ISSN 2587-1366

TUJE

TURKISH JOURNAL OF ENGINEERING

VOLUME 3 ISSUE 3 JULY 2019

TURKISH JOURNAL OF ENGINEERING (TUJE)

Turkish Journal of Engineering (TUJE) is a multi-disciplinary journal. The Turkish Journal of Engineering (TUJE) publishes the articles in English and is being published 4 times (January, April, July and October) a year. The Journal is a multidisciplinary journal and covers all fields of basic science and engineering. It is the main purpose of the Journal that to convey the latest development on the science and technology towards the related scientists and to the readers. The Journal is also involved in both experimental and theoretical studies on the subject area of basic science and engineering. Submission of an article implies that the work described has not been published previously and it is not under consideration for publication elsewhere. The copyright release form must be signed by the corresponding author on behalf of all authors. All the responsibilities for the article belongs to the authors. The publications of papers are selected through double peer reviewed to ensure originality, relevance and readability.

AIM AND SCOPE

The Journal publishes both experimental and theoretical studies which are reviewed by at least two scientists and researchers for the subject area of basic science and engineering in the fields listed below:

Aerospace Engineering
Environmental Engineering
Civil Engineering
Geomatic Engineering
Mechanical Engineering
Geology Science and Engineering
Mining Engineering
Chemical Engineering
Metallurgical and Materials Engineering
Electrical and Electronics Engineering
Mathematical Applications in Engineering
Computer Engineering
Food Engineering

PEER REVIEW PROCESS

All submissions will be scanned by iThenticate® to prevent plagiarism. Author(s) of the present study and the article about the ethical responsibilities that fit PUBLICATION ETHICS agree. Each author is responsible for the content of the article. Articles submitted for publication are priorly controlled via iThenticate ® (Professional Plagiarism Prevention) program. If articles that are controlled by iThenticate® program identified as plagiarism or self-plagiarism with more than 25% manuscript will return to the author for appropriate citation and correction. All submitted manuscripts are read by the editorial staff. To save time for authors and peer-reviewers, only those papers that seem most likely to meet our editorial criteria are sent for formal review. Reviewer selection is critical to the publication process, and we base our choice on many factors, including expertise, reputation, specific recommendations and our own previous experience of a reviewer's characteristics. For instance, we avoid using people who are slow, careless or do not provide reasoning for their views, whether harsh or lenient. All submissions will be double blind peer reviewed. All papers are expected to have original content. They should not have been previously published and it should not be under review. Prior to the sending out to referees, editors check that the paper aim and scope of the journal. The journal seeks minimum three independent referees. All submissions are subject to a double blind peer review; if two of referees gives a negative feedback on a paper, the paper is being rejected. If two of referees gives a positive feedback on a paper and one referee negative, the editor can decide whether accept or reject. All submitted papers and referee reports are archived by journal Submissions whether they are published or not are not returned. Authors who want to give up publishing their paper in TUJE after the submission have to apply to the editorial board in written. Authors are responsible from the writing quality of their papers. TUJE journal will not pay any copyright fee to authors. A signed Copyright Assignment Form has to be submitted together with the paper.

PUBLICATION ETHICS

Our publication ethics and publication malpractice statement is mainly based on the Code of Conduct and Best-Practice Guidelines for Journal Editors. Committee on Publication Ethics (COPE). (2011, March 7). Code of Conduct and Best-Practice Guidelines for Journal Editors. Retrieved from http://publicationethics.org/files/Code%20of%20Conduct_2.pdf

PUBLICATION FREQUENCY

The TUJE accepts the articles in English and is being published 4 times (January, April, July and October) a year.

CORRESPONDENCE ADDRESS

Journal Contact: tuje@mersin.edu.tr

CONTENTS

Volume 3 – Issue 3

ARTICLES

PREDICTING THE ENERGY PRODUCTION OF A ROOFTOP PV PLANT BY USING DIFFERENTIAL EVOLUTION ALGORITHM
Ali Yıldız **106**

MASS CONCRETE CONSTRUCTION USING SELF-COMPACTING MORTAR
İ. Raci Bayer, Lutfullah Turanli and P. Kumar Mehta **110**

DETERMINATION OF MALIGNANT MELANOMA BY ANALYSIS OF VARIATION VALUES
Ahmet Kürşat Esim, Hilal Kaya and Veysel Alcan **120**

AN INTEGRATED APPROACH BASED ON THE TAGUCHI METHOD AND RESPONSE SURFACE METHODOLOGY TO OPTIMIZE PARAMETER DESIGN OF ASBESTOS-FREE BRAKE PAD MATERIAL
Funda Kahraman and Banu Sugözü **127**

MECHANICAL AND WEAR CHARACTERISATION OF QUARRY TAILINGS REINFORCED A6063 METAL MATRIX COMPOSITES
Stephen Durowaye, Olatunde Sekunowo, Babatunde Bolasodun, Isaac Oduaran and Ganiyu Lawal **133**

MODELLING AND OPTIMIZATION OF THE EXAM INVIGILATOR ASSIGNMENT PROBLEM BASED ON PREFERENCES
Dilek Küçük Matcı and Ilgın Acar **140**

USER-ORIENTED FILE RESTORATION FOR OPERATING SYSTEMS
Hüseyin Pehlivan **149**

Turkish Journal of Engineering

Turkish Journal of Engineering (TUJE)
Vol. 3, Issue 3, pp. 106-109, July 2019
ISSN 2587-1366, Turkey
DOI: 10.31127/tuje.466953
Research Article

PREDICTING THE ENERGY PRODUCTION OF A ROOFTOP PV PLANT BY USING DIFFERENTIAL EVOLUTION ALGORITHM

Ali Yıldız *[1]

[1] Mersin University, Engineering Faculty, Department of Electrical and Electronics Engineering, Mersin, Turkey
ORCID ID 0000-0003-3904-6017
yildiz@mersin.edu.tr

* Corresponding Author
Received: 03/10/2018 Accepted: 04/11/2018

ABSTRACT
In this study, a simple and plain closed-form mathematical expression has been obtained to precisely estimate the monthly production of a rooftop photovoltaic (PV) plant installed in Adana, Turkey. The proposed model is developed by utilizing the Differential Evolution (DE) optimization algorithm based on the PV plant's 5-year (August 2013 – July 2018) real measurement data. The PV plant is a grid-connected rooftop solar PV system located at Kıvanç Textile in Adana, Turkey. The PV system is equipped with an online monitoring system that provides real time data. The study shows the actual energy production is 730 MWh/year on average for 5 years. In order to test the robustness and precision of the present model, it has been compared with the long-term real measurement data of the PV system. The key benefit of the model is giving a convincing prediction of the future production of the PV panel in a simple way. It also does not require any further information other than time. Average percentage error was reached as small as 7.4% for 5-year data.

Keywords: *Energy Production Estimation, Rooftop PV Plant, Optimization, Long Term Performance, Adana*

1. INTRODUCTION

During last decades, diminishing of fossil fuel sources and rising of energy demand have motivated the scientist and engineers to center their studies on renewable energy sources (Bayhan *et al*., 2018; Arslan *et al*., 2016; Bulut *et al*., 2007).

The main renewable energy resources are currently wind, wave, geothermal and solar. Solar energy has the highest potential among those resources. Current technology allows humankind to use solar energy potential by converting it to process heat or converting directly to electric energy by using photovoltaic panels. Another advantage of solar energy is that it is available all over the world.

For the future energy sources, solar energy can be accepted as a key alternative. Therefore, the solar energy is being severely considered for meeting major part of the energy demand in Turkey, and the world. Based on these facts, engineers should properly estimate future solar production of existing PV panels to project their future energy demands. Many studies have been presented in recent years on this issue for Turkey's different cities (Bulut *et al*., 2007; Ecevit *et al*., 2002; Hepbaslı *et al*., 2002; Kaygusuz *et al*., 1999; Kaygusuz *et al*., 1999; Ogulata *et al*., 2002; Togrul *et al*., 1999; Togrul *et al*., 2002; Ulgen *et al*., 2002).

The energy production of PV systems depends upon so many parameters. (Almorox, 2011) sorted those parameters as astronomical factors (distance between Earth and Sun, solar declination, solar constant, and hour angle); geometrical factors (azimuth of the location, surface tilt angle, solar altitude, and solar azimuth); geographical factors (latitude, longitude, and altitude); meteorological factors (atmospheric pressure, cloudiness, temperature and total sunshine duration, air and soil temperature, relative humidity, evaporation, and precipitation number of rainy days, etc.) and physical factors (albedo, water vapor content, scattering of air molecules, scattering of dust and other atmospheric constituents); These parameters are neither deterministic nor completely random; so it is very hard to express all variances mathematically to estimate energy production of PV systems.

Moreover, it is essential to have the data of weather changes throughout the year as well as possible. For this reason, many studies have been carried out to improve models for prediction of typical weather factors such as solar radiation, temperature, and relative humidity both hourly and daily (Knight *et al*., 1991).

In the present study, a simple model to accurately calculate the monthly production of a rooftop PV plant installed in Adana-Turkey was developed by using the DE optimization algorithm. However, in this study only time series was used as an input to predict the energy production of the plant for the sake of simplicity of the model. Since the model is predicting the 30-day production, it is not easy task to have future information about environmental factors such as astronomical, meteorological, and physical factors.

2. MODELLING THE PREDICTION OF FUTURE ENERGY PRODUCTION

The PV plant is a 499,20 kWp grid-connected rooftop solar PV system located at Kıvanç Textile in Adana, Turkey. It was installed on 12.000 m^2-roof of fabrication building. It was the biggest rooftop PV plant of Turkey at the installation time. The plant consists of 3840 Mitshubishi 130 Wp thin film tandem modules with two orientations 10° east and 10° west while north south orientation is 20°. Orientation of the PV panels can be seen in Fig. 1. The PV array is configured in a way that the system contains 40 Fronius inverters which has 12 kWp AC output capacity.

The PV system is equipped with an online monitoring system that provides real time data. The study shows the actual energy production is 730 MWh/year on average for 5 years.

Fig. 1. Orientation of PV plant, panels faces to south.

From studies in the literature (Bulut *et al*., 2007; Ecevit *et al*., 2002; Hepbaslı *et al*., 2002; Kaygusuz *et al*., 1999; Kaygusuz *et al*., 1999; Ogulata *et al*., 2002; Togrul *et al*., 1999; Togrul *et al*., 2002; Ulgen *et al*., 2002), it can be concluded that the energy production of PV systems is a function of time since solar radiation changes distinctively from season to season. Therefore, in this study a new model is created only depending on time parameter based on 5-year measurement of the PV plant.

Daily measurement data can be seen in Fig. 2 from July 2013 to July 2018. It can be easily seen that the data has DC and AC components.

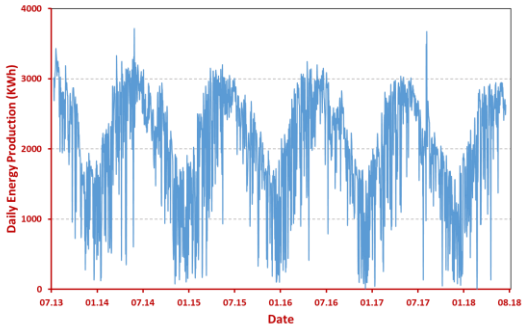


Fig. 2. Daily production of the PV plant.

Since the data has also very significant noises all over the production time, 30-day total production values have been used instead of daily production as can be seen in Fig. 3.

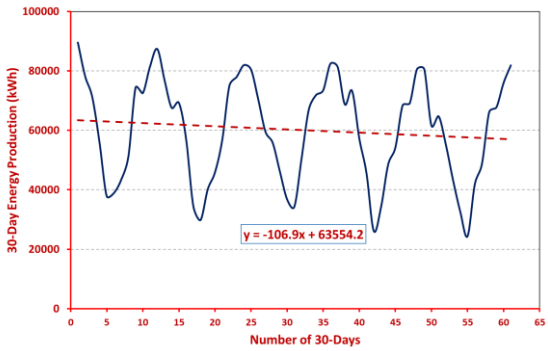


Fig. 3. 30-Day total production of the PV plant.

Fig. 3 shows that the noise is minimized and both DC and AC components are very explicit. DC components have been reached by using regression of the whole data. DC component also gives the degradation of the PV plant. When subtract the DC component from the data we can have AC component which gives the seasonal dependency of the production of the PV system. Because the seasonal dependency is periodic, we choose a basic model with sinusoidal functions for AC component.

2.1. Problem Formulation

The DE algorithm, which was developed to optimize a problem by repeatedly trying to make better a potential solution with regard to a required measure of quality, was used as the optimization tool. The DE algorithm has been preferred to have a solution for many numerical problems with a good performance (Ak *et al*., 2018; Storn *et al*., 1997). It reaches to correct global minimum point of the cost function very fast when compared to other optimization tools. For the optimization, many expressions ranging from simple to complicated ones were tested while modeling the energy production of the system. However, it was figured out that the predictions of simpler models, which have less number of coefficients, are not well agreed with measurements and the more complicated models, which have more coefficients, make just a little progress in the predictions at the expense of the increased complexity. Therefore, the ideal model, which establishes a simple expression and delivers satisfying estimations, was chosen as:

$$P(t) = P_0 - St + \left[c_1\, Cos\big(D(c_2 Mt + c_3)\big) + c_4\, Sin\big(D(c_5 Mt + c_6)\big)\right] \quad (1)$$

P_0: 63554.2 (starting point of DC linear regression)
S: -106.9 (performance drop of the PV system by the time, degradation of the PV plant)
t: Time (30 days)
D: π/180 (radian degree conversion constant)
M: 360/(73/6) (periodicity conversion constant)

Where c_1, c_2, … c_6 are the coefficients to be optimally found by the DE algorithm to minimize the following average percentage error (APE).

$$APE = \frac{100\sum\left|\frac{P_m - P_c}{P_m}\right|}{Total\ Number\ of\ Data} \quad (2)$$

Here, P_m and P_c are measured and calculated power values that is obtained by Eq. (2) respectively. Each coefficient of the model is determined by DE algorithm to optimize its influence based on the measurement data. The all coefficients found by DE optimization algorithm were given in Table 1.

Table 1. Coefficients determined by DE algorithm

c_1	c_2	c_3	c_4	c_5	c_6
25200	1	4	-4300	3	-66.667

By substituting determined coefficients into Eq. (2), developed formula can be simply described as:

$$P(t) = 63554.2 - 106.9t + [25200Cos\big(D(Mt + 4)\big) - 4300Sin(D(3Mt - 66.667))] \quad (3)$$

2.2. Comparison of the Results

Fig. 4 presents the estimations of the derived expression and 5-year real production data. As it can be seen from the graph, the estimations are very close to the measurements within 7.4 % APE. This good self-consistent agreement between predictions of the developed formula and real measurements shows the success of the new model.

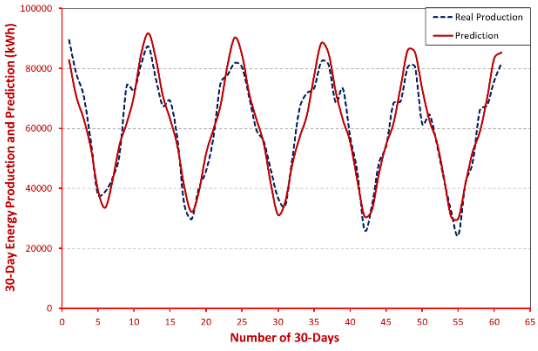


Fig. 4. Real production and prediction of the model.

3. CONCLUSION

In this contribution, a computationally efficient and plain closed-form formula has been attained to estimate the production of energy in a 30-day period for a rooftop PV plant located at Adana, Turkey. The model has been selected so that it has 2 parts; the first part represents the DC component of the energy production while the second part represents the AC part of it. DC part is easy to obtain with the help of linear regression method of the whole production data. However, AC part strongly depends on seasons since the solar radiation changes drastically from winter to summer. Differential Evolution algorithm is used for determining the constants of the model. Because

the seasonal solar radiation is not constant from year to year, the model overestimates the production of the energy for some years especially for summer and winter seasons. Nevertheless, the model is very successful during the seasonal transitions for the spring and fall seasons as can be seen in Fig. 4. When we compare the model with whole data set, it achieves a significant consistency and the APE is reached as 7.4 % which is very satisfactory error level for long term estimation process.

The fundamental contribution of the model is yielding a satisfying estimation of the energy production for a 30-day period with a plain and simple expression. Since the proposed expression accurately estimates energy production for a real PV plant, it can be reliably used to predict the long term future productions for the plants that have similar local specifications. Since the model only takes time series as an input, there is no need to have the further environmental information. One can simply calculate the energy production even with a hand calculator or with a excel sheet in a couple of easy steps.

ACKNOWLEDGEMENTS

The author thanks to Kıvanç textile for providing real time measurement data.

REFERENCES

Ak. C., Yildiz A. (2018) "A novel closed-form expression obtained by using differential evolution algorithm to calculate pull-in voltage of mems cantilever." *Journal of microelectromechanical systems*, Vol. 27, No. 3, pp. 392-397.

Almorox, J. (2011). ''Estimating global solar radiation from common meteorological data in Aranjuez, Spain.'' *Turkish Journal of Physics*, Vol. 35, pp. 53-64.

Arslan, G., Bayhan, B. (2016). "Solar energy potential in Mersin and a Simple model to predict daily solar radiation." *Mugla Journal of Science and Technology*, Special Issue, pp. 1-4.

Bayhan B., Arslan G. (2018) "Applicability of solar and wind energy technologies for a non-residential building." *Turkish Journal of Engineering*, Vol. 2, No.1, pp. 27-34.

Bulut, H., Buyukalaca. O. (2007). "Simple model for the generation of daily global solar-radiation data in Turkey." *Applied Energy*, Vol. 85, No. 5, pp. 477-491.

Ecevit A., Akinoglu B. G., Aksoy B. (2002) "Generation of a typical meteorological year using sunshine-duration data." *Energy*, Vol. 27, No: 10, pp. 947-954.

Hepbaslı A., Ulgen K. (2002) "Prediction of solar-radiation parameters through the clearness index for Izmir, Turkey." *Energy Source*, Vol. 24, No. 8, pp. 773-785.

Kaygusuz K. (1999) "The comparison of measured and calculated solar radiations in Trabzon, Turkey." *Energy Source*, Vol. 21, No. 4, pp. 347-353.

Kaygusuz K., Ayhan T. (1999) "Analysis of solar radiation data for Trabzon, Turkey." *Energy Convers Manage*, Vol. 40, No. 5, pp. 545-556.

Knight K. M., Klein S. A., Duffie J. A. (1991) "A methodology for the synthesis of hourly weather data." *Sol Energy*, Vol. 46, No. 2, pp.109-120.

Ogulata R. T., Ogulata S. N. (2002) "Solar radiation on Adana, Turkey." *Appl Energy*, Vol. 71, No. 4, pp. 351-358.

Storn R., Price K. (1997) "Differential evolution-A simple and efficient heuristic for global optimization over continuous spaces." *J. Global Optim*., Vol. 11, no. 4, pp. 341-359.

Togrul I.T., Onat E. (1999) "A study for estimating solar radiation in Elazıg using geographical and meteorological data." *Energy Convers Manage*, Vol. 40, No.14, pp. 1577-1584.

Togrul I. T., Onat E. (2002) "Global solar-radiation over Turkey: comparison of predicted and measured data." *Renew Energy*, Vol.25, No. 1, pp. 55-67.

Ulgen K., Hepbaslı A. (2002) "Comparison of solar-radiation correlations for _Izmir, Turkey." *Int J Energy Res*, Vol. 26, No. 5, pp. 413-430.

Turkish Journal of Engineering

Turkish Journal of Engineering (TUJE)
Vol. 3, Issue 3, pp. 110-119, July 2019
ISSN 2587-1366, Turkey
DOI: 10.31127/tuje.462548
Research Article

MASS CONCRETE CONSTRUCTION USING SELF-COMPACTING MORTAR

İ. Raci Bayer *[1], Lutfullah Turanli [2] and P. Kumar Mehta [3]

[1] Department of Civil Engineering, Kırıkkale University, Kırıkkale, Turkey
ORCID ID 0000 – 0003 – 0827 – 9168
iraci.bayer@csb.gov.tr

[2] Department of Civil Engineering, Ankara Yıldırım Beyazıt University, Ankara, Turkey
ORCID ID 0000 – 0001 – 5304 – 091X
lturanli@ybu.edu.tr

[3] Department. of Civil and Environmental Eng., University of California at Berkeley, USA
ORCID ID 0000 – 0003 – 1570 – 3617
mehta@ce.berkeley.edu

* Corresponding Author
Received: 21/09/2018 Accepted: 28/11/2018

ABSTRACT
Risk of thermal cracking is a well-known problem in mass concrete structures. One of the steps in mitigating this problem is to keep the initial placement temperature of concrete as low as possible. From this standpoint, Preplaced Aggregate Concrete (PAC) is advantageous because the friction among the coarse aggregate particles during the mixing operation causes the initial temperature of concrete to rise. Making PAC by the conventional method of grout injection requires special equipment and skilled workers. An alternative method using a self-compacting mortar for grouting preplaced aggregate was investigated in this study. Using different concrete compositions made with locally available cementitious materials and aggregate fractions, massive concrete cubes, each 1 m. were prepared. The specimens were tested for thermal and mechanical properties. In addition to low temperature rise that is suitable for mass concrete, the PAC method produced concrete specimens with compressive strength and modulus of elasticity values comparable to conventional concrete.

Keywords: *Preplaced Aggregate Concrete, Mass Concrete, Self-compacting Concrete, High Volume Pozzolan, Heat of Hydration, Thermal Cracks*

1. INTRODUCTION

In massive structures, it is well known that the combination of heat produced by cement hydration and relatively poor heat dissipation conditions results in a large temperature rise in concrete within a few days after placement. Subsequent cooling of concrete to ambient temperature often causes cracking. A primary requirement in the design and construction of mass concrete structures is that the structure should remain a monolithic, free from cracks, therefore every effort to control the temperature rise is made through proper selection of materials, mix-proportions, curing conditions, and construction practice.

Precooling of fresh concrete is a commonly preferred method of minimizing the subsequent temperature drop. For this purpose, chilled aggregates and/or ice shavings are often specified for making concrete mixtures. During the mixing operation, the latent heat needed for melting of ice is withdrawn from other components of the concrete mixture, providing a very effective way to lower the temperature (Mehta and Monteiro, 2005).

The temperature of freshly placed concrete at a given time depends on the placement temperature and the amount of temperature rise subsequent to placement. The temperature rise that occurs after the placement is controlled by the heat of cement hydration. Thus, the first step to be taken to control the temperature rise is the selection of a cementitious material with low heat of hydration. For this purpose, use of high volume of pozzolans is an effective way to prevent thermal cracks. When this is not adequate, a network of pipes is installed in the formwork before the concrete placement, so that cooled water may be circulated later in this pipe network. Some of the heat generated as a result of hydration is thus removed and excessive rise of concrete temperature is avoided. This procedure, called post cooling, is effective in controlling the risk of thermal cracking, but it is quite cumbersome and costly (ACI Committee 207, 2008).

Preplaced Aggregate Concrete (PAC) is defined as the concrete produced by placing coarse aggregate in a form and later injecting a cement-sand grout to fill the voids (ACI Committee 116, 2008). PAC technology was first used in 1937 by Louis S. Wertz and Lee Turzillo during rehabilitation of Santa Fe railroad tunnel near Martinez, California. They were grouting the voids in the concrete at crown areas. In order to reduce the consumption of grout, they thought of filling most of the spaces in formwork with coarse aggregate, before the grouting operation. The resulting concrete showed good performance and attracted the attention of other researchers. Professor Raymond E. Davis, from University of California at Berkeley, developed grout mixtures and basic procedures to make this new method more widely applicable. As a result of his studies, he determined many of the unique properties of PAC, which is also known as two-stage concrete (ACI Committee 304, 2008). Since a large volume of the two-stage concrete is occupied by coarse aggregate, it is likely that variations in coarse aggregate content will significantly influence the strength of PAC (Abdelgader, 1996). Because of this, the commonly used formulae for compressive strength of traditional concrete are not valid for PAC (Rajabi and Moaf, 2017). Although the modulus of elasticity can be satisfactorily predicted from the compressive strength value for normal concrete, there is no background information for the evaluation of elastic modulus in the case of PAC (Abdelgader and Górski, 2002). However, there are some equations in the literature offered for PAC to derive the static modulus of elasticity as a function of the compressive strength (Abdelgader and Górski, 2003). Similarly, the equations used for ordinary concrete to predict the tensile strength (splitting) from the compressive strength value are not applicable for PAC, but there are some relations between the tensile strength and the compressive strength of PAC offered in the literature (Abdelgader and Ben-zeitun, 2004; Abdelgader and Elgalhud, 2008). Good quality PAC grout can be prepared in high-speed mixers and because of the unique equipment involved, PAC technology is underutilized. However, there are some grout mix designs which can be followed to achieve similar results by using conventional mixing equipment and recently developed super-plasticizing admixtures (Nowek *et al*., 2007).

In this study, the authors assumed that application of PAC method in combination with the use of high volume pozzolan could eliminate the need for the cumbersome and costly post cooling operation for mass concrete construction. Making PAC by the conventional grout injection method requires special equipment and skilled workers. An alternative method, using a self-compacting mortar for grouting was investigated in this study. Four concrete cube specimens, each 1 m in volume and having different mix designs, were cast; three cubes by PAC method with the new approach based on the use of self-compacting grout, and one cube as conventional concrete for reference purposes.

2. EXPERIMENTAL STDUY

2.1. Materials

Four sizes of aggregate fractions, namely 0 - 3 mm, 7 - 15 mm, 15 - 30 mm and 30 - 60 mm, were used. They include both fine and coarse aggregates, from the same geological region (nearly equal content of monzonite-monzodiorite and meta-andesite minerals). The specific gravity, absorption capacity values, and sieve analyses results of aggregates are shown in Table 1 and Table 2, respectively.

Table 1. Specific Gravity and Absorption Capacity of the Aggregates

Aggregate	Specific Gravity		Absorption Capacity
Class, mm	Dry	SSD	(%)
0-3	2.59	2.61	0.98
7-15	2.66	2.68	0.79
15-30	2.66	2.69	1.02
30-60	2.7	2.71	0.6

Table 2. Sieve Analysis of the Aggregates

Sieve No. (mm)	% Passing 0-3 mm	7-15 mm	15-30 mm	30-60 mm
63.5	100	100	100	100
50.8	100	100	100	69.26
28.1	100	100	100	14.85
25.4	100	100	93.4	0
19.05	100	100	70.42	0
12.7	100	100	19.37	0
9.53	100	85.41	1.54	0
4.76	99.81	28.94	0	0
2.38	94.68	0	0	0
1.19	60.17	0	0	0
0.59	33.6	0	0	0
0.297	16.48	0	0	0
0.149	8.19	0	0	0

The same type of a Portland cement and a blended Portland cement, which are commercially available in Turkey were used in all concrete mixtures. The Portland cement is classified as CEM I 42.5 according to Turkish Standards (similar to Type I - Ordinary Portland Cement according to ASTM Standard C 150). The blended cement is classified as CEM II/B-P 32.5 according to TS EN 197-1 (similar to Portland-Pozzolan Cements according to ASTM Standard C 595), and contains about 35% (by weight) pozzolanic addition (EN TS 197-1, 2012; ASTM C 150, 2012; ASTM C 595, 2000).

The fly ash was obtained from a thermal power plant in Turkey, and it is similar to ASTM Class F fly ash, C 618 (ASTM C618-15, 2015). The brick powder was obtained from a factory in Manisa (Turkey), which collects the broken bricks from brick factories. It is well known that pieces of clay bricks were widely used in Roman and Byzantine structures to obtain hydraulic mortars. Crushed and/or powdered bricks were utilized for load-bearing and water-resistant purposes (Degryse *et al*., 2002; Moropolou et. al, 2000; Baronio and Binda, 1997). Recent studies have shown that ground clay brick when added to cement in sufficient amount increased the resistance of mortar against chemical attacks (Afshinnia and Poursaee, 2015; Kishar *et al*., 2013; Turanlı *et al*., 2003).

The PAC specimens are designated according to their dominant cementitious component, such as FA for Fly Ash, BP for Brick Powder, BC for Blended Cement, and CC for Conventional Concrete. The difference between the middle and the surface temperatures of the specimens was monitored by thermocouples located in the specimens until the temperatures reached a steady-state. Compressive strength and modulus of elasticity were determined from the core specimens taken at the ages of 28, 90, 180, 360, 540 and 720 days. Minimum three cores were tested at each age and the diameter of the cores were 250 mm.

The physical properties and the chemical analysis of all the cementitious materials used are given at Table 3 and Table 4, respectively. As chemical admixtures, the same superplasticizer and air-entraining admixture were used in all concrete mixtures**.**

Table 3. Physical Properties of the Cementitious Materials

Property		Portland Cement	Blended Cement	Fly Ash	Brick Powder
Specific Gravity		3.12	2.78	2.1	2.64
Blaine Fineness (m^2/kg)		392	480	289	400
W/C for NC		0.26	0.27	0.5	-
45 μm sieve residue (%)		-	-	27	37
Strength Activity Index (%)	7 days	-	-	82	66.2
	28 days	-	-	89	80.3
Setting Time (min)	Initial	108	110	-	-
	Final	162	165	-	-
Compressive Strength (Mpa)	3 days	27.6	15.4	-	-
	7 days	39	19.6	-	-
Heat of Hydration (cal/g)	28 days	47.7	34.1	-	-
	90 days	-	38.8	-	-
	3 days	67.5	45.9	-	-
	7 days	74.9	49	-	-
	28 days	92.1	74	-	-
	90 days	-	83.1	-	-

Table 4. Chemical Analysis of the Cementitious Materials

% by weight	Portland Cement	Blended Cement	Fly Ash	Brick Powder
CaO	62.56	41.93	3.34	3.94
SiO_2	20.47	31.31	58.4	62.7
Al_2O_3	5.68	8.94	18.8	17.1
Fe_2O_3	3.08	4.19	10.6	6.84
MgO	1.8	3.23	4.52	2.25
SO_3	3.22	2.09	1.75	0.84
K_2O	0.95	-	1.86	-
Na_2O	0.3	-	0.22	-
Cl^-	0.014	-	-	-
P_2O_5	-	-	0.25	-
Mn_2O_3	-	-	0.22	-
Free CaO	0.98	-	-	-
Loss on Ignition	2.49	5.46	0.77	2.67

Table 5. Mixture Proportions

		Amount (kg/m^3)										
		Portland Cement	Blended Cement	Fly Ash	Brick Powder	Super-plasticizer	Air-entrainer	Water	Aggregate (mm) 0-3	7-15	15-30	30-60
Specimen	CC	115	-	173	-	2.88	0.576	144	538	365	457	465
	FA	115	-	173	-	5.76	0.576	144	576	-	-	1420
	BP	118	-	-	177	5.9	0.59	148	590	-	-	1420
	BC	-	175	113	-	5.76	0.576	144	576	-	-	1420

The mixture proportions of the four concrete types are shown in Table 5. The same coarse and fine aggregates gradations were used for all PAC specimens. As the volume of the preplaced coarse aggregate is determined by the volume of the formwork used, the same amount of aggregate was used for all specimens. The amount of fine aggregate was fixed as twice the amount of cementitious material, therefore, it differed in each concrete type due to the fact that the cementitious material required to fill the volume of voids differed according to its specific gravity. Note that all the specimens, including the reference specimen, contained 60% (by weight) mineral admixture. The gradation of the aggregate for the reference specimen was arranged to have a well-graded aggregate with a maximum aggregate size of 60 mm. The W/C ratio was kept the same (0.5) for all concrete types. Also, all the specimens contained 0.2% by mass of air-entraining admixture. While the superplasticizer content was 0.2% by mass for the PAC specimens, it was 1% by mass for the reference concrete in order to have 75 mm slump which is appropriate for filling all the voids within the preplaced aggregate.

2.2. Research Significance: A New Method for Making PAC

In the conventional method of making PAC, first the coarse aggregate is placed into the forms, and then the grout is injected through the pipes inserted in the coarse aggregate mass by special pumps, till the injected grout rises up to the top surface of the preplaced coarse aggregate mass. Despite the many advantages of making concrete by this way as compared to conventional concrete, in many situations this method requires special equipment and experienced workers. Self-compacting concrete (SCC) technology is a new and technique that was developed for placement of concrete mixtures especially where the closely-spaced reinforcement does not allow concrete compaction by vibration. In SCC, the rheology of concrete is improved by the use of new generation of superplasticizers and chemical admixtures such that the concrete mixture is very fluid but does not segregate (Demir *et al*., 2018; Akgüngör *et al*., 2017). It fills all the voids in the formwork without requiring any high-pressure injection or vibration. The authors used SCC for making PAC mixtures of this study. By using a suitable superplasticizer and a large volume of pozzolanic admixtures, the rheology of the mortar was controlled such that it could fill without segregation the voids when placed over the preplaced coarse aggregate. Thus, this method can be called as making PAC by a self-compacting grout (PAC-SCG).

This way of making PAC has been found to be much more advantageous compared to the conventional injection method, since it does not require any special equipment, such as pumps and insert pipes, or any skilled workers. The grout prepared by a conventional drum mixer, when placed at the top of the preplaced coarse aggregate, penetrates into voids by gravity. The question about this method was how can one be sure whether all the voids have been filled? This can be simply checked by the amount of grout consumed; because, if the volume of voids within the preplaced coarse aggregate mass is known, the amount of grout required to fill these voids can be determined easily before starting the operation. The volume of voids within a preplaced coarse aggregate can easily be determined by filling a container of known volume with that coarse aggregate gradation in SSD state, and then adding water to the container till it overflows. In such a test, it is obvious that the volume of the water consumed is equal to the volume of voids within the aggregate mass; and proportioning this volume to the volume of the container used for the test, the actual volume of voids can easily be calculated by using the volume surrounded by the formwork.

In this research, it has been observed that the amount of grout required to fill all the voids within the preplaced coarse aggregate mass, which had been determined before starting to place the grout, was totally consumed when the grout appeared at the top of the aggregate mass, which means that all the voids within the preplaced coarse aggregate mass have been completely filled. As a final check on the uniformity of the void-filling operation, concrete cores can be obtained and visually examined. Photograph of a cross-section of the concrete core prepared by the new PAC-SCG method is shown in Fig. 1, and a similar photograph of concrete prepared by the conventional PAC method used for restoration of Barker Dam under the supervision of Professor Raymond E. Davis about 60 years ago, is shown in Fig. 2. When these photographs are compared, it is seen that the voids in the preplaced aggregates have successfully been filled by the SCG technique.

Fig. 1. A sample of PAC prepared by the new SCG method

Fig. 2. A sample of PAC prepared by the conventional method for the restoration of Barker Dam

2.3. Test Methods

All the concrete tests were carried out at the Materials of Construction Laboratory of the Civil Engineering Department of Middle East Technical University, Turkey, and in accordance with the ASTM Standard Test Methods.

Prevention of thermal cracking is an important goal in mass concrete construction, therefore large concrete samples are used to monitor temperature changes in the proposed mix design. During the preparation of the specimens, four thermocouples were placed in each specimen, two of them at the surface and two of them in the middle of the specimen, as shown by Fig. 3. The reason of using two thermocouples at the same location of the specimen was to insure the continuity of temperature measurement in case one of the thermocouples stopped working. Temperature measurements were discontinued when they reached a steady-state.

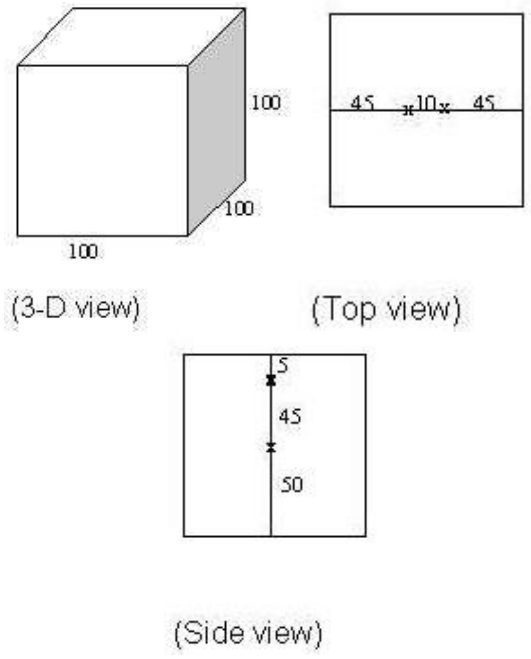


Fig. 3. The location of the thermocouples in the concrete cube specimens (the dimensions are given in cm)

Since it was not possible to carry such big specimens to the curing room of the laboratory, a special tent was built for these specimens, and they were kept in this tent till the formworks were removed. The formworks were removed after the task of monitoring the concrete temperatures had ended with the reaching of steady temperature values, which took approximately one to two weeks, depending on the type of the concrete mix. The ambient temperature and the relative humidity inside the tent were monitored by digital equipment, and an attempt was made to keep them as constant as possible.

When the concrete temperature reached a steady-state, the formwork was removed, and rate of strength development was monitored by taking core samples and testing them under compression at ages 28, 90, 180, 360, 540 and 720 days. Modulus of elasticity values were also determined for the same specimens.

3. RESULTS AND DISCUSSION

3.1. Temperature Measurements

The results of temperature measurement at the surface and middle of concrete specimens are shown in Fig. 4. It can be observed that the maximum peak temperature, 41 C, occurred in concrete CC, and the minimum, 37 C occurred in concrete BC. From standpoint of the risk of thermal cracking, the peak temperature difference is more important than the peak temperature itself (Mehta and Monteiro, 2005). The maximum peak temperature difference 11.5 C occurred in the case of concrete CC, which is acceptable for mass concrete; however, this is higher than the corresponding temperature difference of 9.5 C with specimen FA containing the same amount of total cementitious materials as concrete CC but using the new method.

The minimum peak temperature difference occurred with the specimens BP and BC, one of which contains brick powder, the other contains a natural pozzolanic mineral admixture. Fly ash is already known as a successful mineral admixture in reducing the heat of hydration, but this study shows that brick powder and interground fly ash in blended cements are also a very suitable for mass concrete due to their effectiveness in reducing the heat of hydration. It is interesting to note that the blended cement containing interground fly ash was more effective in reducing the heat of hydration than separately batched fly ash, the explanation of which requires further investigation.

While evaluating the specimens under the topic of temperature measurements, the other two important criteria are the rate of hydration heat evolution, which is represented by the time of occurrence of the peak temperature difference, and the length of the duration for the temperatures to get stabilized, which represents the rate of heat release of the specimen. Evaluating the specimens from these criteria points of view, the shortest duration for the temperatures to get stabilized has occurred with the specimen CC, which also showed the maximum peak temperature difference in minimum time. On the other hand, the longest durations were recorded for concrete specimens BP and BC which also showed minimum peak temperature differences with the slowest rates of heat release. As a result it can be concluded that, as the peak temperature increases (which means that as the amount of heat of hydration increases), the time

required to reach this peak value decreases (which means that the rate of heat evolution increases), and the length of duration for the temperatures to get stabilized decreases. The time to reach the peak temperature value is important, because during the first few days following placement the rate of cooling or heat removal can be as high as possible since the elastic modulus of concrete is relatively low. The strength and the elastic modulus generally increase rapidly until after the peak in concrete temperature has been experienced. When concrete has become elastic, it is important to have temperature drop as slowly as possible (ACI Committee 207, 2008).

Photographs of cross-section of concrete specimens, both without and with magnification (1 to 25) are presented in Fig. 5 and Fig. 6, respectively. None of the specimens showed thermal cracking.

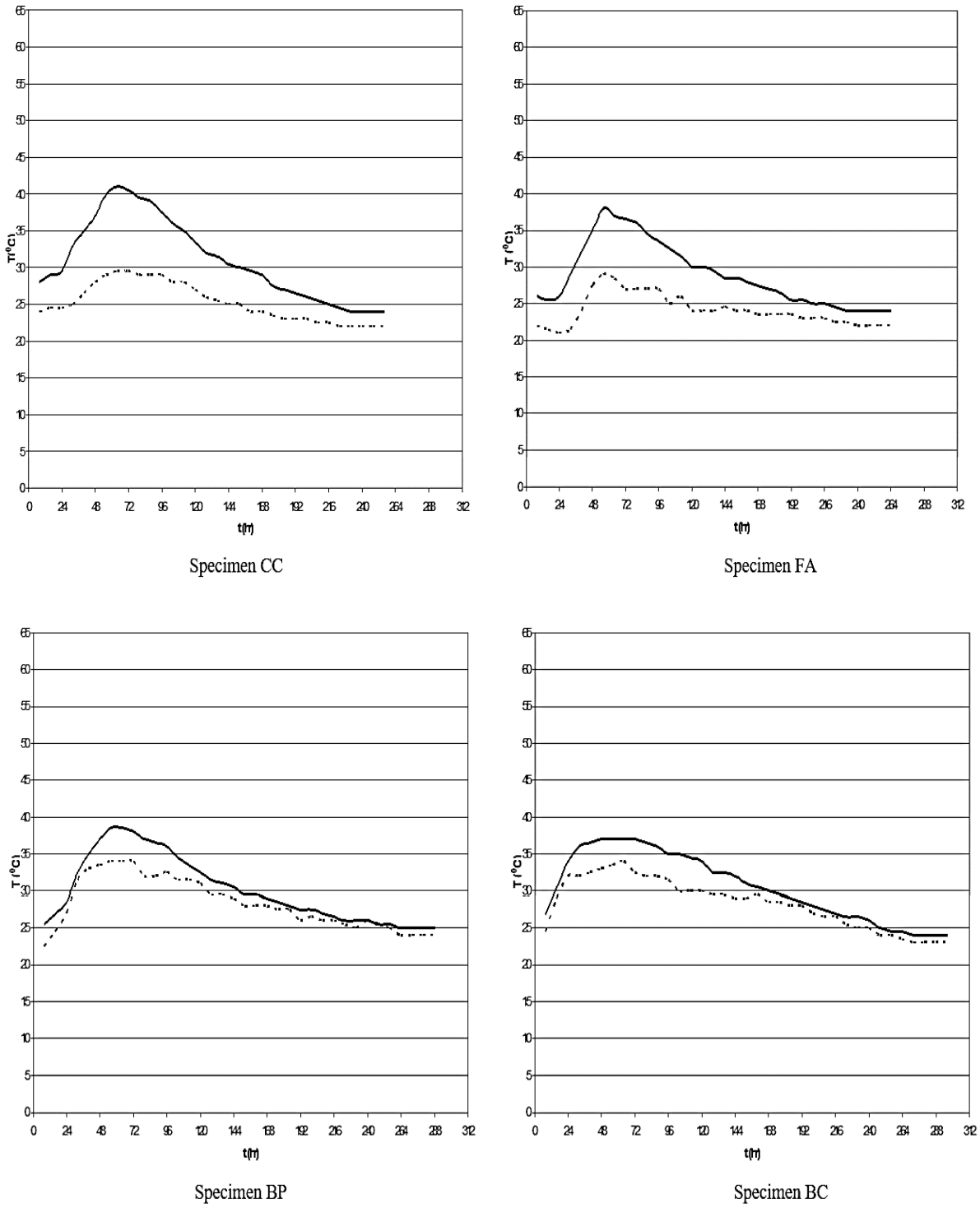
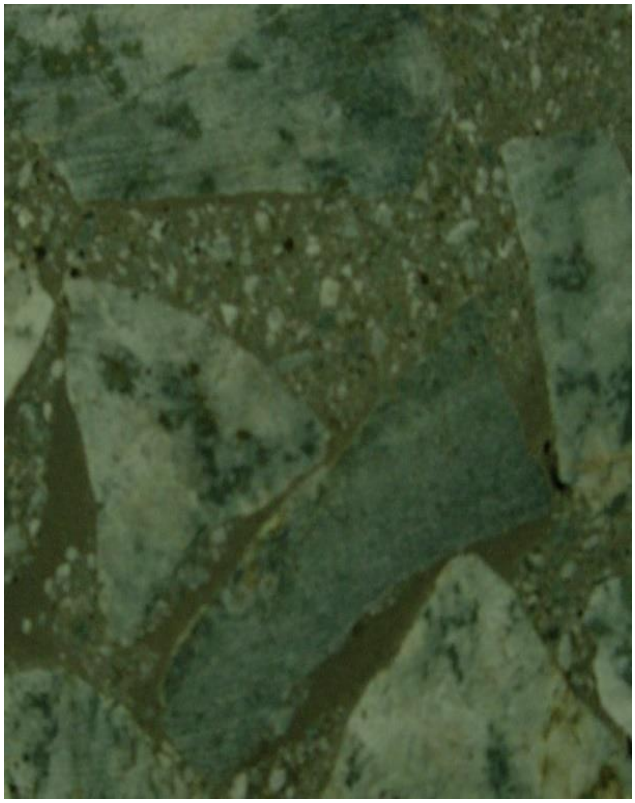
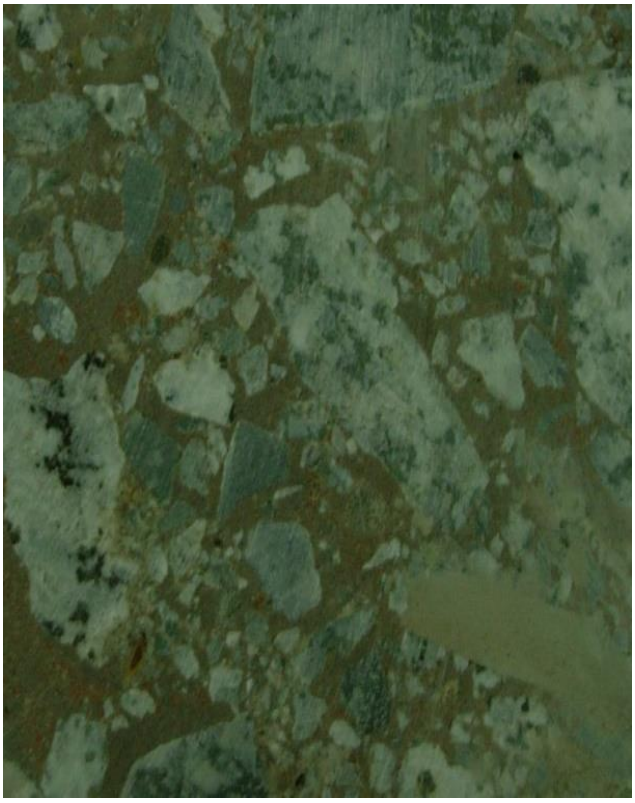


Fig. 4. Concrete Temperature (——— middle, - - - - - surface)

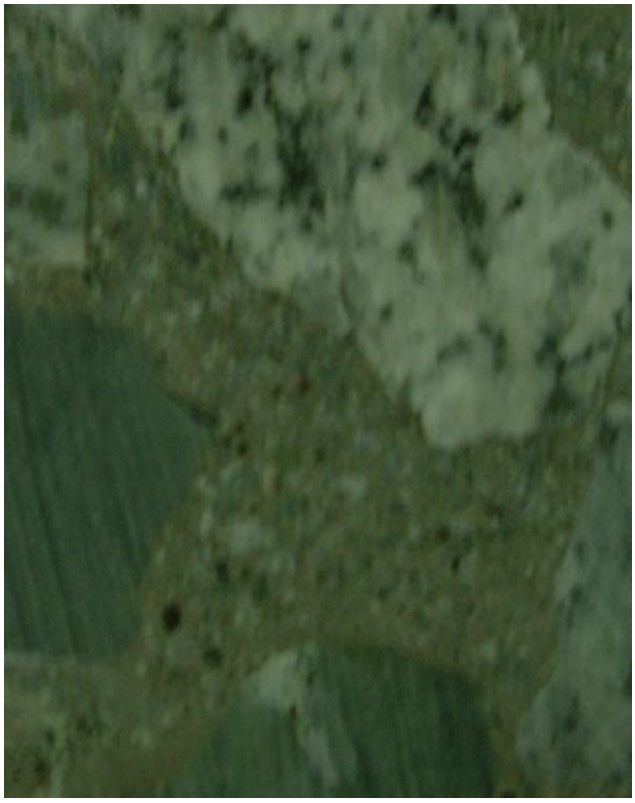
Specimen CC



Specimen FA



Specimen BP



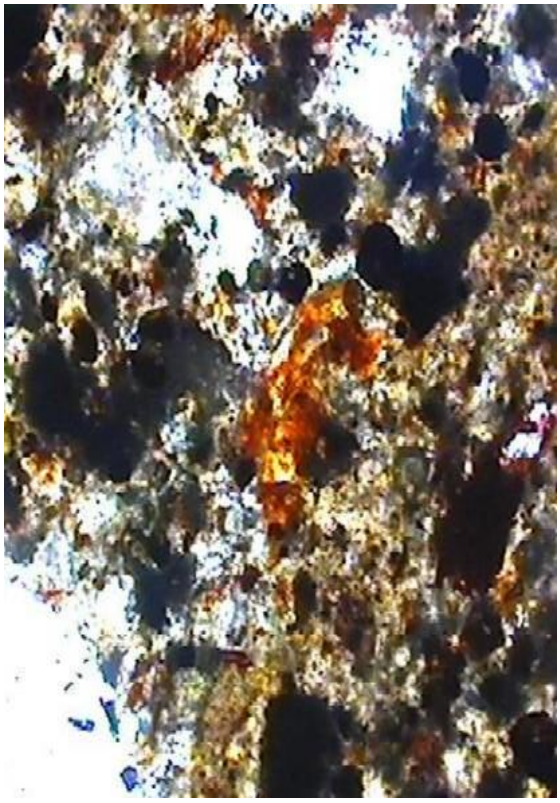
Specimen BC

Fig. 5. The Appearance of the Specimens at Macro Scale (1 to 1)

Specimen CC

Specimen FA

Specimen BP



Specimen BC

Fig. 6. The Appearance of the Specimens at Micro Scale (1 to 25)

3.2. Mechanical Properties

Table 6 shows concrete core compressive strength and elastic modulus data at ages 28, 90, 180, 360, 540 and 720 days. The concrete specimens produced by PAC-SCG method gave strength values equal or greater than those of the reference specimen CC at all ages. The concrete specimens FA and BC gave strength values that were close to each other at all ages, which is not surprising because the total percentage of pozzolanic addition was same and the mineral admixture was fly ash in both the specimens. The surprising result was with brick-powder concrete (BP) because at early ages this concrete gave strength values close to FA and BC; however, at later ages (180 days and more) much higher strength values were obtained.

In regard to the elasticity moduli, according to the published literature the modulus of elasticity of PAC is usually more than half of that of conventional concrete (Waterways Experiment Station, 1954; Davis, 1958; Davis, 1960).

Table 6. Mechanical Properties

		Specimen			
		CC (Mpa)	FA (Mpa)	BP (Mpa)	BC (Mpa)
28 days	σ	15.6	16.5	15.9	14.6
	E	22565	16033	15879	11408
90 days	σ	18.7	19.9	19.9	18.9
	E	26837	25930	19874	17825
180 days	σ	20.6	21.9	24	21.3
	E	28571	26662	25816	20499
360 days	σ	22.9	23.4	27.8	23.1
	E	29335	29944	30909	22678
540 days	σ	23.7	24.4	30.8	24.2
	E	30287	31602	34375	24204
720 days	σ	24.5	25.1	32.9	24.8
	E	31952	32591	36446	25223

σ: Compressive Strength (MPa) ; E: Modulus of Elasticity (MPa)

In this study, this was the case for the elastic moduli up to 180 days, however, the values were either similar or slightly higher for FA and BP at later ages than 180 days. The formulae used for conventional concrete to predict the modulus of elasticity from the compressive strength do not provide reasonable results for conventional PAC specimens (Abdelgader and Górski, 2002; Abdelgader and Górski, 2003). However, for the PAC-SCG specimens, when the expected modulus of elasticity values calculated from the corresponding compressive strength values by the formula offered by ACI (ACI Committee 318, 2015) taken into account, it is seen that after 90 days, the observed values are either more or less similar to predicted values, or even higher.

4. CONCLUSIONS

A new method of making preplaced aggregate concrete has been investigated for mass concrete construction. For this purpose, four cubes of concrete, each with a volume of 1 m and with different mix designs, were prepared, and the specimens have been examined from thermal and mechanical properties points of view. As a result of the experiments carried out, the following conclusions have been achieved:

1. The maximum peak temperature difference of 11.5 C was obtained with reference concrete specimen CC. This value does not indicate any risk of thermal cracking; because it is generally experienced that slowly cooled massive concrete structures can withstand a 20 C drop in temperature without cracking (Mehta and Monteiro, 2005**)**. All other PAC concrete specimens made with SCG method showed even a lower peak temperature difference. Thus it can be concluded that the PAC-SCG method is very effective in mitigating the risk of thermal cracking in mass concrete.

2. The PAC method using the new technology of self-compacting mortar for grouting the pre-placed aggregate (PAC-SCG), is even more effective in mitigating the risk of thermal cracking by lowering the maximum peak temperature as a result of reduced initial temperature of freshly placed concrete.

3. The PAC-SCG method produced concrete specimens with compressive strength and modulus of elasticity values comparable to that of conventional concrete at early ages, and somewhat higher values at later ages. Although the tensile strength of the PAC specimens were not determined experimentally, by applying the formula suggested by Abdelgader *et al.*, 2006 it can be concluded that PAC is also more advantageous from tensile strength point of view than conventional concrete. This is important in mass concrete to avoid the tensile cracks caused by temperature changes.

4. It is already known from the feasibility study of the restoration of Barker Dam that the cost of mass construction by PAC method is less than that of by conventional concrete (Baumann, 1948). It is obvious that the cost of mass construction can be reduced further by the new approach offered in this paper, since the needs for special equipment and skilled labor are eliminated.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Middle East Technical University for providing financial support under project BAP 2004-03-03-04, TEKAR Technical Research Ltd. for supplying the equipment. The authors also thank ERG Construction Co. for supplying the materials used in this research, and Professor M. Cemal Göncüoğlu (Department of Geological Engineering, Middle East Technical University) for providing the micrographs of the specimens.

REFERENCES

Abdelgader, H. S., 1996. Effect of quantity of sand on the compressive strength of two-stage concrete. *Mag.Concr. Res.* 48(177): 353-360

Abdelgader, H. S., and Górski, J., 2002. Influence of grout proportions on modulus of elasticity of two-stage concrete. *Mag.Concr. Res.* 54(4): 251-255

Abdelgader, H. S., and Górski, J., 2003. Stress-strain relations and modulus of elasticity of two-stage concrete. *J. Mater. Civ. Eng.* 15(4): 329-334

Abdelgader, H. S., and Ben-zeitun, A. E., 2004. Effect of grout proportions on tensile strength of two-stage concrete measured by split and double-punch tests. *Structural Concrete* 5(4): 173-177

Abdelgader, H. S., Ben-Zeitun, A. E., Saud, A. F. and Elgalhud, A. A., 2006. Mechanical Behavior of Two-Stage (Pre-placed Aggregate) Concrete. *The 10th Arab Structural Engineering Conference*, Kuwait

Abdelgader, H. S. and Elgalhud, A. A., 2008. Effect of grout proportions on strength of two-stage concrete. *Structural Concrete* 9(3): 163-170

ACI Committee 116 (116R-00), 2008. *Cement and Concrete Terminology Part I*: ACI Manual of Concrete Practice

ACI Committee 207 (207.1R-05), 2008. *Mass Concrete Part II*: ACI Manual of Concrete Practice

ACI Committee 304 (304.1R-92), 2008. *Guide for the Use of Preplaced Aggregate Concrete for Structural and Mass Concrete Applications Part II*: ACI Manual of Concrete Practice

ACI Committee 318 (ACI 318-14), 2015. *Commentary on Building Code Requirements for Structural Concrete*: American Concrete Institute

Afshinnia, K., and Poursaee, A., 2015. The potential of ground clay brick to mitigate Alkali–Silica Reaction in mortar prepared with highly reactive aggregate *Const. Build. Mater*. 95: 164-170

ASTM C 150, 2012. *Standard specification for Portland cement*, American Society of Testing and Materials, West Conshohocken, PA: ASTM International

ASTM C 595, 2000. *Standard Specification for Blended Hydraulic Cements*. American Society of Testing and Materials, West Conshohocken, PA: ASTM International

ASTM C618-15, 2015. *Standard Specification for Coal Fly Ash and Raw or Calcined Natural Pozzolan for Use in Concrete*, American Society of Testing and Materials, West Conshohocken, PA: ASTM International

Akgüngör, A., Sevim, Ö., Kalkan, İ. and Demir, İ., 2018. Restrained Shrinkage Cracking of Self-Consolidating Concrete Roads. *Sci. Eng. Compos. Mater.* 43:112

Baumann, P., 1948 Use of Prepacked Aggregate Concrete in Major Dam Construction. *ACI Journal Proceedings*, 20(3): 229-236

Baronio, G., and Binda, L., 1997. Study of the Pozzolanicity of Some Bricks and Clays. *Const. Build. Mater.* 11(11): 41-46

Davis, H. E., 1958. High-Density Concrete for Shielding Atomic Energy Plants. *ACI Journal Proceedings* 54(11): 965-977

Davis, R. E., 1960. Prepakt method of concrete repair. In *Journal Proceedings* (Vol. 57, No. 8, pp. 155-172).

Degryse, P., Elsen, J. and Waelkens, M., 2002. Study of Ancient Mortars from Sagalassos (Turkey) in view of Their Construction. *Cem. Concr. Res.* 32: 1457-1463

Demir, İ., Sevim, Ö. and Tekin, E., 2018. The Effects of Shrinkage-reducing Admixtures Used in Self-Compacting Concrete on its Strength and Durability, *Const. Build. Mater.* 172: 153-165

EN T S 197-1, 2012. *Cement-Part 1: Composition, specifications and conformity criteria for common cements*, Ankara: Turkish Standards Institute

Kishar, E. A., Ahmed, D. A., Mohammed, M. R. and Noury, R., 2013. Effect of calcium chloride on the hydration characteristics of ground clay bricks cement pastes. *Beni-Suef university journal of basic and applied sciences*, 2(1): 20-30

Mehta, P. K., and Monteiro, P. J. M., 2005. *Concrete: Microstructure, Properties, and Materials.* New York: McGraw-Hill

Moropolou, A., Çakmak, A. S. and Lohvyn, N., 2000. Earthquake Resistant Construction Techniques and Materials on Byzantine Monuments in Kiev. *Int. J. Soil Dyn. Earthquake Eng.* 19: 603-615

Nowek, A., Kaszubski, P., Abdelgader, H. S. and Górski, J., 2007. Effect of admixtures on fresh grout and two-stage (Pre-placed aggregate) concrete. *Structural Concrete* 8(1): 17-23

Rajabi, A. M. and Moaf, F. O., 2017. Simple empirical formula to estimate the main geomechanical parameters of preplaced aggregate concrete and conventional concrete. *Constr. Build. Mater.* 146: 485-492

Turanlı, L., Bektaş, F. and Monteiro, P. J. M., 2003. Use of Ground Clay Brick as a Pozzolanic Material to Reduce the Alkali-Silica Reaction. *Cem. Concr. Res.* 33: 1539-1542

Waterways Experiment Station (U.S.)., 1954. Investigation of the Suitability of Prepakt Concrete for Mass and Reinforced Concrete Structures: *Appendices A and B, report.* Vicksburg, Mississippi. (digital.library.unt.edu/ark:/67531/metadc853010/: accessed September 20, 2018), University of North Texas Libraries, Digital Library, digital.library.unt.edu; crediting UNT Libraries Government Documents Department.

Turkish Journal of Engineering

Turkish Journal of Engineering (TUJE)
Vol. 3, Issue 3, pp. 120-126, July 2019
ISSN 2587-1366, Turkey
DOI: 10.31127/tuje.472328
Research Article

DETERMINATION OF MALIGNANT MELANOMA BY ANALYSIS OF VARIATION VALUES

Ahmet Kürşat Esim [1], Hilal Kaya [2] and Veysel Alcan *[3]

[1] Ankara Yıldırım Beyazıt University, Faculty of Engineering and Natural Science, Department of Computer Engineering, Ankara, Turkey
ORCID ID 0000 – 0002 – 3458– 2411
kursatesim@gmail.com

[2] Ankara Yıldırım Beyazıt University, Faculty of Engineering and Natural Science, Department of Computer Engineering, Ankara, Turkey
ORCID ID 0000 – 0003 – 4787 – 105X
hilalkaya@ybu.edu.tr

[3] Tarsus University, Faculty of Technology, Department of Software Engineering, Mersin, Turkey
ORCID ID 0000 – 0002 – 7786– 8591
alcan@tarsus.edu.tr

*Corresponding Author
Received: 13/11/2018 Accepted: 18/12/2018

ABSTRACT
Melanoma is a serious disease associated with mutation-based cancer cells. Genetic structure and hereditary condition play important role to understand the underlying reasons of the diseases caused by Deoxiribole Nucleic Acid (DNA). In order to identify mutation carriers and to analyze disease, researchers tend to find various gene determinations methods. Nowadays, Next Generation Sequencing (NGS) is emerging as a valuable and powerful platform to detect gene-based diseases by entiring human genome. In this study, we aimed to propose a bioinformatics application workflow to distinguish between insertions/deletions and somatic/germline mutations, by using NGS methods. We carried this study out on a data set containing 100 human genomes data (20 training, 80 testing) for the detection of Malignant Melanoma. We found that the results of diagnosis performance were 92.50% accuracy, 94.03% precision, 96.92% sensitivity and 95.45% F1 score. These results show the potential for proposed application based on NGS to improve Melanoma detection.

Keywords: *Next Generation Sequencing, Burrows-Wheeler Transform, Variant Calling, Malignant Melanoma*

1. INTRODUCTION

Melanoma is generally the most serious form of skin cancer. It tends to trigger genetic defects causing skin cells to proliferate and form malignant tumors when unrepaired Deoxyriba Nucleic Acid (DNA) damage occurs. These tumors are caused by pigment-producing melanocytes in the basal layer of the epidermis. When cancer begins to change and starts to grow, healthy cells form a mass called tumor and are out of control (Fig1). In general, about 8% of those diagnosed with new Melanoma have first-degree relatives with Melanoma. In a much smaller percentage of 1% to 2%, there are two or more close relatives having Melanoma.

Fig. 1. A general pattern of Malignant Melanoma

Familial Melanoma is a genetic or hereditary condition which means that the risk of Melanoma can pass from one generation to another in a family. To date, familial Melanoma has been primarily associated with two genes called cyclin-dependent kinase inhibitor 2A (CDKN2A) and cyclin-dependent kinase 4 (CDK4). A mutation (change) in one of these genes gives the person an increased risk of Melanoma and plays an important role in transferring the disease to future generations when it affects the genetic sequence. Variations such as dysplastic nevus (mole), sun exposure, freckled skin and red hair occur as risk factors in genes are associated with Melanoma tendency with known genetic mutations (Audibert *et al.*, 2018).

In order to identify mutation carriers in the melanocortin 1 receptor (MC1R) and CDKN2A genes, the fastest, most comprehensive and delicate method is so important for detecting this type of disease. Thus, researchers tend to find various gene determinations and to conduct any studies on mutation-based cancer cells associated with the genetic structure (Rihtman *et al.*, 2016). Nowadays, Next Generation Sequencing (NGS) is emerging as a valuable and powerful platform to detect gene-based diseases by entiring human genome. The simultaneous processing of each part of a DNA molecule separated into millions of pieces taken from a single sample is called NGS. The detection of the gene architecture behind the diseases can be analyzed by using gene mapping methods (Vogelstein *et al.*, 2013).

In this study, it was aimed to develop a bioinformatics application workflow to introduce both single nucleotide variants (SNV) and small diagnoses in order to obtain the molecular diagnosis of Melanoma by using NGS methods to distinguish between insertions/deletions and somatic/germline mutations.

2. METHODS

2.1. Experimental Data

In this study, an application workflow was created to prepare the data to be used to detect the tumor gene sequence. An Sh file (a scripted file for the Unix Bourne-Again type-SHell (Fileinfo (sh file), 2018)) is prepared using the Biotechnology National Information Center (NCBI) ftp servers. By the help of this Sh file, it is possible to download Melanoma and healthy reference samples from anywhere and anytime.

Sequence Read Archives (SRA) is the file type required to convert files in the toolkit. fastqz is used as a compression tool for FASTQ files. SRA and compressed files refer to softlinked, a symbolic path that represents the abstract position of another file (Cyberciti, 2018). A FASTQ record contains a sequence of quality scores for each nucleotide. Tab seperated value (TSV) files that store a data table in which data columns are separated by tabs are created in this way (Fileinfo (Tsv File), 2018). By this file, it is checked whether the sample examined has malignant tumor or not. After preparing the sample, the healthy reference sample is downloaded and converted to the required file types such as FASTA, fastfai, and dict. All required files are downloaded from Illumina's FTP servers, and then files are prepared to assist in targeting and locating the appropriate unit.

SRA tools dict, amp and files are created and the indexing process is also made from the FASTA file. Then, by Burrows-Wheeler Alignment (BWA) tools, the rules were created, the required files were taken. The sa file was created with Fastfai and Burrows-Wheeler Transformation (BWT). The pac file was created using the BWT and sa file. Indexing the genome to avoid any problems on NGS data is an important task. In order to speed up genomic searching, indexing operations on genome size, length and number can be done.

2.2. Next Generation Sequencing (NGS)

The sequencing logic behind NGS technology is similar to the Sanger Sequence. DNA polymerase strand accelerates the incorporation of genome-labeled nucleotides into the DNA template during DNA replication. During each cycle, the nucleotides are identified by fluorophore tags. Instead of sequencing a single piece of DNA, NGS executes this process in parallel between millions of DNA fragments and offers many advantages over Sanger Sequencing. Because NGS carries out multiple gene evaluation procedures in a single test, it eliminates the need for multiple tests to identify the causal mutation. This approach, while saving time, provides a more economical solution, reducing the risk of overlooking valuable clinical specimens. In addition, it provides about 5% high sensitivity in the detection of DNA mutations when compared with the existing methods (Kearse *et al.*, 2012).

NGS also allows zooming to the next target regions with the help of Ribo Nucleic Acid (RNA) sorting to discover the original RNA variants, allowing all genes to be sequenced quickly. Thus, gene expression analysis allows analysis of epigenetic factors, such as DNA methylation and DNA-protein interactions, for the study

of tumor types with digital messenger RNAs (mRNAs), gene-related cancer samples and somatic derivatives (Griffiths-Jones *et al*., 2006).

All NGS platforms require a library obtained by amplification or ligation with special adapter arrays. These adapter arrays provide a universal array of sequences for hybridization and sequencing pieces to the sorting cores, allowing the use of the library. By this sample preparation process, necessary data amounts for NGS were created and used. Each library fragment is raised to a solid surface by a covalently bonded DNA linker that hybridizes library adapters. This amplification process generates DNA clusters, each of which originates from a single library segment, and each cluster moves with an individual sequencing reaction. The order of each cluster is read optically from the repeated nucleotide loops. The raw data is obtained at the end of the sequencing run. This raw data is a collection of DNA sequences generated in each cluster. It may be necessary to perform further analysis to obtain meaningful results from these data (Zhang *et al*., 2011).

The use of molecular profiling exceeds the limitations of conventional solid tumor and surrounding tissue classification methods based on the morphology of tumor cells. Molecular profiling is critical to identify and characterize unique somatic mutations occured in cancer cells. Tumor profiling using NGS focuses on a preselected subset of genes / gene panel. These panels contain genes known to be involved in cancer and allow simultaneous evaluation of all potentially causing genes. Tumor profiling using NGS follows a simple workflow that can be easily scaled to hundreds of samples, allowing clinical laboratories to process and respond to samples. The targeted gene sequence analyzes more than one gene in a single assay. By optimizing the use of limited tissue samples by reducing the need for sequential testing, it provides accurate identification of rare variants in heterogeneous tumor samples (Kearse *et al*., 2012).

This study was started by downloading the mini version of the open source package management system and the environmental management system of Conda (Gao *et al*., 2017). Snakemake workflow management system was used to create transaction files. Snakemake workflows that are essentially the rules defined by extended scripts created with Python. These rules describe how to create output files using input files (Leipzig, 2016).

2.3. Proposed Application Pipeline

In this study, a multilayered architecture was used for the proposed application workflow (Fig. 2). This architecture contains multiple nested layers. Each rule has its own task identity, and many process steps work according to these task identities and threads. Faster operations were achieved by using Cloud systems. Working performance was three times higher than PCs.

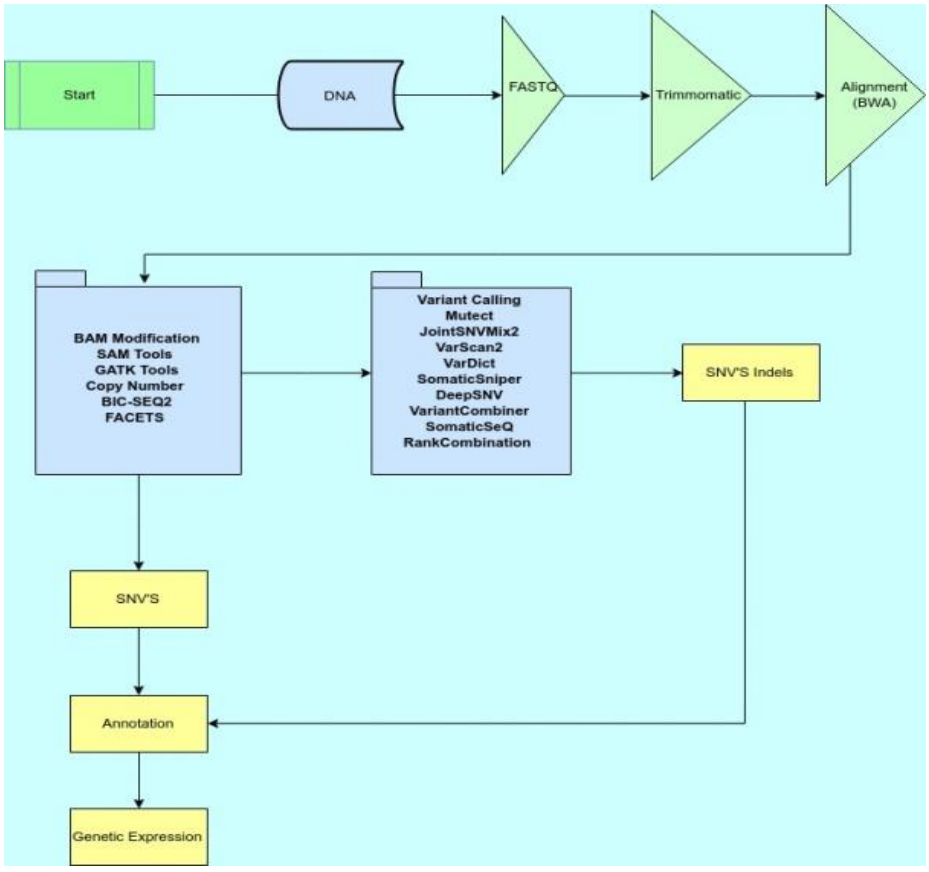


Fig.2. Proposed pipeline

BWA is a software package used to map different sequences corresponding to a large reference genome, such as the human genome, with the features such as long reading support and split alignment etc. (Michele *et al*., 2004). In this study, a BWA file was created and the indexing process was repeated.

To improve the overall quality of the skin data required for this study, cleaning and correcting the data were important pre-processing steps. Trimmomatic produces one or two DNA reading files by reading DNA from the left end and one from the right end, because a pair of reading files have been created using paired end mode (Fileinfo (Tsv File), 2018).Trimmomatic reference file called Truseq2pe was used to transmit fa.Fastq1 and Fastq2 files to reverse the fastq file. Clip trim rule was used to crop these two files and the reading file was created. For this purpose, there were two cut and soft connected fastq files. Additionally, there were two trimmed fastq files, which were fixed-linked files. After this procedure, the genomic analysis toolkit (GATK) was used for BAM modification. GATK is a programming framework designed and structured to facilitate the development of efficient and robust analysis tools for next-generation DNA sequencers using the MapReduce functional programming philosophy. GATK provides a small but rich set of data access patterns covering most of the needs of the analysis tool (McKenna *et al*., 2010).

Picard, a set of command-line tools, such as SAM / BAM / CRAM and VCF, were used to sort and process data and formats and process them with high efficiency (Ogasawara *et al*., 2016). SAM tools provide various utilities to sort, merge, index, create alignments in a preposition format, and handle alignments in Sam format (Li *et al*., 2009).

Using tracers at the end of the reading, which are below a certain threshold, create strips of cut quality adopted under the cropping threshold.

BWA-MEM is also used to deal with long reading and split alignment and is generally recommended for high-quality queries because it is faster and more accurate (Sipos *et al*., 2012).

Polymerase chain reaction (PCR) duplicates that are sets of pairs that have an unaligned start and an unaligned end and are suspected of non-independent measurements of a sequence. They are sampled from the same DNA template and violate the assumptions of the variant call. Acting as a real variant may lead to false positive variant calls, and errors in the sample are propagated to all pairs (Ebbert *et al*., 2016). Creation of SNPs and indentations are likely to be missed. Using many alignments, it may be needed to re-align the original readings to see that it is absolutely impossible to find it with a single read. This realignment-secondary alignment is part of the work of GATK haplotypecaller (Moore *et al*., 2010).

Trimmomatic-paired end-mode reading pairs can be protected, and library preparation to use additional information in better-matched readings can also be performed (Bolger *et al*., 2014).

A .txt file has been prepared to create quality data and quality mapping on the next process steps. In this file; genomic results containing BAM statistics include various statistics on genes with the total base amount found in external results and external window file types, how many of them are matched, how many of them have been processed separately, how many pairs are found in the DNA, and the quality of this data. Thus, a BAM file that has been fixed, sorted, merged and cleared from pairs has been sent to the Variant Calling phase.

2.3.1. Variant Calling

A variant call is a result of a nucleotide difference in a given position in a genome or transcriptome, usually accompanied by an estimate of the variant frequency and a measure of confidence. In this study, DNA-Seq variant calling methodology developed by Cornell Biotechnology Institute was used (Bukowski *et al*., 2017).

The presence of SNVs from NGS results that account for about 90% of all genetic variations between humans and the variant call is defined and a base is replaced by another base. SNV is only a variation in a single nucleotide without any frequency limitation. An SNP (single nucleotide polymorphism) is a single nucleotide change occured in more than 1% of a population of a single base pair difference from a reference. SNV is a special mutation while SNP is a shared mutation between a community. The variant calling is used to perform SNP genotyping, which identifies rare SNPs in a population, as well as to determine somatic SNV for each individual.

Variant calling algorithms are based on quality scores assigned to individual key calls in each widely read sequence. These scores are estimations per base of the error propagated by the sorting machine but are produced scores subject to systematic technical errors from various sources. Basic quality score recalibration (BQSR) is a process in which machine learning is applied to empirically model these errors and adjust the quality scores. This increases the accuracy of variant calls and helps to obtain more accurate basic attributes (Hsu *et al*., 2017).

Single nucleotide polymorphisms due to point mutations and single nucleotide variations; structural variations may also occur due to deletion, replication, insertion, inversion and translocation. Copy number variants (CNVs) are submicroscopic structural variations that occur due to deletion, replication and replicator transposition. Inversion is a segment of DNA in which the orientation is reversed relative to the rest of the chromosome. Translocation is a local change in the position of the chromosome segment in a genome without any change in total DNA content. In addition, Segmental uniparental dysomy as uniparental disomy describes a case where a pair of homologous chromosomes or portions of a chromosome is derived from a single parent in a diploid individual (Bulmer, 1971).

There are two SNV classes in the literature which are constitutional and germline mutations. These mutations are inherited from the parents and are found in every cell and somatic mutations that occur during the life of an individual. When looking for a rare disease, germline variations of SNV are used. For diseases not included in the literature, the contribution of somatic mutations according to disease status is examined by comparing the tumor with the normal samples.

Realign InDels (insertion/deletion) method describes the possible indices in all readings and the results given against these targets. This step is required to avoid false

readings because some of the indices are misaligned during the initial alignment (Sun *et al*., 2017).

BAM Recalibration method is used to recalibrate the basic quality scores of the sequence-synthesis readings in an aligned BAM file. The quality scores of each reading in the BAM files after recalibration are more accurate because the reported quality score reduces the likelihood of misinterpreting the reference genome (McCormick *et al*., 2015).

VarScan; Illumina, SOLiD, Life/Pgm, Roche/454 etc. are platform-independent mutation search tools for targeted exhaust and whole genome reordering data. VarScan uses an intuitive / statistical approach to search for variants that meet the reading depth, basic quality, variant allele frequency, and desired thresholds (Koboldt*et al*., 2009). VarScan calls using a statistical test of somatic variants (SNPs and indices) based on an intuitive method and the number of aligned reads that support each allele. In somatic mode, VarScan reads the pileup files (linked file type) from the normal file and the tumor simultaneously (Koboldt *et al*., 2013).

BCF tools manipulate variant calls in Variant Call Format (VCF) and binary equivalent Bam-Calibration File (BCF). All commands work clearly with both uncompressed and Bgzf compressed VCFs and BCFs (Walker *et al*., 2014).

MuTect method was developed at Broad Institute for reliable and accurate identification of somatic point mutations in the NGS data of genomes (Larson *et al*., 2011). Somatic Sniper aims to identify single nucleotide positions that are different between tumor and normal BAM files. It takes a normal BAM file and compares these two files to determine the differences (Kroigard *et al*., 2016).

Strelka method for somatic SNV helps identifying small indices by sequencing the data of paired tumors and normal samples. In addition, this method utilizes a new Bayesian approach, which represents the continuous allele frequencies for both tumor and normal samples, while utilizing the normal expected genotype structure (Saunders *et al*., 2012). Strelka2 serves a rapid and accurate variant search tool for the analysis of germline variation in small cohorts and somatic variation in tumor / normal samples. Thus, data types can be found faster and safer (Koboldt *et al*., 2013).

2.3.2. Single Nucleotide Polymorphism Calling (SNP Calling)

NGS data, that commonly produced by genetic and genomic studies, attaches great importance to correct calling of SNP and genotypes. In NGS methods, a whole genome of the human or random parts of the targeted regions are randomly added to its short readings that are sequenced and then aligned or assembled to a reference genome. 'SNP calling', which aligns the DNA fragments of one or more persons to the reference genome, is used to identify variable fields. SNP calling is intended to determine which positions are polymorphisms or at least one of the bases is different from the reference sequence (Nielsen *et al*., 2011). In this study, "Hands-on Tutorial on SNP Calling" source from Plant Genome and Systems Biology Group/Pgsb was used (Haberer *et al*., 2018).

2.3.3. Genome Annotation

Genome annotation on nucleotide sequence carries out an important process to encode the gene positions and functions of a genome and to identify non-coding regions. Analysis of the DNA sequence by genome annotation software tools allows the discovery and mapping of genes, exons-introns, regulatory elements, repeats, and mutations (Peter *et al*., 2009). Human intervention is still necessary in order to produce better and meaningful results because this is an NP-hard problem. Annovar is the tool that users use to functionally identify various genes, explain genetic variants, and update information. Online Genome Scanner, hosted by the University of California, Santa Cruz, incorporates human genomes including mouse, worm, fly, yeast data as well as human genome reference names / versions of hg18-hg19-hg38. A list of variants with chromosomes starting position, end position, reference nucleotide and observed nucleotides can be accessed (Yang *et al*., 2015).

Gene-based explanation identifies whether SNPs are the most common genetic diversity among humans, while the region-based explanation describes variants in specific genomic regions. A filter-based annotation defines variants documented in specific databases. In addition, from the exhaust data, identifying the candidate gene list for Mendelian diseases, it helps to collect the nucleotide sequence at any user-specific genomic positions.

2.4. Evaluation of Classification Performance

In order to measure the classification performance of proposed application, confusion matrix was used. A confusion matrix contains information about actual and predicted results obtained by a classification system. Performance of such systems is commonly evaluated using the data in the matrix. The following table shows the confusion matrix for a two class classifier.

Table 1. Confusion Matrix

		Predicted Class	
		Yes	**No**
Actual Class	**Yes**	TP	FN
	No	FP	TN

Accuracy, precision, recall and F1 score evaluation criteria can be calculated with the formulas Eq. (1), Eq. (2), Eq. (3) and Eq. (4) respectively, taking all the correct classified samples into account.

$$\textbf{Accuracy} = \frac{\textbf{TP+TN}}{\textbf{TP+TN+FN+FP}} \quad (1)$$

$$\textbf{Precision} = \frac{\textbf{TP}}{\textbf{TP+FP}} \quad (2)$$

$$\textbf{Recall} = \frac{\textbf{TP}}{\textbf{TP+FN}} \quad (3)$$

$$\textbf{F1 Score} = \textbf{2} * \frac{\textbf{Precision}*\textbf{Recall}}{\textbf{Precision+Recall}} \quad (4)$$

Where, TP = true positive, TN = true negative, FP = false positive, and FN = false negative.

3. RESULTS AND DISCUSSION

We designed a pipeline with open-source technologies to investige a practical and up-to-date study in this field. In this study, 100 human genome data was used. 20% of the data was used for system training; 80% of the data was used as the test data. Table 2 shows the distribution of these data by gender, average age in patient or healthy groups.

Table 2. Human genome data

Test Data			
Group	**Male**	**Female**	**Age (Mean)**
Patient	40	25	30
Healthy	5	10	25
Total	45	35	27.5

Evaluation of classification data with accuracy, precision, recall and F1 score criteria is shown in Table 3.

Table 3. The results of classification performance of the proposed application

Classification Performence			
Accuracy	Precision	Recall	F1 Score
92.50%	94.03%	96.92%	95.45%

The results show that the proposed application for the determination of Melanoma has very well classification performance. We used DNA mapping and gene fragment detection for determination of Malignant Melanoma like previous studies (Laila *et al*, 2016; Andrea *et al*., 2018). However, we have also individually determined the disease in the target patients by creating a personal set of rules. We found higher accuracy rates (92,5%) than previous studies (Laila *et al*., found 70% accuracy rate and Andrea et al found 66% accuracy rate) when compared in classification performance.

To identify the tumor gene sequence, an application workflow was created using NGS technologies and a reference sample. Since this system is running on the cloud systems simultaneously, it provides a faster environment and larger DNA mapping capacity when compared to PC calculations. Thus the performance of the study was greatly increased. In previous studies, only mutated genes were detected but the results of the studies were not evaluated. Within the scope of this study, a workflow which is thought to be shared as open source and which is able to classify the gene mutations in all types of cancers in the human genome, is proposed. The evaluation of the developed workflow with various criteria and the comparison results with similar studies on the same disease in the literature provide promising and evident results.

There are some limitations of this study. Our data set has limited availability. Larger and more diverse testing data sets from patients with various genetic backgrounds could help on confirming our results.

3. CONCLUSION

The results of our study demonstrate that NGS is capable of a highly accurate diagnostic classification of Melanoma. We aimed to support this research area by the strength of the computer science and make the processes faster and more efficiently, and in the future to transform it into an expert system and bring a new breath to the field of medicine. In the future, it is expected that the modust computer-aided methods will entirely support clinical decisions and clinical practices through including all molecular genetic factors in diagnosis and treatment process of Melanoma, by analyzing DNA maps.

REFERENCES

Andrea, F., Franz, J.H., Tobias, S., *et al*. (2018). "Next-generation-sequencing of advanced Melanoma: Which genetic alterations have an impact on systemic therapy response?" *J Clin Oncol*, Vol.36, No.15, suppl. e21557-e21557.

Audibert, C., Stuntz M., and Glass, D. (2018). "Treatment Sequencing in Advanced BRAF-Mutant Melanoma Patients: CurrentPractice in the United States", *Journal of Pharmacy Technology*, Vol.34, No.1, pp. 17–23.

Bolger, A.M., Lohse M., Usadel B. (2014). "Trimmomatic: a flexible trimmer for Illumina sequence data", *Bioinformatics*, Vol. 30, No.15, pp. 2114-2120.

Bukowski, R, Guo, X, Lu Y, et al. (2017) “Construction of the third-generation Zea mays haplotype map.” *Gigascience*. Vol. 7, No. 4, pp. 1-12.

Bulmer, M. G. (1971). "The Effect of Selection on Genetic Variability," *The American Naturalist*, Vol.105, No. 943, pp. 201-211.

Ebbert, M.T., Wadsworth, M.E., Staley, L.A., et al. (2016) "Evaluating the necessity of PCR duplicate removal from next-generation sequencing dat aand a comparison of approaches", *BMC Bioinformatics*. Vol.17, No. 7, pp. 239.

Gao, J., Wan C., Zhang H., et al.(2017) "Anaconda: AN automated pipeline for somatic COpyNumber variation Detection and Annotation from tumor exome sequencing data*", BMC Bioinformatics*, Vol.18, No.1, pp:436.

Griffiths-Jones, S., Grocock, R. J., Dongen, S., Bateman, A., Enright, A. J. (2006). "miRBase: microRNA sequences, targets and gene nomenclature", *Nucleic Acids Research*, Vol.34, No.1, pp.140–144.

Haberer, G., Spannagl, M., "Hands-on Tutorial on SNP Calling" Plant Genome and Systems Biology Group/PGSB (Access Date: 01.11.2018)

Hsu, Y.C., Hsiao, Y.T., Kao, T.Y., Chang, J.G., Shieh, G.S. (2017). *Detection of Somatic Mutations in Exome Sequencing of Tumor-Only Samples*", Scientific Reports 7, 15959.

Kearse, M., Moir, R., Wilson, A., et al. (2012). "Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data." *Bioinformatics*, Vol.28, No.12, pp. 1647–1649.

Koboldt, D. C., Steinberg, K. M., Larson, D. E., Wilson, R. K., & Mardis, E. R. (2013). "The next-generation sequencing revolution and its impact on genomics." *Cell*, Vol.155, No.1, pp. 27-38.

Koboldt, D.C., Chen, K, Wylie, et al. (2009). "VarScan: variant detection in massively parallel sequencing of individual and pooled samples." *Bioinformatics* Vol. 25, No. 17, pp. 2283-2285.

Kroigard, A.B.,Thomassen M., Lænkholm A-V., Kruse T.A., Larsen M.J. (2016) "Evaluation of Nine Somatic Variant Callers for Detection of Somatic Mutations in Exomeand Targeted Deep Sequencing Data." *PLoS One*. Vol.11, No.3, e0151664.

Laila, N., Kelsey, M., Jordan, R., (2016). "Cytology Sample Based Next-Generation Sequencing for Metastatic Melanoma: A Feasible and UsefulTool", *Journal of the American Society of Cytopathology*, Vol.5, No.5, p. 67.

Larson, D.E., Harris C.C., Chen K., et al. (2011). "Somatic Sniper: identification of somatic point mutations in whole genome sequencing data." *Bioinformatics*. Vol.28, No.3, pp. 311-7.

Leipzig, J. (2016). "A review of bioinformatic pipeline frameworks", *Brief Bioinform*. Vol.18, No. 3, pp. 530-536.

Li, H., Handsaker, B., Wysoker, A., et al. (2009). "The Sequence Alignment/Map format and SAM tools", *Bioinformatics*. Vol. 25, No. 16, pp. 2078-2079.

McCormick, R.F., Truong, S.K., Mullet, J.E. (2015). "RIG: Recal ibration and inter relation of genomic sequence data with the GATK", *G3 (Bethesda)*. Vol.5, No. 4, pp. 655-665.

McKenna, A., Hanna, M., Banks, E., et al. (2010). "The Genome Analysis Toolkit: A Map Reduce framework for analyzing next-generation DNA sequencing data." *Genome Research*. Vol.20, No. 9, pp. 1297-1303.

Moore, J.H., Asselbergs, F.W., Williams, S.M. (2010) "Bioinformatics challenges for genome-wide association studies." *Bioinformatics*, Vol.26, No.4, pp. 445–455.

Nielsen, R., Paul, J. S., Albrechtsen, A., Song, Y. S. (2011). "Geno type and SNP calling from next-generation sequencing data" *Nat Rev Genet*. Vol.12, No. 6, pp. 443-451.

Ogasawara, T., Cheng, Y., Tzeng, T-H.K. (2016) "Sam2bam: High-Performance Framework for NGS Data Preprocessing Tools." *PLoS One*. Vol. 11, no.11,. e0167100.

Peter, J. A., Cock, T.A., Jeffrey, T., et al. (2009). "Biopython: freely available Python tools for computational molecular biology and bioinformatics." *Bioinformatics*, Vol.25, No. 11, pp. 1422–1423.

Rihtman, B., Meaden, S., Clokie, M.R., Koskella, B., Millard, A.D. (2016). "Assessing Illumina technology for the high-through put sequencing of bacteriophage genomes." *PeerJ*. Vol. 4, e2055.

Saunders, C.T., Wong, W.S., Swamy, S., Becq, J., Murray, L,J,, Cheetham, R.K. (2012). "Strelka: accurate somatic small variant calling from sequenced tumor-normal sample pairs." *Bioinformatics*, Vol.28, No. 14, pp. 1811-1817.

Sipos, B., Massingham, T., Stütz, A.M., Goldman N. (2012) "An Improved Protocol for Sequencing of Repetitive Genomic Regions and Structural Variations Using Muta genes is and Next Generation Sequencing." *PLoS One* Vol. 7, No. 8, e43359.

Sun, Z., Bhagwate, A., Prodduturi, N., Yang, P., Kocher, J. P.A. (2017) "Indel detection from RNA-seq data: tool evaluation and strategies for accurate detection of actionable mutations", *Briefings in Bioinformatics*, Vol. 18, No. 6, pp. 973–983.

Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S, Diaz LA, Kinzler KW. (2013). "Cancer genome landscapes." *Science*. Vol. 339, No. 6127, pp. 1546-1558.

Walker, B.J., Abeel, T., Shea, T., et al. (2014) "Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement." *PLoSOne*. Vol. 9, No. 11, e112963.

Yang, H., Wang, K. (2015). "Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR", *NatProtoc*. Vol.10, No.10, pp. 1556-66.

Zhang J., Chiodini, R., Badr, A., Zhang, G. (2011). "The impact of next-generation sequencing on genomics." *J Genet Genomics*. Vol. 38, No. 3, pp.95-109.

Turkish Journal of Engineering

Turkish Journal of Engineering (TUJE)
Vol. 3, Issue 3, pp. 127-132, July 2019
ISSN 2587-1366, Turkey
DOI: 10.31127/tuje.479458
Research Article

AN INTEGRATED APPROACH BASED ON THE TAGUCHI METHOD AND RESPONSE SURFACE METHODOLOGY TO OPTIMIZE PARAMETER DESIGN OF ASBESTOS-FREE BRAKE PAD MATERIAL

Funda Kahraman [1] and Banu Sugözü *[2]

[1] Tarsus University, Technology Faculty, Mechatronics Engineering Department, Mersin, Turkey
ORCID ID 0000 – 0003 – 4333 – 4943
fundakahraman33@hotmail.com

[2] Mersin University, Engineering Faculty, Mechanical Engineering Department, Mersin, Turkey
ORCID ID 0000 – 0002 – 7798 – 2677
banusugozu@mersin.edu.tr

* Corresponding Author
Received: 06/11/2018 Accepted: 05/01/2019

ABSTRACT
Braking pressure and sliding velocity can affect wear behavior of brake pad material, so sliding velocity and braking pressure is often regarded as the key factor in defining quality performance of this material. For this reason, Taguchi Method has been used to screen the variables that have significant effects on the wear rate. Furthermore, an integrated approach of Taguchi Method and Response Surface Methodology has been employed to improve an empirical model for predicting wear rate and optimization of control factors. It has been observed that an integrated approach based on Taguchi Method and Response Surface Methodology gives 15.45% lower wear rate as compared to optimal results obtained from Taguchi method.

Keywords: *Taguchi method, Response Surface Methodology, Analysis of Variance, Wear*

1. INTRODUCTION

The brake friction material composition (resin, fibers, fillers, friction modifiers, abrasives) plays a crucial role in getting the effective brake pad performance. Besides size, type and amount of added reinforcement, other important parameters influencing the tribological properties are the manufacturing conditions (pressure, temperature, time), and testing conditions (pressure, speed, temperature). The brakes require friction composites composed of eco-friendly components. The friction composites also have stable and high friction coefficient, low wear rate, low cost and no noise. This is because play important roles in brake performance such as disc wear, stopping distance, pedal fell and brake induced vibrations (Aleksendrić and Barton, 2009). It usually includes 10-15 different raw materials to carry out comfortable brake performance at wide range of sliding speeds, humidity and applied pressure (Aleksendric and Duboka, 2006). The researchers have been investigated friction and wear behaviors of brake system to observe the effect of components in the brake composites. A considerable number of articles related to brake pad materials have been published (Sugozu *et al.*, 2018, Gurunath and Bijwe, 2007, Sugozu *et al*., 2016, Bijwe *et al.*, 2012, Sugozu *et al*., 2014a, Kumar *et al*., 2012, Sugozu *et al*., 2014b, Kukutschová *et al*., 2010, Djafri *et al*., 2014, Sugözü, 2015, Ertan and Yavuz, 2010). Recently, several researchers have been used statistical methods such as artificial neural networks, grey relational analysis, Taguchi design method, fuzzy logic method, response surface methodology, genetic, differential evolution and particle swarm algorithms to determine optimum brake performance of different brake pad materials (Aleksendrić and Barton, 2009, Aleksendric and Duboka, 2006, Zaharudin *et al*., 2011, Xiao and Zhu, 2010, Ficici *et al*., 2014, Mutlu, 2009, Anoop *et al*., 2009, Singh *et al*., 2014).

In this study, the effect of sliding velocity and braking pressure on the dry wear behavior of asbestos free brake pad material has been investigated by using integrated approach of Taguchi Method (TM) and Response Surface Methodology (RSM). The optimal results obtained from TM-RSM integrated approach and TM have been compared and discussed. Significant improvement in wear rate has been obtained by using TM-RSM integrated approach.

2. EXPERIMENTAL PROCEDURES

In this study, solid-state technique was used to prepare the brake pad materials (15% Phenolic resin, 10% Steel Fibers, 15% Copper particles, 10% Cashew dust, 4% Brass Particle, 10% Graphite, 29% Barite, 7% Alumina). Brake pad specimens were manufactured approximately 25.4 mm in diameter by a conventional procedure. All constituents were weighed and then mixed for 3–4 min for obtaining a homogeny dispersion. The mixture was hot pressed at 150 °C and 15 MPa load for 10 min. The experimental study was completed by following the procedure shown in Fig. 1. Wear tests were performed under unlubricated condition using a pin on disc tester with GG 20 gray cast iron as the counter face with a hardness of 200 HB and the average surface roughness of 1.50 µm. Contact temperatures between samples and disc were measured as 50-175 °C during wear test. Specific wear rate was calculated by using Eq. (1).

$$v = \frac{m_1 - m_2}{2 \cdot \pi \cdot R_d \cdot n \cdot f_m \cdot \rho} \quad (1)$$

where, v is specific wear rate (cm^3/Nm), m_1 is mass of brake pad before testing (g), m_2 is mass of brake pad after testing (g), n is total revolution, R_d is radius of disc (m), f_m is average friction force (N), ρ is density of brake pad (g/cm^3) (TS 555, 1992).

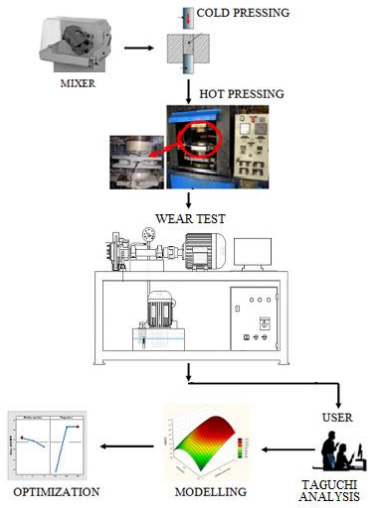


Fig. 1. A schematic diagram of the experimental study

Sliding velocity and braking pressure has been selected as input parameters, and wear rate has been selected as output parameter. Input parameters and their levels have been shown in Table 1. Taguchi analysis was performed by using Minitab 17 software.

Table 1. Factors and their levels for Taguchi design

Factors	Levels		
	Level 1	Level 2	Level 3
Braking pressure (*kPa*)	500	600	700
Sliding velocity (*m/s*)	6	7.5	9

RSM with CCD was used for designing and analyzing experiment by using The Design of Expert 7 software. CCD was used to obtain a mathematical model of wear rate (cm^3/Nm) as a function of the sliding velocity (m/s) and braking pressure (kPa). The optimal value of each parameter obtained from Taguchi analysis has been considered as a central value as "0" for the CCD matrix (Yadav, 2017). The range of each parameter was coded in five levels (−1.41, −1, 0, 1, 1.41). The levels of input factors are shown in Table 2. The relationship between the actual and coded values of factors was calculated by using Eq. (2).

$$x_1 = \frac{S - 500}{100} \qquad x_2 = \frac{V - 9}{3} \quad (2)$$

where x_1 is braking pressure, x_2 is sliding velocity.

Table 2. Factors and their levels for CCD

Factors/Levels	-1.41	-1	0	+1	+1.41
Braking pressure (*kPa*)	360	400	500	600	641
Sliding velocity (*m/s*)	5	6	9	12	13

3. RESULTS AND DISCUSSION

3.1. Optimization of Wear Rate Using Taguchi Method

TM is a very effective statistical technique to deal with responses affected by number of variables. TM has been used to analyze and optimize single performance characteristic of manufacturing process. TM uses specially constructed tables named as "orthogonal array" to design of experiments and using of these orthogonal arrays reduces the number of experiments. Three categories for S/N ratios is suggested depending on the type of characteristic; smaller is better, nominal is the best, and higher is better.

Since lower wear rate, wear loss, surface roughness, tool wear and cutting force are desirable, smaller values are always preferred in this case (Asiltürk and Akkuş, 2011, Phadke, 1995).

In the present study, the S/N ratio was chosen according to the criterion the smaller the better, in order to minimize the response. The loss function of the smaller the better for response was calculated using Eq. (3).

$$\eta = -10\log\left[\frac{1}{n}\sum_{i=1}^{n} y_i^2\right] \tag{3}$$

where η is the S/N ratio, n is the number of trials, y_i is the experimental value of the i th trial in experiments.

Table 3. Orthogonal array of Taguchi for wear rate and results

Trial no	A (*kPa*)	B (*mm/s*)	*v* (cm^3/Nm)	S/N (*dB*)
1	500	6	0.116	18.71
2	500	7.5	0.113	18.94
3	500	9	0.110	19.17
4	600	6	0.190	14.42
5	600	7.5	0.180	14.89
6	600	9	0.172	15.28
7	700	6	0.255	11.87
8	700	7.5	0.250	12.04
9	700	9	0.248	12.11

The experiments were performed according to L_9 array based on Taguchi method. Each factor level combination was repeated three times and a total of 27 tests were concluded and the mean values were reported. Experimental conditions and results have been presented in Table 3. The average test results for the wear rate values were transformed into a signal-to-noise (S/N) ratio by using Eq. (2) (Table 3). The mean S/N ratio for each level of the control factors for wear rate was calculated and listed in Table 4. Mean of S/N ratios variation for each level has been shown in Fig. 2. It is seen from Fig. 2 and Table 4 that the first level of A factor, and the third level of B factor are higher and therefore the combination of parameters is A_1B_3. Consequently, the optimum wear condition of the brake pad material for wear rate was determined 500 kPa for the braking pressure and 9 mm/s for the sliding velocity.

Table 4. Mean S/N ratios (dB) for wear rate

Control factors/ Levels	1	2	3	Δ	Rank
A	18.94	14.86	12.01	6.93	1
B	15.00	15.29	15.52	0.43	2

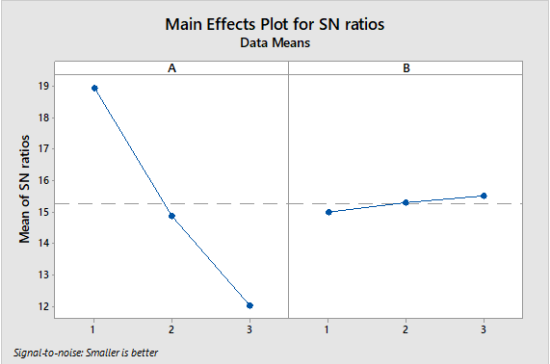


Fig. 2. Mean of S/N graph for wear rate

Further, relative effect of each design parameter is determined by the ANOVA test. Table 5 shows the ANOVA values for the experimental results. The most significant parameter affecting on wear rate was determined that braking pressure (99.28%) at 95% confidence level. Furthermore, F ratios and their percent contribution were taken into consideration to identify the significance level of the control variables.

Table 5. ANOVA results for wear rate

Source	A	B	Error	Total
Degrees of freedom	2	2	4	8
Sum of squares	0.02857	0.00016	0.00004	0.02877
Mean square	0.01428	0.00008	0.00001	
F-value	1266.63	7.16		
P	0.000	0.048		
Contribution (%)	99.28	0.562	0.156	100

3.2. Confirmation Test

Confirmation test was performed to verify experimental conclusions using optimal level of control variables. The predicted value of S/N ratio can be computed by Eq. (4). Table 6 shows that the comparison of the initial, actual and predicted values.

$$\hat{\eta} = \eta_m + \sum_{i=1}^{j}(\eta_i - \eta_m) \tag{4}$$

where $\hat{\eta}$ is the predicted value of S/N ratio, η_m is the

total mean of S/N ratio, η_i is the mean of S/N ratio, j is the number of the main design parameters the affect the quality characteristic.

The predicted optimal wear rate (v_p) was calculated Eq. (5) by considering individual effects of the factors (A_1B_3), and their levels. Optimal wear rate was computed as 0.118 cm^3/Nm. Percentage error (%) was determined as 7.27 for the actual wear rate, 0.83 for the S/N ratio. The improvement in wear rate (38.88%), mean quality loss (28.74%) has been obtained as compared to the initial set value of control variables. Based on the confirmation experiment results, wear rate decreased 1.64 times.

$$v_p = T_v + (A_1 - T_v) + (B_3 - T_v) \quad (5)$$

where Tv is the wear rate total mean value.

Table 6. Results of the confirmation experiments

		Level	v (cm^3/Nm)	S/N (dB)
Initial combination		A_2B_2	0.180	14.89
Optimal combination	Experiment	A_1B_3	0.110	19.17
	Prediction	A_1B_3	0.118	19.33
Error (%)		-	7.27	0.83
Improvement (%)			38.88	28.74

In practice, the quality losses between initial and optimal combination are calculated by Eq. (6). It was calculated as 37% for wear rate. Thereby, quality losses for the wear rate was reduced 63% by using Taguchi method.

$$\frac{L_{opt}(y)}{L_{ini}(y)} = \left[\frac{1}{2}\right]^{\Delta\eta/3} \quad (6)$$

where $L_{opt}(y)$ and $L_{ini}(y)$ are optimal and initial combinations, respectively. Δ_η is the difference between S/N ratios of optimal and initial combinations.

3.3. Prediction of Wear Rate Using RSM Technique

RSM is a statistical technique for the modeling and analysis of problems in which a response of interest is influenced by several variables and objective is to optimize this response. CCD is an experimental design in RSM for developing a predictive model for the response variable. This design consists of a factorial portion and axial portion and a central point (Myer and Montgomery, 2002, Sagbas *et al*., 2009).

The first order and second order model polynomial model can be expressed by the general Eq. (7) and Eq. (8), respectively.

$$y = \beta_0 + \sum_{j=1}^{k} \beta_j x_j + \varepsilon \quad (7)$$

$$y = \beta_0 + \sum_{j=1}^{k} \beta_i x_j + \sum_{j=1}^{k} \beta_{ij} x_j^2 + \sum_{i}^{k-1} \sum_{j}^{j} \beta_{ij} x_i x_j + \varepsilon \quad (8)$$

where y: response, β_0: constant, β_1 ,…, β_k: regression coefficient, x_1, x_2 ,…, x_j: input factors, k: number of input factors, i=1,2, …, k-1 and j=1,2, …, k, Ɛ: random error

Sliding velocity and braking pressure were considered as model variables and wear rate as a response variable. The Design Expert 7 software was used for designing and analyzing experiment. CCD was used to obtain a mathematical model for wear rate (cm^3/Nm) as a function of the sliding velocity (m/s) and braking pressure (kPa). Experimental levels for input parameters were selected according to a CCD. The plan of experiments was made by randomizing the experiments to avoid accumulation of errors. The experiments were conducted based on randomized run number. Three replications of each factor level combinations were conducted and average values were reported.

The predicted and experimental (actual) wear rate values of based on the CCD rotatable design are presented in Table 7.

The relationship between wear rate and control factors is expressed as Eq. (9).

$$v = 0.11 + 0.032 x_1 + 0{,}013 x_1^2 \quad (9)$$

where; v is the estimated wear rate, x_1 is the coded factor that represents the braking pressure.

Table 7. Actual and predicted wear rate (cm^3/Nm) values

Std	Run	P (kPa)	V (m/s)	Actual v $*(10^{-6})$	Predicted $v*(10^{-6})$	% Error
13	1	500	9	0.120	0.110	8.33
10	2	500	9	0.112	0.110	1.78
5	3	360	9	0.085	0.093	9.41
6	4	640	9	0.175	0.180	2,85
12	5	500	9	0.110	0.110	0.00
4	6	600	12	0.162	0.150	7.40
1	7	400	6	0.110	0.100	9.09
11	8	500	9	0.111	0.110	0.90
7	9	500	5	0.118	0.130	10.16
3	10	400	12	0.100	0.093	0.16
2	11	600	6	0.180	0.170	5.55
9	12	500	9	0.115	0.110	4.34
8	13	500	13	0.105	0.110	4.76

ANOVA has been performed to test the applicability of improved model for the experimental data fitted in the model or not. Results of ANOVA for wear rate is presented in Table 8. Results indicate that x_1 and x_1^2 are significant model terms. Other model term is not significant. This insignificant model term is removed in developed model.

Table 8. ANOVA results for improved model

Source	SS	DF	MS	F value	Prob>F
Model	9.79E-003	5	1.96E-003	20.33	0.0005 significant
x_1	8.40E-003	1	8.40E-003	87.19	< 0.0001*
x_2	2.69E-004	1	2.69E-004	2.79	0.1388
x_1^2	1.09E-003	1	1.09E-003	11.30	0.0121*
x_2^2	7.40E-005	1	7.40E-005	0.77	0.4098
$x_1 x_2$	1.60E-005	1	1.60E-006	0.17	0.6959
Residual	6.75E-003	7	9.64E-005		
Lack of fit	6.09E-004	3	2.03E-004	12.46	0.1690 not significant

*significant factors at 5% significance level

As seen from the Table 8, braking pressure has the largest F-value which indicates the stronger influence on performance characteristics. Similar data is also provided by the contribution (%) in Table 5.

The F calculated value is greater than F table value ($F_{0.05,2,5}$=5.79) and hence the quadratic model can be said quiet adequate. Also, mean absolute percentage error (MAPE) and coefficient of correlation (R^2) were used to the developed model as the performance criterion. It was found to be as 4.98% for the wear rate. The value of Adj-R^2 between predictive values and experimental results is obtained 89% for the wear rate. This values showed that the developed quadratic model fit well with experimental results and it can be significantly used to predict wear rate.

3.4. Optimization Using TM-RSM Integrated Approach and Desirability Functions

After building the regression model, a numerical optimization technique using desirability functions analysis (DFA) described by Derringer and Suich can be used to optimize the response. Single response optimization determines how input parameters affect desirability of individual response. The objective of optimization is to find the best settings that minimize a particular response. Desirability ranges from zero to one for any given response. A value of one represents the ideal case, while zero indicates that one or more responses fall outside the desirable limits (Myer and Montgomery, 2002). DFA was also used control variables to find the best combination of control variables in the combined approach of TM-RSM. In this analysis, the goal used for the wear rate is "minimize" and the goal used for the control factors are "within range". Desirability value was obtained as 0.92. Consequently, the optimum wear rate 0.093 cm^3/Nm was obtained at 400 kPa for the braking pressure and 11 mm/s for the sliding velocity

Comparison of optimization techniques have been presented in Table 9. It has been observed that DFA based on TM-RSM integrated approach gives (15.45%) lower wear rate than result obtained from TM.

Table 9. Comparison of optimization techniques

Optimization technique	Wear rate (cm^3/Nm)		Absolute Error (%)
	Optimal	Predicted	
Taguchi approach	0.110	0.118	7.27
TM-RSM integrated approach	0.093	0.091	2.15
Improvement (%)	15.45	-	-

4. CONCLUSION

This study focuses on TM and TM-RSM integrated approach for investigating of the wear behavior of the asbestos-free brake pad materials. Following conclusions based on this study can be drawn from modeling and optimizing:

• The percentage contribution of braking pressure is as (99.28%), which follows by sliding velocity (0.562%).

• Braking pressure is obtained the most significant factor affecting wear rate by ANOVA.

• The percentage improvement in quality loss is (28.74%) at optimal parameters combination obtained by TM as compared to the initial set parameters.

• The developed quadratic model capable to predict the wear rate with MAPE (4.98%).

• The optimal combination obtained by TM is as braking pressure 500 kPa and sliding velocity 9 mm/s.

• The optimal combination obtained by TM-RSM integrated approach is as braking pressure 400 kPa and sliding velocity 11 mm/s.

• Optimum wear rate value obtained by DFA based on TM-RSM integrated approach is better (15.45%) than as compared to optimum value of TM approach.

REFERENCES

Aleksendrić, D., Barton, D.C. (2009). "Neural network prediction of disc brake performance." *Tribology International*, Vol. 42, No. 7, pp. 1074-1080.

Aleksendric, D., Duboka, Č. (2006). "Prediction of automotive friction material characteristics using artificial neural networks-cold performance." *Wear*, Vol. 261, No. 3-4, pp. 269-282.

Anoop, S., Natarajan, S., Babu, S.K. (2009). "Analysis of factors influencing dry sliding wear behaviour of Al/SiC p–brake pad tribosystem." *Materials & Design*, Vol. 30, No. 9, pp. 3831-3838.

Asiltürk, I., Akkuş, H. (2011). "Determining the effect of cutting parameters on surface roughness in hard turning using the Taguchi method." *Measurement*, Vol. 44, No. 9, pp. 1697-1704.

Bijwe, J., Aranganathan, N., Sharma, S., Dureja, N., Kumar, R. (2012). "Nano-abrasives in friction materials-influence on tribological properties." *Wear*, Vol. 296, No. 1-2, pp. 693-701.

Djafri, M., Bouchetara, M., Busch, C., Weber, S. (2014). "Effects of humidity and corrosion on the tribological behaviour of the brake disc materials." *Wear*, Vol. 321, No. 1, pp. 8-15.

Ertan, R., Yavuz, N. (2010). "An experimental study on the effects of manufacturing parameters on the tribological properties of brake lining materials." *Wear*, Vol. 268, No. 11-12, pp. 1524-1532.

Ficici, F., Durat, M., Kapsiz, M. (2014). "Optimization of tribological parameters for a brake pad using Taguchi design method." *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, Vol. 36, No. 3, pp. 653-659.

Gurunath, P.V., Bijwe, J. (2007). "Friction and wear studies on brake-pad materials based on newly developed resin." *Wear*, Vol. 263, No. 7-12, pp. 1212-1219.

Kukutschová, J., Roubíček, V., Mašláň, M., Jančík, D., Slovák, V., Malachová, K., Filip, P. (2010). "Wear performance and wear debris of semimetallic automotive brake materials." *Wear*, Vol. 268, No. 1-2, pp. 86-93.

Kumar, K.R., Mohanasundaram, K.M., Arumaikkannu,

G., Subramanian, R. (2012). "Analysis of parameters influencing wear and frictional behavior of aluminum–fly ash composites." *Tribology Transactions*, Vol. 55, No. 6, pp. 723-729.

Mutlu, I. (2009). "Artificial neural network modelling of non-asbestos brake lining performance boric acid in brake pad." *Information Technology Journal*, Vol. 8, No. 3, pp. 389-402.

Myer, R.H., Montgomery, D.C. (2002). *Response surface methodology: process and product optimization using designed experiment*, John Wiley and Sons, New York, USA.

Phadke, M.S. (1995). *Quality engineering using robust design*, Prentice Hall PTR, Upper Saddle River, NJ, USA.

Sagbas, A., Kahraman, F., Esme, U. (2009). "Modeling and predicting abrasive wear behaviour of poly oxy methylenes using response surface methodolgy and neural Networks." *Metalurgija*, Vol. 48, No. 2, pp. 117-120.

Singh, K.H., Kumar, A., Kumar, R. (2014). "Optimization of Quality and Performance of Brake Pads Using Taguchi's Approach." *International Journal of Scientific & Engineering Research*, Vol. 5, No. 7, pp. 632-639.

Sugözü, İ. (2015) "Investigation of using rice husk dust and ulexite in automotive brake pads." *Materials Testing*, Vol. 57, No. 10, pp. 877-882.

Sugozu, I., Can, I., Oner, C. (2014). "The effect of borax on the friction performance of an automotive brake lining." *Materials Testing*, Vol. 56, No. 5, pp. 362-368.

Sugozu, I., Can, I., Oner, C. (2014). "Investigation of using Calabrian pine cone dust and borax in brake pads." *Industrial Lubrication and Tribology*, Vol. 66, No. 6, pp. 678-684.

Sugozu, I., Mutlu, I., Sugozu, K.B. (2016). "The effect of colemanite on the friction performance of automotive brake friction materials." *Industrial Lubrication and Tribology*, Vol. 68, No. 1, pp. 92-98.

Sugozu, I., Mutlu, I., Sugozu, K.B. (2018). "The effect of ulexite to the tribological properties of brake lining materials." *Polymer Composites*, Vol. 39, No. 1, pp. 55-62.

TS 555 (1992). *Highway Vehicles-Brake System-Brake Pads for Friction Brake*, Turkish Standard Institute, Ankara, Turkey.

Xiao, G., Zhu, Z. (2010). "Friction materials development by using DOE/RSM and artificial neural network." *Tribology International*, Vol. 43, No. 1-2, pp. 218-227.

Yadav, R.N. (2017). "A hybrid approach of Taguchi-Response Surface Methodology for modeling and optimization of Duplex Turning process." *Measurement*, Vol. 100, No. 1, pp. 131–138.

Zaharudin, A.M., Talib, R.J., Berhan, M.N., Budin, S., Aziurah, M.S. (2011). "Taguchi method for optimising the manufacturing parameters of friction materials." *Proceedings of Regional Tribology Conference*, RTC, Langkawi Island, Malaysia, pp. 246-251.

Turkish Journal of Engineering

Turkish Journal of Engineering (TUJE)
Vol. 3, Issue 3, pp. 133-139, July 2019
ISSN 2587-1366, Turkey
DOI: 10.31127/tuje.490509
Research Article

MECHANICAL AND WEAR CHARACTERISATION OF QUARRY TAILINGS REINFORCED A6063 METAL MATRIX COMPOSITES

Stephen Durowaye *[1], Olatunde Sekunowo [2], Babatunde Bolasodun [3], Isaac Oduaran [4] and Ganiyu Lawal [5]

[1] University of Lagos, Faculty of Engineering, Department of Metallurgical and Materials Engineering, Lagos, Nigeria
durosteve02@yahoo.com, sdurowaye@unilag.edu.ng
ORCID ID 0000-0003-4787-5675

[2] University of Lagos, Faculty of Engineering, Department of Metallurgical and Materials Engineering, Lagos, Nigeria
osekunowo@unilag.edu.ng
ORCID ID 0000-0003-6787-5688

[3] University of Lagos, Faculty of Engineering, Department of Metallurgical and Materials Engineering, Lagos, Nigeria
bbolasodun@unilag.edu.ng
ORCID ID 0000-0002-2720-5933

[4] University of Lagos, Faculty of Engineering, Department of Metallurgical and Materials Engineering, Lagos, Nigeria
ioduaran@unilag.edu.ng
ORCID ID 0000-0003-4433-2704

[5] University of Lagos, Faculty of Engineering, Department of Metallurgical and Materials Engineering, Lagos, Nigeria
glawal@unilag.edu.ng
ORCID ID 0000-0003-1452-4270

* Corresponding Author
Received: 30/11/2018 Accepted: 08/01/2019

ABSTRACT
The viability of utilizing cheaply sourced quarry tailings (QTs) particles as reinforcing materials for the development of low cost high performance aluminium matrix composites for structural and industrial applications was explored. Aluminium alloy (A6063) was reinforced with 2 – 10 wt. % of 80 µm QTs particulates using stir casting method. Microstructural, mechanical and wear characterisations were carried out on the composites. Micro hardness and ultimate tensile strength of the composites were improved with addition of reinforced particulate in the base matrix alloy. Addition of quarry tailings reinforcements also decreased the wear rate of the composites from 245 mm/Nm^3 to 180 mm/Nm^3 for 10 N and from 260 mm/Nm^3 to 135 mm/Nm^3 for 15 N. Optical and SEM examinations revealed uniform distributions of the QT particles in the molten matrix.

Keywords: *Aluminium Matrix Composites, Quarry Tailings, Characterisation*

1. INTRODUCTION

Aluminium still remains the most utilised metal as matrix material in the development of metal matrix composites (MMCs). This is due to the desirable property combination which it exhibits thereby making it very suitable for a wide range of engineering applications. Aluminium matrix composites (AMCs) have been explored for varied technical uses due to the high specific strength and stiffness, good wear and corrosion resistance, good mechanical and thermal properties exhibited among others (Alaneme and Bodurin, 2011; Alaneme and Sanusi, 2015).

Silicon carbide, alumina, boron carbide, titanium carbide and tungsten carbide are among the mostly used conventional ceramic materials in the development of AMCs. However, high cost and limited supply of conventional ceramic reinforcing materials especially in the developing countries are among the problems associated with the development of discontinuously reinforced aluminium matrix composites (Bodurin *et al.*, 2015). Hence, studies have been carried out to find inexpensive and effective substitutes to the relatively high cost conventional ceramic particulates with the hope of still maintaining their high performance level in service applications. Aluminium was reinforced with natural resource by-products such as fly ash, silica sand, and red mud to produce AMCs for properties enhancement (Surappa *et al*., 2008; Rohatgi *et al*., 2010).

A large number of composite materials have been developed by reinforcing metallic matrix alloy with high strength, high modulus and brittle ceramic phase particles. Uniform dispersion of reinforcement materials in the metal matrix offers improvement in strength elastic modulus, corrosion and wear resistance of resultant composites (Durowaye *et al*., 2018). Due to these properties, particulate reinforced composites have found wide range of applications in the automobiles and aerospace industries compared to monolithic alloys. Particulate reinforced composites have been considered as suitable alternatives to traditional un-reinforced monolithic alloys.

The search for low cost options in AMCs production has led to a number of efforts tailored at utilizing industrial and agro waste products as reinforcing materials. This has led to the development of hybrid composites consisting of an agro-waste derivative combined with synthetic ceramic materials or an industrial waste combined with synthetic reinforcement (Bodurin *et al.*, 2015). However, the final properties of the composites depend on individual properties of reinforcement and matrix alloy. There are lots of agro waste that are being studied but the limitation to their use is the preparation which involves drying, burning, conditioning, sieving before use as reinforcements and these processes increase the production time, cost and affect the production rate.

One of the industrial by-products that is readily available is quarry tailings (QT). QT is generated from quarrying activities and is obtained as solid waste during crushing of stones to obtain aggregates. It exhibits low density and contains ceramic compounds such as silica (SiO_2) and alumina (Al_2O_3). It is available in commercial quantity at very low cost and studies on its use as reinforcement have not been extensive. QT has been researched on as partial replacement of cement in concretes but very little work has been carried out concerning its use as reinforcement for AMCs production. It has been used in the production of lightweight building materials because of its low density. It also offers the possibility of providing cost competitive reinforcement in composite production because the cost of procuring 500 grams of silicon carbide (SiC) is twice that of QT. Most of the limitations of producing particulate reinforced AMCs could be remedied with the use of quarry tailings. Hence, this study explores the use of QT particles as suitable reinforcement and substitute to conventional ceramic materials to produce aluminium based composite with high performance levels.

2. MATERIALS AND METHODS

2.1. Materials

The A6063 alloy used for this study was obtained from the Nigerian Aluminium Extrusion Company (NIGALEX) Oshodi, Lagos, while the quarry tailings used as the reinforcement was obtained from a quarry company in Akure, Ondo state, Nigeria. The picture of the pulverized and sieved quarry tailings is shown in Fig. 1. The chemical compositions of the aluminium alloy and the quarry tailings were determined using the X-ray fluorescence (Mini Pal 4 ED-XRF) machine and are presented in Tables 1 and 2 respectively.

Fig. 1. Picture of 80 µm quarry tailings particles

Table 1. Chemical composition of A6063 alloy

Element	Si	Fe	Cu	Mn	Mg	Ti	Ni	Zn	Al
Weight (%)	0.366	0.351	0.026	0.013	0.514	0.014	0.0053	0.120	98.56

Table 2. Chemical composition of quarry tailings

Compound	SiO_2	Al_2O_3	Fe_2O_3	CaO	MgO	K_2O	TiO	BaO
Weight (%)	64.97	14.86	6.27	4.92	2.68	3.09	1.14	0.001

2.2. Materials Preparation and Composite Samples Production

The A6063 alloy was cut into small sizes for ease of weighing. The quarry tailings were pulverized in a ball mill and sieved to 80 μm particle size. The quantities of the A6063 alloy to be melted were calculated based on the size of the mould that was used. The measurement was done using a digital weighing balance model UW1020H. The proportion of the materials mixture is presented in Table 3. The quarry tailings particles were preheated in an oven at a temperature of 250^0 C for an hour in order to remove moisture and improve wettability with the molten A6063 metal alloy. The A6063 alloy was melted at 750^0 C in a gas fired crucible furnace. The molten metal alloy was allowed to cool to 600^0 C to reach a semi-solid state. Stirring of the melt was carried out at this temperature for about 2 minutes before pouring appropriate quantities (2, 4, 6, 8 and 10 wt. %) of the preheated quarry tailings into the molten alloy and also stirred for about 5 minutes. The composites slurry was afterwards heated to 800^0 C and stirred at 300 rpm for 10 minutes before pouring into a cylindrical metallic mould of diameter 16 mm and length 60 mm. Manual stirring was done using a long stainless steel rod to avoid clustering and to achieve good dispersion of the particles in the melt. The cast samples (Fig. 2) were allowed to cure at room temperature for 24 hours after which they were removed from the moulds.

2.3. Microstructural Characterisation

The microstructure and the chemical compositions of the phases present in the test samples were examined using an X-SUPREME 8000 by Oxford instruments, USA.

2.4. Tensile and Ductility Characterisation

The test samples were prepared using qualilathe-210–CNC lathe machine and an Instron universal testing machine was used to determine the tensile property of the samples at room temperature based on ASTM 8m-15a standard. circular tensile test piece with one step grip was used with dimension of 5 mm diameter and gauge length of 30 mm. The samples were stressed to fracture while the data generated was used to evaluate the tensile property. the elastic deformation behaviour of the specimens was evaluated by the percentage elongation suffered during tension.

2.5. Hardness Characterisation

The hardness behaviour of the composites was determined using Wolpert micro hardness tester 930 in accordance with ASTME 384 standard. A direct load of 5 kg was then applied on each of samples for about 10 secs. The micro hardness tester was placed on the sample to make indentation at three different positions in the sample and the average reading values were selected.

Table 3. Proportion of the materials mixture

Samples	Aluminium		Quarry tailings		Total (g)
	(g)	(wt. %)	(g)	(wt. %)	
1	384.15	100	-	-	384.15
2	376.47	98	7.69	2	384.16
3	366.60	96	15.27	4	381.87
4	356.82	94	22.78	6	379.60
5	347.12	92	30.18	8	377.30
6	337.51	90	37.51	10	375.02

Fig. 2. Picture of some of the composite cast samples

2.6. Impact Characterisation

A V-notch was cut on each sample, using a Hounsfield notching machine ensuring that the notch screw is set at a depth of 2 mm so that the cutter just touches the piece. Each test piece was broken with a pendulum on the Hounsfield balanced impact machine and the energy absorbed in fracturing is measured. The test was performed on different samples of the composite and average value of each result were taken.

2.7. Wear Characterisation

The wear characterisation of the composites was evaluated using a Din abrasion tester FE05000. Cylindrical pin samples of 10 mm diameter and 20 mm height were machined from the 16 mm rod. The surface of each sample was prepared with 220 grit SiC abrasive paper. The test was carried out using applied loads of 10 N and 15 N, sliding speed of 0.32 m/s, sliding distance of 1 m and a constant time of 2 minutes. The rubbing action between the sample and the abrasive wheel driving the rotation motion of the machine resulted in the generation of loosed composites debris from the samples surface. The weight of the samples was measured before and after each test using a digital weigh balance with model number UW102H and the wear rate was determined using ''Eq. (1)'' (Agunsoye *et al*., 2018).

$$R = \frac{\Delta W}{SA} \quad (1)$$

where,

R is the wear rate, ΔW is the weight difference of the sample before and after the test in mg, S is total sliding distance (m) and A is the applied load (N).

3. RESULTS AND DISCUSSION

3.1. Composition and Microstructure

As shown in Table 1, the major constituent of A6063 alloy is aluminium with 98.56 wt. % while other elements are in traces. The result also indicated that silica (SiO_2) has the highest composition of 64.97 wt. % in quarry tailings followed by alumina (Al_2O_3) with 14.86 wt. % while BaO has the least with composition of 0.001 wt. % indicating that it is a ceramic compound. Fig. 3 is the SEM micrograph of the unreinforced A6063 cast sample revealing pores and Al matrix. Fig. 4 shows the SEM micrograph of cast sample reinforced with 2 wt. % quarry tailings. It shows the uniformly distribution of quarry tailings particulates in the aluminum matrix. Fig. 5 shows the SEM micrograph of cast sample reinforced with 10 wt. % quarry tailings. It reveals a homogeneous distribution of quarry tailings particulates in the aluminum matrix. There is the presence of $Al_2O_3{\cdot}SiO_2$ (mullite) a hard ceramic compound in Figs. 4 and 5 which enhanced the hardness of the samples.

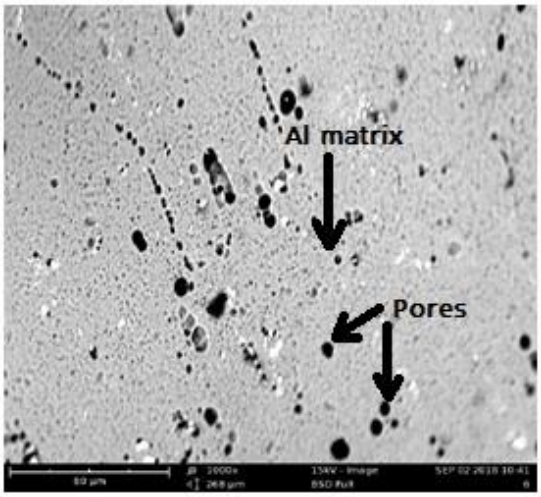


Fig. 3. SEM micrograph of unreinforced A6063 cast sample

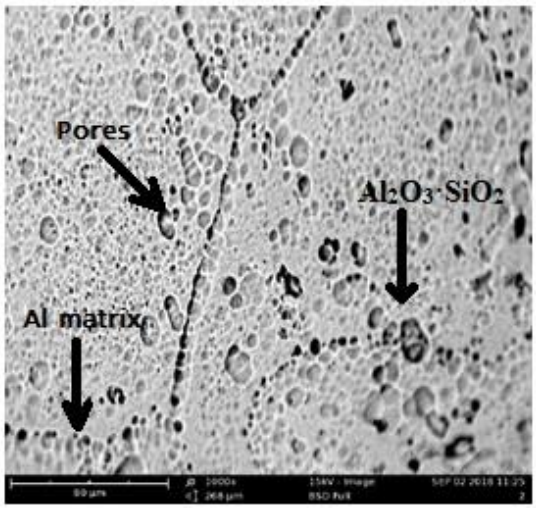


Fig. 4. SEM micrograph of cast sample reinforced with 2 wt. % quarry tailings

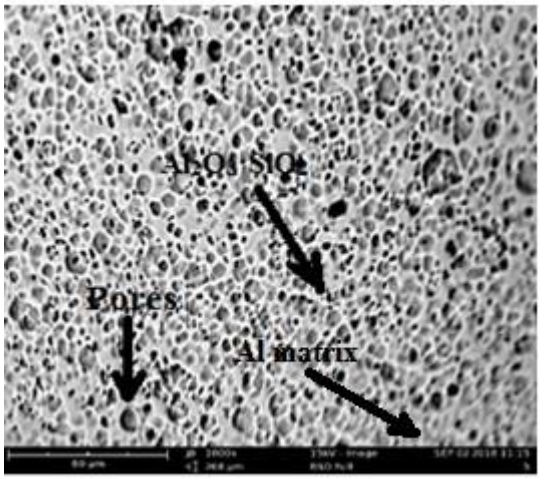


Fig. 5. SEM micrograph of cast sample reinforced with 10 wt. % quarry tailings

3.2. Tensile Strength

The tensile strength of the composites increased steadily with increase in quarry tailings addition as shown in Fig. 6.

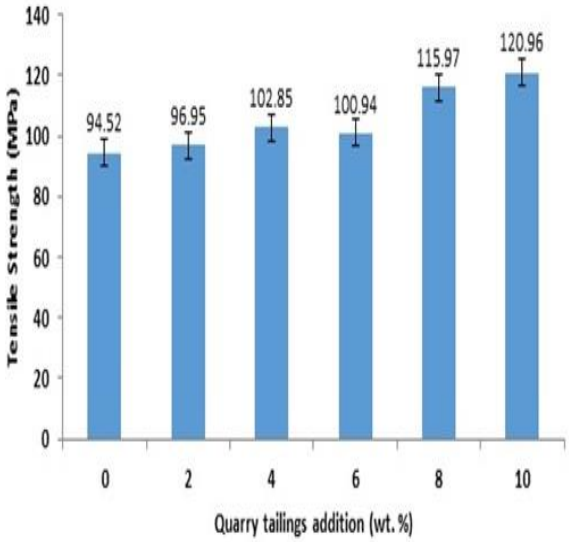


Fig. 6. Effect of quarry tailings on the tensile strength of the composites

The unreinforced A6063 cast sample exhibited the lowest tensile strength value of 94.52 MPa while sample 6 reinforced with 10 wt. % quarry tailings exhibited the highest tensile strength value of 120.96 MPa which is 28 % increase over the unreinforced. The increase in tensile strength is attributed to the applied load transfer to the strongly bonded quarry tailings reinforcements in Al matrix which increased dislocation density near matrix-reinforcement interface. Also, the grain refining strengthening effect as well as the uniform distribution of the quarry tailings particulates in the aluminum matrix enhanced the tensile strength of the composites.

3.3. Ductility

There was a steady decrease in elongation of the composites as the weight percentage of the quarry tailings increased as shown in Fig. 7. The unreinforced A6063 cast sample exhibited the highest elongation value of 16.442 % while sample 6 reinforced with 10 wt. quarry tailings exhibited the lowest elongation value of 6.057 %. The gradual reduction in ductility (elongation) of the composites was due to the resistance offered to the flow ability of aluminium matrix by the quarry tailings particulates as the weight of the quarry tailings particles increased. Similar results were obtained by Ramesh *et al*. (2013).

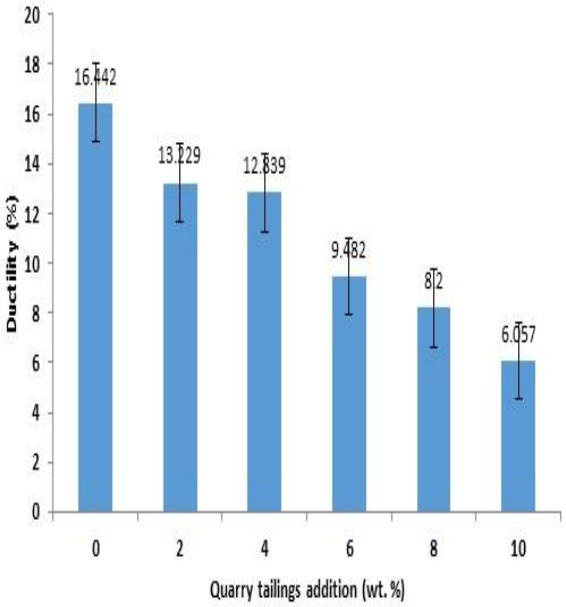


Fig. 7. Effect of quarry tailings additions on the ductility of the composites

3.4. Hardness

There was a steady increase in the hardness of the composites as the weight percentage of the quarry tailings increased as shown in Fig. 8. The unreinforced A6063 cast sample exhibited the lowest hardness value of 34.16 HV while sample 6 reinforced with 10 wt. % quarry tailings exhibited the highest hardness value of 49.95 HV. The presence of $Al_2O_3{\cdot}SiO_2$ (mullite) a hard ceramic compound from the quarry tailings offered resistance to surface plastic deformation during indentation. This rendered its inherent property of hardness to the soft matrix which significantly increased the hardness value. Again, the uniform dispersion of the quarry tailings particles and strong interfacial bonding between the reinforcement and the matrix alloy also enhanced the hardness of the composites. Similar observation was also reported by (Kalaiselvan *et al*., 2011).

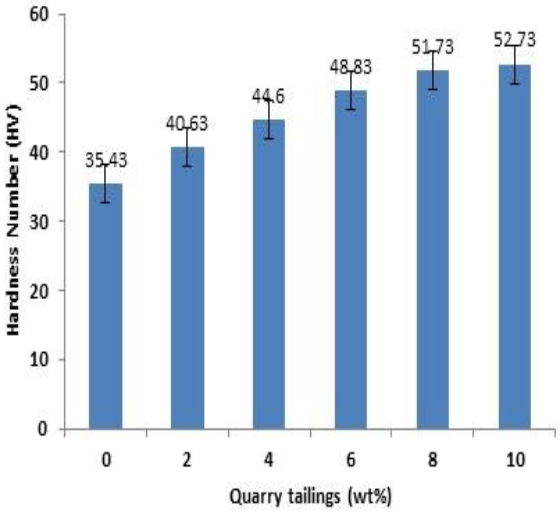


Fig. 8. Effect of quarry tailings additions on the hardness of the composites

3.5. Impact Energy

Fig. 9 shows a steady decrease in the amount of energy absorbed by the composites prior to fracture. The unreinforced sample exhibited the highest energy of 39.44 J while sample 6 reinforced with 10 wt. % quarry tailings exhibited the lowest impact energy value of 27.61 J. The gradual decrease in the amount of impact energy of the composites is due to the increased presence of hard and brittle quarry tailings particles in the soft Al-matrix alloy.

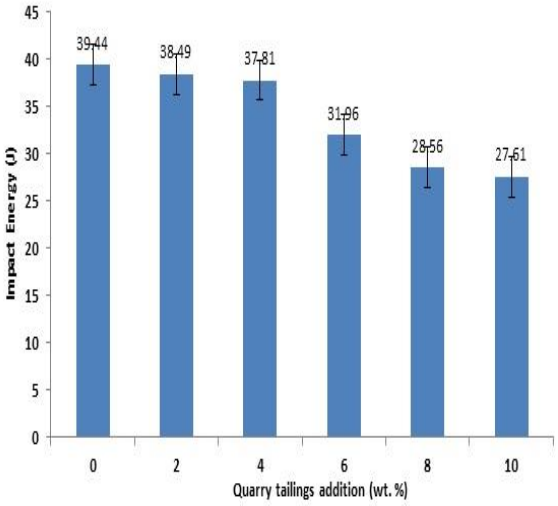


Fig. 9. Effect of quarry tailings additions on the impact energy of the composites

3.6. Wear Characteristics

As illustrated in Fig. 10, the unreinforced sample exhibited the highest wear rate of 245 mm/Nm^3 and 260 mm/Nm^3 at 10 N and 15 N respectively while sample 6 reinforced with 10 wt. % quarry tailings exhibited the lowest wear rate of 180 mm/Nm^3 and 225 mm/Nm^3 at 10 N and 15 N respectively.

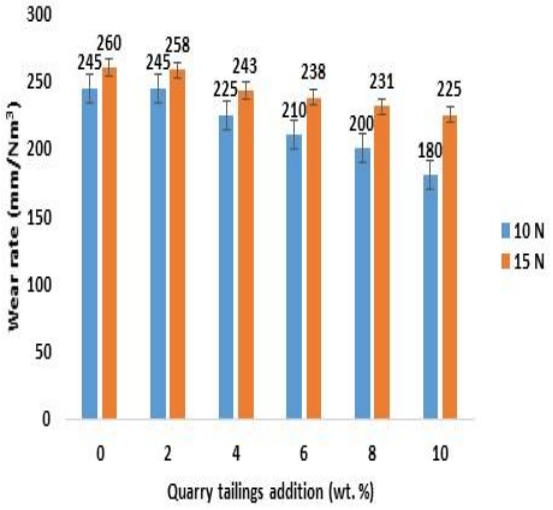


Fig. 10. Effect of quarry tailings on wear rate of the composites

The wear rate of the composites decreased with increasing quarry tailings addition which indicated the effect of concentration of quarry tailings reinforcement in reducing wear. This observation was similar to the one reported by Aigbodion and Akadike, (2010). Increased wear rate at higher load of 15 N was due to increased plastic deformation of the composites in the friction surface as a result of increase in the shear force between asperities and the ploughing force of asperities. Similar observation was also reported by Djafri *et al*. (2014) and Yawas *et al*. (2016).

The wear resistance of the composites is attributed to the presence of hard materials like alumina, silica and CaO in the quarry tailings. These acted as hard solid particles which improved the wear resistance. The strong interfacial bonding between the quarry tailings particles and the matrix alloy also enhanced wear resistance ability of the composites.

4. CONCLUSION

Quarry tailings reinforced A6063 composites have been successfully developed by the stir casting method. The 10 wt. % quarry tailings reinforced aluminium matrix composites exhibited desirable mechanical properties in terms of tensile strength (120.96 MPa) and hardness (49.95 HV). This composite also exhibited the low wear rate of 180 mm/Nm^3 and 225 mm/Nm^3 at 10 N and 15 N respectively. There was significant improvement in tensile strength and hardness of the composites with the addition of quarry tailings particulates. The uniform dispersion of the quarry tailings particles in the aluminium matrix coupled with the strong interfacial bonding between the quarry tailings particles and the matrix alloy enhanced the mechanical properties and wear resistance ability of the composites. This study has shown that quarry tailings has the potential of serving as reinforcement substitute for the expensive conventional ceramic materials in the development of AMCs.

REFERENCES

Agunsoye, J. O., Sefiu, A. B., Bamigbaiye, A., Odunmosu, K. A. and Akinboye, I. (2018). ‘‘Recycled ceramic tile composite for automobile applications, a comparative study with Nissan Jeep Cherokee brake pad.’’ *Engineering and Applied Science Research*, Vol. 45, No. 3, pp. 180-187.

Aigbodion, V. S. and Akadike, U. (2010). ‘‘Development of asbestos – free brake pad using bagasse.’’ *Tribology in Industry*, Vol. 32, No. 1, pp. 45–50.

Alaneme, K. K. and Bodurin, B. (2011). ‘‘Corrosion behavior of alumina reinforced aluminium (6063) metal matrix composites.’’ *Journal of Minerals & Materials Characterization & Engineering*, Vol. 10, pp. 96-100.

Alaneme, K. K. and Sanusi, K. O. (2015). ‘‘Mechanical and wear behaviour of rice husk ash alumina graphite hybrid reinforced aluminium based composites.’’ *Engineering Science and Technology International*, Vol. 18, pp. 416-422.

Bodunrin, M. O, Alaneme, K. K. and Chown, L. H. (2015). ‘‘Aluminium matrix hybrid composites: a review

of reinforcement philosophies; mechanical, corrosion and tribological characteristics.'' *Journal of Materials Research and Technology*, Vol. 4, pp. 434-445.

Djafri, M., Bouchetara, M., Busch, C. and Weber, S. (2014). ''Experimental study of the tribological behaviour of materials of brake.'' *Mechanika*, Vol. 20, No. 4, pp. 420 – 425.

Durowaye, S. I., Sekunowo, O. I., Lawal, G. I. and Orji, J. (2018). ''Effects of ZnO and graphite particulates on mechanical and wear characteristics of Al-Si alloy hybrid composites.'' *Usak University Journal of Engineering Sciences (UUJES)*, Vol. 1, No. 1, pp. 19-27.

Kalaiselvan, K., Murugan, N. and Parameswaran, S. (2011). ''Production and characterization of AA6061–B_4C stir cast composite.'' *Materials and Design*. Vol. 32, No. 7, pp. 4004-4009.

Ramesh, M., Karthikeyan, T., Kumaravel, A., Senthil Kumaran, S. and Raja V.L. (2012). ''Metallurgical and mechanical behavior of basalt particulate reinforced aluminium (A356) matrix composites.'' *European Journal of Scientific Research*, Vol. 95, pp. 524-532.

Rohatgi, P. K., Schultz, B. F, Daoud, A. and Zhang, W. W. (2010). ''Tribological performance of A206 aluminum alloy containing silica sand particles.'' *Tribology International*, Vol. 43, pp. 455-466.

Surappa, M. K. and Sudarshan, (2008). ''Dry sliding wear of fly ash particle reinforced A356 Al composites.'' *Wear*, Vol. 265, pp. 349-360.

Yawas, D. S., Aku, S.Y. and Amaren, S.G. (2016). ''Morpho*logy and Properties of Periwinkle Shell Asbestos Free Brake Pa*d.'' *Journal of King Saud University – Engineering Sciences*, Vol. 28, pp. 103–109.

Turkish Journal of Engineering

Turkish Journal of Engineering (TUJE)
Vol. 3, Issue 3, pp. 140-148, July 2019
ISSN 2587-1366, Turkey
DOI: 10.31127/tuje.467003
Research Article

MODELLING AND OPTIMIZATION OF THE EXAM INVIGILATOR ASSIGNMENT PROBLEM BASED ON PREFERENCES

Dilek Küçük Matcı *[1] and Ilgın Acar [2]

[1] Eskişehir Technical University, Research Institute of Earth and Space Sciences, Eskişehir, Türkiye
ORCID ID 0000-0002-4078-8782
dkmatci@anadolu.edu.tr

[2] Eskişehir Technical University, Department of Industrial Engineering Eskişehir, Türkiye
ORCID ID 0000-0001-9775-5386
ipoyraz@eskisehir.edu.tr

* Corresponding Author
Received: 03/10/2018 Accepted: 11/02/2019

ABSTRACT
The assignment process of invigilators to the exams could be difficult and time-consuming process when there are too many exam centers and the invigilator who will assign. In particular, due to the constraints of assignment according to the characteristics of the invigilators, keeping the assignment process taking into account the officer's preferences and the fact that the test centers have a certain capacity makes the problem very complicated. In this study, the situation was considered as an assignment problem. To solve this complicated situation, we formulated the problem, with the constraints of the capacity of exam centers, type of the invigilator, type of centers and other restrictions. An algorithm for model implementation was developed and automation suitable to the algorithm was written in C# language. The model was then applied to the case of exam invigilators assignment at Open Distance Faculty of Anadolu University. All the assignment process time is reduced from weeks to several minutes.

Keywords: *Examiner Assignment Problem, Optimization, Mathematical Modeling, Automated Assignment Systems*

1. INTRODUCTION

The assignment of personnel in appropriate positions in terms of their capabilities and characteristics increases staff satisfaction, optimizes resources utilization and increases performance. However, assigning personnel to appropriate positions in the context of a number of constraints is a complex problem. The assignment problems examine the distribution of a certain number of resources among the various targets. These problems have been subject to many scholarly works (Diaz 2001, Burkard 2002, Loiola, de Abreu *et al*. 2007, Oncan 2007, Pentico 2007, P. Tapkan 2008, K.Bhunia 2010, G. Liu 2011). The solution of problems in many different areas of life such as nurse assignment (Liang and Turkcan 2016, Wang, You *et al*. 2017), invigilator assignment (Hanum, Romliyah *et al*. 2015), airway fleet assignment (Subramanian, Scheff *et al*. 1994, Özener, Matoğlu *et al*. 2017), traffic assignment (Patriksson 2015, Poole and Kotsialos 2016), shift assignments(Hojati 2010) has been frequently researched in the literature.

The exam invigilator assignment problems, a type of assignment problem, are a problem that universities meet more than once each semester. There are several ways to solve this problem. One of these ways is the solving of the problem manually by a person who has a lot of knowledge and experience about the subject. In this case, the probability of making a mistake will be significant and the time to be spent according to the magnitude of the problem will also increase. Another solution method to the exam assignment problems is the using of mathematical models. In this case, the problem will be mathematical programming model, then the solution will be realized with the help of necessary algorithms. For example, Elizondo (Elizondo 1994) examined the problem of the assignment of n number of students to m number of schools in his study. The objective of their proposed model was to minimize the total distance. As a similar problem to invigilator assignment problem, Elloumi (Elloumi, Kamoun *et al*. 2014) studied the problem of the assignment of exams and classes. The objective of their model was to minimize the number of students that were not assigned. Karimzadeh and Zhai (Karimzadehgan and Zhai 2012) modeled a committee review team assignment problem as an integer programming problem and solved it by using an algorithm that was developed by them. In another study, a model was developed to assign an invigilator to exams for the Carleton University in Ottawa. The system combines special heuristics, genetic algorithm framework and ready software tools (Awad and Chinneck, 1998). Taha and Mansour (Mansour and Taha, 2015), used a solution for the proctor assignment problem based on the Bee Colony meta-heuristic algorithm. They stated that their proposed method maximized the preferences of proctors and balanced the workload of proctors.

In this study, we propose a new method and decision tool to assign invigilators to the exam centers. To accomplish this, an assignment system was designed that takes into account many constraints such as the invigilators' preferences, the detailed of previous exams and invigilators, the features of the centers.

The study started with the database creation process. This database contains information such as supervisors' titles, years of work, exams they have already assigned. While ensuring that the exams are carried out without errors, the comfort and satisfaction of the assigned invigilator are also taken into consideration. For this reason, to be re-created before each exam, a preference table consisting of centers where the invigilators would like to work is added to the database. In addition, since the characteristics of the examination centers are not the same in terms of transportation, comfort and social facilities, tables that are classified according to the characteristics of the centers are added to the database. In addition, the table has been added to the database to show which supervisors have been assigned to the centers in previous exams so that each supervisor can be equally assigned to different centers (Fig. 1). Then, a mathematical model was developed to provide the constraints of the problem. An algorithm that realizes this model has been developed and an automation written in C # language was presented to the university.

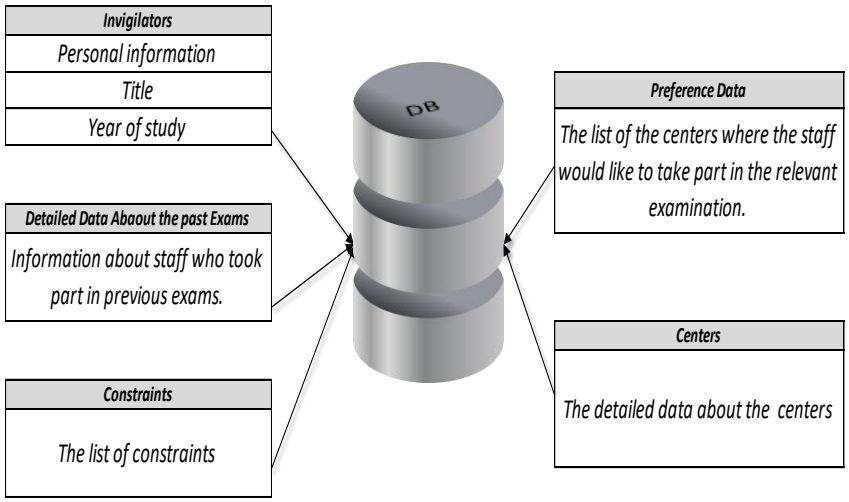


Fig. 1. The details of the created database

This study is an example of solving large-scale staff assignment problems that differ from standard supervisor assignment or exam scheduling problems. It provides very important tips on how to incorporate special cases that may be encountered in real life problems into models and how algorithms can be created in this direction. In the study, many special cases are mathematically modeled, such as the capacity constraints of the centers, detailed information about the invigilators to be assigned, the centers they prefer, the characteristics of the centers, the number of tasks the officials have previously received. Another difference from the other studies examined in the literature is the dimensional nature of the problem being studied in the study. The problem that is examined as an example is the process of assigning invigilator to examination centers to check exams that a university with an open teaching system has done throughout the country. Therefore, the characteristics of the test centers such as transportation, accommodation, number of schools to be audited are included in the appointment procedure.

2. MATERIAL AND METHODS

There are 1671 invigilators and 83 centers considered in this problem. The information on how many invigilators will be assigned to the center (the capacity of the center), the information about the invigilator (the title of the invigilator, the year of seniority, if the invigilator assigned to the center previously, etc.) are stored in the database. And the database in which the obtained data is stored, the created model and the assignment system are shown in Fig. 2.

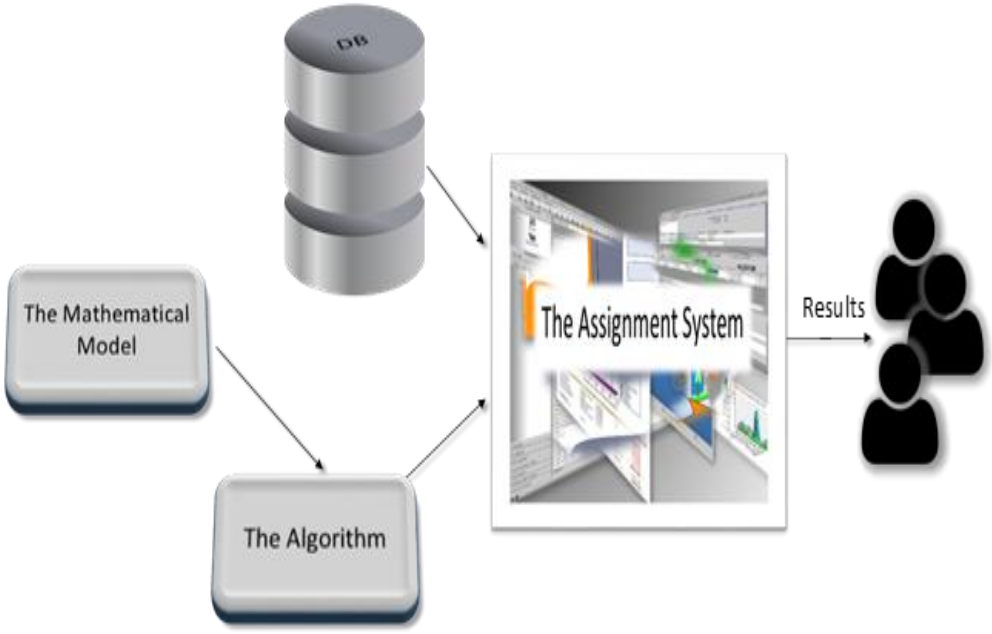


Fig. 2. The overview of the system

In this problem, all centers are divided into six groups according to their characteristics. The highest score was given to the group with the best features and the lowest score was given to the group with the worst features. In this sense, the other groups have scored from one to six. Then, in the past years, according to the information about which invigilators was assigned to which center, the score of the group where each center is located is determined as the score of invigilators. These points are used to give priority to the invigilators with low scores in order to perform a fair assignment. To reflect this in the mathematical model, the minimization of the total score of the invigilators is considered an objective function in the model. The assumptions of the problem are as follows:

- The invigilators should not be assigned to the center to which they were previously assigned;
- The invigilators that were assigned during the previous exam should not be assigned in the current exam.

2.1. The Proposed Mathematical Model

In this problem, the objective is to find the correct invigilator group in order to minimize the total score under the constraints. In the proposed model there are seven constraints, one decision variable(Ygm) and two indices that represent invigilators(g) and the centers(m).

Notation:

Model indices, parameters, and decision variables are given below:

Indices:

g=1…G, invigilator. There are G units of invigilators in the problem.

m=1…M, centers: There are M units of centers in the problem.

Parameters:

P_g : The total score of the invigilator.

$Pr_g = \begin{cases} 1, & \text{If the title of the invigilator g is proffesor.} \\ 0, & \text{otherwise} \end{cases}$

$\mathbf{A_g} = \begin{cases} \mathbf{1}, & \textbf{If the title of the invigilator g is r. assistant.} \\ \mathbf{0}, & \textbf{otherwise} \end{cases}$

S_m= The number of the invigilator that must be assigned to the centre *m*.

$\mathrm{T_{gm}}$ = The state of the invigilator *g* to choose the centre *m*.

B(g,m) = The state of whether the grouping of centres agrees with the groupings of the centres that invigilators were assigned previously.

$\mathbf{D_g}$ = The semester of the exam that the semester of the last assignment of the invigilator g.

Decision Variable:

$\mathbf{Y_{gm}} = \begin{cases} \mathbf{1}, & \textbf{If the inv. g is assigned to the center m} \\ \mathbf{0}, & \textbf{otherwise} \end{cases}$

Model

$$Min\ Z=\sum_{m=1}^{M}\sum_{g=1}^{G}\left(P_g * Y_{gm} * T_{gm} * B(g,m)\right) \quad (1)$$

$$\sum_{m=1}^{M} Y_{gm} \leq 1 \qquad (g=1,...,G) \quad (2)$$

$$\sum_{g=1}^{G} Y_{gm} = S_m \qquad (m=1,...,M) \quad (3)$$

$$\sum_{g=1}^{G} Pr_g Y_{gm} \geq 1 \qquad (m=1,...,M) \quad (4)$$

$$\sum_{g=1}^{G} A_g Y_{gm} \geq 1 \qquad (m=1,...,M) \quad (5)$$

$$\sum_{g=1}^{G} D_g Y_{gm} = 0 \qquad (m=1,...,M) \quad (6)$$

$$Y_{gm} \in \{0,1\} \quad (7)$$

The total score of the invigilator $\boldsymbol{P_g}$ is calculated with the data stored in the database. This data includes the centers where the officers were previously assigned and the scores of these centers.

The objective function (1) minimizes the total score of the invigilator that is chosen and assigned to center m. The parameter Tgm is used for preventing the assignment of invigilator that did not choose the center with a big value, such as 1000. Constraint (2) ensures that each invigilator cannot be assigned more than one center. In constraint (3), the number of invigilators that is assigned to the center must be equal to the capacity of the center. Constraint (4) states that at least one invigilator who is professor must be assigned to each center, while constraint (5) ensures that at least one invigilator who is research assistant must be assigned to each center. The constraints (4) and (5) ensures that (upon management's request), at least one experienced, and one inexperienced invigilator should be assigned to each center. The constraint (6) ensures that the invigilator cannot be assigned in the same period as the previous assignment.

2.2 Proposed Algorithm

In order to apply the constraints, we developed an algorithm. The general approach of the algorithm is to give priority to the invigilator with the smallest score.

The constraint (6) ensures with the first step of the algorithm. In this step, the procedure is starts with the selecting the invigilator whose last assignment period is different from the present exams period. To agree with the constraints (4) and (5), which stated that at least one professor and one research assistant must be assigned to each center, the algorithm consisted of three main parts; i.e., find a professor, find a research assistant, and find other invigilators. The invigilator assigned to a center with the help of the algorithm are removed from the processing pool. Thus, this striker is not assigned to another center. In this way, an attendant whose mathematical model is developed can be assigned to a maximum number of centers. Thus, the constraint (2) in the developed mathematical model is ensured.

Each main part has own three iterations. In Fig. 3, there is the algorithm to find a professor to the centers.

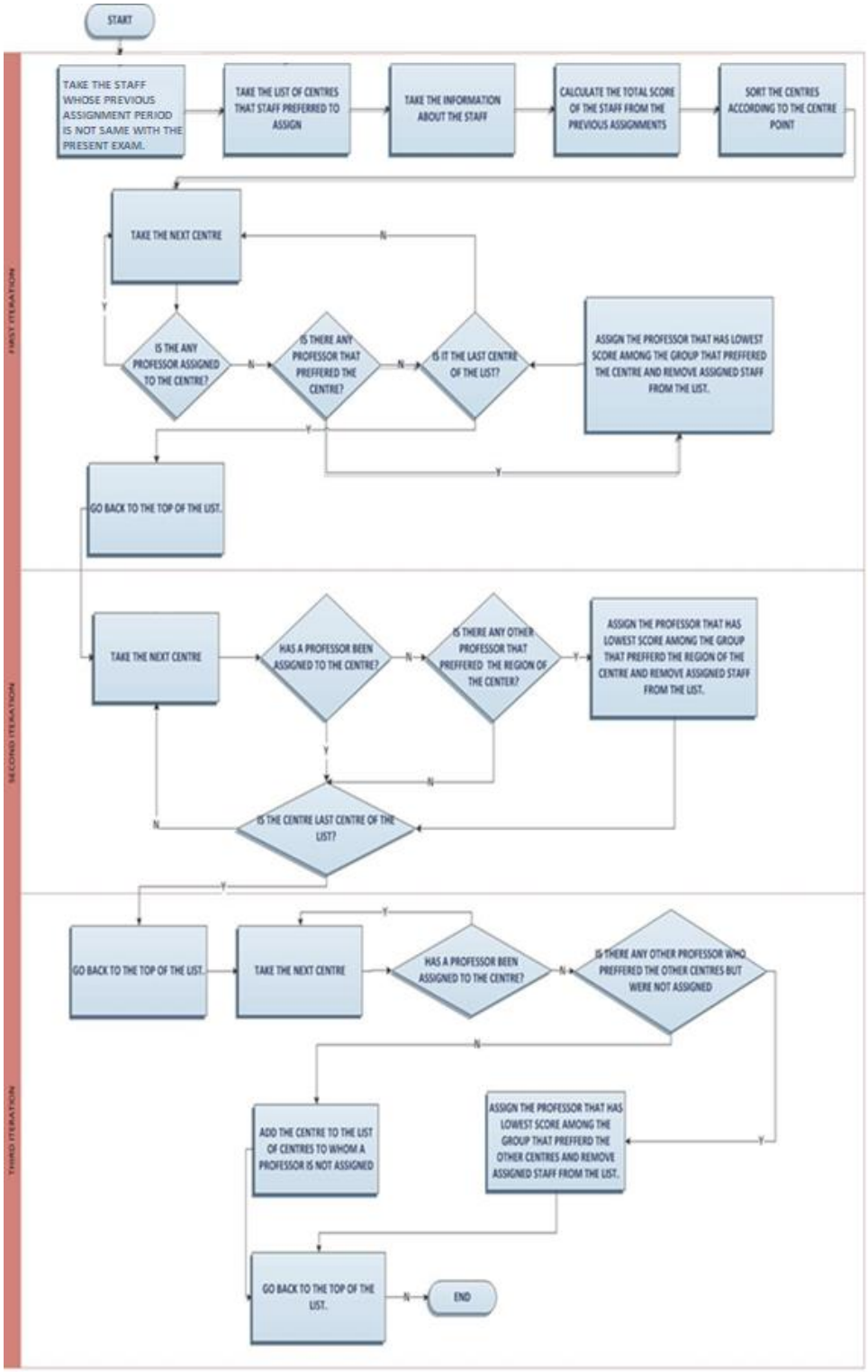


Fig 3. Find an Invigilator Who are a Professor

In the first iteration, a search was conducted for appropriate invigilators among the group that prefers that center, and this process was repeated for each center. In the second iteration, if an insufficient number of invigilators was assigned to the center, the appropriate invigilator was searched for among the invigilators that preferred another center in the same group as the original center but were not assigned it in the first iteration. This process was also repeated for each center. In the last iteration, if there was still an insufficient number of invigilator assigned to the center, the appropriate invigilator was searched for among the invigilators that were not assigned in the first two iterations.

After the assignment of invigilators who are a professor, the assignment of other invigilators who are the research assistant is performed. In Fig. 4, there is the algorithm to

find invigilators, who are the research assistant, to the centers. It is ensured that the assigned officers are removed from the transaction pool by the help of algorithm. In this way, an attendant whose mathematical model is developed can be assigned to a maximum number of centers. After the assignment of invigilators who are research assistant, the assignment of other invigilators is performed. In Fig. 5, there is the algorithm to find an invigilator to the centers.

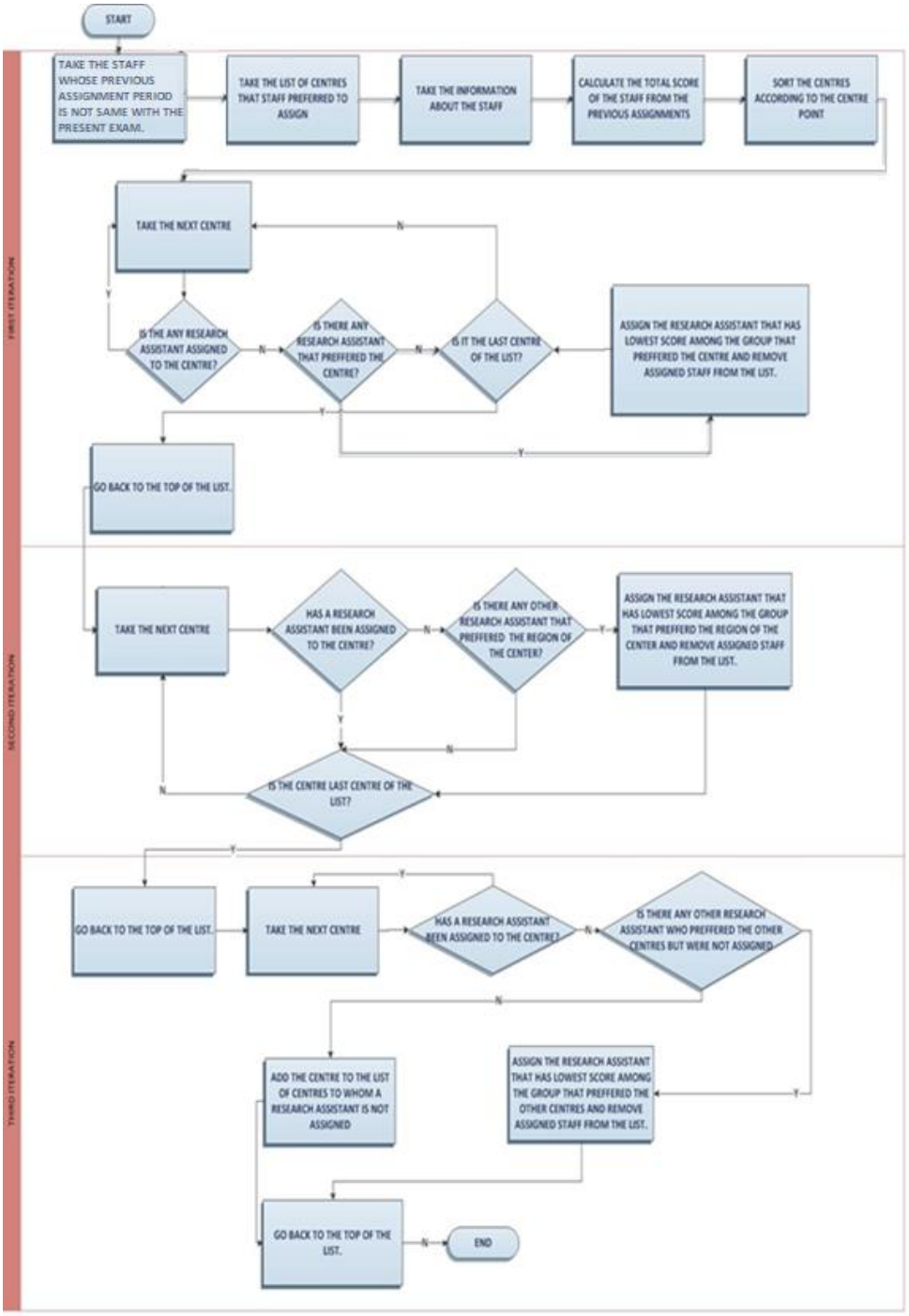


Fig. 4. Find an invigilator Who are a Research Assistant

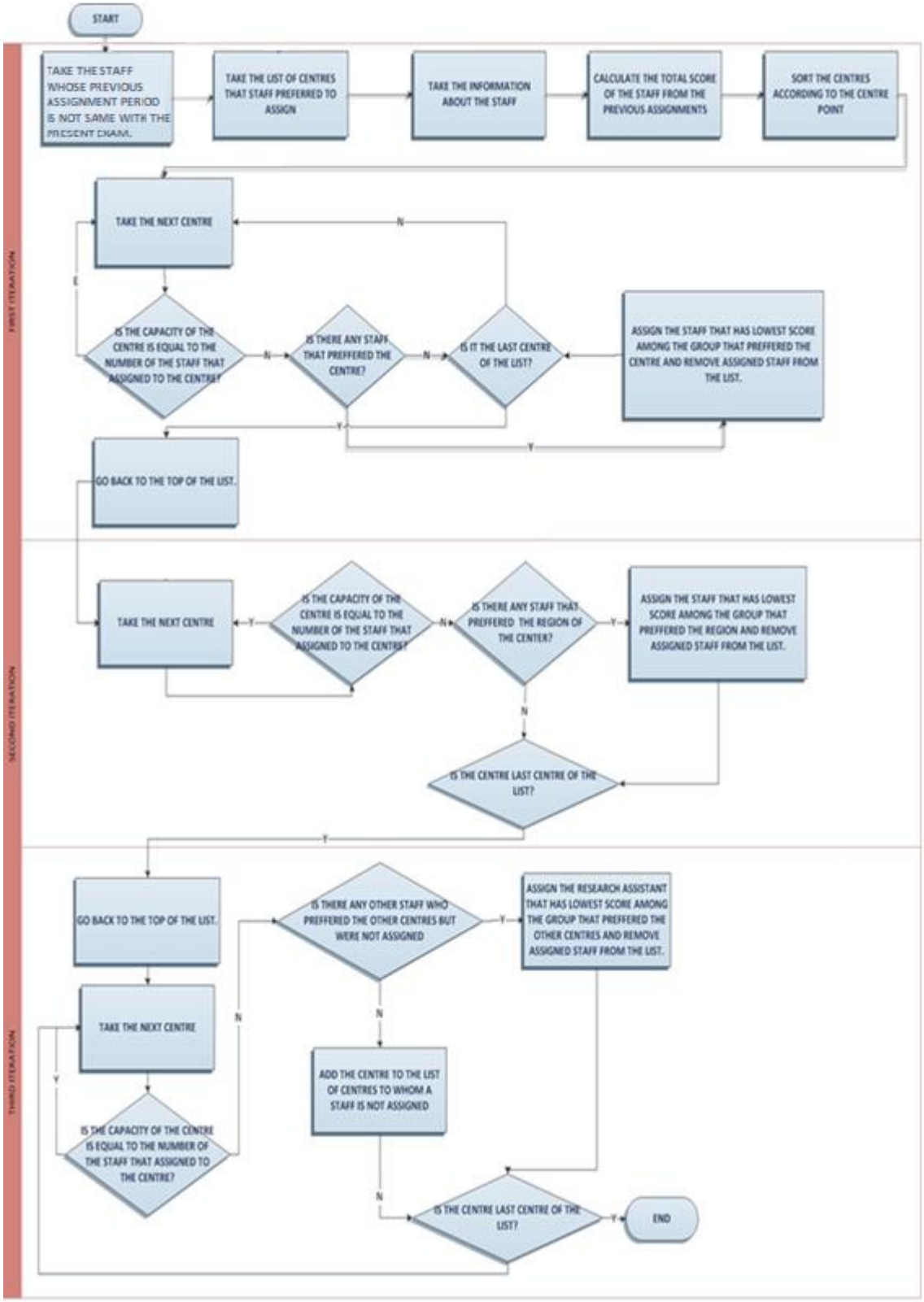


Fig. 5. Find an invigilator to the centers

3. RESULTS AND DISCUSSIONS

The assignment software was developed in the C# language in accordance with the algorithm. Thus, the result of the problem is directly linked to the preferences of the invigilator, and different preference scenarios are developed and the model was run again for these new situations. The profile of each invigilator member is simulated using three choices and these three choices are randomly created. While creating the choices for the

centers, the information regarding which invigilators were assigned previously is removed from the pool. The total score of each member of invigilators is calculated with the data from previous assignments. In the first scenario, when the invigilator prefers three centers from three different center groups, whereas in the second, an invigilator selects two centers from three different groups. In the third scenario, the invigilator selects six centers from three different groups. The evaluation of the assignment is shown in Table 1.

Table 1. Assignment Evaluation Data

	The Number of Assigned) Invigilator	The Number of invigilator assigned to their Preference	The Rate of Assignment to Preferences
Scenario 1	379	326	0,86
Scenario 2	379	362	0,97
Scenario 3	379	379	1

In the first scenario, there are 379 invigilators was assigned to the centers and 326 of these invigilators were assigned to their preferred centers. In the second scenario, 362 of 379 invigilators was assigned to their preferred centers. And in the third scenario, all of the 379 invigilators was assigned to their preferred center.

4. CONCLUSION

In this study, an invigilator assignment for an open distance education exam problem was examined as a real-life problem. The invigilator assignment operations were taken as a whole, and the whole process was then revised. After reviewing the process flow, the assignment problems in the literature were revised, and the problem was resolved to be an assignment problem. The mathematical model was proposed to solve our problem. An algorithm was designed and coded with C# to ensure that end users could use the model actively. The current invigilator assignment system that used in the university has some problems, such as the time and manpower needed to perform the assignment, difficulty in evaluating the preferences of invigilators, failure to equitably assignment of invigilators, and inability to follow changes in the system.

The prepared system assigns the invigilators to the appropriate centers taking account of the constraints and past scores. Minimizing the total scores of the assigned invigilator was considered as the success criterion of the assignment. The program was operated according to different preferred scenarios and the following results were obtained. Making an informed choice seems to play an important rate in increasing the success rate. In order to increase the success rate, the invigilator should be informed about their overall scores while expressing their preferences. The output of the developed program varies according to the sorting list of examination centers. Assignments were made by using 1000 different sets where the centers are listed. Sorting was created by different methods, which selected the best results out of this assignment. In future studies, different sorting techniques could be studied to achieve best results.

The developed system has shown a significant success in terms of time efficiency. The final cut of the developed system takes a period of approximately three minutes to run. With this program, assignment process time has been reduced from weeks to minutes. Before using this program, the preferences of the invigilator were not taken into consideration. While some of the invigilators were assigned to the same centers in all exams, the other invigilator could not have assigned in any exam. The use of this program brought a fair assignment system.

REFERENCES

Burkard, R. E. (2002). "Selected topics on assignment problems." *Discrete Applied Mathematics* 123(1-3): 257-302.

Diaz, E. F. v. J. (2001). "A Tabu Search Heuristic for the generalized assignment problem." *Europian Journal of Operational Research* 132: 22-28.

Elizondo, R. (1994). *Solving very large scale school/student assignment problems*, Rice University.

Elloumi, A., *et al*. (2014). "The classroom assignment problem: Complexity, size reduction and heuristics." *Applied Soft Computing* 14: 677-686.

G. Liu, S. W., Y. Zhang (2011). "New Prediction Algorithm for English Class Personnel Assignment Problem." *Advances in Computer Science, Intelligent System and Environment* 104: 95-98.

Hanum, F., *et al*. (2015). "Exam invigilators assignment problem: a goal programming approach." *Applied Mathematical Sciences* 9(58): 2871-2880.

Hojati, M. (2010). "Near-optimal solution to an employee assignment problem with seniority." *Annals of Operations Research* 181(1): 539-557.

K.Bhunia, J. M. V. A. (2010). "Penalty approaches for Assignment Problem with single side constraint via Genetic Algorithms." *Journal of Mathematical Modelling and Application* 2(2): 60-86.

Karimzadehgan, M. and C. X. Zhai (2012). "Integer linear programming for Constrained Multi-Aspect Committee Review Assignment." *Information Processing & Management* 48(4): 725-740.

Liang, B. and A. Turkcan (2016). "Acuity-based nurse assignment and patient scheduling in oncology clinics." *Health care management science* 19(3): 207-226.

Loiola, E. M., *et al*. (2007). "A survey for the quadratic assignment problem." *European Journal of Operational Research* 176(2): 657-690.

Oncan, T. (2007). "A survey of the generalized assignment problem and its applications." *Infor* 45(3): 123-141.

Özener, O. Ö., *et al*. (2017). "Solving a large-scale integrated fleet assignment and crew pairing problem." *Annals of Operations Research* 253(1): 477-500.

P. Tapkan, L. Ö. v. A. B. (2008). "Arı Algoritması ve Genelleştirilmiş Atama Problemi: Farklı Komşuluk Yapılarının Karşılaştırılması." *Endüstri Mühendisliği Dergisi* 21(2): 2-13.

Patriksson, M. (2015). *The traffic assignment problem: models and methods*, Courier Dover Publications.

Pentico, D. W. (2007). "Assignment problems: A golden anniversary survey." *European Journal of Operational Research* 176(2): 774-793.

Poole, A. and A. Kotsialos (2016). "Swarm intelligence algorithms for macroscopic traffic flow model validation with automatic assignment of fundamental diagrams." *Applied Soft Computing* 38: 134-150.

Subramanian, R., *et al*. (1994). "Coldstart: fleet assignment at Delta Air Lines." *Interfaces* 24(1): 104-120.

Wang, Z., *et al*. (2017). "The ABCD patient classification tool for nurse-to-patient assignment to improve nursing workload balance: a multi-center study." *International Journal of Clinical And Experimental Medicine* 10(6): 9502-9507.

Turkish Journal of Engineering

Turkish Journal of Engineering (TUJE)
Vol. 3, Issue 3, pp. 149-156, July 2019
ISSN 2587-1366, Turkey
DOI: 10.31127/tuje.498878
Research Article

USER-ORIENTED FILE RESTORATION FOR OPERATING SYSTEMS

Hüseyin Pehlivan *[1]

[1] Karadeniz Technical University, Engineering Faculty, Department of Computer Engineering, Trabzon, Turkey
ORCID ID 0000-0002-0672-9009
pehlivan@ktu.edu.tr

* Corresponding Author
Received: 18/12/2018 Accepted: 18/02/2019

ABSTRACT
Folders such as recycle bin are a crucial component of wide working environments like operating systems. In current operating systems, such facilities are implemented either in no user-oriented fashion or very poorly. Various intrusion detection mechanisms are developed to prevent any damage, but very few offers the repair of the user's file system as an additional level of protection. This paper presents how to build a recycle bin mechanism for Unix operating systems entirely at the user level. The mechanism involves the control of system resources in a more intelligent way. Programs thus are running under greater control, monitoring and analyzing their resource requests. The idea is based on the interception of a particular class of system calls, using tracing facilities supported by many Unix operating systems. This provides better high level information, and presents efficient techniques to prevent foreign or untrustworthy programs from doing any irreparable damage. A program called trash has been constructed and experimented to investigate potential consequences of the recycle bin mechanism. The experiments highlight possible overheads imposed on the system. The paper also performs a comparative analysis of the trash program with some related approaches and tools

Keywords: *Recycle Bin, Operating Systems, System Calls, Process Tracing, Restoration*

1. INTRODUCTION

Recycle bin mechanisms are usually provided for file recovery purposes and directly employed by users. User-oriented recovery in operating systems has become very least concern among past studies, possibly due to reasons such as the lack of necessary system facilities. The system-wide integrity protection is accomplished by third party programs that, once an anomalous state has been detected, remove the anomaly from the system, restoring the original state. The anomaly removal and restoration capabilities of these commercial programs fail to completely reverse the effects of an anomalous program (Passerini *et al*., 2009).

In programs such as editors, users are presented with recovery commands (undo, redo, etc.) to meet their preservation and restoration requirements. Given an operating system, undo and redo commands, which are rather useful for smaller environments, do not seem very functional to bring back any destructed file seamlessly, which is mainly caused by the multi-tasking nature of operating systems and possible dependencies between user commands. Two forthcoming studies are made on Windows, where untrusted programs are monitored, logging their operations, (Hsu *et al*., 2006; Paleari *et al*., 2010). Using the logs, it can completely remove malware programs and reliably restore the infected data. However, it is a recovery facility that is provided for a single user environment.

In modern operating systems, notably POSIX compliant ones, recycle bin functionality is integrated into a desktop environment and its file manager. Typical examples are Microsoft Windows with Windows Explorer and GNOME with Nautilus. In such environments, an overwritten file does not usually go to the recycle bin. In fact, the recoverability of files deleted by a program depends on its level of integration with a particular desktop environment. Low-level utilities can bypass this layer entirely and delete files immediately. In Windows or Unix operating systems, a file removed through a DOS or terminal window is not placed in the recycle bin and so is not recoverable.

The aim of the paper is to provide a user-oriented restoring facility for Unix operating systems. Unlike the Windows recycle bin mechanism, the restoring facility operates in a more intelligent way, dealing with dynamically deleted files as well. In fact, it basically deals with only file-related damage of either commands issued or applications managed by a user. The damage imposed by the system itself as a result of situations such as a crash, however, have to be handled by the administrator.

The described mechanism, called *trash* (TRAcing SHell), is implemented using C++, completely at the user level without having to modify the existing system internals. It consists of two separate subprograms which monitor user activities and restore destructed files. The monitoring program (*monitor*) is started as a daemon by the administrator, and controls all users' activities. It can handle all existing commands and programs, and any new ones which have been installed recently. The restoring program (*restorer*) runs as an ordinary program, and can be configured for each different user, with their own restoring requirements.

Like all usual recovery implementations, the *trash* mechanism needs to find out what changes each user command makes to the file system. Shells in common use never know which file operations commands pass on to the kernel to execute, since the commands make these requests with system calls. The *trash* daemon must itself monitor file requests of commands. Fortunately, many Unix operating systems provide standard tracing facilities such as the /proc file system which is used by *truss*-like applications and various intrusion detection systems. With these facilities, system calls are intercepted and then resumed after the appropriate information is extracted.

2. RELATED WORK

Approaches of system protection can be broadly divided into two categories; system restoration and intrusion detection. System restoration based approaches maintain some specific checkpoints in the system and roll back a user-selected collection of actions. Approaches of intrusion detection can be further divided into anomaly detection and misuse detection. Anomaly detection based approaches first construct a profile that describes normal behaviours and then detect deviations from this profile. In contrast, misuse detection based approaches define and look for precise sequences of events that damage the system.

System restoration is based on information stored or gathered during normal interaction with a program. Almost all programs are being currently equipped with restoring facilities. In the environments controlled by self-contained programs such as editors and painters, the provision of recovery support is relatively easy due to paucity and known effects of operations. Wide environments presented by operating systems are open to every kind of operations, and thus the effects of system programs must be handled in a more intelligent way. Therefore different environments involve differentiating types of restoration, which is basically associated with the number of users working in them. Small environments use data chunks for restoration, while wide ones do data files for restoration. Restoration in small environments is inherently user-oriented and supplied as a result of recovery commands such as undo and redo. There are many recovery models introduced for such environments where interactions with a single user (Vitter, 1984; Spenke and Beilken 2003; Brown and Patterson, 2003) or multiple users (Choudhary and Dewan, 1992; Berlage and Genau, 1993) are supported. Typical examples of those models are history undo model (Stallman, 1986), selective undo model (Prakash and Knister, 1992) and object-based undo model (Zhou and Imamiya, 1997). Restoration in wide environments is either system-oriented or user-oriented. As in Unix operating systems, system-oriented restoration is carried out via backup tapes accessible by only the system administrator. As in Windows operating systems, user-oriented restoration is done via a filestore maintained in the system.

Anomaly detection techniques address the existence of an intrusion by considering any abnormalities in user or system behaviour as a potential attack (King and Chen, 2003; Qiao *et al*., 2002; Christodorescu and Jha, 2004). In order to learn normal user or system behaviours, most techniques analyze program behaviours. Typical examples of analyzing program behaviours are the N-gram and FSA-based algorithms

(Sun *et al*., 2005; Wu *et al*., 2016; Yu *et al*., 2005), which are usually applied to server programs. Both algorithms characterize normal program behaviours in terms of sequences of system calls. A sequence of system calls that have not been observed under normal operation of programs is treated as anomalous program behaviour. The N-gram algorithm breaks a system call sequence into substrings of a fixed length N, and then stores these substrings (called N-grams) in a table. The FSA-based algorithm maintains state-related information (the program state in the point of each system call) as system calls is made by a process under normal execution, where the system calls correspond to transitions in FSA (finite-state automata). The performance of these algorithms for three popular servers (FTP, HTTP, NFS) can be found in (Sekar *et al*., 2001).

Misuse detection techniques model known attacks using patterns and detect them via pattern-matching (Abed *et al*., 2015; Anandapriya and Lakshmanan, 2015; Chen *et al*., 2016; Creech and Hu, 2014; Jose *et al*., 2018; Liu *et al*., 2018). These techniques rely on a wide variety of observable data such as system-call data, preventing intrusions from either local or remote users as a result of evaluating the legitimacy of their activities. Many solutions to detecting and preventing intrusions are based on the interception of system calls. A recent comprehensive review is conducted to assess the advantages and drawbacks of intrusion detection techniques proposed in the literature (Ramaki *et al*., 2018).

3. CHARACTERISTICS OF RESTORATION

In this work, a conventional distinction is made between state and file restorations. The restoration of a state occurs as a result of undoing or redoing user commands. The names and contents of files together determine the system state, as well as other components such as directories and access permissions. A change in file names is considered to move the system into another state. So a state recovery means that each component of a state affected of the execution of a command is reinstated. File restoration is usually different from state restoration, which is associated with the contents of files only, excluding file names or permissions.

In practice, this distinction helps provide more useful functionalities from the perspective of a user's requirement. The user would mainly expect a utility to be able to bring the files with the correct contents back. In restoration, since the contents are just important, there is less information stored to carry out file protection and restoration.

On the other hand, there are some difficulties with ensuring the applicability of recovery commands to all situations. One important difficulty is the requirement of controlling concurrency, where programs interleave. Indeed, even if concurrency is controlled desirably, it may leave some programs irrecoverable, restricting the usability of the recovery mechanism. For example, consider the following situation which is led to by concurrent execution of two programs, `P1` and P2, entered in separate command-lines:

```
10> Delete fileA 5 P1
11> Create fileA 11 P1
12> Read fileA 11 P2
13> Create fileB 13 P2
14> Read fileB 13 P1
```

This does not enable both programs to be undone separately. Thus it is impossible to get back the old version of `fileA` (version 5). There is no elegant way to cope with this situation, because the user is currently allowed to specify only one command to undo from the history list at a time. This involves considering each atomic file operation individually.

There is a close relationship between file protection and safe command execution. The best way of protecting a file system is to ensure that every system command executes securely. Provision of secure execution of commands seems to require restricting the environment. There is a lot of work associated with file protection, which have concentrated on safe command execution, as given in Section 2. Many of previous works provide users with restricted environments for safe execution of programs. Any destroyed file cannot be retrieved. However, the *trash* mechanism aims not to restrict the working environment.

4. DESIGN AND IMPLEMENTATION

The *trash* program is designed in two components (subprograms), namely *monitor* and *restorer*, which run as separate processes in tracing and restoration modes respectively. The *monitor* component both controls one user's activities as a result of tracing user-serving programs and stores information required for restoration in a directory named *trash* under the user's home directory. The kind of restoration information stored depends entirely on intercepted system calls, whose effects on the system may vary considerably, and so each system call is handled individually. The *restorer* component handles the restoration requirements, bringing old versions of files back.

The owner of these components is the system administrator, and thus users are not allowed to directly write to the *trash* directory. This restriction is required to protect restoration information against other programs executed by users, intentionally or not. Besides, for more efficient protection, a system group named *restore* is created, which consists of the name of users to employ the *trash* mechanism, and the group ownership of *restorer* is changed to that. In this way restoration is made possible only through the *restorer* program.

In order to control all activities of a particular user, it is not adequate to trace only login shells for recovery purposes. Unix operating systems provide various levels of remote access. For each level, there are many tools which allow users to manage their accounts remotely, as described in Table 1.

Table 1 also gives the names of daemon processes serving the specified tools, which are system-dependent. For example, X servers which allow running Unix desktop environments such as KDE and GNOME, interact with display managers (daemons) such as kdm and gdm, respectively. The file-related effects made through all such tools need monitoring properly.

Table 1. Unix remote access tools and daemons

Tools	Instances	Daemons
Telnet/SSH client programs	AxeSSH, FiSSH, PuTTy	telnetd, sshd
FTP/SFTP programs	AceFTP, SmartFTP, WinSCP3	ftpd, sftpd
X Window System servers	X-Win32, eXceed, WeirdX	kdm, gdm
Web servers	Apache, Java Web Server	httpd, webservd

Unix environments are full of many utilities used for text formatting and program developing purposes that produce numerous temporary or permanent files, which we call "generated files". To hold these specific files, they can use some subdirectories under the system's root directory (e.g., /tmp and /var) or the user's home directory (e.g., ~/.netscape and ~/.ssh). In the current implementation, all files only in home directories are protected by default, except some language and program specific files with the extensions ".aux", ".log", and ".o". However, each user can individually change the default configuration via the menu provided by *restorer* program.

The *trash* program stores restoration information in two files named *.trash* and *.conf* under ~/.trash. The first keeps the record of file deletions, while the latter contains the names of files and directories that are not to be protected by *trash*. File names can be specified explicitly or using some extensions. Other entries in ~/.trash are copies of files stored as a whole with a number extension at the end which is incremented sequentially.

Typical scenario of storing information might be as follows. Whenever *trash* intercepts the *open* system call, it checks the first argument to see whether it points to a pathname under the user's home directory. If the file is owned by other users or belongs to a standard shared library, for example, the *open* is allowed to go ahead, in which case nothing is stored. Otherwise, the second argument is checked to see which operation *open* is to perform on the file. In the case of a *deletion* operation, for example, a complete copy of the file is taken. (For *open*, a *deletion* means modification of the file as a result of flags such as `O_WRONLY`, `O_RDWR` and `O_TRUNC`.) Then the system call is released to resume as normal. The other system calls, such as *creat*, *rename*, *unlink*, *symlink* and *link*, can also perform deletion operations.

The text-based versions of the *monitor* and *restorer* programs are currently implemented. During tracing, *monitor* uses each user's own configuration to gather and store information. There is no graceful way to individually monitor user tasks via separate processes due to the fact that a single process can serve more than one user. Therefore, only one instance of it is running in the system, without spawning one monitoring child process for each user. On the other hand, each user executes her/his own instance of the *restorer* program from the command line which accesses the .trash directory to obtain restoring information.

All these preparations make the restoration of files quite simple. The effect of the *restorer* program consists only of the movement or exchange of file versions between the current state and the *.trash* directory of previous versions. For the restoration of a particular file that has been previously deleted by a user task, the *restorer* program firstly checks to see if an overwrite occurs. If there is no overwrite operation, it brings the deleted file back to the current state and removes the related line of the *.trash* file. Otherwise it warns the user.

5. CONCURRENCY CONTROL

An important issue to handle file restoration is concurrency. Subject to their use of system calls, processes execute independently of each other and share system resources arbitrarily, which can cause non-determinism. In order to reduce the usual non-determinism problems caused by uncontrolled sharing of resources to a minimum, *trash* has to handle concurrency in a more sophisticated way than the operating system or ordinary shells. Fortunately, the /proc interface provides a means of monitoring all user processes and their descendants via only one tracing process. The *monitor* program behaves like a tracing process, using the *poll* call to listen to process events. To efficiently monitor running processes, a process table named `procT` is defined by the following C++ structure:

```
int nprocs;
int current_time;
struct pollfd Pollfds[MAXPROCS];
struct processTable {
    int         time;
    pid_t       pid;
    int         procfd;
    prstatus_t *pstatus;
} procT[MAXPROCS];
```

where the structure Pollfds are used by *poll* to perform efficient control of process events concurrently. Each new process, which is represented by *pid*, is inserted into the table with a unique *time* entry, using the value *current_time*. Note that parent-child relations do not require maintenance in the table. All processes are supposed to be at the same level. To detect and record their concurrent file accesses, a second table `fileT`) is maintained as follows:

```
static int nfiles;
struct fileTable {
    int nprocfds;
    int procfd[MAXPROCS];
    char *path;
} fileT[MAXFILES];
```

where the attempt of opening a file adds a new entry to the table. All processes that hold a single file open are kept together within the same entry.

The time during which a file remains open is usually determined by *open*-like calls and *close*. However, on seeing a *close* call, We cannot simply assume that the process is finished with the related file. One reason is the duplications of file descriptors. In this way, a process can have some files open even after *close*, as a result of performing system calls such as *dup* and *fcntl*.

The monitoring of system calls that copy existing

descriptors is not enough to detect all files a process holds open. There is another situation where files might remain open. It is associated with inheritance structures of processes. When the `FD_CLOEXEC` flag of a file descriptor created by a parent process is clear, the file remains open across *fork* and *exec* calls, which means that its child processes can inherit those file descriptors. This complicates the issue of keeping the track of file descriptors for each process.

In fact it is not easy and practical to monitor processes for file descriptors information, because it imposes numerous dynamic checks on run-time environments of processes. We do not deal with file descriptors much. For simplicity, it is assumed that a process creating or inheriting a file descriptor holds the related file open until it terminates.

File open activities of processes provide the most valuable information for *trash*. Each opening operation does not correspond to a backup. For the first process that opens a file for writing for the first time, the file is copied to ~/.trash. As far as the file remains open, all later accesses to it by different processes would have no effect on the restoring mechanism (that is, no new copies are made). Only after all the related processes end, a new process that tries to open the same file for writing would cause a backup.

6. OVERHEADS AND PERFORMANCE

In order to measure the whole overheads caused by the *trash* program, the overhead is examined in terms of runtime and space usage. Runtime overheads occur with both system call interception and execution of detection/storing code. Space overheads are caused by files stored in ~/.trash.

Table 2. CPU time and disk space overheads of the *trash* program

Program	Not Monitored	Monitored	Overhead
cp	35ms	37ms	5%
latex	160ms	182ms	13%
xterm	390ms	439ms	12%
emacs	483ms	528ms	9%
mozilla	2476ms	2552ms	3%
eclipse	9342ms	9986ms	7%
smc	12350ms	13015ms	5%

With root privileges we monitor three system daemons (e.g. sshd, sftp-server and dtlogin), actually grandchild processes through them. The webservd daemon could not be monitored satisfactorily, due to the absence of the web content that interacts with the file system intensively. Table 2 shows the results for certain programs which operate on Sun Solaris Sparc machine with 2 processors with 1.28GHz each, 4GB of RAM and 4 Ultra160 SCSI hard disks with 73GB each. The machine is daily connected by about 34 users remotely through the tools given in Table 1. To get around the effects of network environment, the connection times are not measured. The CPU time is only given for some system programs. For the long-running programs given in Table 2, the time measurement is restricted by the point the user input is asked (*smc* is the Solaris management console program).

Compared to techniques for interception of system calls within the kernel, user-level mechanisms tend to incur significantly higher overheads. This is unavoidable because, for each system call, there are additional context switches between a process that is handling a particular user task and another process that is intercepting its system calls. The number of context switches depends on how many system calls of interest a program issues. Given the file system operations, each user program usually need to perform a small number of system calls of interest to *trash*, causing a few context switches. For instance, the Unix command "rm *" performs one *unlink* call for each file in the current directory. The fact that file system activities of most Unix commands are very few introduces trivial overheads in the short terms.

Furthermore, to find out the name and parameters of a system call intercepted, the monitoring process is required to access the monitored process memory, incurring the overhead of system call interception. After detecting and decoding a system call, possible storing operations are executed, incurring the overhead of execution of detection/storing code.

Disk space overheads are also explored on the machine with the specifications described above, for the period of one month. For a particular user, the size of disk space is increased with 19% of currently occupied one. For another user, the increase is only 12%. These overheads are imposed by file storing operations, which may lead to more space overheads. As a solution to keeping space usage down, all stored files might be compressed, but this possibly leads to an increase in storing time, which is beyond the subject of this paper.

On the other hand, compared to traditional recovery mechanisms, the restoration technique needs less amount of space in the command-independent fashion. To illustrate, consider the following situation, assuming that `fileB` and `fileC` does not exist in the current state:

```
P1: rename(fileA, fileB)
P2: rename(fileB, fileC)
```

where the *rename* calls are executed through two separate user programs. For purposes of recovery, fileA and `fileB` have to be saved so that they can be separately reinstated later. For restoration, there is no need to save any file, because the contents of `fileA` remain unchanged, even if it is eventually renamed as `fileC`.

We also explored that some programs (e.g. an editor or painter) can overwrite a file many times during execution. In this case, the program carries out many deletions of possibly the same file, causing intermediate copies of the file to be stored. Most of these deletions have no effect on restore operations. So it is unnecessary for *trash* to keep track of them all. Restoration information needs only contain at most one copy for a single file managed by a program, which corresponds to the contents of the file at the time when the program is called.

7. COMPARATIVE ANALYSIS

Most literature studies that analyze the trace of

system calls focus on intrusion detection systems. Although system call data is an instrumental artefact of the kernel, it can be modelled to support decision making activities of these systems at program level. Thus, the anomaly detectors use various modelling techniques for assessing the behaviour of processes via the sequence of system calls and their arguments. The techniques are typically based on sequential features (SF), frequency-based features (FF), argument-based features (AF) and hidden-markov models (HM). Table 3 compares our study with some previous work in terms of the modelling techniques and the type of intrusion detection, which can be supervised or not.

Table 3. Comparison of system call-based modeling techniques

Reference	Detec.	Tech.	Supr.
Anandapriya *et al*., 2015	Anomaly	SF	Yes
Creech *et al*., 2014	Anomaly	SF	Yes
Gupta *et al*., 2015	Misuse	SF	No
Xie *et al*., 2014	Anomaly	FF	No
Haider *et al*., 2015	Anomaly	FF	No
Hoang *et al*., 2009	Hybrid	HM	Yes
Hu *et al*., 2009	Anomaly	HM	No
Sekar, 2001	Anomaly	AF	No
Mutz, 2006	Anomaly	AF	No
Our approach	Hybrid	AF	Yes

Another comparison is conducted based on different performance criteria such as scalability (Scal.), space complexity (Space), time complexity (Time) and detection robustness (Rbst.). Table 4 shows the analysis results for various detection systems.

Table 4. Comparison of various detection systems based on some performance criteria

Reference	Scal.	Space	Time	Rbst.
Fuse *et al*., 2017	Low	High	Low	Low
Yolacan *et al*., 2014	High	High	High	High
Hu *et al*., 2009	High	Low	Low	Low
Zhou *et al*., 2008	High	High	High	High
Zhang *et al*., 2006	High	Low	High	High
Hoang *et al*., 2009	Low	Low	High	High
Our approach	High	Low	Low	High

The final comparison of our approach is made with three other restoration approaches and three commercial malware detectors including Nod32 Anti-Virus, Panda Anti-Virus, and Kaspersky Anti-Virus, which is evaluated in (Passerini, 2009). The restoration operations are performed on three system resources of files, registry keys and/or processes. The performance of the approaches and tools is classified as good, average and poor categories. The results are shown in Table 5.

Table 5. Comparison of some restoration approaches and tools

Approach/Tool	OS	File	Reg.	Proc.
Nod32	Windows	Good	Good	Poor
Panda	Windows	Avg.	Avg.	Avg.
Kaspersky	Windows	Avg.	Good	Poor
Hsu *et al*., 2006	Windows	Good	Good	NA
Paleari *et al*., 2010	Windows	Good	Good	Good
Webster *et al*., 2018	Linux	Good	NA	NA
Our approach	Unix	Good	NA	NA

8. CONCLUSION AND FUTURE WORK

We have described a recycle bin mechanism on Unix for repairing the file system damages in a user-oriented fashion. The mechanism deals with all situations which threaten the integrity and security of the file system, giving a chance the user to restore unintentional deletions of files. It achieves these goals by monitoring individual user activities, storing destructed files, and using the restoration information to eliminate their effects. We examined its overheads for possible file-related operations of programs, concluding that the CPU and storage overhead caused by the utility is acceptable. This conclusion is also supported by the results of the comparative analysis made with some other approaches and tools.

The main subject of future work is on maintaining the recycle bin efficiently. To achieve this goal, one issue is to determine how long file copies are preserved in the user's disk area. Some users can require file preservation to be in effect during multiple login sessions. To avoid long-term file preservations, an optimum duration must be determined for most utilization of the facility. This raises another issue, which is the disk space usage of stored files. Using the disk quota allowed for each user as an additional parameter, we are currently working on the recycle bin to make it occupy less disk space and to keep it covering some particular sessions.

REFERENCES

Abed, A. S., Clancy, C. and Levy, D. S. (2015). "Intrusion Detection System for Applications Using Linux Containers", *International Workshop on Security and Trust Management*, pp. 123-135.

Anandapriya, M. and Lakshmanan, B. (2015). "Anomaly Based Host Intrusion Detection System using semantic based system call patterns", Proc., *9th International Conference on Intelligent Systems and Control (ISCO)*, pp. 1-4.

Berlage, T. and Genau, A. (1993). "From undo to multi-user applications", *Proc., Vienna Conference on Human–Computer Interaction*, Vienna, Austria, Sept 20-22, pp. 213-224.

Brown, A. B. and Patterson, D. A. (2003). "Undo for operators: Building an undoable e-mail store" *Proc., 2003 USENIX Annual Technical Conference*, pp. 1-14.

Chen, C. M, Guan, D. J. and Huang, Y. Z and Ou, Y. H. (2016). "Anomaly network intrusion detection using Hidden Markov Model", *International Journal of Innovative Computing, Information and Control*, Vol. 12, No. 2, pp. 569-580.

Choudhary, R. and Dewan, P. (1992). "Multi-user undo/redo", *Technical Report TR125P*, Computer Science Department, Purdue University.

Christodorescu, M. and Jha, S. (2004). "Testing malware detectors", *Proc., 2004 ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA)*, New York, NY, USA, pp. 34-44.

Creech, G. and Hu, J. (2014). "A Semantic Approach to Host-Based Intrusion Detection Systems Using Contiguous and Discontiguous System Call Patterns", *IEEE Transactions on Computers*, Vol. 63, pp. 807-819.

Fuse, T. and Kamiya, K. (2017). "Statistical anomaly detection in human dynamics monitoring using a hierarchical dirichlet process Hidden Markov Model", *IEEE Transactions on Intelligent Transportation Systems*, Vol. 18, No: 11, pp. 3083–3092.

Gupta, S. and Kumar. P. (2015). "An immediate system call sequence based approach for detecting malicious program executions in cloud environment", *Wireless Personal Communications*, Vol. 81, No. 1, pp. 405–425.

Haider, W., Hu, J. and Xie. M. (2015). "Towards reliable data feature retrieval and decision engine in hostbased anomaly detection systems", *IEEE 10th Conference on Industrial Electronics and Applications (ICIEA)*, Auckland, New Zealand, pp. 513–517.

Hoang, X. D., Hu, J. and Bertok. P. (2009). "A program-based anomaly intrusion detection scheme using multiple detection engines and fuzzy inference", *Journal of Network and Computer Applications*, Vol. 32, No. 6, pp.1219–1228.

Hsu, F., Chen, H., Ristenpart, T., Li, J., and Su, Z. (2006). "Back to the future: A framework for automatic malware removal and system repair", *Proc., 22nd Annual Computer Security Applications Conference*, ACSAC '06, IEEE Computer Society, Washington, DC, pp. 257–268.

Hu, J., Yu, X., Qiu, D. and Chen, H.-H. (2009). "A simple and efficient hidden markov model scheme for host-based anomaly intrusion detection", *IEEE network*, Vol. 23, No. 1, pp. 42–47.

Jose, S., Malathi, D. Reddy, B. Jayaseeli, D. (2018). "Anomaly Based Host Intrusion Detection System Using Analysis of System Calls", *International Journal of Pure and Applied Mathematics*, Vol. 118, No. 22, pp. 225-232.

King, S. and Chen, P. M. (2003). "Backtracking intrusions", *Proc., 19th ACM Symposium on Operating Systems Principles (SOSP)*, pp. 223-236.

Liu, M., Xue, Z. Xu, X., Zhong, C. and Chen. J. (2018). "Host-Based Intrusion Detection System with System Calls: Review and Future Trends", *ACM Computing Surveys*, Vol. 51, No. 5, pp 1-36.

Mutz, D., Valeur, F., Vigna, G. and Kruegel, C. (2006). "Anomalous system call detection". *ACM Transactions on Information and System Security (TISSEC)*, Vol. 9, No. 1, pp. 61–93.

Paleari, R., Martignoni, L., Passerini, E., Davidson, D., Fredrikson, M., Giffin, J., and Jha, S. (2010). "Automatic generation of remediation procedures for malware infections", *Proc., 19th USENIX Security Symposium*, August 11-13, Washington, DC, pp. 419-434.

Passerini, E., Paleari, R. and Martignoni, L. (2009). "How Good Are Malware Detectors at Remediating Infected Systems?", *Proc., 6th International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*, July 09-10, Como, Italy, pp. 21-37.

Prakash, A. and Knister, M. J. (1992). "A framework for undoing actions in collaborative systems", *Technical Report CSE-TR-125-92*, Computer Science and Engineering Division, The University of Michigan, Ann Arbor.

Qiao, Y., Xin, X. W., Bin, Y. and Ge, S. (2002) "Anomaly intrusion detection method based on HMM", *Electronics Letters*, Vol. 38, No. 13, 2002, pp. 663-664

Ramaki, A. A., Rasoolzadegan, A. and Jafari, A. J. (2018) "A Systematic Review on Intrusion Detection based on the Hidden Markov Models", *Statistical Analysis and Data Mining*, Vol. 11, No. 3, pp. 111-134

Sekar, R., Bendre, M., Dhurjati, D. And Bollineni, P. (2001). "A Fast Automaton-Based Method for Detecting Anomalous Program Behaviors", Proc., *IEEE Symposium on Security and Privacy*, pp. 144-155.

Spenke, M. and Beilken, C. (1991). "An overview of GINA—the generic interactive application", *Proc., User Interface Management and Design*, D. A. Duce *et al.*, Eds. Springer-Verlag, New York, NY, USA, pp. 273-293.

Stallman, R. (1986). "GNU Emacs manual, Version 17", *Free software foundation. Inc*.

Sun, W., Liang, Z., Venkatakrishnan, V. N. and Sekar,R. (2005). "One-way isolation: An effective approach for realizing safe execution environments", *Proc., Network and Distributed Systems Symposium (NDSS)*, pp. 265-278.

Webster, A., Eckenrod, R. and Purtilo, J. (2018). "Fast and Service-preserving Recovery from Malware Infections Using CRIU", *Proc., 27th USENIX Security Symposium*, August 15-17, Baltimore, MD, USA, pp. 1199-1211.

Vitter, J. Z. (1984). "US&R: A new framework for redoing", *IEEE Software*, Vol. 1, No. 4, pp. 39-52.

Wu, F., Wu, D. and Yang, Y. (2016). "A Network Intrusion Detection Algorithm Based on FSA Model", *4th International Conference on Machinery, Materials and Computing Technology*, pp. 615-621.

Xie, M., Hu, J., Yu, X. and Chang. E. (2014). "Evaluating host-based anomaly detection systems: Application of the frequency-based algorithms to adfa-ld", *International Conference on Network and System Security*, Berlin, Heidelberg, Springer, pp. 542–549.

Yolacan, E. N., Dy, J. G. and Kaeli, D. R. (2014). "System call anomaly detection using multi-HMMs, Software Security and Reliability-Companion (SERE-C)", *IEEE 8th International Conference on Software Security and Reliability-Companion*, San Francisco, CA, pp. 25–30.

Yu, F., Xu, C., Shen, Y., An, J., and Zhang, L. (2005). "Intrusion detection based on system call finite-state automation machine". *IEEE International Conference on Industrial Technology*, pp. 63-68.

Zhang, J., Liu, Y. and Liu, X. (2006). "Anomalous detection based on adaboost-HMM", *IEEE 6th World Congress on Intelligent Control and Automation (WCICA)*, pp. 4360–4363.

Zhou, C. and Imamiya, A. (1997). "Object-based nonlinear undo model", *Proc., 21th International Computer Software and Applications Conference*, COMPSAC, pp. 50-55.

Zhou, X., Peng, Q. K. and Wang, J. B. (2008). "Intrusion detection method based on two-layer HMM", *Application Research of Computers*, Vol. 3, No. 1, 75.

CONTENTS

PREDICTING THE ENERGY PRODUCTION OF A ROOFTOP PV PLANT BY USING DIFFERENTIAL EVOLUTION ALGORITHM
Ali Yıldız 106

MASS CONCRETE CONSTRUCTION USING SELF-COMPACTING MORTAR
İ. Raci Bayer, Lutfullah Turanli and P. Kumar Mehta 110

DETERMINATION OF MALIGNANT MELANOMA BY ANALYSIS OF VARIATION VALUES
Ahmet Kürşat Esim, Hilal Kaya and Veysel Alcan 120

AN INTEGRATED APPROACH BASED ON THE TAGUCHI METHOD AND RESPONSE SURFACE METHODOLOGY TO OPTIMIZE PARAMETER DESIGN OF ASBESTOS-FREE BRAKE PAD MATERIAL
Funda Kahraman and Banu Sugözü 127

MECHANICAL AND WEAR CHARACTERISATION OF QUARRY TAILINGS REINFORCED A6063 METAL MATRIX COMPOSITES
Stephen Durowaye, Olatunde Sekunowo, Babatunde Bolasodun, Isaac Oduaran and Ganiyu Lawal 133

MODELLING AND OPTIMIZATION OF THE EXAM INVIGILATOR ASSIGNMENT PROBLEM BASED ON PREFERENCES
Dilek Küçük Matcı and Ilgın Acar 140

USER-ORIENTED FILE RESTORATION FOR OPERATING SYSTEMS
Hüseyin Pehlivan 149

ISSN 2587-1366

TURKISH JOURNAL OF ENGINEERING

VOLUME 3 ISSUE 3 JULY 2019