.76	6.40
80	1.66	4.81	6.47
90	1.7	4.87	6.57
100	1.74	4.91	6.65

Source: Authors, (2020).

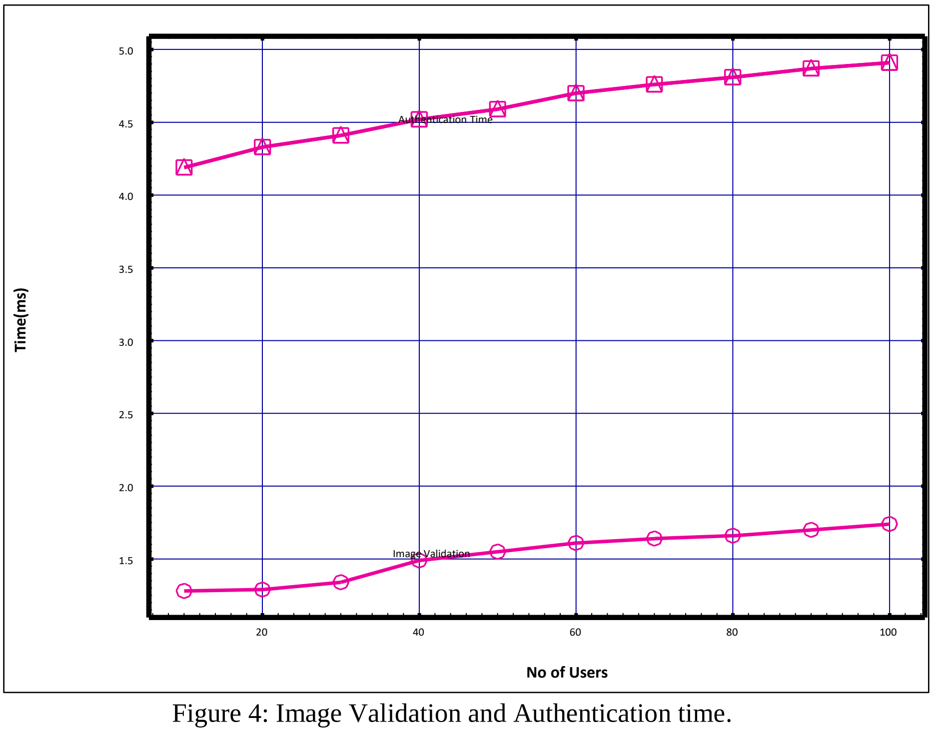


Figure 4: Image Validation and Authentication time.
Source: Authors, (2021).

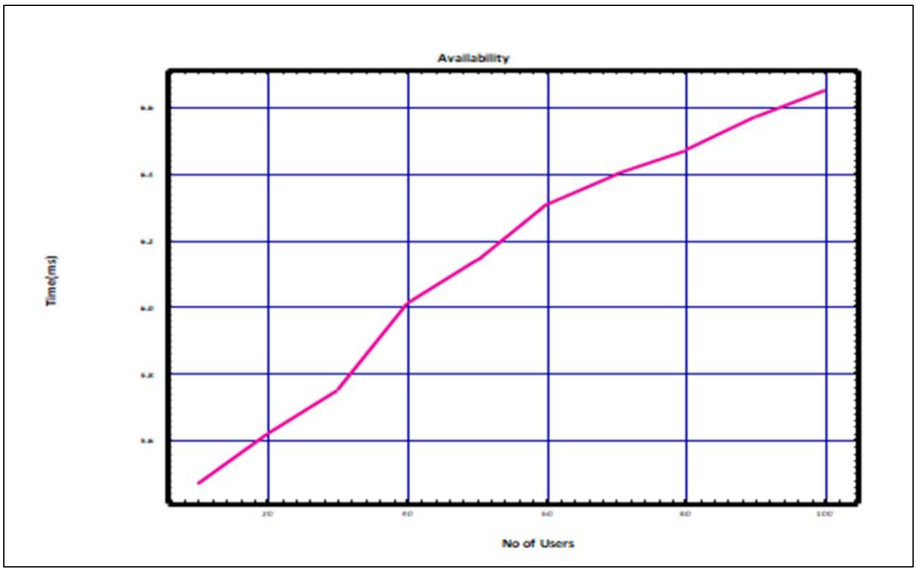


Figure 5: Graph of Availability.
Source: Authors, (2021).

Integrity of the system is tested by checking the similarity value between the symbolic credentials stored in the database and the recent symbolic information collected when the user attempted accessing the system. If there is no discrepancy in the encrypted and the new data, the image is successfully validated. Each successful process is rated based on the criteria incorporated into the classification algorithm. Parts of the determining factors are the pattern of the image, environment and the level of interference in the vicinity where the image is captures.

In Table 2, The similarity threshold for 10 to 100 users are 92.88236 to 99.9980 respectively. The more the number of users in the system the higher the stability and accuracy.

In Figure 6, the integrity graph, the increase in the similarity threshold as more people access the system. The data become stable and the system was not interrupted in any way. The flow of the graph shows that if the number of users increases at 10 folds, there was no case of unauthorized attempt to disrupt the functionality of the system.

Confidentiality is evaluated in this study by the confidence level, which is the percentage of confidence the system has in the user it authenticates. The confidence thresholds is protected using HSA to reconfirm the user privileges to the cloud. SABAC re-evaluates the confidence level of the user at each access point services. The confidence level ranges from 0% to 100%. The higher the rating, the greater the confidence SABAC system has in the user. This authentication system has a minimum trust level of 95%.

Table 2: Integrity.

Number of Users	Similarity Threshold
10	92.88236
20	95.68957
30	95.9957
40	98.49678
50	98.79778
60	98.8965
70	99.8750
80	99.9856
90	99.98975
100	99.9980

Source: Authors, (2020).

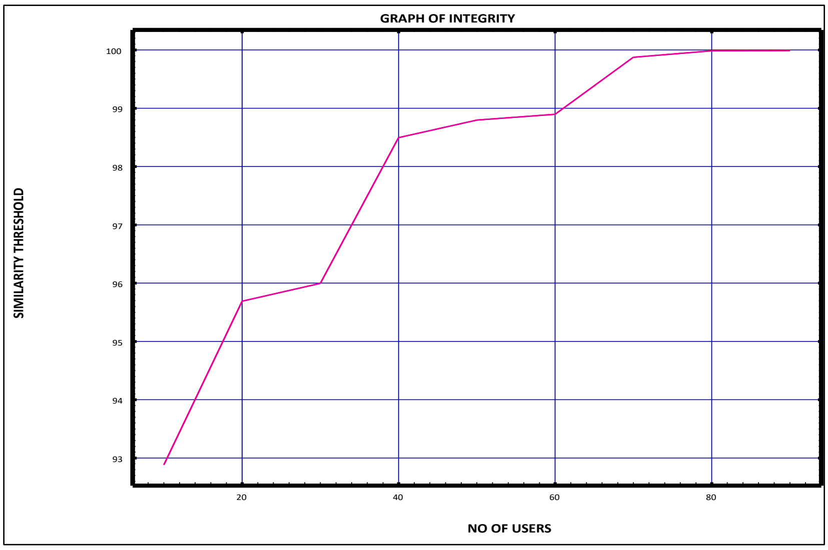


Figure 6: Graph of Integrity.
Source: Authors, (2021).

Table 3: Confidentiality.

Number of Users	Confidence Level
10	99.9999
20	99.9997
30	99.9997
40	99.9995
50	99.9994
60	99.9994
70	99.9994
80	99.9993
90	99.9993
100	99.9993

Source: Authors, (2021).

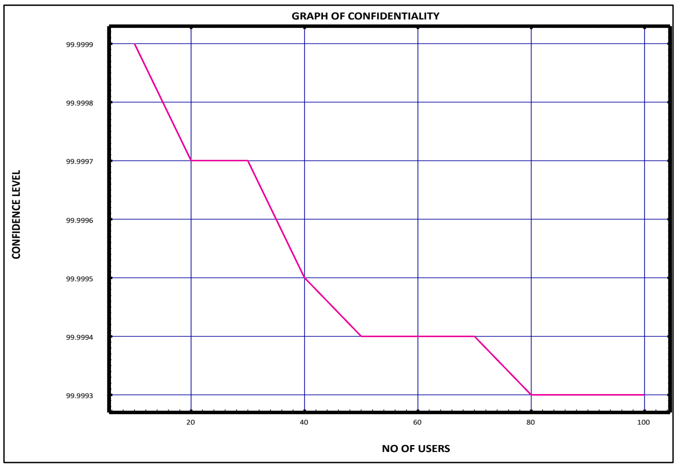


Figure 7: Confidentiality Graph.
Source: Authors, (2021).

The average confidence level for each group is as presented in the Table 3. The presented data shows that the confidence level of SABAC ranges between 99.9993 and 99.9999. This implies that SABAC confidentiality measure is very high and acceptable. It should be observed that as number of users increase, the confidence level tends towards a limit of 99.9990. Figure 7 presents graph for confidentiality.

V. CONCLUSIONS

The value of security cannot be overlooked in cloud computing. It is becoming more and more essential for Internet applications and services. Cloud computing usage is increasing, and it is rapidly becoming main-stream infrastructure. Cloud computing's adoption and use come with major security concerns. As a result, SABAC was built to fix the security problems that come with cloud usage. For successful cloud protection, the developed system was implemented with multifactor attributes fortified with image code. As a result, SABAC is an excellent solution for using cloud services in a safe environment. The use of SABAC as an authentication mechanism would enhance cloud resource security and restore confidence in cloud adoption and use.

VI. AUTHOR'S CONTRIBUTION

Conceptualization: Iyabo Felicia Oyeyinka.
Methodology: Iyabo Felicia Oyeyinka and Sunday Idowu.
Investigation: Iyabo Felicia Oyeyinka and Sunday Idowu.
Discussion of results: Iyabo Felicia Oyeyinka and Sunday Idowu and Afolashade Kuyoro.
Writing – Original Draft: Iyabo Felicia Oyeyinka.
Writing – Review and Editing: Iyabo Felicia Oyeyinka and Sunday Idowu.
Resources: Iyabo Felicia Oyeyinka.
Supervision: Sunday Idowu and Afolashade Kuyoro.
Approval of the final text: Iyabo Felicia Oyeyinka and Sunday Idowu and Afolashade Kuyoro.

VII. REFERENCES

[1] Sandeep, B., "Cloud Computing." University Printing House, Cambridge CB2 8BS, United Kingdom One Liberty Plaza, 20th Floor, New York: NY 10006, USA 477 Williamstown Road, Port Melbourne, Vic 3207, Australia 4843/24, 2nd Floor, Ansari Road, Daryaganj, Delhi – 110002, India 79 Anson Road, #06–04/06, Singapore 079906, 2017.

[2] RSA, "The Current State of Cybercrime: An Inside Look at the Changing Threat Landscape," 2014.

[3] Teri, S., "Data Insecurity on the Rise." The Verison Business Data Breaches Investigation Report, Financial Management Magazine. www.fm-magazine.com, 2020.

[4] Vincent, C. H., Kuhn, D. R., & Ferraiolo, D. F., "Attribute-Based Access Control, "National Institute of Standards and Technology. CSDL Issue No. 02, pp 85-88, ISSN: 0018-9162, http://doi.ieeecomputersociety.org/10.1109/MC..33, 2015.

[5] Xin, J., "Attribute-Based Access Control Models and Implementation in Cloud Infrastructure as a Service," The University of Texas AT SAN ANTONIO College of Sciences Department of Computer Science, 2014.

[6] Vincent, C. H., Kuhn, D. R., & Ferraiolo, D. F. "Guide to Attribute-Based Access Control (ABAC) Definition and Considerations," NIST Special Publication, 2013.

[7] Mikko, K.,"Enforcing Role-Based Access Control with Attribute-Based Cryptography for Environments with Multi-Level Security Requirements." Aalto University publication series DOCTORAL DISSERTATIONS. Department of Information and Computer Science, Finland, 2016.

[8] Khalid, Z.B., Ram, K., & Ravi, S.," Constraints Specification in Attribute-Based Access Control, Institute for Cyber Security," Department of Computer Science, the University of Texas at San Antonio, 2013.

[9] Richard, K. D. & Edward, C. J.," Adding Attributes to Role-Based Access Control," 2010.

[10] Jin, X., Krishnan, R., & Sandhu, R., "A Unified Attribute-Based Access Control Model Covering DAC, MAC, and RBAC," .In DBSec, 2012.

[11] Sandhu, R. S. & Samarati, P., "Access Control Principle and Practice," Communications Magazine, IEEE, vol 32. no 9, pp 40-48, 1994.

[12] Parminder, S., & Sarpreet, S., "Cross Bread Role-based Access Control for Extended Security At Az-ure in Cloud Computing," International Journal of Application or Innovation in Engineering & Management (IJAIEM): ISSN 2319 - 4847 www.ijaiem.org Email: editor@ijaiem.org, editorijai-em@gmail.com, 2013.

[13] Masound, B., Mohammad, S.K., Soheil, L., & Azizallah, R., "A New Semantic Role-based Access Control Model for Cloud Computing," The Ninth International Conference on Internet and Web Applica-tions and Services. ISBN: 978-1-61208-361-2, 2014

[14] Alshehri, S., "Toward Effective Access Control Using Attributes and Pseudoroles," The Ph.D. Program in Computing & Information Sciences B. Thomas Golisano College of Computing & Information Sciences Rochester Institute of Technology, 2014.

[15] Kim, W., "Cloud Computing Trends: State of the Cloud Survey," 2015.

[16] Huihui, Y., "Cryptographic Enforcement of Attribute-Based Authentication." The University of Agder, Faculty of Engineering and Science, 2016.

[17] Djellali, B., Belarbi, K., Chouarfia, A., & Lorenz, P., "User Authentication Scheme Preserving Anonymity for Ubiquitous Devices. Security and Communication Networks, vol 8, no17, 2016.

[18] Oyeyinka, F. I., Awodele, O., Kuyoro, S. & Oyeyinka, I. K.,. "The Design of Hybrid Cloud Migration Techniques," 2nd National conference and 4th Induction ceremony of Nigeria Women in Information Technology, Abuja, Nigeria, 2016..

[19] Seyed, M. D., & Sara, N., "Authentication Techniques in Cloud Computing: A Review," International Journal of Advances Research in Computer Science and Software Engineering, vol 7, no 1, 2017.

[20] Ako, M. A.,. "Advanced Encryption Standard (AES) algorithm to Encrypt and Decrypt Data," Department of Applied Mathematics and Computer Science, Eastern Mediterranean University-Cyprus, 2017.

[21] Imran, K., "An Introduction to the applications of cloud computing technology in academic libraries," International Journal of Library Management and Services. ISSN:2349-6347; vol 4, no 2, pp 15-24, 2018

[22] Yu, S., Yusuke, N., Noboru, K., & Kazuo, O., "Improved Collision Attacks on MD4 and MD5," 2007.

[23] Rajeev, S., & Geetha, G., "Cryptographic Hash Functions: A Review. School of Computer Science," Lovely Professional University Phagwara, Punjab. International Journal of Computer Science (IJCSI), ISSN: 1694-0814, vol 9, no 2, pp 2, 2012.

[24] Yu, S., & Kazumaro, A., "Finding Pre-images in Full MD5 Faster Than Exhaustive Search," 2009.

[25] Kanickam, L. H. S., & Jayasimman J., "Comparative Analysis of Hash Authentication Algorithms and ECC Based Security Algorithms in Cloud Data," Asian Journal of Computer Science and Technology ISSN: 2249-0701, vol 8, no 1, pp 53-61, 2019.

Journal of Engineering and Technology for Industrial Applications

ITEGAM-JETIA

Manaus, v.7 n.29, p. 47-51. May/Jun, 2021

DOI: https://doi.org/10.5935/jetia.v7i29.751

ISSN ONLINE: 2447-0228

RESEARCH ARTICLE OPEN ACCESS

DETERMINATION OF FACTORS AFFECTING THE CONSUMPTION BEHAVIOR OF ORNAMENTAL PLANT CONSUMERS: ŞANLIURFA SAMPLING OF TURKEY

Mustafa Hakki Aydoğdu*[1] and Necla Yildizoğullari[2]

[1, 2] Department of Agricultural Economics, Faculty of Agriculture, Harran University. Sanlıurfa, Turkey.

[1] http://orcid.org/0000-0002-4945-5239, [2] http://orcid.org/0000-0001-9372-9068

Email: *mhaydogdu@hotmail.com, neclahulagu00@gmail.com

ARTICLE INFO

Article History
Received: May 05th, 2021
Accepted: June 18th, 2021
Published: June 30th, 2021

Keywords:
Ornamental plants,
Consumer behavior,
Factors affecting consumption,
Şanlıurfa-Turkey.

ABSTRACT

Ornamental plants are a sector that meets the spiritual and visual needs of people, improves the physical properties of the space they are used in, and contributes significantly to the economy and employment with the added value it creates. Şanlıurfa has an important potential for ornamental plants due to its geothermal resource and seasonal characteristics. The purpose of this study is to determine the factors that affect the consumption behavior of ornamental plant consumers in Şanlıurfa. The material of this research consists of data obtained from face-to-face surveys with consumers selected by a simple random sampling method in Şanlıurfa, and the surveys were conducted in 2020. Multinomial probit, negative binomial regression, and Tobit models were used in the analyzes. According to the results, the gender, marital status, income, employment status, and education level of the participants were determined as effective factors at different statistical significance levels. It was determined that age is not an effective factor. When the income level rises by one level, the ornamental plant expenditure increases by 19.93 TL annually, and when the education level rises by one level, it increases by 17 TL. The spread of ornamental plant cultivation should be encouraged in Şanlıurfa. This study is the first kind on this subject in Şanlıurfa.

I. INTRODUCTION

Plants, especially ornamental plants, have been used as a source of nutrition, food, healing, social and religious ritual, aesthetic, functional, economical, industrial, landscape, decor, morale and motivation, together with the history of humanity until now [1]–[21]. Ornamental plants, which are a sub-branch of crop production in agricultural activities, differ from other agricultural product varieties with their soothing, relaxing, happy, and peaceful results that meet the spiritual and visual needs of people instead of food needs [15]. The term ornamental plants used in agricultural production is an inclusive term and they are collected in four subgroups as cut flowers, potted plants, outdoor and indoor ornamental plants, and natural flower bulbs, and cut flowers have the largest share among ornamental plants in Turkey [16]. Ornamental plants and floriculture is a non-essential necessity item, has become a commercial sector and an important for employment, which is constantly increasing in importance due to its contribution to human psychology, satisfying people's aspirations for the environment, nature, and natural environment in places where industrialization and urbanization are intense [2], [3], [5], [17], [22]. In addition to its positive effects on people due to its aesthetic properties, ornamental plants have become a globally growing sector that softens and improves the physical properties of the place, contributes significantly to the economy and employment with the added value it creates [3]–[5], [8], [12], [17], [19], [21].

Şanlıurfa is the most important province in terms of agricultural production potential, which is included in the Southeastern Anatolia Project (GAP). [23]–[25]. Şanlıurfa has an important potential in terms of ornamental plant cultivation due to its geothermal resources and seasonal characteristics, and also has endemic and geophyte varieties peculiar to the region [4], [19],

[21]. Although Turkey has an important potential in terms of ornamental plants and floriculture, scientific studies in this field are not common enough [1], [3]. The purpose of this study is to determine the factors that affect the consumption behavior of ornamental plant consumers in Şanlıurfa, Turkey.

II. MATERIALS AND METHODS

The main material of this research is the data obtained from face-to-face surveys with ornamental plant consumers in Şanlıurfa, and consumers were determined by a simple random sampling method based on voluntary participation. The population of Şanlıurfa was 2.12 million people according to the results of the address-based population registration system in 2020 [26]. The sampling volume of the research was determined as 384 according to the table of population sizes and sample volumes according to tolerable sampling error [27], [28] with 95% confidence limit and 5% margin of error. To stay on the safe side, 390 surveys based on voluntary participation in the field were conducted in 2020. The data obtained from the surveys were transferred to Excel, analyzed in SPSS, and interpreted. The average and distribution range of numerical data such as age or income of the participants was determined by performing frequency analyzes based on the research questions. The main analyzes were made by regression analysis. Regression analysis is a statistical method used to determine the relationship between two or more variables. It is used to predict models based on dependent or explained variables, and cumulative normal distribution is used in probit analysis [29], [30]. Multinomial probit analysis was performed to determine the characteristics that distinguish consumers in purchasing ornamental plants and flowers. The number of flowers consumers received and gave during the year was analyzed by negative binomial counting data regression. In the analysis of the consumption of ornamental plants and flowers, the Tobit model, which is widely used to explain models with limited dependent variables, was used. The model is a linear regression model with non-negative dependent variables and is based on variables that do not take negative values such as price, wage, and expenditure [30], [31].

III. RESULTS AND DISCUSSIONS

70% of the consumers participating in the research were female and 30% were male. The low number of males in the number of participants is because the research subject was not very interesting by male participants due to the norm culture of the research field. Participants were between the ages of 18 and 70, with an average age of 28.4 years. 53% of the participants are single, 40% are married and 7% are widows or divorced. 41% of the participants are university graduates, 34% are high school graduates and 25% are secondary school and below. 35% of the participants are salaried employees, 7% are self-employed, others are non-working, housewives, retired, and students. The monthly income of 37% of the participants was between 1501 and 5000 Turkish Lira (TL), 17% of them have an income of 5001 TL or more, and the remaining income was 1500 TL and below (The average of 1 USD was 7.01 TL in 2020). The monthly average income of the participants was determined as 3415 TL. Participants' annual average purchase expenses of ornamental plants and flowers are calculated as approximately 126 TL, with a minimum of 10 TL and a maximum of 3500 TL. This calculated average value corresponds to 3 per thousand of the monthly income of the consumers and within the solvency. 55% of the participants buy ornamental plants and flowers only on special occasions and 22% for celebrations.

The most purchased varieties are indoor ornamental plants, bouquets, and arrangements. Most purchases are on mother's day, and the appearance and meaning of the variety purchased are the most important reasons for selection. In a study conducted in Turkey, it was determined that the most flowers were taken on Valentine's Day and then on mother's day [3]. The most purchased place is flower shops, followed by online purchases. Consumers buy flowers on special occasions (Mother's Day, Father's Day, Valentine's Day, New Year, etc.), regularly, seasonally, and for other reasons (such as birthdays, celebrations, promotions, and motivations). While most flowers are bought on special occasions, the rate of those who receive flowers regularly is 9%. The purchasing frequency of consumers is given in Table 1.

Table 1: Consumers' purchase frequency and distribution of ornamental plants and flowers.

Purchase period	Frequency	%
Only on special occasions	213	54,6
Regularly	35	9,0
Seasonally	48	12,3
Other reasons	94	24,1
Total	390	100,0

Source: Authors, (2021).

Multinomial probit analysis was performed to determine the socio-economic characteristics that distinguish consumers from these aspects. In the analysis, the dependent variable is the frequency of purchasing ornamental plants and flowers and the reference category is other reasons, and the analysis results are given in Table 2.

Table 2: Multinomial probit analysis results.

	Variables	Coefficient	Standard error	t	p
Only on special occasions	Age	0.006217	0.011618	0.54	0.593
	Married	0.24683	0.22422	1.1	0.271
	Female	0.100582	0.226898	0.44	0.658
	Unemployed	-0.44574**	0.217892	-2.05	0.041**
	Income	-0.21645***	0.081045	-2.67	0.008***
	Constant	0.936262**	0.397963	2.35	0.019**
Regularly	Age	0.013395	0.015053	0.89	0.374
	Married	0.052466	0.293036	0.18	0.858
	Female	-0.14616	0.288534	-0.51	0.612
	Unemployed	-0.6105**	0.306604	-1.99	0.046**
	Income	-0.00767	0.10235	-0.07	0.940
	Constant	-0.75833	0.514241	-1.47	0.140
Seasonally	Age	0.015073	0.013095	1.15	0.250
	Married	0.205916	0.265599	0.78	0.438
	Female	0.011336	0.281147	0.04	0.968
	Unemployed	0.098927	0.26596	0.37	0.710
	Income	-0.03515	0.097138	-0.36	0.717
	Constant	-0.95262	0.48298	-1.97	0.049

Wald chi-square (15) = 23.59; Log-likelihood = -434.84773; p= 0.0724.
Shows the statistical level of importance by ** 5%, ***1%.

Source: Authors, (2021).

In terms of the multinomial probit model estimation results in Table 2, according to those who bought flowers for other reasons:

- Employees are more likely to buy flowers "only on special occasions".
- The higher the income, the more likely it is to buy flowers "only on special occasions".
- Employees are more likely to purchase flowers "regularly".

The results of marginal effects of the multinomial probit model estimation are given in Table 3.

Table 3: Marginal effects of multinomial probit model estimation.

Variables	Only on special occasions	Regularly	Seasonally	Other reasons
Age	0.000	0.001	0.002	-0.002
Married	0.054	-0.014	0.010	-0.050
Female	0.039	-0.025	-0.004	-0.010
Unemployed	-0.104*	-0.050	0.068*	0.086*
Income	-0.063***	0.014	0.013	0.035**

Shows the statistical level of importance by *10%, **5%.

Source: Authors, (2021).

According to the marginal effects in Table 3:

- Employees are 10.4% more likely to buy flowers only on special occasions than non-employees.
- Unemployed are 6.8% more likely to purchase flowers seasonally than employees.
- Those who do not work are 8.6% more likely to buy flowers for other reasons than employees.
- When income increases by one unit, the probability of buying flowers only on special occasions decrease by 6.3%.
- When income increases by one unit, the probability of buying flowers for only other reasons increases by 3.5%.

The number of flowers that consumers receive throughout the year has been analyzed by negative binomial counting data regression, and the results are shown in Table 4. Alpha testing confirms that negative binomial regression should be used instead of using Poisson regression. It is observed that the number of flowers received by the participant is higher in females than in males, less in unemployed than in employees, and more in married than unmarried. Age, education level, and income level do not affect the number of flowers that the consumer receives.

Table 4: Negative binomial regression estimation results (Dependent variable: Number of flowers received).

Variables	Coefficient	Standard error	t	p	Marginal effect
Age	-0.01096	0.006963	-1.57	0.115	-0.0390685
Education Level	-0.04410	0.054033	-0.82	0.414	-0.1571822
Income	0.043754	0.053289	0.82	0.412	0.1559676
Female	0.398642	0.137064	2.91	0.004***	1.321226
Unemployed	-0.639790	0.131522	-4.86	0.000***	-2.117342
Married	0.315144	0.129026	2.44	0.015**	1.164704
_constant	1.488746	0.331726	4.49	0.000***	
/lnalpha	0.048980	0.091881			
alpha	1.050199	0.096494			

LR test of alpha=0; chibar2 (01) = 1133.40; Prob>= chibar2 = 0.000

Shows the statistically level of importance by ** 5%, ***1%.

Source: Authors, (2021).

According to the marginal effect values in Table 4:

- Females receive 0.39 more flowers than males.
- An employee receives 0.64 fewer flowers than a non-employee.
- Married people receive 0.39 more flowers than unmarried ones.

The number of flowers given by consumers throughout the year has also been analyzed by negative binomial counting data regression, and the results are given in Table 5. The alpha test concluded that negative binomial regression should be used. The number of flowers given is less in females than in males, more in those who are married than in those who are not married, and in those who are not working, less than in those who work. The number of flowers given does not differ in terms of age, education level, and income level.

Table 5: Negative binomial regression estimation results (Dependent variable: Number of flowers received).

Variables	Coefficient	Standard error	t	p	Marginal effect
Age	-0.008540	0.007475	-1.14	0.253	-0.0085429
Education Level	0.009947	0.056078	0.18	0.859	0.0099469
Income	0.082996	0.052858	1.57	0.116	0.0829955
Female	-0.163360	0.134669	-1.21	0.225	-0.1633629
Unemployed	-0.633110	0.136186	-4.65	0.000***	-0.6331074
Married	0.293222	0.134053	2.19	0.029**	0.2932222
_constant	1.626030	0.345427	4.71	0.000***	1.62603
/lnalpha	0.129022	0.086675			
alpha	1.137715	0.098611			

LR test of alpha=0; chibar2(01) = 1485.12; Prob>= chibar2 = 0.000

Shows the statistically level of importance by ** 5%, ***1%.

Source: Authors, (2021).

According to the marginal effect values in Table 5:

- Females give 0.16 fewer flowers than males.
- The employee gives 0.63 fewer flowers than the unemployed.
- Married people give 0.29 more flowers than unmarried ones.

The Tobit model was used to determine the consumer characteristics that determine the annual flower expenditure, and the results are given in Table 6. Consumers with zero flower spending prevent the linear regression to meet the "neutrality of estimators" condition. For this reason, the Tobit model, which is suitable for situations where the dependent variable can also take the value of 0, is used.

Table 6. Tobit model estimation results (Dependent variable: Flower expenditure).

Variables	Coefficient	Standard error	t	p	Marginal effect
Constant	−74.9356	79.6677	−0.9406	0.3469	
Age	1.13181	1.59777	0.70840	0.4787	0.73997
Income	30.4797***	11.4958	2.65100	0.0080***	19.927
Married	21.1025	31.1717	0.6770	0.4984	13.860
Female	−17.8489	31.1998	−0.5721	0.5673	-12.072
Unemployed	−46.6962	30.5238	−1.530	0.1261	-30.571
Education level	26.0066**	12.6644	2.0540	0.0400**	17.003
Sigma	262.343	10.0887			
Constant	−74.9356	79.6677	−0.9406	0.3469	

Shows the statistical level of importance by ** 5%, ***1%.

Source: Authors, (2021).

According to the Tobit model estimation results, as income and education level increase, flower expenditure increases. According to the marginal impact values, when the income category rises by one level, flower expenditure increases by 19.93

TL annually and 17 TL when the education category rises by one level.

IV. CONCLUSIONS

Ornamental plants have started to be among the non-compulsory but widely preferred needs that people are increasingly paying attention to due to reasons such as population growth, urbanization, and industrialization. According to the results of the research, while the ornamental plants and flowers of the consumers participating in the survey were at most 54.6% on special occasions, the rate of those who buy them regularly was 9%. According to the multinomial probit model estimation results, employees' probability of getting flowers “only on special occasions” increases ($p<5\%$). The higher the income, the higher the probability of getting flowers "only on special occasions" ($p<1\%$). Employees are more likely to buy flowers “regularly” ($p<5\%$). In terms of marginal effects, employees are 10.4% more likely to buy flowers only on special occasions than non-employees. Non-working people are 6.8% more likely to purchase flowers seasonally than employees, and 8.6% more likely to purchase flowers for other reasons. When income increases by one unit, the probability of purchasing flowers only on special occasions decrease by 6.3%, while the probability of purchasing flowers increases by 3.5% for other reasons. It is seen that the number of receive flowers is higher in women than in men ($p<5\%$), less in non-working than in employees ($p<1\%$), and more in married than in unmarried ($p<5\%$). Age, education level, and income level do not affect the number of flowers received ($p>10\%$). According to the marginal effect values, women received 0.39 more flowers than men, 0.64 fewer than non-employed, and 0.39 more flowers than unmarried women. The number of flowers given is less in women than in men, more in those who are married than in those who are not married, and in those who are not working, less than in those who work. The number of flowers given does not differ in terms of age, education level, and income level ($p>10\%$). According to the marginal effect values, females give 0.16 more flowers than males, employees 0.63 less than unemployed and married ones give 0.29 more flowers than unmarried ones. According to the model estimation results of ornamental plant and flower expenditures, as income and education level increase, flower expenditure increases. According to the marginal impact values, when the income level rises by one level, flower expenditure increases by 19.93 TL annually and 17 TL when the education level rises by one level.

Şanlıurfa has an important potential for growing ornamental plants due to its endemic and geophyte varieties and seasonal characteristics. The widespread use of ornamental plant cultivation in Şanlıurfa should be encouraged due to the foresight that the prices will decrease with the widespread production and therefore the consumption amount will increase among consumers. This study is the first study on this subject in Şanlıurfa. This study was derived from the master's thesis entitled “Determination of Ornamental Plants Potential: The Case of Şanlıurfa Province” completed by the second author under the supervision of the corresponding author.

V. AUTHOR’S CONTRIBUTION

Conceptualization: Mustafa Hakkı Aydoğdu and Necla Yıldızoğulları.
Methodology: Mustafa Hakkı Aydoğdu.
Investigation: Mustafa Hakkı Aydoğdu and Necla Yıldızoğulları.
Discussion of results: Mustafa Hakkı Aydoğdu and Necla Yıldızoğulları.
Writing – Original Draft: Mustafa Hakkı Aydoğdu and Necla Yıldızoğulları.
Writing – Review and Editing: Mustafa Hakkı Aydoğdu.
Resources: Mustafa Hakkı Aydoğdu and Necla Yıldızoğulları.
Supervision: Mustafa Hakkı Aydoğdu.
Approval of the final text: Mustafa Hakkı Aydoğdu and Necla Yıldızoğulları.

VI. REFERENCES

[1] H. S. Akduman, “İzmit'teki park ve bahçelerde yetiştirilen süs bitkileri,” Yüksek Lisans Tezi, Kocaeli Üniversitesi Fen Bilimleri Enstitüsü, Koaeli, Türkiye, 2014.

[2] S. Ay, “Süs bitkileri ihracatı, sorunları ve çözüm önerileri: Yalova ölçeğinde bir araştırma,” *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, vol. 14, no. 3, pp. 423–443, 2009.

[3] K. Aydinşakir, B. Sayın, M. A. Çelikyurt, and Ö. Karagüzel, “Antalya ili kesme çiçek perakendeciliğinin analizi,” *Akdeniz Üniversitesi Ziraat Fakültesi Dergisi*, vol. 27, no. 2, pp. 75–82, 2016.

[4] M. H. Aydoğdu and N. Yıldızoğulları, “Şanlıurfa'ki süs bitkileri üreticilerinin mevcut yapıları ve sektöre bakışları üzerine bir araştırma,” *Journal of Disciplines Economics & Administrative Sciences Studies*, vol. 7, no. 28, pp. 313–321, 2021.

[5] Y. Bulut, E. Akpınar, and H. Yılmaz, “Erzurum kentinin kesme çiçek tüketim potansiyelinin belirlenmesi ve çözüm önerileri,” *GOÜ Ziraat Fakültesi Dergisi*, vol. 24, no. 2, pp. 7–11, 2007.

[6] M. Cançelik, “A field research on the medical aromatic plant selling enterprises (Aktar/Attar) in Şanlıurfa,” in *Current Researches in Economics and Administrative Sciences*, Y. A. Unvan and İ. Serbestoğlu, Eds. Cetinje-Montenegro: Ivpe Cetinje, 2020, pp. 25–48.

[7] B. Cengiz, M. Sabaz, and M. Sarıbaş, “The use of some natural crataegus L. (Hawthorn) taxa from western Black Sea Region of Turkey for landscape applications,” *Fresenius Environmental Bulletin*, vol. 20, no. 3, pp. 938–946, 2011.

[8] B. Cengiz, P. K. Dağlı, and S. Yiğittekin, “Peyzaj ekonomisi açısından peyzaj ve süs bitkileri fidanlık işletmelerine yönelik sektörel bir analiz,” *Bartın Orman Fakültesi Dergisi*, vol. 19, no. 2, pp. 50–62, 2017.

[9] D. Çiçekdemir, “Türkiye'deki süs bitkileri pazar payının araştırılması: Bursa ili örneğinde,” Yüksek Lisans Tezi, Namık Kemal Üniversitesi Fen Bilimleri Enstitüsü, Tekirdağ, Türkiye, 2010.

[10] A. Eşitken, Y. Çelik, A. T. Polat, and Z. Karakayacı, “Konya'da dış mekân süs bitkileri, iç mekân süs bitkileri, kesme çiçekler ve çiçek soğanları yetiştiriciliği yatırımlarına yönelik fizibilite çalışması,” 2012. https://www.kalkinmakutuphanesi.gov.tr/assets/upload/dosyalar/dsyfme8f4720201712200pm.pdf (accessed May 05, 2021).

[11] B. Gençer, “Dünya'da ve Türkiye'de kesme çiçek sektörü pazarlama organizasyonları ve tüketici eğilimleri,” Doktora Tezi, Namık Kemal Üniversitesi Fen Bilimleri Enstitüsü, Tekirdağ, Türkiye, 2014.

[12] O. Karagüzel *et al.*, *Sekizinci beş yıllık kalkınma planı bitkisel üretim özel ihtisas komisyonu (süs Bitkileri alt komisyon raporu)*. Ankara,Turkey: DPT: 2645-ÖIK:653, 2000.

[13] M. K. Kıymaz, “Dış mekân süs bitkileri üretim işletmelerinde maliyetleme ve raporlama,” Doktora Tezi, Sakarya Üniversitesi Sosyal Bilimler Enstitüsü, Sakarya, Türkiye, 2017.

[14] İ. Kızılkan, “Türkiye'de süs bitkileri üretim ve pazarlama sektörünün sorunları ve çözüm önerileri,” Yüksek Lisans Tezi, Ankara Üniversitesi Fen Bilimleri Enstitüsü, Ankara,Türkiye, 2016.

[15] H. A. Onay, “Türkiye'de süs bitkileri üretim ve pazarlama sektörünün sorunları ve çözüm önerileri,” Doktora Tezi, Ankara Üniversitesi Fen Bilimleri Enstitüsü, Ankara,Türkiye, 2008.

[16] İ. Özkan, “Süs bitkileri yetiştiriciliğinde damla sulama uygulamalarında karşılaşılan sorunlar ve çözüm önerileri,” Yüksek Lisans Tezi, Namık Kemal Üniversitesi Fen Bilimleri Enstitüsü, Tekirdağ, Türkiye, 2009.

[17] (SÜSBİR) Süs Bitkileri Üreticileri Birliği, “Süs bitkileri sektör raporu,” 2019. http://www.susbir.org.tr/yeni/belgeler/raporlar/susbir-sektor-raporu-2019.pdf (accessed May 04, 2021).

[18] M. Temel, “Yalova’da süs bitkileri işletmelerinin yapısı, üretim olanakları, sorunları ve çözüm önerileri,” Yüksek Lisans Tezi, Namık Kemal Üniversitesi Fen Bilimleri Enstitüsü, Tekirdağ, Türkiye, 2011.

[19] İ. H. Ünlü, “Şanlıurfa kent dokusuna uygun karasal iklim koşullarına dayanıklı çok yıllık süs bitkilerinin belirlenmesi,” Yüksek Lisans Tezi, Mustafa Kemal Üniversitesi Fen Bilimleri Enstitüsü, Hatay, 2014.

[20] K. Yazıcı, “Tokat ili ekolojik koşullarında dış mekân süs bitkileri sektör analizi ve uygun üretim alanlarının CBS ile belirlenmesi,” Doktora Tezi, Ege Üniversitesi Fen Bilimleri Enstitüsü, İzmit, Türkiye, 2015.

[21] N. Yıldızoğulları, “Süs bitkileri potansiyelinin belirlenmesi: Şanlıurfa ili örneği,” Yüksek Lisans Tezi, Harran Üniversitesi Fen Bilimleri Enstitüsü, Şanlıurfa, Türkiye, 2020.

[22] A. Çelik, E. Torun, and A. Bobat, “Süs bitkilerine ilgi gösteren kişilerin genel profilin belirlenmesi,” Erdemli-Mersin, Türkiye, Oct. 2010, pp. 339–347.

[23] M. H. Aydoğdu, “Evaluation of willingness to pay for irrigation water: Harran Plain sampling in GAP region -Turkey,” *Applied Ecology and Environmental Research*, vol. 14, no. 1, pp. 349–365, 2016.

[24] G. Sevinç, M. H. Aydoğdu, M. Cançelik, and M. R. Sevinç, “Farmers’ attitudes toward public support policy for sustainable agriculture in GAP-Şanlıurfa, Turkey,” *Sustainability*, vol. 11, no. 23, p. 6617, 2019.

[25] H. P. Doğan, M. H. Aydoğdu, M. R. Sevinç, and M. Cançelik, “Farmers’ willingness to pay for services to ensure sustainable agricultural income in the GAP-Harran Plain, Şanlıurfa, Turkey,” *Agriculture*, vol. 10, no. 5, p. 152, May 2020.

[26] TURKSTAT, “Adrese dayalı nüfus kayıt sistemi sonuçları, 2020,” 2021. https://data.tuik.gov.tr/Bulten/Index?p=Adrese-Dayali-Nufus-Kayit-Sistemi-Sonuclari-2020-37210 (accessed May 05, 2021).

[27] Y. Yazıcıoğlu and S. Erdoğan, *SPSS uygulamalı bilimsel araştırma yöntemleri*. Ankara, Türkiye: Detay Yayıncılık, 2014.

[28] N. Bayram, *Sosyal bilimlerde SPSS ile ileri veri analizi*. Ankara, Türkiye: Ezgi Kitapevi, 2014.

[29] Ş. Kalaycı, *SPSS Uygulamalı çok değişkenli istatistik teknikleri*. Ankara, Türkiye: Asil Yayın Dağıtım, 2014.

[30] M. Mert, *SPSS STATA yatay kesit veri analizi bilgisayar uygulamaları*. Ankara: Detay Yayıncılık, 2016.

[31] Ş. Koç and M. Şahin, “Tobit model ve bir uygulama,” *KSÜ Doğa Bilimleri Dergisi*, vol. 21, no. 1, pp. 73–80, 2018.

Journal of Engineering and Technology for Industrial Applications

ISSN ONLINE: 2447-0228

ITEGAM-JETIA

Manaus, v.7 n.29, p. 52-61. May/Jun, 2021

DOI: https://doi.org/10.5935/jetia.v7i29.740

RESEARCH ARTICLE OPEN ACCESS

FRAMEWORK FOR TOTAL PRODUCTIVE MAINTENANCE FOR AN SME

Norman Gwangwava*[1], Goabaone A. Baile[2], Pageal Dikgale[3] and Ketsile Kefhilwe[4]

[1, 2, 3, 4] Mechanical, Energy & Industrial Engineering, Boswana International University of Science and Technology, Palapye, Botswana.

[1] http://orcid.org/0000-0002-1970-2171, [2] http://orcid.org/0000-0002-4370-8119, [3] http://orcid.org/0000-0003-3478-212X, [4] http://orcid.org/0000-0003-2634-7064

Email: *gwangwavan@biust.ac.bw, amosbaile@gmail.com, dikgalepageal@gmail.com, 212koffikeita@gmail.com

ARTICLE INFO

Article History
Received: March 09th, 2021
Accepted: May 26th, 2021
Published: June 30th, 2021

Keywords:
Total Productive Maintenance,
Overall Equipment Effectiveness,
Performance Index,
Maintenance Management,
Six TPM Losses.

ABSTRACT

Total Productive Maintenance (TPM) is a worldwide maintenance program for plant equipment and machinery. It is a proactive maintenance that optimizes Overall Equipment Effectiveness (OEE) to get rid of the six (6) big losses. The research was inspired by observations made over a six (6) months period at Company X, a Small and Medium Enterprise (SME). Observations made were that the SME produced a lot of defective items, some requiring reworks, the SME had poor maintenance plans, poor relationship between departments, resulting in low morale of workers, frequent breakdowns of machines, as well as low production rate. The main objective of the research was to design a framework that would identify and address the aforementioned problems, resulting in an optimized OEE rate. The research data was obtained through various methodologies, including observation, questionnaire as well as interviews among the company employees - technicians, operators, and maintenance engineers. The data was analysed using Microsoft excel performance dashboards as well as TPM templates. Based on the findings, a framework was designed and developed that sought to address the aforementioned problems at the company. The study improved the OEE of machines and processes through the implementation of TPM approach at Company X.

I. INTRODUCTION

Many SMEs and other established companies experience a lot of scrap due to machine failure or equipment breakdowns and frequent production stops which affects productivity. Frequent breakdown or slow running of equipment and machines could be reduced by the use of Total Productive Maintenance (TPM) philosophy. Case study Company X, a Small and Medium Enterprise (SME) in the manufacturing sector has been faced with producing a lot of defective items, poor maintenance plans, poor relationship between departments, frequent breakdown of machines, as well as low production rate.

The research intends to develop a framework for TPM implementation in the SME manufacturing sector, focusing on Company X. Other interventions involve creating a working environment that is clean and well-organized using the 7S approach, as well as evaluating OEE factors on the TPM tool recommended for the company. The research achieves its purpose by focusing on improving maintenance methods and reducing setup times, job changeovers, downtimes, scraps as well as waste. These operations are sectioned by the quality and maintenance departments of the company.

The TPM approach operates according to the thinking that anybody in a facility has to take part in maintenance alternatively, than just the maintenance team. TPM is defined as a methodology created with the aim of increasing production efficiency by implementing efficient equipment maintenance, having a methodical nature of implementation (TPM pillars), which promotes the involvement of all employees as a way to increase their sustainability and effectiveness [1], [2]. TPM implementation is a change management process and its goal is to improve core business processes. Among things emphasised by TPM are proactive and preventative protection in order to maximize the operational affectivity of equipment. TPM puts emphasis on empowering operators to assist preserve their equipment.

Proper and regular maintenance of production equipment and machines results in high operational efficiency of equipment (fewer breakdowns, few production stoppages, few defects, and zero accidents), hence increased production [3]. Productivity in manufacturing is very important as it measures how efficiently production inputs such as labour and capital are being used in an organisation to produce a certain level of output. Productivity increase is a target for many manufacturing firms because as it increases it allows the firm to make more products without increasing the costs. This helps manufacturing firms to benefit from economies of scale. Almost all industrial manufacturing processes are carried out with the useful resource of machines, as a result of which each production-oriented company is mostly structured on its machinery.

Maintenance downtime is included in manufacturing scheduling, and in many cases, turns into a critical part of the manufacturing process [4]. TPM assigns the responsibility for preventive and routine maintenance to the same humans who operate that character equipment. The result is that the human beings most familiar with the specific machines are put in charge of the machine's care. TPM is constructed on the 5S foundation, which creates high-quality administrative centre company and standardized methods to enhance safety, quality, productivity and employee attitudes.

II. LITERATURE REVIEW

II.1 TOTAL PRODUCTIVE MAINTENANCE (TPM)

TPM is a broad maintenance program which includes an idea for keeping up plant and equipment. A typical TPM program in an organisation expands employee morale and employment fulfilment, thus bringing maintenance at the center as a vital and indispensably critical part of the business. TPM implementation methodology follows a prescriptive process in which the steps that companies must follow to implement the process and achieve intended benefits are detailed [5]. Organisations ultimately achieve sustainable and profitable growth in the long term by monitoring three – Availability, Performance, and Quality. These three key parameters are a direct contribution from each machine of the overall production system and the efficiency of the operators responsible for the machines [6]. Keeping a health check on the three parameters lead to optimal levels for productive plant efficiency.

TPM strategy help improve the competitiveness and economic benefits of a manufacturing or service organization [7]. According to [8], a program to turn the staff into TPM workers under the slogan "My machine my responsibility, I receive well, I deliver well" increased the production standards on average of 5%. There is shared accountability for tools that encourage higher involvement by plant floor workers when TPM is implemented.

TPM philosophy closely resembles another popular Japanese philosophy, Total Quality Management (TQM). The two programs share many common tools such as employee empowerment, benchmarking, documentation, etc, for implement and optimization [9]. Some similarities between the two programs are stated below:

1. Upper level management should show total commitment to the program in both TPM and TQM
2. Employees must be empowered to initiate corrective action, and
3. TPM and TQM require a long range outlook as the programs may take a year or more to implement and are an on-going process. Change management is required as well.

Notable differences between TQM and TPM are summarized in Table 1 below.

Table 1: Differences between TQM and TPM.

Category	TQM	TPM
Goal	Quality (Output and effects)	Equipment (Input and cause)
Means of attainment	Systematize the management. It is software oriented	Employees participation and it is hardware oriented
Target	Quality for PPM	Elimination of losses and wastes.

Source: Authors, (2021).

Overall TPM objectives include:

- Maintain an accident free environment,
- Increasing the operator involvement,
- Maximizing the Reliability of machine,
- Improving the Quality and Reduce cost,
- Focus on Maintainability engineering,
- Improving trouble solving via team,
- Upgrading every operator,
- Motivating the operator,
- Increasing the OEE.

The term PQCDSM is used to denote TPM performance indicators. These are known as productivity (P), quality (Q), cost (C), delivery (D), safety (S) and morale (M) [10]. Specific targets for the key performance indicators (KPIs) [11] are provided in Table 2.

Table 2: TPM key performance indicator targets.

KPI	Target(s)
P – Productivity	80% minimum OPE (Overall Performance Efficiency). 90% OEE (Overall Equipment Effectiveness)
Q – Quality	90% reduction in process defect rate. 75% reduction in customer returns/claims.
C – Cost	30% production costs reduction.
D – Delivery	50% reduction in finished goods and Work in Progress (WIP).
S – Safety	Zero shutdown accidents. Zero pollution incidents. Zero accident environment
M – Morale	5 to 10 times up employee improvement suggestions. Develop Multi-skilled and flexible workers.

Source: Authors, (2021).

II.I1 ORGANIZATION STRUCTURE FOR TPM IMPLEMENTATION

The procedure for introducing TPM in an organization undergoes through four major stages – Stage A: Preparatory Stage, Stage B: Introduction Stage, Stage C: Implementation, and Stage D: Institutionalisation [12]. OEE implementation starts with management awareness of total productive manufacturing and their commitment to focus the factory work force on training in teamwork and cross-functional equipment problem solving. Details of each stage are given below [12].

STEP A - PREPARATORY STAGE:

STEP 1 - Announcement by Management to all about TPM introduction in the organization:

Proper understanding, commitment and active involvement of the top management is needed for this step. Senior management should have awareness programmes, after which announcement is made to all. The awareness is done through in-house publications, announcements, and display on notice boards.

STEP 2 - Initial education and propaganda for TPM:

Training is to be done based on the need. Some individuals need intensive training and some just an awareness. Take TPM mainline personnel to places where TPM is already successfully implemented.

STEP 3 - Setting up TPM and departmental committees:

Committees should cater for TPM pillars and related needs.

STEP 4 - Establishing the TPM working system and target:

Each area is benchmarked and targets for achievement set up.

STEP 5 - A master plan for institutionalizing:

Next step is implementation leading to institutionalizing wherein TPM becomes an organizational culture. Achieving PM award is the proof of reaching a satisfactory level.

STEP B - INTRODUCTION STAGE

This is a ceremony and all should be invited. Suppliers should be made aware of the demand for quality from them. Related companies and affiliated companies who are potential customers are also invited. Some may learn from the new TPM certified company and some can help achieve more. Customers receive the communication that their supplier cares for quality output.

STAGE C - IMPLEMENTATION

In this stage eight activities are carried which are called eight pillars in the development of TPM activity. Of these four activities are for establishing the system for production efficiency, one for initial control system of new products and equipment, one for improving the efficiency of administration and are for control of safety, sanitation as working environment.

STAGE D - INSTITUTIONALISING STAGE

By all the activities one would has reached maturity stage. Now is the time for applying for PM award. The company should also think of challenging levels to which TPM movement can be taken. Figure 1 shows the plant-wide TPM structure.

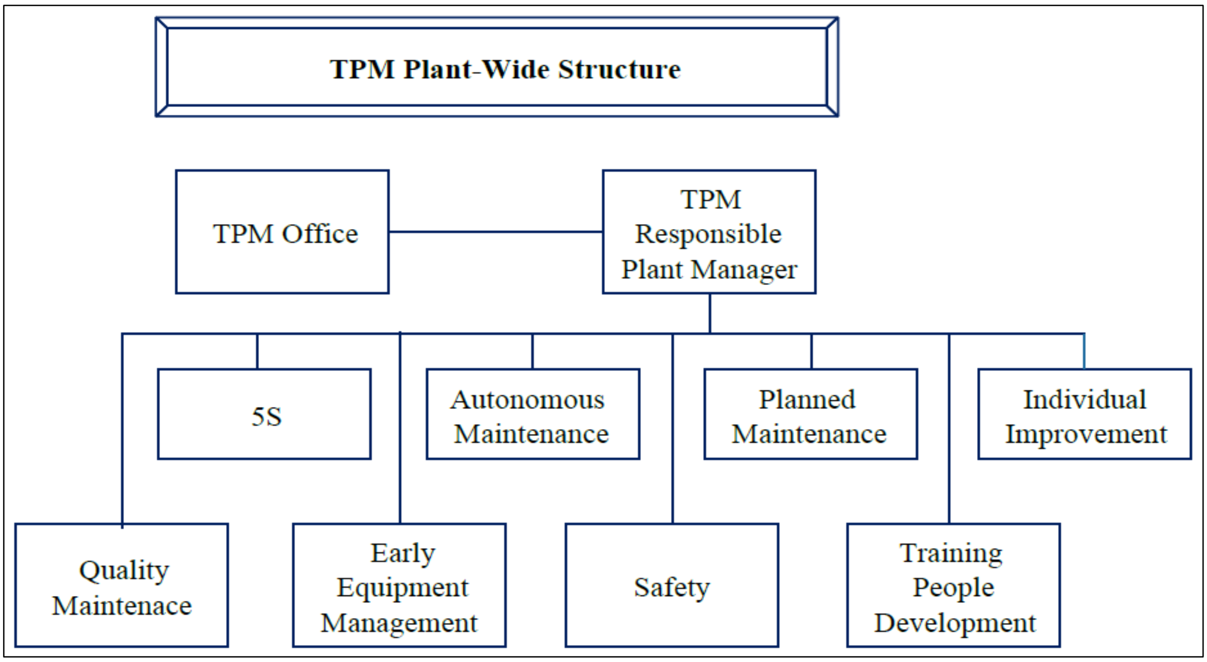


Figure 1: TPM company-wide structure.
Source: Author redrawn from [13].

II.1I 8 PILLARS OF TPM

TPM includes eight supporting activities that are centred on proactive and preventative techniques for enhancing plant and equipment reliability. These supporting activities are listed below [14]:

- Autonomous Maintenance
- Process and Machine Improvement.
- Preventative Maintenance
- Early Management of New Equipment
- Process Quality Management
- Administrative Work
- Education and Training
- Safety and Sustained Success

Figure 2 illustrates the TPM house built on the 5S foundation and supported by 8 pillars.

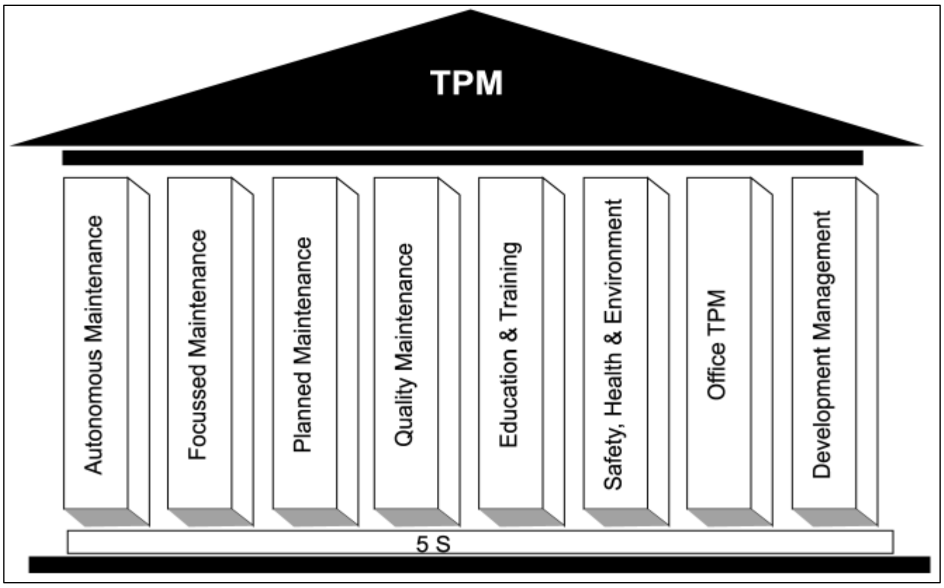


Figure 2: Eight pillars of TPM implementation (suggested by JIPM).
Source: [13].

II.1II 7 S

5S is the foundation of TPM. An organisation establishing a TPM program will center its attention on establishing the 5S basis and creating an autonomous maintenance plan [15]. The 5s system derives its name from the five Japanese words which define the process: seiri, seiton, seiso, seiketsu and shitsuke. The English translation is: sort, set in order, shine, standardize and sustain. Organization, cleanliness and standardization are the guiding principles behind the 5S system. This results in overall workplace cleanliness, created by removing waste from the work area, promoting internal organization and enhancement of visual communication. Later the 5S methodology was then developed to include the "Safety" aspect and it was called 6S. Recently another S (Spirit/Support) was added to the 6S framework and this formed the latest 7S methodology. The 7th S, Spirit/Support, seeks to enhance team consistent cohesion, motivation, and cooperation from top down and up top in the Organizational hierarchy [16], [17]. Table 3 below shows the 7S methodology words and meaning.

Table 3: 7S Word and Definitions.

7S Word	Meaning
SORT	Dispose all useless and waste materials from the workplace
SET IN ORDER	Put everything in the appropriate places for quick access.
SHINE	Make sure that the workplace is clean and tidy at all times.
STANDARDISATION	Make a habit out of keeping order in the workplace at all times.
SUSTAIN	Practice 6S daily without ceasing.
SAFETY	Use all appropriate safety equipment and adhere to all safety requirements.
SPIRIT/SUPPORT	Willingness to cooperate as part of a team. An additional piece to make explicit the reliance on the people factor and the need to continually keep it in mind as other steps are undertaken.

Source: Authors, (2021).

The main objective of the 7S system is to grow the value added to each worker. To grow the added value, manufacturers must create ordered and well adjusted production lines based on the principles of the 7S's: organization and order. A way to improve precision for the entire production and maintenance work in the factory is to give people an extremely organized work environment where a big part of their work is controlled visually. The visual workplace is an ideal one with no defects and no anomalies. The first step in creating a visual workplace is the 7S organization.

II.1V OVERALL EQUIPMENT EFFECTIVENESS (OEE) AND SIX BIG LOSSES

OEE (Overall Equipment Effectiveness) is a metric that is at the core for measuring manufacturing productivity in TPM [18], [19]. It identifies the percentage of planned manufacturing time that is used for actual production. OEE was developed to help TPM programs accurately track productive and strive to attain "perfect production". An achievement of 100% OEE rating means an organisation is achieving 100% Quality (produces good parts only), 100% Performance (Quickest rate), and 100% Availability (Plants runs non-stop). This means it's almost unrealistic to achieve 100% OEE. TPM has the standards of 90% Availability, 95% Performance efficiency, and 99% Quality rate [20]. The overall goal of TPM is to raise the OEE measure, with 85% being the world class measure.

There are six equipment losses identified within TPM that are used to calculate your OEE [21], [22]:

Availability

- Unplanned stops.
- Setup and Adjustments.

Performance

- Small stops.
- Slow running.

Quality

- Defects.
- Reduced yield.

These losses must be routinely observed, studied, computed, and analysed graphically so that the organisation can understand and monitor OEE. Employees must be trained so that they acquire ability to identify and prioritize losses, through practical sessions involving seven steps [23]. In order to effect continuous improvement, an organisation must compare the expected and current OEE measures. This picture will give the organisation the drive to improve its maintenance policy. OEE consists of three underlying components, each of which maps to one of the TPM dreams set out at the start of this topic, and every of which takes into account a distinctive kind of productivity loss [12].

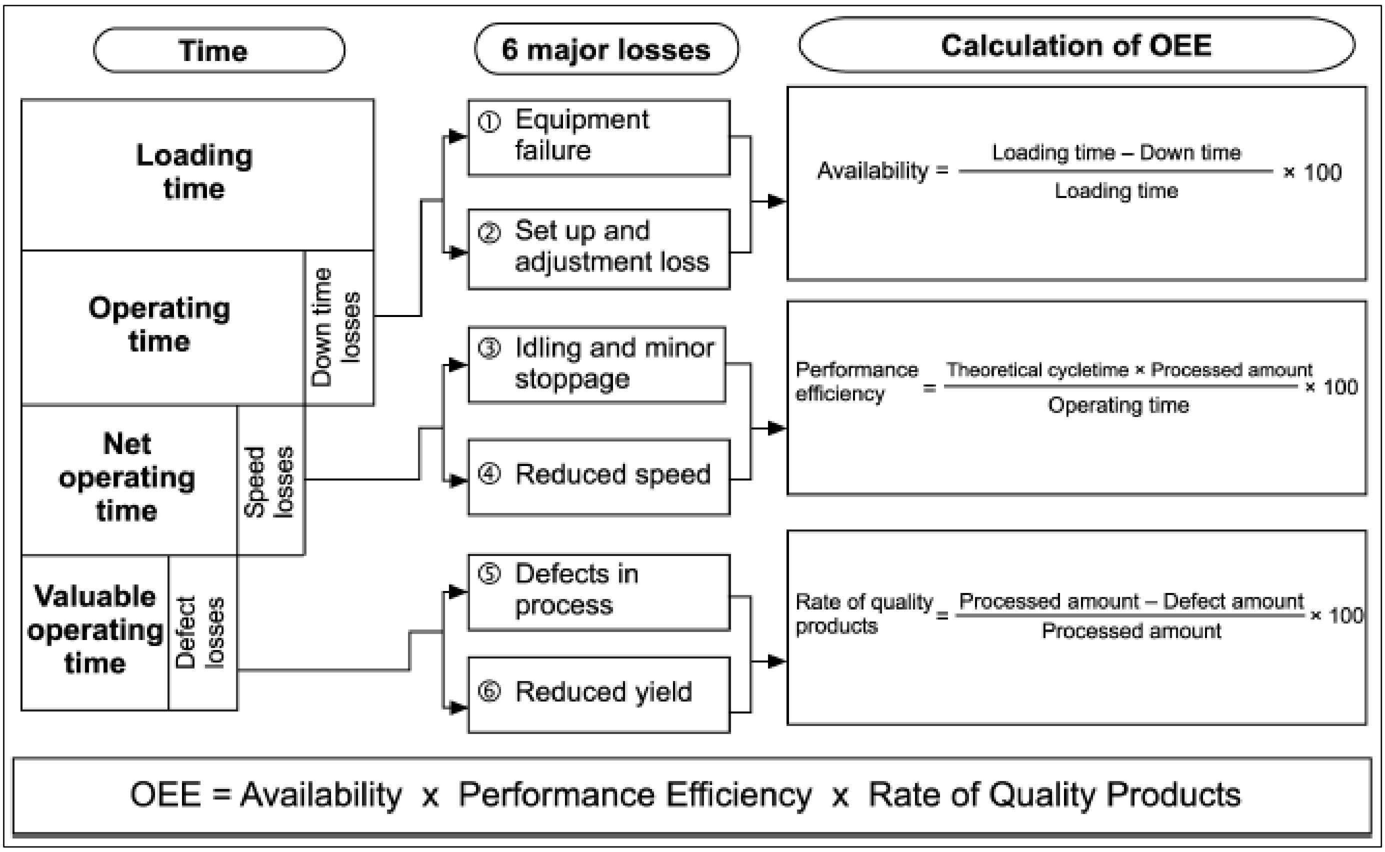


Figure 4: Calculation of OEE based on six major production. losses.
Source: [13] and [24].

OEE is calculated by obtaining the product of availability of the equipment, performance efficiency of the process and rate of quality products [25-27], [29]. The OEE calculation process is depicted in Figure 4 and also summarised below:

OEE=Availability (A)×Performance efficiency(P)×Rate of quality (Q); (1)

Availability refers to the ratio of loading time minus downtime and loading time. Performance refers to the ratio of processed amount times theoretical cycle time and operating time.

Quality refers to the ratio of processed amount minus defect amount and processed amount. The calculations of Availability, Performance and Quality are as follows:

Availability (A) = [(Loading time-Downtime) ÷ Loading time] × 100, (2)

Loading time refers to the running time after the removal of planned activities that affect production [30].

Performance efficiency (P) = [Processed amount ÷ (Operating time ÷ Theoretical cycle time)] × 100, (3)

Theoretical cycle time refers to the shortest cycle time that can be achieved under optimal conditions [30].

Rate of quality (Q) = [(Processed amount-Defect amount) ÷ Processed amount] ×100 (4)

III. MATERIALS AND METHODS

II.1 METHODOLOGY

The study uses field research through data collection over a horizon of 6 months in which the information is condensed in data tables. An SME in automotive manufacturing is used as a case study for the quantitative research approach as an initiative to innovate their processes and research new technologies that help increase their productivity.

An open-ended questionnaire was used for data acquisition inside the firm along with a series of interviews, observations on the production process, and monitoring the machines or equipment. The interview process was done by asking directly to the related stakeholders at the company. The questions used in the questionnaire and interviews were based on knowledge of lean manufacturing principles, production time per unit, bottleneck activity, steps to distribute load at bottleneck, automation level, quality control measure in the firm, industry layout, machine downtime, repair time, maintenance policy, etc.

Follow up questions were asked further which were strictly based upon the responses of the participants. Based on these responses conclusions were drawn through current OEE performance of the SME. By analyzing the current OEE performance and maintenance practices, the state of the firm was determined and then studies were conducted for the implementation of the TPM concept through an appropriate model for SMEs. Secondary data were obtained through a company audit so as extract historical data for the firm, such as downtime, the amount of production, the number of defects, non-productive time, the amount of damage to the machine, the standard repair time, product prices, component costs, and labor costs.

Calculations begin by finding OEE values comprising of three factors - availability, performance, and quality values. The three values are compared with world-class standard values to see the most significant factor. The next step is calculations for the six big losses to find out the big mistakes that impact on availability, performance, and quality.

Evaluation of TMP strategies and general maintenance policy for the SME was carried out in order to overcome the problem of low OEE values that did match with world-class standards. The overall research methodology used in the study is shown in Figure 5.

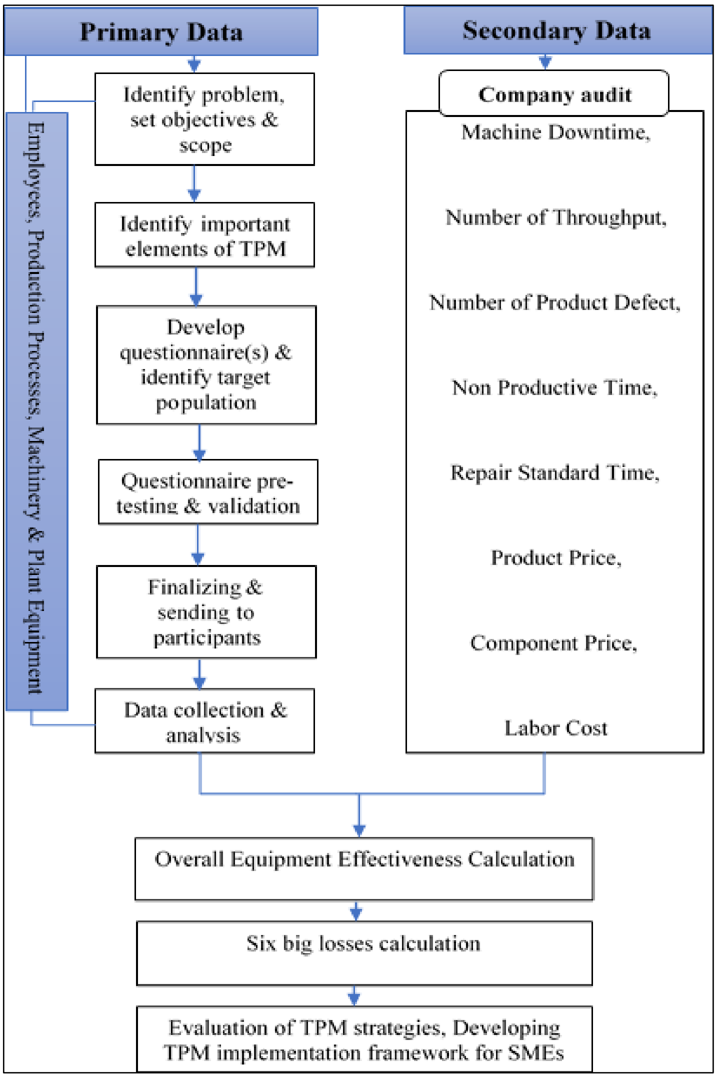


Figure 5: Research methodology.
Source: Authors, (2021).

III.I1 COMPANY PROFILE

The case study company is an SME automotive parts manufacturing company. The company has a fixed production irrespective of market demand. The data collected is tabulated in Table 4 below.

Table 4: Production line profile.

Psudo Company Name	Company X
Age of the company	16 years
Number of employees	36
Number of processes	11
Planned production time	570 minutes per day (inclusive of breaks)
Run time (Available production time)	500 minutes per day (excludes breaks)
Lead time	10 days
Ideal cycle time	1.5
5S foundation	Implemented
Scheduled maintenance	60 minutes duration

Source: Authors, (2021).

II.1II PROBLEMS IDENTIFIED

Based on the methodology used, the problems which led to various types of wastes in the company were identified and listed below:

1. Excess inventory - raw materials, work-in-process, finished goods,
2. Improper management of inventory and tools,
3. Industry works on push system,
4. Delay in the shipment of the orders,
5. Low level of automation,
6. Outdated machinery increases the level of pollution in the firm's environment,
7. Machinery is outfitted for product (consumes too much energy, huge and bulky),
8. Frequent breakdowns of machines,
9. Unbalanced production line,
10. Low production rate,
11. No proper movement of the workers and goods,
12. Improper utilization of floor space,
13. Loading and unloading of raw material and finished goods is a slow process due to space constraint,
14. Lack of commitment from top management,
15. Work attitude from middle management, which is supervisors etc.
16. Lack of dedication by shopfloor workers,
17. Poor relationship between departments, resulting in low morale of workers,
18. Safety measures are inadequate.

IV. ANALYSIS OF DATA AND IDENTIFIED SOLUTIONS

IV.1 OEE CALCULATION

Table 5 represents a seven days sample data set used to calculate Availability, Performance Efficiency and Rate of Quality values for the company. the average availability value of 82.56% with values ranging from 65.22%-87.35%. The average performance value is 90.83%, ranging from 65.22%-87.35%, and the average quality value is 95.04%, ranging between 91.15%-97.67%. Table 6 shows availability, performance, quality, and OEE values over the six months period between March – August 2019.

Table 5: TPM Dataset over seven days.

Item	D1	D2	D3	D4	D5	D6	D7
Shift (min)	570	570	570	570	570	570	570
Breaks	T(20) L(50)	T(20) L(50)	T(20) L(50)	T(20) L(50)	T(20) L(50)	T(20) L(50)	T(20) L(50)
Planned production time	500	500	500	500	500	500	500
Downtime	35	30	40	36	40	35	45
Run time	465	470	460	464	460	465	455
Ideal cycle time	1.5	1.5	1.5	1.5	1.5	1.5	1.5
Total count	450	490	450	400	420	400	400
Rejects count	16	20	13	10	9	10	7
Good count	434	470	437	390	411	390	393

Source: Authors, (2021).

Table 6: Availability, Performance Efficiency, Quality Rate and OEE values over six moths.

Month	Availability (A)	Performance Efficiency (P)	Quality Rate (Q)	OEE
March	81.32	75.33	96.56	59.15
April	77.54	70.56	90.23	49.37
May	63.33	68.89	86.65	37.80
June	87.43	80.12	83.78	58.69
July	74.61	69.87	91.32	47.61
August	70.73	73.43	84.77	44.03
Average	**75.83**	**73.03**	**88.89**	**49.44**
World class	**>=90**	**>=95**	**>=99**	**>=85**

Source: Authors, (2021).

The company performance shows relatively low values against the world class standards. This is caused by equipment failures, idling, minor stoppages, and reduced yield. It can be seen that the OEE value is far below the world-class standard. Figure 6 shows the graph comparing actual compny performance against the world class performance metrics. Corresponding six big losses will help to expose the ultimate causes of low company performance.

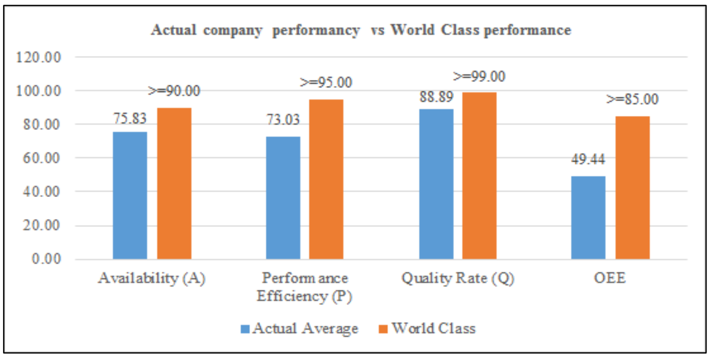


Figure 6: Actual company performance vs World Class performance.

Source: Authors, (2021).

IV.1 6 BIG LOSSES CALCULATION

The next effort after calculating OEE is to identify six big losses factors. The factors are grouped into Avilability (A), Performance Efficiency (P), and Quality (Q). The data obtained from the company is shown on a graph in Figure 7. The graph shows that unplanned stops contribute the largest loss factor of 19.61%. This heavily impacts on plant and equipment availability. This indicates that the company needs a sound maintenance strategy in order to boost availability.

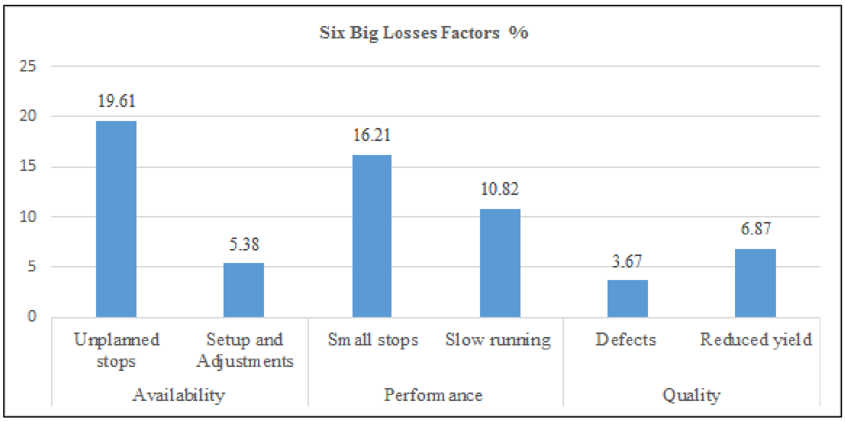


Figure 7: Six big losses.

Source: Authors, (2021).

IV.1 TPM IMPLEMENTATION FRAMEWORK

Based on various claims from literature, TPM has strong stern effects in manufacturing performance. Some case studies have proved that successfully implementing TPM brings out invaluable impacts to the overall performance of the organization or a company. TPM has shown significant improvements of the ranges 30-40% improvement in Overall Equipment Effectiveness, a 45% improvement in manufacturing output, 55-75% reduction in accidents as well as 70-80% reduction in defects & rework, 15% reduction in power costs as well as 75% reduction in breakdowns, downtimes [31-33].

Considering the above benefits, TPM was proposed as a tool for improving OEE and associated metrics for the case study company. A ladder model approach is proposed as a suitable framework for the company as illustrated in Figure 8.

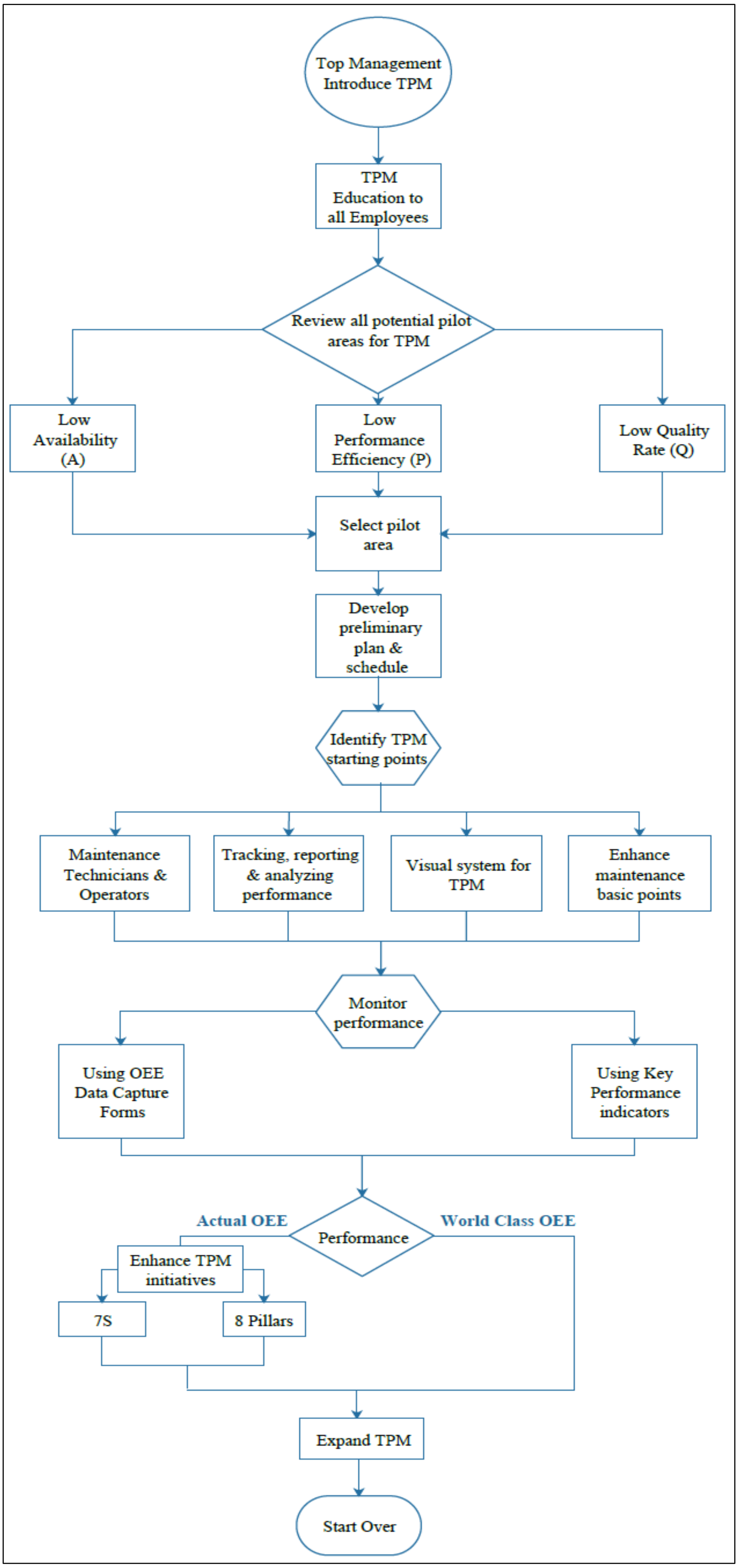


Figure 8: TPM framework.
Source: Authors, (2021).

V. CONCLUSIONS

Overall Equipment Effectiveness (OEE) value for the case study firm was 49.44% over the six months period of observation. Values for availability, performance, and quality over the same period are 75.83%, 73.03%, and 88.89% respectively. All the metrics rank well below the world class standards. The company did not have TPM in place at the time of performance measurement. Further analysis of the six big losses revealed that unplanned stops constituted the highest loss factor of 19.61%, followed by small stops (16.21%), slow running (10.82%), reduced yield (6.87%), setup & adjustments (5.38%), and defects (3.67%). Since the biggest loss contributor affected plant availability, an efficient maintenance strategy is being recommended. We propose a ladder model TPM framework suitable for the manufacturing based case study company. The framework emphasize top management approach, company-wide education of TPM philosophy, prioritization of specific plant equipment,and starting points. Performance is monitored using TPM data capturing forms and computing contributing performance metrics. Comparison against world class performance is emphasised so as to gain drive for improvement. 7S and 8 pillars are recommended bases for company-wide TPM enhancement. The whole framework views TPM strategy as a tool for continuous improvement, hence the last stage is prescribed as a 'start over' phase. The framework helps the company to expand the TPM program across all the processes, as well as stive for world class performance.

VI. AUTHOR'S CONTRIBUTION

Conceptualization: Norman Gwangwava.
Methodology: Norman Gwangwava.
Investigation: Goabaone A. Baile, Pageal Dikgale, and Ketsile Kefhilwe.
Discussion of results: Norman Gwangwava, Goabaone A. Baile, Pageal Dikgale, and Ketsile Kefhilwe.
Writing – Original Draft: Norman Gwangwava.
Writing – Review and Editing: Norman Gwangwava.
Resources: Norman Gwangwava, Goabaone A. Baile, Pageal Dikgale, and Ketsile Kefhilwe.
Supervision: Norman Gwangwava.
Approval of the final text: Norman Gwangwava, Goabaone A. Baile, Pageal Dikgale, and Ketsile Kefhilwe.

VII. ACKNOWLEDGMENTS

Special acknowledgements go to the case study company for providing their plant for TPM research work.

VIII. REFERENCES

[1] E. de Souza, M. Fortes, and G. de Lima, "Application based on fuzzy logic to evaluate implementation of TPM in industries", JETIA, vol. 6, no. 22, pp. 35-41, Apr. 2020.

[2] N. E. Candra, A. Susilawati, and W. Setiady, "Implementation of total productive maintenance (TPM) to improve sheeter machine performance", In MATEC Web of Conferences 2017, Vol. 135, p. 00028). EDP Sciences.

[3] A. Shagluf, A. P. Longstaff, S. Fletcher, "Maintenance strategies to reduce downtime due to machine positional errors", InProceedings of Maintenance Performance Measurement and Management (MPMM) Conference 2014. Imprensa da Universidade de Coimbra.

[4] O. M. Enofe and G. Aimienrovbiye, "Maintenance impact on production profitability-A case study", 2010.

[5] Villegas, L., & Velez, A. (2014). Implantation of Total Productive Maintenance-TPM-scenarios of corporative fusion. 12th Latin American and Caribbean Conference for Engineering and Technology. "Excellence in Engineering to Enhance a Country's Productivity". Guayaquil, Ecuador.

[6] Matos, R. (2012). Development of a Total Productive Maintenance (TPM) Program in the Special mixture's area of a Miller. *Faculty of Engineering Industry. "Engineering Magazine",* 19(3), 66.

[7] Reyes, J., Alvarez, K., Martínez, A., & Guamán, J. (2018). Total productive maintenance for the sewing process in footwear. Journal of Industrial Engineering and Management, 11(4), 814-822. https://doi.org/10.3926/jiem.2644

[8] S. Sahoo and S. Yadav, "Influences of TPM and TQM Practices on Performance of Engineering Product and Component Manufacturers", Procedia Manufacturing, vol. 43, pp. 728-735, Jan. 2020.

[9] Sharma R, Singh J, Rastogi V. The impact of total productive maintenance on key performance indicators (PQCDSM): a case study of automobile manufacturing sector. International Journal of Productivity and Quality Management. 2018;24(2):267-83.

[10] Diky Herdiawan, Erry Rimawan,. (2020). ANALYSIS OF THE PILAR TPM EFFECT ON PRODUCTIVITY TO IMPROVING THE COMPETITIVENESS OF TEXTILE INDUSTRIES IN INDONESIA. *International Journal of Advanced Science and Technology*, *29*(7s), 5256-5268. Retrieved from http://sersc.org/journals/index.php/IJAST/article/view/25933

[11] Wakjira, M. W., & Singh, A. P. (2012). Total productive maintenance: A case study in manufacturing industry. Global Journal of Research in Engineering, 12(1-G).

[12] Ahuja, I.P.S. and Khamba, J.S. (2008), "Total productive maintenance: literature review and directions", *International Journal of Quality & Reliability Management*. Vol. 25 No. 7, 2008 pp. 709-756. DOI 10.1108/02656710810890890

[13] I. Idris and R.A.Sari. Implementation of eight pillars the Total Productive Maintenance at Water Supply Company, in proceedings of 1st International Joint Conference Indonesia-Malasya-Bangladesh-Ireland, 2015.

[14] Georgescu D, Constantin M. 7s's system foundation of total productive maintenance house for lean organization. Scientific Bulletin of the" Petru Maior" University of Targu Mures. 2009;6:131.

[15] Joshi, A. A., 2015. A Review on Seven S (7S) as a tool of Workplace Organization. International Journal of Innovations in Engineering and Technology (IJIET), 6(2), pp. 19-25.

[16] Mahlaha, K., Sukdeo, N. & Mofokeng, V., 2020. A Lean 7S methodology framework to improve efficiency and organizational performance: A review study in an SME organization. Dubai, UAE, International Conference on Industrial Engineering and Operations Management.

[17] Muchiri, P., & Pintelon, L. (2008). Performance measurement using overall equipment effectiveness (OEE): literature review and practical application discussion. *International journal of production research*, *46*(13), 3517-3535.

[18] Borkowski, S., Czajkowska, A., Stasiak-Betlejewska, R., & Borade, A. B. (2014). Application of TPM indicators for analyzing work time of machines used in the pressure die casting. Journal of Industrial Engineering International, 10(2), 1-9.

[19] Nakajima, S., (1988) Introduction to Total Productive Maintenance, Productivity Press Inc, Portland, p.21.

[20] Sethia, C. S., Shende, P. N., & Dange, S. S. (2014). Total productive maintenance-a systematic review. *International journal for scientific research & development (8)*, 124-127.

[21] Filscha N, Meilily A, Hendy T. Total productive maintenance policy to increase effectiveness and maintenance performance using overall equipment effectiveness. Journal of Applied Research on Industrial Engineering. 2019;6(3):184-99.

[22] Khan, F.J., & Quazi, T.Z. (2014). Implementation of Kobetsu Kaizen pillar in Improving Overall Equipment Effectiveness of Machine. International Journal of Engineering Sciences & Research Technology, 3(7), 562-570.

[23] Prabowo HA, Suprapto YB, Farida F. The evaluation of eight pillars total productive maintenance (TPM) implementation and their impact on overall equipment effectiveness (OEE) and waste. Sinergi. 2018;22(1):13-8.

[24] Ljungberg, O. (1998), “Measurement of overall equipment effectiveness as a basis for TPM activities”, International Journal of Operations & Production Management, Vol. 18 No. 5, pp. 495-507.

[25] Dal, B., Tugwell, P. and Greatbanks, R. (2000), “Overall equipment effectiveness as a measure for operational improvement: a practical analysis”, International Journal of Operations & Production Management, Vol. 20 No. 12, pp. 1488-502.

[26] Tsarouhas, P. (2019). Improving operation of the croissant production line through overall equipment effectiveness (OEE). International Journal of Productivity and Performance Management, 68(1), 88-108. https://doi.org/10.1108/IJPPM-02-2018-0060

[27] Tian Xiang, Z,, & Jeng Feng, C. (2021). Implementing total productive maintenance in a manufacturing small or medium-sized enterprise. Journal of Industrial Engineering and Management, 14(2), 152-175. https://doi.org/10.3926/jiem.3286

[28] Singh, R.K., Clements, E.J., & Sonwaney, V. (2018). Measurement of overall equipment effectiveness to improve operational efficiency. International Journal of Process Management and Benchmarking, 8(2), 246-261. https://doi.org/10.1504/IJPMB.2018.090798

[29] Garza-Reyes, J.A., Eldridge, S., Barber, K.D., & Soriano-Meier, H. (2010). Overall equipment effectiveness and process capability measures. *International Journal of Quality & Reliability Management*, 27(1), 48-62. https://doi.org/10.1108/02656711011009308

[30] P. Gupta and S. Vardhan, “Optimizing OEE, productivity and production cost for improving sales volume in an automobile industry through TPM: a case study”, International Journal of Production Research, vol. 54, no.10, pp. 2976-2988, 2016.

[31] K. M. Joshi and D. V. Bhatt, “A modified TPM framework for Indian SMEs”, International Journal of Advanced Research in Engineering and Technology, vol. 9, no. 6, pp. 1-14, Nov. 2018.

[32] S. C. Nwanya, J. I. Udofia, and O. O. Ajayi, “Optimization of machine downtime in the plastic manufacturing”, Cogent Engineering, vol. 4, no. 1, pp. 1335444, Jan. 2017.

Journal of Engineering and Technology for Industrial Applications

ITEGAM-JETIA

Manaus, v.7 n.29, p. 62-68. May/Jun, 2021

DOI: https://doi.org/10.5935/jetia.v7i29.753

ISSN ONLINE: 2447-0228

RESEARCH ARTICLE OPEN ACCESS

EVALUATION OF DISTRIBUTION TRANSFORMERS BASED ON POWER QUALITY PARAMETERS IN FRONT OF DIFFERENT LOAD PROFILES

Salvador Paes Ribeiro Junior[1], Ricardo Chagas Tolentino[2], Marcio Zamboti Fortes*[3] and Guilherme Gonçalves Sotelo[4]

[1, 2, 3, 4] Federal University of Fluminense - UFF. Niterói, Rio de Janeiro, Brazil.

[1] http://orcid.org/0000-0002-8484-271X, [2] http://orcid.org/0000-0001-8284-8600, [3] http://orcid.org/0000-0003-4040-8126, [4] http://orcid.org/0000-0002-5393-7206

Email: salvadorjunior.rj@gmail.com, ricardo-tolentino@hotmail.com, *mzamboti@id.uff.br, gsotelo@id.uff.br

ARTICLE INFO

Article History
Received: May 14th, 2021
Accepted: June 23th, 2021
Published: June 30th, 2021

Keywords:
Power quality,
Power Transformers,
Electrical Equipment.

ABSTRACT

This article presents a study to monitor the power quality supplied by a power distribution concessionaire operating in the State of Rio de Janeiro, Brazil. Three electric power transformers of the same nominal power installed from a Brazilian company were selected as case studies. The data were collected and analyzed to evaluate the power quality according to the formulation regulated by the Electricity Distribution Procedures in the National Electric System - PRODIST, which is the Brazilian legislation in force for the energy distributors. The experiment consists of checking the supply voltage, the power factor, harmonic distortion, voltage unbalance, and frequency variation. Two of the three transformers analyzed showed a failure in the power quality supplied in the frequency variation indicator and all transformers failed in the Power Factor indicator. Thus, this research verifies the importance of monitoring power quality parameters to ensure that the energy obtained by consumers has sufficient characteristics for the proper functioning of electrical equipment.

I. INTRODUCTION

It is well known that the focus on power quality (PQ) is partly attributable to the fact that the electricity sector is being reformulated for the consumer market. Consumers are more demanding and the main product purchased is electricity. Consumers prefer to achieve energy with sufficient quality parameters at the lowest possible cost. In this case, both regulatory agencies, such as ANEEL in Brazil, as well as the market itself, encourage electric system operators to provide information on operating conditions or to provide details about events that have occurred and that have affected consumers [1].

“Power quality” is a term that has different meanings. From the customer's point of view, a PQ problem can be defined as any power condition that causes the equipment to malfunction or becomes unusable. From the concessionaire's point of view, PQ problems can be seen as non-compliance with several pre-established standards [2].

Still, in the understanding of PQ, the Institute of Electronics and Electronics Engineers (IEEE) defines power quality as "the concept of providing energy and adequate grounding for sensitive electronic equipment". Electrical equipment is significantly affected by PQ and can be severely affected if electrical energy is outside the standards [3].

Research shows that there are expenses associated with the loss of PQ. For more than a decade, the costs associated with disrupting manufacturing processes in the United States have been estimated at approximately $ 10 billion. At the same time, in Europe, the estimated cost associated with various interferences reaches around 1.5% of Gross Domestic Product (GDP) [1].

Still in this context, Petr et al. [4], presented an article to help engineers choose the best solution to improve the quality of energy in a specific energy network at the distribution level. The article also presents the characteristics of an active shunt harmonic compensator, which is a very modern PQ controller that can be used in several cases, or in combination with other controllers.

When searching the literature for works related to distribution transformers aimed at PQ, research such as Brida Neto [5] is found that sought to evaluate the loss of electrical energy through the conversion of medium and low voltage due to the insufficient size of the distribution transformer.

In the planning aspect of the power distributor, there is research that finds that the economic management of distribution transformers contributes to the economic planning of power utilities as they are abundant in electrical systems [6].

There are more specific researches, which are important for the knowledge of the study of transformers as is the case presented by Gonen [7]. In this work, it is possible to understand several issues, such as the applicability and characteristics of transformers, PQ, and smart grid concepts with the use of transformers.

Following this research plan, the present work aims to analyze the data collected for three transforming units of the same nominal power that differ by the characteristics of loading in the aspect of non-technical losses (NTL) and named as TR1, TR2, and TR3. This research analyzes performed and check the supply voltage, the power factor, harmonic distortion, voltage unbalance, voltage fluctuation, frequency variation, and short-term voltage variation.

The characteristics of these transformers are:

TR1 – Transformer with a less incidence of NTL (<10%).

TR2 - Transformer with a high incidence of NTL (> 30%).

TR3 - Transformer with an average incidence of NTL (Between 10% and 30%).

To evaluate de PQ index PRODIST parameters - Module 8 will be established in its 10 reviews, approved by the National Electric Energy Agency (ANEEL) [8].

In research carried out in academic databases, it is noticed that Brazil has a good contribution to this theme. The researched documents include Papers, Articles, Letters, Reviews, and Proceedings Paper published between the years 2015 and 2020.

The research was carried out based on the keyword "Energy". The result of this initial research placed Brazil in the 15th position in the ranking of the countries that most publications on the subject. Brazilians are responsible for 2.524% of publications, thus confirming the concern and interest that this topic arouses among Brazilian researchers.

Then, the word "quality" was added to the search field and the previous settings were repeated. Thus, Brazil rose to the eleventh position, and the percentage of documents published to the world increased slightly. This became 3.582%.

Thus, it is hoped that the research presented in this text can contribute to Brazilian research and add to other international works related to the PQ issue in distribution systems since this problem is common to all.

II. THEORETICAL BASIS

II.1 SUPPLY VOLTAGE

According to Rocha [9], the steady-state voltage is assessed by obtaining two performance indicators: the relative transit time index (DRP) of the unstable voltage and the relative transit time index (DRC) of the critical voltage, where *nlp* and *nlc* indicate the maximum value between the phases of the number of measurements within the stable and critical range.

$$DPR = \frac{npl}{1008} x\ 100\% \quad (1)$$

$$DPC = \frac{npc}{1008} x\ 100\% \quad (2)$$

Santiago [10] highlighted the importance of analyzing voltage permanently. His research details a strategy for voltage control that involves nonlinear programming. Such research contributed to the stress profiles in the steady state, of his experiment, to be within the parameters allowed by PRODIST.

II.2 POWER FACTOR

According to PRODIST - Module 8 in its revision 10, the value of the power factor (PF) must be calculated according to the registered active and reactive power (P, Q) or the corresponding energy (EA, ER) using (3) and (4) [8].

$$FP = \frac{P}{\sqrt{P^2+Q^2)}} \quad (3)$$

Or

$$FP = \frac{EA}{\sqrt{EA^2+ER^2)}} \quad (4)$$

The reference values are: for the consumer unit or connection between distributors with a voltage lower than 230 kV, the PF at the connection point must be between 0.92 and 1.00 inductive or 1.00 and 0.92 capacitive, according to current regulations.

Knowledge of the PF is so important that distributors punish customers who demand high reactive power. This happens because the companies that make up the electric sector need to make greater investments in power generation equipment, expanding the capacity of the transmission and distribution lines, and transformers for transmission and conversion of reactive power. The voltage will affect the stability of the network [11].

II.3 HARMONIC DISTORTION FACTOR

According to PRODIST, harmonic distortion is a phenomenon related to the deformation of voltage and current waveforms concerning the sine wave of the fundamental frequency. The limits must be included between the intervals in Table 1.

Table 1: Limits of harmonic distortions (in % of the fundamental voltage).

Indicator	Nominal Voltage (kV)		
	Vn≤1.0	1.0<Vn<69	69≤Vn<230
DTT95%	10.00%	8.20%	5.00%
DTTp95%	2.50%	2.00%	1.00%
DTTi95%	7.50%	6.00%	4.00%
DTT395%	6.50%	5.00%	3.00%

Source: [8]

Harmonic distortion is a phenomenon related to the deformation of voltage and current waveforms concerning the fundamental sinusoidal frequency wave [8].

In this case, in the context of the PQ, phenomena related to harmonic distortion have been emphasized in recent years. Some of the main effects of voltage and current waveform distortion are heating equipment, reduced component life, overvoltage, reduced energy supplied by the equipment, malfunction of various equipment, impact on the PF, and interference telecommunications lines [12].

II.4 UNBALANCE MAGNITUDE OF SUPPLY VOLTAGE

Voltage unbalance is an indicator of PQ that has been extensively explored in research. [12] used this parameter to practically analyze some PQ disorders in a given power supply in the power distribution network.

According to [9] voltage unbalance is a disturbance related to the lack of quality in the supply of electricity and this unbalance can be caused mainly by the uneven distribution of single-phase loads in the system (such as electric arc furnaces and welding machines) and defects in equipment applied in the distribution, for example, fuse blown in the phase of the three-phase capacitor bank.

According to the Energy Distribution Program in the National Electric System (PRODIST), the following definition exists: "Voltage unbalance is a phenomenon that is characterized by any difference in amplitude between the three-phase voltages of a given three-phase system and/or 120º. The electrical difference between the phase voltages of the same system.

PRODIST - Module 8 in its revision 10 defines the unbalance factor as the relationship between the negative sequence voltage module and the positive sequence voltage module [8].

$$FD\% = \frac{V-}{V+}100\% \qquad (5)$$

where:

V- = Negative sequence voltage module.

V+ = Positive sequence voltage module.

FD% = unbalance factor.

However, since most measuring devices are limited in obtaining the voltage angle, (6) and (7) can also be used:

$$FD\% = 100\ \sqrt[2]{\frac{1-\sqrt[2]{3-6\beta}}{1+\sqrt[2]{3-6\beta}}} \qquad (6)$$

Being:

$$\beta = \frac{V_{ab}{}^{4} + V_{bc}{}^{4} + V_{ca}{}^{4}}{(V_{ab}{}^{2} + V_{bc}{}^{2} + V_{ca}{}^{2})^{2}} \qquad (7)$$

Where V_{ab} , V_{bc} e V_{ca} are line voltages modules.

And the unbalance voltage limits corresponding to FD95% are 3.0% for V_n less or equal 1.0 kV and 2.0% for V_n between the interval of 1 kV and 230 kV.

II.5 FREQUENCY

Frequency is an important parameter to assess the operational characteristics of electrical systems. According to PRODIST, the distribution system and the generation facilities connected to it must, under normal operating conditions and permanently operate within the frequency limits between 59.9 Hz and 60.1 Hz.

When disturbances occur in the distribution system, the generation facilities must ensure that the frequency returns, in the time interval of 30 (thirty) seconds after the transgression, to the range of 59.5 Hz to 60.5 Hz, for allowing the recovery of the load-generation balance.

If there is a need to cut generation or load to allow the recovery of the load-generation balance, during disturbances in the distribution system, the frequency:

a) may not exceed 66 Hz or be less than 56.5 Hz in extreme conditions;

b) can remain above 62 Hz for a maximum of 30 (thirty) seconds and above 63.5 Hz for a maximum of 10 (ten) seconds;

c) may remain below 58.5 Hz for a maximum of 10 (ten) seconds and below 57.5 Hz for a maximum of 05 (five) seconds.

Camporez et al. [13] considers that the frequency variation represents the frequency varies from a predefined value. The value is defined at 60 Hz and according to PRODIST - Module 8 in its revision 10, the system under normal conditions must operate between 59.9 and 60.1 Hz.

Bandeira et al. [14] used this parameter to analyze the electrical PQ of a photovoltaic system connected to the grid. The experiment allowed us to observe that the collected data were within the standards allowed by PRODIST.

The frequency shift is the deviation from the fundamental frequency value, usually caused by large loads in connection and disconnection step and this event affect the power system.

III. METHODOLOGY

PRODIST - Module 8 in its revision 10 adopted as consultation and parameter setting. This establishes the procedures related to the PQ, and its scope reaches all consumers connected to the distribution network.

This work aims to evaluate the PQ parameters of distribution transformers against different load profiles, classifying them by NTL in these transformers, three (3) NTL indexes were selected:

Loss index 1: Transformer with NTL up to 10% (low index); Loss index 2: Transformer with a NTL greater than 30% (high rate); Loss index 3: Transformer with a NTL between 10% and 30% (average rate), the transformers have the same nominal power.

The methodology consists of selecting the transformers, collecting data, organizing the information, selecting the PRODIST parameters, and analyzing, verifying how much each loss index influences the behavior of the equipment regarding the PQ (Figure 1).

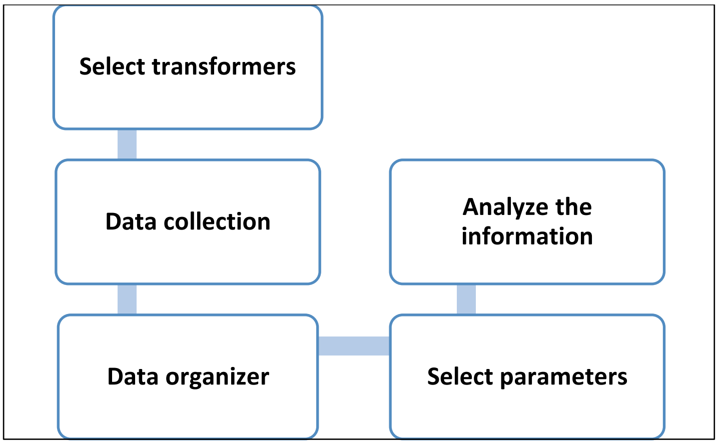


Figure 1: Methodology sequence.
Source: Authors, (2021).

The PQ parameters adopted in the experiment are:

- Supply voltage;
- PF;
- Harmonic distortion;
- Unbalance voltage; and
- Frequency variation

IV. MEASUREMENTS

To make the necessary analysis proposed in this document possible, the data memory for each type of load of the transformers was obtained in addition to what is indicated in PRODIST (1008 valid records in most indicators). In this research, equipment data were acquired in 168 hours of equipment monitoring with data recorded every 120s, exceeding 5000 valid data records.

The equipment used for data acquisition was of the model: PowerNET P-600 from IMS manufacter with measurement class A according to IEC61000-4. The ambient temperature in the location where the transformers are installed showed a maximum of 30ºC and a minimum of 18ºC during the measurement period, so we consider that there was a minimal intervention of the external ambient temperature in the measurements performed.

At the end of the data acquisition, the analyzes reported in item 5 started.

V. RESULTS ANALYSIS

Based on the PQ indicators mentioned in item 2 and on data acquired in the measurements considered for this research the analysis was started. In this item, the results obtained regarding the impacts of the load imposed on the transformers by the amount of NTL in the equipment stand out and show the behaviour of each of the quality indicators that could be measured in the measurements performed.

The first indicator analyzed is the Permanent voltage regime, within the PRODIST limits, there was no violation of the indicator in any of the analyzed equipment, according to the data in Figure 2.

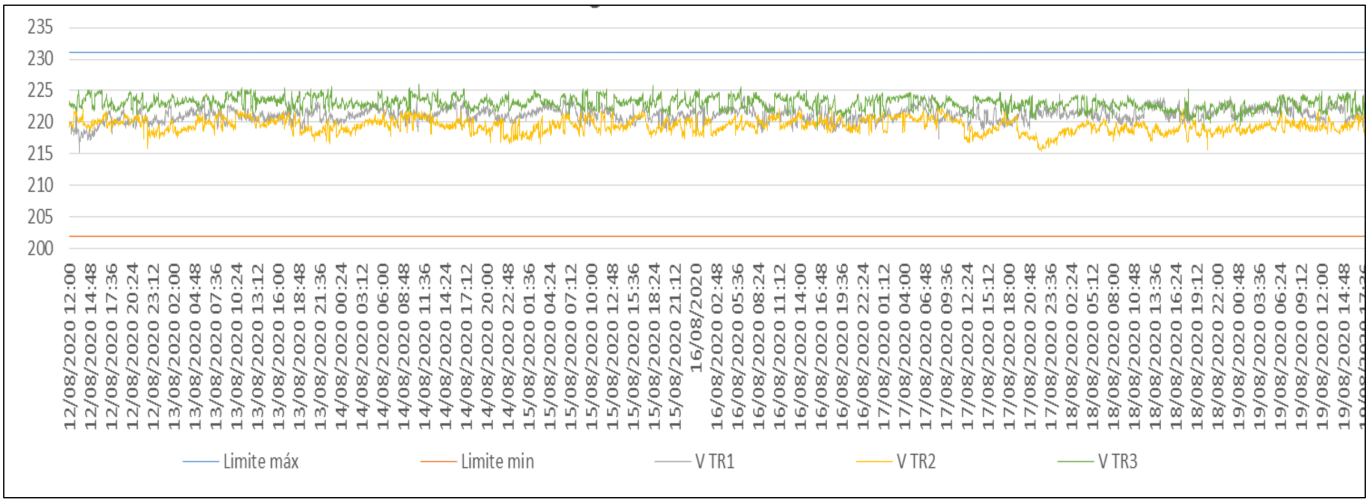


Figure 2: Supply voltage indicator graphic.
Source: Authors, (2021).

However, some peculiarities can be observed between the three transformers and their different loads. We can highlight that transformer 1 with the lowest NTL index presents a smaller variation in this indicator, despite not being the highest average voltage during the measurement period. This higher average is shown in the measurement of transformer 3 and we can also perceive a period of "sinking" of voltage presenting a lower voltage than the average of the period in transformer 3. This period is highlighted in the graph. However, nothing that would interfere with the good performance of transformers in this indicator.

The next indicator is the voltage unbalance which is also in compliance with the limits determined in PRODIST when monitoring transformers.

Analyzing the result of each transformer in this indicator, it can be seen that transformer 1, with a lower rate of NTL, showed greater variation in voltage unbalance to the other transformers, being, as we can analyze by Figure 2, higher values of concentrated unbalance at times considered off-peak, that is, with smaller loads.

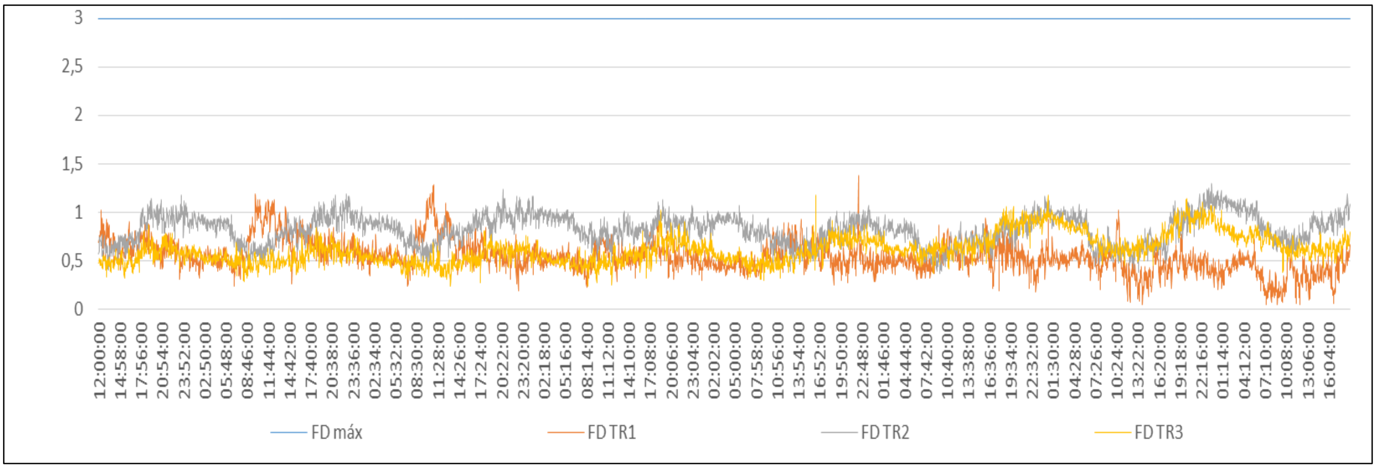


Figure 3: Unbalance factor indicator graphic.
Source: Authors, (2021).

In opposite to what was observed in transformer 1, transformer 2, which has the highest rate of NTL, showed a large variation in the voltage unbalance indicator, but a variation with regularity, with the greatest unbalance during peak hours and at dawn, and it also maintains the highest average unbalance factor value among the monitored transformers. Another observation regarding transformers 1 and 2 with the lowest and highest non-technical loss index, respectively, perceived that to the unbalance factor in the times when the FD is higher in one transformer, it is lower in the other and they continue alternating during the monitoring period.

Transformer 3 presents, as shown in Figure 3, a greater balance in the value of the unbalance factor.

It appears that there were distinct impacts on the voltage unbalance indicator in each of the measured transformers, but we cannot affirm that this presented unbalance factor was caused directly by the NTL index that is acting in each one, the loads profiles other than the monitored equipment cause this variation in voltage unbalances, another point to note is that the measuring equipment did not exceed the limits defined in PRODIST for the indicator in this item presented, thus not characterizing any impacts on the PQ according to this indicator.

The indicator evaluated later is the PF, an indicator linked to active and reactive power (inductive and/or capacitive) consumed by the loads demanded in each equipment.

This is the indicator that suffered the most impact on the monitored equipment during long periods and at different times it was below the minimum index indicated by PRODIST.

As shown in Figure 4, we verified that in transformer 1, the impact of loading the equipment is the lowest in this indicator, with the equipment with the lowest rate of NTL being evaluated, and even so it presented itself outside the ideal indicator impacting the PQ in this equipment.

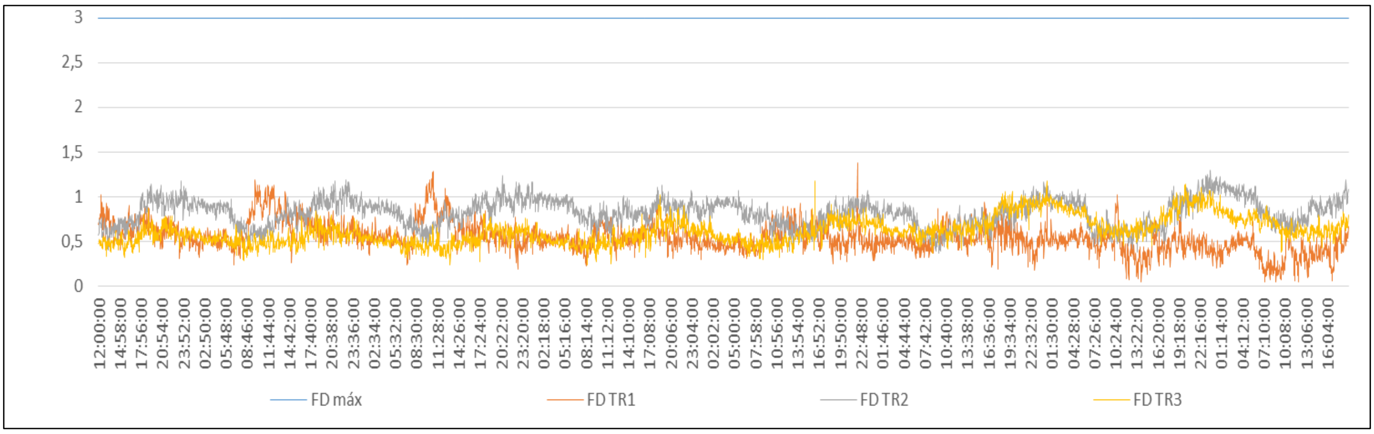


Figure 4: PF indicator graphic.
Source: Authors, (2021).

In transformer 2 the impact of the load on the equipment is perennial in the PF indicator, because during practically the entire monitoring period the indicator remains below the ideal, showing that the NTL interfere a lot in this indicator as we can see in Figure 4.

Following the analysis of the PF, the measurements performed show, according to Figure 4, that transformer 3 was the one that had the most significantly impacted performance and with a great variation of the PF value, varying between very high values, close to 1.00 and very low values, below 0.70.

In this indicator, it was verified that the PRODIST PF limits were extrapolated in all three measured equipment and the PQ supplied by them will be affected by the characteristics of their loads.

In the Total Harmonic Distortion (THD) indicator, there was no extrapolation of the indicator limits, as shown graphically by Figures 5, 6, and 7. Recalling that the distributor is directly responsible for monitoring and managing this indicator vis-à-vis regulatory agents.

This is the indicator resulting from the calculation performed based on PRODIST taking into account the fundamental voltage of the monitored transformers, in which there was no violation of the PQ.

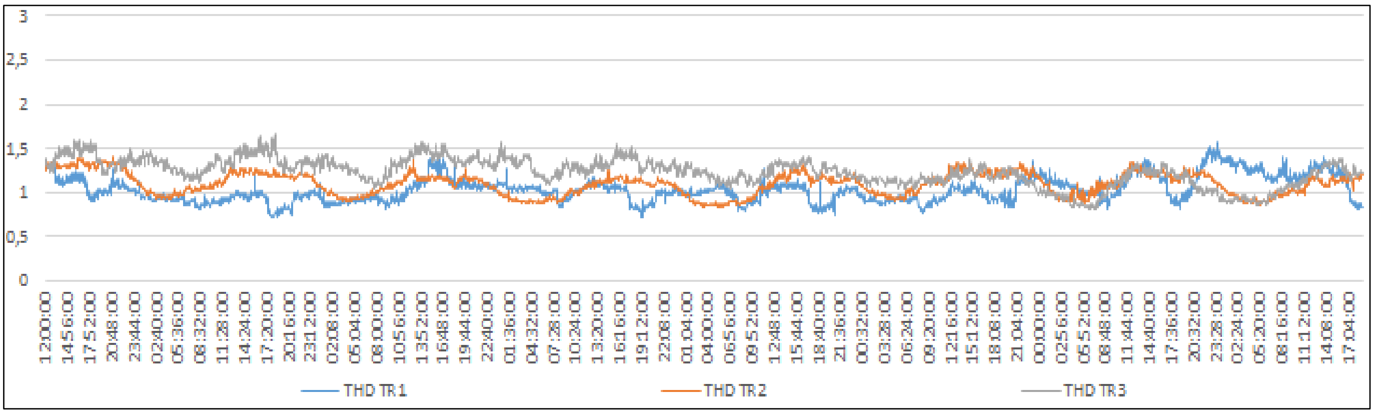


Figure 5: THD voltage phase “a” indicator graphic.
Source: Authors, (2021).

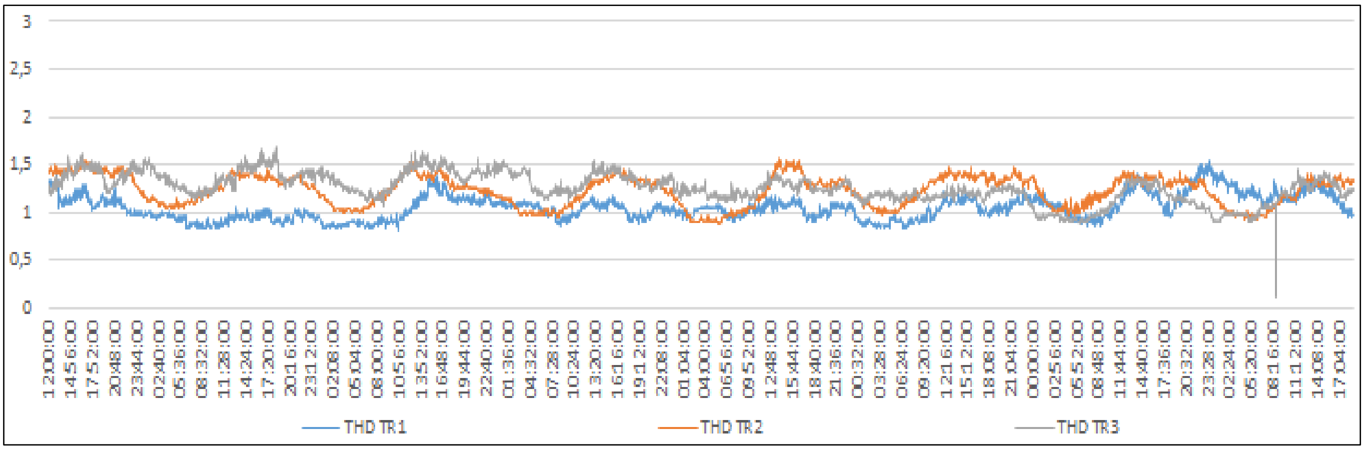


Figure 6: THD voltage phase “b” indicator graphic.
Source: Authors, (2021).

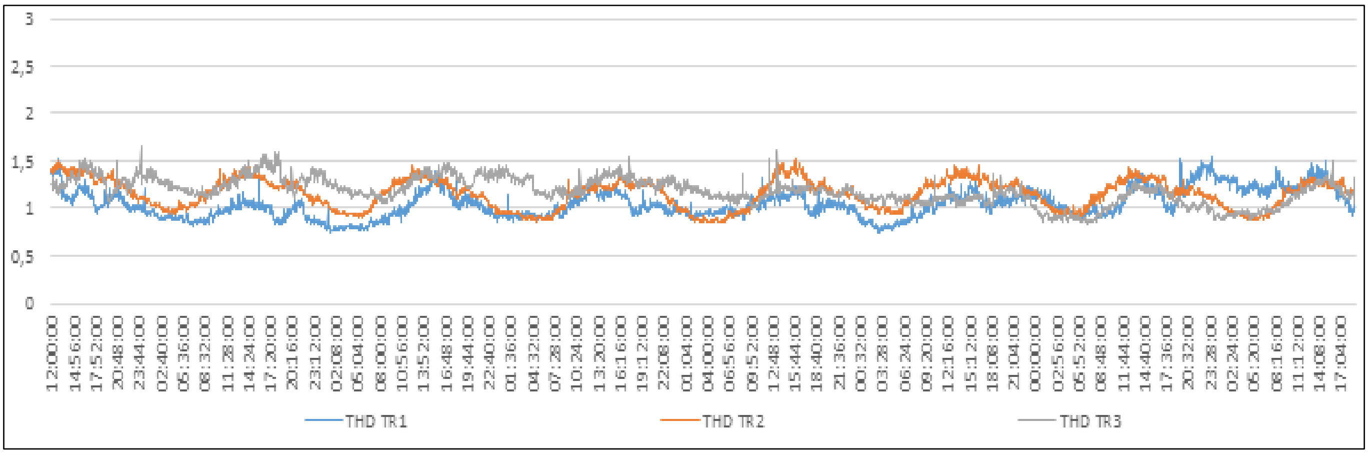


Figure 7: THD voltage phase “c” indicator graphic.
Source: Authors, (2021).

In the Variation of Frequency indicator, Figure 8 illustrates the transformers results, in this regard, there were some moments when the value of the frequency exceeded the minimum and maximum values of PRODIST limits.

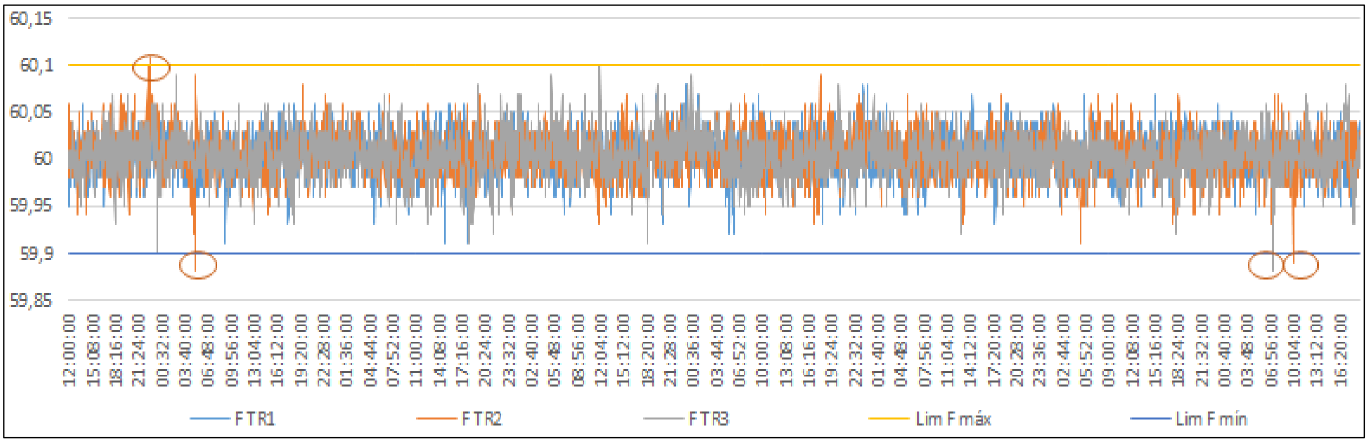


Figure 8: Frequency variation indicator graphic.
Source: Authors, (2021).

We can see in the graph that transformer 2, in this regard, presented in a moment extrapolation of the upper limit and for two moments of the lower limit of the frequency.

In transformer 3, the lower limit of the indicator was extrapolated.

Therefore, this work presents graphically 05 of the 07 indicators of PQ contained in PRODIST – Module 8, illustrating the importance of monitoring these parameters in the guarantee of the supplied PQ. It should be noted that in the next item.

VI. CONCLUSIONS

Five items of PQ assessment were measured and among these results were obtained outside the limit parameters determined in PRODIST in 02 of these indicators, which were the PF indicator, significant impacts, and the Frequency Variation indicator.

The transformers that suffered the impacts were mainly transformers 2 and 3 with high (above 30%) and average (between 10 and 30%) NTL respectively, results that reinforce the main idea of this document to analyze the impacts of these load profiles at the PQ, these two transformers presented a failure in the PQ supplied in the Frequency Variation indicator.

Specifically, in the PF indicator, all transformers failed in the supplied PQ, remaining with performance below the ideal.

In general, the impact of each load profile and not only evaluating the limits of each indicator but it is also concluded that the loads on each of the transformer's impact on other PRODIST indicators even if they remain within the limits of the indicators, they are the cases of the Permanent voltage regime and voltage unbalance factor indicators that showed the greatest impact on the transformer that presented the highest rate of NTL. Table 2 presents a summary of the results of each of the indicators of PQ monitored in the equipment of this study.

Table 2: Summary results of PQ indicators.

Indicator	TR1	TR2	TR3
Steady State Voltage	Indicator – OK	Indicator – OK	Indicator – OK, however it had impacts during peak hours
Imbalance Factor	Indicator – OK	Indicator – OK	Indicator – OK
Power Factor	Violated Indicator	Violated indicator, during almost any monitoring period	Violated indicator, with significant impacts in this indicator
Total Harmonic Distortion	Indicator – OK	Indicator – OK	Indicator – OK
Frequency variation	Indicator – OK	Violated indicator	Violated indicator

Source: Authors, (2021).

It is worth mentioning that PRODIST has seven (7) PQ indicators and a deeper analysis can be supported by the results obtained in this document and the impacts that PQ may suffer from the influence of NTL in distribution transformers.

This work presented for a case study of a distributor a critical analysis of the indicators of PQ considering a concern aspect to the distributors that are the high value of NTL. It is perceived that actions to reduce these losses will directly impact the energy quality parameters, that is, less non-technical loss, better service of supplying quality energy to the final consumer.

VII. AUTHOR'S CONTRIBUTION

Conceptualization: Salvador Paes Ribeiro Junior, Marcio Zamboti Fortes.
Methodology: Salvador Paes Ribeiro Junior, Ricardo Chagas Tolentino, Marcio Zamboti Fortes.
Investigation: Salvador Paes Ribeiro Junior, Ricardo Chagas Tolentino.
Discussion of results: Ricardo Chagas Tolentino, Marcio Zamboti Fortes, Guilherme Gonçalves Sotelo.
Writing – Original Draft: Salvador Paes Ribeiro Junior, Ricardo Chagas Tolentino.
Writing – Review and Editing: Marcio Zamboti Fortes, Guilherme Gonçalves Sotelo.
Resources: Salvador Paes Ribeiro Junior.
Supervision: Marcio Zamboti Fortes.
Approval of the final text: Salvador Paes Ribeiro Junior, Ricardo Chagas Tolentino, Marcio Zamboti Fortes, Guilherme Gonçalves Sotelo.

VIII. REFERENCES

[1] S.M. Deckmann and J.A. Pomilio. "Avaliação da qualidade da energia elétrica," [Online]. Available: http://www. dsce. fee. unicamp. br/antenor/pdffiles/qualidade/b5. pdf. [Accessed 13 05 2021].

[2] R.E. Brown. Electric power distribution reliability. CRC press, 2017. ISBN 978-0849-3756-82

[3] C. Sankaran. "Power Quality",CRC press, 2017. ISBN: 978-813-5183-6050.

[4] P. Cernek, M. Brejcha, J. Hájek and J. Pígl. "A Comparison of Power Quality Controllers," Acta Polytechnica, vol. 52, n. 5, pp.22-25, 2012.

[5] R. Brida Neto. "Perdas de Energia em Transformadores de Distribuição: Análise Técnica e Econômica em Função do Carregamento", Final Project Engineering Electric Course, UNISUL, Tubarão, SC Brazil, 2018.

[6] N. Kalache. "Gerenciamento de transformadores de distribuição operando em redes inteligentes", Doctorate Thesis, Universidade Estadual Paulista, Ilha Solteira, SP, Brazil, 2016.

[7] T.Gonen. "Electric power distribution engineering". CRC press, 2015.ISBN 978-1-4822-0705-7.

[8] ANEEL. "Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional – PRODIST- Módulo 8", 2020. [Online]. Available: <https://www.aneel.gov.br/modulo-8>. [Accessed 29 08 2020].

[9] G.V.S. Rocha. "Sistema inteligente de suporte a análise da qualidade da energia elétrica em regime permanente," Master Dissertation, Universidade Federal do Pará, Belém, PA, Brazil. 2017.

[10] P.M. Santiago. " Cálculo e análise computacional de campos elétricos gerados por linhas de transmissão de alta tensão em regime permanente," In: XVI CEEL, UFU, 2018. ISSN: 2178-8308.

[11] I.S. Machado. "Software para auxiliar na correção do fator de potência," Sinergia, vol.17, n.1, pp.15-21, 2016.

[12] G.G. Santos et al. "Uma proposta para definição de áreas de sensibilidade às distorções harmônicas," In: XIV CEEL, UFU, 2016. ISSN 2178-8308.

[13] H.A.F. Camporez et al. "Protótipo para análise da qualidade e controle da demanda pelo lado do consumidor em smart grid", 2016.

[14] S.L. Bandeira, R.G. Vieira and M.I.S. Guerra. "Análise da qualidade de energia elétrica de um sistema fotovoltaico conectado à rede", 2018.