APPLIED COMPUTER SCIENCE

The Journal is a peer-reviewed, international, multidisciplinary journal covering a broad spectrum of topics of computer application in production engineering, technology, management and economy.

The main purpose of Applied Computer Science is to publish the results of cutting-edge research advancing the concepts, theories and implementation of novel solutions in computer technology. Papers presenting original research results related to applications of computer technology in production engineering, management, economy and technology are welcomed.

We welcome original papers written in English. The Journal also publishes technical briefs, discussions of previously published papers, book reviews, and editorials. Especially we welcome papers which deals with the problem of computer applications in such areas as:

- manufacturing,
- engineering,
- technology,
- designing,
- organization,
- management,
- economics,
- innovations,
- competitiveness,
- quality and costs.

The Journal is published quarterly and is indexed in: BazTech, Cabell's Directory, Central & Eastern European Academic Source (CEEAS), CNKI Scholar (China National Knowledge Infrastucture), DOAJ (Directory of Open Access Journals), EBSCO, ERIH PLUS, Index Copernicus, J-Gate, Google Scholar, Scope Database, Scopus, TEMA Technik und Management.

Letters to the Editor-in-Chief or Editorial Secretary are highly encouraged.

CONTENTS

Applied Computer Science, vol. 18, no. 4, pp. 5–15
doi: 10.35784/acs-2022-25

Submitted: 2022-06-16 | Revised: 2022-09-25 | Accepted: 2022-10-04

Keywords: clonal evolution, mutation waves, yeast evolution, numerical modelling

Jarosław GIL [0000-0001-9228-9937]*, *Andrzej POLAŃSKI* [0000-0002-1793-9546]*

APPLICATION OF GILLESPIE ALGORITHM FOR SIMULATING EVOLUTION OF FITNESS OF MICROBIAL POPULATION

Abstract

In this study we present simulation system based on Gillespie algorithm for generating evolutionary events in the evolution scenario of microbial population. We present Gillespie simulation system adjusted to reproducing experimental data obtained in barcoding studies – experimental techniques in microbiology allowing tracing microbial populations with very high resolution. Gillespie simulation engine is constructed by defining its state vector and rules for its modifications. In order to efficiently simulate barcoded experiment by using Gillespie algorithm we provide modification – binning cells by lineages. Different bins define components of state in the Gillespie algorithm. The elaborated simulation model captures events in microbial population growth including death, division and mutations of cells. The obtained simulation results reflect population behavior, mutation wave and mutation distribution along generations. The elaborated methodology is confronted against literature data of experimental evolution of yeast tracking clones sub-generations. Simulation model was fitted to measurements in experimental data leading to good agreement.

1. INTRODUCTION

Experimental evolution techniques for populations of microbial organisms is a fast-developing area of scientific research, which provides important measurement data concerning scenarios and parameters of adaptive haploid evolution of microbes under different conditions. It has wide areas of applications both in cellular biology where it leads to advances in understanding adaptation mechanisms in microbial populations, and in applications of evolutionary genetics to other research fields such as epidemiology, virology or cancer research. Fast advances in molecular biology give impulses to extending volumes and resolutions of available measurement data in experimental microbial evolution. Implementing techniques, such as high throughput sequencing, labelling and barcoding in evolving microbial cultures allow obtaining detailed data on kinetics and dynamics of their evolution.

* Department of Computer Graphics, Vision and Digital Systems, Silesian University of Technology, Gliwice, Poland, jaroslaw.gil@polsl.pl

Recent publications in the field of experimental evolution of microbes (Bush et al., 2020; Kinnersley et al., 2021; Blundell et al., 2019; Nguyen Ba et al., 2019; Levy et al., 2015) provide datasets of significant size and accuracy enabling the study of genomic mechanisms and genetic forces behind their adaptation mechanisms to environmental conditions. Among the referenced experimental techniques an approach based on barcoding of microbial DNA (Blundell et al., 2019; Nguyen Ba et al., 2019) is of special interest due to possibility of tracing the large number of sub-populations/clones of microbial population over the whole experiment time range.

Important aspects of the study of experimental microbial evolution is characterization of roles and strengths of genetic forces, replication, mutation, genetic drift and selection (Neher, 2013) in scenarios of their adaptation to external conditions (Beckman & Loeb, 2005; Kvitek & Sherlock, 2013). In order to understand strengths, inter-plays between genetic forces and their dynamics, mathematical models of adaptive evolution of haploid populations are developed. Mathematics and computer modeling behind adaptive evolution lead to expressing the increase of adaptive potential of a microbial population as the result of propagation of the fitness wave, driven by emergence of favorable mutations in the DNA replication process and selection against less fitted species (Desai & Fisher, 2007). Mathematical modeling of dynamics of adaptive evolution is a wide area of research including approaches such as deterministic modeling using partial differential equations (Fisher-Kolmogorov equations) e.g., (Wang et al., 2019), or stochastic modeling of processes of cellular replication and death and occurrence of (driver) mutations, by using Markov birth - death processes (Baar et al., 2016; Castillo & Virgilio, 2015) branching processes (Wild, 2011; Yakovlev, Stoimenova & Yanev, 2008; Castillo & Virgilio, 2015) or multitype branching processes (Bozic et al., 2010; McFarland, Mirny & Korolev, 2014; Nguyen Ba et al., 2019; Levy et al., 2015).

When laws describing probabilities of evolutionary events become complicated the approach to modeling by using stochastic simulations can be very useful (Foo, Leder & Michor, 2011). Stochastic simulations are very flexible and can capture arbitrary scenarios of adaptation evolution. The basic approach to (forward) stochastic simulation of cellular evolution is by using Gillespie's algorithm (Gillespie, 2001; Marchetti, Priami & Thanh, 2017) where the state of the simulated process is defined by mutation profiles of cells and possible events are drawn randomly on the basis of probability distributions dependent on the state of the process.

In this study we present the implementation of the Gillespie's algorithm for evolutionary simulations tailored to the scenario of experimental yeast evolution with adaptation to external conditions, quantified by using the barcoding technique (Nguyen Ba et al., 2019; Levy et al., 2015). Novelty of our approach is adjusting (fitting) Gillespie simulation algorithm to the barcoded evolution experimental scenario by suitable changes of its construction. We describe the approach to fitting evolutionary parameters, intensities of cellular/microbial birth and deaths, intensities of mutations and mutations evolutionary advantage/fitness parameter to observational data. We show results of simulation algorithm versus published data on baker yeast (saccharomyces cerevisiae) growth (Levy et al., 2015).

2. MODEL DESCRIPTION

Elaborating simulation system for microbial evolution relies on defining stochastic models of events in the cellular lifetime: cell death, cell division or mutation, which occur during cell division. Methodology for obtaining timing and orderings of events relies on the use of the Gillespie algorithm, which provides a way to obtain suitable probability distributions of times of events on the basis of the state of the processes. Basic approach is to simulate events cell by cell, iterating through whole (Marchetti, Priami & Thanh, 2017). First improvement is considering all cell events with occurring time less than tau (tau leap algorithm (Nguyen Ba et al., 2019)) (Cao, Gillespie & Petzold, 2006). However, tau leap modification still does not significantly improve simulation time for larger populations. The binned version of Gillespie algorithm, elaborated in this study, makes possible to consider microbial population size above one million of cells/microbes/viruses/bacteria with reasonable simulation times.

2.1. Binned Gillespie algorithm

Tab. 1. Simulation parameters assumed as (Levy et al., 2015) experiment conditions. Whole population is divided into groups which represents barcoded cells

Model parameters	
mutation effect	2.5 – 15% by Poisson distribution
initial population	50’000 × 100 cells
population capacity	$5 \cdot 10^6$
mutation probability	20%
simulation cycles	1000

The basic Gillespie’s algorithm modification by binning means grouping microbial cells by one independent characteristic which can describe a group of cells. For that purpose, number of mutations in the cell was chosen. Simulation begins with large population of cells with no mutation. In each cycle some of them can obtain new mutation which cause cells differentiation. Every cycle of simulation consists of three steps: generating random number of dying and dividing cells based on accurate probabilities and tau step, calculating number of mutating cells and population actualization. Probabilities are given by exponential distribution (1), (2).

$$P(death) = 1 - exp(-tau * sum(population)/capacity) \quad (1)$$

$$P(divide) = 1 - exp\big(-tau/(1 - fitness)\big) \quad (2)$$

Each mutation provides a change in fitness coefficient which has impact on the intensity of cellular division process. Calculation of fitness for single cell is easy – mutation provides percentage change in cell structure. To calculate it for one bin we assume that one mutation has impact on mean bin fitness (3). Mutating cells have the same fitness in the bin, new mutations have additional effect. Change of cells adjustment is obtained on the basis of probabilities in the Poisson distribution.

$$fitness(t+1) = (fitness(t) * n_{cells} + mutations * mutation\ effect)/n_{cells} \quad (3)$$

2.2. Simulating barcoding experimental data

In barcoding experiments DNA strands acquire markers allowing for their unique identification. At present, barcoding technique allows introducing hundreds of thousands of markers to populations of microorganisms to trace their evolution. Barcoding is particularly useful in studies of clonal evolution. To track separate clones lineages, it is assumed that population is grouped by clone reference barcodes introduced to the microbial population. In (Levy et al., 2015) experiment was performed based on tracking specific barcodes firstly inserted into microbial cell genomes.

Except bin sizes fitness parameter based on (3) and mutation wave in the form of mean mutation number were calculated. In figure 1 basic information about population evolution in time is presented. In the left side single lineage sizes (a) and sizes distributions (c) are presented. In the right-side mutation (d) and fitness wave (b) (the track of population adjustment) are presented. Model parameters are presented in table 1. For simulation parameters described in (Levy et al., 2015) were implemented.

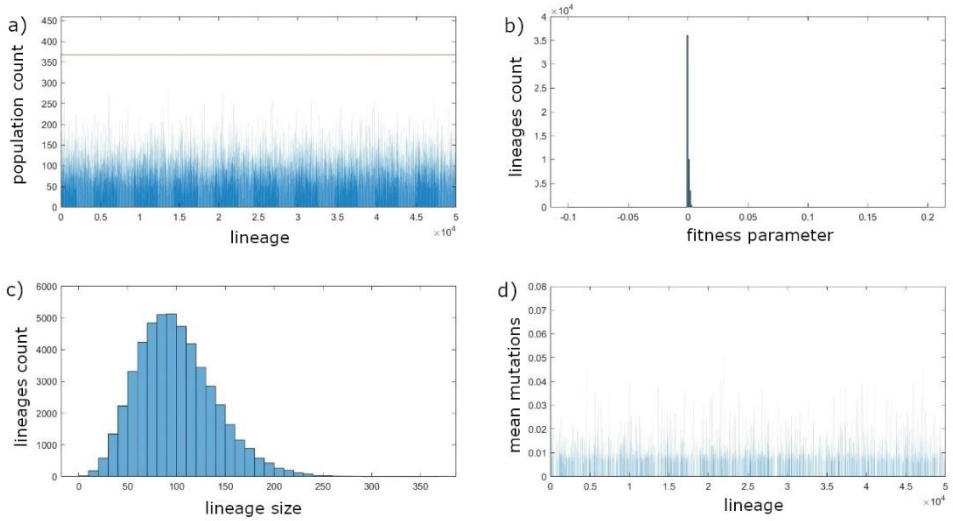


Fig. 1. Lineage binned algorithm result according to one generation, a) linages sizes with highlighted maximum value, b) whole population fitness wave, c) linage sizes distribution, d) mean numbers of mutations along population lineages.

Simulation time is greater for that modification than binned by mutation number because of greater number of separate bins. The complexity of algorithm is about O(n) because number of iterations depends on bin number. The basic algorithm complexity was about $O(n^3)$ cause of multiple iterations and comparisons through whole population cell by cell. For binning by mutation number simulation time is flexible because of changing number of bins. For binning by lineage characteristic number of bins are steady - one cycle time is approximately same.

Two separate simulations were made. First one with assumed parameters in table 1, second one with different initial population and population capacity. Lineages size were providing as experimental data from (Levy et al., 2015) as like simulation parameters were based on.

Pseudocode of our algorithm which shows its basic construction is listed below.

```
SET initial parameters
   {
   SET initial population size as 5'000'000
   SET capacity as 5'000'000
   SET mutation probability as 0.2
   SET tau step as 0.005
   SET step number as 1000
   }

CREATE initial population
   {
   CREATE population array as 50'000 groups of 100 cells
   CREATE mutation array as zeros for 50'000 cell groups
   CREATE fitness array as ones for 50'000 cell groups
   SET mutation effect for each group from 0.025 to 0.15
   }

WHILE step is less than step number
   {
   IF step is multiplier of 10
      {
      SHOW population plot
      SHOW mutation plot
      SHOW fitness plot
      }

   UPDATE population
      {
      CALCULATE death probability (1-exp (-tau*(cells number / capacity))) for whole
population

         FOR group in population
         {
            CALCULATE division probability (1-exp(-tau/(1-fitness))) for group
            CALCULATE number of dying and dividing cells
            SET mutating cells as 0.2 dividing cells

            SUBSTRACT number of dying cells from group
            ADD number of dividing cells to group

            CALCULATE new mean of mutation number for group ((mutation number * cells + new
mutations)/cells)
            CALCULATE new fitness for group ((fitness * cells + new mutations * mutation
effect)/cells)
         }
      }
   }
```

Listing 1. Binned Gillespie algorithm pseudocode

3. RESULTS AND DISCUSSION

Two simulation experiments were made:

1. Simulation basing on 50’000 lineages consisting of 100 clones and parameters described in table 1 – similar to (Levy et al., 2015),
2. Simulation basing on data provided by (Levy et al., 2015) – 500’000 lineages with different initial clones number.

3.1. 1st Experiment

To estimate model accuracy several simulations were pursued and the best suited parameters where chosen. In figure 2 simulated trajectories of lineages evolution are presented. Red, blue and white colors correspond to calculated fitness of single bins. From 50’000 introduced bins about 10% became extinct and few grew to much larger groups. For adaptive lineages the growth rate was almost about 1 magnitude in small time step – about 100 seconds. In theory cell fitness should provide more division events – as results show. Because of no initial differentiation between lineages, small amount of them adapt gaining sufficient value of fitness coefficient.

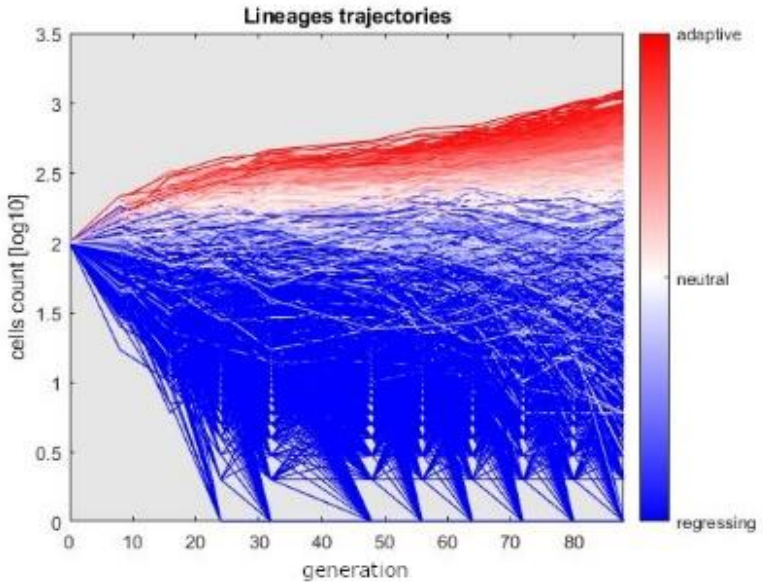


Fig. 2. Simulated lineages trajectories (for lineages greater than 100 cells): red color – lineages with good adaptation, blue color – bad fitness lineages; white color explains neutral fitness effect lineages

Growth of adaptive lineages causes increase in cell death probability. Cells with lower fitness mostly die, which results in lineages extinction. Some clones dominate population imposing smaller genetic variety inside the whole population. In figure 3 linage size distributions, with fitness factor of lineage marked by blue/orange colors, are presented. Adaptive lineages are growing as shown in the figure - lineages distribution changes its mean value from 10^2 to about $10^{2.5}$. The high value bars on figure near zero value represents extinct lineages.

Population differentiation proceeds as assumed. Few lineages gathered beneficial mutation (adaptive lineages) which results in higher fitness factor and rise in lineage size. Some of clones mutations had small effect on its fitness causing lineage extinction. Mostly mutation provided group adjustment establishing lineage size at steady level.

Further simulations should cause more lineage extinctions because of fast growth of adapted ones. Death probability, which depends, by logistic relation, on the whole population size, will be shaped mostly by the biggest groups. Extended simulation time could provide information about dominant clones. Fast lineage growth would cause extinction of small clones. Genetic variety should be very small and population should be composed of genetic clones.

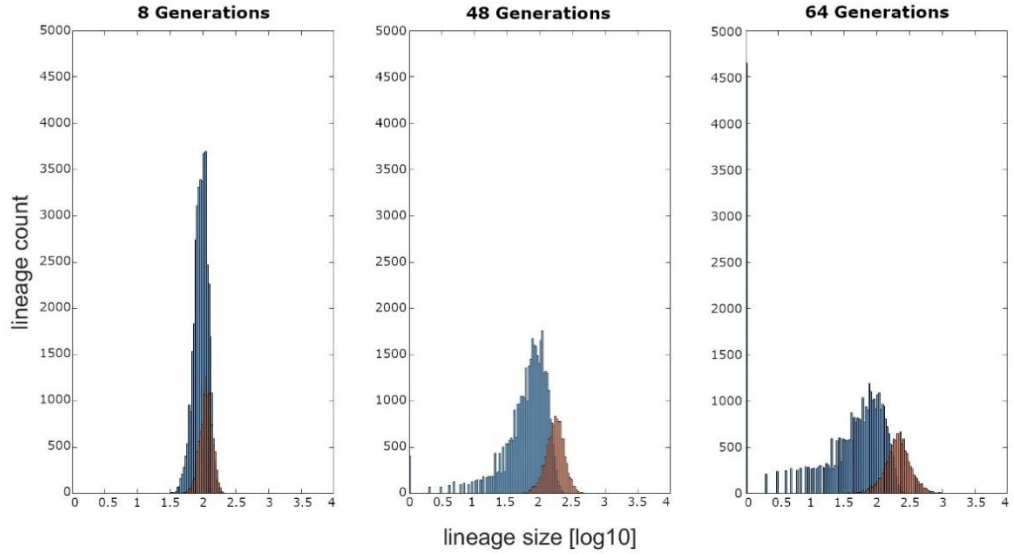


Fig. 3. Lineage size distribution in simulated data: blue part means neutral and regressing lineages, red part represents adaptive lineages

3.2. 2nd Experiment

Basing on data provided by (Levy et al., 2015) simulation was performed assuming that initial lineages are varied. Initial lineage size and its fitness was set according to experimental data. Mutation effect on group fitness was same as in 1st experiment – 2.5-15% described by Poisson distribution. Provided data is not reflection of simulation. Research group was gathering information about cells by selecting random number from population. Our goal was to obtain similar characteristic of population growth. In figure 4 simulated lineages trajectories are presented. Differences between figure 2 and figure 4 follow from different initial model state. When groups are assumed to be equal at simulation beginning fitness of whole population is distributed as mutation factor. At simulation begin both division and death probability are dependent on whole population – in further cycles dependency changes. When initial size of lineages is randomized death probability is mostly dependent on high size lineages. More cells also provide more division and mutation events which causes lineage fitness increase. Big groups are growing, gaining good adaptation and causing extinction of smaller or less adapted lineages.

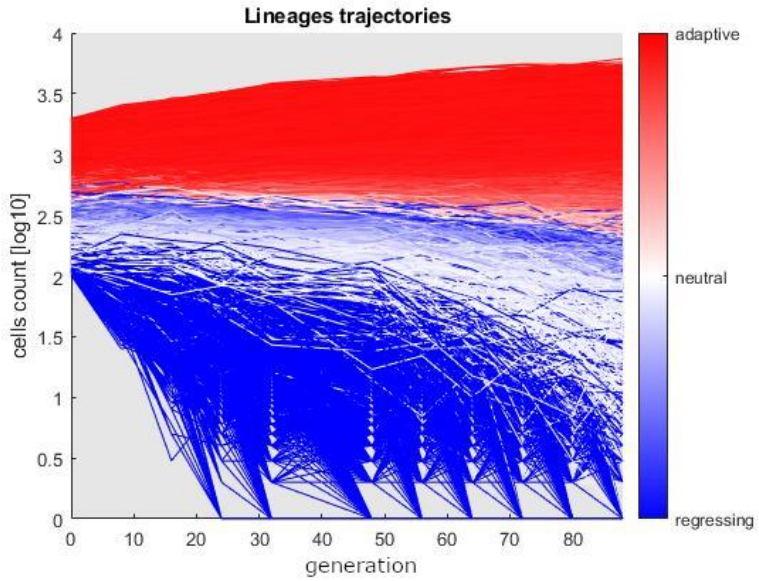


Fig. 4. Simulated lineages trajectories based on experimental data (for lineages greater than 100 cells): red color – lineages with good adaptation, blue color – bad fitness lineages; white color explains neutral fitness effect lineages

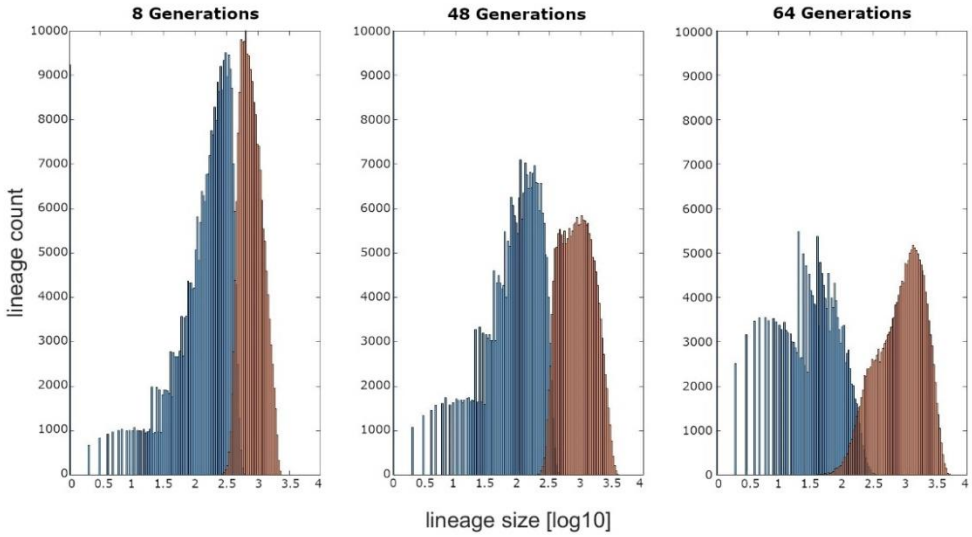


Fig. 5. Lineage size distribution in experimental data: blue part means neutral and regressing lineages, red part represents adaptive lineages

In figure 5 lineage size distribution and on figure 6 its fitness factor is presented to demonstrate that at initial stage of simulation the groups were different from each other. Real experiment shows that adaptive lineages should grow at exponential rate when gaining appropriate size and fitness factor. Further cycles of population evolution should show that. At the right-hand side of the picture the change of lineage sizes is shown. It creates some kind of evolution front so it is possible that population would rise much faster in next steps.

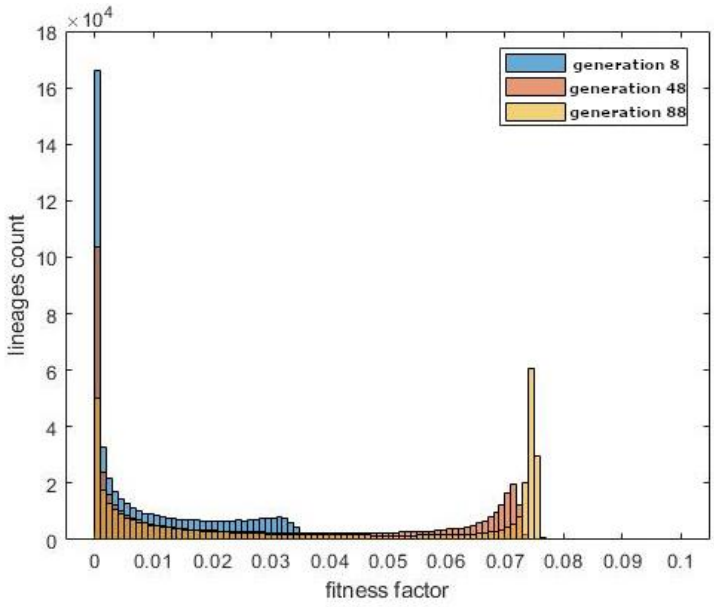


Fig. 6. Fitness change

Changing in lineage size and in its fitness is caused by multiple mutations. Some of them could provide good adaptation and increase in fitness factor. Figure 6 describes how lineages adapt in simulation time. High fitness factor can be explained by gaining beneficial mutation by a lineage. Low value of fitness mostly occurs in lineages for which mutation effects were neutral or have no mutations at all. Differentiation inside population is visible and caused by multiple mutations in big groups and small number of mutations in small ones.

3.3. Summary

Computational experiments give insights into evolution of fitness of microbial population. Microbial populations are adapting to environmental conditions by the process of asexual evolution. Due to the lack of recombination force, the whole microbial population can be partitioned into clones – subpopulation of identical or similar genetic profiles. Adaptation is fast because of cell replications and advantageous mutation.

The evolutionary process is studied by the clonal evolution theory. In this paper clonal evolution is numerically simulated by appropriately defined Gillespie simulation engine presented in subsection 2.2. Stochastic simulations reproduce clonal evolution scenarios with fast adaptation as observed in many biological systems, such as population of bacteria, some fungi, or in the processes of mitotic evolution of cellular subpopulations of organisms, importantly in cancer cellular populations development. Arising many genetic duplicates provides very fast, nearly chaotic, population growth with lack of genetic variety. Cell mutation causes differentiation between subpopulations. Mutations can cause positive, negative and neutral effect. Positive effect is observed when mutation persist in many cells, negative causes cell death. For good interpretation mutation effect valiant allelic frequency (VAF) coefficient is needed to be defined. Positive effect mutation is observed for high values of VAF, neutral and negative for low.

Population behavior is similar in both experiments pursued in this study – lineages sizes increase, some of them extinct and some stay on steady level. Gaining by lineage beneficial mutation causes good adaptation of clone as a result of fitness factor increase. Simulation parameters are dependent mostly on big groups of clones which are replacing smaller ones. Death probability increases while population growth, division probability changes because of mutations. Big clone groups easily compensate cells death because of good adaptation while smaller ones mostly regress and extinct. The beneficial mutations are visible because adaptive lineage size distribution moves to higher values.

4. CONCLUSION

The main conclusion of this study is that microbial experimental evolution traced with high resolution by using barcoding technique can be efficiently reproduced by using Gillespie simulation engine. In contrast to branching process from (Levy et al., 2015) our model contains direct mechanism of whole microbial population growth or decline. Our simulation algorithm is computationally efficient, it scales well with large population sizes. Our computations prove that results of simulations can easily reconstruct observations obtained in the microbial experimental evolution scenario. Results of our simulations algorithms shown in Figures 2-5 show good agreement with experimental data. Moreover, Gillespie simulation technique is very flexible, so it will easily cover possible variations in measurement scenarios in the future versions of experimental evolution researches.

Acknowledgements

This publication was supported by the Department of Graphics, Computer Vision and Digital Systems, under research project for young scientists (Rau6, 2022), Silesian University of Technology (Gliwice, Poland).

REFERENCES

Baar, M., Coquille, L., Mayer, H., Hölzel, M., Rogava, M., Tüting, T., & Bovier, A. (2016). A stochastic model for immunotherapy of cancer. *Scientific Reports*, *6*(1), 24169. https://doi.org/10.1038/srep24169

Beckman, R. A., & Loeb, L. A. (2005). Negative Clonal Selection in Tumor Evolution. *Genetics*, *171*(4), 2123–2131. https://doi.org/10.1534/genetics.105.040840

Blundell, J. R., Schwartz, K., Francois, D., Fisher, D. S., Sherlock, G., & Levy, S. F. (2019). The dynamics of adaptive genetic diversity during the early stages of clonal evolution. *Nature Ecology & Evolution*, *3*(2), 293–301. https://doi.org/10.1038/s41559-018-0758-1

Bozic, I., Antal, T., Ohtsuki, H., Carter, H., Kim, D., Chen, S., Karchin, R., Kinzler, K. W., Vogelstein, B., & Nowak, M. A. (2010). Accumulation of driver and passenger mutations during tumor progression. *Proceedings of the National Academy of Sciences*, *107*(43), 18545–18550. https://doi.org/10.1073/pnas.1010978107

Bush, S. J., Foster, D., Eyre, D. W., Clark, E. L., De Maio, N., Shaw, L. P., Stoesser, N., Peto, T. E. A., Crook, D. W., & Walker, A. S. (2020). Genomic diversity affects the accuracy of bacterial single-nucleotide polymorphism–calling pipelines. *GigaScience*, *9*(2), giaa007. https://doi.org/10.1093/gigascience/giaa007

Cao, Y., Gillespie, D. T., & Petzold, L. R. (2006). Efficient step size selection for the tau-leaping simulation method. *The Journal of Chemical Physics*, *124*(4), 044109. https://doi.org/10.1063/1.2159468

Castillo, F., & Virgilio, N. (2015). *Stochastic Modeling of Cancer Tumors using Moran Models and an Application to Cancer Genetics* [Thesis, Rice University]. https://scholarship.rice.edu/handle/1911/87795

Desai, M. M., & Fisher, D. S. (2007). Beneficial Mutation–Selection Balance and the Effect of Linkage on Positive Selection. *Genetics*, *176*(3), 1759–1798. https://doi.org/10.1534/genetics.106.067678

Foo, J., Leder, K., & Michor, F. (2011). Stochastic dynamics of cancer initiation. *Physical Biology*, *8*(1), 015002. https://doi.org/10.1088/1478-3975/8/1/015002

Gillespie, D. T. (2001). Approximate accelerated stochastic simulation of chemically reacting systems. *The Journal of Chemical Physics*, *115*(4), 1716–1733. https://doi.org/10.1063/1.1378322

Kinnersley, M., Schwartz, K., Yang, D.-D., Sherlock, G., & Rosenzweig, F. (2021). Evolutionary dynamics and structural consequences of de novo beneficial mutations and mutant lineages arising in a constant environment. *BMC Biology*, *19*(1), 20. https://doi.org/10.1186/s12915-021-00954-0

Kvitek, D. J., & Sherlock, G. (2013). Whole Genome, Whole Population Sequencing Reveals That Loss of Signaling Networks Is the Major Adaptive Strategy in a Constant Environment. *PLOS Genetics*, *9*(11), e1003972. https://doi.org/10.1371/journal.pgen.1003972

Levy, S. F., Blundell, J. R., Venkataram, S., Petrov, D. A., Fisher, D. S., & Sherlock, G. (2015). Quantitative evolutionary dynamics using high-resolution lineage tracking. *Nature*, *519*(7542), 181–186. https://doi.org/10.1038/nature14279

Marchetti, L., Priami, C., & Thanh, V. H. (2017). *Simulation Algorithms for Computational Systems Biology*. Springer International Publishing. https://doi.org/10.1007/978-3-319-63113-4

McFarland, C. D., Mirny, L. A., & Korolev, K. S. (2014). Tug-of-war between driver and passenger mutations in cancer and other adaptive processes. *Proceedings of the National Academy of Sciences*, *111*(42), 15138–15143. https://doi.org/10.1073/pnas.1404341111

Neher, R. A. (2013). Genetic draft, selective interference, and population genetics of rapid adaptation. *Annual Review of Ecology, Evolution, and Systematics*, *44*(1), 195–215. https://doi.org/10.1146/annurev-ecolsys-110512-135920

Nguyen Ba, A. N., Cvijović, I., Rojas Echenique, J. I., Lawrence, K. R., Rego-Costa, A., Liu, X., Levy, S. F., & Desai, M. M. (2019). High-resolution lineage tracking reveals travelling wave of adaptation in laboratory yeast. *Nature*, *575*(7783), 494–499. https://doi.org/10.1038/s41586-019-1749-3

Wang, C.-H., Matin, S., George, A. B., & Korolev, K. S. (2019). Pinned, locked, pushed, and pulled traveling waves in structured environments. *Theoretical Population Biology*, *127*, 102–119. https://doi.org/10.1016/j.tpb.2019.04.003

Wild, G. (2011). Inclusive Fitness from Multitype Branching Processes. *Bulletin of Mathematical Biology*, *73*(5), 1028–1051. https://doi.org/10.1007/s11538-010-9551-2

Yakovlev, A. Y., Stoimenova, V. K., & Yanev, N. M. (2008). Branching Processes as Models of Progenitor Cell Populations and Estimation of the Offspring Distributions. *Journal of the American Statistical Association*, *103*(484), 1357–1366. https://doi.org/10.1198/016214508000000913

Applied Computer Science, vol. 18, no. 4, pp. 16–27
doi: 10.35784/acs-2022-26

Submitted: 2022-06-24 | Revised: 2022-09-25 | Accepted: 2022-10-01

Keywords: Meteorological data, M5 model tree, Linear model functions, Gradient boosting, Logistic Model trees

Sheikh Amir FAYAZ [0000-0001-6606-0864]*, *Majid ZAMAN* [0000-0003-1070-8195]**
Muheet Ahmed BUTT [0000-0002-8059-0180]* , *Sameer KAUL* [0000-0003-0911-0073]*

HOW MACHINE LEARNING ALGORITHMS ARE USED IN METEOROLOGICAL DATA CLASSIFICATION: A COMPARATIVE APPROACH BETWEEN DT, LMT, M5-MT, GRADIENT BOOSTING AND GWLM-NARX MODELS

Abstract

Rainfall prediction is one of the most challenging task faced by researchers over the years. Many machine learning and AI based algorithms have been implemented on different datasets for better prediction purposes, but there is not a single solution which perfectly predicts the rainfall. Accurate prediction still remains a question to researchers. We offer a machine learning-based comparison evaluation of rainfall models for Kashmir province. Both local geographic features and the time horizon has influence on weather forecasting. Decision trees, Logistic Model Trees (LMT), and M5 model trees are examples of predictive models based on algorithms. GWLM-NARX, Gradient Boosting, and other techniques were investigated. Weather predictors measured from three major meteorological stations in the Kashmir area of the UT of J&K, India, were utilized in the models. We compared the proposed models based on their accuracy, kappa, interpretability, and other statistics, as well as the significance of the predictors utilized. On the original dataset, the DT model delivers an accuracy of 80.12 percent, followed by the LMT and Gradient boosting models, which produce accuracy of 87.23 percent and 87.51 percent, respectively. Furthermore, when continuous data was used in the M5-MT and GWLM-NARX models, the NARX model performed better, with mean squared error (MSE) and regression value (R) predictions of 3.12 percent and 0.9899 percent in training, 0.144 percent and 0.9936 percent in validation, and 0.311 percent and 0.9988 percent in testing.

* Department of Computer Sciences, University of Kashmir, J&K, India, skh.amir88@gmail.com, ermuheet@gmail.com, kaulsameer78@gmail.com
** Directorate of IT & SS, University of Kashmir, J&K, India, zamanmajid@gmail.com

1. INTRODUCTION

Rainfall forecasting is useful in preventing floods, which saves lives and property. In reality, it aids in the management of the water supply. Rainfall data from the past has aided farmers in better managing their crops, resulting in increased economic growth in the country. Weather forecasting is difficult for climatological scientists because of the variability in rainfall timing and volume. Precision rainfall modeling is becoming one of the most pressing topics in hydrology, as early warning of extreme weather can help prevent natural disasters and losses if forecasts are made quickly and accurately. Forecasting is one of the most difficult problems for experts in a variety of domains, including meteorological data mining (Yang et al., 2007), and statistical forecasting (Pucheta et al., 2009). A typical question in such situations is how to analyze the past and make future forecasts. Even over a short period, the characteristics needed to predict rainfall are extremely complicated and delicate.

Traditional approaches that apply statistical techniques to examine the link between rainfall, geographic coordinates, and other atmospheric parameters have been used to forecast rainfall for years (like pressure, temperature, wind speed, and humidity). Rainfall, on the other hand, is difficult to anticipate due to its complexity, such as its nonlinearity (Wu & Chau, 2013). As a result, Singular Spectrum Analysis, Empirical Mode Decomposition, and Wavelet analysis, among other techniques, have been used to minimize non-linearity (Xiang et al., 2018). However, the mathematical and statistical models used need a lot of computational power (Singh & Borah, 2013; Singh et al., 2015) and can be time-consuming with little impact.

Weather prediction has gotten more successful in recent years in resolving a topic that has perplexed mankind for centuries, yet precise and timely weather forecasting remains a problem for scientists. Weather and weather forecasting are two subjects that practically everyone is interested in. Weather forecasting may therefore be anticipated utilizing many applicable machine learning methodologies in current era where machine-learning techniques are applied in every industry.

On various raw datasets in various places, numerous machine-learning algorithms have been constructed to forecast the amount of rainfall. In one of our studies (Kaul et al., 2023), comparisons between DT, DDT, and RF were done. We intended to apply more algorithms in order to select the best one in terms of accuracy measurement based on this comparison. This encourages us to work with the same dataset and assess the efficacy of different methods. This research compares the performance of models based on Logistic Model Tree, Gradient Boosting, GWLM-NARX and M5 Model Tree Networks to that of the original decision tree model and a projected model that will be the outcome of using an Automated Machine Learning tool (Mohd, Butt & Baba, 2022).

This paper is organised as follows: Section 1 provides a brief explanation of rainfall and its influence on agriculture, the environment, and the different machine learning techniques used to predict rainfall in JK. Section 2 is a quick overview of the literature. In section 3, numerous geographical setting and the satellite map photos have been evaluated which determines the climate of JK. Section 4 contains a description of the content and dataset, while sections 5 and 6 outline the algorithmic framework and technique. Section 7 outlines the experimental setup and assessment of several machine-learning methods. Finally, section 8 concludes the report with recommendations for the future.

2. REVIEW OF LITERATURE

The majority of weather prediction research is done using numerical approaches. The primary focus of this study will be on prior research based on various classical and ensemble machine learning algorithms utilized in rainfall prediction. Decision trees (DT), logistic model trees (LMT), M5 model trees (M5-MT), Gradient Boosting, and GWLM-NARX models are among these techniques.

Authors (Adnan, 2021), provided a comparison analysis of four machine learning techniques for rainfall modelling. The capabilities of OPELM, MARS, and the M5 model tree in daily precipitation modelling is the topic of this research. It was discovered that accuracy improves dramatically, with RMSE and MAE improvements of more than 90% in most cases. MARS-K also surpasses the other options tested in this study.

A research on groundwater level predictions using MARS and M5 Model Tree machine learning techniques was proposed by (Rezaie-balf et al., 2017). The data spans almost ten years, from August 1996 to July 2006. Validation of the models is done using the parameters utilized in this study. Validation was done using statistical performance assessment parameters such as RMSE, NNSE, and Coefficient of Determination.

Mohd et al. (Mohd, Butt & Baba 2020) developed a time series prediction model use the GWLM-NARX model. With rainfall data from the preceding period as input and results derived using the NARX model's GWLM algorithm, this model was employed as an adaptive forecast model.

(Fayaz, Zaman & Butt, 2021b) use a stepwise machine learning technique to estimate rainfall in India's Kashmir area. They used an LMT technique in their research, where the leaf node predicts model functions using logistic-regression approaches. The data for their analysis came from the Indian meteorological department in Pune, and it covered the years 2012 to 2017. Season, temperature at various intervals, humidity from 12 a.m. – 3 p.m., and rainfall were some of the variables studied. The study finishes with a comparison analysis in which the performance of several traditional and ensemble techniques is compared to that of the LMT, demonstrating that the LMT's accuracy measure is far superior to the other models utilized in the study.

(Fayaz, Kaul, Zaman & Butt, 2022) again used the same labelled data set as was used in (Zaman & Butt, 2012; Fayaz, Zaman & Butt, 2021a). The use of an ensemble distributed decision tree for rainfall prediction is defined in this work. The dataset was separated into three portions in this analysis depending on the station id. The performance of each decision tree was calculated after each decision tree was formed. A final accuracy was estimated based on the voting technique of the three smaller decision trees. The resulting accuracy was then compared to the accuracy of the original decision tree. The accuracy of the distributed decision tree was found to be much lower than the accuracy of the original decision tree.

Kaul et al. (2022) performs the comparative study on same set of data used in (Zaman & Butt, 2012; Fayaz, Zaman & Butt, 2022b, 2022c). In this research a comparsion has been made between Decision tree, Distributed decision tree (DDT) and Random forest (RF) on the geographical dataset. This research concluded that the decision tree performs better accuracy results in comparsion to the DDT and RF.

Furthermore, various network models for monthly rainfall rate forecasting and climate change were proposed in (Fayaz, Zaman & Butt, 2021c, 2022a) and the proposed models' performance was found to be extremely effective. The findings of the trial indicate that the accuracy rates will improve.

Since then, numerous writers have used the tabular dataset to estimate rainfall using a range of classical and ensembled methodologies, and some of them are described in this paper. From the literature, we may deduce that no one technique outperforms others on diverse types of datasets.

3. GEOGRAPHICAL SETTING AND CLIMATE OF J&K

In the state summary, the climatology of the UT of "Jammu and Kashmir" is described in terms of various meteorological parameters such as temperature, rainfall, rainfall variability, pressure, winds, relative humidity, clouds, weather hazards, and so on, followed by a detailed description of the climate of each district taking geographical and topographic characteristics into account as shown below (Figure 1).

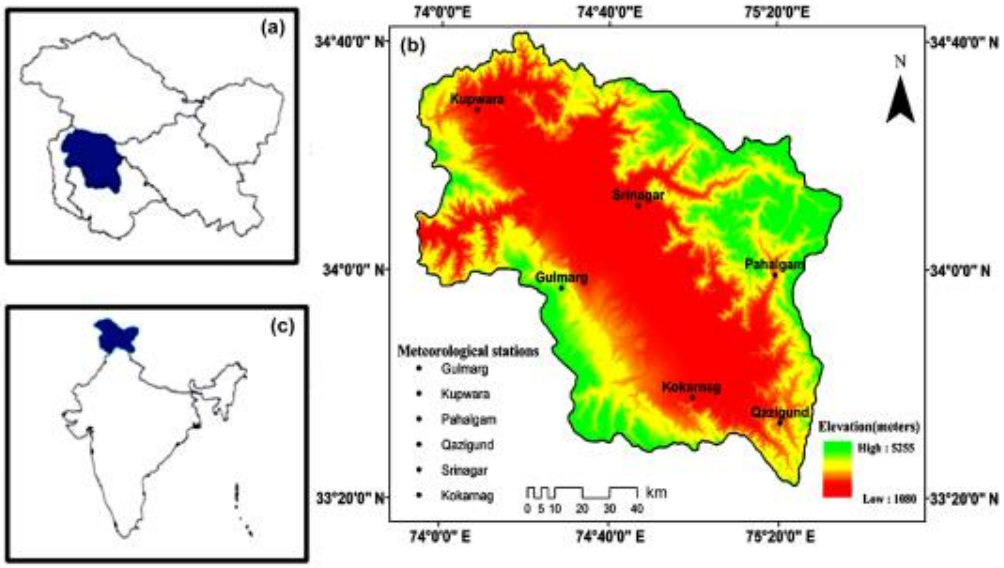


Fig. 1. Orographic and Geographical setting of the Kashmir valley (Zaz, 2019)

The Indian Union territory of Jammu and Kashmir, which is located in the Himalayas, may be divided into two climatic regions: Jammu and Kashmir. The winter capital is Jammu, whereas the summer capital is Srinagar. The greatest time to visit the 'Kashmir' region is between April and October, when the weather is nice in the summer and frigid in the winter. During the summer, the valley is blanketed in blooms, and the orchards are overflowing with fruit. Winter, from October to March, is the greatest season to explore the 'Jammu' region, with maximum temperatures about 18 degrees Celsius and minimum temperatures as low as 4 degrees Celsius on some days. The weather in Kashmir is nice, with temperatures ranging from 14 to 30 degrees Celsius. Although some days might be a little hot, the evenings are typically comfortable.

4. DATASET DESCRIPTION

Kashmir is classified as a temperate zone. As a result, the data were separated into four seasons to compute seasonal means: winter (December to February), spring (March to May), summer (June to August), and autumn (September to November). The seasonal rainfall in centimeters for each season is depicted in the graphs below figure (figure 2) (winter, spring, autumn, and summer) (Zaz et al., 2019).

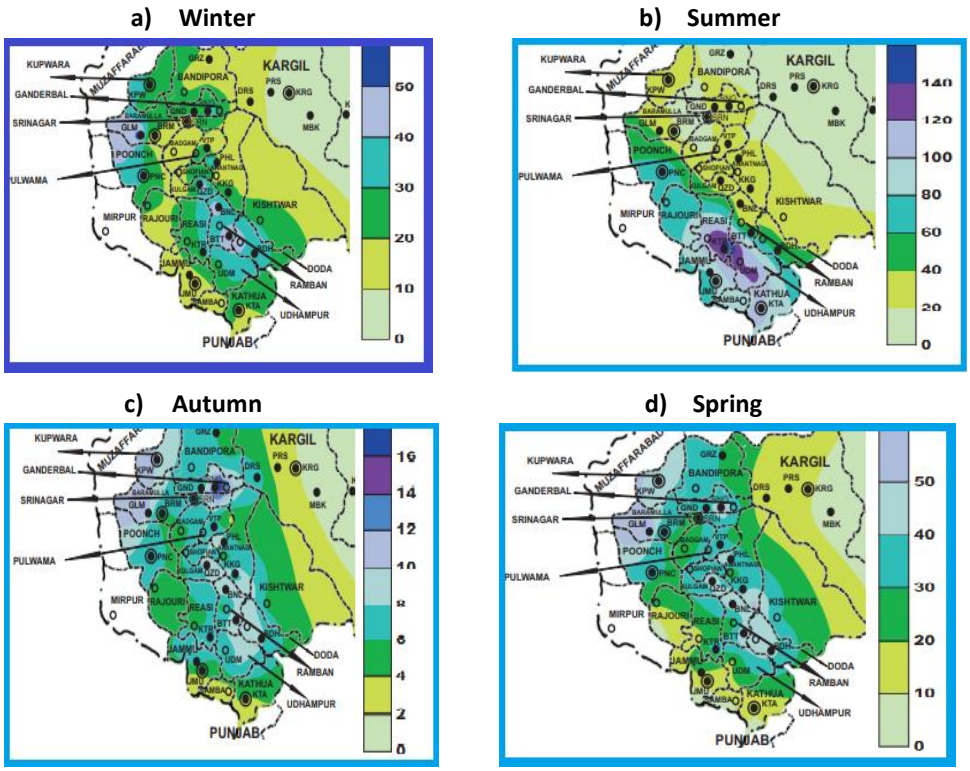


Fig. 2. Average Seasonal rainfall (cm) for each season in Kashmir province

For all six sites, the India Meteorological Department supplied data for five years (2012–2017) of daily precipitation, maximum and lowest temperatures, and humidity measurements at various time intervals. Central zone (34.0837° N, 74.7973° E), North zone (34.0837° N, 74.7973° E), and South zone (34.0484° N, 74.3805° E) are the three primary zones of Kashmir province. Figure 3 shows the overall structure of the dataset used in this study (Zaman & Butt, 2012).

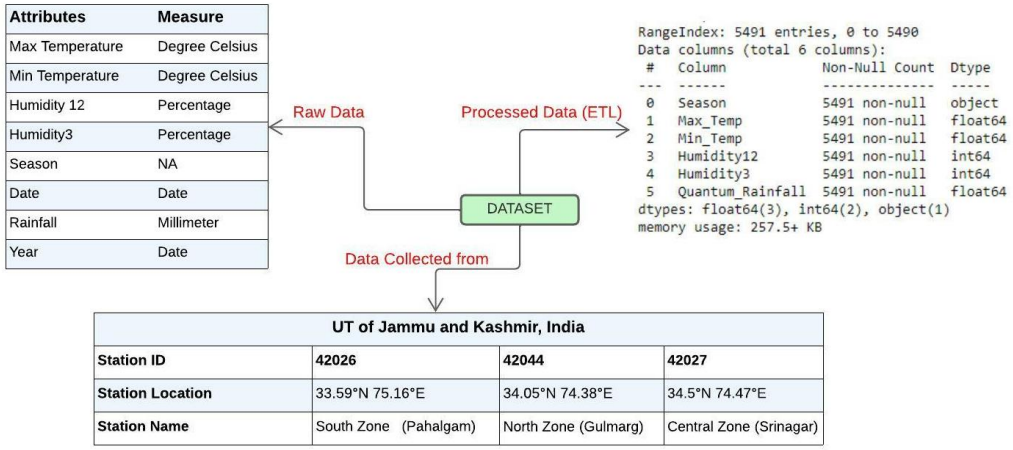


Fig. 3. Meteorological dataset of Kashmir province

To establish long-term trends and turning points of meteorological parameters with statistical significance, statistical tests such as kurtosis, cumulative deviation, and t test were used. The statistics of the data used in this study is shown in a tabular form (Table 1 and Table 2).

Tab. 1. Statistical properties of data

Attributes	t-test	Mean Difference	CI (Lower Bound)	Std. Dev	Skewness	CI (Upper Bound)	Kurtosis
Max Temp (°C)	151.8693	18.0409	17.808	8.80	-0.24	18.2738	-0.86
Min Temp (°C)	63.2285	6.3435	6.1469	7.43	0.02	6.5402	-0.84
Humid12 (%)	247.434	60.2723	59.7947	18.0	0.21	60.7498	-0.73
Humid3 (%)	396.5277	75.6416	75.2676	14.1	-0.76	76.0156	0.40
Rf (mm)	22.5342	2.7579	2.518	9.07	7.75	2.9979	99.4

Tab. 2. Correlation matrix and P value of the geographical attributes

First Column	Second Column	Correlation value	P value
Max	Min	0.879289995	0
Max	Hum12	-0.301475201	9.52E-116
Max	Hum3	-0.259730257	2.34E-85
Max	Rf	-0.190566402	4.56E-46
Min	Hum12	-0.141257364	7.16E-26
Min	Hum3	-0.107901977	1.08E-15
Min	Rf	-0.035400555	0.008704384
Hum12	Hum3	0.961517542	0
Hum12	Rf	0.008394317	0.534008313
Hum3	Rf	0.001847154	0.891152955

We employed around 6000 instances of meteorological data in this work, which included five parameters: humidity at 12 a.m., humidity at 3 p.m., maximum temperature, minimum temperature, and one goal parameter rainfall, which determines the amount of rain. Several metrics are employed in the model training, validation, and hyperparameter search to measure the validity of the machine learning algorithm's predictions. The best measure is determined totally by the task at hand. In binary and multiclass classification tasks, accuracy and kappa metrics are among the most often employed metrics. The proportion of properly categorized observations in relation to the total number of predictions is called accuracy, and kappa is the normalized accuracy value in relation to the predicted percentage of hits (Fayaz, Zaman, Kaul & Butt 2022).

5. ALGORITHMS FRAMEWORK

Several algorithms were tested for goodness of fit in this research. These algorithms include:

1. Decision trees: The decision tree (DT) is a data aggregation approach proposed in (Zaman & Butt, 2012) that is regarded as one of the most precise general-purpose tools. It entails making many judgments on samples from a data set obtained by random sampling with replacement (Banday et al., 2022).
2. Logistic Model Trees: At the leaves, Logistic Model Trees (lmt) mix model trees with logistic regression procedures. The logistic regression models that may identify important features in the data are built via a stage wise fitting approach (Aguasca-Colomo, Castellanos-Nieves & Méndez, 2019).
3. M5 Model trees: The M5 model tree is made up of two steps: a traditional decision tree and a linear regression function. To begin, the regression tree is constructed using the decision tree induction procedure. The standard deviation at each node will be determined to assess the predicted reduction in error for the splitting criterion. This node splitting in M5 will continue until there are very few instances left. Second, after constructing the normal regression tree, internal sub nodes are pruned and replaced with the regression plane rather than constant values (Niu & Zhang, 2015).

Gradient boosting: Gradient boosting is a flow process in which the original data used for prediction is given to the base model, which performs the first prediction. The error will be computed once this predicted output is compared to the actual output. The next decision tree is created based on the error, with only independent parameters considered and residuals for target parameters employed (error) (Barrera-Animas, 2022).

Several frameworks are now utilized to deal with predictive models, such as python packages such as TensorFlow, pytorch, Keras, or Scikit-learn. We utilized the caret software (Classification and Regression Training) in this study. Caret is an interface that combines many machine-learning tools into a single framework, making data preparation, training, optimization, and validation of predictive models easier, as well as native support for parallel computations.

GWLN-NRAX model: The GWLM method is a hybridization of the grey wolf optimization (GWO) and levenberg-marqueret (LM) algorithms with nonlinear autoregressive model (NARX) that is utilized for effective and adaptable rainfall forecast (Fayaz, Zaman & Butt, 2021c).

6. METHODOLOGY

In this paper, we show how the model is applied to meteorological data of kashmir province utilizing three traditional and ensemble approaches which includes Decision tree (DT), logistic model trees (LMT), Gradient Boost (GBoost), GWLM-NARX and M5 model trees (M5-MT). The datasets in these methods have been separated into (70-15-15) percent training, validation and test sets, respectively. This data splitting was done in Python using the sklearn split model. All the above three models follow the same basic approach in the implementation processes and we have provided a brief discussion of the machine learning strategies utilized in the prediction model development. These different machine learning methods are implemented on same set of geographical data of Kashmir province which includes different parameters like humidity at different intervals, temperatures and seasons and most importantly the target parameters rainfall.

Our approach includes the following steps, which result in an adequate prediction model: 1. training (apply a machine-learning algorithm to the training data set so that the model learns), 2. validation (predict the error of a statistical model with new data), and 3. prediction approach. Figure 4 shows the flowchart for the proposed methodologies (Aftab, Butt & Zaman, 2018; Afolayan, Ojokoh & Falaki, 2016).

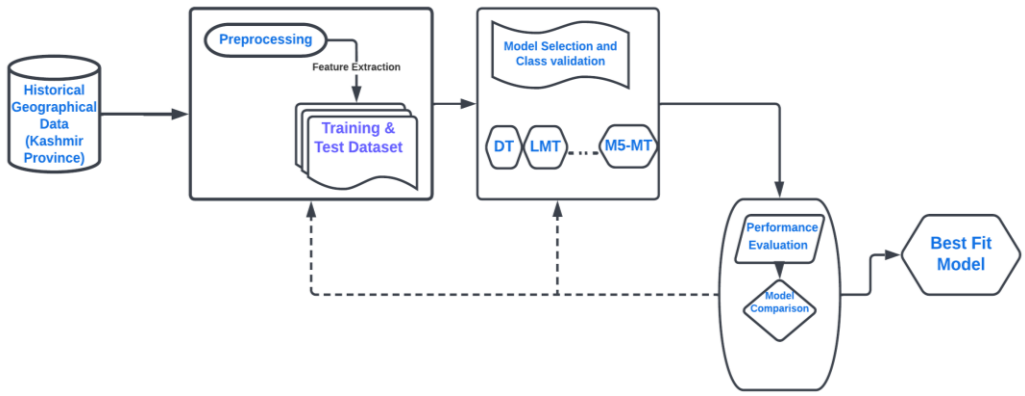


Fig. 4. Proposed Methodology

7. EXPERIMENTAL SETUP

In this work, the researchers used cutting-edge technologies on geographical datasets to test the algorithm with the greatest overall performance and accuracy. When the accuracy measure of the LMT and Gradient Boost (GBoost) were compared, it was shown that there is some increase in performance when compared to the original decision tree. Also, in the case of NARX and M5-MT models, the performance measure appears to be similar with lower mean absolute error rate (MAE), but the key difficulty with the methods is that training the data takes a long time because these models deal with continuous data streams. Table 3 displays a snapshot of the results, including accuracy, precision, recall values, and several other computations. The ML techniques chosen enable the development of a prediction model capable of representing the patterns contained in the training data set and generalizing them to new findings as shown in below table.

Tab. 3. Algortihms used with various statistical measures

Algorithms	Accuracy	Error	No of Classification Rules	Cohen Kappa	Precision
Original Decision Tree	80.12%	19.87%	51	0.456	0.812
LMT	87.23%	12.77%	10	0.102	0.893
Gradient Boost	87.51%	12.49%	–	0.073	0.914
M5-MT	R2= 0.478; MAE = 1.689; MSE = 6.726; RMSE = 2.593; MSD = 0.844				
GWLM-NARX	Regression (R) Testing: 0.9988%; Validation: 0.9936%; Training: 0.9899% MSE Testing: 0.311% ; Validation: 0.144%; Training: 3.12%				

The following graphical depiction (Figure 5) of geographical data from the Kashmir region aids in the simple visualization of the conclusions obtained. Other approaches are also highly efficient, but they require a huge amount of training data to train on in order to predict a relatively little amount of test data. Figure 5 is the visual representation of the table 3 where the maximum precision value of gradient boosting can be seen as 0.914 and highest accuracy level of 87.51%. Furthermore the accuracy level of LMT and GB remains head to head and remains much better than Decision tree.

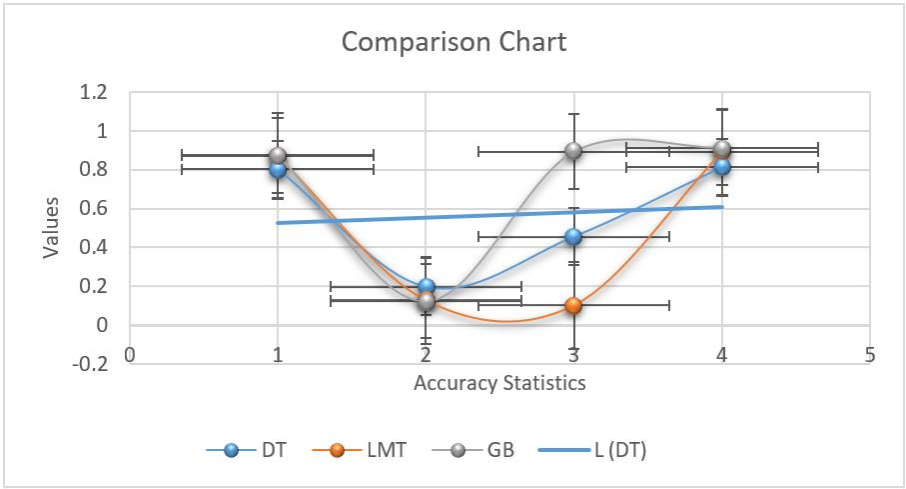


Fig. 5. Line graph: defines the accuracy statistics of each algorithm

Determining the optimum model is not an easy task; there are many of approaches, each with its own set of features and different parameters that must be changed. An unbalance in the probabilities of the observed classes is one of the classification difficulties that can have a major detrimental influence on the model's efficiency. An unbalanced data set is one in which the categorization categories are not roughly evenly represented. A potential solution to such class imbalance is to reconstruct the original the training data in a way that mitigates these issues.

To prevent having a detrimental influence on the prediction models, imbalances in the frequencies of the observed classes were handled. Because of the random sample, the minority class is the same size as the majority class. Predictors were preprocessed in order to interact with the ML algorithm or increase their performance. In order to predict how the

model would perform with unknown data, the dataset was partitioned into 80-20 training and testing ratio respectively (Dhamodaran & Lakshmi, 2021; Altaf, Butt & Zaman, 2022).

8. CONCLUSIONS AND FUTURE SCOPE

In this paper, we utilized and compared various well-known ML systems for rainfall prediction. This research compares the performance of various machine-learning machines and discusses potential applications. This proposal includes prediction models that are both accurate and easy to understand. This research is driven by the need to simplify and improve the process of rainfall prediction, as well as to solve the difficulties that existing solutions involve. As a result, the following are the primary contributions of this paper: a) Generating and comparing rainfall forecast models using various machine learning approaches; and b) Determining if the combination of local meteorological factors, and the algorithms utilized affects the predictive algorithm's accuracy.

The total accuracies achieved by the original Decision tree, LMT, and Gradient boosting models are 81.12%, 87.23%, and 87.51%, respectively. As a result, we can conclude that the LMT and Gradient Boosting models show significant improvement in predicting the class labels as compared to original decision tree, and they show head-to-head accuracy and prediction and can thus be considered as promising techniques for the prediction of rainfall in temperate zones such as Kashmir province. Furthermore, in case of M5-MT and GWLM-NARX models where continuous data was taken into consideration, NARX model was proven to be better with mean squared error (MSE) and regression value (R) predictions of 3.12 percent and 0.9899 percent in training, 0.144 percent and 0.9936 percent in validation, and 0.311 percent and 0.9988 percent in testing.

Since we concluded that GWLM-NARX performs better results based on the geographical dataset of the Kashmir province, the performance on other threshold datasets, such as academic data, health data, and other geographic data, has not been determined. It will be a future suggestion of this study to check the performance of these implemented algorithms on a wide range of datasets.

Conflicts of Interest

The authors have no conflicts of interest to declare.

REFERENCES

Adnan, R. M., Petroselli, A., Heddam, S., Santos, C. A. G., & Kisi, O. (2021). Comparison of different methodologies for rainfall–runoff modeling: machine learning vs conceptual approach. *Natural Hazards*, *105*(3), 2987–3011.

Afolayan, H. A., Ojokoh, B. A., & Falaki, S. O. (2016). Comparative analysis of rainfall prediction models using neural network and fuzzy logic. *International Journal of Soft Computing and Engineering*, *5*(6), 4–7.

Aftab, S., Ahmad, M., Hameed, N., Bashir, M. S., Ali, I., & Nawaz, Z. (2018). Rainfall prediction using data mining techniques: A systematic literature review. *International journal of advanced computer science and applications*, *9*(5), 143–150.

Aguasca-Colomo, R., Castellanos-Nieves, D., & Méndez, M. (2019). Comparative Analysis of Rainfall Prediction Models Using Machine Learning in Islands with Complex Orography: Tenerife Island. *Applied Sciences*, *9*(22), 4931. https://doi.org/10.3390/app9224931

Altaf, I., Butt, M. A., & Zaman, M. (2021). A Pragmatic Comparison of Supervised Machine Learning Classifiers for Disease Diagnosis. In *2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA)* (pp. 1515–1520). IEEE. https://doi.org/10.1109/ICIRCA51532.2021.9544582

Banday, I.R., Zaman, M., Quadri, S.M.K., Fayaz, S.A., Butt, M.A. (2022). Big data in academia: A proposed framework for improving students performance. Revue d'Intelligence Artificielle, Vol. 36, No. 4, pp. 589–595. https://doi.org/10.18280/ria.360411

Barrera–Animas, A. Y., Oyedele, L. O., Bilal, M., Akinosho, T. D., Delgado, J. M. D., & Akanbi, L. A. (2022). Rainfall prediction: A comparative analysis of modern machine learning algorithms for time-series forecasting. *Machine Learning with Applications*, *7*, 100204. https://doi.org/10.1016/j.mlwa.2021.100204

Dhamodaran, S., & Lakshmi, M. (2021). Comparative analysis of spatial interpolation with climatic changes using inverse distance method. *Journal of Ambient Intelligence and Humanized Computing*, *12*(6), 6725–6734. https://doi.org/10.1007/s12652-020-02296-1

Fayaz, S. A., Kaul, S., Zaman, M., & Butt, M. A. (2022). An adaptive gradient boosting model for the prediction of rainfall using ID3 as a base estimator. *Revue d'Intelligence Artificielle*, *36*(2), 241–250. https://doi.org/10.18280/ria.360208

Fayaz, S. A., Zaman, M., & Butt, M. A. (2021a). To ameliorate classification accuracy using ensemble distributed decision tree (DDT) vote approach: An empirical discourse of geographical data mining. *Procedia Computer Science*, *184*, 935–940. https://doi.org/10.1016/j.procs.2021.03.116

Fayaz, S. A., Zaman, M., & Butt, M. A. (2021b). An application of logistic model tree (LMT) algorithm to ameliorate Prediction accuracy of meteorological data. *International Journal of Advanced Technology and Engineering Exploration*, *8*(84), 1424–40.

Fayaz, S. A., Zaman, M., & Butt, M. A. (2021c). A hybrid adaptive grey wolf Levenberg-Marquardt (GWLM) and nonlinear autoregressive with exogenous input (NARX) neural network model for the prediction of rainfall. *International Journal of Advanced Technology and Engineering Exploration*, *9*(89), 509–522. https://doi.org/10.19101/IJATEE.2021.874647

Fayaz, S. A., Zaman, M., & Butt, M. A. (2022a). Numerical and Experimental Investigation of Meteorological Data Using Adaptive Linear M5 Model Tree for the Prediction of Rainfall. *Review of Computer Engineering Research*, *9*(1), 1–12.

Fayaz, S. A., Zaman, M., & Butt, M. A. (2022b). Knowledge Discovery in Geographical Sciences—A Systematic Survey of Various Machine Learning Algorithms for Rainfall Prediction. In *International Conference on Innovative Computing and Communications* (pp. 593–608). Springer.

Fayaz, S. A., Zaman, M., & Butt, M. A. (2022c). Performance Evaluation of GINI Index and Information Gain Criteria on Geographical Data: An Empirical Study Based on JAVA and Python. In *International Conference on Innovative Computing and Communications* (pp. 249–265). Springer.

Fayaz, S. A., Zaman, M., Kaul, S., & Butt, M. A. (2022). Is Deep Learning on Tabular Data Enough? An Assessment. *International Journal of Advanced Computer Science and Applications*, *13*(4), 2022. http://dx.doi.org/10.14569/IJACSA.2022.0130454

Kaul, S., Fayaz, S. A., Zaman, M., & Butt, M. A. (2022). Is decision tree obsolete in its original form? A burning debate. *Revue d'Intelligence Artificielle*, *36*(1), 105–113.

Kaul, S., Zaman, M., Fayaz, S. A., & Butt, M. A. (2023). Performance Stagnation of Meteorological Data of Kashmir. In *International Conference on Innovative Computing and Communications. Lecture Notes in Networks and Systems* (vol. 471). Springer. https://doi.org/10.1007/978-981-19-2535-1_63

Mohd, R., Butt, M. A., & Baba, M. Z. (2020). GWLM–NARX: grey wolf levenberg–marquardt-based neural network for rainfall prediction. *Data Technologies and Applications*, *54*(1), 85–102. https://doi.org/10.1108/DTA-08-2019-0130. 2020.

Mohd, R., Butt, M. A., & Baba, M. Z. (2022). Grey Wolf-Based Linear Regression Model for Rainfall Prediction. *International Journal of Information Technologies and Systems Approach*, *15*(1), 1-18.

Niu, J., & Zhang, W. (2015). Comparative analysis of statistical models in rainfall prediction. In *2015 IEEE International Conference on Information and Automation* (pp. 2187-2190). IEEE.

Pucheta, J. A., Cristian, M. R. R., Martín, R. H., Carlos, A. S., Patiño, H. D., & Benjamín, R. K. (2009). A feed-forward neural networks-based nonlinear autoregressive model for forecasting time series. *Comput y Sistemas*, *14*(4), 423–435.

Rezaie-balf, M., Naganna, S. R., Ghaemi, A., & Deka, P. C. (2017). Wavelet coupled MARS and M5 Model Tree approaches for groundwater level forecasting. *Journal of hydrology*, *553*, 356–373.

Singh, P., & Borah, B. (2013). Indian summer monsoon rainfall prediction using artificial neural network. *Stochastic Environmental Research and Risk Assessment*, *27*(7), 1585–1599.

Singh, U., Chauhan, S., Krishnamachari, A., & Vig, L. (2015). Ensemble of deep long short term memory networks for labelling origin of replication sequences. In *2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)* (pp.1–7). IEEE. http://dx.doi.org/10.1109/DSAA.2015.7344871

Wu, C., & Chau, K.-W. (2013). Prediction of rainfall time series using modular soft computing methods. *Engineering Applications of Artificial Intelligence*, *26*(3), 997–1007. https://doi.org/10.1016/j.engappai.2012.05.023

Xiang, Y., Gou, L., He, L., Xia, S., & Wang, W. (2018). A SVR–ANN combined model based on ensemble EMD for rainfall prediction. *Applied Soft Computing*, *73*, 874–883. https://doi.org/10.1016/j.asoc.2018.09.018

Yang, Y., Lin, H., Guo, Z., & Jiang, J. (2007). A data mining approach for heavy rainfall forecasting based on satellite image sequence analysis. *Comput Geosci*, *33*(1), 20–30.

Zaman, M., & Butt, M. A. (2012). Information translation: a practitioners approach. In *World Congress on Engineering and Computer Science* (WCECS).

Zaz, S. N., Romshoo, S. A., Krishnamoorthy, R. T., & Viswanadhapalli, Y. (2019). Analyses of temperature and precipitation in the Indian Jammu and Kashmir region for the 1980–2016 period: implications for remote influence and extreme events. *Atmospheric Chemistry and Physics*, *19*(1), 15-37. https://doi.org/10.5194/acp-19-15-2019

Applied Computer Science, vol. 18, no. 4, pp. 28–52
doi: 10.35784/acs-2022-27

Submitted: 2022-06-27 | Revised: 2022-12-09 | Accepted: 2022-12-11

Keywords: EEG, Eye Tracker, engagement, IGA, Advergaming

Konrad BIERCEWICZ [0000-0003-1635-4527]*,
Mariusz BORAWSKI [0000-0001-8470-8329]*,
Anna BORAWSKA [0000-0001-6680-1052]*, *Jarosław DUDA* [0000-0001-9971-3075]*

DETERMINING THE DEGREE OF PLAYER ENGAGEMENT IN A COMPUTER GAME WITH ELEMENTS OF A SOCIAL CAMPAIGN USING COGNITIVE NEUROSCIENCE TECHNIQUES

Abstract

Due to the popularity of video games in various applications, including both commercial and social marketing, there is a need to assess their content in terms of player satisfaction, already at the production stage. For this purpose, the indices used in EEG tests can be used. In this publication, a formula has been created based on the player's commitment to determining which elements in the game should be improved and for which graphic emblems connected with social campaigns were more memorable and whether this was related to commitment. The survey was conducted using a 2D platform game created in Unity based on observations of 28 recipients. To evaluate the elements occurring in the game at which we obtain a higher memory for graphic characters, a corresponding pattern was created based on player involvement. The optimal Index for moving and static objects and the Index for destruction were then selected based on the feedback. Referring to the issue of graphic emblems depicting social campaigns should be placed in a place where other activities such as fighting will not be distracted, everyone will be able to reach the level where the recently placed advertisement is. This study present the developed method to determine the degree of player's engagement in particular elements in the game using the EEG and to explore the relationship between the visibility of social advertising and engagement in a 2D platform game where the player has to collect three keys and defeat the ultimate opponent.

1. INTRODUCTION

Since the time of the mass market entry, computer games have been used as an advertising tool. In the 1980s, they were based on modifications of available games and served to entertain a small, limited group of people. In those days, product placement in computer games was limited by technological possibilities. However, over time, the attractiveness of games, their quantity, the graphic possibilities have improved and the

* Finance and Management Faculty of Economics, University of Szczecin, Szczecin, Poland, konrad.biercewicz@usz.edu.pl, mariusz.borawski@usz.edu.pl, anna.borawska@usz.edu.pl, jaroslaw.duda@usz.edu.pl

number of players around the world has increased, which has attracted potential advertisers. Initially, computer games were aimed mainly at computer enthusiasts, but now they have become a mass product, aimed at different types of players, in different age ranges, with different interests. A breakthrough product that changed the player's profile was The Sims, launched in 2000, which attracted women and girls. This has opened up a new space for advertising activities (*Lokowanie produktów w grach komputerowych – Prawo własności intelektualnej*, b.d.).

In 2015, profits from game advertising alone amounted to $2.75 billion. Profits are expected to increase to $5.05 trillion in 2020. (Statista, 2017). We distinguish two groups of promotion tools that can be used in computer games (Gałuszka, 2016):

- IGA (in-game advertising) means using the game as an advertising medium;
- Advergaming (advertising + games) are advertising games created specifically to advertise a particular product or service.

There are different ways of using the game as a means of advertising. The easiest way is to place advertising billboards. In 3D games, advertising billboards are placed according to the same rules as in the real world. An example is the 2008 advertising campaign of President Barack Obama, which used 18 games, including Need for Speed: Carbon. In the game, the player drives a car and the presidential candidate's ads are displayed on large billboards along the road (Yenigun, 2012).

Advergaming, on the other hand, is a custom-made game that aims to advertise certain goods or services. In 2006 Burger King launched the Sneak King game created by Blitz Games. The game was sold with a meal at a special price of 3.99$. In total, three games were created, which were not well received by critics, but generated huge revenues. They made a significant contribution to the company's sales growth. The main character of the game was the royal toy Burger King, and his task was to deliver burgers to hungry people (Hyman, 2007).

The discussed ways of presenting content in games also apply to social campaigns. Advergaming is a more common method used because among the benefits of this approach we can mention (The motion monkey, 2017):

- combining advertising and entertainment,
- interactivity – the recipient of an advertising message is a participant, not a passive viewer,
- creating positive associations,
- long life of the computer game (compared to standard advertising).

Besides, it is important that all the elements in the game are associated with the goal of the campaign. As a result, the recipient can memorize most of the presented material in a shorter time. An example of such a game created for a social campaign is Food Force. It was published by the UN World Food Programme (WFP) in 2006. Players take part in hunger prevention missions. In this way, they learn about the problems associated with hunger and WFP activities (United Nations World Food Programme, 2017).

A campaign does not always have to involve spectacular ventures like the Food Force game – sometimes even a small element, if well prepared, can make a difference. Regardless of the scale of the actions taken, the overriding goal of each campaign is to change attitudes or behaviour among its recipients. To achieve this goal, it is extremely important to prepare the visual identification of the campaign. One of the important elements of such identification

may be the campaign emblem – a simple image (sign or graphic), clearly linked to the promoted views or behaviours, which will be easy to remember. This method is used in in-game advertising.

Therefore, in the case of the in-game advertising method, it is necessary to place static ads appropriately and the player should reach such a level of commitment to the game that he plays it for as long as possible. Playing the game longer will give us a higher probability that social advertising will be noticed and thus remembered. The situation is similar in the case of TV commercials - once displayed, an advertisement will be remembered by a handful of people, but if it is repeated several dozen times, the number of recipients who will remember it will increase. What matters in the game is "its life expectancy". The number of views of the advertisement depends on the player's engagement, because the more the player gets involved, the longer he plays it, and consequently, the longer he plays it, the more often he watches the advertisement (Hofman-Kohlmeyer, 2017). The second aspect is the location of advertising (Gałuszka, 2016). It should be placed in such a way that it is well visible and at the same time does not affect the game. Besides, social campaign placement in the game attracts attention and makes the virtual space closer to reality - the game is more realistic (Kiedy reklama flirtuje z gamingiem..., 2015). The research presented in the article concerns the player's engagement, while the second problem concerning the placement of advertisements requires separate research.

1.1. Measuring Engagement

Player engagement is one of the dimensions of gaming experience and can be associated with many concepts, such as (Filsecker & Kerres, 2014; Schoenau-Fog, 2011): flow (Chen, 2007; Csikszentmihalyi, 1990), gameflow (Sweetser & Wyeth, 2005), presence (Lombard & Ditton, 1997; Tamborini & Skalski, 2006), immersion (Brown & Cairns, 2004; Jennett et al.., 2008; Mcmahan, 2003), pleasure (Costello & Edmonds, 2009), motivation (Iacovides et al., 2011; Przybylski, Scott Rigby & Ryan, 2010; Yee, 2006), enjoyment (Ijsselsteijn et al., 2008), arousal (Ravaja et al., 2006) and fun (Koster & Wright, 2004). Therefore, for the recipient to feel as many positive emotions as possible, it is necessary to maintain the player's commitment and concentration at a certain level, e.g. by introducing unexpected action phrases that will encourage him to continue exploring further areas of the game. First of all, to assess whether the game does not discourage the participant, it is necessary to conduct research on the participant's engagement in the game. Also, the growing community of video players creates demand among game developers or researchers for examining player engagement.

It may seem that evaluating a player's engagement in a video game is quite an easy task. We can think that, because there are ready-made tools on the market, such as Unity Analytics (Unity - Analytics, b.d.) and Game Analytics (GameAnalytics, b.d.), that analyze players behaviour. Thanks to them we are able to monitor, among others:

- Onboarding – whether players use mechanisms such as tutorials or initial levels?
- Progression – whether players pass the game levels?
- Economics – does the game economy work as expected?
- Design accuracy – does the designed game work the right way?
- Application accuracy – are all areas of the application used as expected? Are there elements that players ignore or do not notice?
- Earnings – are the strategies for making money optimal?

However, they are not able to check in real-time e.g. what the player is looking at and what emotions accompany him.

It is necessary to look for such methods of engagement research that will allow determining the level of engagement at any time in the game, while not being dependent on other factors. This would be possible with the use of cognitive neuroscience techniques. They are becoming more and more useful because they allow us to get to know the current state of the brain. This task is facilitated by the indexes calculated based on the recorded signals. In the literature on the subject, numerous indices of engagement can be found, which will be presented later in this chapter. They allow us to know the level of human engagement in a given activity in a given moment.

New developments in Brain-Computer Interfaces (BCI) using wireless electroencephalographic systems (EEG) provide recording and access to neuronal activity, enabling the computer to retrieve and analyse information from brain waves. It has been demonstrated that EEG has the ability to determine the engagement of the user. The frequency bands are determined from the EEG signal using the spectral method. Besides, the EEG can be used to change the game scenario for the player (Hondrou & Caridakis, 2012). For example, the detection of boredom will cause changes in the game to make it more challenging whereas the detection of anxiety will cause the game to slow down or decrease the levels of difficulty.

Using the EEG and eye tracker device we can determine the preferences of the player, as well as the moment of the game which is not very interesting, and we can improve it to make the player fully active in the game. New EEG devices are increasingly being used outside of medicine and are finding more and more new applications.

1.2. Related works

Using the EEG to measure the commitment of tasks is not a new concept. Pope (Pope, Bogart & Bartolome, 1995) built a system to control the level of automation of tasks based on whether the operator had increased or decreased his engagement. Freeman (Freeman et al., 1999) extended this system by evaluating the performance of each task with the use of absolute values of commitment. Berk (Berka et al., 2007) has invented a more accurate and effective method for people to interact with technology, with the ability to develop more productive work environments that increase motivation and productivity. The results suggest that the commitment measured using the EEG reflects information gathering, visual processing, and attention allocation. Smith and Gevins (Smith & Gevins, 2005) used a flight simulator to study the reactions of the human brain to low, medium, and high difficulty exercises. Studies have shown increased activity of the frontal lobe waves together with decreased activity of parietal lobe alpha waves during demanding tasks. In turn, Yamada (Yamada, 1998) measured the activity of theta waves along with blinking of the eye and discovered that children playing video games had higher activity of theta waves during more frequent blinking. These results suggest that interesting tasks cause higher activity of theta waves, while the task inhibits the activity of blinking eyes. Kamzanova (Kamzanova et al., 2011) compared the sensitivity of a series of EEG engagement indices by examining time pressure individuals performing tasks of varying degrees of stress to determine which one was most effective. McMahan (McMahan, Parberry & Parsons, 2015) investigated in Super Meat Boy game whether there is a connection between engagement and arousal in events of death and general entertainment. The results of their research suggest that by combining engagement data with arousal data, we can establish

thresholds indicating when a player has left the flow state. On the other hand, Ewing (Ewing, Fairclough & Gilleade, 2016) investigated the sensitivity of EEG power in the (front) theta and (parietal) alpha bands to changing levels of demand for play. Besides, he also conducted a study that assessed the adaptive performance of Tetris in terms of system behavior and user experience. Vourvopoulos's (Vourvopoulos i in., 2017) research focuses on the impact of how gaming experience has on modulating brain activity, as an attempt to systematically identify elements that contribute to high BCI control and that can be used in the design of a neurogame.

The above-mentioned research (McMahan, Parberry & Parsons, 2015; Vourvopoulos et al., 2017) is examining player engagement, but it focuses on topics related to dependencies or BCI. Therefore, the article attempts to solve the following problems:

- to examine which elements in the game should be modified on the basis of the developed method based on the engagement index
- whether placing banners next to engaging elements affects their memorization

The aim of the article is to present the developed method to determine the degree of player's engagement in particular elements in the game using the EEG (engagement indexes) and to explore the relationship between the visibility (an Eye Tracker was used to check visibility) of social advertising and engagement in a 2D platform game where the player has to collect three keys and defeat the ultimate opponent. This will allow determining at which obstacles the advertisement was more memorable and what should be improved in the game so that the player can spend more time in it.

2. MATERIALS AND METHODS

The experiment was attended by 32 people, but only 28 people joined the main survey. Four people were not taken into account because the specified target group had to be mainly occasional players and those who liked 2D platform games. The results of survey are presented in Tab. 1. In addition, respondents had to list three things that they like in playing games (story – about 70%, collecting items – about 40%), and also whether they like 2D platform games – everyone answered positively.

Tab. 1. Results of the survey

Question	Answers	Number of people
How often do you play computer games?	Every weekend	8
	Occasionally	7
	Rarely	13
What player would you describe yourself?	Commonly known gamer	3
	Reactional player	25
What's your preference for the way you play?	Single Player	21
	Multiplayer	7

2.1. Description of game

The game was downloaded from the Unity Asset Store ("2D Game Kit") (2D Platformer - Asset Store, b.d.) and adapted for testing in the Unity engine. Before the start of the game, there were instructions on how to move (using arrows or WASD and key space for jumping) and on the goal that need to be achieved, i.e. three keys had to be collected, which guaranteed the entrance to the room where the last opponent was located, the so-called the boss (Fig. 1).

Fig. 1. A screenshot of the game showing the fight with the boss

In addition, the game also features emblems that show the various real situations to be prevented (see

Tab. ***2***). The emblems were in the background, so the player did not have to avoid them. They were used in the experiment because often in computer games, due to the resolution of the graphics, very complex advertisements are not clear enough to be noticed properly (Chang et al., 2010; Yang et al., 2006). There are also social campaigns, using simple symbols to promote their activities (Rubin, b.d.).

The purpose of the emblems' placement is to examine whether such signs in a typical game are noticeable and whether this type of advertising is effective for use in a social campaign.

Tab. 2. Description of signs

Name of sign	Emblem	Description
Pregnant woman		Prohibition of drinking alcohol during pregnancy
Key		No driving on alcohol
Alcohol		Prohibition of drinking alcohol
Stop		Shows the inability to drive under the influence of alcohol
Bottle		Ban on drinking alcohol
Syringe		Ban on using drugs
Tap		Ban on drinking tap water
Drug		Ban on using drugs
Key2		Prohibiting drinking when driving
Selfie		Prohibiting taking pictures

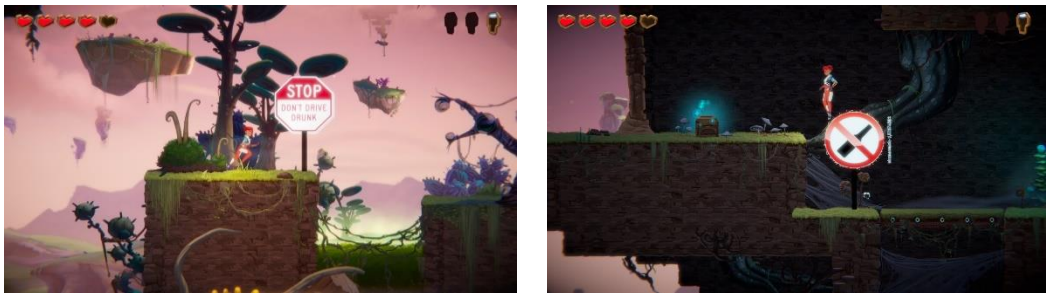
On the first level (Zone 1) there is a sign – “pregnant woman” (Fig. 2) when moving on to the next level.

Fig. 2. Social advertisements on level 1 – Zone 1

On the second level, there are two graphic emblems (Fig. 3): the "key" and "alcohol" – at the doorway to the last level of Zone5 where the boss is waiting to be defeated. You have to collect three keys to getting there. Each key collected is followed by a close-up on the door.

Fig. 3. Social advertisements on level 2 – Zone2

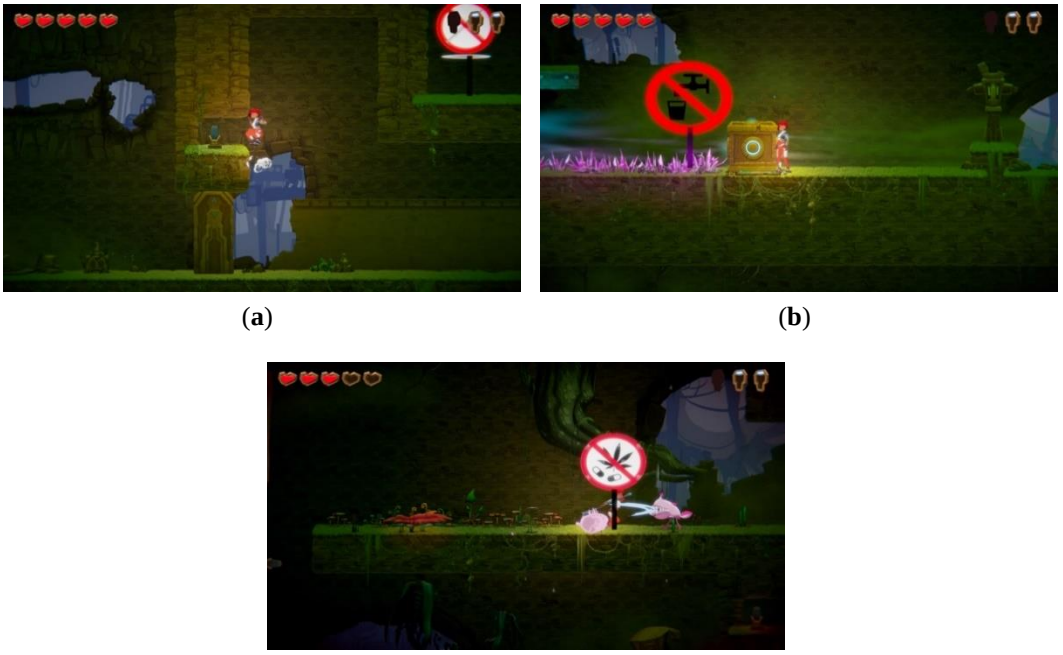
On the third level, there are two signs – "stop" and "bottle" (Fig. 4) – the first one (a) was located near water and the second one (b) near column destruction.



(**a**) (**b**)

Fig. 4. Social advertisements on level 3 – Zone 3

On the fourth level, there are three graphics (Fig. 5). The first one – "syringe" was placed in a safe place when opening the door (a), the second – "tap" in the danger of spikes (b), and the third – "drug" when fighting the Chomper monsters (c).



(**a**) (**b**)

(c)

Fig. 5. Social advertisements on level 4 – Zone4

Two prohibitions have been placed on the last level – "selfie" and "key2" immediately after entering the level (Fig. 6).

Fig. 6. Social advertisements on level 5 – Zone5

2.2. Test procedure of the experiment

EEG data were collected from 28 healthy people (4=female, 24=male), the average age was 23 years. The persons were informed about the course of the examination. They then signed their consent to participate in the study and were seated in a comfortable chair with access to the keyboard and mouse. The next step was to put on the cap and connect the electrodes to the participant's scalp and connect them to the data recorder of the participant's brain.

The cap (g.Nautilus Research Wearable EEG Headset) with 24 electrodes placed in AF3, AF4, F3, F4, F7, F8, FC5, FC6, P7, P8, T7, T8, O1, O2, P3, C3, C4, Pz, Fz, Cz, FPz, Fp1 P4, POz, and 3 reference electrodes: AFz, FCz, CPz was used (see: Fig. 7). The channels have been distributed according to the 10–10 system, the international EEG electrode distribution system (Jurcak, Tsuzuki & Dan, 2007). The electrodes required a dampened socket to improve conductivity. In order to check whether the EEG electrodes are in good contact with the scalp, impedance values were measured with the g.Recorder program. The sampling frequency was 500 Hz.

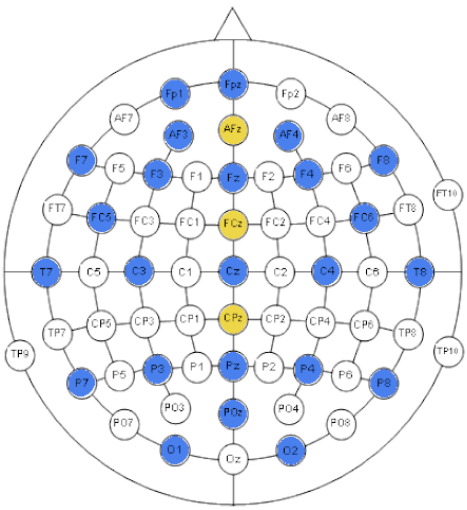


Fig. 7. Sensor location on headsets

After the above steps have been taken, a study was started. Before the game, there was information shown what the game will be about, its goal, and how to move around. Then, after clicking the “Play” button, a black screen appeared, lasting 60 seconds, during which the participant silenced himself (see: Fig. 8). The recorded signals during the play were used to calculate the EEG indicators. On the basis of the respondent's engagement, concentration, and response (in relation to the game and the remembered ads), a comparison of responses and indicators was made. This made it possible to create a pattern based on engagement, more precisely specifying the relationship between engagement for a given element and the engagement that accompanied the whole level. The difficulty in defeating monsters or obstacles) in the game needs to be improved to keep the engagement in the game at a certain level. Also, it made it possible to investigate which social advertisements have been remembered and whether this is related to engagement. the effectiveness of social advertising placed.

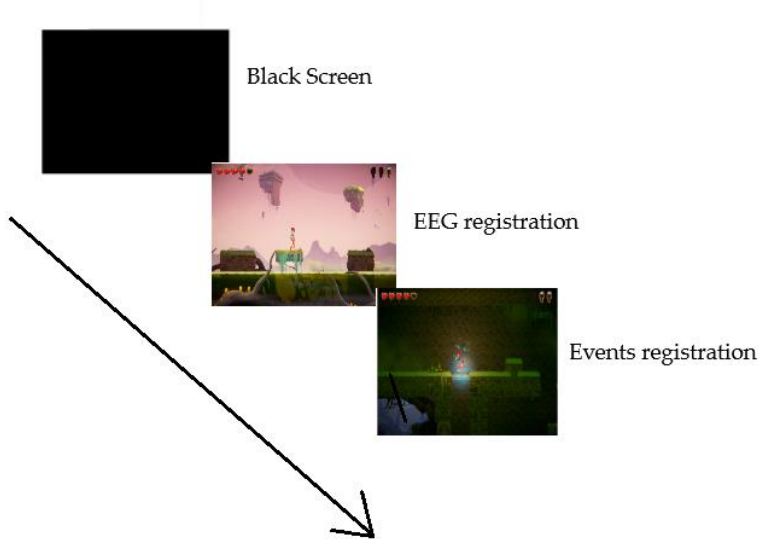


Fig. 8. Chronological order of events while playing a computer game

Each participant's game was recorded at a resolution of 1920×1080 using programmed in-game registration. During the game, screenshots were taken at a frequency of 3 shots per seconds. Each screenshot generated a timestamp for the EEG data to determine the start and end position of each section. Screenshots were saved for later use during the data analysis phase. In addition to the EEG, the study used the Eye Tracker (Eyetribe eye-tracker with a frequency of 30 points per second) to track image elements that were particularly important to the respondent.

2.3. Game survey

After the game was finished, an interview was conducted about the impressions of playing the game, which stage in the game was the hardest, which elements should be improved, what was the engagement during the levels and the fight. The last questions focused on remembering the social advertisements emblems.

2.4. Measures

Measuring the level of engagement and memory of a player is one of the elements determining their experience while playing a computer game. In particular, it can be used to determine player preferences. For this purpose, the relevant indicators have been calculated, as indicated in Tab. 3. These indices were selected for the study because they were used by their developers to study player engagement in computer games or simulations. They were then normalized against the reference image - the black screen, and the spider that appeared on the screen.

Tab. 3. Description of the indices used in the test

Index number	Formula	Counting method
Index 1 (McMahan et al., 2015)	Beta-3 / (Alpha-2 + Theta)	Average registration value of all electrodes on the head.
Index 2 (McMahan et al., 2015)	Theta / Alpha-2	Average registration value from electrodes placed on the frontal lobe of Theta and parietal lobe of Alpha.
Index 3 (McMahan et al., 2015)	Theta	The average value of registration from electrodes placed on the frontal lobe of Theta.
Index 4 (Kamzanova et al., 2011)	Beta-4 / (Alpha-3 + Theta)	Average registration value from electrodes: F3, F4, F7, F8, Cz, P3, Pz, P4.
Index 5 (Kamzanova et al., 2011)	Beta-5 / (Alpha-4 + Theta)	Average registration value from electrodes: F3, F4, F7, F8, Cz, P3, Pz, P4.
Index 6 (Chaouachi & Frasson, 2012)	Beta-6 / (Alpha-5 + Theta)	Average registration value from electrodes: P3, C3, Pz, Fz, Cz, FPz.
Index 7 (Lee et al., 2010)	(SMR + Mid Beta) / Theta	Power index from electrode Fp1.

3. RESULTS

All data were analysed using Matlab R2019a. The analysis of the EEG signal started with filtering the bandwidth removing the disturbances of the power network, i.e. frequencies above 50 Hz. In addition, the signal was detrended and filtered using the Fieldtrip library. The EEG spectral signal was then analysed using a Morse wave, which calculated an average peak frequency of half a second in a frame (Lilly & Olhede, 2010, 2012; Wachowiak et al., 2018). However, in order to calculate the Alpha, Beta and Theta frequencies, the signal has been divided into appropriate bands (Tsipouras, 2019), see Tab. 4.

Tab. 4. Representation of frequency bands Alpha, Beta, Theta

Bandwidth	Frequency [Hz]
theta	4–8
alpha2	7–13
alpha3	8–13
alpha4	8–10.9
alpha5	11–13.9
SMR	12–15
Mid beta	15–20
beta3	13–25
beta4	13–22
beta5	14–19.9
beta6	20–29.9

The next step was to analyse the variance to see if there is a significant difference between player engagement and concentration. The ANOVA statistical test was used for this purpose. The engagement was divided into seven subgroups - Index1, Index2, up to Index7. This creates different conditions and allows to use repeated ANOVA. The results of repeated ANOVA measurements using the indicator as an object factor for the general concentration-dependent variable showed a significant difference between the engagement indices ($F(6,580)=18.37$; $p= 1.42*10^{-19}$; $p < 0.05$). The Turkey-Kramer statistical test was used to investigate between them. The test showed that there is a statistically significant difference between all groups (Tab. 5).

Tab. 5. Results of Tukey-Kramer statistical test

Index Number	Index Number	Lower limit for the 95% confidence interval of the true difference in average	The difference between the estimated averages	Upper limit for the 95% confidence interval of the true difference in average	Value of p
1	7	0.1234	0.2980	0.4725	1.0058e-05
2	7	0.1755	0.3501	0.5247	1.0565e-07
3	7	0.0637	0.2383	0.4129	0.0011
4	7	0.1498	0.3244	0.4989	9.2470e-07
5	7	0.3331	0.5077	0.6822	3.7064e-08
6	7	0.3635	0.5380	0.7126	3.7064e-08

In addition, it was examined whether there is a dependence of the commitment on concentration (

Tab. **A1**), using the CHI-Squared independence test. The analysis showed six engagement indices that there is no relationship between engagement and concentration (Tab. 6). The statistical value for each index is greater than the value of the Chi-Squared distribution for 4 degrees of freedom, which is $\chi_4^2 = 9{,}48773$.

Tab. 6. Results of the CHI-Squared statistical test.

Index	Statistical value CHI Square	Number of degrees of freedom
Index1	150,0472	4
Index2	142,0259	4
Index3	130,8962	4
Index4	126,3576	4
Index5	121,7529	4
Index6	152,8122	4

The results obtained from the analysis made it possible to create a formula (1), which is based on the engagement index

$$Assessment\ of\ the\ element = \frac{Engagement\ to\ the\ element}{Number\ of\ events} * \frac{1}{Engagement\ at\ a\ given\ level} \quad (1)$$

First, the player's level of engagement with the item in the game is determined. The result is then divided by the amount of engagement that accompanies the test person at a given level of the game in order to determine what was the average engagement at that particular moment. And finally, it was divided by the number of occurrences of a given element. This resulted in an average engagement for a particular element in the game. Taking into account whether the engagement level is low/medium/high (as will be determined later in the article), we can determine whether a given element in a 2D platform game should be corrected or removed.

Then, appropriate thresholds for low, medium and high levels of engagement (the breakdown that was used in Tables 7 to 14) were determined for each of the engagement indices. The intervals were calculated as follows:

1. The 10th and 90th percentile have been calculated for the engagement observed for each player at Zone1, Zone3, Zone4 and Zone5 levels.
2. The arithmetic mean was determined for all players from the results obtained in step 1.
3. For each level (Zone1, Zone3, Zone4 and Zone5) the difference between the 90th percentile and the 10th percentile has been calculated.
4. The obtained range was divided into 3 in order to obtain three ranges – for low, medium and high engagement

Zone2 level was not included in the calculation because there were no obstacles or enemies – it was the so-called interim.

Formula created and calculated thresholds (**Błąd! Nie można odnaleźć źródła odwołania.**) of the engagement at a given level allowed for the evaluation of the following elements in the game:

- Moving Platform moving up and down, sideways, with spikes (Tab. C1). In most cases, the indices indicate average engagement. Index4 is the most distinct.
- Spikes - a static object, the hero loses his life if he enters it (
- Tab. **C2**). All the indices are consistent and indicate the average engagement when spike-bypassing players.
- Destruction of the column - a static column which, after being destroyed, leads to the next stage in the game (Tab. **C3**). Only Index2 indicates low exposures during the destruction of the column, the others indicate medium.
- Moving a box - moving an item triggers an action in the game, e. g. closure of the water gap (Tab. C4). Index4 indicates low engagement when moving the box, others indicate medium.
- Standing in a designated place causing the platform to move (
-). As with moving a box, only Index4 indicates low engagement when standing on a platform that is moving.
- Attack from "gun" – Key Down O (Tab. **C6**). Nearly all the indices indicate average engagement on shooting. The exception is the Zone4 and Index4 level, where it indicates high engagement.
- Attack by monsters – Boss (
- Tab. **C7**), Chomper, Spritter (
- Tab. **C8**). The indices for the attack by regular Chomper show mainly high level of engagement, while for Spritter mainly medium level. Referring to the last opponent, the so called "Boss" – you can see that attacking by Ball is at a medium level, Light at a high level, while Throwgranade is hard to say.

The final stage was to examine the visibility and memory of the social campaign's emblems placed in the game. For this purpose, the number of times an average person looked at a banner was calculated (Tab. 7). For the analysis of this objective, the eye-tracker and Matlab software were used, with the following scheme:

1. The eyesight of the test person was traced using EyeTracker.
2. File in txt format was generated, containing the viewing coordinates x, y and time unit
3. The Matlab program analyzed where:
   a. first of all, the emblem was found in every screenshot,
   b. secondly, within the designated area, it was calculated how many times the person looked and how much time person spent looking at the sign.

This made it possible to determine which location for the emblem in the game is the best and how much time it takes to remember the social campaign (Tab. 7), as well as whether we pay more attention to banner ads in case of greater engagement. This shows that most people noticed the emblem representing a pregnant woman – 25 people and the average time of looking at the advertisement was 2 seconds.

Tab. 7. Average time of looking at the sign.

Name of the advertisement	Average number of glances at the emblem	Average time of glances at the emblem	How many people have seen the sign on the basis of the survey	Level	How many people have reached the level
Pregnant woman	72	2,311378	25	Zone1	28
Key	27	0,879667	14	Zone2	28
Alcohol	13,73077	0,44902	10	Zone2	28
Stop	17,23077	0,565488	12	Zone3	28
Bottle	16,80769	0,527995	17	Zone3	28
Syringe	2,538462	0,081886	10	Zone4	22
Tap	47,26923	1,539069	20	Zone4	22
Drag	9,692308	0,323447	11	Zone4	22
Selfie	2,576923	0,08306	5	Zone5	9
Key2	5,153846	0,173572	13	Zone5	9

In addition, the memorization rate (Tab. 8) was calculated by dividing the number of people who saw the mark by the total number of participants in the study (see: formula (2)).

$$Memorable = \frac{Number\ of\ people\ seen\ the\ sign\ on\ the\ bassi\ of\ the\ survey}{28} \quad (2)$$

Tab. 8. Emblem memorization rate expressed as a percentage

Name of the advertisement	Memorization rate
Pregnant woman	89%
Key	50%
Alcohol	36%
Stop	43%
Bottle	61%
Syringe	36%
Tap	71%
Drag	39%
Selfie	18%
Key2	46%

4. DISCUSSION

The main aim of the study was to develop a method to determine the degree of player engagement in the individual elements of the game. In addition, additional objectives have been achieved, among which we can mention:

- to investigate which obstacles to social advertising have been more memorable,
- indication which elements in the game should be corrected.

It was helpful to create a formula (1) to evaluate the elements in the game and to determine which index should be used on the basis of comparing the results obtained with the opinion of the respondents. The constructed formula is based only on the engagement itself, because examining the engagement ratio for a given element to the engagement that is at a given level allows to determine whether a given element contributes significantly to the player's engagement. As a result, we have the answer as to whether we should correct or remove an item in the game. The ratio itself, by mathematical definition, is the quotient of one value to another, which is intended to indicate the identity or relative size difference of the two quantities. Therefore, the formula created seems to be quite good enough in mathematical as well as functional terms, as will be discussed later in this chapter.

Matlab software and survey data were used for analysis. As a result, the engagement in elements such as up-and-down, sideways, spiked moving platforms, for which Index4 should be used (Tab. C1,

Tab. **C2**), was assessed, as it is the only one reflecting the correct opinions according to the respondents. Interestingly, the only biggest difficulty in the game for people participating in the survey was jumping from platform to platform. The low engagement turned out to be the case with PushableBox, PreasurePad and Column Destruction. And these elements should certainly be improved, or their placement limited because there were too many of them - PushableBox and PreasurePad. For these two elements Index4 (Tab. C4,

), should be used for the assessment and for Column Destruction Index2 (

Tab. **C3**). In the game, it was the only item that had to be destroyed and it did not cause any difficulties, as evidenced by the commitment obtained. It seems that this element should be removed and not corrected, because some of the people surveyed about 40% were wrong in this case or wasted time because they did not know if they could destroy it. For the remaining elements, it was not clearly indicated which Index should be used, and this is the case for an object with a medium level of engagement (

Tab. **C2**) – all Indexes indicated this way. A similar situation is for the attacking monsters - Chomper, Spritter (

Tab. **C8**) and the final opponent, the so-called final boss (Tab. **C7**), who attacks with different powers. And the same is true to indicate which method of fighting is the best - even or with firearms. According to the respondents, the melee fight was better than shooting, but the engagement in both cases was the same, i. e. average, as shown by the formulated formula. From the results, it is not possible to determine which index should be used for this purpose, because in case of melee combat all indexes fit, and for shooting all but 4.

Referring to the results of social advertising, it was found that most people, because almost 90% of the respondents remembered the sign on which the pregnant woman was placed. The reason for this is to place the sign when jumping from platform to platform when, as confirmed by the survey, the player's engagement was high. In addition, from the platform where the sign was located, the next level was moved to the next level, which could also affect the memory. Moreover, the average time of looking at the sign was 2 seconds, which could also be relevant. In addition, the majority of respondents, i.e. 70% remembered the sign showing the ban on using drinking tap water, which was in the place where you jumped on the platform to avoid spikes. It can be assumed that the combination of medium engagement in the case of a spike element in the game and high for a moving platform helps to remember social advertising. On the other hand, the characters located on the horizontal Zone2, which was a transition level, were remembered by 43% of people on average. It would seem that there will be many more because each time the key is lifted, there is a close-up on the door and the graphic emblems next to it. This fact may also be influenced by the short time of looking at the advertisements, but also by the lack of action at this level. Least people, because only 5 remembered the sign representing the ban on taking pictures. He was at the door on the last level. This is because only 9 people have reached this level. On the other hand, 13 people indicated an advertisement with a ban on driving under the influence of alcohol on Zone5, which is apparently not true, but this may be due to the fact that a fairly similar sign was placed at the door, but on Zone2 level. Low memorization efficiency is also for a sign representing drugs at Zone4 level. Probably because the monsters absorbed more attention than the sign in the background that didn't stand out.

5. CONCLUSIONS

The results of research aimed at evaluating the elements in the game are presented, and at which we obtain greater memory of graphic emblems. In this achievement, it was decided to create an appropriate pattern based on the player's commitment. Then, on the basis of the opinions, the optimal Index4 for moving and static objects and Index2 for destruction were selected, as they best represent the opinion of the respondents.

It should be taken into account that these findings are based on one type of game and that further research will be needed to extend the results of the methodological approach to assessing which elements in the game need to be improved and at which ads need to be placed, not only by analysing the player's engagement but also by adding further indicators from other categories, such as emotions. However, these results support the view that Index 2 and 4 can be judged to be the best index to illustrate these elements in the game and for a high level of commitment, ads should be placed on the item to be remembered. Additionally, graphic emblems depicting social campaigns should be placed in a place where other activities such as fighting will not distract attention, everyone will be able to reach the level where the last placed advertisement is and the graphics of each advertisement should stand out from the others. As a result, we will get a better result for remembering among the respondents.

Conclusions on graphic emblems should be understood in the context of certain constraints. This is due to the fact that not all respondents have reached the final level and therefore the analysis is in some cases based on low sample size. At this point, a formula

was created for evaluating in-game elements based on the player's engagement and it was checked which index should be used for moving elements, static objects, and when destroying objects. The next step will be to prepare a revised game to establish the reliability of the selected Index in order to ensure that current performance is not an anomaly due to the current number of people who have reached a given level.

Author Contributions

K.B.: Conceptualization, software, validation, formal analysis, writing–original draft preparation, writing–review and editing; M.B, A.B.: Conceptualization, methodology, writing–review and editing, supervision; J.D.: software, validation, visualization. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

REFERENCES

2D Platformer—Asset Store. (b.d.). *Unity Asset Store*. Retrieved September 2, 2019 from https://assetstore.unity.com/packages/essentials/tutorial-projects/2d-platformer-11228

Berka, C., Levendowski, D. J., Lumicao, M. N., Yau, A., Davis, G., Zivkovic, V. T., Olmstead, R. E., Tremoulet, P. D., & Craven, P. L. (2007). EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, Space, and Environmental Medicine*, *78*(5 Suppl), B231–244.

Brown, E., & Cairns, P. (2004). A Grounded Investigation of Game Immersion. *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (pp. 1297–1300). The ACM Digital Library. https://doi.org/10.1145/985921.986048

Chang, Y., Yan, J., Zhang, J., & Luo, J. (2010). Online In-Game Advertising Effect: Examining the Influence of a Match Between Games and Advertising. *Journal of Interactive Advertising*, *11*, 63–73. https://doi.org/10.1080/15252019.2010.10722178

Chaouachi, M., & Frasson, C. (2012). Mental Workload, Engagement and Emotions: An Exploratory Study for Intelligent Tutoring Systems. *Intelligent Tutoring Systems. ITS 2012. Lecture Notes in Computer Science* (vol 7315). Springer Heidelberg. https://doi.org/10.1007/978-3-642-30950-2_9

Chen, J. (2007). Flow in Games (and Everything else). *Commun. ACM*, *50*(4), 31–34. https://doi.org/10.1145/1232743.1232769

Costello, B., & Edmonds, E. (2009). A Tool for Characterizing the Experience of Play. *Proceedings of the Sixth Australasian Conference on Interactive Entertainment*, *2*, 1–10. https://doi.org/10.1145/1746050.1746052

Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. Harper Perennial Modern Classics.

Ewing, K. C., Fairclough, S. H., & Gilleade, K. (2016). Evaluation of an Adaptive Game that Uses EEG Measures Validated during the Design Process as Inputs to a Biocybernetic Loop. *Frontiers in Human Neuroscience*, *10*, 223–223. https://doi.org/10.3389/fnhum.2016.00223

Filsecker, M., & Kerres, M. (2014). Engagement as a Volitional Construct: A Framework for Evidence-Based Research on Educational Games. *Simulation & Gaming*, *45*(4–5), 450–470. https://doi.org/10.1177/1046878114553569

Freeman, F. G., Mikulka, P. J., Prinzel, L. J., & Scerbo, M. W. (1999). Evaluation of an adaptive automation system using three EEG indices with a visual tracking task. *Biological Psychology*, *50*(1), 61–76.

Gałuszka, D. (2016). Nowy wymiar reklamy-in-game advertising oraz advergaming. *Kultura i Historia*, *29*, 33–49.

GameAnalytics. (n.d.). *GameAnalytics*. Retrieved April 5, 2020 from https://gameanalytics.com/

Hofman-Kohlmeyer, M. (2017). Komunikacja marketingowa w grach komputerowych—Współczesne kierunki badań. *Studia Ekonomiczne*, *328*, 70–82.

Hondrou, C., & Caridakis, G. (2012). Affective, Natural Interaction Using EEG: Sensors, Application and Future Directions. W I. Maglogiannis, V. Plagianakos, & I. Vlahavas (Eds.), *Artificial Intelligence: Theories and Applications* (s. 331–338). Springer Berlin Heidelberg.

Hyman, P. (2007). *Burger King has it their way with advergame sales*. Hollywood Report.

Iacovides, I., Aczel, J., Scanlon, E., Taylor, J., & Woods, W. (2011). Motivation, Engagement and Learning through Digital Games. *IJVPLE*, *2*, 1–16. https://doi.org/10.4018/jvple.2011040101

Ijsselsteijn, W., van den Hoogen, W., Klimmt, C., De Kort, Y., Lindley, C., Mathiak, K., Poels, K., Ravaja, N., Turpeinen, M., & Vorderer, P. (2008). Measuring the experience of digital game enjoyment. *Proceedings of Measuring Behavior* (pp. 88–89). Noldus.

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies*, *66*(9), 641–661. https://doi.org/10.1016/j.ijhcs.2008.04.004

Jurcak, V., Tsuzuki, D., & Dan, I. (2007). 10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems. *NeuroImage*, *34*(4), 1600–1611. https://doi.org/10.1016/j.neuroimage.2006.09.024

Kamzanova, A. T., Matthews, G., Kustubayeva, A. M., & Jakupov, S. M. (2011). EEG indices to time-on-task effects and to a workload manipulation (Cueing). *World Academy of Science, Engineering and Technology*, *80*, 19–22.

Kiedy reklama flirtuje z gamingiem... (2015, March 20). *Marketer+ przewodnik po marketingu*. https://marketerplus.pl/teksty/artykuly/reklama-flirtuje-gamingiem/

Koster, R., & Wright, W. (2004). *A Theory of Fun for Game Design*. Paraglyph Press.

Lee, C., Kwon, J., Hong, J., & Lee, D. (2010). *A Study on EEG based Concentration Power Index Transmission and Brain Computer Interface Application* (s. 537–539). Springer. https://doi.org/10.1007/978-3-642-03882-2_142

Lilly, J. M., & Olhede, S. C. (2010). On the Analytic Wavelet Transform. *IEEE Transactions on Information Theory*, *56*(8), 4135–4156. https://doi.org/10.1109/TIT.2010.2050935

Lilly, J. M., & Olhede, S. C. (2012). Generalized Morse Wavelets as a Superfamily of Analytic Wavelets. *IEEE Transactions on Signal Processing*, *60*(11), 6036-6041. https://doi.org/10.1109/TSP.2012.2210890

Lokowanie produktów w grach komputerowych—Prawo własności intelektualnej. (b.d.). Retrieved April 3, 2020 from https://www.pwi.us.edu.pl/kategorie/prawo-reklamy/253-lokowanie-produktow-w-grach-komputerowych?jjj=1586621566826&jjj=1586626966769

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, *3*(2).

Mcmahan, A. (2003). Immersion, engagement, and presence: A method for analyzing 3-D video games. *The Video Game Theory Reader* (pp. 67–86). Routledge, Taylor & Francis Group.

McMahan, T., Parberry, I., & Parsons, T. D. (2015). Evaluating Player Task Engagement and Arousal Using Electroencephalography. *Procedia Manufacturing*, *3*, 2303–2310. https://doi.org/10.1016/j.promfg.2015.07.376

Pope, A. T., Bogart, E. H., & Bartolome, D. S. (1995). Biocybernetic system evaluates indices of operator engagement in automated task. *Biological Psychology*, *40*(1–2), 187–195. https://doi.org/10.1016/0301-0511(95)05116-3

Przybylski, A. K., Scott Rigby, C., & Ryan, R. (2010). A Motivational Model of Video Game Engagement. *Review of General Psychology*, *14*, 154–166. https://doi.org/10.1037/a0019440

Ravaja, N., Saari, T., Salminen, M., Laarni, J., & Kallinen, K. (2006). Phasic Emotional Reactions to Video Game Events: A Psychophysiological Investigation. *Media Psychology*, *8*(4), 343–367. https://doi.org/10.1207/s1532785xmep0804_2

Rubin, B. (b.d.). *Policies & Programs to Reduce Distracted Driving*. Retrieved September 9, 2020 from https://www.government-fleet.com/156182/policies-programs-to-reduce-distracted-driving

Schoenau-Fog, H. (2011). The player engagement process—An exploration of continuation desire in digital games. *Proceedings of DiGRA 2011 Conference: Think Design Play*.

Smith, M., & Gevins, A. (2005). Neurophysiologic monitoring of mental workload and fatigue during operation of a flight simulator. *Proceedings of SPIE - The International Society for Optical Engineering* (5797). Society of Photo-Optical Instrumentation Engineers. https://doi.org/10.1117/12.602181

Statista. (2017). *Video games advertising spending worldwide from 2010 to 2020 (in billion U.S. dollars)*. https://www.statista.com/statistics/238140/global-video-games-advertising-revenue/

Sweetser, P., & Wyeth, P. (2005). GameFlow: A Model for Evaluating Player Enjoyment in Games. *Computers in Entertainment*, *3*(3), 3–3. https://doi.org/10.1145/1077246.1077253

Tamborini, R., & Skalski, P. (2006). The role of presence in the experience of electronic games. *Playing video games: Motives, responses, and consequences* (pp. 225–240). Lawrence Erlbaum Associates Publishers.

The motion monkey. (2017). *UK Advergame Design & Development*. https://www.themotionmonkey.co.uk/advergames/

Tsipouras, M. G. (2019). Spectral information of EEG signals with respect to epilepsy classification. *EURASIP Journal on Advances in Signal Processing*, *2019*(1), 10. https://doi.org/10.1186/s13634-019-0606-8

United Nations World Food Programme. (2017). *Food Force*. https://web.archive.org/web/20050605073447/http://www.food-force.com/

Unity—Analytics. (b.d.). *Unity*. Retrieved April 5, 2020 from https://unity3d.com/unity/features/analytics

Vourvopoulos, A., Bermudez, I., Badia, S., & Liarokapis, F. (2017). EEG Correlates of Video Game Experience and User Profile in Motor-imagery-based Brain—Computer Interaction. *The Visual Computer*, *33*(4), 533–546. https://doi.org/10.1007/s00371-016-1304-2

Wachowiak, M., Smolikova-Wachowiak, R., Johnson, M., Hay, D., Power, K., & Williams-Bell, F. (2018). Quantitative feature analysis of continuous analytic wavelet transforms of electrocardiography and electromyography. *Philosophical Transactions of The Royal Society A Mathematical Physical and Engineering Sciences*, *376*, 20170250. https://doi.org/10.1098/rsta.2017.0250

Yamada, F. (1998). Frontal midline theta rhythm and eyeblinking activity during a VDT task and a video game: Useful tools for psychophysiology in ergonomics. *Ergonomics*, *41*(5), 678–688. https://doi.org/10.1080/001401398186847

Yang, M., Ewoldsen, D., Dinu, L., & Arpan, L. (2006). The Effectiveness of „in-Game" Advertising: Comparing College Students' Explicit and Implicit Memory for Brand Names. *Journal of Advertising*, *35*, 143–152. https://doi.org/10.2753/JOA0091-3367350410

Yee, N. (2006). Motivations for play in online games. *Cyberpsychology & Behavior : The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society*, *9*(6), 772–775. https://doi.org/10.1089/cpb.2006.9.772

Yenigun, S. (2012). *Presidential Campaigns Rock The Gamer Vote: NPR*. https://www.npr.org/2012/10/01/162103528/presidential-campaigns-rock-the-gamer-vote?t=1582567815999

Appendix A

Tab. A1 Distribution of engagement by the group

Index1

		Low	Medium	High
		<= -0,64	(-0,64; -0,02]	> -0,02
High	> 0,69	0	1	7
Medium	(-0,17; 0,69]	4	33	5
Low	<= -0,17	5	25	3

Index2

		Low	Medium	High
		<= -0,64	(-0,64; -0,02]	> -0,02
High	> 0,34	0	6	12
Medium	(-0,33; 0,34]	4	48	3
Low	<= -0,33	5	5	0

Index3

		Low	Medium	High
		<= -0,64	(-0,64; -0,02]	> -0,02
High	> 0,34	0	4	9
Medium	(-0,34; 34]	3	46	6
Low	<= 0,34	6	9	0

Index4

		Low	Medium	High
		<= -0,64	(-0,64; -0,02]	> -0,02
High	> 0,23	0	11	6
Medium	(-0,27; 0,23]	5	41	9
Low	<= -0,27	4	7	0

Index5

		Low	Medium	High
		<= -0,64	(-0,64; -0,02]	> -0,02
High	> 0,53	7	4	1
Medium	(0,06; 0,53]	2	40	4
Low	<= 0,06	0	15	10

Index6

		Low	Medium	High
		<= -0,64	(-0,64; -0,02]	> -0,02
High	> 0,65	8	0	1
Medium	(0,08; 0,65]	1	49	2
Low	<= 0,08	0	10	12

Appendix B

Tab. B1 Distrubution of engagement by level in game

Zone1

	Index1	Index2	Index3	Index4	Index5	Index6
High	> 0,35	> 0,39	> 0,35	> 0,38	> 0,69	> 0,70
Medium	(-0,45; 0,35]	(-0,42; 0,39]	(-0,47; 0,35]	(-0,41; 0,38]	(-0,16; 0,69]	(-0,14; 0,70]
Low	<= -0,45	<= -0,42	<= -0,47	<= -0,41	<= -0,16	<= -0,14

Zone3

	Index1	Index2	Index3	Index4	Index5	Index6
High	> 0,38	> 0,46	> 0,35	> 0,39	> 0,70	> 0,68
Medium	(-0,40; 0,38]	(-0,36; 0,46]	(-0,44; 0,35]	(-0,38; 0,39]	(-0,13; 0,7]	(-0,10; 0,68]
Low	<= -0,40	<= -0,36	<= -0,44	<= -0,38	<= -0,13	<= -0,10

Zone4

	Index1	Index2	Index3	Index4	Index5	Index6
High	> 0,43	> 0,46	> 0,36	> 0,45	> 0,76	> 0,71
Medium	(-0,32; 0,43]	(-0,33; 0,46]	(-0,45; 0,36]	(-0,35; 0,45]	(-0,10; 0,76]	(-0,11; 0,71]
Low	<= -0,32	<= -0,33	<= -0,45	<= -0,35	<= -0,10	<= -0,11

Zone5

	Index1	Index2	Index3	Index4	Index5	Index6
High	> 0,64	> 0,62	> 0,56	> 0,51	> 0,90	> 0,75
Medium	(-0,08; 0,64]	(-0,14; 0,62]	(-0,27; 0,56]	(-0,20; 0,51]	(0,0012; 0,90]	(-0,08; 0,75]
Low	<= -0,08	<= -0,14	<= -0,27	<= -0,20	<= 0,0012	<= -0,08

Appendix C

Tab. C1 Assessment of engagement for the Moving Platform

	MovingPlatform			
	Zone1	**Zone3**	**Zone 4**	**Zone5**
Index1	high	medium	medium	medium
Index2	high	medium	medium	medium
Index3	średnie	medium	medium	medium
Index4	high	high	low	medium
Index5	medium	medium	medium	medium
Index6	medium	medium	medium	medium

Zone4		
	MovinPlatform – sideways	**MovingPlatform with spikes**
Index1	medium	medium
Index2	medium	high
Index3	medium	medium
Index4	medium	high
Index5	medium	medium
Index6	medium	medium

Tab. C2 Assessment of engagement for the Spikes

Zone4	
	Spikes
Index1	medium
Index2	medium
Index3	medium
Index4	medium
Index5	medium
Index6	medium

Tab. C3 Assessment of engagement for the destruction of the column

Zone1	
	CollisionEnter – DestructibleColumn_Whole
Index1	medium
Index2	low
Index3	medium
Index4	medium
Index5	medium
Index6	medium

Tab. C4 Assessment of engagement for PushableBox

Zone4	
	PushableBox
Index1	medium
Index2	medium
Index3	medium
Index4	low
Index5	medium
Index6	medium

Tab. C5 Assessment of engagement for PreasurePad

Zone4	
	PreasurePad
Index1	medium
Index2	medium
Index3	medium
Index4	low
Index5	medium
Index6	medium

Tab. C6 Assessment of engagement for attack from the gun

	KeyDown – O		
	Zone3	**Zone4**	**Zone5**
Index1	medium	medium	medium
Index2	medium	medium	medium
Index3	medium	medium	medium
Index4	medium	high	medium
Index5	medium	medium	medium
Index6	medium	medium	medium

Tab. C7 Assessment of engagement when the monsters Boss attacked

	Zone5 – Boss		
	Ball	**Ligh**	**Throwgrandae**
Index1	medium	high	high
Index2	medium	high	high
Index3	medium	high	medium
Index4	medium	high	high
Index5	medium	high	medium
Index6	medium	high	high

Tab. C8 Assessment of engagement when the monsters Chomper, Spritter attacked

	Attack Chomper		Attack Spritter
	Zone3	**Zone4**	**Zone4**
Index1	high	high	medium
Index2	high	high	medium
Index3	low	low	medium
Index4	high	low	low
Index5	high	high	medium
Index6	high	high	medium

Applied Computer Science, vol. 18, no. 4, pp. 53–67
doi: 10.35784/acs-2022-28

Submitted: 2022-09-07 | Revised: 2022-11-10 | Accepted: 2022-12-03

Keywords: singular value decomposition, JPEG2000, JPEG, discrete wavelet transform, discrete cosine transform

Edyta ŁUKASIK [0000-0003-3644-9769]*, *Emilia ŁABUĆ***

ANALYSIS OF THE POSSIBILITY OF USING THE SINGULAR VALUE DECOMPOSITION IN IMAGE COMPRESSION

Abstract

In today's highly computerized world, data compression is a key issue to minimize the costs associated with data storage and transfer. In 2019, more than 70% of the data sent over the network were images. This paper analyses the feasibility of using the SVD algorithm in image compression and shows that it improves the efficiency of JPEG and JPEG2000 compression. Image matrices were decomposed using the SVD algorithm before compression. It has also been shown that as the image dimensions increase, the fraction of eigenvalues that must be used to reconstruct the image in good quality decreases. The study was carried out on a large and diverse set of images, more than 2500 images were examined. The results were analyzed based on criteria typical for the evaluation of numerical algorithms operating on matrices and image compression: compression ratio, size of compressed file, MSE, number of bad pixels, complexity, numerical stability, easiness of implementation.

1. INTRODUCTION

The problem of data compression is a major one these days: as the capabilities of computing machines increase, so does the amount of data that is collected and processed. For example, the images are showcases for websites that will or will not attract a customer to take a look at the offer presented. Not only the good quality of the images displayed, but also the time it takes for them to load must be taken into account. It is therefore natural to be concerned with reducing the size of stored files, which has a significant impact on file transfer times. Many publications have been made on data compression. This topic is repeatedly discussed at scientific conferences (Jackson & Hannah, 1993; Jinchuang, Yan & Wenli, 2009; Nasri et al., 2010; Xiao et al., 2011) and in publications (Gandhi, Patel & Prajapati, 2015; Hoffman, 1997; Pu, 2005; Salomon, Motta & Bryant, 2007; Short, Manohar & Tilton, 1994; Shukla & Prasad, 2011; Wayner, 1999).

* Department of Computer Science, Faculty of Electrical Engineering and Computer Science, Lublin University of Technology, Lublin, Poland, e.lukasik@pollub.pl
** University of Maria Curie-Skłodowska, Faculty of Biology and Biotechnology, Institute of Biological Sciences, Lublin, Poland, emilia.labuc@mail.umcs.pl

Research conducted in 2018 shows that the ratio of images to all data transmitted over the network has been growing rapidly over recent years, with over 70% of data transmitted over the network in 2019 being images (Chen et al., 2020; Karwowski, 2019). A literature survey was conducted to determine the current state of knowledge on the topic of image compression and data compression in general. This has made it possible to identify image compression problems that have not yet been sufficiently researched, implemented or described. Such an issue is, among others, the possibility of using the matrix singular value decomposition (SVD) algorithm for image file compression, hence the author's desire to do research in this field.

There are publications indicating the expediency of such a solution (Cao, 2006; Compton & Ernstberger, 2020; Swathi et al., 2017), however, they are superficial, conducted on single images, without studying the real impact of SVD application on compression ratio, so those publications do not fully cover the subject. The main purpose of image compression is to reduce the space required for image storage so it is possible to minimize the amount of hardware needed to store images, which ultimately means a significant decrease in expenditure on storage of such data (Pratt, Kane & Andrews, 1969; Shih et al., 2012). It is also worth mentioning that thanks to smaller file size, the time needed for its transmission is reduced, meaning less bandwidth used which, of course, also reduces costs while increasing productivity. It is the standard now to store in databases and transmit compressed images. The most popular formats used today for image compression are: JPEG, PNG, GIF and TIFF used to exchange files between applications or computer platforms (Karwowski, 2019). Less commonly used image file compression methods are for example: PIXAR, HEIF, JBIG, PCX, PGF, XPM, EXIF, JPEG XL, JFIF, WEBP or WBMP (Miano, 1999; Murray & VanRyper, 1996).

Data compression is the process of encoding, restructuring or otherwise modifying data to reduce its size. Essentially it involves re-encoding information using fewer bits than its original representation (Hoffman, 2012; Sayood, 2022). An image can be defined as a discrete function that assigns to the coordinates of a pixel its color (and in some image file formats, also an alpha channel, defining transparency). Each pixel can have a different color than all the others. If an exact reproduction of the original image is to be achieved, a certain minimum amount of information has to be stored. Finding this minimum is the task of lossless compression (Arps & Truong, 1994; Gong et al., 2018). Lossless compression is used wherever there is a need for an exact reproduction of the original data. However, there is not always a need for perfect image reconstruction, because man does not have a perfect sight organ. Data that cannot be seen by humans can be omitted from a digital image recording.

Lossy compression is based on this assumption. In the literature (Mammeri, Hadjou & Khoumsi, 2012; Nixon & Aguado, 2019; What's the difference between 'visually lossless' and real lossless and what does this mean for future encodes?, n.d.) it is possible to come across the concept of visually lossless compression, it means de facto lossy compression, in which the differences of the compressed image in relation to the original image are negligible in visual assessment. One of the most significant factors affecting compression is the presence of redundancy in the data (Davies, 2017; Lu & Guo, 2016; Parekh, 2021). Redundancy can be related to coding (using fewer codewords than the optimal number causes coding redundancy), prediction of pixel values based on the values of all neighboring pixels. Visual redundancy, on the other hand, is related to the fact that

the human eye cannot process every frequency band. The pixels in the screen are in a grid form, so the image can be represented as a matrix of data.

Singular value decomposition (SVD) allows to approximate a data set with a large number of dimensions by reduced to the minimum number of dimensions (Dumka et al., 2020). By applying the decomposition to the matrix representing the image, redundancy can be used to the maximum extent: only the repeating part can be eliminated, so that the integrity of the image as a whole remains unchanged. The main objective of the paper is a comprehensive analysis of the possibility of using the matrix decomposition algorithm on singular values in image compression. The author has put forward the following research hypotheses:

The application of the SVD algorithm on the image matrix increases the efficiency of its compression by the JPEG2000 method. Applying the SVD algorithm on the image matrix increases the efficiency of its compression by the JPEG method. As the image dimension increases, the fraction of eigenvalues used for good quality image reconstruction decreases.

2. MATERIALS AND METHODS

2.1. Data

The data used to test the image compression algorithms are 2561 image files in the two most popular formats, specific to data containing image information: .BMP and .JPG. Most of the collected files are images obtained from Wikipedia under the Creative Commons license. Other images used for testing come from the author's private collection and include photos taken with different equipment and of different quality, screenshots, and digital art. Among the collected data there are binary images, monochromatic, as well as color images of different sizes. The color images vary, some are stored in high color palette. Most, however, are stored in true color with 24-bit depth, where each pixel is stored in 3 bytes. This is currently the most commonly used color depth. The smallest image is 100×67 pixels, meaning it contains information about 6700 pixels. Its size in .BMP format is 20154 bytes. The largest image examined has dimensions of 10200× 14039 pixels, containing information on more than 143 million pixels. The size of this image in .BMP format is 429 593 538 bytes. The compressed images are stored in two formats: .JPG and .JP2. The .JP2 extension, now rather rarely used, is a file format that has been compressed using the JPEG2000 method. This method allows both lossy and lossless compression (skipping the quantization step). .JP2 files are therefore larger than .JPG files.

2.2. SVD

Singular Value Decomposition is a factorization of the matrix A into three special matrices such that: $A = U \cdot \Sigma \cdot V^T$ is usually used to reduce the dimension of the original data (Britanak, Yip & Rao, 2007; Jankowska & Jankowski, 1988; Kostrikin, 2004). The SVD can be applied to image and signal processing or robotics. It is useful wherever there is a need to reduce the dimension of the original data: in statistics, in geographic data inversion, or in approximation theory. Theoretically, it can be applied to the calculation of inverse matrices (although there are better algorithms) (Cormen et al., 2005). The SVD

decomposition is sometimes used in statistics for factor analysis, more specifically as an alternative to PCA (Principal Component Analysis) (Stewart, 2001). There are publications confirming he possibility of using SVD decomposition in Big Data (Wayner, 1999; Dhawan, 2011). The algorithm for finding SVD involves: computing $A \cdot A^T$, finding the eigenvalues of $A \cdot A^T$, square rooting it and placing the diagonal Σ. The matrices U and V are calculated from the normalized eigenvectors of the matrices $A \cdot A^T$ and $A^T \cdot A$ respectively.

2.3. JPEG

Generally, a JPEG file can be encoded in various ways. The encoding process usually involves several steps as shown in Fig. 1. The most interesting step is the application of the DCT (Discrete Cosine Transform). The individual pixel values are replaced by an average value within the block and an average that determines the frequencies of change within the block, both averages are expressed as floating point numbers, so although the DCT transform is reversible, some information is lost due to a fair amount of rounding. The DCT increases the number of bits needed to store the pixel data because the DCT coefficients are floating point numbers that take up more space in the computer's memory (up to 16 bits instead of the standard 8, the number of bits depends on the accuracy of the DCT calculation). Such a temporary size increase is not a problem in most JPEG implementations, because usually only a small part of the image is stored in the full DCT form. The next step - quantization reduces these values back to 8 bits.

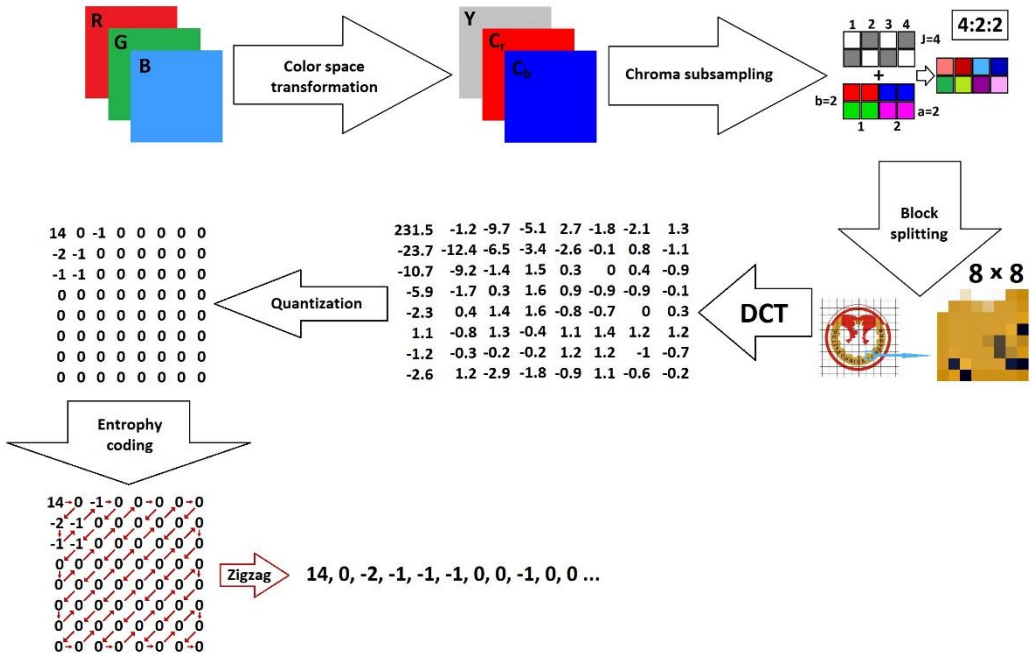


Fig. 1. JPEG codec example

2.4. JPEG2000

JPEG2000 is an image compression standard and encoding system developed in 2000 by the JPEG committee. Files using JPEG2000 are saved in the .jp2 format. It was originally intended to replace the previously used JPEG standard, but sluggishness of digital camera manufacturers, as well as web application developers (they didn't want to use the standard until it became more widespread) halted the development of JPEG2000 (Tanwar, Ramani & Tyagi, 2018).

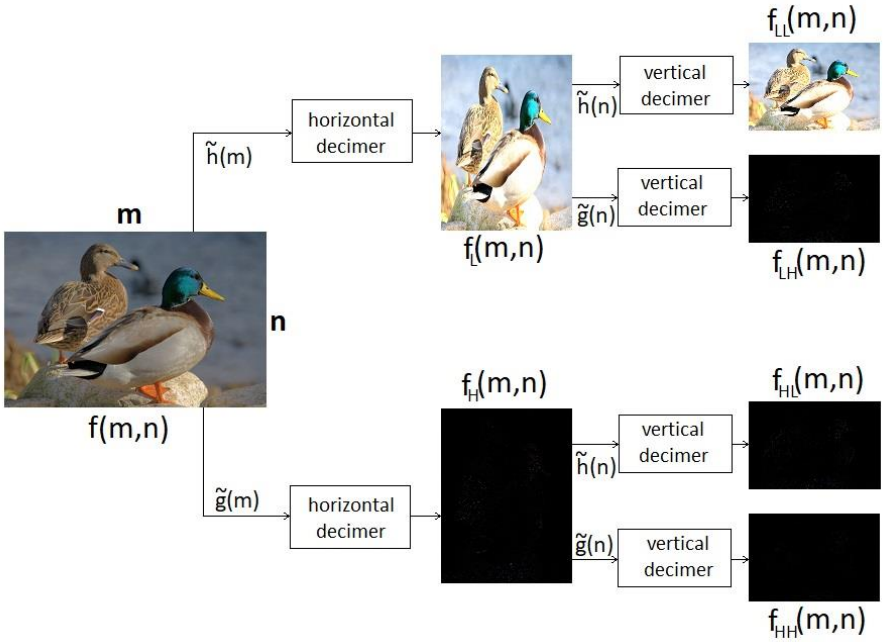


Fig. 2. Figure illustrating the idea of DWT operation on an image matrix

As a result, .JP2 is one of the less popular image formats. JPEG2000 is more flexible than JPEG because it allows both lossy and lossless compression. However, lossless compression is less efficient because it takes up more disk space. The individual compression steps of the JPEG2000 method are very similar to those of the JPEG method. The most significant difference is the use of DWT (Discrete Wavelet Transform) instead of DCT. JPEG2000 uses two types of wavelet transform. For lossless compression it is CDF (Cohen-Daubechies-Feauveau) 5⁄3 wavelet, for lossy it is CDF 9⁄7 wavelet. CDF filters built on the basis of the lifting scheme allow reversible integer transforms to be designed using biorthogonal wavelet coefficients. In this way, the error occurring in the image reconstruction is negligible and the results of the transformations are close to operations performed on a set of real numbers.

The idea of the DWT is well presented in Fig. 2. The original image is described by a binary function: as f(m,n), where m is responsible for the horizontal dimension, and n is for the vertical. Functions $\tilde{f}(m)$ and $\tilde{g}(m)$ are actually low-pass and high-pass filters. Thus, it is apparent that the image function is first filtered, and then followed by decimation, which is the halving of the horizontal dimension. Further on, both components of the image

are filtered again and then vertical decimation takes place. The resulting components are in a child-parent relationship with the output image. In the next step, this action is performed on the part containing the most information, that is, the upper left corner. Usually no more than 10 layers are created in this way.

In the images: Fig. 4, Fig. 3, Fig. 6 the effect of applying the discrete wavelet transform to the duck image can be observed. The color transformation was omitted in DWT implementation. The DWT was calculated based on RGB components and the image composed of individual blocks was saved to a new file. The image shows that most of the information about the image is in the top left block. The other blocks are almost black, containing trace amounts of information.

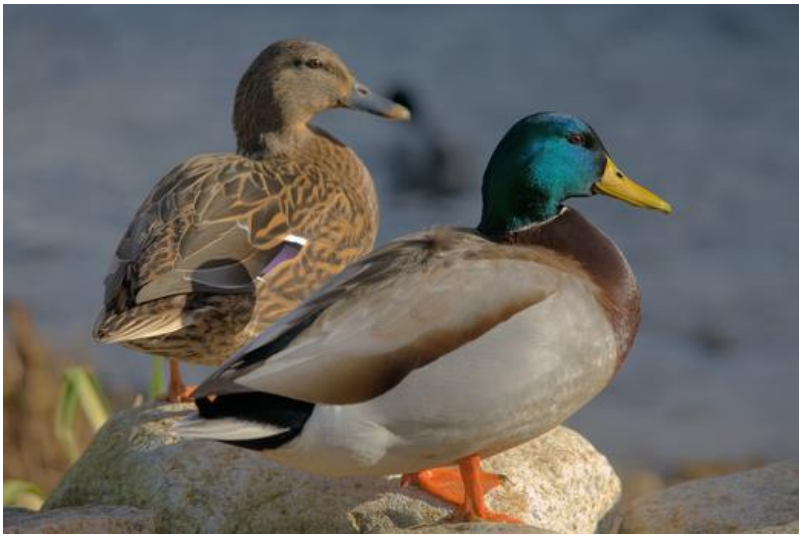


Fig. 4. Original image

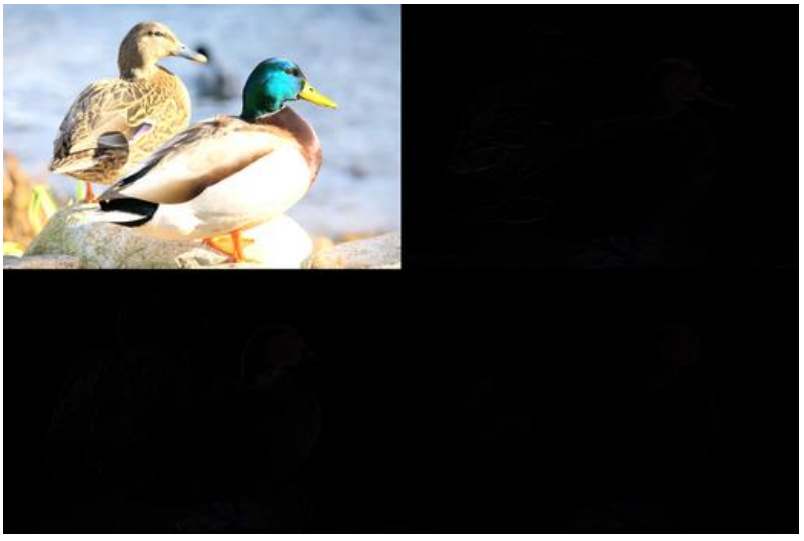


Fig. 5. Image after using first level DWT

Fig. 6. Image after using second level DWT

Tab. 1. Summary of compression parameters on example images using methods: JPEG, JPEG2000

R_{orig} [B]	Width	Height	R_{JPEG} [B]	k_{JPEG} [%]	R_{JP2} [B]	k_{JP2} [%]	R_{JP2L} [B]	k_{JP2L} [%]
15116598	2592	1944	940098	93.78	8066550	46.64	8605524	43.07
180054	300	200	16278	90.96	121873	32.31	125334	30.39
204534	320	213	25953	87.31	165910	18.88	167198	18.25
24860214	3840	2158	710875	97.14	5180921	79.16	7327314	70.53
720054	600	400	58383	91.89	471683	34.49	486422	32.45
921654	640	480	113024	87.74	733174	20.45	738477	19.87
819894	640	427	95444	88.36	628631	23.33	637605	22.23
1255254	697	600	57546	95.42	500026	60.17	581918	53.64
1279254	800	533	73059	94.29	502030	60.76	591067	53.80
1279254	800	533	147850	88.44	980861	23.33	994869	22.23
499554	500	333	36821	92.63	243678	51.22	275048	44.94
13517454	2600	1733	339097	97.49	3039933	77.51	4120015	69.52
156006	228	228	10778	93.09	75535	51.58	83638	46.39
23887926	2304	3456	329120	98.62	3556418	85.11	5854096	75.49
5992758	1632	1224	191748	96.80	1702784	71.59	2188998	63.47
4316454	1598	900	153640	96.44	1062431	75.39	1341009	68.93

In the literature e.g. (Anutam & Rajni, 2014), as well as on Adobe's (Bovik, 2009) website, one can find the information that lossy JPEG2000 allows compressing the image up to two times better than regular JPEG with the same compressed image parameters. This is not true, as shown in Tab. 1.

2.5. MSE

$$MSE = \frac{\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\sum_{k\in\{R,G,B\}}\left[V_O^k(i,j) - V_S^k(i,j)\right]^2}{3 \cdot M \cdot N} \quad (1)$$

MSE was calculated according to the formula (1), where M is width of image (number of pixels), N is height of image (number of pixels), V_O^k the value of the k-th color component of the pixel lying at the position with coordinates (i, j) in the original image, V_S^k the value of the k-th color component of the pixel lying at the position with coordinates (i, j) in the compressed image.

2.6. Compression methods tested and verification of the result

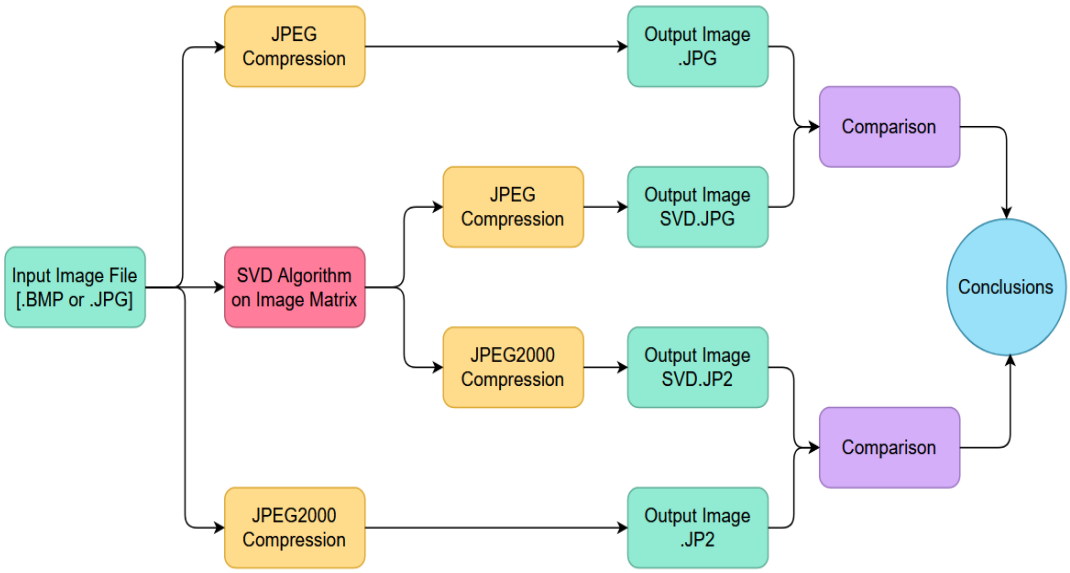


Fig. 7. Diagram showing the idea of proceeding with the verification of results

The Fig. 7 diagram illustrates the results verification procedure. Each bitmap has been compressed using JPEG and JPEG 2000 methods with and without SVD algorithm.

The comparison shown in the diagram refers to an image which was SVD processed prior to compression and one for which no SVD was used. The size of the JPEG-compressed image was compared with that of the previously SVD-applied image (same for JPEG2000 and SVD+JPEG2000).

The statistical analysis summary of the comparisons for the JPEG and SVD+JPEG methods is included in section 3.1, for JPEG2000 and SVD+JPEG2000 in section 3.2. Conclusions are provided in chapter 4.

2.7. Compression evaluation criteria

Tab. 2. Compression evaluation criteria

	JPEG	SVD+JPEG	JPEG2000	SVD+JPEG 2000
Lossless compression	No	No	Yes	No
Lossy compression	Yes	Yes	Yes	Yes
Flexibility	Low	High	High	Very high
Pessimistic time complexity	$O(n^2)$	$O(n^3)$	$O(n^2)$	$O(n^3)$
Expected time complexity	O(n log(n))	$O(n^3)$	O(n log(n))	$O(n^3)$
Memory complexity	O(n log(n))	$O(n^2)$	O(n log(n))	$O(n^2)$
Ease of implementation(Python)	Easy	Medium	Easy	Medium

It is not necessary to run the program to determine the above criteria. They result directly from the properties of the individual algorithms. Detailed analyses have been carried out to measure file size, compression ratio or to determine the MSE. The results will be presented in the subsequent section.

2.8. Hardware & Software

Tab. 3. Parameters of the equipment used to conduct the tests

Processor	AMD Ryzen 7 4800H with Radeon Graphics
Clock signal	2.90 GHz
Number of cores	8
Logic processors	16
Installed RAM	32.0 GB
GPU 1	NVIDIA GeForce GTX 1650
GPU 2	AMD Radeon (TM) Graphics
Operating system type	64-bit operating system, processor x64
Version of Operating System	Windows 11 Home 21H2

The important factor affecting the time of the calculations is the hardware parameters on which they are performed. In this case it was a Lenovo Legion 5 15ARH05 laptop with the parameters as shown in Tab. 3.

The second important factor is the version of the programming language used to create the image compression application: Python version 3.10.

2.9. Statistical methods used to test the hypotheses

To test the hypothesis whether the use of the SVD algorithm significantly affects the efficiency of the JPEG method, a group of 2561 different images were compressed using the methods: JPEG and SVD+JPEG. For each image, the size of the file after compression was measured for both methods. Two equinumerous series of data were obtained, which

were naturally combined into pairs. Then, for each of them, the one-sided Wilcoxon test for pairs of observations with alternative hypothesis "less" was carried out at a fixed significance level of $\alpha = 0.05$. The demonstration of the impact of the SVD algorithm on the efficiency of the JPEG2000 method was carried out in a similar way. The tests were performed in the RStudio environment (R version 4.2.0 – Vigorous Calisthenics).

In order to check whether the fraction of eigenvalues required to reconstruct the image in good quality decreases as the image dimension increases, the image matrix was decomposed using the SVD algorithm. The image was then reconstructed a hundred times using successively 1%, 2%, ..., 100% of the found eigenvalues of the image matrix. The desired level of image quality was determined based on the MSE. For each compressed image, the MSE value was calculated for each fraction of the eigenvalues. The image with the MSE parameter closest to the determined MSE <= 0.5 was selected. In this way, coordinates of the points were obtained: fraction of specific eigenvalues used for reconstruction, minimum of width and height of the image, which were plotted on a graph. This study was performed on a group of 50 randomly selected images in .BMP format. The graph was generated in a Python program using matplotlib library.

3. RESULTS

The following images were created by compressing the original image using different eigenvalue fractions. The fraction of eigenvalues used to reconstruct them and their sizes after compression are shown above the individual compressed images.

Fig. 8 presents result of compression small image, 100×67 pixels. Only a small part of the results of the tests performed are included in Tab. 4.

Fig. 9 shows the result of compression performed on a medium image. Fig. 9 presented MSE depending on the value of the image's compression ratio.

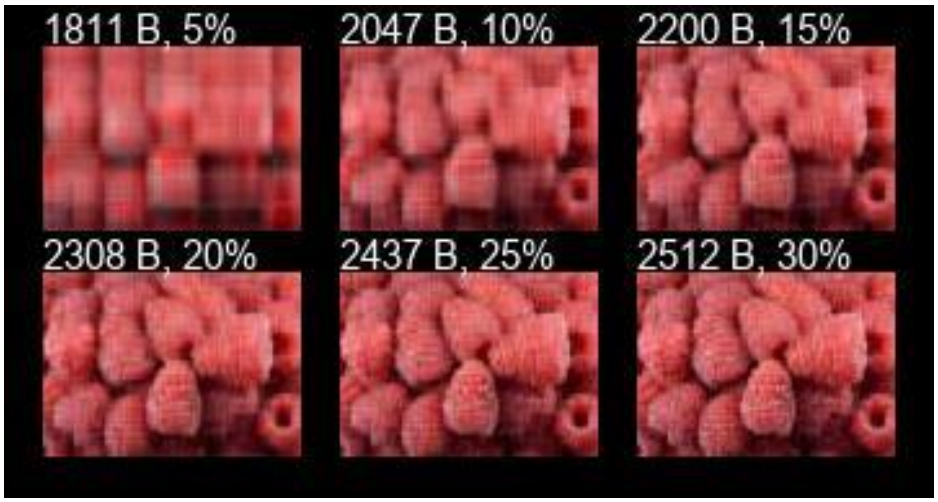


Fig. 8. Image compression using eigenvalue fractions: 5–30%. Information about the original image: dimension: 100×67, file size .BMP: 20 154 B

Tab. 4. Example results of compressing a .BMP image into .JPG format using SVD and different eigenvalue fractions

k	n	Time of compression [µs]	Size of compressed file [B]	Avg number of bad pixels	Max number of bad pixels	MSE	Compression [%]
0.01	1	39126.7000	1498	72.1179	300.0582	916.2527	92.57
0.05	3	31761.8000	1811	54.0602	242.9183	533.8225	91.01
0.10	7	28343.7000	2047	36.3769	215.8520	243.0560	89.84
0.15	10	30138.7000	2200	29.3871	177.6175	158.9510	89.08
0.20	13	28359.3000	2308	24.2515	152.8387	107.5304	88.55
0.25	17	26359.9000	2437	19.4205	120.2190	69.0755	87.91
0.30	20	26305.8000	2512	16.8861	100.7209	52.1632	87.54
0.35	23	27013.5000	2561	14.7831	78.0716	39.5877	87.29
0.40	27	26235.9000	2625	12.2134	64.1442	27.0916	86.98
0.50	34	26323.1000	2721	8.5490	46.4939	13.2923	86.50
0.60	40	33556.9000	2773	5.8081	35.4037	6.2488	86.24
0.70	47	27329.6000	2804	3.4723	23.3793	2.2468	86.09
0.80	54	32545.1000	2810	1.6939	11.7402	0.5417	86.06
0.90	60	27339.0000	2806	0.6080	4.1410	0.0787	86.08
0.95	64	33509.5000	2811	0.1766	1.7116	0.0092	86.05
1.00	67	33555.9000	2809	0.0000	0.0000	0.0000	86.06

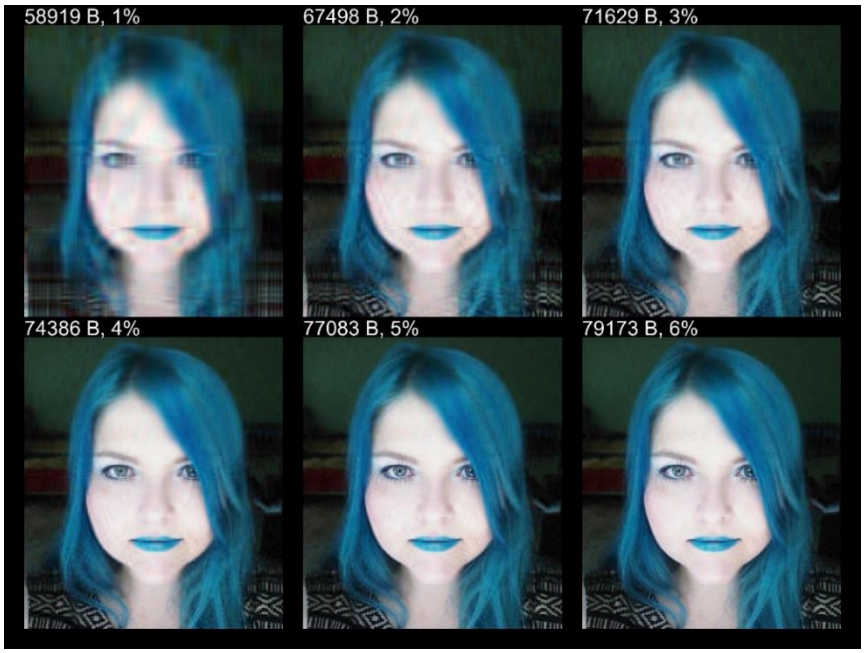


Fig. 9. Image compression using eigenvalue fractions: 1-6%. Information about the original image: dimension: 960×1081, file size .JPG: 243 728 B

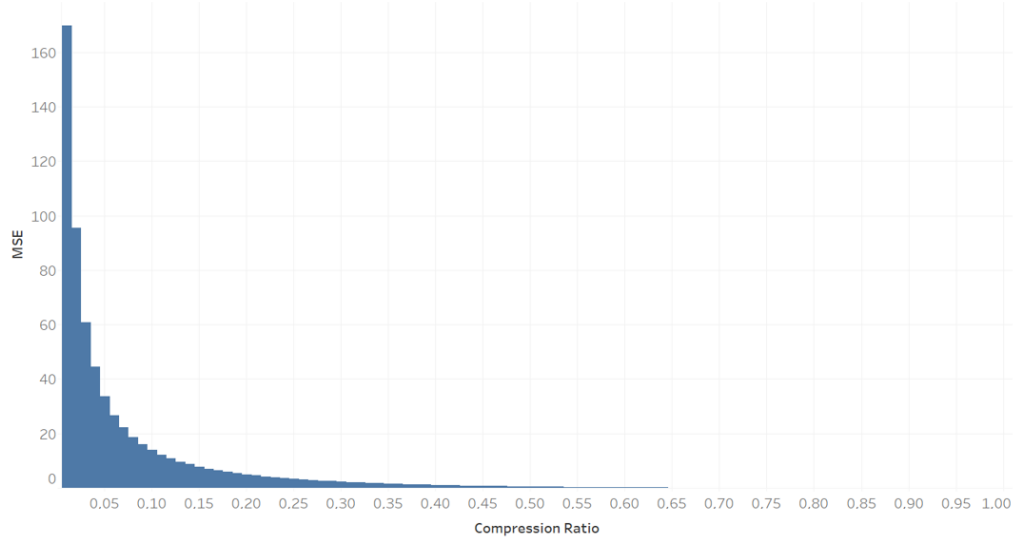


Fig. 10. MSE depending on the value of the compression ratio using SVD+JPEG for image showed on figure 8

3.1. Applying the SVD algorithm on the image matrix increases the efficiency of its compression by the JPEG method

Null hypothesis: The size of the SVD+JPEG compressed file is larger than the size of the same JPEG compressed file.

Alternative hypothesis: The size of the file compressed by SVD+JPEG method is smaller than the size of the same file compressed by JPEG method.

The result of Wilcoxon signed rank test with continuity correction:

data: size$SVDJPG and size$JPG, V = 111090, p-value < 2.2e-16.

The p-value is less than the set significance level, so the null hypothesis should be rejected in favour of the alternative hypothesis: The size of the file compressed with the SVD+JPEG method is smaller than the size of the same file compressed with the JPEG method.

3.2. The application of the SVD algorithm on the image matrix increases the efficiency of its compression by the JPEG2000 method

Null hypothesis: The file size compressed by the SVD+JPEG2000 method is larger than the size of the same file compressed by the JPEG2000 method.

Alternative hypothesis: The size of the file compressed by the SVD+JPEG2000 method is smaller than the size of the same file compressed by the JPEG2000 method.

The result of Wilcoxon signed rank test with continuity correction:

data: size$SVDJPG2000 and size$JPG2000, V = 81092, p-value < 2.2e-16.

The p-value is less than the established significance level, so the null hypothesis should be rejected in favour of the alternative hypothesis: The size of the file compressed by SVD+JPEG2000 is smaller than the size of the same file compressed by JPEG2000.

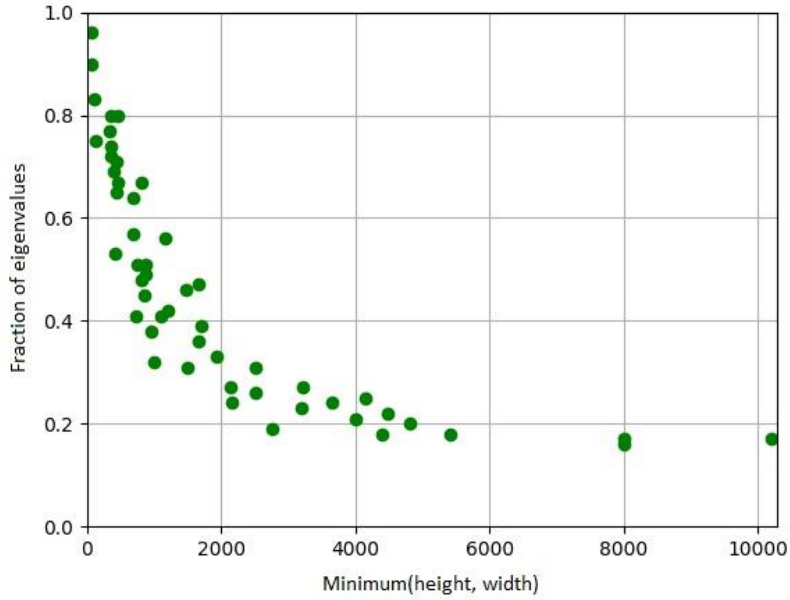


Fig. 11. Fraction of eigenvalues to be used for image reconstruction based on minimum of height and width, with fixed MSE <= 0.5

As the image dimension increases, the fraction of eigenvalues to be used for good quality image reconstruction decreases.

4. CONCLUSIONS

In conclusion, this paper analyses the possibility of using SVD matrix decomposition in image compression and discusses the performance, strengths and weaknesses of the JPEG and JPEG2000 algorithms. The analysis performed confirmed the accuracy of the hypotheses. It can therefore be concluded that the use of the SVD algorithm before compression with the JPEG and JPEG2000 methods significantly reduces the size of the output file. As presented on Fig. 8, Fig. 9 and Fig. 11, the fraction of eigenvalues to be used for image reconstruction, based on image size, decreases with the increase of dimension of image. Minimum of height and width was taken to show that relation, because it determines total number of eigenvalues, which is the same as total number of singular values of image matrix. Summarizing, JPEG is based on DCT, while JPEG2000 on DWT. In the authors opinion, the full potential of using the SVD algorithm in image compression would only be explored when creating a new standard for image file compression, in which SVD decomposition of image matrices would be one of the compression steps. In order to investigate the full power of SVD it would be necessary to write a custom image compressor.

REFERENCES

Anutam, & Rajni. (2014). Comparative Analysis of Filters and Wavelet Based Thresholding Methods for Image Denoising. *Computer Science & Amp; Information Technology (CS &Amp; IT)* (pp. 137–148). https://doi.org/10.5121/csit.2014.4515

Arps, R., & Truong, T. (1994). Comparison of international standards for lossless still image compression. *Proceedings of the IEEE*, *82*(6), 889–899. https://doi.org/10.1109/5.286193

Bovik, A. C. (2009). *The Essential Guide to Image Processing* (1st ed.). Academic Press.

Britanak, V., Yip, P. C., & Rao, K. (2007). CHAPTER 4 – Fast DCT/DST Algorithms. *Discrete Cosine and Sine Transforms. General Properties, Fast Algorithms and Integer Approximations* (pp. 73–140). Academic Press. https://doi.org/10.1016/b978-012373624-6/50006-0

Cao, L. (2006). *SVD applied to digital image processing*. Division of Computing Studies, Arizona State University Polytechnic Campus.

Chen, Y., Mukherjee, D., Han, J., Grange, A., Xu, Y., Parker, S., Chen, C., Su, H., Joshi, U., Chiang, C. H., Wang, Y., Wilkins, P., Bankoski, J., Trudeau, L., Egge, N., Valin, J. M., Davies, T., Midtskogen, S., Norkin, A., de Rivaz, P., Design, A., & Liu, Z. (2020). An Overview of Coding Tools in AV1: the First Video Codec from the Alliance for Open Media. *APSIPA Transactions on Signal and Information Processing*, *9*(1), e6. https://doi.org/10.1017/atsip.2020.2

Compton, E. A., & Ernstberger, S. L. (2020). Singular Value Decomposition: Applications to Image Processing. Lagrange College. *Journal of Undergraduate Research*, *17*, 99–105.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2005). *Wprowadzenie do algorytmów*. Wydawnictwo Naukowe PWN.

Davies, E. R. (2017). *Computer Vision: Principles, Algorithms, Applications, Learning*. Elsevier Gezondheidszorg.

Dhawan, S. (2011). *A Review of Image Compression and Comparison of its Algorithms*. https://www.semanticscholar.org/paper/A-Review-of-Image-Compression-and-Comparison-of-its-Dhawan/034dcf50d99bbd9870c5c2e67201f6d792f96a5f

Dumka, A., Ashok, A., Verma, P., & Verma, P. (2020). *Advanced Digital Image Processing and Its Applications in Big Data* (1st ed.). CRC Press.

Gandhi, T., Patel, H., & Prajapati, D. (2015). *Image Compression Using Fractal: Image compression based upon the self-similarities present in the image*. LAP LAMBERT Academic Publishing.

Gong, L., Deng, C., Pan, S., & Zhou, N. (2018). Image compression-encryption algorithms by combining hyper-chaotic system with discrete fractional random transform. *Optics &Amp; Laser Technology*, *103*, 48–58. https://doi.org/10.1016/j.optlastec.2018.01.007

Hoffman, R. (1997). *Data Compression in Digital Systems*. Springer Publishing.

Hoffman, R. (2012). *Data Compression in Digital Systems*. Springer Publishing.

Jackson, D., & Hannah, S. (1993). Comparative analysis of image compression techniques. *1993 (25th) Southeastern Symposium on System Theory* (pp. 513–517). IEEE. https://doi.org/10.1109/ssst.1993.522833

Jankowska, J., & Jankowski, M. (1988). *Przegląd metod i algorytmów numerycznych*. Wydawnictwa Naukowo-Techniczne.

Jinchuang, Z., Yan, T., & Wenli, F. (2009). Research of image compression based on Wireless visual sensor networks. *4th International Conference on Computer Science & Education* (pp. 353–356). IEEE. https://doi.org/10.1109/iccse.2009.5228430

Karwowski, D. (2019). *Zrozumieć kompresję obrazu: podstawy technik kodowania stratnego oraz bezstratnego obrazów*. Damian Karwowski.

Kostrikin, A. I. (2004). *Wstęp do algebry cz. 1 i cz. 2 Podstawy algebry*. Wydawnictwo Naukowe PWN.

Lu, Z., & Guo, S. (2016). *Lossless Information Hiding in Images* (1st ed.). Syngress.

Mammeri, A., Hadjou, B., & Khoumsi, A. (2012). A Survey of Image Compression Algorithms for Visual Sensor Networks. *International Scholarly Research Notices*, *2012*, 760320. https://doi.org/10.5402/2012/760320

Miano, J. (1999). *Compressed Image File Formats: JPEG, PNG, GIF, XBM, BMP*. Addison-Wesley Professional.

Murray, J. D., & VanRyper, W. (1996). *Encyclopedia of Graphics File Formats*. O'Reilly & Associates.

Nasri, M., Helali, A., Sghaier, H., & Maaref, H. (2010). Energy-efficient wavelet image compression in Wireless Sensor Network. *2010 International Conference on Wireless and Ubiquitous Systems* (pp. 1–7). IEEE. https://doi.org/10.1109/icwus.2010.5670430

Nixon, M., & Aguado, A. (2019). *Feature Extraction and Image Processing for Computer Vision* (4th ed.). Academic Press.

Parekh, D. (2021, April 25). *Image Compression Standards | Digital Image Processing* [Video file]. YouTube. https://www.youtube.com/watch?v=6IuKH7IGspU

Pratt, W., Kane, J., & Andrews, H. (1969). Hadamard transform image coding. *Proceedings of the IEEE*, *57*(1), 58–68. https://doi.org/10.1109/proc.1969.6869

Pu, I. M. (2005). *Fundamental Data Compression* (1st ed.). Butterworth-Heinemann.

Salomon, D., Motta, G., & Bryant, D. (2007). *Data Compression: The Complete Reference*. Springer Publishing.

Sayood, K. (2002). *Lossless Compression Handbook*. Elsevier Gezondheidszorg.

Shih, C. W., Chu, H. C., Chen, Y. M., & Wen, C. C. (2012). The effectiveness of image features based on fractal image coding for image annotation. *Expert Systems With Applications*, *39*(17), 12897–12904. https://doi.org/10.1016/j.eswa.2012.05.003

Short, M. N., Manohar, M., & Tilton, J. C. (1994). Planning/Scheduling Techniques for VQ-Based Image Compression. Science Information Management and Data Compression Workshop. *1994 Science Information Management and Data Compression Workshop* (pp. 95–104). US Government.

Shukla, K. K., & Prasad, M. V. (2011). *Lossy Image Compression: Domain Decomposition-Based Algorithm*s. Springer Publishing.

Stewart, G. W. (2001). *Matrix Algorithms: Volume 2, Eigensystems* (1st ed.). SIAM: Society for Industrial and Applied Mathematics.

Swathi, H. R., Sohini, S., Surbhi, & Gopichand, G. (2017). Image compression using singular value decomposition. *IOP Conference Series: Materials Science and Engineering*, *263*, 042082. https://doi.org/10.1088/1757-899x/263/4/042082

Tanwar, S., Ramani, T., & Tyagi, S. (2018). Dimensionality Reduction Using PCA and SVD in Big Data: A Comparative Case Study. *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering* (pp. 116–125). Springer. https://doi.org/10.1007/978-3-319-73712-6_12

Wayner, P. (1999). *Compression Algorithms for Real Programmers*. Elsevier Gezondheidszorg.

What's the difference between 'visually lossless' and real lossless and what does this mean for future encodes? (2019, May 18). *Video Production Stack Exchange*. Retrieved May 2022 from https://video.stackexchange.com/questions/27656/whats-the-difference-between-visually-lossless-and-real-lossless-and-what-doe

Xiao, F., Zhang, P., Sun, L. J., Wang, J., & Wang, R. C. (2011). Research on image compression and transmission mechanism for wireless multimedia sensor networks. *2011 International Conference on Electrical and Control Engineering* (pp. 788–791). IEEE. https://doi.org/10.1109/iceceng.2011.6057601

Applied Computer Science, vol. 18, no. 4, pp. 68–81
doi: 10.35784/acs-2022-29

Submitted: 2022-10-01 | Revised: 2022-11-20 | Accepted: 2022-12-03

Keywords: ANN, the compressive strength, RCA, MLP, RBF

Monika KULISZ [0000-0002-8111-2316]*, *Justyna KUJAWSKA* [0000-0002-4809-2472]**,
Zulfiya AUBAKIROVA [0000-0003-0322-7692]***,
Gulnaz ZHAIRBAEVA [0000-0002-5102-397X]***,
Tomasz WAROWNY [0000-0002-1500-9812]****

PREDICTION OF THE COMPRESSIVE STRENGTH OF ENVIRONMENTALLY FRIENDLY CONCRETE USING ARTIFICIAL NEURAL NETWORK

Abstract

The paper evaluated the possibility of using artificial neural network models for predicting the compressive strength (Fc) of concretes with the addition of recycled concrete aggregate (RCA). The artificial neural network (ANN) approaches were used for three variable processes modeling (cement content in the range of 250 to 400 kg/m^3, percentage of recycled concrete aggregate from 25% to 100% and the ratios of water contents 0.45 to 0.6). The results indicate that the compressive strength of recycled concrete at 3, 7 and 28 days is strongly influenced by the cement content, %RCA and the ratios of water contents. It is found that the compressive strength at 3, 7 and 28 days decreases when increasing RCA from 25% to 100%. The obtained MLP and RBF networks are characterized by satisfactory capacity for prediction of the compressive strength of concretes with recycled concrete aggregate (RCA) addition. The results in statistical terms; correlation coefficient (R) reveals that the both ANN approaches are powerful tools for the prediction of the compressive strength.

1. INTRODUCTION

Machine learning methods have been constantly developing in recent times. One of the methods of machine learning are artificial neural networks (ANN) that are used in various areas of life and science (Machrowska et al., 2020a, 2020b; Karpiński, 2022; Szabelski, Karpiński & Machrowska, 2022; Rymarczyk et al., 2021; Szala et al., 2021; Pytka et al., 2022; Rymarczyk et al., 2019). ANN is one of the important artificial intelligence technique

* Lublin University of Technology, Faculty of Management, Department of Organisation of Enterprise, Poland, m.kulisz@pollub.pl
** Lublin University of Technology, Faculty of Environmental Engineering, Department of Biomass and Waste Conversion into Biofuels, Poland, j.kujawska@pollub.pl
*** D. Serikbayev East Kazakhstan Technical University, School of Architecture and Construction, Kazakhstan, Aubakirova.zulfiya@mail.ru, gulnaz.w@mail.ru
**** Lublin University of Technology, Faculty of Management, Department of Quantitative Methods in Management, Poland, t.warowny@pollub.pl

inspired on the study of biological neural networks that can be applied to studies in the area of construction where there is a database of a problem and the ANN model learns by example (Dantas, Leite & Nagahama, 2013; Parichatprecha & Nimityongskul, 2009). It has been proven to be a powerful modeling technique for complex and nonlinear problems with strong proposed to learning and function approximation (Dahmoune et al., 2015; Hammoudi et al., 2019). At present, the topic of the use of neural networks in various fields of human activity is extremely popular. Many methods and methodologies based on the application of ANN are used in the construction industry for optimization, control, problems of identification and forecasting (Das, Swetapadma & Panigrahi, 2019; Orosa et al., 2019; Pezeshki & Mazinani, 2019; Fei, Youfu & Xuejun, 2019).

In recent years, researchers have begun using artificial neural networks to determine the properties of building materials, including: predicting performance of lightweight concrete with granulated expanded glass and ash aggregate (Kurpinska & Kułak, 2019), designing the composition of cement stabilized rammed earth (Anysz & Narloch, 2019), studying adiabatic temperature rise reflecting hydration degree of concrete (Han et al., 2018), and predicting the compressive strength of cement-based materials exposed to sulfate attack (Chen et al., 2018).

The work presented here evaluates the feasibility of using a neural network model to predict the performance of recycled aggregates (RCAs). ANN techniques are rarely used to predict the performance of RCA and concretes in general due to their complex composition. Topçu and Seridemir (2008) predicted the compressive and splitting tensile strength of recycled aggregate concentrate containing silica fume have been developed at the age of 3, 7, 14, 28, 56 and 96 days. The values are closer to the experimental results obtained from training and testing for in artificial neural networks (Topçu & Saridemir, 2008). Chopra et al. employed an ANN model to predict the compressive strength of concentrate. It was deduced that the best training algorithm is 'Levenberg-Marquardt' algorithm that attains more than 95% on average prediction accuracy (Chopra, Kumar & Kumar, 2015). Atici (2011) applied multiple regression analysis and an artificial neural network in estimating the compressive strength of concrete that contains various amounts of blast furnace slag and fly ash. He showed that the application of an artificial neural network to the prediction of the compressive strength in admixture concrete of various curing times shows great potential in terms of inverse problems, and it is suitable for calculating nonlinear functional relationships, for which classical methods cannot be applied (Atici, 2011).

This study aimed at predicting and modeling the compressive strength of a concrete containing recycled concrete aggregates following 3, 7 and 28 days for different ranges of cement content, percentage of recycled concrete aggregate and the ratios of water contents.

2. MATERIALS AND METHODS

Type I ordinary portland cement was used as a binder content for the experiment. The chemical compositions were illustrated in Table 1. Sand (NS), i.e. crushed limestone with nominal size of 4 mm was used as well; the sand was dried at 105°C. Crushed granite (NSA) was used as concrete aggregate with specific gravity 2.7 and nominal size 19 mm in normal concrete and the recycled concentrate aggregate (RCA) with attached mortar, nominal size 19 mm was used as a replacement of the concrete aggregate. The RCA was obtained from

the demolition of an old building and had undergone a crushing process to obtain the required nominal size. The grading of both types of concrete aggregate complied with the grading limits for the crushed-rock aggregate in BS 882:1992 (Gjorv & Sakai, 2014).

2.1. Experimental

The mix design of the concrete was done according to the DoE method, which was targeted at compressive strength of 25 MPa at the 28th day (British Standards Institution, 1988).

Tab. 1. Composition of CEM II B-V 32,5R (CEM II/B-V 32,5 R, n.d.)

Properties	Unit	CEM II B-V 32,5R
Specific surface	($cm^2 \cdot g^{-1}$)	4237
Initial setting time	(min)	243
Compressive strength		
after 2 days	(MPa)	20.3
after 28 days	(MPa)	45.7
Density	($g \cdot cm^{-3}$)	2.83
SO_3 content	(%)	2.28
Chloride ion content	(%)	0.06
Na_2O content	(%)	1.09

The mixture compositions of all mixes are presented in Table 2. Notably, there are five types of mixtures prepared by replacing the concrete aggregate with the RCA at 25%, 50%, 75% and 100% of the total concrete aggregate content. The percentage of replacement was calculated based on the total weight of the concrete aggregate content.

Tab. 2. Mixture proportion for 1 m^3 of concentrate [kg/m^3]

Mixture	% Repl. of recycled aggregates	Cement [kg/m^3]	Water [kg/m^3]	NSA	RCA	Sand
Mix 1	25	250	150	858	286	762
Mix 2	50	250	150	564	564	762
Mix 3	75	250	150	279	836	743
Mix 4	100	250	150	0	110	734
Mix 5	25	400	180	770	257	685
Mix 6	50	400	180	507	507	676
Mix 7	75	400	180	250	751	667
Mix 8	100	400	180	0	988	659
Mix 9	25	350	230	105	105	661
Mix 10	50	350	230	177	177	661
Mix 11	75	350	230	370	70	661
Mix 12	100	350	230	360	320	625
Mix 13	25	350	150	105	105	661
Mix 14	50	350	150	177	177	661
Mix 15	75	350	150	370	70	661
Mix 16	100	350	150	360	320	625

All specimens were cast under laboratory condition and demolded at 24 ± 2 hours after mixing; afterwards, they were fully submerged in water at a temperature of 25 ± 2°C until the age of testing. The testing program introduced the determination of the compressive strength and ultra-sonic pulse velocity test, while the durability was tested through the shrinkage and expansion test, the ratios of water contents and gas permeability test. Testing was carried out in accordance to the British Standard testing procedures.

2.2. Compressive strength test

The compressive strength test was performed according to BS EN 12390-3:2009 using three cubes with the dimensions of 100 mm × 100 mm × 100 mm to obtain an average value (British Standards Institution, 2009). This test was carried out on the specimens at the age of 3, 7 and 28 days.

2.3. Neural network simulation

Modeling was performed using artificial neural networks, via Statistica Neural Networks software. The input neurons were cement content, %RCA and the ratios of water contents, and the output neuron was Fc after 3, 7 and 28 days. In connection with modeling Fc at three time points, three types of models were analyzed, the diagram of which is shown in Figure 1.

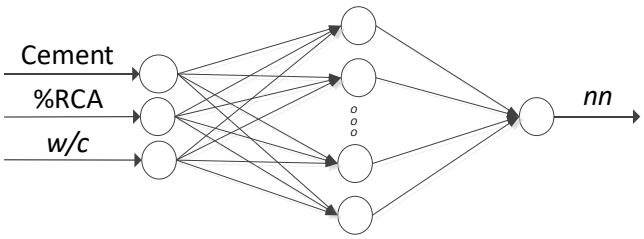


Fig. 1 Schematics of the artificial neural network, where *nn* – Fc after 3, 7 and 28 days

Two types of neural networks were used for modeling: MLP and RBF. The multi-layer perceptron (MLP) is one of the most popular. Characterized by a layered arrangement of neurons and a unidirectional flow of data (from input to output) without feedback. The training of MLP-type networks is possible by using the backward error propagation method. Radial basis function (RBF) networks are a special type of artificial neural networks. They are unidirectional three-layer networks consisting of an input layer, a hidden layer and an output layer. In the hidden layer, there are radial basis functions that correspond to hidden neurons (Karpiński et al., 2022a, 2022b).

In the case of MLP networks, the learning algorithm – BFGS gradient was used, and different activation functions were tested, including: linear, exponential, logistic, tanh and sinus. For RBF networks, the learning algorithm is RBFT, and the activation functions are: Gaussian distribution (hidden neurons) and linear function (output neuron). Networks with one hidden layer were modeled, with a change in the number of neurons in the hidden layer (2–10). The input data set was divided into 75%–25% (learning data – validation data). Due to the small data set, test data was omitted.

The selection of networks was based on indicators such as learning and validation quality as well as learning and validation errors. For each Fc model, 200 networks were learned after 3, 7 and 28 days, from which one of each type was selected.

Learning and validation quality is defined as the correlation coefficient for these sets, calculated according to equation (1):

$$R(y', y^*) = \frac{cov(y', y^*)}{\sigma_{y'}\sigma_{y^*}} \qquad R \in <0,1> \tag{1}$$

where: $\sigma_{y'}$ – standard deviation of reference values,
σ_{y^*} – standard deviation of predicted values,
$cov(y', y^*)$ – covariance.

The errors are defined as the sum of the squared differences between the set values and the values obtained at the outputs of each output neuron, according to the formula (2):

$$Err = \sum_{i=1}^{n}(y_i' - y_i^*)^2 \tag{2}$$

where: n – number of cases in a given set,
y_i' – actual value of Fc for the given set for the *i*-th observation;
y_i^* – predicted value of Fc for the given set for the *i*-th observation.

3. RESULTS AND DISCUSION

3.1. Compressive strength

Table 3 shows the obtained results of the tested mixtures. The results show that the compressive strength of recycled concrete after 3, 7 and 28 days significantly changes under the influence of the cement content and the addition of the recycled concentrate aggregate. Among all the samples, the concrete with 25% RCA addition achieves the highest strength, followed by 50%, 75%, 100% RCA addition.

Tab. 3. Experimental values for compressive strengths of 3, 7 and 28 days for the tested mixtures

Mixtures	Cement [kg/m^3]	%RCA	w/c	Fc [MPa]		
				3	7	28
Mix 1	250	25	0.65	21.1	25.9	27.1
Mix 2	250	50	0.65	18.9	22.1	22.7
Mix 3	250	75	0.65	20.2	22.2	22.9
Mix 4	250	100	0.65	16.8	25.1	26.2
Mix 5	400	25	0.45	32.1	37.7	42.4
Mix 6	400	50	0.45	30.2	36.3	36.3
Mix 7	400	75	0.45	27.4	35.2	36.0
Mix 8	400	100	0.45	21.5	34.6	34.7
Mix 9	350	25	0.65	34.1	36.5	37.0
Mix 10	350	50	0.65	23.8	27.2	29.1
Mix 11	350	75	0.65	19.8	24.1	26.0
Mix 12	350	100	0.65	18.2	22.3	29.2
Mix 13	350	25	0.45	28.8	34.2	37.8
Mix 14	350	50	0.45	27.0	32.4	33.3
Mix 15	350	75	0.45	24.3	31.5	32.4
Mix 16	350	100	0.45	19.4	31.2	31.3

4.2. Modeling results

The results of the obtained modeling with the parameters of the obtained networks are shown in Table 4. The best parameters for MLP networks for Fc modeling after 3 days were obtained for a network with six neurons in the hidden layer, after 7 days and after 28 days for 4 neurons in the hidden layer. In the case of the RBF networks for Fc modeling, after 3 days the best results were obtained for a network with seven neurons in the hidden layer, after 7 days for six neurons, and after 28 days for 7 neurons in the hidden layer. The quality of both learning and validation for all networks exceeds 0.97. In addition, Table 4 shows the R-correlation coefficients (for the entire dataset) between the test data and the modeling data. By analyzing the R-correlation, it can be concluded that the cross-correlation between the experimental data and the data predicted for the networks of both networks is at a very high level (above 0.97).

Tab. 4. Network parameters obtained as a result Fc after 3, 7 and 28 days of modeling.

Modeled Fc	3 days		7 days		28 days	
Network Name	MLP 3-6-1	RBF 3-7-1	MLP 3-4-1	RBF 3-6-1	MLP 3-4-1	RBF 3-7-1
Quality (Training)	0.9907	0.9714	0.9932	0.9951	0.9725	0.9881
Quality (Validation)	0.9941	0.9961	0.9988	0.9989	0.9952	0.9943
Err (Training)	0.2496	0.7526	0.1106	0.1363	0.7565	0.2946
Err (Validation)	0.7861	0.1276	0.0244	0.0422	0.3292	0.3004
Learning algorithm	BFGS 168	RBFT	BFGS 3339	RBFT	BFGS 97	RBFT
Activation (hidden)	Sinus	Gaussian	Logistic	Gaussian	Gaussian	Gaussian
Activation (output)	Exponential	Linear	Exponential	Linear	Sinus	Linear
R(i) correlation	0.9876	0.9773	0.9962	0.9908	0.9903	0.9813

For a more detailed comparison of the results of modeling RBF and MLP networks and real Fc data after 3, 7 and 28 days, the following figures show correlation plots of these relationships – for Fc after 3 days (Figure 2a), Fc after 7 days (Figure 2b), Fc after 28 days (Figure 2c).

Analysis of the following graphs confirms that for both types of RBF and MLP networks, the quality of these models is at an acceptable level. Therefore, it can be concluded that artificial neural networks are a suitable tool for predicting the Fc after 3, 7 and 28 days.

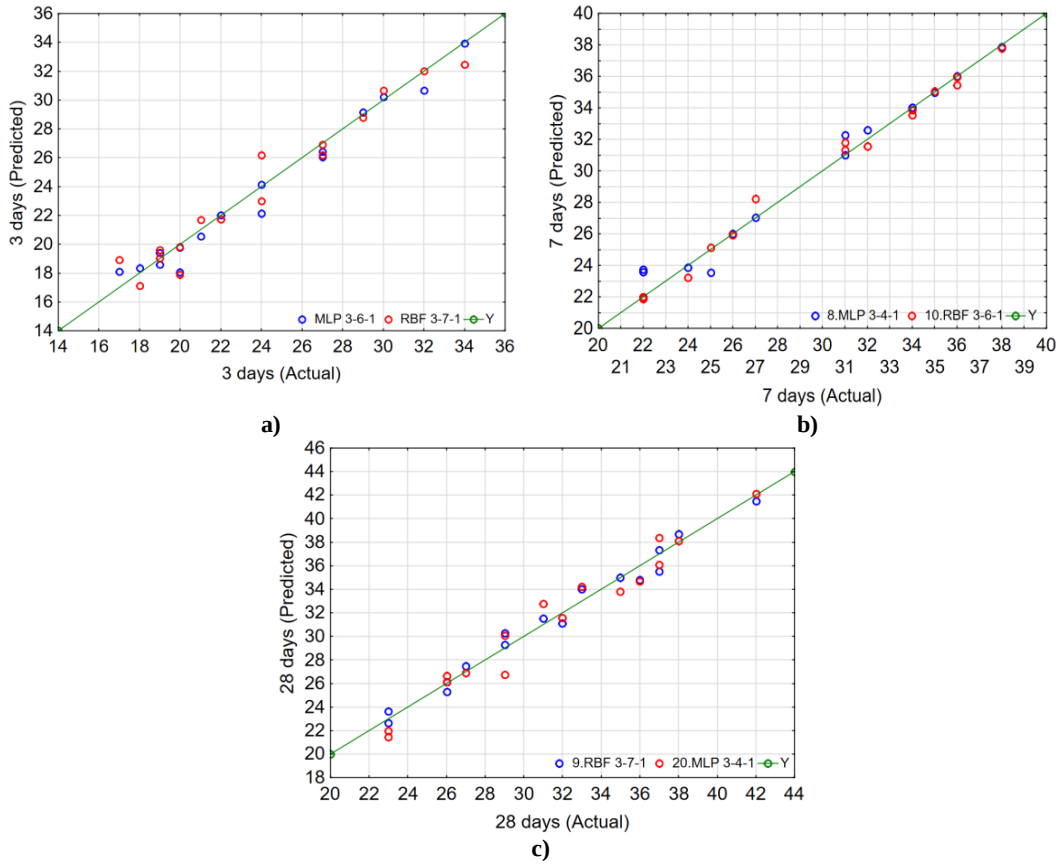


Fig. 2. Correlation graph of comparison between the modeling and actual results of the Fc after 3, 7 and 28 days for MLP and RBF networks

As a result of the modeling, it was possible to predict Fc after 3, 7 and 28 days, using the trained networks by entering the input data into Statistica. The results of the networks are shown for the following figures, for Fc after 3 days depending on %RCA and cement content for MLP network (Fig. 3a) and RBF network (Fig. 3b) and depending on w/c and cement content for MLP network (Fig. 3c) and RBF network (Fig. 3d), for Fc after 7 days depending on %RCA and cement content for MLP network (Fig. 4a) and RBF network (Fig. 4b), and w/c and cement content for MLP network (Fig. 4c) and RBF network (Fig. 4d), as well as for Fc after 28 days depending on %RCA and cement content for MLP network (Fig. 5a) and RBF network (Fig. 5b), and w/c and cement content for MLP network (Fig. 5c) and RBF network (Fig. 5d).

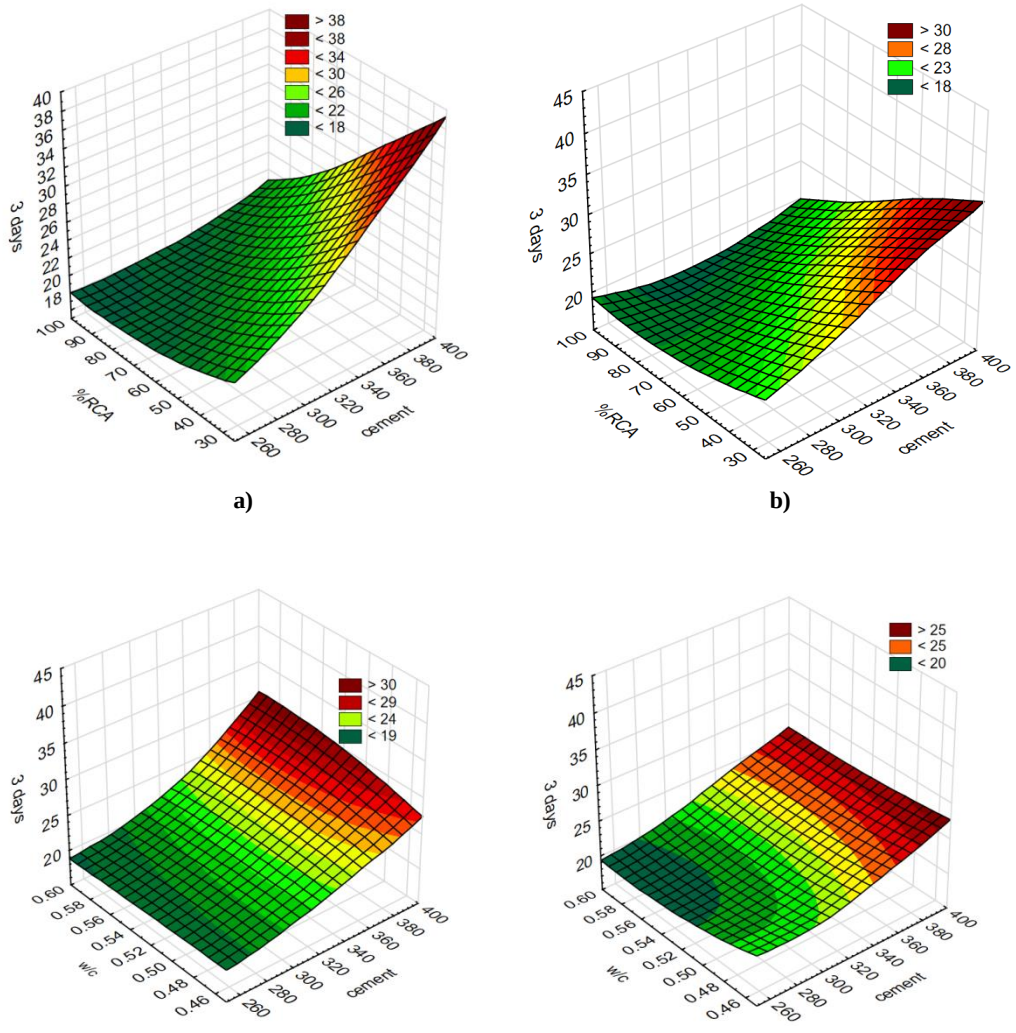


Fig. 3. The network performance results for Fc after 3 days depending on %RCA and cement content for MLP (a) and RBF (b) networks as well as w/c and cement content for MLP (c) and RBF (d) networks

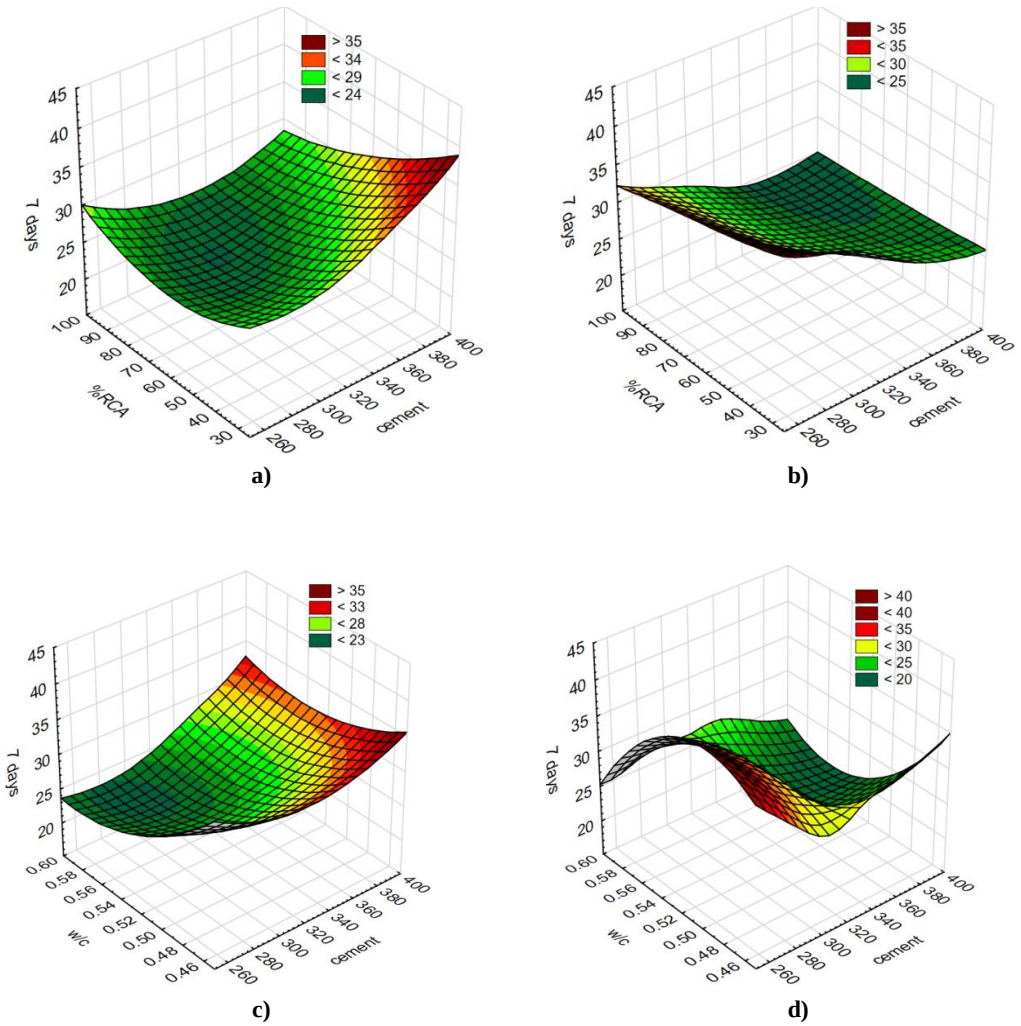


Fig. 4. The network performance results for Fc after 7 days depending on %RCA and cement content for MLP (a) and RBF (b) networks as well as w/c and cement content for MLP (c) and RBF (d) networks

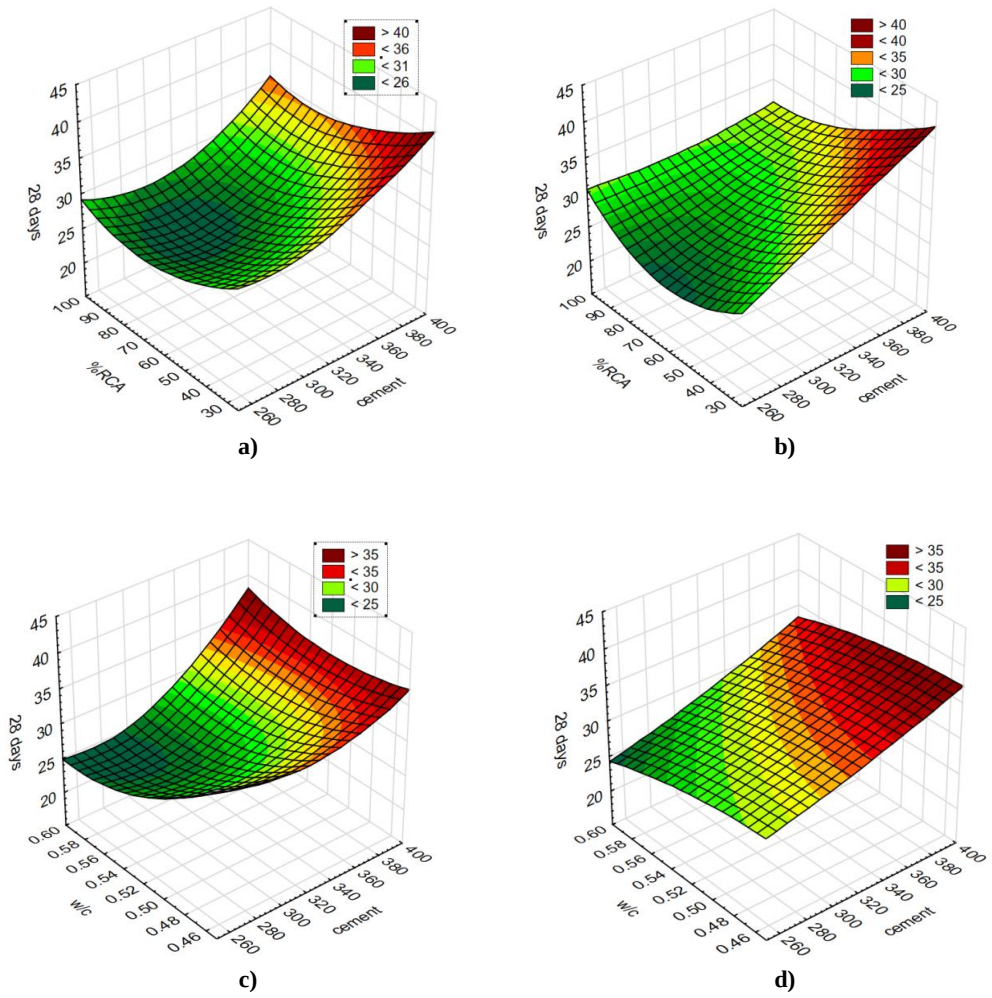


Fig. 5. The network performance results for Fc after 28 days depending on %RCA and cement content for MLP (a) and RBF (b) networks as well as w/c and cement content for MLP (c) and RBF (d) networks

In this study, an artificial neural network was developed to evaluate the compressive strength properties of recycled concrete aggregate based on the input variables, which were: cement content, %RCA, and the w/c ratio. The results of the modeling of compressive strengths after 3, 7 and 28 days and the prediction made enable to conclude that the MLP and RBF networks obtained have a satisfactory ability to predict these values. This is confirmed, among other things, by the R-correlation value of 0.97, the high quality of learning and validation of the network at 0.97, and the learning and validation errors. Comparing the experimental data and simulated values of compressive strengths after 3, 7 and 28 days, it can be concluded that the relative error value does not exceed 15%, which indicates that the network is well trained.

There is little work in the available literature on predicting the compressive strength of RCA concretes, where in most previous studies were particularly centric about high-performance concrete (HPC) containing blast furnace slag (BFS), fly ash (FA) and superplasticizer. The networks created for these types of concrete achieved R^2>90% (Yeh & Lien, 2009; Chou et al., 2011; Deepa, Kumari & Sudha, 2010; Atici, 2011; Erdal, Karakurt & Namli, 2013; Qmran et al., 2016).

The obtained results in statistical terms do not differ much in that respect. Hammoudi et al. (2019) predicted the compressive strength of a concentrate containing RCA after 7, 28 and 56 days. The input data included: content cement, %RCA, slump. For the employed model, he obtained the correlation coefficient of 0.98% (Hammoudi et al., 2019).

Naderpour et al. used an artificial neural network to evaluate the strength properties of recycled aggregate concrete based on input variables: water-cement ratio, water absorption, fine aggregate, recycled concrete aggregate, natural concrete aggregate, water-total material ratio, and 28-day compressive strength. He achieved lower values of correlation coefficient than in the presented study. The correlation values of their neural network for training, validation and testing reached 0.903, 0.89 and 0.829 respectively (Naderpour, Rafiean, & Fakharian, 2018).

The use of ANN model, which based on experimental results showed that it is useful and efficient model to predict the compressive strength. Wider application of ANN methods will facilitate determining the composition of concretes with recycled aggregate addition and manufacturing of new building materials.

4. CONCLUSIONS

The main objective of the study was to present an ANN model for predicting the compressive strength of concrete containing recycled concrete aggregate following 3, 7 and 28 day. The following were used as input date: different ranges of cement content, %RCA and water content ratios.

The following conclusions can be drawn in connection with the research conducted on training artificial neural networks:

- For Fc modeling after 3, 7 and 28 days for both types of RBF and MLP networks, the quality of the models is at an acceptable level. In the case of MLP networks for individual networks the quality of training and validation were respectively – for Fc after 3 days the quality of training it was 0.9907, validation was 0.9941, for Fc after 7 days the quality of training was 0.9932, validation was 0.9988, and for Fc after 28 days the quality of training was 0.9725, validation was 0. 9952. In the case of RBF networks for individual networks, the training and validation quality were as follows: for Fc after 3 days the training quality – 0.9714, validation – 0.9961, for Fc after 7 days the training quality – 0.9951, validation – 0.9989, while for Fc after 28 days the training quality – 0.9881, and validation – 0.9943.
- The networks obtained by modeling Fc after 3, 7 and 28 days show satisfactory predictive ability, as evidenced by the obtained correlation values R. These are $R_{MPL-3\ days} = 0.9876$, $R_{MPL-7\ days} = 0.9962$, $R_{MPL-28\ days} = 0.9903$, $R_{RBF-3\ days} = 0.9773$, $R_{RBF-7\ days} = 0.9908$, $R_{RBF-28\ days} = 0.9813$. Thus, it can be concluded that artificial

neural networks are an effective tool that can be used to predict compressive strengths after 3, 7 and 28 days.

- The trained networks show the relationships between the input data (cement content, %RCA and the ratios of water contents) and the output data (Fc after 3, 7 and 28 days), allowing the determination of the corresponding values of the analyzed indicators after the input of the set parameters into the network .
- A model to predict the compressive strength of concretes with recycled coarse aggregates can be the basis for creating optimal concrete compositions with RCA. It will save time and effort, as well as eliminate the costs that are incurred when manufacturing new construction materials.

Funding

This research was funded by the Polish Ministry of Science and Higher Education, grant numbers: FD-20/IM-5/061/2022 and FD-NZ-062/2022.

Conflicts of Interest

The authors declare no conflict of interest.

REFERENCES

Anysz, H., & Narloch, P. (2019). Designing the composition of cement stabilized rammed earth using artificial neural networks. *Materials*, *12*(9), 1396. https://doi.org/10.3390/ma12091396

Atici, U. (2011). Prediction of the strength of mineral admixture concrete using multivariable regression analysis and an artificial neural network. *Expert Systems with Applications*, *38*(8), 9609–9618. https://doi.org/10.1016/j.eswa.2011.01.156

British Standards Institution. (1988). *British DOE (Department of Environment) Method (1988) Design of Normal Concrete Mixes*.

British Standards Institution. (2009). *BS EN 12390-3:2009 Testing hardened concrete. Compressive strength of test specimens.* https://bsol.bsigroup.com/en/Bsol-Item-Detail-Page/?pid=000000000030164906

CEM II/B-V 32,5 R. (n.d.). *Lafarge*. Retrieved June 10, 2022 from https://www.lafarge.pl/cement-cem-iib-v-325-r-disabled-page

Chen, H., Qian, C., Liang, C., & Kang, W. (2018). An approach for predicting the compressive strength of cement-based materials exposed to sulfate attack. *PLoS ONE*, *13*(1), 1–17. https://doi.org/10.1371/journal.pone.0191370

Chopra, P., Kumar, R., & Kumar, M. (2015). Artificial Neural Networks for the Prediction of Compressive Strength of Concrete. *International Journal of Applied Science and Engineering*, *13*, 187–204.

Chou, J., Chiu, C., Farofura, M., & Al-Tharawa, I. (2011). Optimizing the Prediction Accuracy of Concrete Compressive Strength Based on a Comparison of Data-Mining Techniques. *Journal of Computing in Civil Engineering*, *25*(3), 1171.

Dahmoune, F., Remini, H., Dairi, S., Aoun, O., Moussi, K., Bouaoudia-Madi, N., Adjeroud, N., Kadri, N., Lefsih, K., Boughani, L., Mouni, L., Nayak, B., & Madani, K. (2015). Ultrasound assisted extraction of phenolic compounds from P. lentiscus L. leaves: Comparative study of artificial neural network (ANN) versus degree of experiment for prediction ability of phenolic compounds recovery. *Industrial Crops and Products*, *77*(0926–6690), 251–261. https://doi.org/https://doi.org/10.1016/j.indcrop.2015.08.062

Dantas, A. T. A., Leite, M. B., & Nagahama, K. J. (2013). Prediction of compressive strength of concrete containing construction and demolition waste using artificial neural networks. *Construction and Building Materials*, *38*, 717–722. https://doi.org/10.1016/j.conbuildmat.2012.09.026

Das, S., Swetapadma, A., & Panigrahi, C. (2019). A study on the application of artificial intelligence techniques for predicting the heating and cooling loads of buildings. *Journal of Green Building*, *14*(3), 115–128. https://doi.org/10.3992/1943-4618.14.3.115

Deepa, C., Kumari, K. S., & Sudha, V. P. (2010). Prediction of the compressive strength of high performance concrete mix using tree based modeling. *International Journal of Computer Applications*, *6*(5), 0975–8887.

Erdal, H. I., Karakurt, O., & Namli, E. (2013). High performance concrete compressive strength forecasting using ensemble models based on discrete wavelet transform. *Engineering Applications of Artificial Intelligence*, *26*(4), 1246–1254. https://doi.org/10.1016/j.engappai.2012.10.014

Fei, X., Youfu, S., & Xuejun, R. (2019). A simulation analysis method based on PSO-RBF model and its application. *Cluster Computing*, *22*(s1), 2255–2261. https://doi.org/10.1007/s10586-018-2596-y

Gjorv, O. D., & Sakai, K. (2014). *Concrete Technology for a Sustainable Development in the 21st Century*. CRC Press. https://doi.org/https://doi.org/10.1201/9781482272215

Hammoudi, A., Moussaceb, K., Belebchouche, C., & Dahmoune, F. (2019). Comparison of artificial neural network (ANN) and response surface methodology (RSM) prediction in compressive strength of recycled concrete aggregates. *Construction and Building Materials*, *209*, 425–436. https://doi.org/10.1016/j.conbuildmat.2019.03.119

Han, Y., Fu, S., Wang, S., & Xie, Z. (2018). Study on Adiabatic Temperature Rise Reflecting Hydration Degree of Concrete. *Advances in Materials Science and Engineering*, *2018*, 435049. https://doi.org/10.1155/2018/1435049

Karpiński, R. (2022). Knee joint osteoarthritis diagnosis based on selected acoustic signal discriminants using machine learning. *Applied Computer Science*, *18*(2), 71–85. https://doi.org/10.35784/acs-2022-14

Karpiński, R., Krakowski, P., Jonak, J., Machrowska, A., Maciejewski, M., & Nogalski, A. (2022a). Diagnostics of Articular Cartilage Damage Based on Generated Acoustic Signals Using ANN—Part II: Patellofemoral Joint. *Sensors*, *22*(10), 3765. http://dx.doi.org/10.3390/s22103765

Karpiński, R., Krakowski, P., Jonak, J., Machrowska, A., Maciejewski, M., & Nogalski, A. (2022b). Diagnostics of Articular Cartilage Damage Based on Generated Acoustic Signals Using ANN—Part I: Femoral-Tibial Joint. *Sensors*, *22*(6), 2176. http://dx.doi.org/10.3390/s22062176

Kurpinska, M., & Kułak, L. (2019). Predicting performance of lightweight concrete with granulated expanded Glass and Ash aggregate by means of using Artificial Neural Networks. *Materials*, *12*(12), 2002. https://doi.org/10.3390/ma12122002

Machrowska, A., Karpiński, R., Jonak, J., Szabelski, J., & Krakowski, P. (2020a). Numerical prediction of the component-ratio-dependent compressive strength of bone cement. *Applied Computer Science*, *16*(3), 88-101. https://doi.org/10.23743/acs-2020-24

Machrowska, A., Szabelski, J., Karpiński, R., Krakowski, P., Jonak, J., & Jonak, K. (2020b). Use of Deep Learning Networks and Statistical Modeling to Predict Changes in Mechanical Parameters of Contaminated Bone Cements. *Materials*, *13*(23), 5419. http://dx.doi.org/10.3390/ma13235419

Naderpour, H., Rafiean, A. H., & Fakharian, P. (2018). Compressive strength prediction of environmentally friendly concrete using artificial neural networks. *Journal of Building Engineering*, *16*(January), 213–219. https://doi.org/10.1016/j.jobe.2018.01.007

Orosa, J. A., Vergara, D., Costa, Á. M., & Bouzón, R. (2019). A novel method based on neural networks for designing internal coverings in buildings: Energy saving and thermal comfort. *Applied Sciences*, *9*(10), 2140. https://doi.org/10.3390/app9102140

Parichatprecha, R., & Nimityongskul, P. (2009). Analysis of durability of high performance concrete using artificial neural networks. *Construction and Building Materials*, *23*(2), 910–917. https://doi.org/10.1016/j.conbuildmat.2008.04.015

Pezeshki, Z., & Mazinani, S. M. (2019). Comparison of artificial neural networks, fuzzy logic and neuro fuzzy for predicting optimization of building thermal consumption: a survey. *Artificial Intelligence Review*, *52*(1), 495–525. https://doi.org/10.1007/s10462-018-9630-6

Pytka, J., Budzyński, P., Tomiło, P., Michałowska, J., Błażejczak, D., Gnapowski, E., Pytka, J., & Gierczak, K. (2022). Measurement of aircraft ground roll distance during takeoff and landing on a grass runway. *Measurement*, *195*, 111130. https://doi.org/10.1016/J.MEASUREMENT.2022.111130

Qmran, B. A., Chen, Q., Asce, A. M., & Jin, R. (2016). Comparison of Different Data Mining Techniques for Predicting Compressive Strength of Environmentally Friendly Concrete. *Journal of Computing in Civil Engineering*, *30*(6), 2208.

Rymarczyk, T., Kłosowski, G., Hoła, A., Sikora, J., Wołowiec, T., Tchórzewski, P., & Skowron, S. (2021). Comparison of Machine Learning Methods in Electrical Tomography for Detecting Moisture in Building Walls. *Energies*, *14*(10), 2777. http://dx.doi.org/10.3390/en14102777

Rymarczyk, T., Klosowski, G., Kozlowski, E., & Tchórzewski, P. (2019). Comparison of Selected Machine Learning Algorithms for Industrial Electrical Tomography. *Sensors*, *19*(7), 1521. https://doi.org/10.3390/S19071521

Szabelski, J., Karpiński, R., & Machrowska, A. (2022). Application of an Artificial Neural Network in the Modelling of Heat Curing Effects on the Strength of Adhesive Joints at Elevated Temperature with Imprecise Adhesive Mix Ratios. *Materials*, *15*(3), 721. http://dx.doi.org/10.3390/ma15030721

Szala, M., Awtoniuk, M., Latka, L., MacEk, W., & Branco, R. (2021). Artificial neural network model of hardness, porosity and cavitation erosion wear of APS deposited Al2O3 -13 wt% TiO2 coatings. *Journal of Physics: Conference Series*, *1736*(1), 012033. https://doi.org/10.1088/1742-6596/1736/1/012033

Topçu, I. B., & Saridemir, M. (2008). Prediction of mechanical properties of recycled aggregate concretes containing silica fume using artificial neural networks and fuzzy logic. *Computational Materials Science*, *42*(1), 74–82. https://doi.org/10.1016/j.commatsci.2007.06.011

Yeh, I. C., & Lien, L. C. (2009). Knowledge discovery of concrete material using Genetic Operation Trees. *Expert Systems with Applications*, *36*(3 PART 2), 5807–5812. https://doi.org/10.1016/j.eswa.2008.07.004

Applied Computer Science, vol. 18, no. 4, pp. 82–95
doi: 10.35784/acs-2022-30

Submitted: 2022-10-09 | Revised: 2022-11-27 | Accepted: 2022-12-21

Keywords: pump, CFD, the head of the pump, power, efficiency

Łukasz SEMKŁO [0000-0002-1737-8793]*, *Łukasz GIERZ* [0000-0003-4040-5718]**

NUMERICAL AND EXPERIMENTAL ANALYSIS OF A CENTRIFUGAL PUMP WITH DIFFERENT ROTOR GEOMETRIES

Abstract

The paper presents a comparative analysis of the operation of two variants of centrifugal pump rotors, a description of the main parameters, and the influence of the blade geometry on the performance characteristics obtained. Rotors have been designed using the arc and point method. Based on the developed 3D CAD models, the rotors were printed using the rapid prototyping method on a 3D printer in FFF (Fused Filament Fabrication) technology, in order to experimentally verify the performance, by placing them on the Armfield FM50 test stand. The analysis part of the CFD includes a fluid flow in Ansys Fluent. The process of creating a flow domain and generating a structural mesh was described, along with the definition of boundary conditions, the definition of physical conditions and the turbulence model. The distribution of pressures and velocities in the meridional sections is shown graphically. The chapter with the experimental analysis contains a description of the measuring stand and the methodology used. The results obtained made it possible to generate the characteristics, making it possible to compare the results received. The results allowed to note the influence of geometry on the behavior of the rotors during operation in the system and to indicate that the arc rotor gets a 7% higher head and 2% higher efficiency than the point method rotor, which gives the basis for its commercial use in industry.

1. INTRODUCTION

During development works, CFD computer analysis is more and more often used, which is validated through experimental studies (Cheah et al., 2007). The application of these two methods opens up a wide range of possibilities, as it clearly shows their convergence. CFD analysis allows us to significantly reduce the costs of designing machines and devices. A lot of research work has already been done in which numerical simulation was used to design rotors. The relationships between rotors and stators were analyzed (Zhu et al., 2011), the rotor stabilization itself (Steinbrecher et al., 2003), pump performance under normal and cavitation conditions (Mousmoulis et al., 2021), and critical velocities during pump operation

* Poznan University of Technology, Institute of Thermal Energy, Faculty of Environmental Engineering and Energy, Poznan, Poland
** Poznan University of Technology, Institute of Machine Design, Faculty of Mechanical Engineering, Poznan, Poland, lukasz.gierz@put.poznan.pl

were also analyzed (Song, Zhang & Zhang, 2022). Many other valuable analyzes were also performed (Barmaki & Eghaghi, 2019; Fan & Piao, 2017). In all of such studies, the focus was on numerical analysis. On the other hand, the research work did not forget about experiments related to pumping systems, pumps, and rotors supported by numerical design and analysis (Kaczmarczyk et al., 2019; Ciocan & Kueny, 2006; Li et al., 2020; Bosioc et al., 2019).

The pump is a machine used to pump the medium from the low pressure to the area of higher pressure area. It usually involves the mechanical conversion of electrical energy into hydraulic energy and energy from other sources may also be converted, including e.g., wind force and human strength (Jędral, 2001). The centrifugal pump uses rotating elements to increase torsion (Kijewski, 1993). The increase in angular momentum causes a pressure difference between the inlet and outlet sides, setting the fluid in motion (Ciałkowski et al., 2015). Centrifugal pumps work in a similar way to all turbines; according to Newton's second principle, the fluid flowing through the rotating working organ increases its energy, which is then partially converted into hydrodynamic energy of the fluid flow (Cengel & Cimbala, 2013). The value of its increase depends on the selected rotor structure and rotational speed (Anderson, 1980). There are two families of centrifugal pumps, rotodynamic and circulating (Polish Standard PN-90 / M-44000).

A centrifugal pump was taken into account for the research because it is widely used both in industry and by individual users. Centrifugal pumps are used to transport fluid in central heating systems and water supply systems. They are widely used in the food industry to transport juices, wort, and non-sticky liquids, used for technological processes. The study attempted to design a rotor of such a pump and performed a simulation and experimental analysis. Nowadays, when attention is paid to energy savings, this work is concerned with determining the efficiency of a rotor designed in two variants (point-arc method) and showing which of these variants will be more energy-efficient.

The purpose of this study was to perform a comparative analysis of two rotor variants dedicated to a centrifugal pump using the simulation and experimental method. It was decided to show the influence of two variants of the rotor geometry on the obtained performance characteristics of the centrifugal pump, with an indication of which structure would have a chance to be implemented in the industry.

2. METHODS

2.1. Pomp project

For research purposes, two rotors were designed with different blade curvature. During the preparation of the 3D CAD model, the arc and point methods were chosen. The arc method assumes that the continuity of meridional velocity changes is omitted and the drawing itself takes place by drawing a single arc between the diameters d_1 and d_2. The second chosen method is the method of marking the course of the shoulder blade line according to C. Pfleiderer. This method, used in centrifugal pumps with a single curvature of the blades, consists in assuming a change in the angle β depending on the radius r in the range from r_1 to r_2 and determining the angle ϑ for a given r and β. The values of r and ϑ are the polar coordinates of a given point of the blade and allow to plot the skeleton of the blade.

Table 1 presents the calculation results of two variants of the designed rotors, where characteristic diameters and angles are indicated. Figure 1 shows the meridional section of the rotor determined based on the results obtained from the calculations (see Table 1). Figure 2 shows the geometric interpretation of the blade curvature based on the point method and shows the angle of blade coverage. This angle was determined at the assumption level of $\theta = 69.39°$. Figure 3 shows the angle of blade coverage determined on the basis of the arc method. This angle was determined at the level of $\theta = 24.85°$. The angles shown in both methods have been calculated. The indicated research results were obtained on the basis of relations available in the literature (Troskolański, 1973).

Tab. 1. Calculation results

Parameter	Symbol	Value
Shaft diameter	d_w	15 mm
Hub diameter	d_p	21 mm
Inner diameter	d_1	46 mm
Inlet diameter	d_w	48 mm
Shockless inflow angle	β_1	22°
Plow angle at the outlet	β_2	30°
Outer diameter	d_2	120 mm
Blade width - leading edge	b_1	6 mm
Blade width - trailing edge	b_2	11 mm

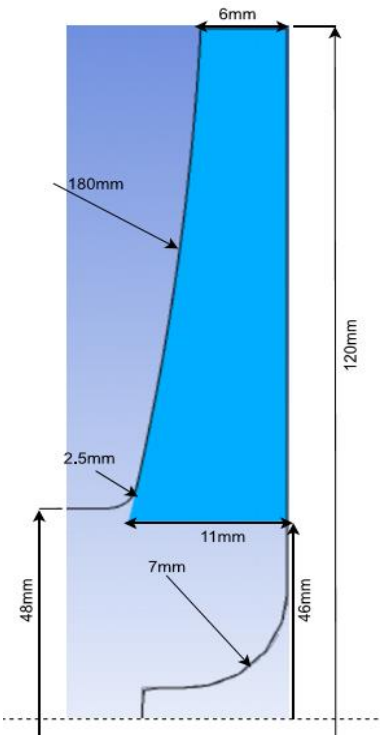


Fig. 1. Rotor meridion cross-section

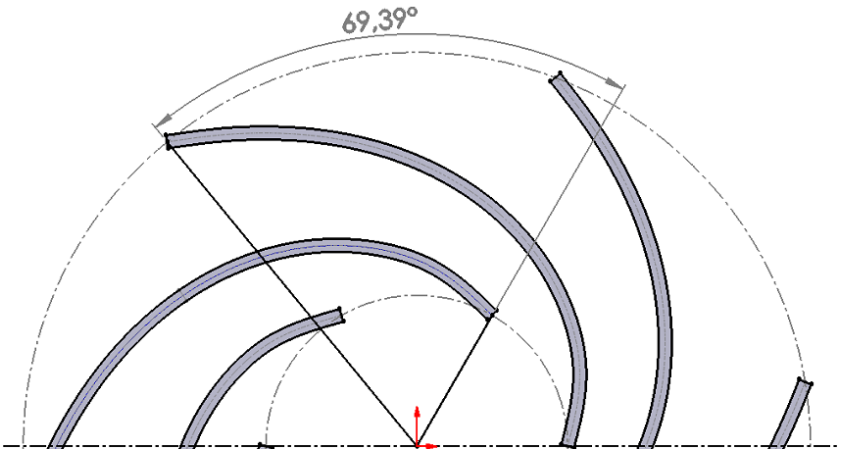


Fig. 2. Blade coverage angle measurement for the point method, θ=69,39°

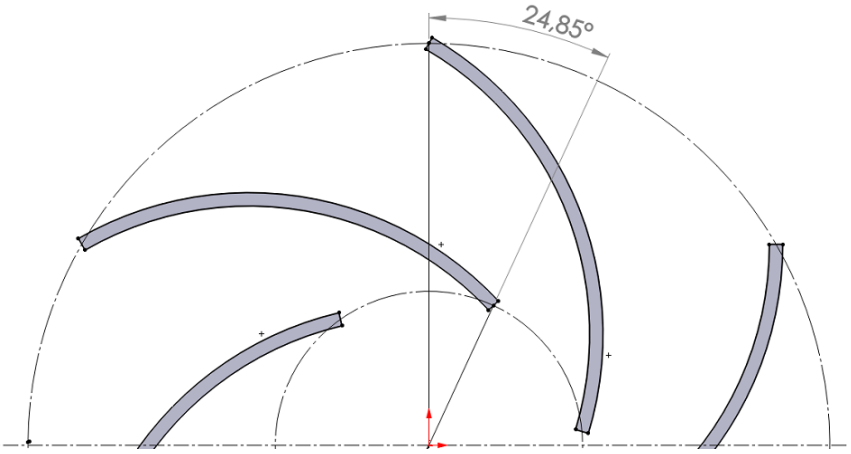


Fig. 3. Blade coverage angle measurement for the arc method; θ=24,85°

2.1. CFD analysis

Numerical analysis was performed in the ANSYS Fluent software (version 19.2). The purpose of the analysis was to compare the behavior of the characteristics of closed rotors with different blade coverage angles, ie, head H, rotor power P, and its efficiency μ. The simulation was performed with the following assumptions: frozen rotor; k-epsilon turbulence model; transported medium – water with a density of 1000 kg/m^3; initial pressure 1 bar; the boundary was taken as p-Total Inlet Mass Flow Outlet. The assumed boundary conditions are shown in Fig. 4. All basic results of the numerical analyzes, i.e., reference diameter, volumetric flow rate, head of the pump H, rotor power P and efficiency μ were collected and presented in Table 2. Based on these data, it can be concluded that the first two parameters are identical, while the next three show differences, where the curved rotor has a higher head of 2.57 m and a lower efficiency of 64.9%, while the point method rotor has a lower head of 2.44 m but a higher efficiency of 66.4%. Figure 5 shows the mesh of one of

the rotors. The tetra net consisted of 230,000 elements. Figure 6 shows the 3D CAD model of both rotors, i.e., made a) with the arc method and b) with the point method. In both rotors, the number of blades is 6 and the main difference between the methods is the radius of curvature of the blades. For the arc method, the radius of curvature is 50 mm, and for the point method it is 60 mm. The value of the radius of curvature of the blades determines their length. So the blades in the arc method are shorter and in the point method the blades are longer. Figure 7 shows the pressure distribution on the discharge side of the pump impeller, while Figure 8 shows the velocity distribution throughout the vane. Based on the numerical analysis using the point method, higher discharge pressure were obtained (Fig. 7). In addition, in the point method, a higher value of the velocity of the fluid in the blade was obtained. It is also possible to observe a uniform velocity distribution in the case of the point method in Figure 8. Analysis of the velocity distribution on the blade for the arc method showed the formation of a local reduction in the velocity value.

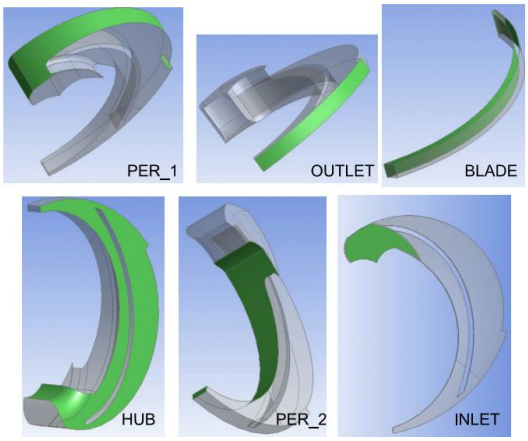


Fig. 4. The assumed boundary conditions

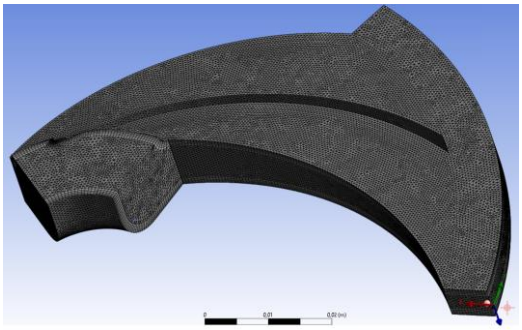


Fig. 5. Calculation grid

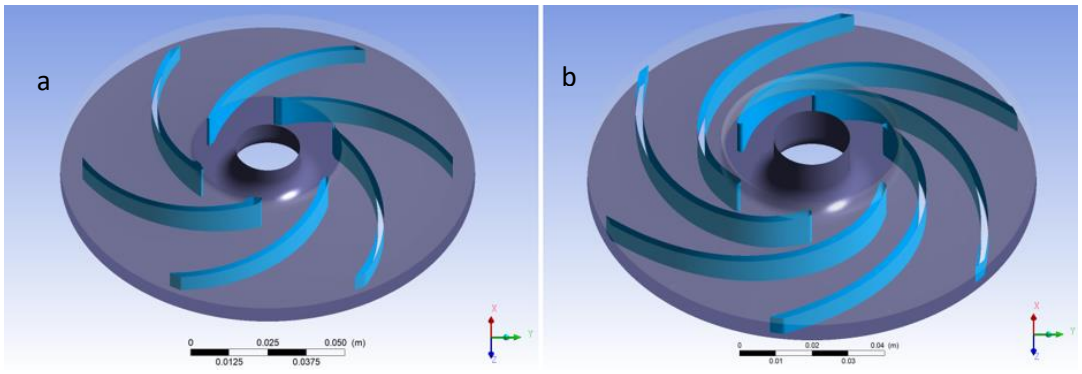


Fig. 6. Isometric view of rotors: a) arc method, b) point method

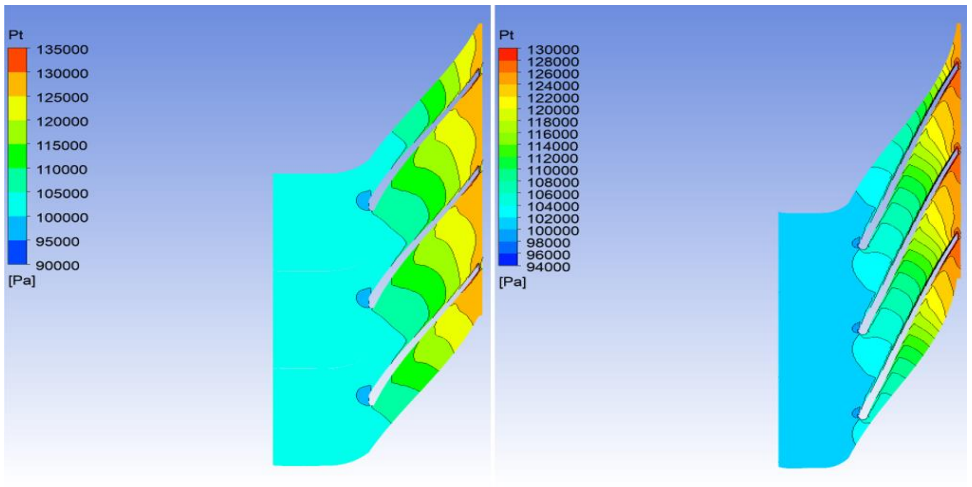


Fig. 7. Discharge pressure P_t: left – arc method; on the right – point method

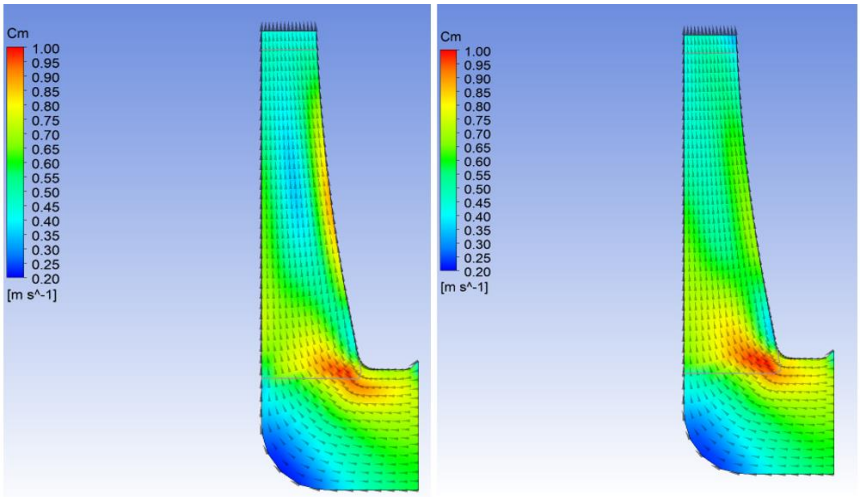


Fig. 8. Velocity vectors on the blade surface: left – arc method; on the right – point method

Tab. 2. The results obtained from the CFD analysis (results for two variants of rotors)

Parameter	Arc method	Point method	Unit
Reference diameter	0.12		m
Volume flow rate	0.0010		m^3/s
Head H	2.57	2.44	m
Power P	25.89	26.28	W
Efficiency μ	64.9	66.4	%

2.2. Experimental analysis

In order to carry out the experiment, the rotors were made using the rapid prototyping method on a 3D printer in FFF (Fused Filament Fabrication) technology. The material used was polylactide PLA, a bioplastic derived from renewable raw materials where additives may be, for example, corn starch or sugar cane.

The tests were carried out on the Armfield FM50 stand. A schematic image of the stand is shown in Figure 9. The system consists of a centrifugal pump driven by an electric motor 4 with a power of 250 W and an acrylic tank 1 with a capacity of 20 liters connected to a small hydraulic system. The water flow through the pump is controlled by a regulating valve in the range of 0 l/min to 96 l/min on the pump discharge side. Manually operated valves at the pump inlet and outlet allow flow control from 0 to 1.6 l/s. The integrated flow sensor allows you to analyze the pump performance and the parameters are controlled using the dedicated FM50 software (Fig. 10), which allows you to record information from the built-in sensors (thermocouple and two pressure sensors), while making the necessary calculations for the needs of the tests. The stand consists of: 1 – reservoir; 2 – control valve on the pump discharge side; 3 – flow meter; 4 – electric motor; 5 – location of the tested rotor; 6 – ball valve on the suction side of the pump; 7 – thermocouple; 8 – base; 9 – tank drain valve; 10 – drainage; 11 – rotor; 12 – drain valve from the rotor body; 13 – pressure measurement on the discharge side; 14 – pressure sensor on the suction side.

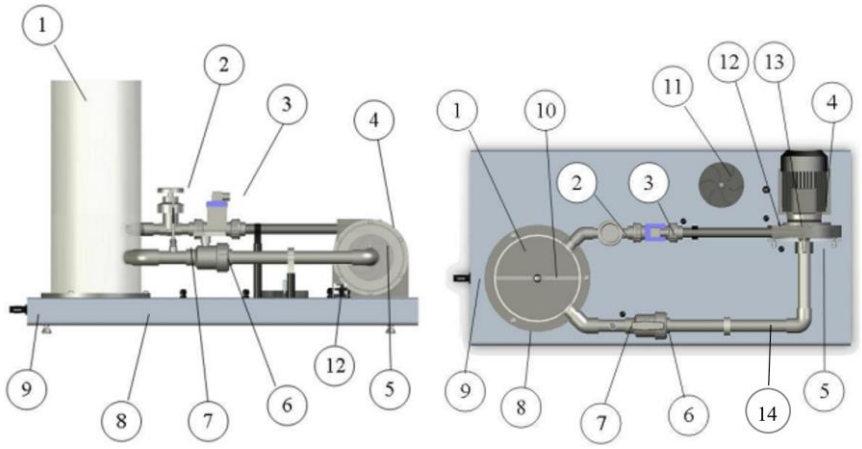


Fig. 9. Diagram of the test stand 1 – reservoir; 2 – control valve on the pump discharge side; 3 – flow meter; 4 – electric motor; 5 – location of the tested rotor; 6 – ball valve on the suction side of the pump; 7 – thermocouple; 8 – base; 9 – tank drain valve; 10 – drainage; 11 – rotor; 12 – drain valve from the rotor body; 13 – pressure measurement on the discharge side; 14 – pressure sensor on the suction side

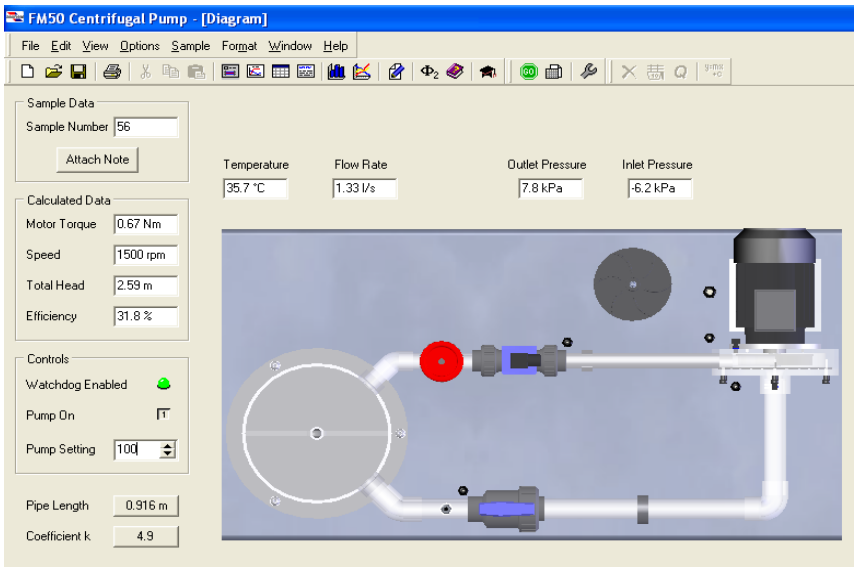


Fig. 10. The Armfield user interface

In order to validate the results of numerical tests (CFD) as part of the planned test program, the discharge part of the pump was throttled through the control valve. This was done by reducing the flow from the maximum value, i.e., 1.0 l/s, in steps of 0.1 l/s, to the set minimum flow value of 0.3 l/s. This procedure was performed with different rotor speed settings: 50%, 60%, 70%, 80%, 90% and 100%. The rotational speeds were similar: 750, 900, 1050, 1200, 1350, 1500 rpm. For each flow tested, the results were recorded (for all the ranges of the pump operation).

On the basis of such a specified test plan, the measurement results presented in Figures 11–18 were obtained. Fig. 11 shows the rotor characteristics for the point method, while Fig. 12 shows the rotor characteristics for the arc method. Thanks to these charts, it can be observed that, in the case of the arc method, it is possible to obtain a higher value of the head by as much as 7% and a better efficiency by 2% compared to the point method. Figures 13–15 show the change of individual characteristics at variable rotational speed for the point method. Figures 16–18 also show the change of individual elements of the pump characteristics at different values of rotational speed for the arc method.

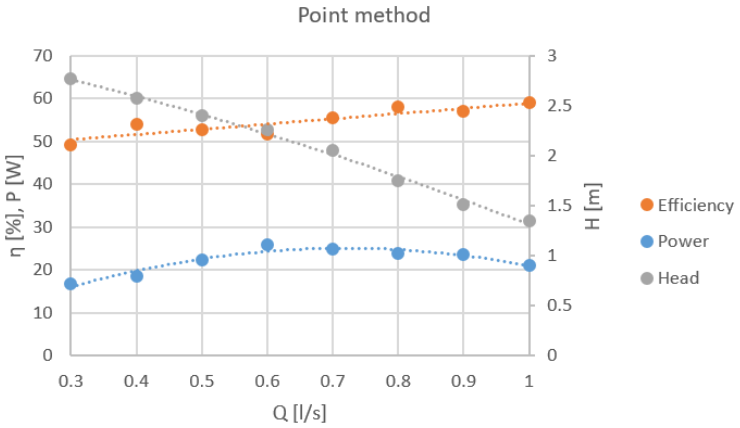


Fig. 11. Pump operation characteristics – point method

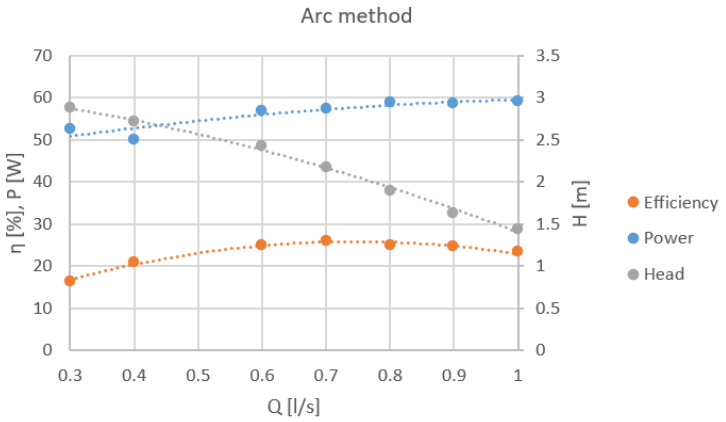


Fig. 12. Pump operation characteristics – arc method

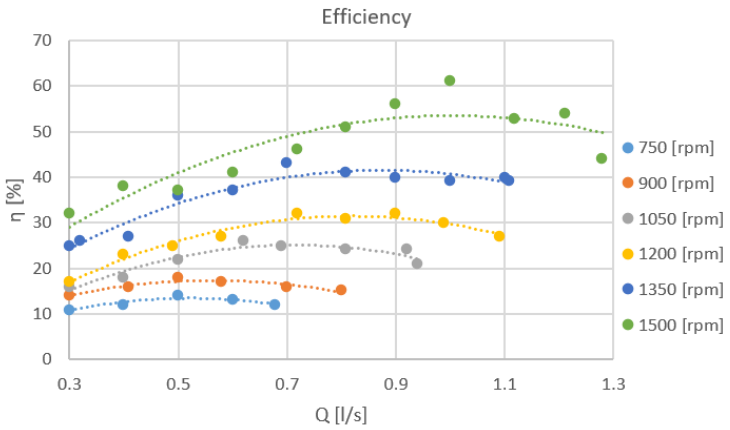


Fig. 13. Efficiency diagram for a rotor designed by the point method at different rotational speeds

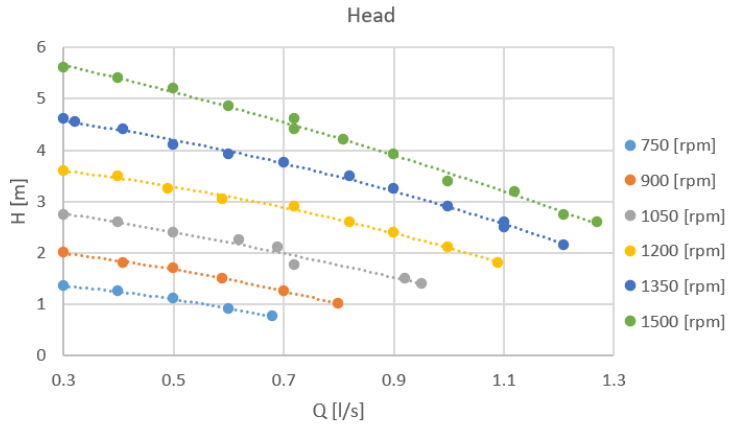


Fig. 14. Diagram of the head of the pump for a rotor designed by the point method at various rotational speeds

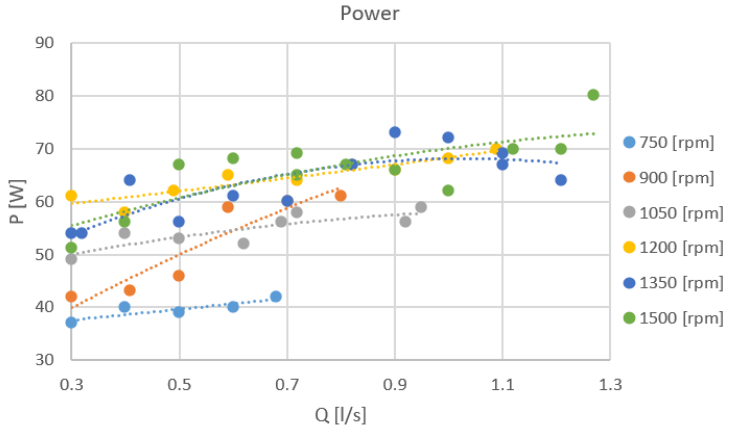


Fig. 15. Power diagram for a rotor designed by the point method at different rotational speeds

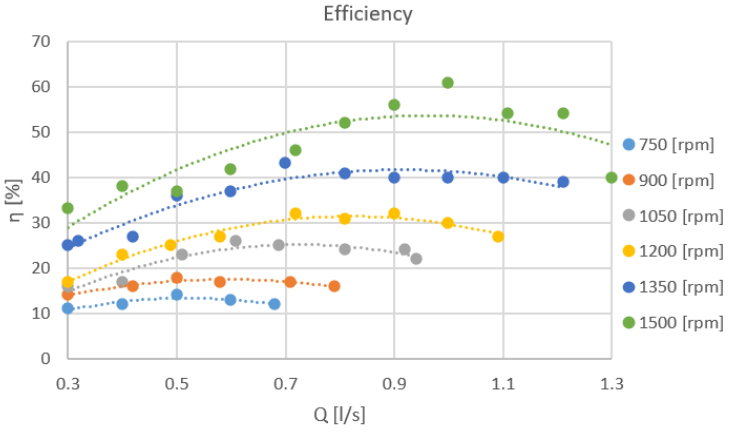


Fig. 16. Efficiency diagram for a rotor designed by the arc method at different rotational speeds

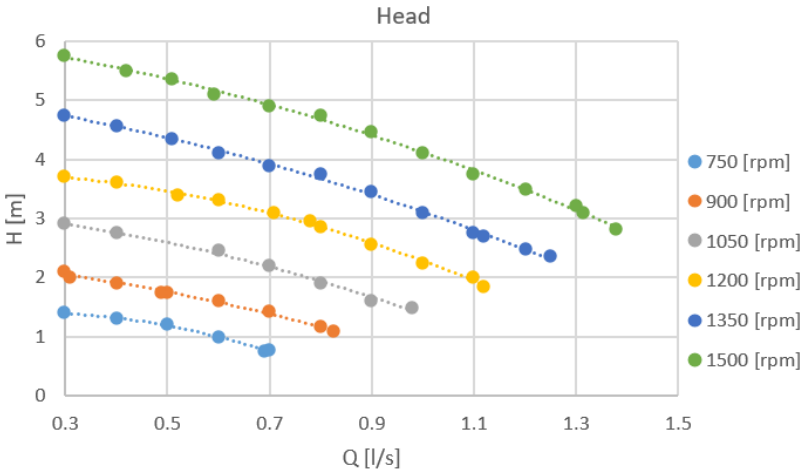


Fig. 17. Diagram of the head of pump for a rotor designed by the arc method at various rotational speeds

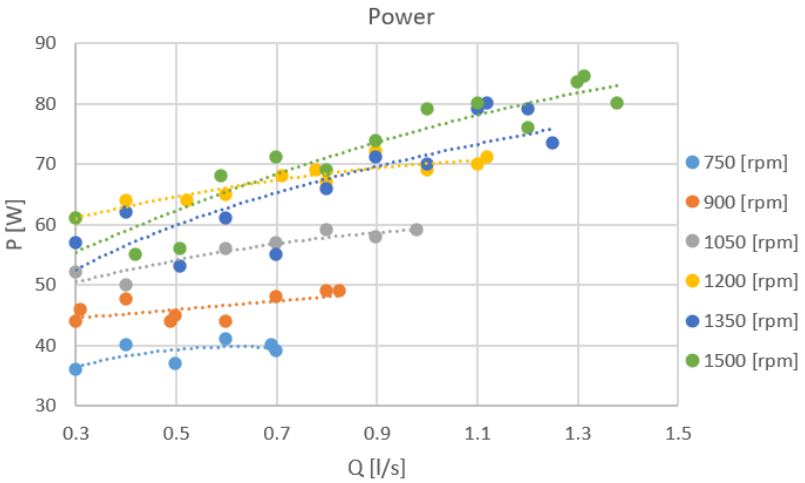


Fig. 18. Power diagram for a rotor designed by the arc method at different rotational speeds

In Figure 13 it can be seen how much influence the change of rotational speed has on the characteristics of the pump. At the lowest speed, the efficiency of the pump, due to the rotor designed by the point method, is from 10 to 14%. Increasing the rotational speed to 1500 rpm increases the efficiency to over 60%. Figure 14 shows the change in the head of pump depending on the rotational speed. At the lowest rotational speed of 750 rpm, the head of the pump H is less than 1.5 m. An increase in the rotational speed causes an increase in the head of pump H up to 5.8 m. Figure 15 shows the change in power generated by the pump. Increasing the engine speed from 750 rpm to 1500 rpm results in a power increase of over 80%.

Almost similar to the point method, the results were obtained for the rotor designed with the arc method. In Figure 16, we can observe an increase in pump efficiency from 10% to over 60%. The same applies for the head of pump. Increasing the rotational speed from 750 rpm to 1500 rpm generates an increase in the head of the pump from less than 1 meter to almost 6 meters. In the case of power in the arc method, the designed rotor increased the power from 35 W to 85 W.

3. DISCUSSION

3.1. Analysis at constant rotational speed

The performed analysis, consisting of the simulation of the flows in the Ansys environment, showed that the head of pump between the rotors (arc and point method) differs by 6%. The higher value of the head of pump of 2.58 m was observed for the rotor with a smaller blade cover angle, i.e. made with the arc method. Despite this, the efficiency of the rotor made by the point method is higher by slightly more than a percentage point. There were no significant differences in pressure distribution and velocity in the meridional sections.

In the experimental analysis the hights of raising at a flow rate of 1 l/s for the closed rotor of the point method and the arc method are 1.35 and 1.44 m. This means that the arc method was characterized by a higher head of pump of approx. 7%, which translates into greater efficiency by 2 percentage points compared to the rotor designed by the second (point) method. The value of the maximum efficiency in the tested flow range for both rotors was identical - although it was obtained for slightly different flows: 0.6l for the point flow and 0.7l for the arc flow. Based on the characteristics for a constant rotational speed of 1050 rpm, a slight influence of the blade shaping method on the performance of the rotor was found. These considerations concern the low specific speed which is 18.8.

Comparing the results of the numerical and experimental analysis at constant rotational speed, it can be seen that the results are convergent. The difference between the results in the case of efficiency in both methods is at the level of 8-9%. The values obtained during the experiment turned out to be lower than the values obtained in the numerical analysis. Interestingly, the situation was reversed in the case of the head of pump. The experimental results are slightly higher than the results obtained in the numerical analysis. The difference is 0.3 meters.

3.2. Analysis at Variable Speed

For the tests performed for variable rotational speed, the efficiency for each rotor was higher the higher the rotational speed. The flow rate corresponding to the maximum efficiency point was the higher the rotational speed. Additionally, higher speeds resulted in an increase in fluctuation and a more irregular course of the curve, making it difficult to estimate the efficiency for points between the flow rates.

The graphs of the head of pump (H) – similar to the efficiency (μ) – show that the increase in rotational speed caused the increase of the tested values. In this case, the changes were more predictable as the increase in rotational speed n contributed to the parallel shift of the curve upward. With the increase of $\dot{Q}$, the difference between successive values of the head of pump H increased – the graph of the course of the hight of increasing increased the slope in relation to the horizontal axis.

The power consumption curve P showed an upward trend with an increase in the flow rate – this applied to each of the tested rotors. Among the characteristics examined, the rotor power P curves had the most irregular course. The rotor designed with the arc method was characterized by the highest power obtained on the pump shaft of the two variants indicated.

3.3. Analysis of experimental studies with CFD

Comparing the experiment (Figures 11 and 12) with numerical calculations, some differences can be noticed. Numerical analysis showed higher values than experimental studies. In the case of the head, the differences in both methods are at the level of 40%. In the case of power, the differences are at the level of 11% for the arc method and 20% for the point method. The smallest difference in results can be observed in the case of efficiency. The difference in results is only 9% comparing the arc method and 11% in the case of the point method.

4. CONCLUSIONS

As part of this study, the performance of a pump was compared, in which two rotors with different parameters (two variants of rotors) were installed. The two types of analyzes performed (simulation CFD and experimental) showed that the designed closed rotors differed slightly from each other for nominal parameters. Both numerical and experimental analysis showed differences in the head of the pump H that did not exceed 7%. In the case of μ efficiency, also the differences were below 2 percentage points for the tested models.

In the experimental analysis, the relative differences between the tested rotors were similar to the results obtained during the simulation tests; however, the absolute values of the head of the pump H were slightly higher by 0.3 meters than indicated by the simulation in the Ansys program. The main reason for these discrepancies is that only the impellers were designed, not the entire pumping system.

In the longer term, the impact of the semi-open impeller on the pump operating parameters should be analyzed, as well as the structure of the entire pump body. Changing the geometry of the pump body could result in increased efficiency.

REFERENCES

Anderson, H. (1980). *Centrifugal Pumps*. Trade and Technical Press

Barmaki, R., & Ehghaghi, M. (2019). Experimental Investigation of a Centrifugal Pump Hydraulic Performance in Hydraulic Transmission of Solids. *Mechanics and Mechanical Engineering*, *23*(1), 259–270. http//doi.org/10.2478/mme-2019-0035

Bosioc, A., Moș, D., Draghici, I., Muntean, S., &Anton, L. E. (2019). Experimental analysis of a pump equipped with an axial rotor with variable speed. *IOP Conference Series: Earth and Environmental Science*, *240*, 032021. https://doi.org/10.1088/1755-1315/240/3/032021

Cengel, Y., & Cimbala, J. (2013). *Fluid Mechanics Fundamentals and Applications*. McGraw Hill.

Cheah, K., Lee, T., Winoto, S., & Zhao, Z. (2007). Numerical Flow Simulation in a Centrifugal Pump at Design and Off-Design Conditions. *International Journal of Rotating Machinery*, *2007*, 083641. http://doi.org/10.1155/2007/83641

Ciałkowski, M., Brodzik, Ł., Wróblewska, A., Frąckowiak, A., Bartoszewicz, J., Joachmiak, M., & Semkło, Ł. (2015). *Mechanika płynów – zbiór zadań z rozwiązaniami*. Wydawnictwo Politechniki Poznańskiej.

Ciocan, G., & Kueny, J.-L. (2006). Experimental Analysis of the Rotor-Stator Interaction in a Pump-Turbine. *23rd IAHR Symposium on Hydraulic Machinery and Systems.* Yokohama, Japan.

Fan, H., & Piao, Y. (2017). Cooling design of an aero-engine fuel centrifugal pump at shut-off. *Advances in Mechanical Engineering*, *9*(6). http://doi.org/10.1177/1687814017709700

Jędral, W. (2001). *Pompy wirowe*. Wydawnictwo Naukowe PWN.

Kaczmarczyk, T., Ihnatowicz, E., Żywica, G., & Kaniecki, M. (2019). Experimental study of the prototype of a Roto-Jet pump for the domestic ORC power plant. *Archives of thermodynamics*, *40*(3), 83–108. http://doi.org/ 10.24425/ather.2019.129995

Kijewski, J. (1993). Maszynoznawstwo. Wydawnictwo Szkolne i Pedagogiczne

Li, W., Ji, L., Shi, W., Yang, Y., Awais, M., Wang, Y., & Xu, X. (2020). Correlation research of rotor–stator interaction and shafting vibration in a mixed-flow pump. *Journal of Low Frequency Noise, Vibration and Active Control*, *39*(1), 72–83. http://doi.org/10.1177/1461348419836530

Mousmoulis, G., Kassanos, I., Aggidis, G., & Anagnostopoulos, I. (2021). Numerical simulation of the performance of a centrifugal pump with a semi-open impeller under normal and cavitating conditions. *Applied Mathematical Modelling*, *89*(2), 1814–1834. https://doi.org/10.1016/j.apm.2020.08.074

Polish Standard PN-90/M-44000. *Przenośniki cieczy.*

Song, H., Zhang, J., & Zhang, F. (2022). Rotor strength and critical speed analysis of a vertical long shaft fire pump connected with different shaft lengths. *Scientific reports*, *12*, 9351. https://doi.org/10.1038/s41598-022-13320-z

Steinbrecher, Ch., Skoda, R., Schilling, R., Müller, N., Breitenbach, A., & Mendler, N. (2003). Numerical Simulation of a Self-Stabilizing Rotor of a Centrifugal Pump. *Proceedings of the ASME/JSME 2003 4th Joint Fluids Summer Engineering Conference* (pp. 71–78). ASME. http://doi.org/10.1115/FEDSM2003-45468

Troskolański, A. T. (1973). *Pompy wirowe*. Wydawnictwo Naukowo-Techniczne.

Zhu, L., Yuan, S., Yuan, J., Zhou, J., Jin, R., & Wang, H. (2011). Numerical simulation for rotor-stator interaction of centrifugal pump with different tongues. *Journal of Agricultural Engineering*, *27*(10), 50–55. http://doi.org/:10.3969/ j.issn.1002-6819.2011.10.009

Applied Computer Science, vol. 18, no. 4, pp. 96–115
doi: 10.35784/acs-2022-31

Submitted: 2022-11-01 | Revised: 2022-12-07 | Accepted: 2022-12-19

Keywords: COVID-19, cough recordings, machine learning, PCA, classification

Elmehdi BENMALEK [0000-0003-1078-1421]*,
Jamal EL MHAMDI [0000-0001-8219-3560]*, *Abdelilah JILBAB* [0000-0002-1577-9040]*,
Atman JBARI [0000-0002-1855-2503]*

A COUGH-BASED COVID-19 DETECTION SYSTEM USING PCA AND MACHINE LEARNING CLASSIFIERS

Abstract

In 2019, the whole world is facing a health emergency due to the emergence of the coronavirus (COVID-19). About 223 countries are affected by the coronavirus. Medical and health services face difficulties to manage the disease, which requires a significant amount of health system resources. Several artificial intelligence-based systems are designed to automatically detect COVID-19 for limiting the spread of the virus. Researchers have found that this virus has a major impact on voice production due to the respiratory system's dysfunction. In this paper, we investigate and analyze the effectiveness of cough analysis to accurately detect COVID-19. To do so, we performed binary classification, distinguishing positive COVID patients from healthy controls. The records are collected from the Coswara Dataset, a crowdsourcing project from the Indian Institute of Science (IIS). After data collection, we extracted the MFCC from the cough records. These acoustic features are mapped directly to the Decision Tree (DT), k-nearest neighbor (kNN) for k equals to 3, support vector machine (SVM), and deep neural network (DNN), or after a dimensionality reduction using principal component analysis (PCA), with 95 percent variance or 6 principal components. The 3NN classifier with all features has produced the best classification results. It detects COVID-19 patients with an accuracy of 97.48 percent, 96.96 percent f1-score, and 0.95 MCC. Suggesting that this method can accurately distinguish healthy controls and COVID-19 patients.

* E2SN, ENSAM de Rabat, Mohammed V University in Rabat, Morocco, elmehdi.benmalek@um5s.net.ma, mhamdi_jamal@yahoo.fr, a_jilbab@yahoo.fr, atman.jbari@ensam.um5.ac.ma

1. INTRODUCTION

The coronavirus disease 2019 (COVID-19) is a pandemic caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). As of July 2022, more than 560 million cases of COVID-19 have been reported in 223 countries, regions, and territories, resulting in more than 6.36 million deaths (World Health Organization). Symptoms of COVID-19 are nonspecific and the presentation of the disease can vary from no symptoms (asymptomatic patients) to severe pneumonia and death. In the majority of cases (about 80 percent), people infected with COVID-19 have mild to moderate symptoms (eg, cough, fever, fatigue) while 14 percent have severe symptoms (eg, dyspnea and hypoxemia), and 6 percent have critical clinical condition (eg, respiratory failure, septic shock, multi-organ failure) (Weng, Su & Wang, 2021).

To control the spread of COVID-19, effective patient screening is essential. So far, the gold standard screening method is the polymerase chain reaction-reverse transcription (RT-PCR) assay which was designed to detect SARS-CoV-2 genetically (Ismail, Deshmukh & Singh, 2021). Unfortunately, the RT-PCR takes time and necessitates the use of medical expertise, which may not be available. On the other hand, some investigations have found that RT-PCR testing has a high percentage of false-positive results (Ai et al. 2020; Yang et al., 2020). Also, the quick detection of this virus is very important, to reduce the likelihood and risk of COVID-19 spreading. As a result, researchers in the fields of virology, medicine, and artificial intelligence (AI) have stepped up to find creative ways to contain this issue. The AI community made major contributions and proposed computer-aided diagnostics systems that can help identify, forecast, and treat COVID-19. With the use of machine learning (ML) technology, computers can mimic human intelligence and find patterns and information in massive amounts of data to comprehend the spread of COVID-19, and speed up research and treatment.

Numerous researchers have suggested using speech signals and medical imaging to automatically detect patients suffering from COVID-19. A dataset comprised of 368 COVID-19 positive patients and 127 other pneumonia cases from two hospitals in China was used by (Wu et al., 2020) to apply ResNet50 with the multi-view fusion approach. They attained 76 percent accuracy, 81.1 percent sensitivity, 61.5 percent specificity, and 81.9 percent AUC. ResNet50 and CT scans were used by (Li et al., 2020) for an automatic coronavirus diagnosis (COVNet). 4536 chest CT samples were used in total (1296 COVID-19 cases, 1325 non-pneumonia observations, and 1735 community-acquired pneumonia cases). For COVID-19 instances, their approach achieved a sensitivity of 90 percent, specificity of 96 percent, and AUC of 96 percent. ResNet-18 architecture has the best overall precision and sensitivity of 98.5 percent and 98.6 percent, using CT scan images when (Benmalek, Elmhamdi & Jilbab, 2021) compared the performances of CT scan and CXR images in the diagnosis of COVID-19. To determine how distinguishable COVID-19 sounds are from those in asthma or healthy controls, (Brown et al., 2020) used coughs and breathing. Their model achieves an AUC of above 80 percent for all tasks. A voice-based approach has been suggested by (Han et al., 2021) to automatically identify those who have tested positive for COVID-19. AUC of 0.79, a specificity of 0.82, and a sensitivity of 0.68 have been achieved. A medical dataset comprising 328 cough sounds from 150 patients split into four classes (Healthy, Asthma, Bronchitis, and COVID-19) was used by (Pal & Sankarasubbu, 2020). With specificity and accuracy of 95.04 percent and 96.83 percent, respectively,

the experiment findings show that their model captures a more robust feature embedding to distinguish between COVID-19 patient coughs and diverse sorts of coughs that are not COVID-19.

Based on several studies that suggest the voice of COVID-19 patients is infected by the disease (Han et al., 2021), we aim in this work to detect COVID-19 by applying the MFCC, different machine learning classifiers, and dimensionality reduction using the PCA. The evolution of the models was done by the confusion matrix, sensitivity, specificity, precision, f1-score, and Matthews correlation coefficient (MCC). The main contribution of this paper is to explore the effectiveness of the ML and AI algorithms in improving disease screening, diagnosis, and monitoring of the COVID-19 pandemic and reducing the need for human involvement in a way that lessens the burden on the healthcare industry.

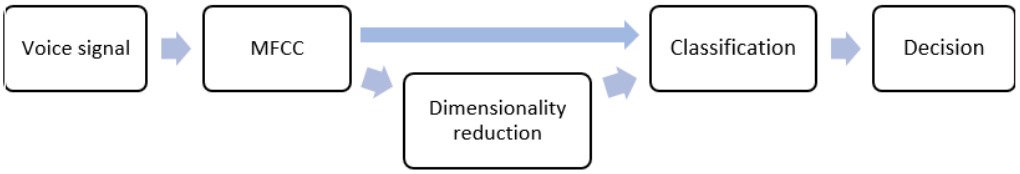


Fig. 1. COVID-19 diagnosis with cough recording diagram

2. METHODS

2.1. Dataset

The data is collected from the Indian Institute of Science Bangalore's Coswara Project (Sharma et al., 2020). The dataset is a collection of vowel (/a/, /e/, and /o/) sustained phonation, a counting exercise, breathing sounds, and cough recordings that we employed in this study. The gathering of records began on April 13th, 2020.

Reaching out to the global human population was the main goal of the data collection strategy. To do so, a website application was developed with a straightforward and interactive user interface. Open the application in a web browser on a computer or mobile device, enter the necessary metadata, without identification data, and then start recording sound samples with the device's microphone. The program is used on average for 5 to 7 minutes. The user was instructed to use a personal device, clean it with sanitizer before and after recording, and keep it 10 cm away from their mouths while they were recording. The dataset is divided into positive COVID-19 cases with 77 observations and 82 healthy controls representing the true negatives. The demographic information, symptoms, and comorbidities are represented in Tab. 1 and 2 for each class.

Tab. 1. Sex and age group of the participants for each class

		Age (years old)				
	Sex	**[20-29]**	**[30-39]**	**[40-49]**	**[50-59]**	**>=60**
Positive	45 Males 32 Females	48	9	6	11	3
Negative	54 Males 28 Females	44	11	7	12	8

Tab. 2. Disease symptoms and comorbidities of the participants for each class

	asthma	cold	cough	loss_of_smell	diabetes	fever	pneumonia
Positive	2	13	19	5	1	14	1
Negative	0	0	0	0	2	0	0

2.2. Mfcc

Mel Scale Frequency Spectral Coefficients are the most widely used parameters in speech processing for automated voice speaker, and language recognition, as well as the detection and classification of speech pathologies (Zheng, Zhang & Song, 2001). The principle for calculating MFCCs comes from psychoacoustic research on the perception of different frequency bands by the human ear. The main interest of these coefficients is to extract relevant information in a limited number based on both production (Cepstral theory) and speech perception (Mels scale) (Muda, Begam & Elamvazuthi, 2010). For this study, we used 14 MFCCs.

MFCCs are used in various speech processing techniques. The basic procedure for developing them is represented in Fig. 2:

- Convert Hertz to Mel Scale,
- Take the logarithm of the Mel representation of the audio,
- Take a logarithmic magnitude and use the discrete cosine transformation,
- This result creates a spectrum on Mel frequencies as opposed to time, thus creating MFCCs.

According to the MFCC calculation, the number of filters, their shape, how they are spaced and whether or not they overlap each other, and how the power spectrum is warped can all have an impact on how well the MFCC performs (Han et al., 2006).

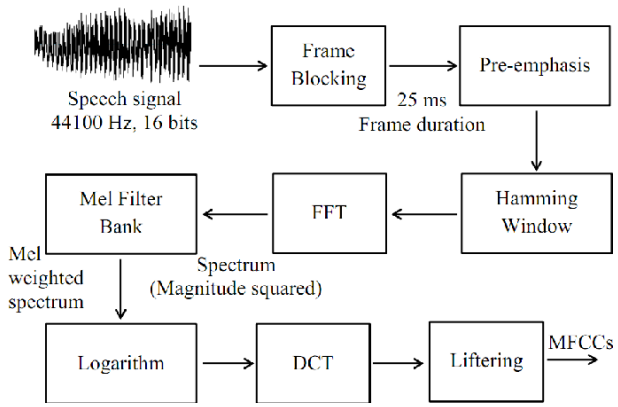


Fig. 2. Block Diagram of the computation steps of MFCC

2.3. Classification algorithms

a) K-NN

the K-NN (K-nearest neighbors) algorithm is a supervised learning method. It can be used for both regression and classification (Indyk & Motwani, 1998). Its operation can be likened to the following analogy “tell me who your neighbors are, I will tell you who you are…”.

To make a prediction, the K-NN algorithm will not calculate a predictive model from a training set as is the case for other machine learning classifiers. Indeed, K-NN does not need to build a predictive model. Thus, for K-NN there is no actual learning phase. This is why it is sometimes categorized as Lazy Learning. To be able to make a prediction, K-NN relies on the dataset to produce a result.

The importance of K is that it affects how accurate and efficient the algorithm is. Other KNN algorithm extensions that have been proposed include the weighted KNN classifier, the K-means KNN classifier, the Shared Nearest Neighbor KNN classifier, and the SVM KNN classifier. These reduce execution time and increase accuracy. A KNN example can be seen in Fig. 3, which includes training samples with two classes: "blue square" and "red triangle." The green circle designates the test sample. These samples are set up in two-dimensional feature spaces, where each feature has its own dimension. To identify whether a test sample belongs to the "blue square" or the "red triangle" class, KNN uses a distance function to identify the test sample's K closest neighbors. The test sample's class can be predicted by finding the majority of classes among the k closest neighbors. Due to the presence of two red triangles, the test sample in this instance is classified to the first class “red triangle” when k = 3. However, when k = 5, it is classified as the second class, "blue square," since there are two red triangles and three blue squares.

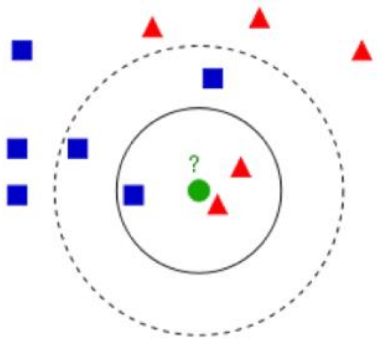


Fig. 3. An example of KNN classification with K neighbors K = 3 (solid line circle) and K = 5 (dashed line circle), the distance measure is Euclidean distance.

b) SVM

Support vector machines or wide-margin separators are a set of supervised learning techniques designed to solve regression or classification problems. They were introduced by V.Vapnik in 1995 in his book "The nature of statistical learning theory", but their first appearance was in 1992 after they were published by (Boser, Guyon & Vapnik, 1992).

The dimensionality of the data and its increased power of generalization, make the SVM more advantageous. SVM is widely used as a binary classifier in most fields. Its objective is to find the optimal boundary that separates two classes with the largest margin between the separation boundary and the support vectors (Fig. 4). SVM could surpass more sophisticated classifiers like deep neural networks for some classification problems, where the model selection in supervised ML should typically be based on the model's suitability for answering the specific question at hand, despite a natural intuition dictates that model complexity equates to model superiority (Pisner & Schnyer, 2020).

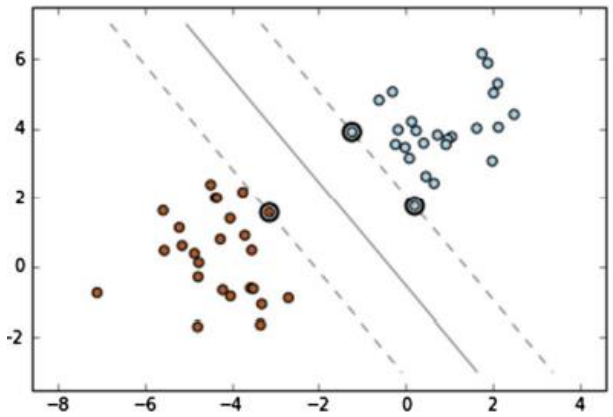


Fig. 4. An infinite number of classifiers can be drawn for the given data but SVM finds the classifier with the largest gap between support vectors. Circles represent the support vectors

When compared to other types of classifiers, SVM's strength and appeal largely come from its capacity to deliver balanced performance – high accuracies that are generalizable – even in situations where the dimensionality of the feature space significantly exceeds the number of observations available for training.

c) Decision tree

Decision trees represent one of the best-known and most-used techniques in classification (Quinlan, 1986). Their success is notably due to their ability to deal with complex classification problems. Indeed, they offer a representation that is easy to understand and interpret, as well as an ability to produce logical classification rules.

A decision tree is made up of:

- decision nodes each containing a test on an attribute,
- branches generally corresponding to one of the possible values of the selected attribute,
- sheets including objects that belong to the same class (Fig. 5).

The use of decision trees in classification problems is done in two main steps:

- the construction of a decision tree from a learning base,
- classification or inference consists in classifying a new instance from the decision tree built in the first step.

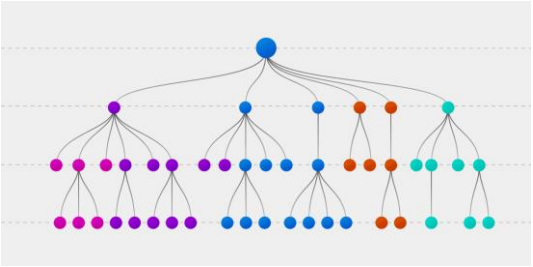


Fig. 5. Decision Tree Structure

With the use of algorithms like decision trees Iterative Dichotomiser 3 (ID3), C4.5, C5.0, and classification and regression trees (CART), problems relating to the classification, regression, clustering, and optimization can be resolved (Anuradha & Velmurugan, 2014). The ID3 and C4.5 algorithms were both created by Ross Quinlan; the latter is an enhanced variation of the former. Only categorical type features are applicable in ID3 decision trees; no numerical type features are allowed. The usage of information gain ratio (IGR) rather than information gain as in ID3 is one of the enhancements in C4.5. Second, pruning can be carried out both during and after tree building. Thirdly, C4.5 is capable of handling attributes with continuous features. Also can handle missing data (Adhatrao et al., 2013).

d) DNN

Multilayer neural networks are used in deep learning to design supervised and unsupervised learning mechanisms. In these mathematical architectures, each neuron performs simple calculations but the input data passes through several layers of calculation before producing an output. The results of the first layer of neurons are used as input for the calculation of the next layer and so on (Fig. 6). It is possible to play on the different parameters of the network architecture: the number of layers, the type of each layer, and the number of neurons that make up each layer (Bengio, 2009).

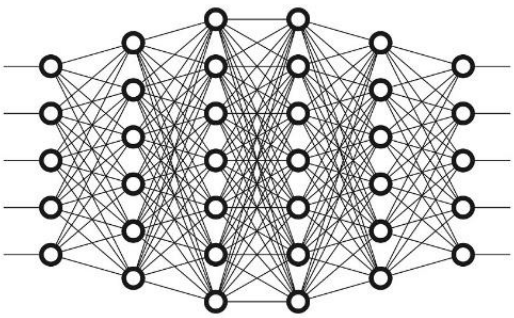


Fig. 6. Deep neural network structure

The multi-layer perceptron, capable of processing nonlinear phenomena, is an example of this type of network. Concretely, the first layers will make it possible to extract simple characteristics, which the following layers will combine to form increasingly complex and abstract concepts. Deep learning algorithms have mostly been used to improve computer

capabilities so that they can comprehend what humans can accomplish, including speech recognition. Since the advent of artificial intelligence five decades ago, speech in particular – the primary means of human communication – has attracted a lot of attention (Singh & Bathla, 2013; Anusuya & Katti, 2009). As a result, it comes as no surprise that speech was one of the first applications of deep learning. To date, a large number of research papers have been published on the use of deep learning in speech-related applications, notably speech recognition (Singh & Bathla, 2013; Anusuya & Katti, 2009; Nassif et al., 2019).

2.4. Dimensionality reduction

Dimension reduction is a process studied in mathematics and computer science. It consists of taking data in a high-dimensional space and replacing them with others in a lower-dimensional space, which still contains most of the information contained in the large set. In other words, we seek to construct fewer variables while retaining as much information as possible.

Principal component analysis, or PCA, is a dimensionality reduction method that is often used to transform a large set of variables into a smaller set that still contains most of the information in the large set. The idea is to transform correlated variables into new uncorrelated variables by projecting the data in the direction of increasing variance. The variables with the maximum variance will be chosen as the principal components. To do this, we must first find a new orthonormal basis in which we will represent our data, such that the variance of these data along these new axes is maximized.

2.5. Model evaluation

The dataset is divided into two groups. For each label, 80 percent of the data are utilized for training, while the remaining 20 percent are used for testing. For the training set, we have 63 positives and 66 negatives. As for the test set, we used 14 positives and 16 negatives. To improve our model's performance, we took all the MFCC extracted from the cough records. Meaning we have done the classification based on the frames of the signals, and the frames are labeled based on the original records. Resulting in 12684 observations, which were divided as follows:

- 7424 negatives (5940 training and 1484 test),
- 5260 positives (4208 training and 1052 test).

The acoustics features of the test set were extracted separately from the training set, so no test frame was used while forming our models, to avoid data leakage.

a) k-fold validation

We use the k-fold cross-validation method to perform cross-validation. In k-fold cross-validation, the input data is divided into k subsets of data (in this work we set k to 5). The machine learning model is trained on all but one subset (k-1), then evaluates the model on the subset that was not used for training. This process is repeated k times, with a different subset reserved for evaluation (and excluded from training) each time. After validating the models with the training set using 5-fold cross-validation, we tested the models with the test set, firstly by frames, then by subjects.

b) Confusion matrix

The confusion matrix is like a summary of the prediction results for a particular classification problem. It compares the actual data for a target variable to that predicted by a model. Correct and false predictions are revealed and distributed by class, which allows them to be compared with defined values.

Also known as a contingency table, the confusion matrix is used to evaluate the performance of a classification model. It shows how confusing a certain model can be when making predictions. In its simplest form, it is a 2×2 matrix. For more complex classification problems, it is always possible to add rows and columns to the basic form.

c) Evaluation metrics

Sensitivity is a measure of how many actual positives have been correctly classified as such (true positive rate, recall):

$$Sensitivity = \frac{TP}{TP + FN} \tag{1}$$

A measure of how many of the negative observations in the data are classified negative is called specificity (also known as true negative rate or selectivity):

$$Specificity = \frac{TN}{TN + FP} \tag{2}$$

Precision is the ratio of true positives to all the subjects predicted to be positive:

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

Accuracy is the proportion of true positives and true negatives, to all the subjects:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

F1 score: a recall and precision-based indicator of test accuracy:

$$F1\ score = \frac{(2 * TP)}{(2 * TP + FN + FP)} \tag{5}$$

The binary classification of a sample's expected and actual class are correlated and measured by the Matthews correlation coefficient (MCC). The scale of this coefficient is defined as: (+1) indicates an accurate prediction. A score of (-1) indicates that the prediction and result are completely inconsistent, whereas a score of (0) provides no useful information.

$$MCC = \frac{[(TP * TN) - (FP * FN)]}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{6}$$

The ROC curve is a graphical representation of the relationship between the sensitivity and specificity of a test for all possible cut-off values. The ordinate represents the sensitivity and the abscissa corresponds to the quantity (1 – specificity). The calculation of the numbers

VP, FP, VN, and FN makes it possible to deduce the sensitivity and specificity of the test for each value obtained. The couples {1 – specificity, sensitivity} are then placed on the curve. Joining them with straight lines leads to a stepped path connecting the lower left corner of the graph (Se = 0 and Sp = 1) to the upper right corner (Se = 1 and Sp = 0).

3. RESULTS

In this section, we present the results achieved by the models using the test set (30 observations/2536 frames):

- 16 negatives (1484 frames),
- 14 positives (1052 frames).

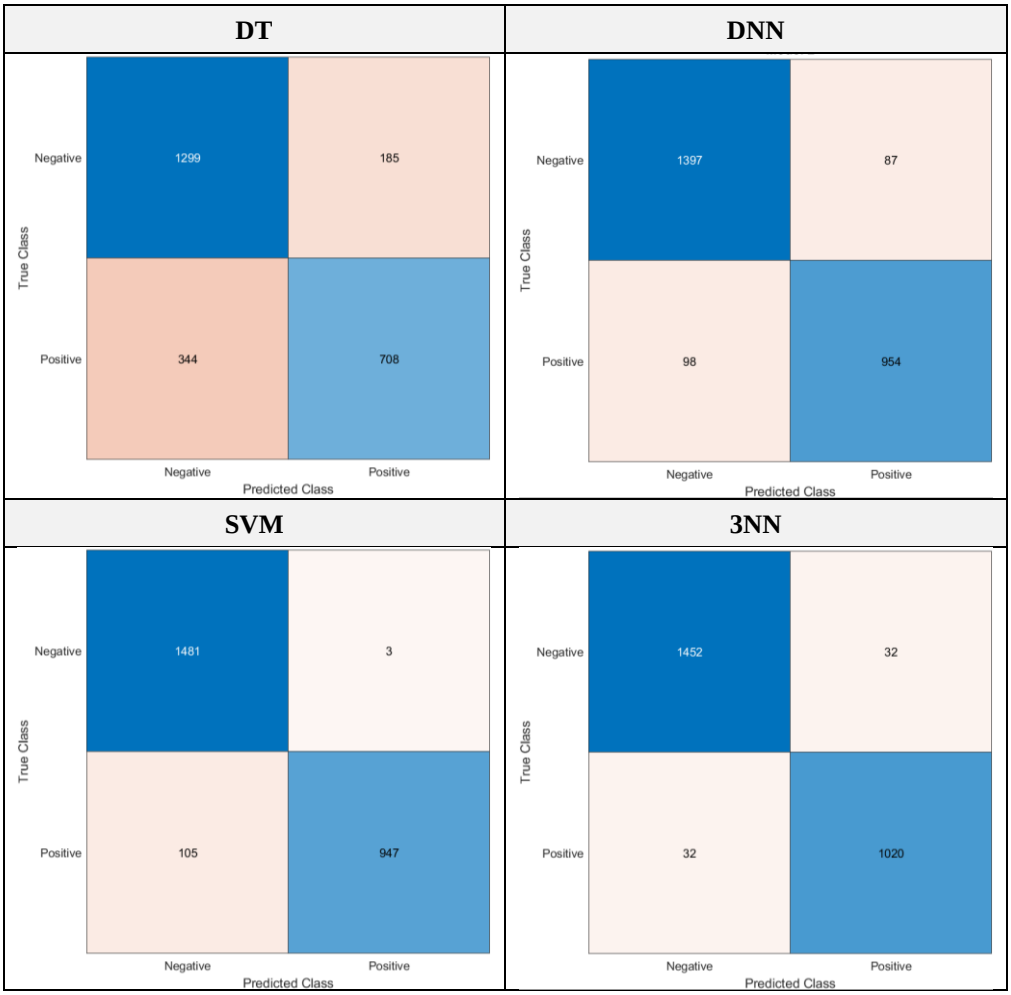
3.1. All features

Table 3 illustrates the confusion matrix obtained using all features. The 3NN, SVM, and DNN classifiers have achieved an excellent performance, they obtained an accuracy of 97.48 percent, 95.74 percent, and 92.7 percent, respectively, whereas the DT has 79.14 percent. The correlation between the sample's expected and actual class is calculated using the MCC; The 3NN and the SVM have an excellent prediction with 0.95 MCC for the 3NN and 0.91 for the SVM.

To evaluate the performances of the model in terms of false negatives and false positives, we calculated the sensitivity, specificity, and precision. With 99.8 percent specificity using SVM, and 96.96 percent for sensitivity using the 3NN, each model can detect the true positive of the respected class, but since in our application the false negatives could be fatal and may cause the spread of the virus, we believe the 3NN has the edge with 96.96 percent versus 90.02 percent achieved by SVM. The false alarms are quantified by the precision. The model with fewer false positives is the SVM with nearly perfect 99.68 percent precision. The F1-score values for each model were calculated to evaluate the overall performances. The 3NN has the best f1-score of 96.96 percent. For this model we have:

- for the negatives, the model has correctly classified 1452 observations from 1484, with 32 false positives,
- as for the class of the positives, 1020 observations have been detected, and 32 were misclassified and considered negatives.

Tab. 3. Confusion matrices using all features

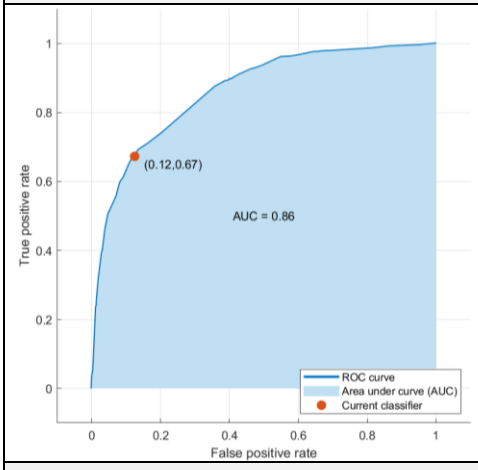
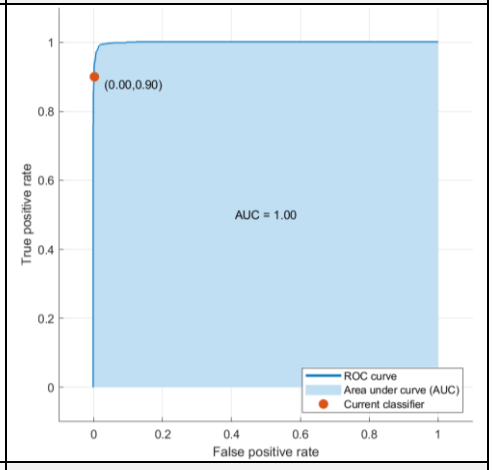
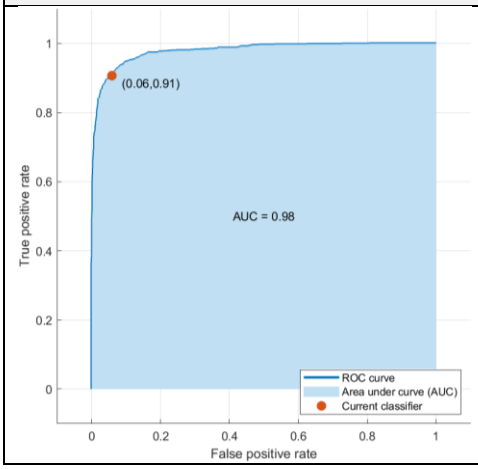
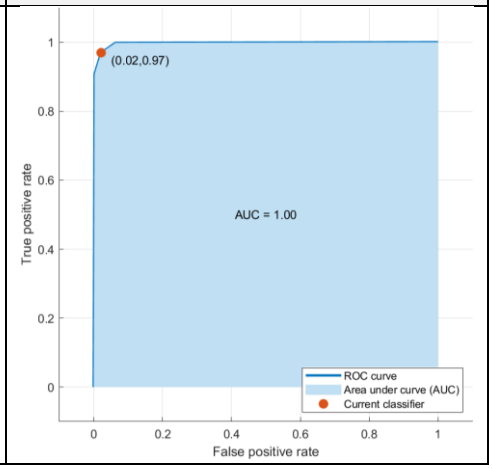


Tab. 4 summarizes the evaluation metrics, and the ROC curves are illustrated in Tab.5 for all the classifiers.

Tab. 4. Evaluation metrics using all features

	Sensitivity (%)	**Specificity (%)**	**Precision (%)**	**Accuracy (%)**	**F1 score (%)**	**MCC**
DT	67.3	87.53	79.28	79.14	72.8	0.56
SVM	90.02	**99.8**	**99.68**	95.74	94.6	0.91
DNN	90.68	94.14	91.64	92.7	91.16	0.85
3NN	**96.96**	97.84	96.96	**97.48**	**96.96**	**0.95**

Tab. 5. ROC curve using all features

DT	SVM
	
DNN	**3NN**
	

3.2. PCA with 95 percent of the variance

We applied the PCA to explain the 95 percent variance. After training, 11 components were kept. Fig. 7 shows the variance per component.

When performing dimensionality reduction by the PCA with 95 percent variance, the 3NN and SVM have very close accuracy (96.92 percent using 3NN and 96.6 percent using the SVM), followed by DNN with an accuracy of 91.09 percent and the DT with 78.08 percent. The SVM has a slight improvement in the MCC with 0.94, whereas the 3NN has 0.95.

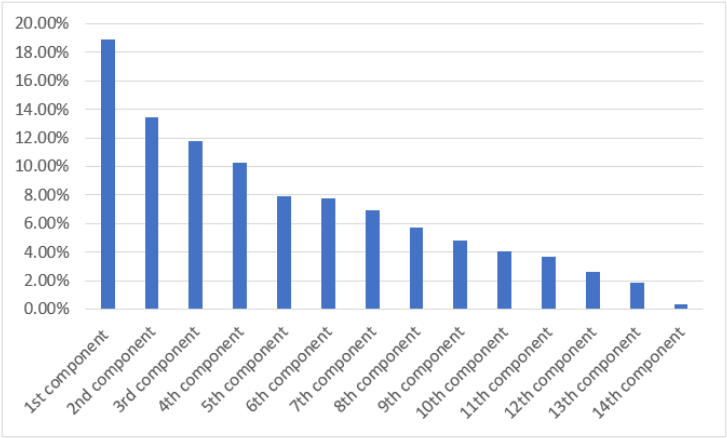
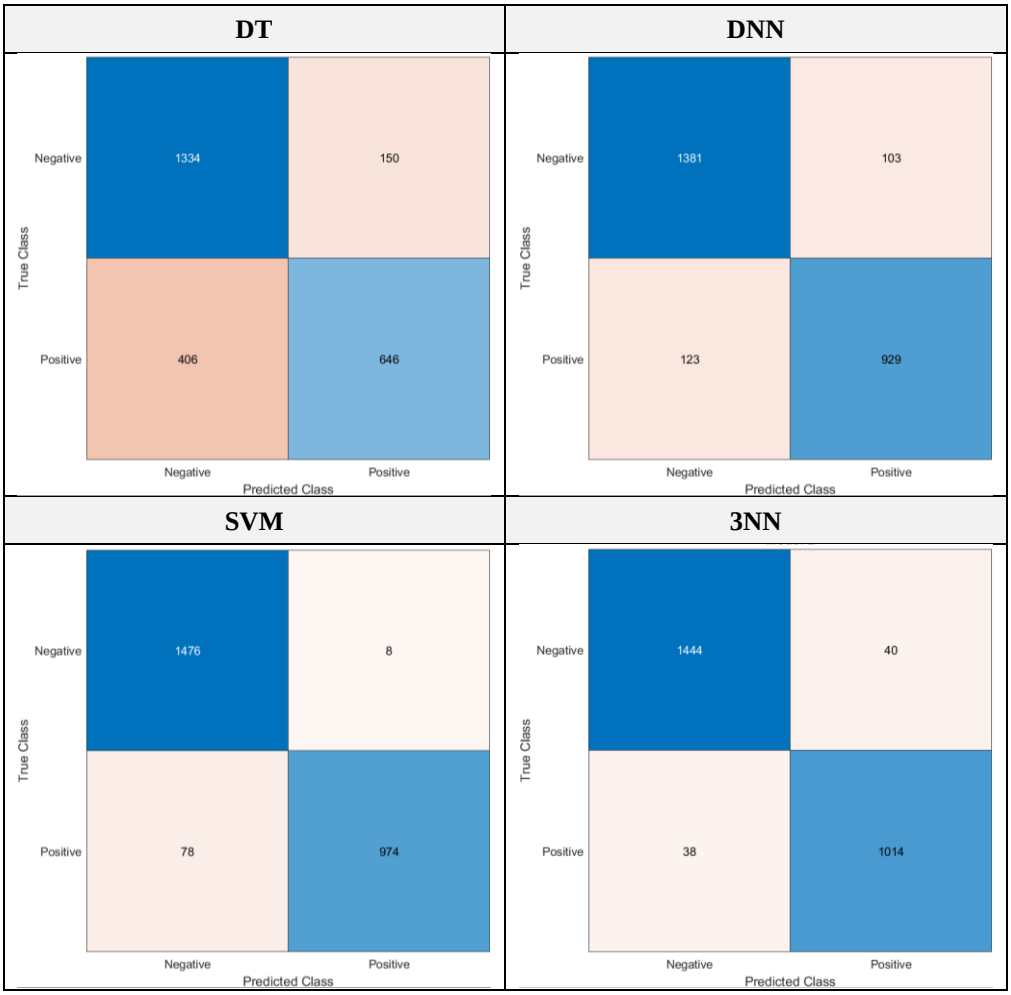


Fig. 7. Variance per component

Tab. 6. Confusion matrices using PCA with 95 percent of the variance



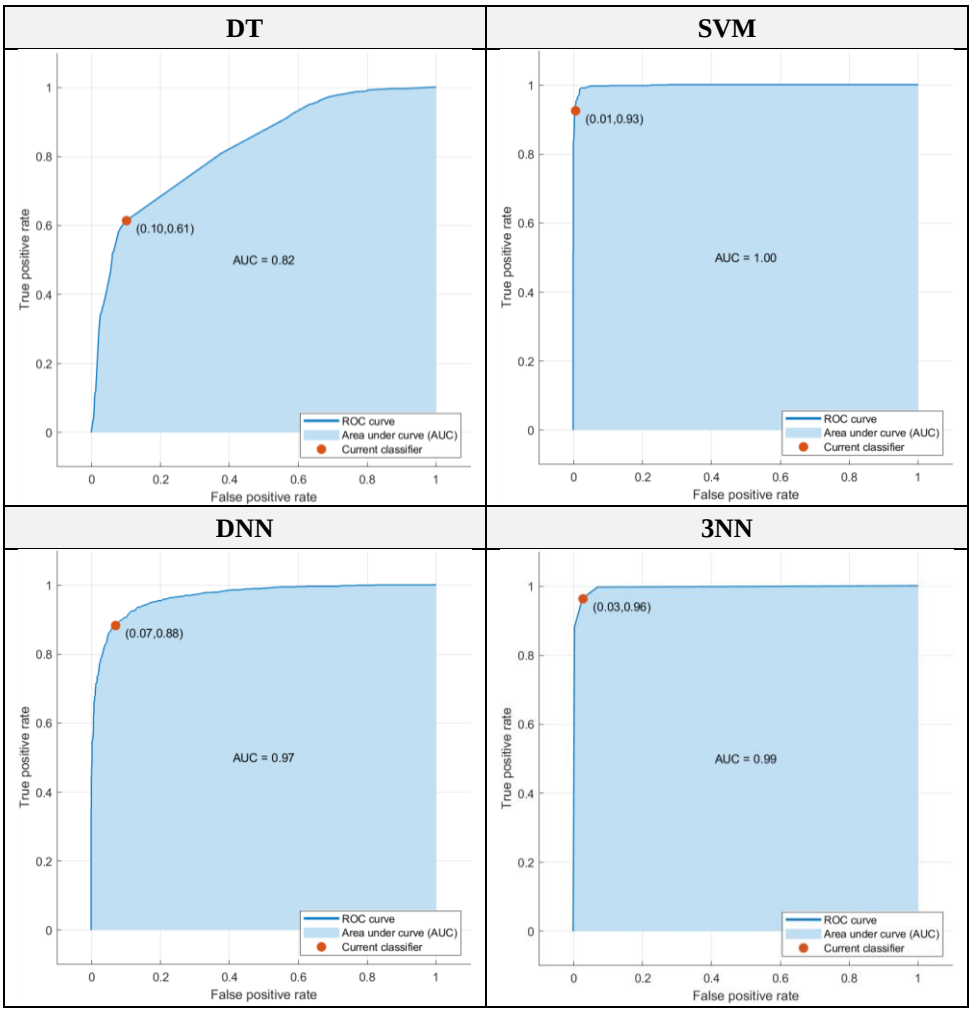
The F1-score values for each technique were calculated to evaluate the overall performances (Tab. 6). The 3NN has the highest f1-score of 97.3 percent. For this model we have:

- for the negatives, the model has correctly classified 1444 observations, with 40 false positives,
- in the class of the positives, 1014 observations have been identified, whereas 38 were misclassified.

Tab. 7. Evaluation metrics using PCA with 95 percent of the variance

	Sensitivity (%)	Specificity (%)	Precision (%)	Accuracy (%)	F1 score (%)	MCC
DT	61.41	89.9	81.16	78.08	69.91	0.54
SVM	92.59	**99.46**	**99.19**	96.6	95.77	0.93
DNN	88.3	93.06	90.02	91.09	89.16	0.82
3NN	**96.39**	97.3	96.2	**96.92**	**96.3**	**0.94**

Tab. 8. ROC curve using PCA with 95 percent of the variance

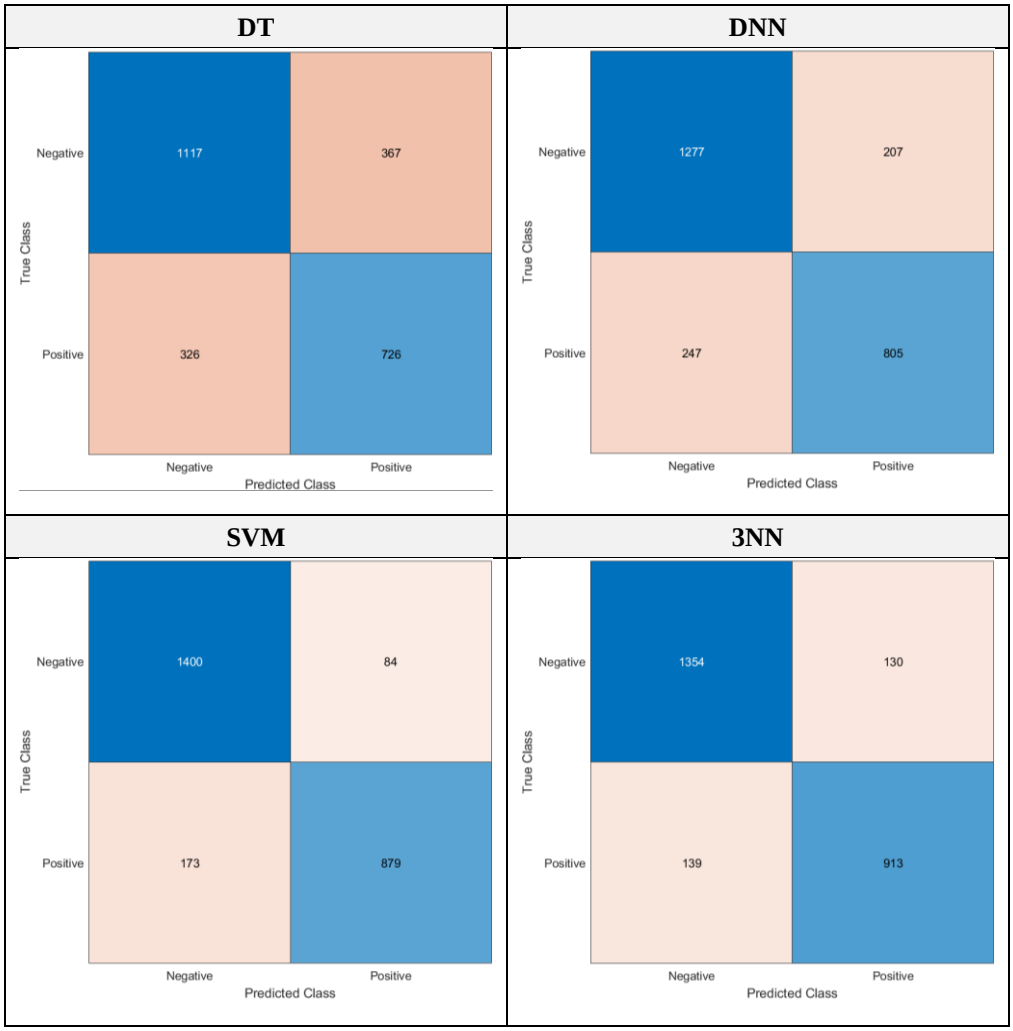


The highest specificity and precision were achieved by the SVM above 99 percent. While the 3NN has a sensitivity and f1-score above 96 percent. The rest of the results are listed in Tab. 7. The ROC curves of each model are presented in Tab.8.

3.3. PCA with 6 components

We have selected the first six principal components calculated, for 70 percent variance. The classification results are given in Tab. 9.

Tab. 9. Confusion matrices using PCA with 6 principal components



Reducing the space features by mapping only 6 principal components, the SVM has a satisfactory accuracy of 89.87 percent, and an F1-score of 87.25 percent, close to the 3NN (89.4 accuracies and 87.16 F1-score). With an MCC of 0.79 for the first and 0.78 for the latter. The SVM has:

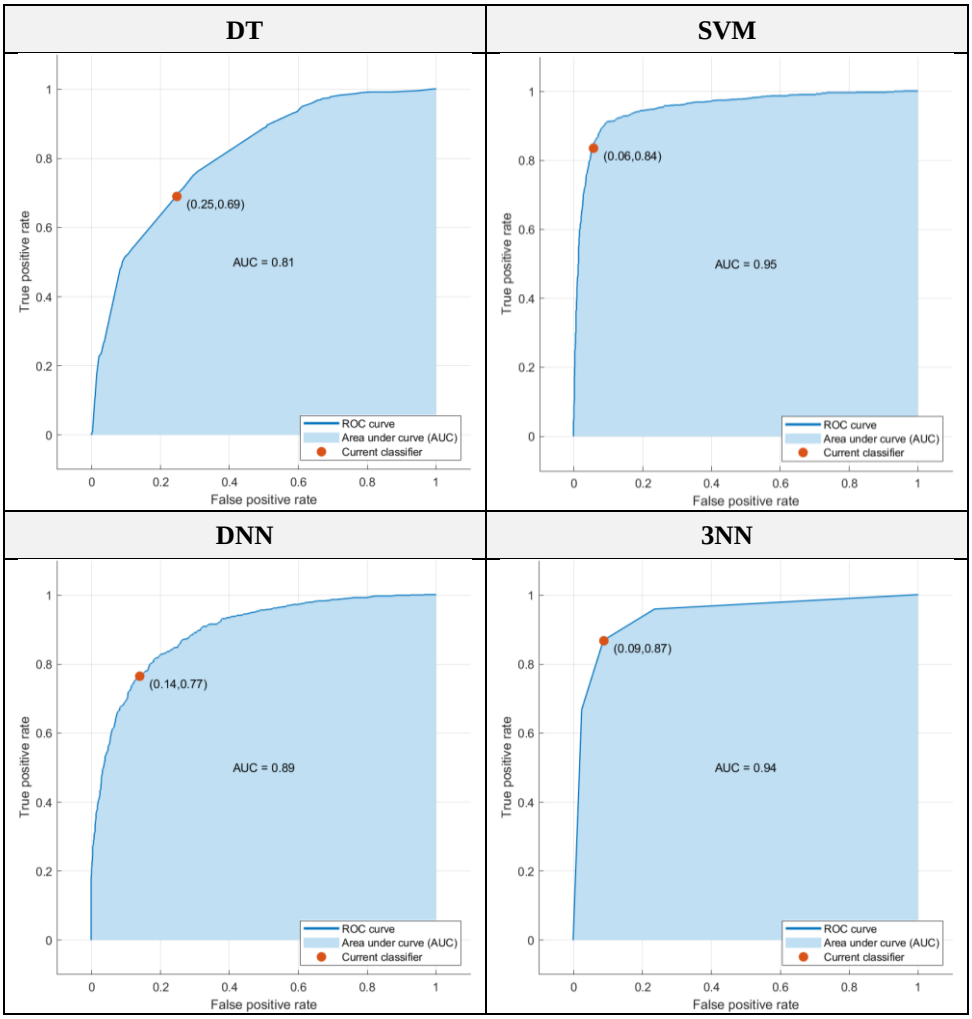
– 1400 observations negatives have been correctly classified, with 84 false positives,
– and 879 positive subjects have been detected, and 173 misclassified.

The highest sensitivity was achieved by the 3NN (86.79 percent). While the SVM has better specificity and precision of 94.34 percent and 91.28 percent, respectively. The evaluation metrics are presented in Tab. 10, and the ROC curves are in Tab. 11.

Tab. 10. Evaluation metrics using PCA with 6 principal components

	Sensitivity (%)	Specificity (%)	Precision (%)	Accuracy (%)	F1 score (%)	MCC
DT	69.01	75.27	66.42	72.67	67.69	0.44
SVM	83.56	**94.34**	**91.28**	**89.87**	**87.25**	**0.79**
DNN	76.52	86.05	79.55	82.1	78	0.63
3NN	**86.79**	91.24	87.54	89.4	87.16	0.78

Tab. 11. ROC curve using PCA with 6 principal components



After testing our models for each frame of the signals, we performed the classification by subjects, using the 3NN with all features. The results show how accurate the model is in predicting the subjects based on the most frequent value predicted by frame for each subject (Fig. 8).

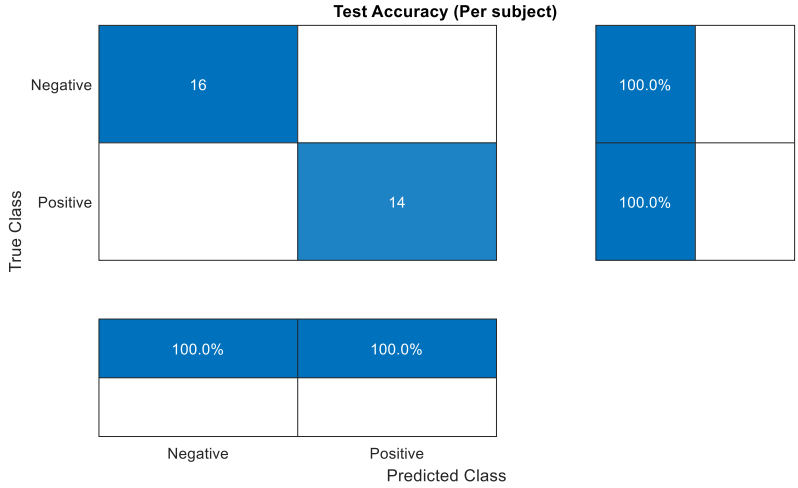


Fig. 8. confusion matrix by subject using 3NN

4. DISCUSSION

The 3NN classifier with all features has produced the best classification results. The model detects COVID-19 patients with an accuracy of 97.48 percent, 96.96 percent f1-score, and 0.95 MCC. This suggests that our approach can accurately distinguish healthy controls from COVID-19 patients. The importance of the AI-based screening method is limiting the spread of the virus, meaning building models with high sensitivity of the positive cases, in our case we were able to achieve 96.96 percent. The specificity of 97.84 percent, and the 96.96 percent precision. Performing PCA with 95 percent variance also gives an accurate prediction for the 3NN and the SVM classifiers. When choosing only 6 components the performance dropped but was still acceptable since the SVM and the 3NN obtained more than 89 percent accuracy, 87 percent f1-score, and 0.78 MCC.

We believe that when compared to previous studies, our findings are extremely encouraging (Tab. 12). Using less computationally expensive classifiers and focusing more on data augmentation through frame analysis, we were able to get excellent results.

Tab. 12. Comparison with other earlier studies for the COVID-19 sound-based diagnosis

Research	Dataset	Sound type	Features	Models /Classifiers	Results (Accuracy)
Coppock et al., 2021	Covid-19-sounds	Cough Breathing	Mel-spectrogram	ResNet	84.6%
Aly, Rahouma & Ramzy, 2022	Coswara Virufy	Coswara Virufy	MFCC – RMS - ZCR Spectral Rolloff/ Centroid/etc.	Deep Model Shallow classifiers	96.4%
Fakhry et al., 2021	Coughvid	Cough	Mel-spectrogram MFCC Clinical features	Multi-Branch Deep Learning Network	91%
Chaudhari et al., 2020	Coswara Coughvid Virufy	Cough	Mel-spectrogram MFCC Clinical features	Ensemble Deep Learning Model	77.1%
Laguarta, Hueto & Subirana, 2020	Opensigma	Cough	MFCC - other biomarkers	ResNet50	97%
Brown et al., 2020	Covid-19-sounds	Cough Breathing	MFCC - Tempo - RMS - ZCR - etc.	VGGish Shallow classifiers	80%
Pahar et al., 2021	Coswara SARCOS	Cough	MFCC - Log Energies - ZCR - Kurtosis.	Resnet50 LSTM	98% 94%
This study	Coswara	Cough	MFCC	3NN	97.48%

These results are encouraging and a controlled clinical trial by medical specialists is required to validate these findings. Furthermore, due to the pandemic's fast and current spread, there is still a scarcity of knowledge about the disease's etiology and progression, as well as the relationship between demographic and clinical data of COVID-19 patients. We focused solely on the effects of COVID-19 infection on voice quality in this study. However, in the future, we hope to investigate the effects of patient data, such as age and gender, the etiopathogenesis of the pandemic, whose symptoms, particularly in the early stages of the disease, are still frequently misinterpreted as other respiratory illnesses, and to identify COVID-19 disorders and enhance the model's accuracy. One of the biggest problems with COVID-19 is the lack of high-quality datasets. because of: (1) closed-source and unpublished datasets; (2) the scattered nature of COVID-19 datasets; and (3) privacy concerns that restrict data sharing. Therefore, cooperation amongst all medical organizations worldwide is necessary to increase the availability and quality of COVID-19 datasets. A link between COVID-19 infection and various medical comorbidities has been observed in the literature. Therefore, the COVID-19 prediction and detection processes must both take into account a patient's history of various illnesses (such as diabetes, liver, renal, heart disease, etc.) in order to create a precise and reliable prediction model. In contrast to working with IoT devices, building complicated ML models, analyzing, and interpreting massive data demand high computational resources. Edge computing and fog computing may therefore be useful in addressing this issue.

5. CONCLUSION

In this paper, we proposed a cough-based detection of the presence of COVID-19 using the main Machine learning classifiers and the PCA. Our objective is to realize a system capable of screening COVID-19 disorder, which could be useful as a pre-screening test as well as for the monitoring of patients' symptoms. The proposed models have been applied to the Coswara database, a crowdsourcing project from the Indian Institute of Science aiming to build a diagnostic tool for COVID-19 using audio recordings. The results have shown that the best accuracy in COVID-19 detection is achieved by mapping all the extracted features directly to the 3NN classifier.

Conflicts of Interest

Funding: Not applicable

Conflicts of interest/Competing interests: On behalf of all authors, the corresponding author states that there is no conflict of interest.

REFERENCES

Adhatrao, K., Gaykar, A., Dhawan, A., Jha, R., & Honrao, V. (2013). *Predicting students' performance using ID3 and C4. 5 classification algorithms*. arXiv preprint arXiv:1310.2071.

Ai, T., Yang, Z., Hou, H., Zhan, C., Chen, C., Lv, W., Tao, Q., Sun, Z., & Xia, L. (2020). Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. *Radiology*, *296*(2), E32-E40. https://doi.org/10.1148/radiol.2020200642

Aly, M., Rahouma, K. H., & Ramzy, S. M. (2022). Pay attention to the speech: COVID-19 diagnosis using machine learning and crowdsourced respiratory and speech recordings. *Alexandria Engineering Journal*, *61*(5), 3487–3500. https://doi.org/10.1016/j.aej.2021.08.070

Anuradha, C., & Velmurugan, T. (2014). A data mining based survey on student performance evaluation system. In *2014 IEEE International Conference on Computational Intelligence and Computing Research* (pp. 1–4). IEEE. https://doi.org/10.1109/ICCIC.2014.7238389

Anusuya, M. A., & Katti, S. K. (2010). *Speech recognition by machine, a review*. arXiv preprint arXiv:1001.2267.

Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and trends® in Machine Learning*, *2*(1), 1–127.

Benmalek, E., Elmhamdi, J., & Jilbab, A. (2021). Comparing CT scan and chest X-ray imaging for COVID-19 diagnosis. *Biomedical Engineering Advances*, 1, 100003. https://doi.org/10.1016/j.bea.2021.100003

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory* (pp. 144–152). The ACM Digital Library.

Brown, C., Chauhan, J., Grammenos, A., Han, J., Hasthanasombat, A., Spathis, D., Xia, T., Cicuta, P., & Mascolo, C. (2020). *Exploring automatic diagnosis of COVID-19 from crowdsourced respiratory sound data*. arXiv preprint arXiv:2006.05919.

Chaudhari, G., Jiang, X., Fakhry, A., Han, A., Xiao, J., Shen, S., & Khanzada, A. (2020). *Virufy: Global applicability of crowdsourced and clinical datasets for AI detection of COVID-19 from cough*. arXiv preprint arXiv:2011.13320.

Coppock, H., Gaskell, A., Tzirakis, P., Baird, A., Jones, L., & Schuller, B. (2021). End-to-end convolutional neural network enables COVID-19 detection from breath and cough audio: a pilot study. *BMJ innovations*, *7*(2), 356–362. https://doi.org/10.1136/bmjinnov-2021-000668

Fakhry, A., Jiang, X., Xiao, J., Chaudhari, G., Han, A., & Khanzada, A. (2021). Virufy*: A multi-branch deep learning network for automated detection of COVID-19*. arXiv preprint arXiv:2103.01806.

Han, J., Brown, C., Chauhan, J., Grammenos, A., Hasthanasombat, A., Spathis, D., Xia, T., Cicuta, P., & Mascolo, C. (2021). Exploring Automatic COVID-19 Diagnosis via voice and symptoms from Crowdsourced Data. In *ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 8328–8332). IEEE.

Han, W., Chan, C. F., Choy, C. S., & Pun, K. P. (2006). An efficient MFCC extraction method in speech recognition. In *2006 IEEE International Symposium on Circuits and Systems* (ISCAS) (pp. 4). IEEE.

Indyk, P., & Motwani, R. (1998). Approximate nearest neighbors: towards removing the curse of dimensionality. In *Proceedings of the thirtieth annual ACM symposium on Theory of computing* (pp. 604–613). The ACM Digital Library.

Ismail, M. A., Deshmukh, S., & Singh, R. (2021). Detection of COVID-19 through the analysis of vocal fold oscillations. In *ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 1035–1039). IEEE.

Laguarta, J., Hueto, F., & Subirana, B. (2020). COVID-19 Artificial Intelligence Diagnosis using only Cough Recordings. In *IEEE Open Journal of Engineering in Medicine and Biology* (vol. 1, 275–281). IEEE. https://doi.org/10.1109/OJEMB.2020.3026928

Li, L., Qin, L., Xu, Z., Yin, Y., Wang, X., Kong, B., Bai, J., Lu, Y., Fang, Z., Song, Q., Cao, K., Liu, D., Wang, G., Xu, Q., Fang, X., Zhang, S., Xia, J., & Xia, J. (2020). Using Artificial Intelligence to Detect COVID-19 and Community-acquired Pneumonia Based on Pulmonary CT: Evaluation of the Diagnostic Accuracy. *Radiology*, *296*(2), E65–E71. https://doi.org/10.1148/radiol.2020200905

Muda, L., Begam, M., & Elamvazuthi, I. (2010). *Voice recognition algorithms using mel frequency cepstral coefficient (MFCC) and dynamic time warping (DTW) techniques*. arXiv preprint arXiv:1003.4083.

Nassif, A. B., Shahin, I., Attili, I., Azzeh, M., & Shaalan, K. (2019). Speech recognition using deep neural networks: A systematic review. *IEEE access*, *7*, 19143–19165. https://doi.org/10.1109/ACCESS.2019.2896880

Pahar, M., Klopper, M., Warren, R., & Niesler, T. (2021). COVID-19 cough classification using machine learning and global smartphone recordings. *Computers in Biology and Medicine*, *135*, 104572. https://doi.org/10.1016/j.compbiomed.2021.104572

Pal, A., & Sankarasubbu, M. (2021). Pay attention to the cough: Early diagnosis of COVID-19 using interpretable symptoms embeddings with cough sound signal processing. In *Proceedings of the 36th Annual ACM Symposium on Applied Computing* (pp. 620–628). The ACM Digital Library. https://doi.org/10.1145/3412841.3441943

Pisner, D. A., & Schnyer, D. M. (2020). Support vector machine.In *Machine learning* (pp. 101–121). Academic Press.

Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, *1*(1), 81–106.

Sharma, N., Krishnan, P., Kumar, R., Ramoji, S., Chetupalli, S. R., Ghosh, P. K., & Ganapathy, S. (2020). *Coswara--a database of breathing, cough, and voice sounds for COVID-19 diagnosis*. arXiv preprint arXiv:2005.10548.

Singh, H., & Bathla, A. K. (2013). A survey on speech recognition. *International Journal of Advanced Research in Computer Engineering & Technology*, *2*(6), 2186–2189.

Weng, L. M., Su, X., & Wang, X. Q. (2021). Pain symptoms in patients with coronavirus disease (COVID-19): a literature review. *Journal of Pain Research*, *14*, 147. https://doi.org/10.2147/JPR.S269206

Wu, X., Hui, H., Niu, M., Li, L., Wang, L., He, B., Yang, X., Li, L. Li, H., Tian, J., & Zha, Y. (2020). Deep learning-based multi-view fusion model for screening 2019 novel coronavirus pneumonia: a multicentre study. *European Journal of Radiology*, *128*, 109041. https://doi.org/10.1016/j.ejrad.2020.109041

Yang, Y., Yang, M., Shen, C., Wang, F., Yuan, J., Li, J., Zhang, M., Wang, Z., Xing, L. Wei, J., Peng, L., Wong, G., Zheng, H., Wu, W., Liao, M., Feng, K., Li, J., Yang, Q., Zhao, J., Zhang, Z., Liu, L., & Liu, Y. (2020). *Evaluating the accuracy of different respiratory specimens in the laboratory diagnosis and monitoring the viral shedding of 2019-nCoV infections*. MedRxiv. https://doi.org/10.1101/2020.02.11.20021493

Zheng, F., Zhang, G., & Song, Z. (2001). Comparison of different implementations of MFCC. *Journal of Computer science and Technology*, *16*(6), 582–589. https://doi.org/10.1007/BF02943243

Applied Computer Science, vol. 18, no. 4, pp. 116–129
doi: 10.35784/acs-2022-32

Submitted: 2022-11-04 | Revised: 2022-12-17 | Accepted: 2022-12-27

Keywords: accessability, efficiency, production process, maintenance, AHP, Fuzzy AHP

Piotr WITTBRODT [0000-0002-6377-4252]*, *Iwona ŁAPUŃKA* [0000-0003-1472-1477] **, *Gulzhan BAYTIKENOVA*[***], *Arkadiusz GOLA* [0000-0002-2935-5003]****, *Alfiya ZAKIMOVA*[*****]

IDENTIFICATION OF THE IMPACT OF THE AVAILABILITY FACTOR ON THE EFFICIENCY OF PRODUCTION PROCESSES USING THE AHP AND FUZZY AHP METHODS

Abstract

Maintenance has a key impact on the efficiency of the production processes because the efficiency of the machines determines the ability of the system to produce in accordance with the assumed schedule. The key element of the system performance assessment remains the availability of technological equipment, which directly translates into the efficiency and effectiveness of the performed production tasks. Taking into account the dynamic nature of manufacturing processes, the proper selection of machinery and equipment for the implementation of specific production tasks becomes an issue of particular importance. The purpose of this research was to determine the impact of technical and non-technical factors on the material selection of machine tools for production tasks and to develop a method of supporting the selection of production resources using the AHP and Fuzzy AHP methods. The research was carried out in a manufacturing company from the automotive industry.

[*] Department of Management and Production Engineering, Faculty of Production Engineering and Logistics, Opole University of Technology,Poland, p.wittbrodt@po.edu.pl
[**] Department of Management and Production Engineering, Faculty of Production Engineering and Logistics, Opole University of Technology, Poland, i.lapunka@po.edu.pl
[***] School of Business and Entrepreneurship, D. Serikbayev East Kazakhstan Technical University, Kazakhstan, gbaytikenova@mail.ru
[****] Department of Production Computerisation and Robotisation, Faculty of Mechanical Engineering, Lublin University of Technology, Poland, a.gola@pollub.pl
[*****] School of Business and Entrepreneurship, D. Serikbayev East Kazakhstan Technical University, Kazakhstan, azakimova@edu.ektu.kz

1. INTRODUCTION

Current market requirements result in an increase in demand for personalized products with unique features and functionalities tailored to specific customer needs (Zubrzycki et al., 2021; Relich & Świć, 2020). Providing a wide range of products is a huge challenge for manufacturing companies, which determines the nature of the organization's activities and functioning, and above all, the type and structure of the manufacturing system (Gola, 2014). Appropriate preparation of this type of production in terms of construction, technology or organization is a significant challenge – moreover, it is not easy to control and manage such type of production (Pizon, Kulisz & Lipski, 2021; Szabelski, Karpiński & Machrowska, 2022).

In general, the implementation of production activities is carried out in two ways: in steady-state and transient conditions (Burduk et al., 2019). In the first case, it is possible to determine future production levels with high probability, and thus it is possible to develop scenarios of remedial actions for potential threats to production (Bałdowska-Witos et al., 2020). In the second case, when working conditions are unstable, it is impossible to develop action scenarios due to the variety of disruptive factors, e.g. damage to equipment and machines (Varela et al., 2018). Currently, due to changing market requirements, the second type of production dominates in enterprises. The reasons for this situation are: (1) a customer who expects personalized products - requires focusing the company's attention on trends in society, responding to the customer's wishes, suggestions and comments, and time pressure - producing the product in the shortest possible time (Swic & Gola, 2013). Meeting the requirements is an indicator of the existence of the company on the market (Vollman et al., 2005; Madu, 2000). The lack of repeatability of the series results in the lack of organizational patterns of the production structure and the accumulation of errors in its organization and management. Therefore, solutions supporting and eliminating these disturbances should be sought (Rakyta et al., 2015). Unfortunately, in enterprises, when considering and trying to prevent faults in unstable production conditions, the aspect of maintenance is often overlooked (Kosicka, Gola & Pawlak, 2019). The role it plays in the enterprise is important, because with the proper functioning of maintenance services, it is possible to both maximize the use of available machine resources and determine their actual states (technical diagnostics of objects), as well as to obtain new data that can be input data, supplementing control systems and production management (advisory systems, expert systems). Supporting production with such systems makes flexible production have the characteristics of mass production, stable and predictable.

The implementation of modern machinery and equipment in the production system of the company (implementation of technical solutions, e.g. automation and robotization, flexible production systems) as well as the introduction of various production philosophies (e.g. Just-in-Time (JiT) or Lean Manufacturing) improved the use of available resources. Despite these activities, the area of maintaining the machine park in constant readiness leaves you unsatisfied, because technical equipment is of key importance for the efficient and effective functioning of production processes and quick order fulfillment (Crespo Márquez et al., 2009; Al-Najjar, 2007).

The purpose of this article is to analyze the impact of technical and non-technical indicators on the material selection of technological equipment for specific production purposes and to present the method of supporting decisions in the selection of resources for the needs of production tasks using the AHP and Fuzzy AHP methods.

2. THE IMPORTANCE OF MAINTENANCE IN A MANUFACTURING COMPANY

The growing pressure to shorten production cycles means that the readiness to work of machines and technical devices remains the key factor determining the possibility of timely execution of orders, which consequently increases the role and importance of maintenance services (Aspinwall & Elgharib, 2013). Therefore, maintenance has an increasing impact on the efficiency and profitability of the company. Production lost due to unplanned machine downtime will never be recovered without additional cost. Disruptions in production processes caused by a failure of a machine or device not only reduce productivity and increase the cost of the product, but also cause the loss of the ability to produce products on time, which in turn translates into the loss of the company's image.

The importance of maintenance is growing due to its role in the maintenance and availability of the machine park and the efficiency of its operation. Traditionally, the approach to maintenance is considered to be an area of additional costs for the enterprise, however, research shows the positive impact of maintenance on the efficiency of the enterprise, its profitability and productivity (Al-Najjar & Algabroun, 2018; Azizi, 2015; Maletic et al., 2014).

It is believed that by applying an effective maintenance policy, production deficiencies and failures can be reduced to a minimum level, the economic indicator of which is acceptable to the company. In the perspective of long-term goals, this type of action can bring significant savings to the organization (Al-Najjar, 2007).

Maintenance is a strategy that involves the identification, investigation, and implementation of many repair, replacement, and inspection decisions. In the initial scope of its operation, it was based on a reactive approach, where no action is taken to prevent failures or detect the beginnings of failures, and the device must be repaired when it fails. It was assumed that the costs related to traffic maintenance are high in relation to its functioning. When mechanization elements were introduced to the industry, the approach to maintenance changed to a preventive approach. Inspections of machines and devices were carried out at specific time intervals (e.g. number of shifts) or other criteria aimed at reducing the probability of failure or deterioration of efficiency (e.g. number of manufactured products). In the following years, when the share of automated and robotic processes increased significantly, the maintenance approach changed to a predictive approach. A number of IT systems for maintenance management based, for example, on monitoring the technical condition of the facility with the use of vibroacoustic signals, have been developed. Currently, due to globalization, dedicated IT systems are not sufficient in the effective use of maintenance. Supporting solutions were and are still being sought, consisting in combining the maintenance department and other departments necessary for the proper functioning of the company (e.g. supply departments, production optimization, organization and management of the company). This approach is called a process-oriented approach. Although different approaches have been developed at different times, in practice more than one approach can be used at the same time, as well as many 'customized maintenance strategies'. Therefore, from the point of view of the user of these systems, it is very important to know which of the available approaches to maintenance is the most cost-effective and corresponds to the technical system in the company's operational activities. The strategies discussed can be found in many scientific publications, and various maintenance concepts

have been developed on their basis, e.g.: Reliability Centered Maintenance (RCM), Total Productive Maintenance (TPM), Condition-Based Maintenance (CBM) and Integrated Logistics Support (ILS) (Sagar & Singh, 2012; Ahmad & Kamaruddin, 2012).

Systemically, the maintenance process is currently perceived as supporting the main processes implemented in the company - mainly production and its cost. However, the complexity of modern production systems and their dependence on a large number of internal and external factors forces a change in the approach to maintenance to a process-oriented approach (Blanchard, 2004). Identification of these factors and determining their importance is one of the basic activities enabling the construction of a model for assessing the added value through the maintenance of the system, organization and management of a production company (Al-Najjar, 2007).

3. ESSENCE AND CHARACTERISTICS OF AHP AND FUZZY AHP METHODS

Analytic Hierarchy Process (AHP) is an effective decision support tool. Using the AHP method, a problem can be solved in a hierarchical manner, where the decision is based on criteria (usually multiple criteria) (Saaty, 1980). The AHP method is based on the concept of a hierarchy of goals and creating binary comparisons between goals of the same level (determining the ranking of the analyzed solutions). The construction of the hierarchy diagram strictly depends on the type of problem under study and maps the hierarchy of goals of this problem. If the problem is complex and contains a number of alternative paths, the structure is more complex, which makes solving the problem more difficult. Most often, a simplified structure is sought, which contains, for example, three levels, where the first level contains only one element, the decision problem. The second level consists of elements - decision criteria, and the third level contains the characteristics of the tested object. A typical, hierarchical structure of the AHP method is shown in Fig. 1.

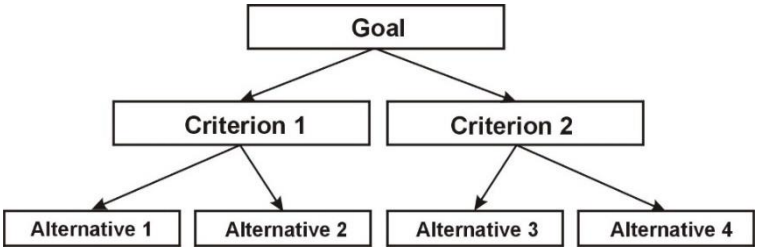


Fig. 1. Typical hierarchy structure of the Analytic Hierarchy Process method

In mathematical terms, the record of a decision problem presents the so-called decision matrix (1).

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{1n} \\ x_{21} & x_{22} & x_{2n} \\ x_{m1} & x_{m2} & x_{mn} \end{bmatrix} \tag{1}$$

where: X – action, strategy, alternative decision or decision,
x_{mn} – the measure of the variant W_m according to the K_n criterion,
m – value of m variant,
n – value of n variant.

Decision matrix is a description of individual variants along with criteria describing those variants to which weight should be assigned (usually in the form of a numerical value). The numerical value of the criterion is a measure of the implementation of the assumed tasks and objectives of individual variants. The scale of the point value is ambiguous and a point scale from 0 to 5 is adopted in simplified tasks, and from 0 to 10 in exact solutions. The assumption of the point criterion is that the minimum value is the worst grade, while the maximum value is the best grade. For example, a comparative rating of 1 might mean "equal", a value of 3 for "slightly greater", a value of 9 for "extremely greater". To solve a decision problem using the AHP method, four steps must be taken. It is necessary to specify the decision problem with its description (hierarchy of the problem), make a comparison in pairs in relation to the criteria, and criteria in relation to the goal (evaluation of the criteria). Then set preferences with regard to the priority of criteria and decision variants and choose the best variant (analysis of results). As a result of the analysis, we obtain the highest ranking value, which is considered to be a solution to the decision problem and is a compromise between different goals or criteria (Bayzit, 2005).

The assessment of the solution made by the decision-maker is ambiguous, depending on his attitude to the task, personal characteristics, function performed, knowledge, data and skills. The result of such an approach is a multifaceted approach to the decision-making problem. The AHP method has a wide and varied application, and the availability of examples in the literature is considerable (Das & Chattopadhyay, 2003).

The Fuzzy Analytic Hierarchy Process (FAHP) method is similar in terms of methodology. Unlike the classical AHP method, it enables a more accurate assessment of linguistic criteria and is similar to human reasoning. It is characterized by the use of fuzzy sets in relation to linguistic criteria and membership functions. Membership functions can take various forms, e.g. triangular or trapezoidal, but in practical applications the first one is most often used – the triangular membership function. The fuzzy values correspond to the start, middle, and end of the triangle, respectively. The fuzzy triangle (TNF) scores for the FAHP method are presented in Table 1.

Tab. 1. Values of fuzzy evaluation triangles (Kutlu, 2012)

Linguistic assessment	Fuzzy value TNF	AHP classic equivalent
Absolute preference	(2; 5/2 ;3)	9
Very clear preference	(3/2; 2; 5/2)	7
Clear preference	(1; 3/2; 2)	5
Slight preference	(1; 1; 3/2)	3
Equal preference	(1; 1; 1)	1
Slight inferiority	(2/3; 1; 1)	1/3
Clear inferiority	(1/2; 2/3; 1)	1/5
Very clear inferiority	(2/5; 1/2; 2/3)	1/7
Indisputable inferiority	(1/3; 2/5; 1/2)	1/9

The calculation scheme is similar to the classic AHP method, except that fuzzy evaluations should be used for pairwise evaluation.

4. FORMULATION OF THE RESEARCH PROBLEM

The entity in which the research was carried out was a manufacturing company belonging to the SME sector, dealing in the production of metal and metal-rubber elements used in the automotive industry. The assortment of manufactured products includes over 300 different products, on average twice a month a new product is launched. The large number of products offered resulted in the problem of the appropriate selection of production stations along with their availability (efficient and ready to work machines and production equipment) so that the production process ran smoothly and efficiently. For the planner, it involves a number of actions and decisions that must be made.

The basic information for production planning must be individual data on the availability of production resources - machines and devices. At a later stage of planning, economic and technological factors of production should be taken into account. Based on this data, the planner selects the optimal solution from among many available other solutions, assessing their impact on the organization and management of the enterprise.

The production system of the enterprise works on the basis of the technological principle. This means that the production space consists of separated, separate technological cells with the same type of machines and production devices (high operational interchangeability). It is possible to distinguish turning, milling, drilling, grinding and manual machining (locksmith) production units, combined with a product assembly station. In addition, the cells were divided into rough and fine machining stations. Individual production stations in a cell show differences in terms of efficiency, cost-intensity and availability. Availability of machines or devices is limited by readiness for operation, planned repairs and unexpected breakdowns. These factors cause disruptions that ultimately affect the organization and proper management of the enterprise.

In order to determine the impact on the enterprise and eliminate disturbances in the availability of machinery and equipment, and to support the planner (often inexperienced) in selecting the appropriate resource for a given technological operation, the state of affairs (resource failure rate) was analyzed and an attempt was made to support the decision-making process using the AHP method (as the basic approach), as well as in its extended variant – FAHP. The goal was to select a resource available immediately and meeting the criterion of failure-free working time for the assumed time frame of expected work in unit production. The technological and economic aspects were omitted, as the available resources are at a similar level of technical advancement in individual production units (small differences in efficiency). As a result, the financial aspect can also be omitted in this example, because the lack of availability of the resource (too long waiting time for the product to be put into production) as well as failures compensate for the potential profit.

In order to check the assumptions of the conducted research, one group of machine tools was selected. The selection was made from a group of 10 machines for milling processing, characterized by comparable technical and economic properties and technical wear (data obtained from the maintenance department). A monthly period of work (twenty working days) was taken into account, in which the machines were used in three shifts, five days a week (the working time fund of a single resource was therefore 2,400 hours).

5. METHOD AND OBTAINED RESULTS

The analysis of the operation of machines in the analyzed period allowed to conclude that the availability of individual machine tools was extremely different, as shown in Table 2 and Figure 2. The fourth column of Table 2 presents an indicator that shows the monthly availability of a given resource. Based on the obtained data, an attempt was made to search for the causes that were responsible for the downtime of the machine tools.

Tab. 2. Machine availability indicators

Resource (milling machine)	Working time [%]	Downtime [%]	Availability indicator
M1	90	10	0,9
M2	89	11	0,89
M3	96	4	0,96
M4	76	24	0,76
M5	85	15	0,85
M6	92	8	0,92
M7	97	3	0,97
M8	69	31	0,69
M9	87	13	0,87
M10	88	12	0,88

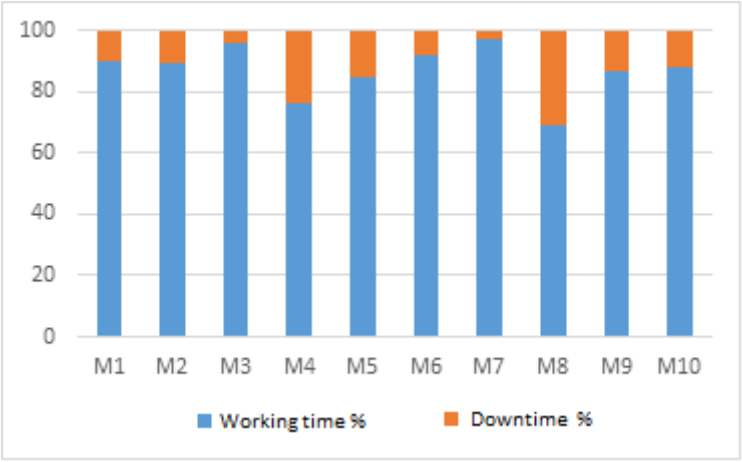


Fig. 2. Percentage of machine downtime in the analyzed period

After determining the reasons for the downtime of the machines, the aspects that had an impact on the organization and management of the company were identified. These reasons have been grouped into areas that include:

- Technical area:
  - failure (mechanical, electrical, hydraulic or pneumatic),
  - change in the quality of the input material of the same production batch,
  - quality of tools used during production tasks,
  - service errors (programming errors or settings),

- machine degradation,
- quality of products,
- work safety,

- non-technical area:
  - the availability of spare parts to remove the failure,
  - the availability of a maintenance worker (mechanic or electrician),
  - the date of completion of the products,
  - the cost of maintaining the machine (e.g. depreciation, inspections, service),
  - work safety.

It was decided that the listed factors will constitute the basic criteria that can be identified with the impact on the enterprise. The planner also has such data and should be guided by the selection of the appropriate workstation for production tasks. For the available criteria, weights were assigned on a scale from 1 to 10 (Table 3). The weight values were consulted with a group of planners and the maintenance department. The individual criteria and the values assigned to them are presented in Table 4.

Tab. 3. Machine availability indicators

Criterion			The value of impact on organization and management
Technical	A1	failure	10
	A2	change in the quality of the input material of the same production batch	3
	A3	quality of tools used during production tasks	2
	A4	service errors	1
	A5	machine degradation	6
	A6	quality of products	4
	A7	work safety	5
Non-technical	P1	the availability of spare parts to remove the failure	8
	P2	the availability of a maintenance worker	5
	P3	the date of completion of the products	3
	P4	the cost of maintaining the machine	3
	P5	work safety	5

The last level of the decision tree – level 3, defines the alternatives. The dilemmas were the available set of machines, defined as M1, M2, ..., M10, forming a decision tree, the solution of which is the appropriate selection of the workplace with the least impact (the least downtime) on the enterprise organization and management system.

The solution to the research problem was carried out using the Matlab program. Due to the large size of the matrix of pairwise comparisons of the AHP and FAHP methods as well as the m-code calculation functions, partial calculations and results are presented (Fig. 3 – AHP method, Fig. 4 – FAHP method, Fig. 5 – synthetic TNF values for criterion A1).

a)

```
function [machine_selection] = Art_AHP()
%% Method analytical hierarchy process (AHP)
%   AUTHOR:
%          P. Wittbrodt, I.Łapuńka, A. Gola
%   CREATED:
%          March, 2022
%% Problem:
% Situation: Select the most suitable machine from the available M1, M2, ..., M10 according to technical and non-technical criteria.
% I will use % AHP with the following:
%        alternatives:  M1, M2, ..., M10
%        criteria:      A1 Malfunction
%                       A2 Change in the quality of input material within the same production batch
%                       A3 Quality of tools used for manufacturing
%                       A4 Operational errors
%                       A5 Machine degradation
%                       A6 Product quality
%                       A7 Occupational safety
%                       P1 Availability of replacement parts needed to repair a malfunction
%                       P2 Availability of maintenance staff
%                       P3 Manufacturing deadline
%                       P4 Machine maintenance costs
%                       P5 Occupational safety
%%-----------------------------------------------

%% Problem formulation:
clear all; close all; clc;
%% Step 1
% Using Scale:
% Weight - numerical scale      Weight - verbal scale
%       1                       equally important
%       3                       little more important
%       5                       much more important
%       7                       far more important
%       9                       absulutely more important
%      2, 4, 6, 8               intermediate values ?
%      1/2, 1/3, 1/4, 1/5,         for inverse relations
%      1/6, 1/7, 1/8, 1/9
```

b)

```
%% Step 2

%% Compare A1
disp('A1');

%     M1    M2    M3    M4    M5    M6    M7    M8    M9    M10
A1 = [1     3     5     5     1     3     3     1/3   5     1    ; ...
      1/3   1     3     5     1/7   3     1/3   3     1/5   1/3  ; ...
      1/5   1/3   1     1/7   1/3   3     5     1     7     1/5  ; ...
      1/5   1/5   7     1     3     1/5   1/3   3     1/5   1/3  ; ...
      1     7     3     1/3   1     1/5   5     1/5   3     5    ; ...
      1/3   1/3   1/3   5     5     1     1/5   5     1/5   1    ; ...
      1/3   3     1/5   3     1/5   5     1     1/3   1/3   3    ; ...
      3     1/3   1     1/3   5     1/5   3     1     1     5    ; ...
      1/5   5     1/7   5     1/3   5     3     1     1     1/7  ; ...
      1     3     5     3     1/5   1     1/3   1/5   7     1    ];

eA1 = calc_eig(A1);
%%
```

c)

```
%% Step 3:
% matrix of eigenvectors calculated above for each criteria eigenvectors: A1, A2, A3, A4, A5, A6, A7, P1, P2, P3, P4, P5
eM  =   [eA1   eA2   eA3  eA4  eA5  eA6  eA7  eP1  eP2  eP3  eP4  eP5];

% obtain scores for each najlepiej dostosowanej maszyny spośród dostepnych M1, % M2, ..., M10 on criteria and massyn comparisons
disp('Scores for: M1, M2, M3, M4, M5, M6, M7, M8, M9, M10')
machine_selection = eM * eC;

%% sub-function on calculating eigenvectors
    function [eigvect ] = calc_eig(M) ...

%% sub-function to normalize a vector (0-1)
    function [normvect ] = calc_norm(M) ...

end
```

Fig. 3. A part of m-code of the AHP method: a) formulation of the problem, b) matrix of pairwise comparisons for criterion A1, c) computational function

a)

```
function [ weights CompMat fuzzyTFN ] = FuzzyAHP( CompMat )
%% Method analytical hierarchy process (AHP)
%   AUTHOR:
%           P. Wittbrodt, I.Łapuńka, A. Gola
%   CREATED:
%           March, 2022
% fuzzy tfn and inverse fuzzy tfn constants
fuzzyTFN = [[1     1     1  ]   [1      1     1  ]
            ...                                 ]];

....
....
%%
% find sum of every l,m,u values for triangular fuzzy number
for i=1:m...

vec = [mExtendAnalysis(1,:)];
mExtendAnalysisSum = sum(reshape(vec,3,[])');

for i=1:m...

%%
% degree of possibility calculation
%              /---
%              | 1    if m2>=m1
%              |
%              | 0    if l1>=l2
% V(M2>=M1) = <
%              |     l1-u2
%              | --------------- otherwise
%              | (m1-u2)-(m1-l1)
%              \---
```

b)

```
%%
% normalised weight calculation
weights = zeros(1,m);
for i=1:m,...
weights = weights/sum(weights);

end
```

Fig. 4. A part of m-code of the FAHP method: a) the sum of l, m, u values for a triangular fuzzy number, b) calculation of weights

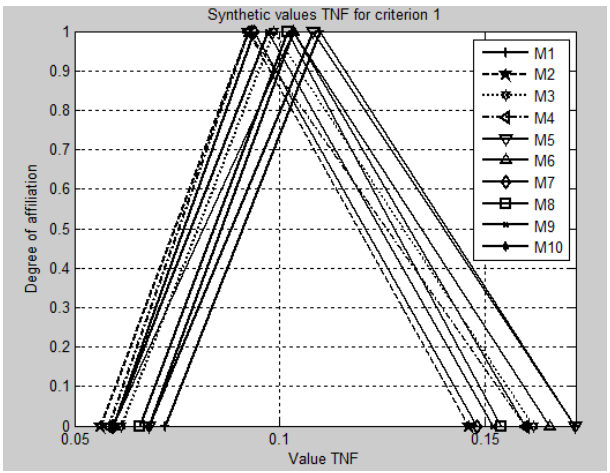


Fig. 5. Synthetic THF values for criterion A1

As shown in Figure 5 – for the A1 criterion, the M5 machine tool is the preferred machine resource, because its values are the largest and the span (width of the set) is the largest.

As a result of the calculations carried out in the Matlab package with the use of m-codes presented in Figures 3–4, a set of values was obtained (for all criteria) for the one-level structure of the problem, which are presented collectively in Table 4 and graphically in Figure 6.

Tab. 3. Comparison of the obtained results with the choice made by the decision maker, using the AHP and FAHP methods

Machine tool	Decision maker	AHP		FAHP	
	Position	Weight	Position	Weight	Position
M1	2	0,111	1	0,119	1
M2	10	0,093	9	0,085	10
M3	6	0,104	5	0,092	8
M4	9	0,073	10	0,090	9
M5	3	0,108	3	0,095	7
M6	8	0,098	7	0,099	5
M7	7	0,096	8	0,097	6
M8	5	0,102	6	0,103	4
M9	4	0,106	4	0,109	3
M10	1	0,109	2	0,111	2
Sum		1,000		1,000	

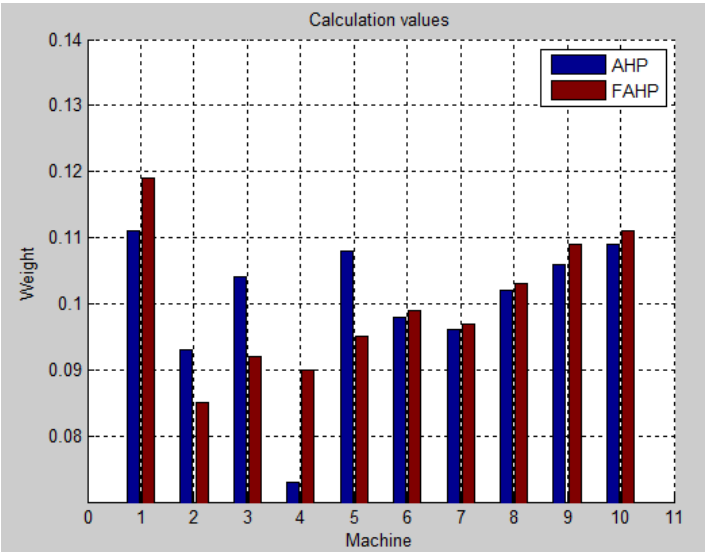


Fig. 5. Bar graph of the results obtained using the AHP and FAHP methods

As shown in Table 5, in the analyzed case, the machine resource M1 should be assigned to the production task in the first place (this results from both the use of the AHP and FAHP methods). The decision maker placed this resource as the second position, preferring the choice of the M10 resource. Nevertheless, the action of the decision-maker and the methods

used should be determined at a satisfactory level, as the presented results are largely convergent. Differences in the obtained results result from the adopted small differences in the values of the criteria weights. The lowest values occur in the M2 and M4 resources, which is caused by the strong impact of the A4 criterion, whose impact on organizations and management is at level 1.

6. CONCLUSIONS

Maintenance is of key importance from the point of view of the efficiency of production processes, and its proper functioning improves the productivity of the company, increases production efficiency, stabilizes the quality of products and timeliness. At the same time, maintenance as a function supporting production reduces downtime and extends the life of production machines and equipment.

This study shows that the decision maker who has the given criteria has the ability to identify faults caused by maintenance for the proper functioning of production and determine the breakdown schedule in the enterprise for individual machines. The correct correlation of the planner's activities and the AHP and FAHP methods allows us to conclude that less experienced employees, based on the support methods, will be able to take corrective actions, preventive actions in the pre-emergency time (before the failure occurs), develop a scenario of actions and appropriately plan the use of production resources. The reaction of maintenance services in such a case is planned and better organized and managed, ensuring minimization of costs for the company and increasing the efficiency of operations. Proper maintenance practice can therefore keep production assets in reliable condition, thereby minimizing production inefficiencies, product defects, downtime, etc.

Maintenance activities are costly, but the lack of these activities is more costly for the company. Companies that will be equipped with a decision support system (e.g. based on the AHP or FAHP method) will be able to improve their maintenance systems and improve their performance. It is important to develop a system supporting the decision-making process, because such a system is the foundation for the efficient and trouble-free operation of the production system, its organization and management.

The research results have a potential practical application, which is demonstrated by a practical example, which is why research on a multidimensional model, taking into account a wider range of criteria and factors, is continued. In the future, the decision support system will be expanded with other decision support methods.

Funding

This work is supported by Lublin University of Technology under grant no. FD-20/IM-5/2022.

Conflicts of Interest

The authors declare no conflict of interest regarding this paper.

REFERENCES

Ahmad, R., & Kamaruddin, S. (2012). An Overview of Time-Based and Condition-Based Maintenance in Industrial Application. *Computers & Industrial Engineering*, *63*(1), 135–149. https://doi.org/10.1016/j.cie.2012.02.002

Al-Najjar, B. (2007). The lack of maintenance and not maintenance which cost: A model to describe and quantify the impact of vibration-based maintenance on company's business. *International Journal of Production Economics*, *107*, 260- 273. https://doi.org/10.1016/j.ijpe.2006.09.005

Al-Najjar, B., & Algabroun, H. (2018). A Model for Increasing Effectiveness and Profitability of Maintenance Performance: A Case Study. In: M. Zuo, L. Ma, J. Mathe & H. Z. Huang (Eds.), *Engineering Asset Management 2016. Lecture Notes in Mechanical Engineering* (pp. 1–12). Springer. https://doi.org/10.1007/978-3-319-62274-3_1

Aspinwall, E., & Elgharib, M. (2013). TPM Implementation in Large and Medium Size Organizations. *Journal of Manufacturing Technology Management*, *24*(5), 688-710. https://doi.org/10.1108/17410381311327972

Azizi, A. (2015). Evaluation Improvement of Production Productivity Performance Using Statistical Process Control, Overall Equipment Efficiency, and Autonomous Maintenance. *Procedia Manufacturing*, *2*, 186–190. https://doi.org/10.1016/j.promfg.2015.07.032

Bałdowska-Witos, P., Piotrowska, K., Kruszelnicka, W., Błaszczak, M., Tomporowski, A., Opielak, M., Kasner, R., & Flizikowski, J. (2020). Managing the uncertainty and accuracy of life cycle assessment results for the process of beverage bottle molding. *Polymers*, *12*(6), 1320. http://doi.org/ 10.3390/polym12061320

Bayazit, O. (2005). Use of AHP in decision-making for flexible manufacturing system. *Journal of Manufacturing Technology Management*, *16*(7), 808–819. https://doi.org/10.1108/17410380510626204

Blanchard, B. (2004). *Logistic Engineering and Management. 5th Ed*. Prentice Hall Inc.

Burduk, A., Musiał, K., Kochańska, J., Górnicka, D., & Stetsenko, A. (2019). Tabu search and genetic algorithm for production process scheduling problem. *Logforum*, *15*(2), 181–189. http://doi.org/10.17270/J.LOG.2019.315

Crespo Márquez, A., Moreu de León, P., Gómez Fernández, J. F., Parra Márquez, C., & López Campos, M. (2009). The maintenance management Framework. A practical view to maintenance management, *Journal of Quality in Maintenance Engineering*, *15*(2), 167–178. https://doi.org/10.1108/13552510910961110

Das, S., & Chattopadhyay, A. B. (2003). Application of the analytic hierarchy process for estimating the state of tool wear. *International Journal of Machine Tools & Manufacture*, *43*(1), 1–6. https://doi.org/10.1016/S0890-6955(02)00168-2

Gola, A. (2014). Economic aspects of manufacturing system design. *Actual Problems of Economics*, *156*(6), 205–212.

Kosicka, E., Gola, A., & Pawlak, J. (2019). Application-based support of machine maintenance. *IFAC-PapersOnLine*, *52*(10), 131–135. http://doi.org/10.1016/j.ifacol.2019.10.033

Kutlu, A. (2012). Fuzzy failure modes and effects analysis by fuzzy TOPSIS-based fuzzy AHP. *Experts Systems with Application*, *39*(1), 61-67. https://doi.org/10.1016/j.eswa.2011.06.044

Madu, C. (2000). Competing through maintenance strategies. *International Journal of Quality and Reliability Management*, *17*(9), 937–948. https://doi.org/10.1108/02656710010378752

Maletic, D., Maletic, C., Al-Najjar, B., & Gomiscek, B. (2014). The Role of Maintenance in Improving Company's Competitiveness and Profitability: A Case Study in A Textile Company. *Journal of Manufacturing Technology Management*, *25*(4), 441–456. https://doi.org/10.1108/JMTM-04-2013-0033

Pizon, J., Kulisz, M., & Lipski, J. (2021). Matrix profile implementation perspective in Industrial Internet of Things production maintenance application. *Journal of Physics: Conference Series*, *1736*(1), 012036. http://doi.org/10.1088/1742-6596/1736/1/012036

Rakyta, M., Fusko, M., Haluska, M., & Grznar, P. (2015). Maintenance support system for Reconfigurable manufacturing systems. *Annals of DAAAM and Proceedings of the International DAAAM Symposium* (pp. 1102–1108). DAAAM International. http://doi.org/ 10.2507/26th.daaam.proceedings.155

Relich, M., & Świć, A. (2020). Parametric estimation and constraint programming-based planning and simulation of production cost of a new product. *Applied Sciences*, *10*(18), 6330. http://doi.org/10.3390/APP10186330

Saaty, T. (1980). *The analytic hierarchy process*. McGraw–Hill.

Sagar, M. K., & Singh, D. (2012). Supplier Selection Criteria: Study of Automobile Sector in India. *International Journal of Engineering Research and Development*, *4*(4), 34–39.

Swic, A., & Gola, A. (2013). Economic analysis of casing parts production in a flexible manufacturing system. *Actual Problems of Economics*, *141*(3), 526–533.

Szabelski, J., Karpiński, R., & Machrowska, A. (2022). Application of an artificial neural network in the modelling of heat cutting effects on the strength of adhesive joints at elevated temperature with imprecisive adhesive mix ratios. *Materials*, *15*(3), 721. http://doi.org/10.3390/ma15030721

Varela, M. L. R., Putnik, G. D., Manupati, V. K., Rajyalakshmi, G., Trojanowska, J., & Machado, J. (2018). Collaborative manufacturing based on cloud, and on other i4.0 oriented principles and technologies: a systematic literature review and reflections. *Management and Production Engineering Review*, *9*(3), 90–99. http://doi.org/10.24425/119538

Vollmann, T. E., Berry, W. L., Whybark, D. C., & Jacobs, F. R. (2005). *Manufacturing Planning and Control for Supply Chain Management. 5th Ed*. McGraw- Hill/Irwin.

Zubrzycki, J., Świć, A., Sobaszek, Ł., Kovac, J., Kralikova, R., Jencik, R., Smidova, N., Arapi, P., Dulencin, P., & Homza, J. (2021). Cyber-physical systems technologies as a key factor in the process of Industry 4.0 and smart manufacturing development. *Applied Computer Science*, *17*(4), 84–99. http://doi.org/10.23743/acs-2021-31