THEORY AND RESEARCH IN ENGINEERING II

VOLUME 2

EDITOR: DOÇ. DR. MİHRİBAN KALKANCI

gece kitaplığı

İmtiyaz Sahibi / Publisher • Yaşar Hız

Genel Yayın Yönetmeni / Editor in Chief • Eda Altunel

Kapak & İç Tasarım / Cover & Interior Design • Gece Kitaplığı

Editör / Editor • Doç. Dr. Mihriban Kalkancı

Birinci Basım / First Edition • © Aralık 2020

ISBN • 978-625-7319-20-1

© copyright
Bu kitabın yayın hakkı Gece Kitaplığı'na aittir.

Kaynak gösterilmeden alıntı yapılamaz, izin
almadan hiçbir yolla çoğaltılamaz.

The right to publish this book belongs to Gece Kitaplığı.
Citation can not be shown without the source, reproduced in any way
without permission.

Gece Kitaplığı / Gece Publishing
Türkiye Adres / Turkey Address: Kızılay Mah. Fevzi Çakmak 1. Sokak
Ümit Apt. No: 22/A Çankaya / Ankara / TR
Telefon / Phone: +90 312 384 80 40
web: www.gecekitapligi.com
e-mail: gecekitapligi@gmail.com

Baskı & Cilt / Printing & Volume
Sertifika / Certificate No: 47083

Theory and Research in Engineering II

Volume 2

<u>EDITOR</u>

DOÇ. DR. MIHRIBAN KALKANCI

gece kitaplığı

CONTENTS

Chapter 29

STATIC CHARACTERISTICS OF THE ELECTRIC ARC FURNACES[1]

Mustafa ŞEKER

1 Assistant Prof.Dr. Mustafa ŞEKER, Sivas Cumhuriyet University, Hafik Kamer Ornek Vocational School of Higher Education, Department of Transportation Services, Rail Systems Management Program, Sivas/Turkey. mustafaseker@cumhuriyet.edu.tr ; Orcid no: 0000-0002-3793-8786

1. Introduction

Iron and steel production is a key factor that determines a country's position in the global industry. The iron and steel industry meets the raw material demands of the countries in many fields, particularly construction, automotive and industrial production. The production of iron and steel is carried out in industrial facilities including induction furnace elements and Electric Arc Furnaces (EAF). Thanks to the developments in the technological field in recent years, EAFs have reached high metal melting capacities. This has resulted in lower melting costs compared to other steel production methods. Today, approximately 40% of global steel production is carried out using EAF production plants. In the future, this proportion is expected to increase to about 50% by 2030 (Schau., H., Stade, 1994; Workshop, 2016). Due to the increasing cost of energy today, consuming the energy in EAFs in the most economical way is very important for both producers and the country's economy. Efficient use of energy in EAFs can be possible by a good understanding of its working principle and making an optimum production plan (Seker, Memmedov, & Huseyinov, 2017).

The first developments in the field of EAF began in the 19th century. Humphry Davy made the first experimental demonstration to introduce the electric arc in 1810. Pepsy conducted studies on welding systems in 1815. On the other hand, William Siemens was impressed by these studies and obtained a patent for the arc type electric furnace in 1978-79. The first commercial electric arc furnace was developed by Paul Heroult (Seker, 2017).

The basic working principle of EAFs is to facilitate the melting of the metals by converting electrical energy into thermal energy. In this melting process, the energy conversion is performed by the electric arc created between the melted metal and the electrodes. The electric arc is characterized by high current and low voltage. With this method, temperatures of about 3000-4000°C can be obtained. Figure 1 shows the cross-section of a typical Alternating Current (AC) EAF used for melting.

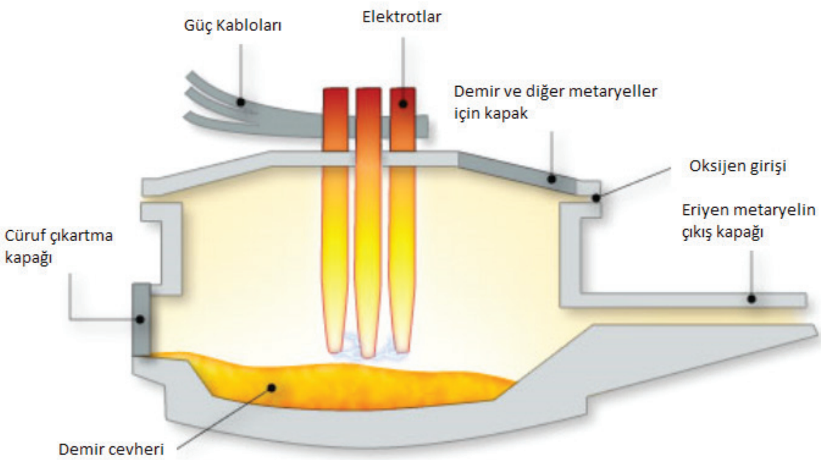


Figure 1. *Cross section and basic schematic representation of Electric Arc Furnace*

Electric arc furnaces mainly consist of three parts (Beaty, 1978);

1. Shell: It consists of steel and sidewalls at the bottom.

2. Hearth: The part consisting of the bottom of the melting pot and **refractories**.

3. Roof: It supports the refractories when one or multiple graphite electrodes are involved, and it can have a water-cooled spherical form.

The EAFs have a non-linear load characteristic. Therefore, they may cause power quality problems in the power grids they are connected to such as harmonics, voltage sags and swells, and reactive power problems (Seker & Memmedov, 2017; Seker, Memmedov, Huseyinov, & Kockanat, 2017). Initially, these problems were neglected for furnaces with small capacity. However, due to the high melting capacities and power variations of EAFs used today, these negativities have reached a level that cannot be neglected.

Depending on the characteristic behavior of the arc created between the melted material and the electrodes, EAFs have four different operating processes: charging, burning, melting, and refining. The operating processes measured during a melting period at an EAF with 60 MVA power are presented in Figure 2. The burning process is the beginning of the melting process. The electrodes are immersed in the molten pool and they are charged to start the melting process. During this process, the electric arc has a short arc length (Seker, Memmedov, Huseyinov, et al., 2017). During the melting process, a large area is formed in the middle of the melted material. The electric arc has a long arc length. Since the metal melted in this process is mobile, the arc length constantly changes. In the refining process, almost all of the metal in the molten pool is melted and the arc length of the electric arc is almost constant. The arc has a stable structure. In this process, the heat generated by the arc keeps the molten metal in a liquid state during its transfer to the ladle furnace (Issouribehere, Issouribehere, & Barbera, 2005)

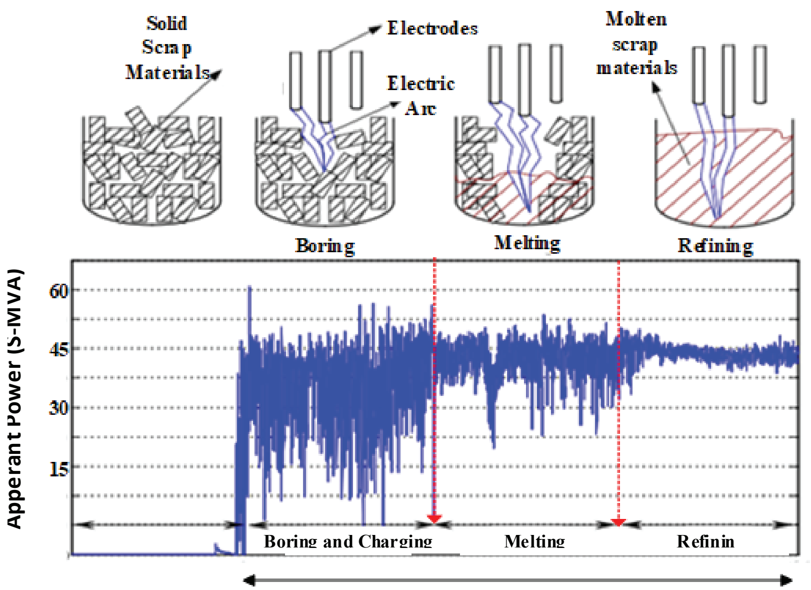


Figure 2. *Apparent power change of 60 MVA EAF in boring, melting and refining processes*

Considering EAFs' energy consumption, they use 53% of the energy for melting in steel production(Göl et al., 2010). Cooling losses, waste gas losses, and settling losses during the melting period constitute total losses. The power change varies depending on the secondary voltage of the furnace transformer. The furnace transformer is usually set to the voltage levels of 10-12 to ensure optimal operating conditions. Also, the furnace performance should be tested at regular intervals and optimum working conditions should be determined when an arc furnace is commissioned. It is aimed to operate EAFs in the maximum efficient way by determining their optimal working conditions. This can be possible by understanding and analyzing the static operating characteristics of EAF well. The operating characteristics of EAFs can be obtained by mathematically expressing the behavior of the electrical circuit elements that make up the system.

This study provides basic information on how to describe the static operating characteristics of an EAF with 60 MVA power. In the second part of the study, the circuit components of the electrical system feeding the electric arc furnace are explained. In the third part, mathematical assumptions used in power calculations in EAFs are explained in detail. The changes in the static characteristics are presented visually in the fourth part. In the last part, the results obtained are given as a general summary.

2. Electrical Circuit Components of the System Feeding the Electric Arc Furnace

Figure 3 shows the circuit components that make up the system feeding the electric arc furnace. As can be seen from the figure, the electrical components that feed the EAF consist of a step-down transformer, EAF transformer, static Var compensation (SVC) system, and harmonic filter groups, busbars, and electrodes(Şeker & Memmedov, 2015). The electrical equivalent of each circuit component constituting this system must be expressed mathematically for defining the static characteristics of the EAF. One of the most effective methods for defining the mathematical equivalent of the system is to create a single-phase equivalent circuit model. A phase equivalent circuit model of the electrical system feeding the Electric Arc Furnace is presented in Figure 4.

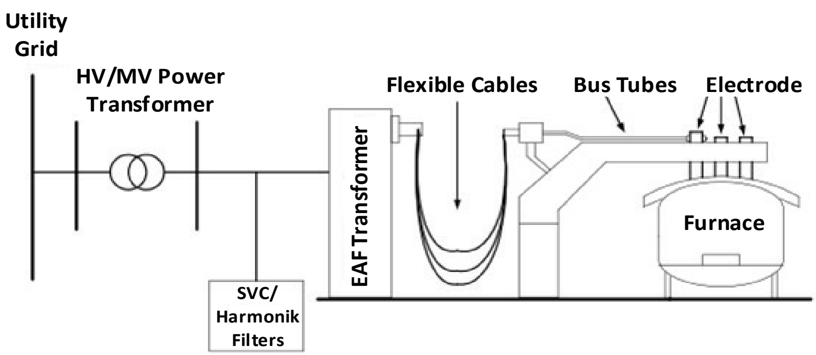


Figure 3. *Electrical circuit elements constituting the EAF system.*

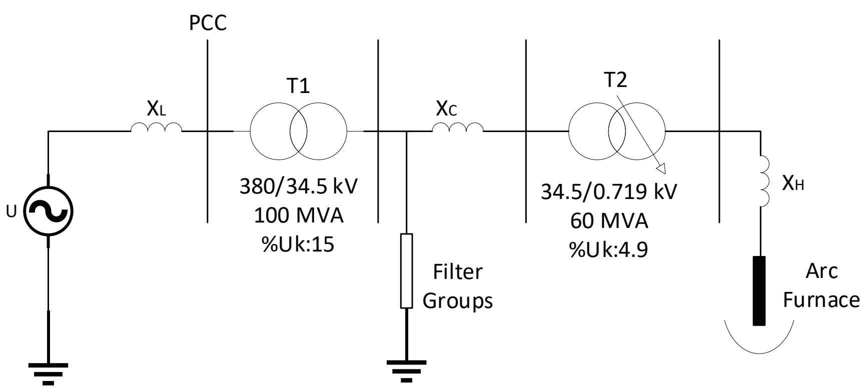


Figure 4. *Equivalent single line diagram of Electric arc furnace.*

3. Power Calculations for Electric Arc Furnace

The system presented in Figure 4 can be addressed in two parts. The equivalent circuit of the primary system includes the resistance in the national grid, step-down transformer, regulators if any, and furnace transformer. The secondary circuit, on the other hand, consists of electrodes and flexible cables.

Usually, the circuit parameters are calculated in the impedance (Z) form. To calculate a phase equivalent circuit model of the EAF shown in Figure 4, it is required to calculate the impedance of the grid, impedance of the step-down transformer, impedance of the furnace transformer, and impedance of the secondary circuit. If the impedance in the system feeding the EAF is shown in the R + jX form, the assumptions presented in equations (1-2) are used in the calculations of the resistance and reactance values;

- For grid impedance, the X/R ratio is assumed to be approximately 10.

$$\frac{X}{R} = \frac{10}{1}; \qquad R = Z.\cos(\theta), \quad X = Z.\sin(\theta) \qquad (1)$$

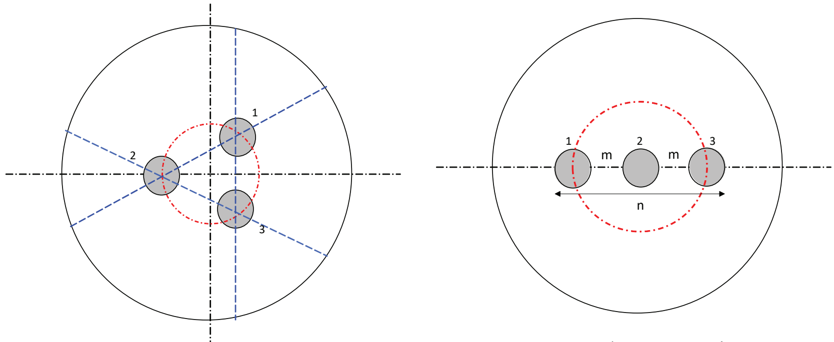
$$\theta = \arctan\left(\frac{X}{R}\right), \qquad \theta = 84{,}29_{o}$$

- The X/R ratio is assumed to be 8 for the impedance values of the transformers.

$$\frac{X}{R} = \frac{10}{1}; \qquad R = Z.\cos(\theta), \quad X = Z.\sin(\theta) \qquad (2)$$

$$\theta = \arctan\left(\frac{X}{R}\right), \qquad \theta = 84{,}29_{o}$$

Thus, the parameters of the grid system feeding the EAF can be easily defined by using these assumptions. However, the major problem is the calculation of the parameters of the secondary circuit that is a part of the EAF transformer. The reactance of the secondary circuit depends on the configuration of the electrodes. Two geometric forms, i.e. coplanar and triangular, are used to calculate the secondary reactance of the electric arc furnace.



A. Triangular Design B. Coplanar Design

Figure 5. *Triangular and Coplanar design of EAF electrodes according to positions.*

A. Triangular Desing

In triangular design, inductance values of three phases are expressed in equations (3-5).

$$L_A = L_S - L_{B_n} + L_{C_n} \quad (3)$$

$$L_B = L_S - L_{A_n} + L_{C_n} \quad (4)$$

$$L_C = L_S - L_{A_n} + L_{B_n} \quad (5)$$

Depending on the inductance values, Self and common inductance values are calculated using equations (6-7).

$$L_{Self} = 2.S.\left(ln\frac{2S}{ds} - 1\right).10^{-9}\ H \quad (6)$$

$$L_{Mutual} = 2.S.\left(ln\frac{2S}{ds} - 1 + \frac{d}{S}\right).10^{-9}\ H \quad (7)$$

In this equations, S, ds and d are defined as fallow;

S ; length of conductor (cm)

ds ; Radius of geometric field

d ; Average geometric distance between conductors.

With the help of the calculated self and common reactance values, the reactance value of each phase is calculated by equations (9-11).

$$L_A = L_{ASELF} - \frac{M_{AB}}{2} + J0.866.M_{AB} - \frac{M_{AC}}{2} - J0.866.M_{AC} \quad (9)$$

$$L_B = L_{BSELF} - \frac{M_{AB}}{2} + J0.866.M_{AB} - \frac{M_{BC}}{2} - JO.866.M_{BC} \tag{10}$$

$$L_C = L_{CSELF} - \frac{M_{BC}}{2} + J0.866.M_{BC} - \frac{M_{AC}}{2} - JO.866.M_{AC} \tag{11}$$

B. Coplanar Design

In the coplenar design, the inductance values of each phase are expressed in equations (12-15).

$$L_A = L_S - L_{B_m} - L_{C_n} \tag{12}$$

$$L_B = L_S - L_{A_m} - L_{C_m} \tag{13}$$

$$L_C = L_S - L_m - L_{A_n} \tag{14}$$

In this expression, the relationship between inductance values is defined as $L_B < L_A = L_C$.

Based on the calculated L value, the reactance resistance of the secondary circuit is defined by the equation (15).

$$X_L = 2.\pi.f.L \tag{15}$$

In this statement;

XL : Reactance resistance of each phase as ohms

f : Utility frequency

L : It is the value of Inductance as Henry (H).

Depending on these assumptions, the circuit impedance of the EAF system consists of five parts.

1) Utility line impedance (R_L+X_L)
2) Stepdown Transformer impedance (R_T+X_T)
3) EAF Transformer impedance (RF+X_F)
4) Seconder circuit impedance (R_S+X_S)
5) Arc resistance (R_A)

The main purpose of EAF is to create an electric arc between electrodes and scrap. This arc is characterized by low voltage and high current. Theoretically, arc is the basic function of an arc furnace. The real power supplied to the furnace is the heat energy generated in the arc. This power is based on the control of Z = U / I arc impedance. In arc furnaces operating with alternating voltage, the arc is extinguished when the current is zero in alternating alternation. When the arc is extinguished, the voltage between the electrode tip and the metal is approximately equal to the secondary voltage of the transformer. This voltage has an effect on ignition of the arc. In order for the arc to be ignited, at least 40 V voltage and 5 kA current are required (Pelfrey D.L., 1980). Steady-state conditions in the electric arc furnace is low power factor. This requires a high reactive energy requirement.

With the help of arc, electrical energy is converted into thermal energy. There is a direct relationship between heat conduction and arc length. If the arc length is short, heat will occur around the electrodes. If the arc length is large, most of the heat will be on the walls of the furnace or towards the furnace cover. The arc length depends on the arc voltage and is directly related to the voltage of the secondary circuit of EAF transformer. In an arc furnace, the arc length is approximately defined by the following equation (16)(Gandhare & Lulekar, 2007).

$$L_a = U_a - 50 \text{ mm} \tag{16}$$

Here; Ua refers to the arc voltage.

If short arc is desired, the voltage value should be kept low. With the increase in arc length, the arc resistance increases and the current decreases and the phase between the current and the voltage changes. If we define the arc resistance as Ra, Ra = 0 represents the short circuit condition of the arc. If Ra = ∞, it indicates the situation in which the arc is extinguished. The power change depending on the arc resistance is shown graphically in Figure 6.

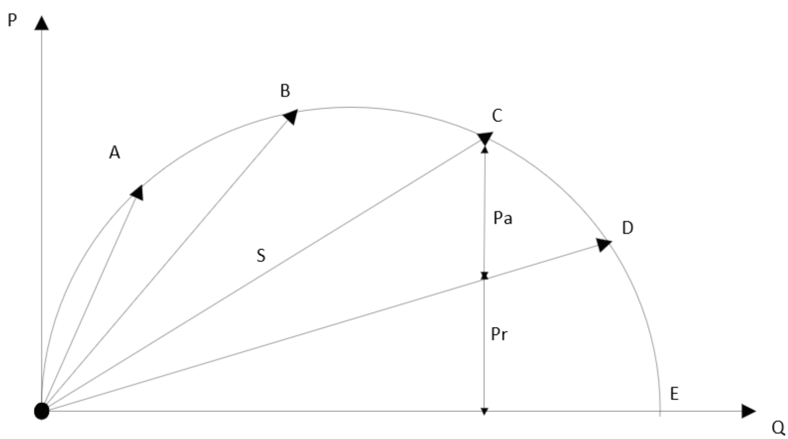


Figure 6. *Power change depending on arc resistance in EAF's.*

In EAF's, the maximum power is obtained when the cos (Φ) value is equal to 0.707 (Maximum power transfer theorem). In this case; active power, reactive power and apparent power could be written as equations in

(17-19) respectively (Bello, 1971; Borrebach, 1969; Montanari, Loggini, Cavallini, Pitti, & Zaninelli, 1994).

$$MW\ (Real\ Power) = \frac{\frac{3.I^2}{10^6} R_{\Phi}}{N} \tag{17}$$

$$MVAR\ (Reactive\ Power) = \frac{\frac{3.I^2}{10^6} X_{\Phi}}{N} \tag{18}$$

$$MVA\ (Apparent\ Power) = \frac{\frac{3.I^2}{10^6} Z_{\Phi}}{N} \tag{19}$$

Here;

I : Secondary circuit current at 0.707 power factor

Z_{Φ}/N : Total impedance for one phase

R/N : Total resistance for one phase

X/N : Total inductance value for one phase.

expresses the values.

Equations (20-22) are used to define the input operating power in the furnace transformer.

$$MVA = \frac{3.I^2(R_F + R_S + R_A)^2 + j(X_F + X_S)^2}{10^6} \tag{20}$$

$$MW = \frac{3.I^2(R_F + R_S + R_A)^2}{10^6} \tag{21}$$

$$MVA = \frac{3.I^2.(X_F + X_S)^2}{10^6} \tag{22}$$

In these equations, parameters are defining as bellow.

I : Seconder circuit current at 0.707 power factor

Rf : Resistance of EAF transformer

Rs : Resistance of seconder circuit

Xs : Reactance of seconder circuit

Ra : Arc resistance

The line current passing through the secondary circuit is calculated with the following expression (23).

$$I_{line} = \frac{E_{\Phi\Phi}}{\sqrt{3}.Z_t} \qquad (23)$$

In these equations;

$E_{\Phi\Phi}$: phase to phase Secondary voltage value

Z_T : Total impedance value of equivalent circuit

The main issue in examining the power characteristics of the EAF is the definition of the powers on the primary side of the EAF transformer. In this case, the power factor (PF) is expressed as equation (24).

$$P.F = \frac{MW.10^6}{E.I.\sqrt{3}} \qquad (24)$$

In (24);

E : primary input voltage between phases

I : primary line current

MW is power value in the primary side of the transformer as Megawatt.

Based on this expression, the power factor (PF) could be defining as Eq. (25);

$$P.F = \frac{(R_F + R_S + R_A)}{(R_F + R_S + R_A)^2 + (X_F + X_S)^2} \qquad (25)$$

In these equations; R_F- Furnace transformer resistance, X_F- Furnace transformer reactance, R_S- Secondary circuit resistance, X_S- Secondary circuit reactance and R_A is arc resistance. Power factor is expressed as cos (Θ) and the angle value is calculated by equation (26).

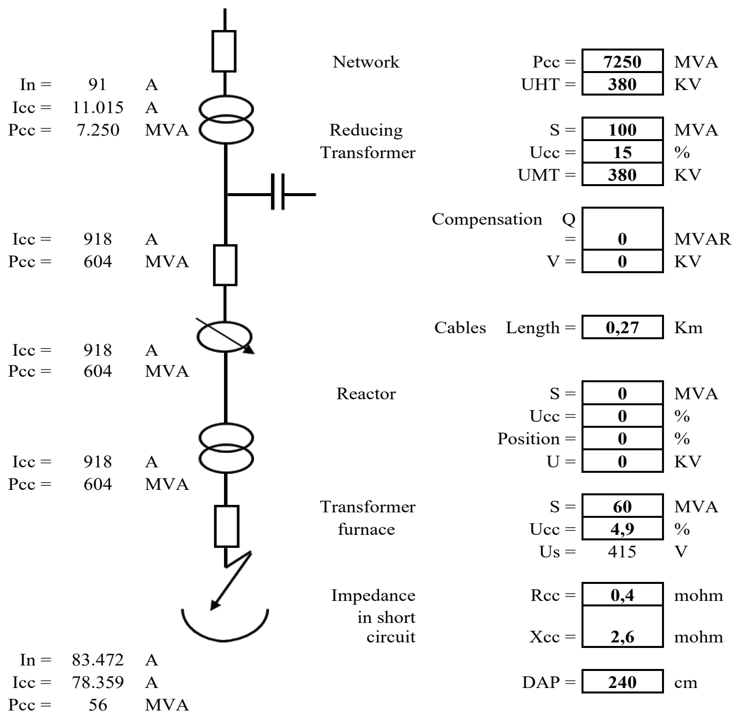
$$\Theta = arctan\left(\frac{MVAR}{MW}\right) \qquad (26)$$

In Eq. (26), MVAR is reactive power as MegaVar at transformer primary side. On the other hand, mw refers to the effective power defined as Megawatt on the primary side of the transformer.

4. Calculation of impedance parameters of EAF

The assumptions presented in Chapter 3 are used to calculate the circuit parameters of the electric arc furnace. The parameters of the system feeding the 60 MVA EAF are shown in Table 1. In this table, it is assumed that the furnace transformer is operated at the 12th stage (Seker, Memmedov, & Huseyinov, 2017).

Table 1. *Circuit parameters of the system feeding the 60 MVA EAF Facility*



Using the mathematical assumptions presented in section 3 and the parameters presented in Table 1, a phase equivalent circuit parameters belonging to the system feeding the EAF are calculated and shown in Table 2.

Table 2. *Impedance parameters calculated based on the tapping values of the EAF transformer (Tap=12).*

	Real parameters		Equivalent Parameters for 719 V Reference voltage	
	R(Ω)	X (Ω)	R (mΩ)	X (mΩ)
Z_{res}	2.845	19.917	0.0034	0.0238
Z_{tab}	30.943	216.60	0.0369	0.2583
Z_{lig}	11	95	0.0000	0.0001
Z_{self}	0	0	0	0
Z_{tf}	16.847	117.927	0.0201	0.1407

In order to calculate the input power of the electric arc furnace, the following procedures are performed (Seker, Memmedov, & Huseyinov, 2017).

1) Operation resistance

$$R_{op} = \sqrt{\frac{(P.F)^2.\left(X_{tot} + X_{utility} + X_{st}\right)^2}{1-(P.F)^2}} \quad (\mathrm{m\Omega}) \tag{27}$$

2) Operation impedance

$$Z_{op} = \sqrt{\left(R_{op}\right)^2 + \left(X_{op}\right)^2} \quad (\mathrm{m\Omega}) \tag{28}$$

3) Operation current (seconder current)

$$I_{op} = \frac{E_{\Phi\Phi}}{\sqrt{3}.Z_{op}} \quad (kA) \tag{29}$$

4) Primer active power in operation current

$$MW_{pri} = \frac{3.\left[R_{op} - \left(R_{utility} + R_{st}\right)^2\right].I_{op}{}^2}{10^3} \tag{30}$$

5) Primer reactive power in operation current

$$MVAR_{pri} = \frac{3.\left(X_f + X_s\right).I_{op}{}^2}{10^3} \tag{31}$$

6) Appearance power in operation current

$$MVA_{pri} = \sqrt{(MW)^2 + (MVAR)^2} \tag{32}$$

The power characteristic curves of the EAF furnace are defined using these defined equations and the circuit parameters calculated for the electric arc furnace load.

5. Static Characteristics of the 60 MVA EAF

The static characteristics of the 60 MVA electric arc furnace defined using the mathematical approaches presented in Chapters 3 and 4 are presented in Figure 7 and Figure 8. The stability index is expressed as the ratio of the reactance value of the system feeding the furnace to the furnace reactance. This ratio can be changed by adjusting the tap settings of the transformers in the system feeding the furnace. The graphical curves presented are defined for the operation of the furnace transformer at the 12th tap for a secondary voltage of 719 V. Similarly, graphics can be defined for all transformer tap values.

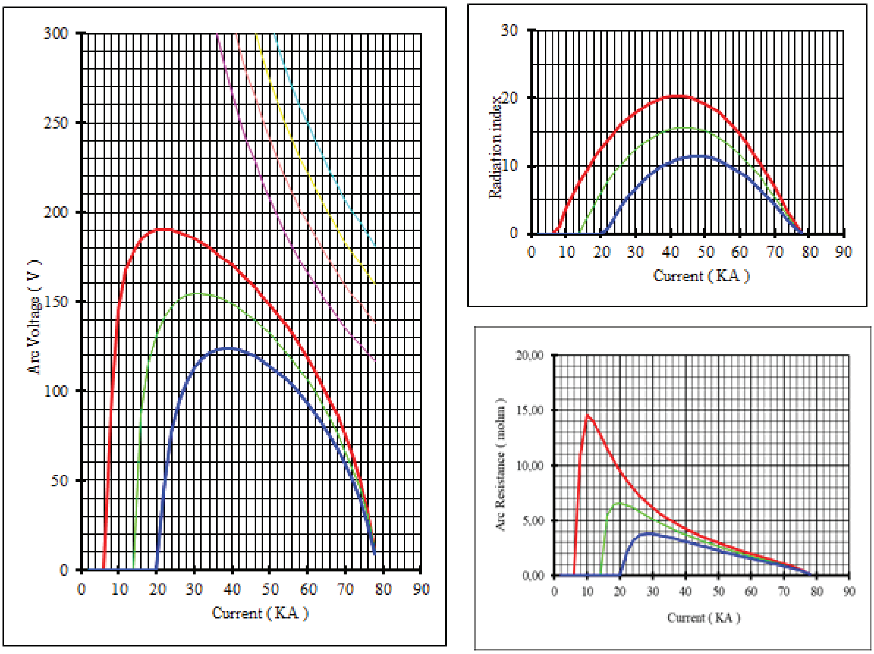


Figure 7. *Static changes of arc voltage, radiation index and arc resistance depending on the stability index (SI) of 0.15, 0.18 and 0.25 (Red: SI = 0.15, Green: SI = 0.18, Blue: SI = 0.25)*

EAF transformatöünün kademe değerlerine bağlı olarak

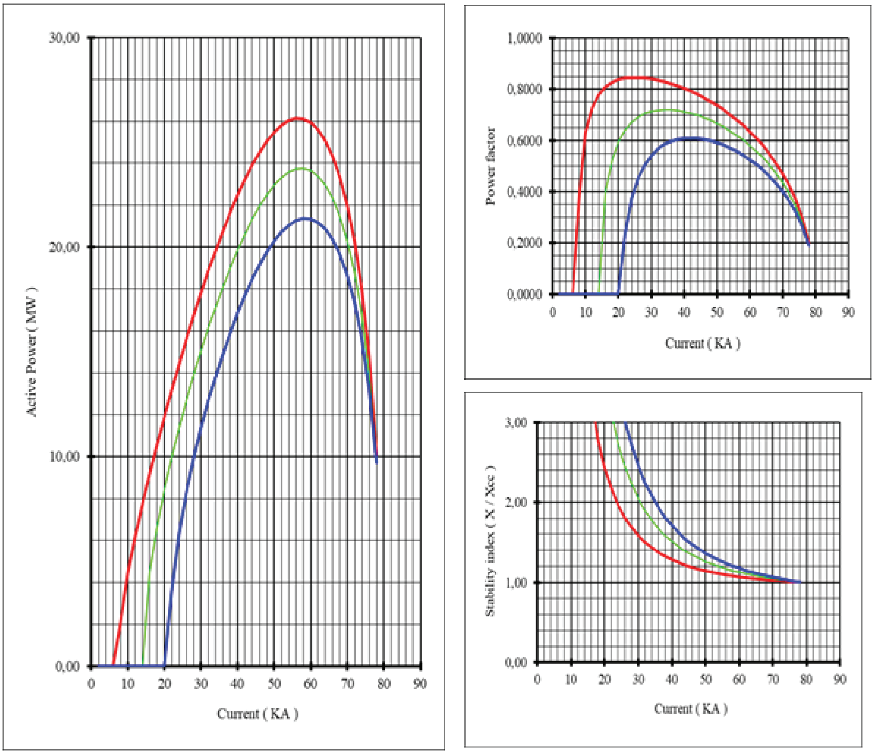


Figure 8. *Static changes of active power, power factor and stability indices depending on current depending on stability index (SI) of 0.15, 0.18 and 0.25. (Red: SI=0.15, Green: SI=0,18, Blue: SI=0,25)*

6. Results

In order to evaluate the power quality problems in electric arc furnaces, it is necessary to define the short circuit currents and optimal operating conditions of EAFs. Determining of these conditions is possible by defining the static power characteristics of EAFs. In this study, the mathematical approaches required to define the static characteristics of 60 MVA EAF are presented. By defining static characteristics, improvement of power quality, determination of short circuit currents and development of protection methods against the effects of these currents can be provided for EAFs. In addition, it is necessary to evaluate the Static characteristics in terms of determining the suitable operating range and transformer tap for optimal operation of EAF. The results show that the power coefficient should be adjusted as 0.707 at the maximum power point in order to ensure optimal operating conditions in the 60 MVA EAF furnace. During the melting process, the voltage level of the transformer should be reduced and EAF transformers should be operated below the melting tap value in the melting process. In melting and other working processes, the X / Xcc stability value of the EAF can be adjusted depending on the transformer tap changes. This will also increase the efficiency of the furnace in terms of optimal use and energy consumption.

References

Bello, J. R. (1971). Fundemental of The Electric Arc Furnace. In *Electric Furnace Proceedings* (p. 219).

Beaty. H. W. (1978). *Standard Handbook for Electrical Engineers* (11th Ed). New York.

Borrebach, E. J. (1969). Maximum Power Operation of Electric Arc Furnace. *Iron and Steel Engineering*, 74.

Gandhare, W. Z., & Lulekar, D. D. (2007). Effect of electrode regulation system on arc parameters in furnace operation. *IECON Proceedings (Industrial Electronics Conference)*, 921–925. Retrieved from https://doi.org/10.1109/IECON.2007.4460303

Göl, M., Salor, Ö., Alboyací, B., Mutluer, B., Çadírcí, I., & Ermis, M. (2010). A new field-data-based EAF model for power quality studies. *IEEE Transactions on Industry Applications*, 46(3), 1230–1242. Retrieved from https://doi.org/10.1109/TIA.2010.2046280

Issouribehere, P. E., Issouribehere, F., & Barbera, G. A. (2005). Power quality measurements and operating characteristics of electric arc furnaces. *2005 IEEE Power Engineering Society General Meeting*, 1, 784–791. Retrieved from https://doi.org/10.1109/pes.2005.1489388

Montanari, G. C., Loggini, M., Cavallini, A., Pitti, L., & Zaninelli, D. (1994). ARC-Furnace model for the study of flicker compensation in electrical networks. *IEEE Transactions on Power Delivery*, 9(4), 2026–2036. Retrieved from https://doi.org/10.1109/61.329535

Mustafa SEKER. (2017). *Elektrik ark fırınlarının elektrik şebekesine etkilerinin deneysel incelenmesi ve bilgisayar destekli modellenmesi*. İnönü Üniversitesi/Fen Bilimleri Enstitüsü, Elektrik Elektronik Anabilimdalı, Doktora Tezi.

Pelfrey D.L. (1980). Specifying Arc Furnace Transformer. pİTTSBURG.

Schau., H., Stade, D. (1994). Mathematical Modeling of Three-Phase Arc Furnace. *Proceedings IEEE ICHPS VI*.

Seker, M., & Memmedov, A. (2017). An experimental approach for understanding v-i characteristic of electric arc furnace load. *Elektronika Ir Elektrotechnika*, 23(3). Retrieved from https://doi.org/10.5755/j01.eie.23.3.18328

Şeker, M., & Memmedov, A. (2015). The modelling of electric arc furnace load with Matlab/Simulink. In *2015 23rd Signal Processing and Communications Applications Conference, SIU 2015 - Proceedings*. Retrieved from https://doi.org/10.1109/SIU.2015.7130390

Seker, M., Memmedov, A., & Huseyinov, R. (2017). Investigation of characteristics of electric arc furnace operation. In *2016 National Conference on Electrical, Electronics and Biomedical Engineering, ELECO 2016*.

Seker, M., Memmedov, A., Huseyinov, R., & Kockanat, S. (2017). Power quality measurement and analysis in electric arc furnace for Turkish electricity transmission system. *Elektronika Ir Elektrotechnika*, 23(6). Retrieved from https://doi.org/10.5755/j01.eie.23.6.19691

Workshop, J. I. (2016). Globolisation and Consolidation in the Steel Industry. In *New Delphi* (p. 17).

Bölüm 30

BIOMEDICAL APPLICATIONS OF ELECTROSPUN NANOFIBER MEMBRANES

Seyedehnegar Arabi[1]

Seyedali Naziri Mehrabani[1]

İsmail Koyuncu[1,2]

Bihter Zeytuncu Gökoğlu[1]

1 [1]National Research Center on Membrane Technologies, İstanbul Technical University, İstanbul, Turkey

[2]Environmental Engineering Department, İstanbul Technical University, İstanbul, Turkey

Corresponding author at: Phone: +90 212 2853473; fax: +90 212 285 6667

E-mail address: bihtzeytuncu@itu.edu.tr (B. Zeytuncu)

Introduction

The advent of membrane science arose from 1861 through dialysis experiments using synthetic membranes [1]. Thanks to the extraordinary properties and extensive applications, membrane technology attracted much attention up to now. Since membrane technology is an interdisciplinary field, numerous investigations have been conducted from different fields to expand the membranes applications in other research areas by developing the structural properties. The cognizance of membrane definition and its nature should be considered essential for the next levels of investigations.

Membranes' classification depends on several parameters. However, they could be classified as biological and synthetic membranes. Synthetic membranes were considered remarkably up to now from its arrival. Membranes' synthesis methods have been investigated widely due to their applications. Different synthesis methods are introduced to improve their properties according to their utilization area. Membranes' synthesis phase inversion (IP), track-etching, stretching, interfacial polymerization, and are mostly used techniques in membrane technology. However, among the various synthesis methods, electrospinning is considered a special process due to its facile procedure and unique properties of electrospun fabricated membranes.

Since they can be produced by several methods, different membrane structures and properties were reported. These membranes' diverse structures and properties result in further applications of membranes in various areas such as water treatment, textile, and biomedical. Membranes' application in this field is attributed to their ability to facilitate the natural mechanisms of human body. Although there are numerous publications regarding membrane technology, this chapter is prepared to summarize the recent membrane technology developments. For this aim, the basic principles of membranes, structure, classification, and synthesis methods are described in this chapter. The electrospinning synthesis method is illustrated in detail as well as their effective parameters. Furthermore, biomedical applications of membranes are explained.

1. Membrane

1.1. Definition and classification

Membranes' description varies from the aspect of its functions and processes. Some of these functions could be separation, purification, and dialysis. Without considering the operations, the membrane could be defined as a semi-permeable barrier between two phases. In other words, the membrane is able to separate two phases and control the transport of some species in the mixture [2]. The minimum work for this operation could be assigned thermodynamically with the free enthalpy of molecules.

As it is defined in Eq. (1), the minimum work should be equal or greater than the free enthalpy of mixing. The separation mechanism takes place due to the chemical potential difference between the two sides, which they are feed and permeate [3]. This chemical potential difference is attributed to the gradient of pressure, temperature, concentration, and electrical potential, as it is shown in Figure 1.

$$W_{min} \geq \Delta G_m = \Delta H_m - T\Delta S_m \quad (1)$$

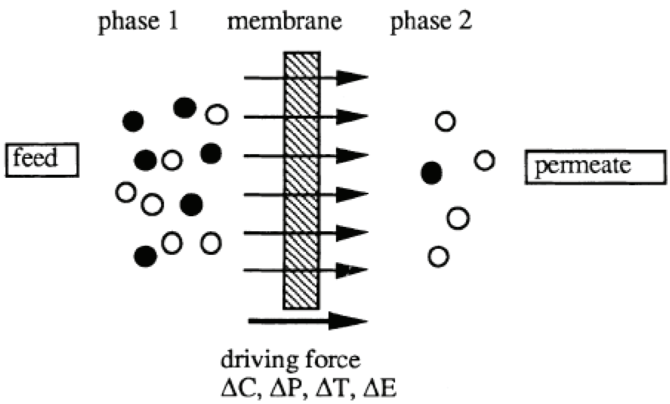


Figure 1. *Schematic of a two-phase separated system by a membrane. Reprinted with permission from [3].*

The investigations of membrane technology regarding the structural properties were done to explore and improve the membranes' performance or efficiency. Without considering membrane classification and processes, some material and structural dependent parameters, such as permeability, retention factor, flux, and separation factor, are considered critical to evaluate the membranes' performance. Permeability is the basic principle of the membrane mechanism based on the interaction between component and membrane material [4] As it is clarified in Eq. 2, permeability is dependent on the solution and diffusion.

$$P = D.S \quad (2)$$

The other parameter, retention, is defined as a retained amount of solute instead of water passing. This transport mechanism is affected by the structure of membranes. Specifically, this retention is attributed to the ratio of molecular size to pore size. The retention equation is defined below where the C_p and C_f are solute concentration in permeate and solute concentration in feed sides, respectively. This retention amount is between 0 and 1, that the retention equal to 1 is affiliated to the complete solute retention [3,5].

$$R = 1 - \frac{c_p}{c_f} \tag{3}$$

The separation factor (α) is considered the permeation rate of two components (A & B). It is defined by the concentration of components in both permeate and feed sides [3]. If we assume X as the concentration of components in the permeate and Y as the concentrations in the feed side, the separation can be calculated in the following as Eq. 4. In this case, the permeate rate of component A is considered to greater than component B. It is important to note that no separation occurs when $\alpha_{A/B}$ and $\alpha_{B/A}$ Equals to 1.

$$\alpha_{A/B} = \frac{Y_A/Y_B}{X_A/X_B} \tag{4}$$

Flux is another principal parameter to compare the performance of membranes. It has a direct proportion with the pore diameter of membranes [6]. It could be assumed that the membrane is a cylindrical capillary pore in which liquids flow through the membrane. Thus, we will have Eq. 5 where μ, l, d, and Δp refers to the liquid viscosity, pore length, pore diameter, and pressure difference, respectively.

$$\text{Pore (q)} = \frac{\pi d^4}{128\ \mu\ l} \cdot \Delta p \tag{5}$$

The number of pores is defined by Eq. 6 where ε refers to porosity :

$$N = \ . \frac{4}{\pi d^2} \tag{6}$$

Clearly, if we know the number of pores (N), the flux can be calculated;

$$\text{Flux (J)} = N. \frac{\pi d^4}{128\ \mu\ l} . \Delta p \tag{7}$$

Since the definition has no effective presentation about the structure or function of membranes, the next subject that should be considered primarily is related to the classification of membranes. Membranes' classification pertains to numerous specifications. Some of these specifications could be related to nature, driving force, structure, transport form, charge, and size. It could be mentioned the passive and active transport for transportation form, homogenous or heterogeneous for structure, neutral or charged membranes to exemplify these specifications for further understanding. Meanwhile, the driving force is diversified into pressure, concentration, and temperature. Besides, membranes can be categorized into porous, nonporous, and liquid membranes according to their separation process or structure. Schematics of these membranes are shown in Figure 2.

Membranes classification related to their nature could be mentioned in two main groups; biological and synthetic. Biological membranes also are assorted to living and non-living ones. On the other hand, synthetic membranes are divided into organic (polymer or liquid) and inorganic (metallic, glass, and ceramics). Synthetic membranes attract much attention due to the tunable parameters during the fabrication. Polymeric membranes are the most considered ones since they have remarkable properties such as stability of thermal and chemical properties, fouling resistance, cost-effect, and ease of fabrication. Numerous researches have been conducted to explore and optimize the synthesis methods of polymeric membranes. Some of these synthesis methods are described in detail [2,3,6].

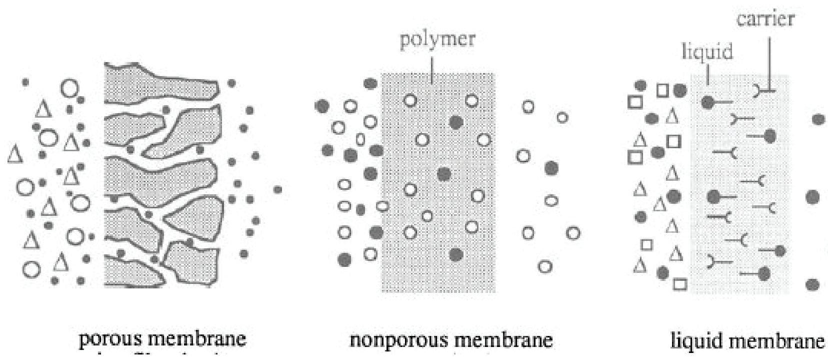


Figure 2. *Schematic drawing of the three basic membranes. Reprinted with the permission from Ref. [2]*

1.2. Synthesis methods

The synthesis of polymeric membranes has been investigated widely due to their excellent performance and applications. Meanwhile, different polymeric structures can be fabricated from various fabrication techniques. Thus, the selection of synthesis methods is dependent on the choice of the material and the preferred structure [7]. Among the various techniques such as phase inversion, interfacial polymerization, track-etching, and stretching, the electrospinning technique is elucidated in detail in the next section thanks to the distinct properties and structure of electrospun polymeric membranes.

2. Electrospinning: an advanced nanofiber fabrication technique

2.1. Introduction to Electrospinning

Since the early 1990s, electrospinning is defined as a primary facile method that has been used for the fabrication of micro/nano-sized fibers to employ in various applications, including water treatment, filtration systems, and biomedical applications like drug delivery systems [8-10].

Compared with other morphologies, specific characteristics of electrospun fibers composed of micro and nanopores like high specific surface area, ideal cell adhesion, and proliferation convinced researchers to employ them in the application were mentioned in the previous part of this section. Moreover, the highly porous structure of electrospun mats originated from micro/nano-sized fibers' involvement, making them one of the superior candidates in tissue engineering. In summary, high porosity, high tensile strength, simple processing, the versatility of the material, simple control of fiber diameter, interconnected pore structure, and high surface/volume ratio are some of the electrospun fibrous matrices incomparable properties that have been persuaded by scientists and researchers in the last decades. As the explained unique features of electrospun fibers in many publications, this book's primary goal is to define the electrospinning process, effects of different parameters on resultant fibrous mats, and its usages due to the extensive demands in various sectors like biomedical applications. Pharmaceutical repositories, tissue engineering, wound healing, sensors, reinforcement, sound absorption, and filtration are some of the biomedical applications of electrospun matrices. It must be mentioned that there have been numerous studies in the field of electrospinning mechanisms, electrospinning conditions, and fibrous mats characteristics. At the same time, modification of micro/nano-sized electrospun fibers during electrospinning process including core-shell, hollow, side-by-side, multilayer, twisted, and porous surface structures which cause to the production of fibrous matrices with more desirable physiochemical attributes had become hot topics in the recent years [11].

2.2. Principle and Mechanism of Electrospinning

As it is shown in Figure 3, a high voltage supply, a dope driven system or a syringe pump, a spinneret, and a grounded metal collector are four fundamental segments of classic electrospinning set up [12]. During the electrospinning procedure, surface tension, Coulombic repulsion force, viscoelastic force, gravity, electrostatic, and air drag force are diverse forces are applied to the charged liquid jet to elongate it in the polymeric fibers form under an electrical field. Commonly, the formation of electrospun fibers through electrospinning is carried out in the following three steps:

1. Onset of the polymer solution's emission after applying forces emit of polymers solution to the syringe pump, rectilinear jet development

2. Making a structure with looping and spiraling pathways and solidification of generated fibers after Evaporation of solvents

3. nanofibers accumulation [13].

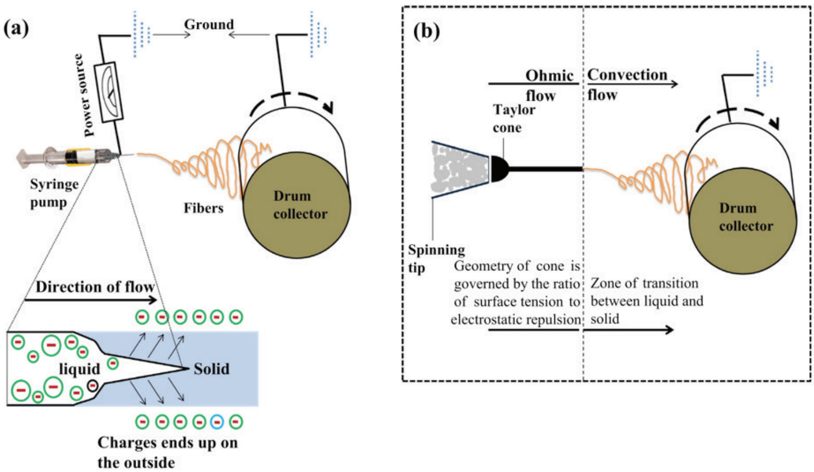


Figure 3. *Schematic depicting electrospinning setup and phenomenon of electrospinning [14].*

In the first step of the electrospinning process, during the exerting of the electrical potential difference between spinneret's tip and grounded collector, the shape of the generated droplet is reshaped into a conical the shape is named Taylor Cone. The distance is moved by the polymer solution's jet after applying electrical force has a critical value that is symmetrical of an applied electrical field, conductivity, dope flow rate, and asymmetrical to other operators like liquid density the current passing through the jet. The fibers' diameter reduction from micrometers to nanometers is a required occurrence that can be controlled by the electrically driven non-axisymmetric bending of the polymer's solution jet at high frequencies. Finally, the produced fibers in the previous step are gathered in the grounded apparatus. Some other factors, such as the design of spinnerets and apparatus parts of electrospinning, are other substantial effective aspects that alter the morphologies and structures of obtained nanofibers [15].

In the following part of this chapter, the significant factors during electrospinning procedure including applied electrical voltage, the solution's follow rate, the distance between grounded collector and needle, polymer concentration, and solution viscosity, the conductivity of polymer solution, the role of utilized solvent, humidity, and temperature will be discussed in detail. Due to these parameters' direct effect on the resulting fibers' characteristics, and to obtain the desired product, understanding all of these parameters is necessary for researchers who use electrospinning in their works.

2.3. Effects of Parameters on Electrospinning

2.3.1. Applied voltage

Flowing current, which originated from high-voltage power storage into a polymer solution, is the reason for converting droplets from spherical to Taylor cone and producing ultrafine nanofibers. The amount of critical applied voltage causes droplet deformation during the electrospinning process is variable between polymers [16]. Based on recent reports, by increasing the applied voltage, we will create nanofibers with the beaded structure correlated to the Taylor cone decline and polymers jet's flowing rate increment [17].

2.3.2. The following rate of solution

The polymeric solution flow rate is another influential factor in specifying the morphology of the electrospun nanofibers. The amount of critical applied voltage to fabricate electrospun nanofibers with a more uniform beaded structure is variable among different polymer solutions. For instance, in polystyrene, the bead formation of electrospun nanofibers is onsetting by increasing the applied voltage above 0.10 mL/min and stoped by lowering the voltage to 0.07 mL/min. The flow rate increment of polymeric solution over the critical applied voltage cause to increase of (1) pore size, (2) diameter of the electrospun nanofibers. The imperfect drying of nanofibers jet during the electrospinning process is why the transformation of produced nanofibers' features. Because of the shifting effects of the applied voltage in electrospinning proceeding, a minimum flow rate in preferred due to the (1) balance maintenance between the leaving polymeric solution and its replacement by another one during jet foundation (2) construction of a permanent and recede jet cone. In other words, because of the effect of addition and reduction of polymeric solution's flow rate, which is a factor influence produced fibers morphology, a minimum flow rate is selected in all electrospinning processes. Also, by applying the minimum voltage, the balance between leaving and receding polymeric jet is preserved until the end of the process.

In some other cases, besides the bead formation, ribbon-like faults formation occurs during the electrospinning process. This type of imperfections that are constructed because of the flow rate growth is related to some of the solvent's features, such as non-evaporation and low stretching factors. As well as, surface charge density is another factor that causes to make changes in the electrospun nanofibers morphology [14].

2.3.3. The distance between the needle's tip and collector

Similar to the mentioned factors, the distance between the tip of the metallic needle and collector affects the fabricated electrospun nanofibers morphology. Furthermore, like the electrical applied voltage, viscosity, and flow rate, by changing the type of the polymeric solution, the distance between the needle tip and collector is shifted in the electrospinning operation. The distance can change the morphology of the electrospun nanofibers straightforwardly due to its dependence on the evaporation rate, deposition time, and whipping or instability [18]. Thus, the critical distance is an essential factor in obtaining smooth and uniform electrospun nanofibers, and any shifting in the distance affects the nanofibers' soft morphology [19]. Based on multiple studies that have been accomplished up to now, by keeping the critical distance small, the produced nanofibers will be more defective and larger, and by increasing the distance bigger, the nanofiber's diameter is reduced [18, 20, 21]. Although, in some cases, changing the critical distance has no effect on the created nanofibers morphology [22].

2.3.4. Polymer concentration and solution viscosity

The stretching of the charged jet plays a significant role in changing the polymeric solution's concentration. This means, by choosing the polymeric jet with a lower concentration in the electrospinning process, the chains of the polymers entangle with each other before achieving the collector and results in the formation of more beaded electrospun nanofibers [23,24]. Likewise, due to the polymeric solution's increment, the viscosity of the solution is gained and causes the growth of polymeric chain entanglement. Meanwhile, by increasing the polymeric chain entanglement, the surface tension is decreased and causes to form electrospun nanofibers with uniform beadles' structure. However, rising the polymeric solution's concentration beyond the critical value causes to dry of the solution and district the needle's tip. These blocks prevent the flow of solution and result in the construction of defective beaded electrospun nanofibers [23]. Figure 4 represents the droplets shape deformation by changing the solutions viscosity.

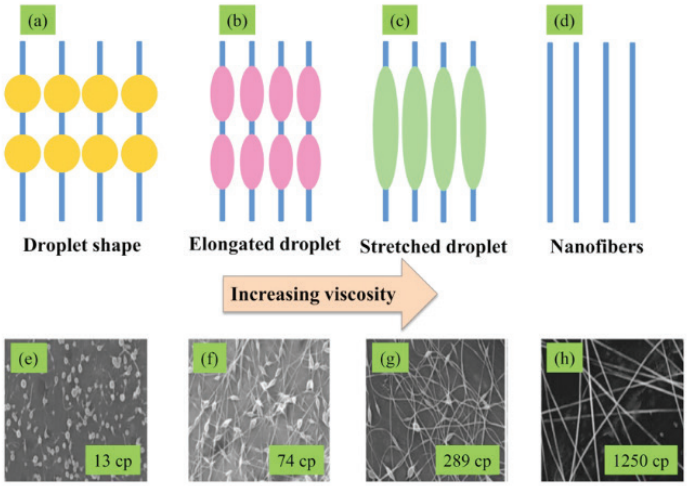


Figure 4. *variation in morphology of electrospun nanofibers of PEO with viscosity: (a-d) schematic and (e-h) SEM micrographs [25-27].*

2.3.5. Role of solvent in electrospinning

The type of utilized solvent in the electrospinning process plays a significant role in fabricating smooth and beaded electrospun nanofibers. In order to choose the type of solvent, solvent characteristics must be considered. For example, the solubility of the solvent is one of the factors that affect the resulted nanofibers in the electrospinning procedure. The solvents with complete solubility are the preferred ones to synthesize electrospun nanofibers. In addition to the solvent's solubility, the boiling point of the solvent is another key factor that influences the electrospinning procedure. Selecting the solvents with a moderate boiling point is suggested because they can vaporize easily and at a high rate from the needle tip to the collector. In contrast to the excellences of choosing the solvents with a moderate boiling point, in some cases, they are not preferred due to their higher evaporation rate, which causes them to block the needle tip by drying in that segment [28,29].

Based on multiple studies, the porosity of the fabricated electrospun nanofibers is affected by the solvent type. Production nanofibers with high porosity are taken place using two different types of solvents, which play various roles during the electrospinning process. One of them is considered as a solvent, and the other one acts as a non-solvent component. Diversity of the evaporation rate of solvent and non-solvent is the reason for the phase separation and produces electrospun nanofibers with highly porous structure [29].

3. Electrospun membranes in biomedical applications

3.1. Introduction to biomedical applications

Due to the stunning properties of membranes, they have been investigated in the biomedical research area almost for two decades. It would be essential to have knowledge of some basic medical principles to understand the membranes' role in this field. The investigations have been conducted to produce and improve properties of membranes such as diameter, alignment, porosity, biodegradability, and surface modification for cell migration, which is the primary mechanism for repairing or regeneration of tissues in types of a blood vessel, nerve, heart, skin, and musculoskeletal system [30]. The primary purpose of the following part of this chapter is to illustrate the recent advancement of the electrospinning process and electrospun fibers in biomedical's sectors. Likewise, the drug-loaded proceeding on the electrospun fibers and various polymers used in multiple pharmaceutical systems will be discussed. Figure 5 indicates the primary applications of electrospun fibers in the biomedical sector.

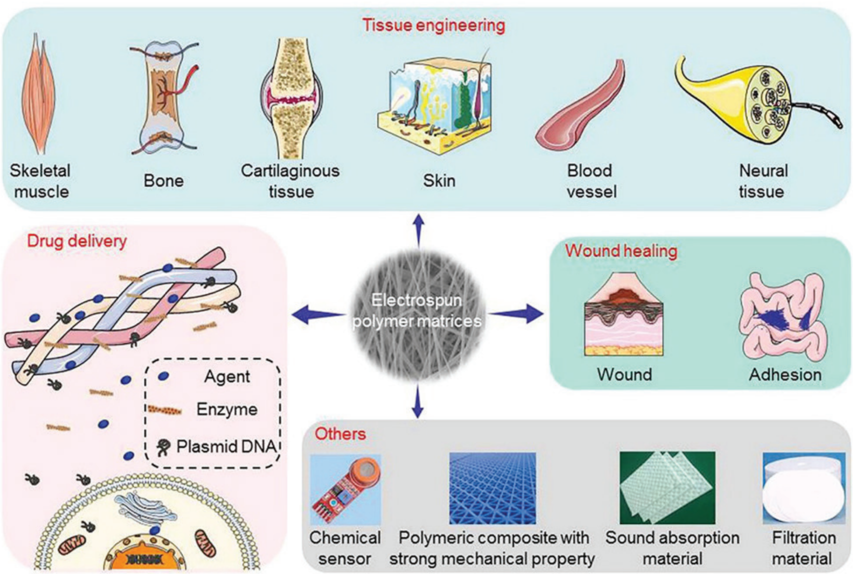


Figure 5. *multiple biomedical applications of electrospun fibers [11].*

3.2. Tissue engineering

It was observed some kinds of wounds are not healed naturally due to the vast lesions. The critical role of nanofibers as biological materials to bridle this issue can be described by their ability to mimic the natural mechanism of the skin by implanting it on a wound site [11,31]. It was considered that they must possess some basic preconditions such as biocompatibility, interconnected 3-dimensional structure, and biodegradability. The biodegradable ability must parallel the rate of tissue regeneration [32-34]. Several methods, such as particulate leaching [35], thermally-induced phase separation [36], molecular self-assembly [37], and electrospinning [38] were reported to obtain these materials. Electrospinning has been recommended highly because of tunable biological and mechanical properties.

Cell migration or cell repair is considered as the main mechanism of healing that is defined as the incorporation of scaffolding materials that assist tissue growth by providing biochemical instruction and mechanical support. These scaffolding materials must have some properties in terms of architecture and composition to mimic the duty of native extracellular matrix (ECM) as well as possible [30,31]. Generally, cells are besieged by ECMs with diverse compositions across tissues. To define the ECMs, they possess structural proteins (collagen), complexes of proteins and polysaccharides for digging the structural proteins, and adhesive glycoproteins for attachment of cells [39]. ECMs are vital for tuning the phenotype and cells' operation within the response to injury [40].

A wide range of natural and synthetic polymeric nanofibers such as PLA [41], PGA [42], PPy [43], PANI [44], PCL [45], PLGA [46], PU [47], etc., were used to mimic the ECMs in tissue regeneration. It might be challenging to assign the exact process of tissue regeneration and scaffolds effect. However, scaffolds act as temporary ECMs for enhancing the recruitment of cells as a growth factor that is essential for the construction of permanent ECMs [30]. Numerous investigations have been done to promote the useful parameters and properties of nanofibers to enhance cell migration in vitro and ex vivo. Alignment and diameter effects of nanofibers are the most crucial parameters.

It was observed that the direction and speed of cell migration are affected by the alignment of PLA. Random nanofibers show shorter distance due to the migration of cells to all possible directions. On the other hand, uniaxial alignment of PLA increases the migration speed thanks to the cell migration along nanofiber. Uniaxial aligned PLA shows the ability to close a 2.25 mm gap in 2 days, while the random PLA stays static after days [48-50]. Figure 6 is represented to understand specifically these explanations. The effect of diameter was reported by Wang et al. They exerted three different PLA diameters (small, intermediate, and large) in their investigations and found that cells are able to migrate to the furthest in large diameter fiber. Some surface functionalization methods such as plasma treatment, graft polymerization, and the wet chemical was used to prevail some limitations like hydrophobicity and adhesion inability of fibers.

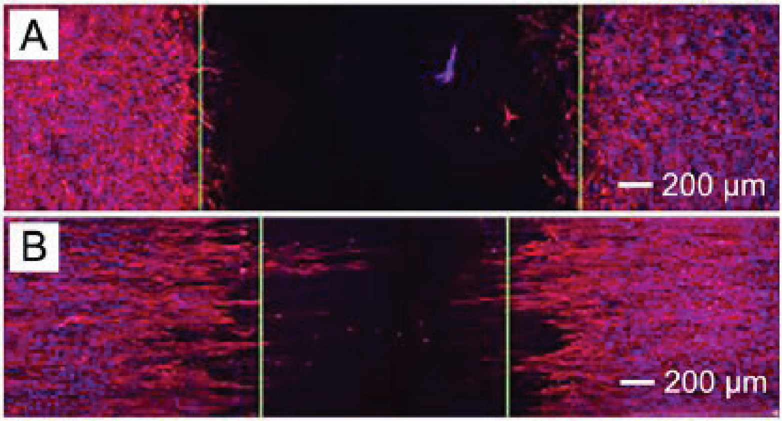


Figure 6. *Fluorescence micrographs showing the migration of astrocytes on (A) random and (B) uniaxially aligned PLA fibers. Reprinted with the permission from [51]*

3.3. Pharmaceutical applications

Generally, the distinct types of nanostructures that can be charged on small/macro-molecules using various formulations, approaches, and technologies to carry pharmaceutical compounds to different human body organs are defined as drug delivery systems. In recent years, liposomes, micelles, Pickering emulsions, dendrimers, and polymeric nanoparticles (NPs) are some of the desirable nanomaterials which have been selected to employ in drug delivery systems. Among the multiple types of drug delivery systems, the nanoparticle (NP) based systems are the most favorable drug delivery systems due to their ability to transport much more drugs to the body's organisms. Polylactic acid (PLA), poly lactic-co-glycolic acid (PLGA), and polyethylene glycol (PEG) are some of the polymers which can be used in nanoparticle-based drug delivery systems. Based on further investigations, increasing the drug dosage in drug delivery systems leads to organize much more controllable systems. By the same token, during the last decade, due to the remarkable features of the polymeric nanofibers synthesized using the electrospinning technique, they have been preferred in biomedical applications such as cancer therapy and drug delivery systems. The simplicity of the electrospinning method, a large ratio of the surface area to volume, high porosity between the interlayer fibers, lower impediment of the mass transfer to the body organs, flexibility, tunable morphology, and high mechanical strength are specified as electrospun fibers characteristics which make them as beneficial NPs to use in biomedical applications in an extensive range [52-54]. Table. 1 shows the various types of electrospun nanofibers' usages in different fields of the biomedical sector.

Table 1. *Electrospun drug-loaded nanofibers in drug delivery applications [55].*

No	Drug incorporated	Polymer
1	Ofloxacin/gellan	PVA
2	Ondansetron hydrochloride	PVA
3	Ibuprofen/carvedilol	PCL
4	Ibuprofen	Polyvinylpyrrolidone (PVP)
5	Caffeine/riboflavin	PVA
6	Donepezil HCL	PVA
7	Clobetasol-17-propionate	Eudragit RS100/PVP/PEO
8	Salmon calcitonin (SCT)	Sodium alginate/PVA
9	Ciprofloxacin HCL	PVA/chitosan/PCL
10	Ketoprofen (KP)	PVA/poly(acrylic acid)/multiwalled carbon nanotubes
11	Asiaticoside (AC)	Cellulose acetate
12	Vitamin A acid/Vitamin E	Cellulose acetate
13	Meloxicam	PVA
14	Curcumin/diclofenac/vitamin B12	Chitosan/phospholipid
15	Tetracycline hydrochloride/ phenytoin sodium	Cellulose acetate/PVA
16	Collagen/salicylic acid (SA)	PVA
17	Amoxicillin	Polyethylene glycol (PEO)/PLGA
18	Cefradine/5-flourouracil	PLGA/gelatin
19	Doxorubicin	PLA
20	Paclitaxel	Polyurethane (PU)/Eudragit L100-55

3.3.1. Drug loaded carriers for medical therapy

Over the last few years, because of the weaknesses of the drug delivery systems based on microspheres, hydrogels, and micelles, drug-containing electrospun fibers have been paid much attention among researchers to employ in biomedical applications [56]. The ease of electrospinning procedure, adaptability of synthesized fibers, and hampering the drug eruption during the first hours causes much more stable systems; drug-containing electrospun fiber matrices are the most desirable drug-delivery systems. As an illustration, placing the selective drugs close by the electrospun fibers' surfaces resulting from the imperfective dissolving of the medicines in the polymeric solution causes rapid diffusion of the drugs into the cells and controls the medicine delivery mechanism. Because of the degradation possibility of the electrospun fibers in drug-delivery systems, the drug release process takes place with zero-order kinetics.

Over and above that, the electrospun nanofibers' alternative properties are an essential factor that makes them one of the excellent tunable and constant drug delivery systems. It means that, by changing the electrospun nanofibers features, we will be able to regulate more effective drug delivery systems. For instance, electrospun fibers with a larger diameter and lower drug concentration show an extended zero-order periodicity during the drug release to the cells. As well as, electrospun nanofibers are classified as PH-responsive drug delivery systems. This means that PH shifting in the specific times can moderate the drug release process in the pathophysiological sicknesses [57].

In the coaxial electrospinning technique, which has been carried out to control drug delivery systems, the inner jet is formed using the solution of the bio-macromolecules and then results to form the outer jet by performing the co-electrospun polymeric solution. Two distinct solutions, including organic polymeric solution and aqueous-based solution, are applied to the coaxial electrospinning system to synthesize the modified fibers using the coaxial electrospinning method. In other terms, the interaction between polymeric solution and the aqueous-based molecules is lowered and leads to sustain the bioactivity factor of the biological molecules [58]. Figure 7 represents the fabrication method of poly(vinyl phenol)/poly(ε-caprolactone) (PVP/PCL) electrospun nanofiber matrices using coaxial electrospinning technique that is categorized as a core-shell fiber membrane. As it is shown in Figure 7, in the coaxial electrospinning system, vancomycin hydrochloride or VAN was employed as a drug. The combination of the graphene oxide (GO) and PVP was utilized as the core, and the PCL is used to prepare the shell part of the desired matrix [59]. In summary, because of the drugs' sustainability and continuous release to the cells, the core-shell electrospun nanofibers are preferred in various drug delivery systems [60].

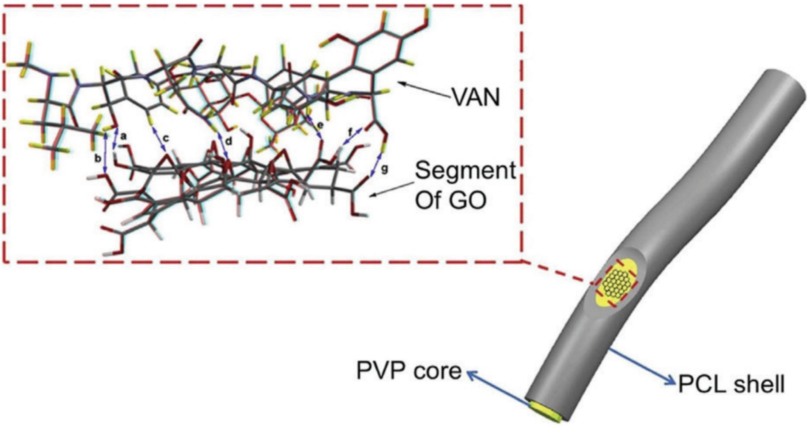


Figure 7. *schematic illustration of PVP/PCL core-shell nanofiber [59].*

3.3.2. Nucleic acid delivery for gene therapy

Owning to the topography resemblance between the synthesized electrospun nanofibers and the natural ECMs, the electrospun nanofibers have been admitted as one of the ideal drug delivery systems [61].

In order to manufacture the electrospun nanofibers with desirable properties, the nanofibers scaffolds must be (1) controllable, (2) sustainable to apply in gene delivery systems. Besides, the particular biomolecules must be wholly preserved by electrospun nanofibers' scaffolds before the entire drug release [62]. The prominent role of the electrospun nanofibers mats controls the charge of the biomolecules, displays the forwarded signals, and regulates tissue engineering by physically supporting the drug release to the desired cells. For the successful drug transmission to the body organs and preserve the nucleic acids from existing enzymes in the human body, viral and non-viral carriers have been selected to carry out the nucleic acid delivery process. After placing the nucleic acids on the fibers mats, the nucleic acids combine with the carriers and then unify with the electrospun nanofibers. The molecules' bioactivity protection, high efficiency of gene transmission, and the kinetic of the gene release are challenges to developing gene therapy using nucleic acids [63]. Furthermore, the electrospinning procedure's multiple parameters such as molecular weight, the concentration of the polymeric solution, and plasmid DNA (pDNA) affect the gene therapy's efficiency. In other words, the electrospinning parameters' increment leads to fabricate electrospun fibers with regulated features and changeable characteristics of the nucleic acid transfection to the desired cells [64].

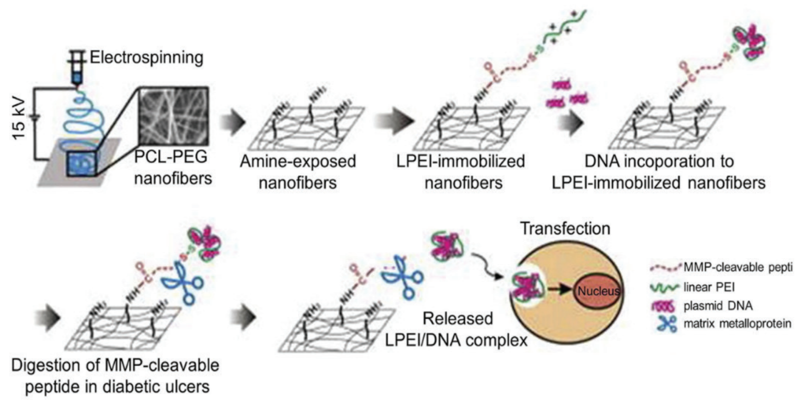


Figure 8 shows the steps of synthesizing the MMP-cleavable electrospun nanofibers for use in local gene therapies.

Figure 8. Schematic diagram of MMP-responsive electrospun nanofibrous matrix for gene delivery [65]

3.3.3. Enzyme carriers for biomedical application

Due to the enzyme's repercussions, including sensing, active packaging, and bio-processing, they are known as one of the promising drug delivery systems in biomedical applications. In the enzyme delivery systems, enhancing the ratio of the polymeric matrix's area to its volume causes the construction of the enzyme delivery systems with the high efficiency and least diffusional restrictions. Like the other types of drug delivery systems, the electrospinning technique is preferred to fabricate the nanofibers to utilize as enzyme carriers to the body organisms. The higher efficiency of the enzymes' catalytic activity during the hydrolyzing olive oil process using lipase biomolecule is the reason for the election of the electrospinning technique to apply in the desired enzyme delivery system. Moreover, composite nanofibrous membranes are the other type of nanofibers that have been recommended by researchers to use in enzyme delivery systems. The composite nanofiber membranes are composed of two parts: (1) enzymes that act as a surfactant, and (2) the polymeric material to confirm the enzyme's activity. Compared to the other types of drug delivery systems, because of the durability of the composite nanofibrous matrices and the possibility of the simple recovery of these systems, they are known as the ideal enzyme delivery systems. Based on recent studies, some of the novel polymeric electrospun nanofiber membranes with a highly porous structure can protect the human body from inflammatory illnesses by cytokine protection from the interstitial body fluids [66].

4. Conclusion

Since 1861, multiple types of membranes have been utilized to separate undesirable elements from the solutions selectively. The functionalization of membrane separation is based on the molecular sieving, chemical, and electrochemical affinity mechanisms. Besides the water treatment applications, membrane technology has been performed in other sectors, including energy, agriculture, foods, medicine, air protection, and biomedical. Tissue engineering, cancer therapy, wound healing, sensors, and drug delivery are some biomedical sectors that have exerted various membranes in various applications. Additionally, phase inversion, interfacial polymerization, track-etching, stretching, and electrospinning are techniques used in the membrane fabrication process.

In the last decade, among the different types of membranes that have been applied in biomedical applications, electrospun nanofiber membranes have become the most popular one. Due to the intrinsic characteristics of the electrospun nanofiber polymeric membranes, they are frequently used in tissue engineering and drug delivery systems. The ease of the electrospinning technique, high porosity of the produced nanofibers, the large ratio of the surface area/body organs, flexibility, and the electrospun's excellent mechanical and thermal properties nanofibers make them one of the best membranes in biomedical applications.

In drug delivery systems, electrospun polymeric nanofibers have been accomplished as the carriers for medical therapy, gene therapy, and enzyme delivery to the organ's cells. Hence, liposomes, micelles, emulations, dendrimers, and polymeric nanoparticles are the other types of nanomaterials that have been used in drug delivery systems. Among the numerous types of polymeric nanoparticles, poly(lactic acid) or PLA, poly lactic-co-glycolic acid (PLGA), and polyethylene glycol (PEG) is the eligible ones in drug delivery systems.

References

[1] Graham, T. (1860). Liquid Diffusion Applied to Analysis. *Proceedings of the Royal Society of London*, *11*, 243-247.

[2] De Jong, J., Lammertink, R. G., & Wessling, M. (2006). Membranes and microfluidics: a review. *Lab on a Chip*, *6*(9), 1125-1139.

[3] Mulder, M., & Mulder, J. (1996). *Basic principles of membrane technology*. Springer science & business media.

[4] Wijmans, J. G., & Baker, R. W. (1995). The solution-diffusion model: a review. *Journal of membrane science*, *107*(1-2), 1-21.

[5] Mehta, A., & Zydney, A. L. (2005). Permeability and selectivity analysis for ultrafiltration membranes. *Journal of membrane science*, *249* (1-2), 245-249.

[6] Baker, R. W. (2012). *Membrane technology and applications*. John Wiley & Sons.

[7] Lalia, B. S., Kochkodan, V., Hashaikeh, R., & Hilal, N. (2013). A review on membrane fabrication: Structure, properties and performance relationship. *Desalination*, *326*, 77-95.

[8] Agarwal, S., Wendorff, J. H., & Greiner, A. (2008). Use of electrospinning technique for biomedical applications. *Polymer*, *49*(26), 5603-5621.

[9] Pham, Q. P., Sharma, U., & Mikos, A. G. (2006). Electrospinning of polymeric nanofibers for tissue engineering applications: a review. *Tissue engineering*, *12*(5), 1197-1211.

[10] Liang, D., Hsiao, B. S., & Chu, B. (2007). Functional electrospun nanofibrous scaffolds for biomedical applications. *Advanced drug delivery reviews*, *59*(14), 1392-1412.

[11] Ding, J., Zhang, J., Li, J., Li, D., Xiao, C., Xiao, H., ... & Chen, X. (2019). Electrospun polymer biomaterials. *Progress in Polymer Science*, *90*, 1-34.

[12] Sun, B., Long, Y. Z., Zhang, H. D., Li, M. M., Duvail, J. L., Jiang, X. Y., & Yin, H. L. (2014). Advances in three-dimensional nanofibrous macrostructures via electrospinning. *Progress in Polymer Science*, *39*(5), 862-890.

[13] Reneker, D. H., & Fong, H. (2006). Polymeric nanofibers: introduction.

[14] Haider, A., Haider, S., & Kang, I. K. (2018). A comprehensive review summarizing the effect of electrospinning parameters and potential applications of nanofibers in biomedical and biotechnology. *Arabian Journal of Chemistry*, *11*(8), 1165-1188.

[15] Liao, Y., Loh, C. H., Tian, M., Wang, R., & Fane, A. G. (2018). Progress in electrospun polymeric nanofibrous membranes for water treatment: Fabrication, modification and applications. *Progress in Polymer Science*, *77*, 69-94.

[16] Laudenslager, M. J., & Sigmund, W. M. (2012). Electrospinning. *Encycl. Nanotechnol., Springer Netherlands, Dordrecht*, 769-775.

[17] Deitzel, J. M., Kleinmeyer, J., Harris, D. E. A., & Tan, N. B. (2001). The effect of processing variables on the morphology of electrospun nanofibers and textiles. *Polymer*, *42*(1), 261-272.

[18] Matabola, K. P., & Moutloali, R. M. (2013). The influence of electrospinning parameters on the morphology and diameter of poly (vinyledene fluoride) nanofibers-effect of sodium chloride. *Journal of Materials Science*, *48*(16), 5475-5482.

[19] Bhardwaj, N., & Kundu, S. C. (2010). Electrospinning: a fascinating fiber fabrication technique. *Biotechnology advances*, *28*(3), 325-347.

[20] Baumgarten, P. K. (1971). Electrostatic spinning of acrylic microfibers. *Journal of colloid and interface science*, *36*(1), 71-79.

[21] Wang, T., & Kumar, S. (2006). Electrospinning of polyacrylonitrile nanofibers. *Journal of applied polymer science*, *102*(2), 1023-1029.

[22] Zhang, C., Yuan, X., Wu, L., Han, Y., & Sheng, J. (2005). Study on morphology of electrospun poly (vinyl alcohol) mats. *European polymer journal*, *41*(3), 423-432.

[23] Haider, S., Al-Zeghayer, Y., Ali, F. A. A., Haider, A., Mahmood, A., Al-Masry, W. A., ... & Aijaz, M. O. (2013). Highly aligned narrow diameter chitosan electrospun nanofibers. *Journal of Polymer Research*, *20*(4), 105.

[24] Pillay, V., Dott, C., Choonara, Y. E., Tyagi, C., Tomar, L., Kumar, P., ... & Ndesendo, V. M. (2013). A review of the effect of processing variables on the fabrication of electrospun nanofibers for drug delivery applications. *Journal of Nanomaterials*, *2013*.

[25] Luzio, A., Canesi, E. V., Bertarelli, C., & Caironi, M. (2014). Electrospun polymer fibers for electronic applications. *Materials*, *7*(2), 906-947.

[26] Zander, N. E. (2013). Hierarchically structured electrospun fibers. *Polymers*, *5*(1), 19-44.

[27] Fong, H., Chun, I., & Reneker, D. H. (1999). Beaded nanofibers formed during electrospinning. *Polymer*, *40*(16), 4585-4592.

[28] Lannutti, J., Reneker, D., Ma, T., Tomasko, D., & Farson, D. (2007). Electrospinning for tissue engineering scaffolds. *Materials Science and Engineering: C*, *27*(3), 504-509.

[29] Sill, T. J., & Von Recum, H. A. (2008). Electrospinning: applications in drug delivery and tissue engineering. *Biomaterials*, *29*(13), 1989-2006.

[30] Xue, J., Wu, T., Dai, Y., & Xia, Y. (2019). Electrospinning and electrospun nanofibers: methods, materials, and applications. *Chemical reviews*, *119*(8), 5298-5415.

[31] Miguel, S. P., Figueira, D. R., Simões, D., Ribeiro, M. P., Coutinho, P., Ferreira, P., & Correia, I. J. (2018). Electrospun polymeric nanofibres as wound dressings: A review. *Colloids and surfaces B: Biointerfaces*, *169*, 60-71.

[32] Boffito, M., Sartori, S., & Ciardelli, G. (2014). Polymeric scaffolds for cardiac tissue engineering: requirements and fabrication technologies. *Polymer international*, *63*(1), 2-11.

[33] Guilak, F., Butler, D. L., Goldstein, S. A., & Baaijens, F. P. (2014). Biomechanics and mechanobiology in functional tissue engineering. *Journal of biomechanics*, *47*(9), 1933-1940.

[34] Orth, P., Zurakowski, D., Alini, M., Cucchiarini, M., & Madry, H. (2013). Reduction of sample size requirements by bilateral versus unilateral research designs in animal models for cartilage tissue engineering. *Tissue Engineering Part C: Methods*, *19*(11), 885-891.

[35] Levengood, S. K. L., & Zhang, M. (2014). Chitosan-based scaffolds for bone tissue engineering. *Journal of Materials Chemistry B*, *2*(21), 3161-3184.

[36] Akbarzadeh, R., & Yousefi, A. M. (2014). Effects of processing parameters in thermally induced phase separation technique on porous architecture of scaffolds for bone tissue engineering. *Journal of biomedical materials research Part B: applied biomaterials*, *102*(6), 1304-1315.

[37] Cheng, T. Y., Chen, M. H., Chang, W. H., Huang, M. Y., & Wang, T. W. (2013). Neural stem cells encapsulated in a functionalized self-assembling peptide hydrogel for brain tissue engineering. *Biomaterials*, *34*(8), 2005-2016.

[38] Dhand, C., Ong, S. T., Dwivedi, N., Diaz, S. M., Venugopal, J. R., Navaneethan, B., ... & Beuerman, R. W. (2016). Bio-inspired in situ crosslinking and mineralization of electrospun collagen scaffolds for bone tissue engineering. *Biomaterials*, *104*, 323-338.

[39] Aamodt, J. M., & Grainger, D. W. (2016). Extracellular matrix-based biomaterial scaffolds and the host response. *Biomaterials*, *86*, 68-82.

[40] Hussey, G. S., Dziki, J. L., & Badylak, S. F. (2018). Extracellular matrix-based materials for regenerative medicine. *Nature Reviews Materials*, *3*(7), 159-173.

[41] Saxena, A. K., Marler, J., Benvenuto, M., Willital, G. H., & Vacanti, J. P. (1999). Skeletal muscle tissue engineering using isolated myoblasts on synthetic biodegradable polymers: preliminary studies. *Tissue engineering*, *5*(6), 525-531.

[42] Riboldi, S. A., Sampaolesi, M., Neuenschwander, P., Cossu, G., & Mantero, S. (2005). Electrospun degradable polyesterurethane membranes: potential scaffolds for skeletal muscle tissue engineering. *Biomaterials*, *26*(22), 4606-4615.

[43] Xiong, G. M., Yuan, S., Wang, J. K., Do, A. T., Tan, N. S., Yeo, K. S., & Choong, C. (2015). Imparting electroactivity to polycaprolactone fibers with heparin-doped polypyrrole: modulation of hemocompatibility and inflammatory responses. *Acta biomaterialia*, *23*, 240-249.

[44] Hosseinzadeh, S., Mahmoudifard, M., Mohamadyar-Toupkanlou, F., Dodel, M., Hajarizadeh, A., Adabi, M., & Soleimani, M. (2016). The nanofibrous PAN-PANi scaffold as an efficient substrate for skeletal muscle differentiation using satellite cells. *Bioprocess and biosystems engineering*, *39*(7), 1163-1172.

[45] Tallawi, M., Rosellini, E., Barbani, N., Cascone, M. G., Rai, R., Saint-Pierre, G., & Boccaccini, A. R. (2015). Strategies for the chemical and biological functionalization of scaffolds for cardiac tissue engineering: a review. *Journal of the Royal Society Interface*, *12*(108), 20150254.

[46] Kumbar, S. G., Nukavarapu, S. P., James, R., Nair, L. S., & Laurencin, C. T. (2008). Electrospun poly (lactic acid-co-glycolic acid) scaffolds for skin tissue engineering. *Biomaterials*, *29*(30), 4100-4107.

[47] Khil, M. S., Cha, D. I., Kim, H. Y., Kim, I. S., & Bhattarai, N. (2003). Electrospun nanofibrous polyurethane membrane as wound dressing. *Journal of Biomedical Materials Research Part B: Applied Biomaterials: An Official Journal of The Society for Biomaterials, The Japanese Society for Biomaterials, and The Australian Society for Biomaterials and the Korean Society for Biomaterials*, *67*(2), 675-679.

[48] Liu, Y., Franco, A., Huang, L., Gersappe, D., Clark, R. A., & Rafailovich, M. H. (2009). Control of cell migration in two and three dimensions using substrate morphology. *Experimental cell research*, *315*(15), 2544-2557.

[49] Mi, H. Y., Salick, M. R., Jing, X., Crone, W. C., Peng, X. F., & Turng, L. S. (2015). Electrospinning of unidirectionally and orthogonally aligned thermoplastic polyurethane nanofibers: Fiber orientation and cell migration. *Journal of Biomedical Materials Research Part A*, *103*(2), 593-603.

[50] Ottosson, M., Jakobsson, A., & Johansson, F. (2017). Accelerated wound closure-differently organized nanofibers affect cell migration and hence the closure of artificial wounds in a cell based in vitro model. *PLoS One*, *12*(1), e0169419.

[51] Hurtado, A., Cregg, J. M., Wang, H. B., Wendell, D. F., Oudega, M., Gilbert, R. J., & McDonald, J. W. (2011). Robust CNS regeneration after complete spinal cord transection using aligned poly-L-lactic acid microfibers. *Biomaterials*, *32*(26), 6068-6079.

[52] Li, W. J., Laurencin, C. T., Caterson, E. J., Tuan, R. S., & Ko, F. K. (2002). Electrospun nanofibrous structure: a novel scaffold for tissue engineering. Journal of Biomedical Materials Research: An Official Journal of The Society for Biomaterials, The Japanese Society for Biomaterials,

and The Australian Society for Biomaterials and the Korean Society for Biomaterials, 60(4), 613-621.

[53] Venugopal, J., Prabhakaran, M. P., Low, S., Choon, A. T., Zhang, Y. Z., Deepika, G., & Ramakrishna, S. (2008). Nanotechnology for nanomedicine and delivery of drugs. *Current pharmaceutical design*, *14*(22), 2184-2200.

[54] Zhang, Y., Liu, S., Wang, X., Zhang, Z. Y., Jing, X. B., Zhang, P., & Xie, Z. G. (2014). Prevention of local liver cancer recurrence after surgery using multilayered cisplatin-loaded polylactide electrospun nanofibers. *Chinese Journal of Polymer Science*, *32*(8), 1111-1118.

[55] Mwiiri, F. K., & Daniels, R. (2020). Electrospun nanofibers for biomedical applications. In Delivery of Drugs (pp. 53-74). Elsevier.

[56] Xue, J., He, M., Liu, H., Niu, Y., Crawford, A., Coates, P. D., ... & Zhang, L. (2014). Drug loaded homogeneous electrospun PCL/gelatin hybrid nanofiber structures for anti-infective tissue regeneration membranes. *Biomaterials*, *35*(34), 9395-9405.

[57] Jiang, J., Xie, J., Ma, B., Bartlett, D. E., Xu, A., & Wang, C. H. (2014). Mussel-inspired protein-mediated surface functionalization of electrospun nanofibers for pH-responsive drug delivery. *Acta biomaterialia*, *10*(3), 1324-1332.

[58] Saraf, A., Baggett, L. S., Raphael, R. M., Kasper, F. K., & Mikos, A. G. (2010). Regulated non-viral gene delivery from coaxial electrospun fiber mesh scaffolds. *Journal of Controlled Release*, *143*(1), 95-103.

[59] Yu, H., Yang, P., Jia, Y., Zhang, Y., Ye, Q., & Zeng, S. (2016). Regulation of biphasic drug release behavior by graphene oxide in polyvinyl pyrrolidone/ poly (ε-caprolactone) core/sheath nanofiber mats. *Colloids and surfaces B: biointerfaces*, *146*, 63-69.

[60] Li, J., Xu, W., Li, D., Liu, T., Zhang, Y. S., Ding, J., & Chen, X. (2018). Locally deployable nanofiber patch for sequential drug delivery in treatment of primary and advanced orthotopic hepatomas. *Acs Nano*, *12*(7), 6685-6699.

[61] Kang, M. S., Kim, J. H., Singh, R. K., Jang, J. H., & Kim, H. W. (2015). Therapeutic-designed electrospun bone scaffolds: mesoporous bioactive nanocarriers in hollow fiber composites to sequentially deliver dual growth factors. *Acta biomaterialia*, *16*, 103-116.

[62] Luu, Y. K., Kim, K., Hsiao, B. S., Chu, B., & Hadjiargyrou, M. (2003). Development of a nanostructured DNA delivery scaffold via electrospinning of PLGA and PLA–PEG block copolymers. *Journal of controlled release*, *89*(2), 341-353.

[63] Benoit, D. S., & Boutin, M. E. (2012). Controlling mesenchymal stem cell gene expression using polymer-mediated delivery of siRNA. *Biomacromolecules*, *13*(11), 3841-3849.

[64] Lee, S. J., Oh, S. H., Liu, J., Soker, S., Atala, A., & Yoo, J. J. (2008). The use of thermal treatments to enhance the mechanical properties of electrospun poly (ε-caprolactone) scaffolds. *Biomaterials*, *29*(10), 1422-1430.

[65] Kim, H. S., & Yoo, H. S. (2010). MMPs-responsive release of DNA from electrospun nanofibrous matrix for local gene therapy: in vitro and in vivo evaluation. *Journal of Controlled Release*, *145*(3), 264-271.

[66] Zhang, J., Zheng, T., Alarçin, E., Byambaa, B., Guan, X., Ding, J., ... & Li, Z. (2017). Porous electrospun fibers with self-sealing functionality: An enabling strategy for trapping biomacromolecules. *Small*, *13*(47), 1701949.

Chapter 31

FUNDAMENTAL CONCEPTS IN SMART MANUFACTURING

***Zeynep Adak*[1]**

1 Dr., Gebze Technical University, Department of Industrial Engineering, 41400, Gebze/Kocaeli, zeynepadak@gtu.edu.tr, ORCID: 0000- 0001-7654- 0773

1. Introduction

There has been vast number of research in smart manufacturing in the last decade. It is a collective term to define an integrated and collaborative manufacturing model which embrace cutting-edge technologies to collect real-time data, transmit and analyze shop-floor data, take predictive actions, respond to dynamic demand, make more accurate forecasts, monitor manufacturing equipment health status, and take proactive actions and have other numerous benefits in return. The main goals in smart manufacturing are reducing costs, increasing product quality and throughput, enabling personalized and smart products, and reaching a resilient and sustainable mode of manufacturing. National strategies and initiatives are introduced to support and facilitate transform to smart manufacturing, among those are *Advanced Manufacturing Partnership* and *Industrial Internet* in US, *Made in China 2025* in China, *Industry 4.0* in Germany*, UK Industry 2050* strategy in England, *Society 5.0* in Japan, *New Industrial France* in France, *Factories of the Future* in EU, and *Manufacturing Innovation 3.0* in South Korea.

Smart manufacturing paradigm is defined by National Institute of Standards and Technology (NIST) – an agency of the US Department of Commerce – as "fully-integrated, collaborative manufacturing systems that respond in real-time to meet changing demands and conditions in the factory, in the supply network, and in customer needs" (NIST, 2017). Other paradigms such as Industry 4.0, cyber-physical production systems and cloud manufacturing were also introduced to define the manufacturing system with the same characteristics.

Six pillars of smart manufacturing were proposed as follows (Kusiak, 2018): (1) Manufacturing technology and processes, (2) Materials, (3) Data, (4) Predictive engineering, (5) Sustainability, and (6) Resource sharing and networking. Built on these pillars a smart factory -a name used to describe an enterprise in smart manufacturing paradigm- is a sustainable and resilient one with adaptive, secure, and on-demand services, shared, digital and integrated manufacturing resources, communicating, autonomous and smart cyber-physical systems (CPS) objects, decreased human-intervention, increased self-optimization, and decentralized real-time decision-making.

A manufacturing system is described by a reference architecture that includes descriptions of components, interactions among them, activities they perform as well as the rules and constraints of those activities. A service-oriented smart manufacturing system architecture was proposed by NIST as given in Figure 1. Moghaddam et al. (2018) proposed 6 characteristics that a reference architecture for smart manufacturing would possess.

Among them are the following:

1. It will be a service-oriented architecture (SOA) model adapted to manufacturing environment.

2. Various capabilities of Internet of Things (IoT) objects (machine, sensor, software, idea, concept) will be standardized and published as services.

3. SOA will offer better interoperability, standardization, and reusability, will have loose-coupling and decreased complexity, and include highly scalable operations.

4. End-to-end and real-time value network coordination would be enabled by SOA. It will also enhance local autonomy and self-reconfigurability in the operative level.

Characteristics associated with smart manufacturing are listed as (Mittal et al., 2017): accuracy, adaptability, agility, asset self-awareness, autonomy, composability, compositionality, context awareness, decentralized, digital presence, distributed, flexibility, fully automated, heterogeneity, information appropriateness, integrability, interoperability, modularity, networkability, proactivity, reliability, resilience, responsiveness, reusability, robustness, scalability, and sustainability.

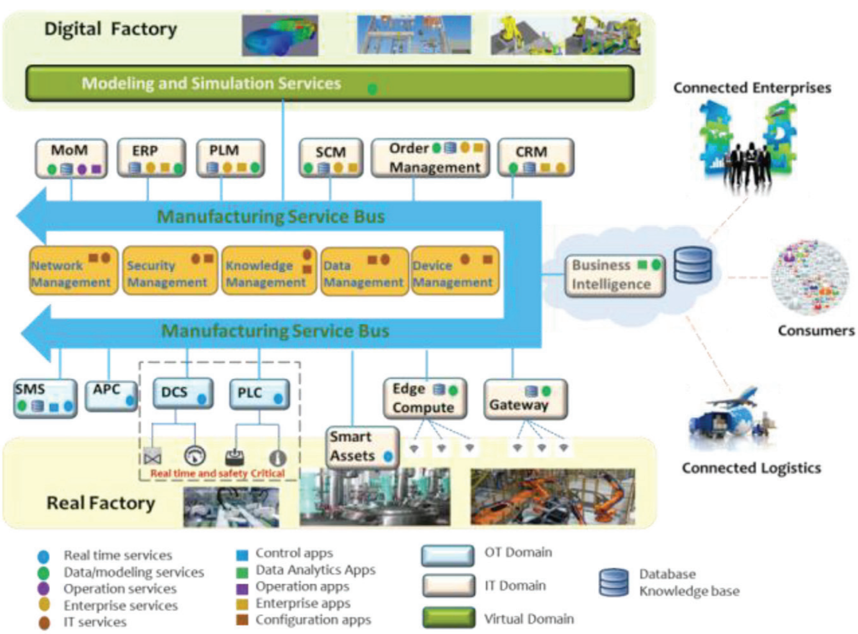


Figure 1. *NIST service-oriented smart manufacturing system architecture (Lu et al., 2016)*

Resiliency and sustainability in manufacturing are two important concepts discussed in the literature. Smart manufacturing paradigm is expected to lead to resilient and sustainable production. A resilient and robust organization is less vulnerable to disruptions in supply chain and the factory. Resiliency attributes include energy, materials, assets, processes, transport, supply-chain and communications, while Kusiak (2019) extended the attributes by adding logistics, efficiency, productivity, capacity, dependability, quality, compatibility, societal values, workforce and sustainability. Sustainability refers to giving the least damage to environment while manufacturing products. Reusability of products and materials, avoiding industrial waste, technologies for reducing carbon emissions and clean energy options are significant at this point. Current manufacturing practices fail to address this issue, while sustainability will be a core characteristic of a smart factory.

Smart manufacturing paradigm is developed on the realm of computing, hence the number of technologies associated is huge. Mittal et al. (2017) reported 38 different technologies. Knowledge about underlying technologies of a system would make it easier to understand its dynamics. In this study, essential concepts and technologies associated with smart manufacturing systems are defined. These include CPS, IoT, Big Data, data-driven analysis, cloud computing, digital twin, augmented reality (AR)/ virtual reality (VR), and direct digital manufacturing (DDM). In the last section, future challenges and perspectives in smart manufacturing are discussed, together with additional concluding remarks.

2. Concepts and technologies

CPS, IoT, Big Data, data-driven analysis, cloud computing, digital twin, AR/VR and DDM consist core elements of a smart manufacturing enterprise in any industry. The technologies are interrelated and have overlapping parts.

2.1. Cyber-Physical Systems (CPS)

CPS defines a system where physical objects and processes are integrated with the cyber computing space. It is a concept embraced by both enterprises and industrial consortia across the globe. “Cyber-physical production system” is a term used interchangeably with smart manufacturing to refer to the very-same production paradigm. Key elements of a CPS include interconnected computations and computing systems, and communication and interaction with physical assets of manufacturing, leading to convergence of virtual and physical worlds.

A 5-level architecture was proposed for CPS in manufacturing as given in Figure 2, which provides a guideline to develop and deploy CPS in an Industry 4.0-factory (Lee et al., 2015). Connection level defines accurate, real-time, and reliable data acquisitions from embedded sensors, controllers or manufacturing management and control systems. Machine or asset-based information is generated from the data in the conversion level, enabling self-awareness in the factory floor. The information in this level is valuable in assessing machine health, measuring performance, and predicting degradation for every machine. Cyber level gathers all the data and information about the individual machines and builds a collective, higher-level information through both machine-machine and time-machine analysis, enabling self-comparison and future predictions. Advanced data-driven analysis is conducted at this level. At the cognition level, knowledge is produced and used to take priority actions and make optimizations. Feedback generated due to the acquired knowledge constitutes the configuration level and enables self-configure and self-adaptive machines. Corrective and preventive actions are decided on at this level.

"Cyber-physical manufacturing services (CPMS)" is a notion proposed by Lu and Ju (2017) to define a distributed architecture of manufacturing ecosystem. It was defined as manufacturing services with cyber and physical parts, where the cyber part represents computing and network services, while the physical part is responsible for material processing, storage, and transportation. Schematic representation of the CPMS-based smart manufacturing is given in Figure 3. The scheme has nodes to represent traditional IT domain services, while the CPMS service nodes are shown as circles which represent processing, storage and transportation services provided by workstations, conveyor belts, robots, cells, shops, plants, factories, and entire enterprises.

2.2. Internet of Things (IoT)

Internet of Things (IoT) refers to an interconnection of *things* through embedded tiny electronics such as sensors and actuators. In a manufacturing setting, these "things" are parts, machines, conveyor belts, work in process, workers, equipment, materials, warehouses, etc., and the term IoT becomes Industrial IoT (IIoT). Internet of People (IoP) and Internet of Services (IoS) are other more specialized networks built in the same manner.

The main aim with establishing an IIoT in a manufacturing enterprise is to achieve reliable and real-time data and information transmission and analysis both horizontally and vertically. This would allow self-awareness, self-adjust and self-compare capabilities among machines and equipment, as well as improved decision making, proactive behavior, predictive analysis and on-fly adaptation to dynamic changes in demand and factory settings.

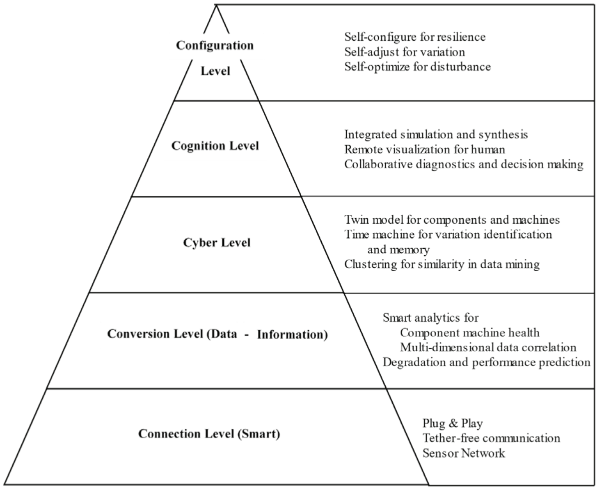


Figure 2. *5C architecture for CPS (Lee et al., 2015)*

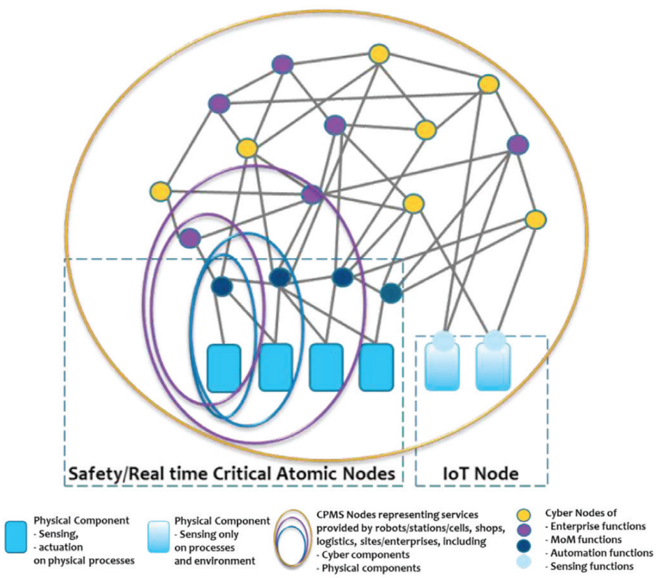


Figure 3. *Cyber-physical manufacturing services framework (Lu & Ju, 2017)*

Three core technologies that would enable IIoT implementation in a smart factory are (Yang et al., 2016): 1) Radio-Frequency Identification (RFID), 2) Wireless Sensor Networks (WSNs), 3) Cloud Computing and

Big Data. RFID technology uses RFID tags to transmit, receive and read information through radio waves. WSNs are networks of wireless autonomous sensors, where every sensor has a radio transceiver to transmit data and receive signals, a microcontroller for computing, an analog circuit for signal processing, an operating system, and a power source, all together creating smart sensors (Yang et al., 2019). Cloud computing and Big Data are explained in the following subsections.

Communication is the baseline to implement IIoT. Wired communication technologies such as the fieldbus and ethernet are only applicable between a shop floor and the management, and the communications between shopfloors on different regions and with data centers are typically established through the internet. However, using internet for communication has security, large delay, and line layout issues, while the large-area wireless communication technologies such as GPRS, GSM, 3G and 4G, as an alternative, have low transmission rates, support small number of node connections, have large time-delay, and are not regarded as reliable and secure (Cheng et al., 2018). An efficient communication technology to be used in a CPMS should (1) have high transmission rate and bandwidth, (2) supply high coverage, (3) has lower communication latency, (4) support large number of node connections, (5) ensure high reliability, and (6) guarantee high communication security (Cheng et al., 2018). 5G as the next generation large-area wireless communication technology is expected to meet the communication needs in an IoT-enabled smart manufacturing enterprise.

2.3. Big data

Massive amounts of data are generated and collected at minimal time frames across the shopfloor in a smart factory. Both the dimensions of that size data and handling of it are referred as the "Big Data" concept. Big Data has exceptionally large volumes, varieties, velocities and most importantly has increased value, which are collectively referred as 4Vs of Big Data. Recently, validation, variability (uncertainty), veracity (quality and credibility), verification, vision and volatility (data lifetime) are introduced additionally to define characteristics of Big Data as a whole.

"Smart manufacturing must embrace big data", as Kusiak (2017) phrased. The technologies that produce and consume data in a smart manufacturing enterprise are automation and process technologies, data storage technologies, digitization technology, cloud computing technology, agent technology and prediction technology (Kusiak, 2019). More specifically, data is collected from embedded sensors and actuators relating to the operating condition and real-time status of machines and other manufacturing resources, or data relates to environmental temperature, humidity, pressure, etc., also management and planning data is coming from traditional IT systems such as ERP, MES and CAx systems. Consumer generated data such as social media profiles, customer comments in web or support data are also in place in smart enterprises to meet customer needs and to increase satisfaction.

There are various aspects where analysis of Big Data would bring value in manufacturing: these include product design, production planning, production control, product quality and maintenance (Qi & Tao, 2018). Analyzing user behaviors, comments and market trend data enable correctly and rapidly identifying customer expectations for product attributes and quality in designing a product. Accurate real-time data on manufacturing resources and capacity allows a more realistic manufacturing program. Real-time manufacturing data also allows process monitoring and control and taking corrective on-time actions. Health monitoring and fault diagnosis are possible with preventive maintenance.

Seven requirements in handling Big Data for a predictive manufacturing are pointed out as (Babiceanu & Seker, 2016): 1) Embed sensors throughout the manufacturing domain and store data, 2) Collect and detect information about manufacturing events by filtering the stored data, 3) Correlate and analyze datasets through real-time data storage capabilities, 4) Generate event-based prediction models and adjust continuously, 5) Calculate analytics to monitor deviations from manufacturing objectives and create alerts, 6) Derive recommendations and automate decisions, and 7) Based on the recommendations and automatic decisions, proactively adapt the processes.

2.4. Data-driven analysis

Not only gathering the large amount of data is a challenge, but also converting it to useful information and to extract the real value in the data is a hard task that an Industry 4.0-based manufacturing enterprise should solve. A manufacturing firm could benefit from data analytics during product design, production planning and process optimization, material distribution and tracking, process monitoring, product quality control and manufacturing equipment maintenance (Tao et al., 2018). Instead of traditional mathematical modelling tools, data-driven analysis methods are rather used in smart manufacturing paradigm. Those methods include statistical inference, machine learning, deep learning, artificial intelligence, and simulation.

Four main types of data analysis are possible (Shao et al., 2014): 1) *Descriptive analytics* answers the question of what is happened or happening? It figures out patterns and trends present in data. 2) *Diagnostic analytics* deals with the reasons of events and low performances. Sensitivity analysis is a tool to be used. 3) *Predictive analytics* make future predictions about performance. What-if scenarios can be used in making predictions. 4) *Prescriptive analytics* figures out the required plan and programs to reach the desired performance. Different types of data analytics tools can be found useful in all the four types of data analysis.

Artificial intelligence is a broad concept to encompass both techniques of machine learning and deep learning. Machine learning refers to ability of software systems to self-adapt and self-update their models according to new input-output relations, hence, to learn. Currently, machine learning techniques are applied in decision support processes, plant and operations health management, data management and product life-cycle management, while other promising areas of application are proposed as manufacturing applications development, linking various stages of product life-cycle development, and cross-domain applications of machine learning (Sharp et al., 2018). Support vector machines, metaheuristic algorithms, neural networks, logistic regression, Markov processes and Linear temporal logic are examples of tools used in machine learning. Deep learning is a more advanced way of learning of machines. In contrast to machine learning, human intervention into modelling process is minimized and a black box approach is utilized to convert inputs to outputs. One important distinction is due to feature identification which is automatically performed by the model itself. Existence of nonlinear connection of multi-layers enables a high hierarchical model structure. Several typical deep learning architectures include Convolutional Neural Network, Restricted Boltzmann Machine, Auto Encoder, and Recurrent Neural Network (Wang et al., 2018).

Simulation is a traditional tool still capable to produce value adding analysis in factories of the future with the opportunities offered by IoT and Big Data. It can produce both diagnostic, predictive and prescriptive analytics (Shao et al., 2014). A verified and validated simulation model of the complete manufacturing enterprise would enable cause detection of performance indicators through sensitivity analysis. Additionally, what-if analysis of different scenarios would allow predictive analytics. Also, possible changes in input parameters and operational policies can be simulated to observe the potential future outcomes and to define the best alternative as a prescriptive analysis.

2.5. Cloud computing

In its most basic definition, cloud computing is to perform computing activities over the Internet. It is a distributed computing paradigm based on abstraction, virtualization, scalability and on demand service. The offered activities include storage, business applications and services, and cloud makes them reachable from anywhere at any time. *Ubiquitous* property is not the only advantage of cloud computing, reduced IT hardware and software costs are the other major gains for the enterprises.

Cloud computing covers Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS) solutions. In SaaS, software applications are installed in cloud -either in servers of the vendor or centrally in a datacenter- and are accessible over the internet, offering cost savings by renting the software and remote software updates. PaaS provides software development applications online enabling design, develop, test, hosting and maintenance aspects, while IaaS offers storage, hardware and network services both in physical and virtual modes. Cloud services can be accessible through web browsers, smartphones or tablets. Typical examples of SaaS, PaaS and IaaS are given in Table 1. Additionally, in a smart manufacturing scheme, Design as a Service, Manufacturing as a Service, Experimentation as a Service, Simulation as a Service, Management as a Service, Maintain as a Service, and Integration as a Service are other proposed cloud service models (Tao et al., 2014) , which are collectively referred as XaaS.

Table 1. *Examples of SaaS, PaaS, and IaaS*

Software as a Service: SaaS	Platform as a Service: PaaS	Infrastructure as a Service: IaaS
Salesforce AppExchange	Amazon Elastic Beanstalk	GoGrid
Microsoft Office 365	Windows Azure Compute	Amazon S3
Google Apps	Google App Engine	Microsoft Azure VM
OnLive	Salesforce	Google Compute Engine
Joyent	Facebook F8	Rackspace Technology
Gmail	Bunzee Connect	MSP on Demand
Facebook, Twitter	Amazon EC2	masterIT

Cloud computing is such an important component so that *cloud manufacturing* term is commonly used as a substitute for smart manufacturing. The characteristics of service-oriented structure, virtualization, multi-tenancy, shared pool of resources, scalability, elasticity, pay-per-use scheme, on-demand customization and user experience make cloud computing shaping smart manufacturing paradigm (Ren et al., 2017). Enterprises do not need to invest in costly software and hardware, instead they can outsource IT systems and infrastructure through service providers in the cloud and pay in a utility-based billing model. These computing resources are available as virtual machines generally, which provide increased computing power, higher utilization with flexible partitioning of physical computing resources and diversified pool of resources with various capabilities. In times of peak computing requirements, the resources can be scaled up to balance the workload, and down again if demand falls, thus allowing energy saving and sustainability. Customization of cloud services is another very useful property that allows users to configure the cloud services based on their needs and with no help from a technical assistant.

One main concern is the adoption of a company to implementing cloud computing solutions and leaving the traditional IT structure. Several adoption models are proposed in the literature, while technology-organization-environment (TOE) framework and diffusion of innovation (DOI) model are the most common ones. TOE explains the three dimensions of innovation adoption, while DOI figures out five attributes of adoption as relative advantage, compatibility, complexity, observability and trialability. Relative advantage, complexity, technological readiness, top management support and company size are statistically found to have influence on adoption of cloud computing (Oliveira et al., 2014).

2.6. Digital twin

Digital twin is the concept of having a virtual model of a physical object created in computing environment. Beyond from being a simple copy of the object, there is a two-way dynamic mapping and synchronization between the two that enables simulation models for judgement, predictive and prescriptive analysis, and optimization. A real-time, continuous interaction between the physical entity and the virtualization is a must and has complete reliance on data transmission. Thus, the three elements: physical world, virtual world, and the interconnection among them all together define the concept of digital twin.

Digital twin is composed of three levels: unit level, system level and system of system level as shown in Figure 4 (Qi et al., 2018). Unit level consists of virtual models of equipment in factory floor that take role in manufacturing activities. System level represents a single production line with machine tools, robot arms, conveyor belts, etc. A one-to-one mapping of the units and the system with respect to shape, status and working principles are carried out. System of system level represents the virtualization of the complete factory floor for cyber-physical management and optimization.

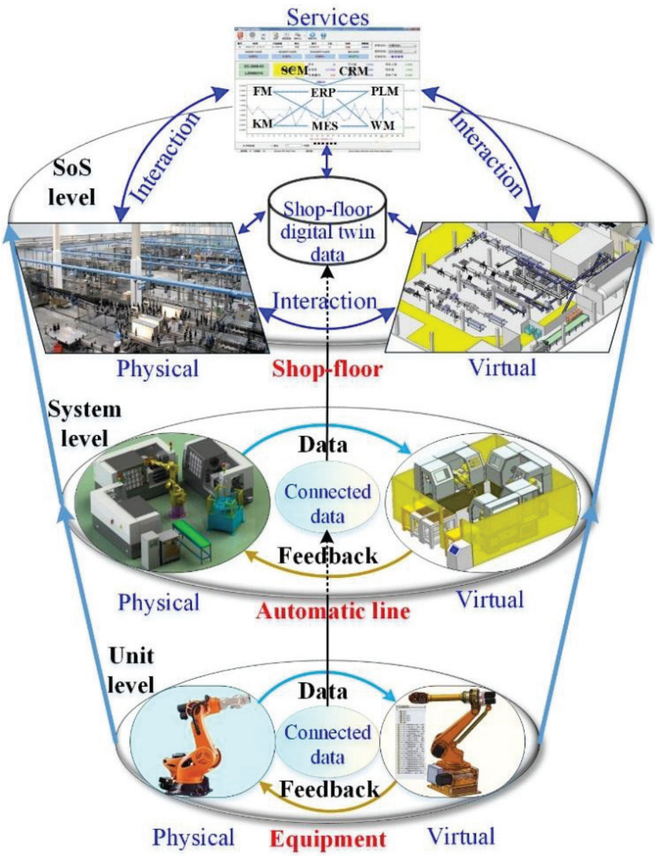


Figure 4. *Three levels of digital twin (Qi et al., 2018)*

Digital twin assumes substantial roles in product design, manufacturing, usage monitoring, maintenance, repair and overhaul (Qi & Tao, 2018). Before a product being manufactured, a virtual model is created based on digital twin which considers both the designer's expectations and the real constraints in the physical world. It is a process of continuous interaction, adjustment, and improvement. It also facilitates product verification in terms of functionality and manufacturability and detect any defects in advance and take action. Following the design phase, the manufacturing of a product is also optimized with the help of a virtualization of the factory floor together with all its elements, resources, processes, and rules. Various manufacturing scenarios are simulated from raw material to end product, and the best performing strategy and production plan are chosen for implementation. However, it is not a hierarchical top-down process and the plan continues to be updated once real manufacturing is started, to enable an executable and optimal manufacturing. After the manufacturing of the product is completed, a virtual copy of the product -product digital twin-is kept in computing environment which holds product-specific real-time data such as its status, use environment, and operating parameters. This would enable both having accurate and predictive information about the product such as its status, performance, fault, and remaining lifetime, and simulating for effects of changing environmental conditions. This health and lifetime awareness about the product leads to smarter maintenance, repair and overhaul activities that would be on-time and less costly. Also, pre-simulations on virtual model for different maintenance strategies enable more executable, effective, and optimal actions.

2.7. Augmented Reality / Virtual Reality

AR/VR refer to technologies and devices that display virtual 3D worlds. They are based on real-time rendering and require high-rate computing with abstracting interfaces and specific graphics software and hardware. Superposition of virtual and real worlds is established to display computer-generated information on exactly right places in the vision of the user. To enable tracking of the real world, infrared, ultrasonic, electromagnetic, GPS, and image processing methods are used.

AR/VR devices and technologies can be found in various forms. AR can be offered through mobile devices such as smartphones and tablets, or can be in form of AR projections consisting of a depth camera and a video projector -DigitalDesk as an example-, or, as another form, AR head-mounted displays (HMD) or VR-HMDs are commonly used wearables -Google Glass and HTC Vive are examples- (Büttner et al., 2017).

One common use of AR/VR research in manufacturing is in product design phase. CAD systems are used together with VR technologies either to allow a direct integration of a product model with the capturing of properties and with planning and fabrication or to create 3D model of a 2D input (Nee et al., 2012). AR, on the other hand, can be used for product prototyping purposes. Design options can be overlayed onto real views to make color, texture, etc. modifications to be reflected on the model and to observe the final result, or mock-up models can be created by assembling body parts (Nee et al., 2012).

AR/VR are used in manufacturing to support manual assembly work, operating machinery and supervising; in logistics to assist manual tasks such as picking, navigation and data management; in maintenance to enable remote assistance and provide instructions; and in worker training applications to facilitate learning through VR headsets, external displays, AR HMD instead of paper manuals (Büttner et al., 2017). Assembly instructions, corrective actions or other relevant information about a manufacturing process can be projected through AR tools onto a virtual desk or directly onto the field of workers as shown in Figure 5. AR devices are also used to assist physically impaired workers. Projecting instructions to guide workers are shown to increase productivity and reduce error rates. AR/VR can be used in logistics to guide workers by highlighting the material to pick from a warehouse by an HMD, or to assist in forklift operations or in factory layout planning through displaying 3D virtual animations on the factory floor for possible options.

Another application of AR/VR technology in manufacturing is proposed as the simulation of a production system (Dangelmaier et al., 2005). A planner with an AR-equipment can walk through the factory floor to identify the resources, their numbers, positions, and availability and to measure distance, times, etc. for a complete and correct reproduction of the system. Also, AR can be used similarly for verification of the simulation model. A close-to-real 3D model of the production system with all its resources (machines, robots, transportation plants, etc.) can be presented to a team of experts through AR HMD, where they can interactively change the simulation parameters to observe the results.

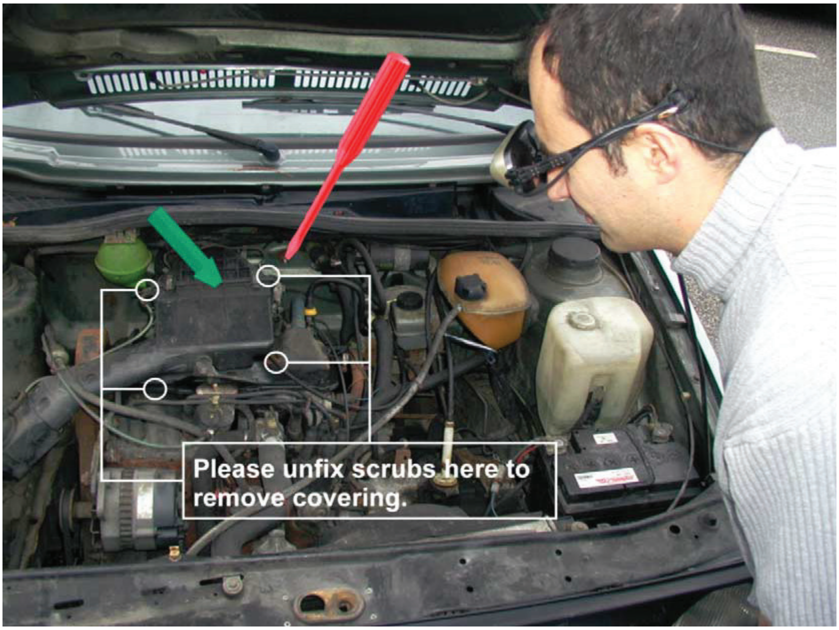


Figure 5. *Instructions are displayed in the worker's field of view with AR HMD (Dangelmaier et al., 2005)*

2.8. Direct digital manufacturing

DDM describes a mode of manufacturing where a 3D CAD model is used to directly fabricate the end product without the need for process planning (Gibson et al., 2015).

Additive manufacturing (AM) is the baseline technology for DDM.

AM uses a 3D model to derive thin cross-sections added layer by layer to create the final product -an example can be seen in Figure 6-. Products with different shapes and complex architectures are manufactured. Powder, liquid and solid materials can be used in AM as raw materials, where the common ones are aluminum, steel alloys, precious metals, plastics in powder form, and in filament form: wood, wax, paper, clay, concrete, sugar and chocolate (Chen et al., 2015). 3D printers are the main enabling technology in additive manufacturing. Main benefits with the use of AM are to have innovative and customized products, increased sustainability, reduced waste, increased productivity, and decreased costs.

The term prosumer is used instead of consumer in DDM paradigm to reflect the designer and producer role of the consumer. The production occurs locally and is based on design preferences of a group of consumers. In contrast to mass customization where there are customized products with several variants, the products have a highly personalized portfolio in DDM with a lot size of one.

Applications of DDM are common in automotive, aerospace and defense, chemical, medical, food, and construction industries (Paritala et al., 2017). Creating prototypes, producing jigs and fixtures, manufacturing concept models, and reproducing parts are the applications in automotive industry. Aerospace and defense industries are the ones where AM technologies are used thoroughly, in which lightweight but complex components are manufactured. In medicine, surgical and diagnostic aids, prosthetics, biomedical implants, regenerative medicine, drug delivery systems, bioprinting, and tissue engineering are the areas of AM use. Lastly, construction engineers can create architectural models by AM instead of hand-build models.

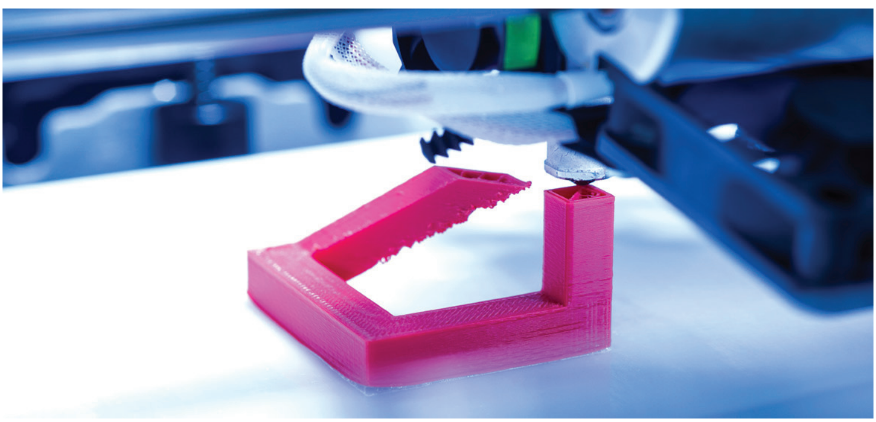


Figure 6. *Example of Additive Manufacturing (NASA, 2016)*

3. Discussion and future perspectives

Manufacturing industry is undergoing a fundamental evolution based on advancements in production technologies, computing capabilities, and information and communication technologies. The ongoing change is very favorable but still there is a long way to go in terms of conceptualization, development, implementation, integration, and adoption of emerging technologies.

Application strategies for Product Life Cycle management need to be developed together with design models, manufacturing steps and reference models and operation and maintenance systems and guidelines (Kang et al., 2016). Real-time perception of information across regions, real-time collaborative optimization of multiple systems, organizing tasks to respond to changing demand, and precise organization and autonomous execution of manufacturing activities are issues need to be addressed to fully reap benefits of smart manufacturing (Qu et al., 2019). CPS systems, IoT, big data and service-oriented computing have been more common in other fields such as transportation, while their adaptation to manufacturing is required to meet the industry's characteristics and requirements.

Another critical issue is a firm's adoption of smart manufacturing technologies. The determinants of adoption are listed as the following (Ghobakhloo, 2020): perception about positive consequences, availability of financial resources to invest in technologies, managerial commitment and support, capability to evaluate existing operations technology, develop upgrade and purchase criteria, upgradeability of existing technology or readiness for new equipment implementation, capability to purposefully leverage the technologies, openness to organizational change, employees' qualification in terms of IT and manufacturing digitalization skills, seamless integration capability, corporate social responsibility policy (relationships with stakeholders and adherence to laws and regulations), strategic road mapping for digitalization, and lastly, a mature and integrated cybersecurity vision.

Security related issues are among vital challenges in smart manufacturing implementation. Using Internet-accessible manufacturing devices widely, deploying wireless networks on the shop floor and implementing cloud solutions have the potential to increase vulnerability of manufacturing and IT systems to cyber-attacks. Furthermore, instead of being a characteristic of smart manufacturing -as commonly mistaken- cyber security should be regarded as a design principle that is to be considered at or before the design stage, pervade all aspects of the system and continuously be reviewed with the emergence of new threats (Tuptuk & Hailes, 2018).

References

Babiceanu, R. F., & Seker, R. (2016). Big Data and virtualization for manufacturing cyber-physical systems: A survey of the current status and future outlook. *Computers in Industry, 81*, 128-137.

Büttner, S., Mucha, H., Funk, M., Kosch, T., Aehnelt, M., Robert, S., & Röcker, C. (2017). *The Design Space of Augmented and Virtual Reality Applications for Assistive Environments in Manufacturing: A Visual Approach* Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments, Island of Rhodes, Greece.

Chen, D., Heyer, S., Ibbotson, S., Salonitis, K., Steingrímsson, J. G., & Thiede, S. (2015). Direct digital manufacturing: definition, evolution, and sustainability implications. *Journal of Cleaner Production, 107*, 615-625.

Cheng, J., Chen, W., Tao, F., & Lin, C.-L. (2018). Industrial IoT in 5G environment towards smart manufacturing. *Journal of Industrial Information Integration, 10*, 10-19.

Dangelmaier, W., Fischer, M., Gausemeier, J., Grafe, M., Matysczok, C., & Mueck, B. (2005). Virtual and augmented reality support for discrete manufacturing system simulation. *Computers in Industry, 56*(4), 371-383.

Ghobakhloo, M. (2020). Determinants of information and digital technology implementation for smart manufacturing. *International Journal of Production Research, 58*(8), 2384-2405.

Gibson, I., Rosen, D., & Stucker, B. (2015). *Additive Manufacturing Technologies: 3D Printing, Rapid Prototyping, and Direct Digital Manufacturing*. Springer New York.

Kang, H. S., Lee, J. Y., Choi, S., Kim, H., Park, J. H., Son, J. Y., Kim, B. H., & Noh, S. D. (2016). Smart manufacturing: Past research, present findings, and future directions. *International Journal of Precision Engineering and Manufacturing-Green Technology, 3*(1), 111-128.

Kusiak, A. (2017). Smart manufacturing must embrace big data. *Nature, 544*(7648), 23-25.

Kusiak, A. (2018). Smart manufacturing. *International Journal of Production Research, 56*(1-2), 508-517.

Kusiak, A. (2019). Fundamentals of smart manufacturing: A multi-thread perspective. *Annual Reviews in Control, 47*, 214-220.

Lee, J., Bagheri, B., & Kao, H.-A. (2015). A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. *Manufacturing Letters, 3*, 18-23.

Lu, Y., & Ju, F. (2017). Smart Manufacturing Systems based on Cyber-physical Manufacturing Services (CPMS). *IFAC-PapersOnLine, 50*(1), 15883-15889.

Lu, Y., Riddick, F., & Ivezic, N. (2016). *The paradigm shift in smart manufacturing system architecture* IFIP International Conference on Advances in Production Management Systems (APMS), Iguassu Falls, Brazil.

Mittal, S., Khan, M. A., Romero, D., & Wuest, T. (2017). Smart manufacturing: Characteristics, technologies and enabling factors. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 233*(5), 1342-1361.

Moghaddam, M., Cadavid, M. N., Kenley, C. R., & Deshmukh, A. V. (2018). Reference architectures for smart manufacturing: A critical review. *Journal of Manufacturing Systems, 49*, 215-225.

NASA. (2016, 11/04/2020). *NASA Collaborates to Progress AM Efforts Both Nationally and Internationally*. Retrieved 12/20/2020 from https://sma.nasa.gov/news/articles/newsitem/2016/10/13/nasa-collaborates-to-progress-am-efforts-both-nationally-and-internationally

Nee, A. Y. C., Ong, S. K., Chryssolouris, G., & Mourtzis, D. (2012). Augmented reality applications in design and manufacturing. *CIRP Annals, 61*(2), 657-679.

NIST. (2017). *Smart manufacturing operations planning and control program*, National Institute of Standards and Technology, Gaithersburg.

Oliveira, T., Thomas, M., & Espadanal, M. (2014). Assessing the determinants of cloud computing adoption: An analysis of the manufacturing and services sectors. *Information & Management, 51*(5), 497-510.

Paritala, P. K., Manchikatla, S., & Yarlagadda, P. K. D. V. (2017). Digital Manufacturing- Applications Past, Current, and Future Trends. *Procedia Engineering, 174*, 982-991.

Qi, Q., & Tao, F. (2018). Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison. *IEEE Access, 6*, 3585-3593.

Qi, Q., Tao, F., Zuo, Y., & Zhao, D. (2018). Digital Twin Service towards Smart Manufacturing. *Procedia CIRP, 72*, 237-242.

Qu, Y. J., Ming, X. G., Liu, Z. W., Zhang, X. Y., & Hou, Z. T. (2019). Smart manufacturing systems: state of the art and future trends. *The International Journal of Advanced Manufacturing Technology, 103*(9), 3751-3768.

Ren, L., Zhang, L., Wang, L., Tao, F., & Chai, X. (2017). Cloud manufacturing: key characteristics and applications. *International Journal of Computer Integrated Manufacturing, 30*(6), 501-515.

Shao, G., Shin, S., & Jain, S. (2014, 7-10 Dec. 2014). Data analytics using simulation for smart manufacturing. Proceedings of the Winter Simulation Conference 2014,

Sharp, M., Ak, R., & Hedberg, T. (2018). A survey of the advancing use and development of machine learning in smart manufacturing. *Journal of Manufacturing Systems, 48*, 170-179.

Tao, F., Cheng, Y., Xu, L. D., Zhang, L., & Li, B. H. (2014). CCIoT-CMfg: Cloud Computing and Internet of Things-Based Cloud Manufacturing Service System. *IEEE Transactions on Industrial Informatics, 10*(2), 1435-1442.

Tao, F., Qi, Q., Liu, A., & Kusiak, A. (2018). Data-driven smart manufacturing. *Journal of Manufacturing Systems, 48*, 157-169.

Tuptuk, N., & Hailes, S. (2018). Security of smart manufacturing systems. *Journal of Manufacturing Systems, 47*, 93-106.

Wang, J., Ma, Y., Zhang, L., Gao, R. X., & Wu, D. (2018). Deep learning for smart manufacturing: Methods and applications. *Journal of Manufacturing Systems, 48*, 144-156.

Yang, C., Shen, W., & Wang, X. (2016, 4-6 May 2016). Applications of Internet of Things in manufacturing. 2016 IEEE 20th International Conference on Computer Supported Cooperative Work in Design (CSCWD),

Yang, H., Kumara, S., Bukkapatnam, S. T. S., & Tsung, F. (2019). The internet of things for smart manufacturing: A review. *IISE Transactions, 51*(11), 1190-1216.

Chapter 32

COMPARISON OF WORD EMBEDDING METHODS FOR TURKISH SENTIMENT CLASSIFICATION

Murat AYDOĞAN[1]

1 Murat AYDOĞAN, Bingol University, Genc Vocational School, Department of Computer Technologies maydogan@bingol.edu.tr

1. INTRODUCTION

The amount of data produced has reached quite large amounts following the developments in informatics and related internet technologies. Especially with the introduction of social media environments, processing vast amounts of data has become difficult, but the processing is necessary. These environments enabled people to learn what others think about any person, place, institution, organization or activity within minutes. People can learn anything in a concise time. In this case, processing and analyzing data become even more important (Ünal, 2015).

Most of this data that needs to be processed consists of text data that has made the ever-important and popular text processing field more attractive in recent years. Considering the recent academic studies, it is seen that the number of text processing studies has increased significantly, but the number of studies specific to the Turkish language is insufficient. Therefore, this study was conducted for the Turkish language using a Turkish dataset (Kılınç et al., 2015).

Sentiment analysis is a text processing task that has become popular in recent years, especially with the introduction of social media platforms and many users (Liu and Zhang, 2012). Companies allocate a large share of their budget to social media research, celebrities, and even politicians research social media and direct their strategies according to social media interactions (Çoban and Özyer, 2018).

One of the most significant factors in the increase and success of text processing studies is the development of word embedding methods. After the successful development of word embeddings methods, there has been a great leap forward in this area, and the performance levels of the tasks have reached satisfactory levels (Amasyalı et al., 2012). Many text-processing researchers have agreed that word embedding methods are the most critical choice to carry out a successful task. Therefore, it is a fundamental issue in terms of production and use (Ayata et. al, 2017).

In this study, a Turkish sentiment analysis study, in which the performances of word embedding methods were compared, analyzed, and evaluated using machine learning algorithms on a Turkish dataset created using the comments that the users share on various shopping sites, was conducted.

2. TEXT PROCESSING

Humans can process only a small amount of unconstructed data, while computers can process a significant number of structured data. Therefore, it is necessary to process unstructured data so that computers can understand it. Text pre-processing task falls within the natural language processing field, and it is the initial stage of text processing (Adalı, 2012).

Text processing is an engineering field that focuses on designing and implementing computer systems whose main purpose is to analyze, understand, interpret, and produce a natural language. The popular research subjects of text processing are listed as follows:

- Text classification
- Sentiment Analysis
- Text summarizing
- Translation between natural languages
- Question and answering asistants, chatbots
- Information Extraction

Text Pre-processing

Text processing starts with the formation of a document collection, which is also called "corpus". Corpus can be described as the set of documents consisting of texts. Text processing solutions aim to identify unknown patterns within the large-scaled corpora. There may be thousands, millions, or even billions of documents in such corpora, nowadays, with the development of big data. Data in the corpus include unconstructed raw data when compared to the classical database structure. These data can be written both in a specific computer language and a natural language. Within a corpus, documents consist of paragraphs, sentences create paragraphs, and words create sentences. Also, the corpus structure may be static or dynamic. If the initial state of the corpus is preserved, it is considered static. However, if new documents are added to a corpus or the existing documents are updated, it is considered dynamic (Krouska, 2016).

Successful text processing studies are based on the advanced data pre-processing methodologies. Text processing relies remarkably on the several pre-processing methods which make the constructed data views out of unconstructed raw data sources either semantically, or via quotation, or with both that it is claimed that these detailed preparation methods determine text processing to a certain extent (Angiani et al., 2016).

Many different text processing methods exist in the field. Almost all methods aim to construct the corpus, or rather the document collection which constitutes the corpus. Various text pre-processing methods are commonly used together to form constructed document views from the raw text data. One way to categorize all the pre-processing data construction methods is to classify them according to their use of algorithms and formal frames for their individual duties.

Tokenization

The first pre-processing step necessary for a corpus study is to divide the character stream into words or “tokens” as called in field literature (Weiss et al., 2010). This step is crucial for text processing. It is challenging to obtain higher-level information from a corpus without acquiring tokens. Moreover, this step should be completed before other more advanced pre-processing stages. The tokenization process can be operated at various levels. A document can be segmented into chapters, sections, paragraphs, sentences, words, and even syllables and characters within a corpus. The most preferred method in text processing studies is the segmentation of a text into sentences and the words.

While the tokenization method is an easy task for a human being, it is challenging for a computer system. This difference depends on the use of specific characters; tokens may be restrictive or non-restrictive in some cases. For example, space, tabs and home characters are always considered as restrictive because they are not recognized as symbols. In general, they are called “white spaces” as a whole. “( )” (parenthesis), “< >” (angle brackets), “!” (exclamation point) and “?” (question mark) characters are always considered as restrictive, as well as tokens. Moreover, “.” (period), “,” (comma), “:” (colon) and “-” (hyphen) characters may be considered as restrictive or non-restrictive depending on the context.

The tokenizer used in the tokenization method should be selected according to the text in question to obtain the best features possible. Otherwise, the obtained tokens may constitute an overload. It is advised to keep in mind that the tokenization process depends on the language. General principles may be applicable to other different languages, however there will be differences in a detailed study.

Stemming

After the character stream's segmentation into the token string, the next step is to convert each symbol, called a token, to standard form. This process is generally referred to as stemming or lemmatization. The necessity of this operation depends on the application in question. For example, if a document will be classified, stemming may provide an advantage. For the frequency-driven classification algorithms, the operation can be convenient as well. However, for other cases, an extra stemming process may not give a significant advantage.

Lexicography

A dictionary is the listing of words and some symbols together. The extent of a dictionary is a criterion that can affect the performance of the developed application. Therefore, it should be free of unnecessary words.

One of the fundamental methods to downsize the dictionary is to remove the *stopwords*, the unnecessary words list, from the corpus. These words generally do not affect the forecast performance of the application. For stopwords, tokens such as "a", "an", "the" in English and words such as "ve", "veya" and pronouns such as "ben" and "sen" in Turkish can be given as examples.

Additionally, downsizing the dictionary by removing words may positively affect the model's performance and increase the forecast performance in some cases as well. As the most frequent words in a corpus are the stopwords, removing these words may enhance the model's performance. Moreover, as some rarest words may have been misspelled, they can also be removed from the corpus.

3. WORD EMBEDDING METHODS

Word embedding methods are described under two main topics: the traditional text embedding methods based on the word weighting formulas and the text embedding methods based on the neural network architecture. Frequency-based word embeddings described as the traditional methods are built on the principle of finding words in texts and recording their frequencies. Word embedding simply is the conversion of words to numerical forms so that the computers can understand (Wang et. al, 2018).

Traditional Word Embedding Method Approaches

TF (Term Frequency) – IDF (Inverse Document Frequency)

The term Frequency method is used for the identification of the term frequency in a document. Inverse Document Frequency method determines the frequency of a word in more than one document and decides its word-class such as a term, a conjunction, or a stopword.

This method is generally used to calculate the weight or the score of words. First, in this step, the frequency of unique words should be determined by counting. In Equation 3, it can be seen that the TF-IDF weight of the *j* word is the adaptive term frequency with the scale factor based on the word's significance. The scale factor is called as the inverse document frequency in Equation 2. Frequency reverses the scaling by calculating the document count (*df(j)*), including the *j* word. Therefore, if the word is included in more than one documents, it shall be considered insignificant, and the scale shall decrease almost to the value of zero. On the other hand, if the word is relatively rare and included in a few documents, the scale factor shall increase and the word shall be considered significant (Peter, 1957).

In Equation 1, term frequency (TF) value is calculated with $tf_{i,j}$. As shown in the equation, this value is determined with *TF = word frequency in the document/word count of the document*.

$$tf_{i,j} = \frac{n_{i,j}}{\sum_k n_{i,j}} \tag{1}$$

IDF value is calculated by using Equation 2. *IDF = log (total document count / document count including the word)*

$$idf\ (w) = \log\left(\frac{N}{df_t}\right) \tag{2}$$

Equation 3 is used to constitute the weights of the TF-IDF matrix.

$$w = tf * idf \tag{3}$$

Bag of Words

According to this method, each sentence in a document is transformed into a matrix in the unique words' size by segmenting the sentences to unique words. Columns of the matrix consist of words in the document (*N*), while the row number is equal to the document number (*D*). As a result, the whole corpus is represented as a matrix with the size of *DxN*.

Nevertheless, Co-occurrence Matrix is based on the principle of similar words' tendency to be in the same sentence. Both the columns and the rows of the matrix formed with this method consist of unique words. Also, the co-occurrence frequencies of the words in columns and rows are located in the matrix cells (Scott and Matwin, 1998).

Modern Word Embedding Method Approaches

Traditional methods have two crucial disadvantages. First, depending on the method, as unique words are located in columns and/or rows, matrix size is equal to word count, and as the value of most of them shall be 0 (zero), a sparse matrix occurs. Second, the semantic affinity of the words cannot be identified. Therefore, neural network-based modern word embedding methods were developed in the field.

The same text can be converted to different numerical values in other ways. Word embedding methods are the techniques preferred for this transformation. When the field literature is reviewed, especially with the latest developments in neural networks, word embedding is considered as one of the most critical points in text processing studies (Erdinç and Güran, 2019).

Word2Vec

Developed by Mikolov in 2013, Word2Vec is a model based on the principle of training words with the neural network. In this model, the target word is estimated according to the initial word inputs. Word2Vec consists of two different models: CBOW (Continuous Bag of Words) and Skip-gram.

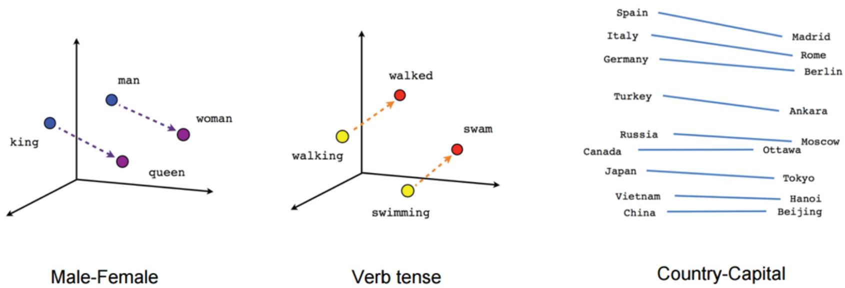


Fig. 1. *Relation captured by word embedding (Mikolov et al., 2013)*

Deciding on how the texts within a corpus will be represented is one of the most critical topics of text processing studies and the most important network input. In the simplest terms, converting texts to numerical data is referred to as word embedding (Mikolov et al., 2013). The same text can be converted to different numerical values in other ways as well. The other word embedding method, which is also from a forecasting-based word embedding method group, the Word2Vec model, was developed by Mikolov (Mikolov et al., 2013) and is based on the principle of training words with the neural network. Word2Vec consists of two different models: CBOW (Continuous Bag of Words) and Skip-gram. In this study, the CBOW model was used.

In the CBOW model, words that are not in the center and within the window size are considered inputs, and the words in the center are to be forecasted as outputs. This operation continues until the end of a sentence. When Figure 2 is examined, w(t) value is the output value in the center and to be forecasted, values shown as w(t-2)…w(t+2) are the output values which are not in the center according to the preferred window size (Le et al., 2014). Skip-gram algorithm, the opposite of this process is done.

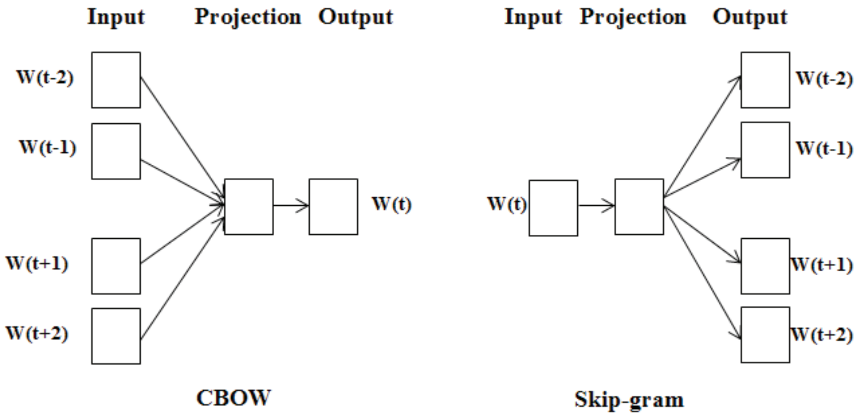


Fig. 2. *Word2Vec method*

Word2Vec algorithm is a neural network-based approach. Equation 4 shows the error function to be optimized during the training.

$$J = -\frac{1}{T}\sum_{t=1}^{T}\sum_{-m\le j\le m}\log P(w_{t+j} \mid w_t) \tag{4}$$

When the operation t=1….T in Equation 4 is applied to all the words in a corpus, vectors forecasting the words neighboring each word with a great chance of success are calculated. While *m* signifies the window size of the target word, w_t shows the word in center and w_{t+j} shows the word in *j* distance.

The most significant advantage of the Word2Vec model is the protection of word affinity resulting from its implementation according to the conditional probability principle and words' positions in a sentence. As seen in Figure 2, after the words were vectorized with Word2Vec, words closest to one word were determined with the cosine similarity formula.

Glove

Glove is another word embedding method that abbreviates the initials of "Global Vectors for Word Representation". It is the second most used method after Word2Vec in the natural language processing field. It is developed by Pennington et al. at Stanford University (Pennington et al., 2014)

This method is less likely to be successful than the Word2Vec method while determining the near-synonyms in a corpus. The glove is also a method that is trained on the global word-word count and therefore enables the statistics to be more effective. Equation 5 shows the error function to be optimized in accordance with the Glove method.

$$J = \frac{1}{2}\sum_{i,j=1}^{W} f\left(P_{ij}\right)\left(u_i^T v_j - \log P_{ij}\right)^2 \tag{5}$$

In Equation 5, *u* and *v* symbols represent the column and row values of the matrix formed with the global word-word counts. While $f\left(P_{ij}\right)$ is the nominal weight function, *W* is the size of the created dictionary.

FastText

FastText is a library within the Gensim structure, developed for the text classification by Facebook AI Research in 2016. It converts texts and words to continuous vectors. Differing from the Word2Vec model, FastText segments words to "n-gram" values consisting of fewer letters than considering each word as an input for the artificial neural network (Joulin et al., 2016).

n in the n-gram statement refers to the repetition grade; in other words, it determines the character number into which a word will be divided regarding the *n*. Also, the word vector of the word is the sum of all these n-gram vectors. After the training is completed, word vectors are calculated for all the n-grams given in the training set. As the occurrence probability of the n-grams of words with low frequency is small, these words can be represented more accurately with the n-gram method (Bojanowski et al., 2016).

4. SENTIMENT ANALYSIS

Sentiment analysis is a popular text processing subject used in academia and industry, analyzing people's analysis, attitudes, and emotions towards products, events, topics, individuals, and organizations. The fundamental task of the sentiment analysis is to detect the components of sentiment expression from texts (Seker, 2016; Kaynar et al., 2016). Generally, in the sentiment analysis studies, sentiment can be classified using two labels as positive and negative, or three labels as positive, negative, and neutral (Medhat et al. 2014).

In the classification stage, commonly used machine learning algorithms such as K- Nearest Neighbors, Naive Bayes, Logistic Regression and Decision Trees and Support Vector Machine were preferred (Balaban, 2015).

K-Nearest Neighbors (KNN)

The algorithm developed by Fix and Hodges is one of the most commonly used methods in the field (Altman,1992). In simplest terms, in the k-nearest neighbor's algorithm, the class of a new example and the Euclidean distance to the examples in the number of k in a sample space is calculated according to the k value designated as a parameter for this example, and the class which has the most examples in k-nearest neighbors is assigned.

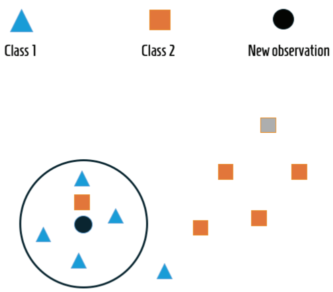


Fig. 3. *K-nn algorithm*

When Figure 3 is examined, it can be seen that the k value was determined as three, when the nearest three neighbors are analyzed, as there are two red class examples and one example of a blue class in these three neighbors, k shall be labelled as a red class in the given an example according to the nearest neighbor algorithm.

Naive Bayes (NB)

The Naive Bayes method, one of the machine learning algorithms and commonly preferred in the text classification projects, is used for data analysis. It is a supervised classification algorithm titled naive. Forecasting operation needs classifying a new case by combining the effects of an independent variable on dependent variables. The naive Bayes method builds the calculations on probability and statistics fundamentals (James, 2003).

$$X = [x_1, x_2, x_3, \dots\ x_n]$$

$$C = [c_1, c_2, c_3, \dots\ c_n]$$

$$P(C_i/x) = \frac{p(x\ /\ C_i\ )\ P(C_i)}{P(x)} \tag{6}$$

Equation 6 is the formula in which the probability values of the states are calculated.

$$\arg max_{ci}\{\ P\ (x|c_i)\ P\ (c_i)\ \} \tag{7}$$

The new situation encountered is assigned to the class of the highest probability value calculated using Equation 7.

Logistic Regression (LR)

The logistic regression analysis's primary purpose is to constitute an acceptable model that can define the optimal relationship between independent and dependent variables using the least number of variables. Logistic regression determines the effect of multiple independent variables presented simultaneously to predict the class of one or the other of the two categories of dependent variables (Tolles and William, 2016).

Decision Trees (DT)

As the name suggests, a Decision Tree means creating a tree-like structure for decision-making. In decision trees, several special terms, which are taken from a real tree structure such as branch and leaf, are used. Moreover, decision nodes and leaf nodes are included in a decision tree. Decision nodes are the properties used for decision-making, classification, and forecasting in a data set, and they can be separated into two or more branches, while leaf nodes store the decisions. Also, the node at the highest point of the tree is called the root node. A specific path is followed from the tree root to the nodes, to make a decision (Safavian and Landgrebe, 1991)

Support Vector Machine (SVM)

Support Vector Machine is a statistical information theory-based classifier. SVM algorithm has two sub-algorithms: linear SVM and non-linear SVM. For the linear SVM algorithm, infinite hyperplanes are created to separate data, and a hyperplane with maximum margin is chosen among all these hyperplanes. On the other hand, the non-linear SVM algorithm is used when the classes cannot be linearly classified, and data are transferred to a larger area. Therefore, data become linearly separable (Cortes et al., 1995). The main advantage of SVM is high accuracy. Today the SVM algorithms are applied in many classification problems ranging from face recognition systems to voice analysis.

5. EXPERIMENTS

During the experiments, firstly preprocessing processes were done and then the model developed with training and testing was first trained and tested. Finally, the performance of the model is evaluated. The structure of the developed model is as shown in the Figure 4.

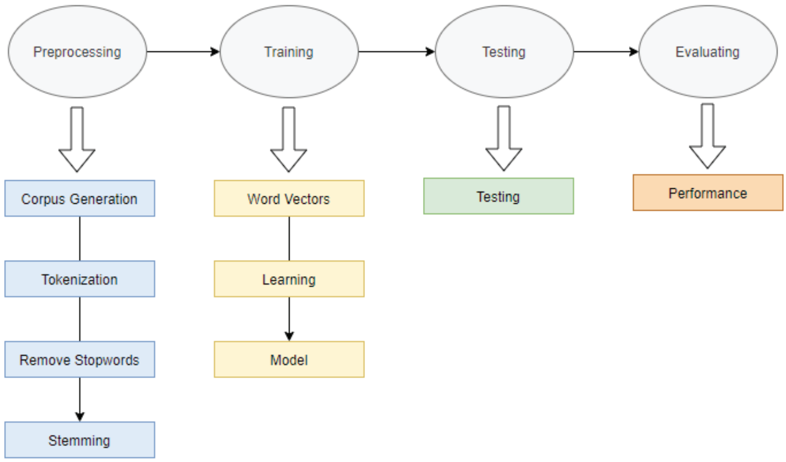


Fig. 4. *Architecture of proposed system*

The training and tests conducted on this study were performed with an Intel i7 3.4 GHz processor and 16 GB of memory.

Dataset

This data set is created for the classification operation to be conducted together with the training word vectors and machine learning algorithms. After developing a crawler with Python, comments of various online shopping websites' users on the products they are interested in are collected. Overall, the data set consists of 90 thousand comments, 30 thousand negative, 30 thousand positive and 30 thousand neutral sentences (Fig 5).

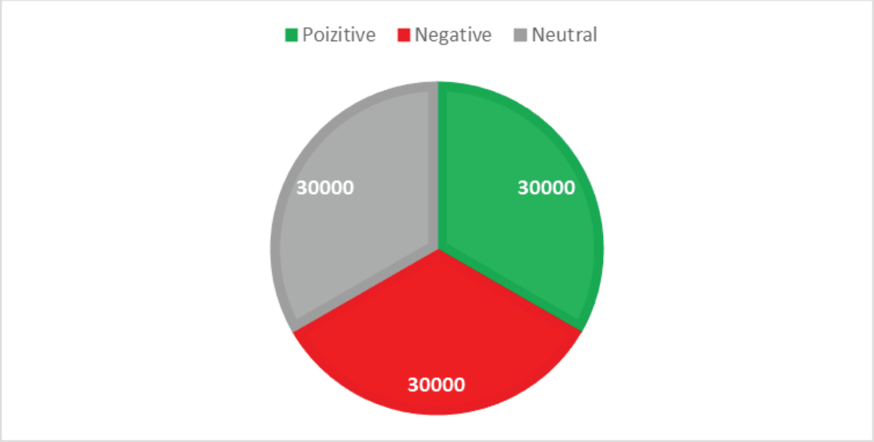


Fig. 5. *Distribution of labeled class*

Labeling is performed according to the users' product points: Point 0 signifies negative, Point 3 signifies neutral, and Point 5 signifies positive. It can be seen in Table 1.

Table 1. *Structure of dataset*

id	Date	Comments	Score	Sentiment
1	01.08.2020 20:54	Hiç beğenmedim tamamen gereksiz	0	Negative
2	01.08.2020 10:08	Ürün idare eder, daha iyi olabilirdi	3	Neutral
3	01.08.2020 16:11	Mükemmel, bayıldım tavsiye ediyorum	5	Positive

The data in the dataset was pre-processed before being transformed into a word vector. The pre- processed steps applied are as follows.

- By reading the texts in the corpus, the texts are divided into sentences by referring to the point mark.
- All sentences have been converted into lowercase letters.
- All characters (numbers, symbols, punctuation marks) from the sentences except letters have been cleared.
- Sentences are divided into words with reference to the spaces between them.

The purpose of this study is to compare the performance of word embedding methods with machine learning algorithms on an original dataset. During the experiments, the vector dimension size was chosen as 100, the window size as 5, and the minimum word count parameter as 5 in word embedding methods (Duwairi and El-Orfali, 2014).

Equations 8, 9, 10 and 11 were used to evaluate the results.

$$Accuracy = \frac{(TN + TP)}{(TN + TP + FP + FN)} \tag{8}$$

$$Precision = \frac{TP}{TP + FP} \tag{9}$$

$$Recall = \frac{TN}{TN + FP} \tag{10}$$

$$F\ Score = 2\ \frac{Precision * Recall}{Precision + Recall} \tag{11}$$

During the experimental studies, the words in the data with positive and negative labels and converted to tokens were analyzed. When the data belonging to the negative class were examined, it was seen that the most common words in these data were "kötü" (English meaning bad), "iade" (English meaning return), "beğenmedim" (English meaning I didn't like). When the same was done for words with positive labels, it was seen that the words with the highest frequency were "iyi" (English meaning good), "güzel" (English meaning beautiful) and "süper" (English meaning super). It can be seen in Fig 6.

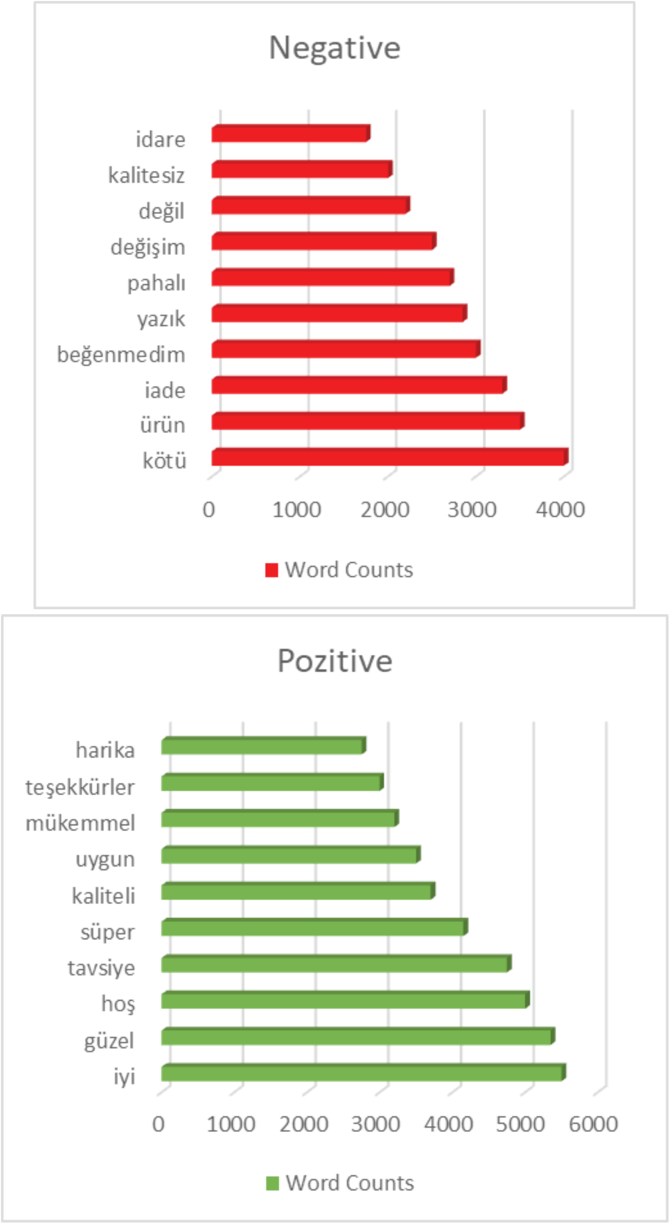


Fig. 6. *Contribution of words for positive and negative sentiment*

During the experiments, the dataset was split according to the ratio random_state and the size of the training data was determined to be 67500 (75% of the dataset). After these data were trained, the test data was determined as 22500 (25% of the data set) to be tested.

Table 2. *Accuracy of modern word embedding methods*

	CBOW	Skip-Gram	Glove	FastText
K-NN	0.79	0.77	0.75	0.76
SVM	0.85	0.85	0.81	0.83
NB	0.82	0.84	0.80	0.80
DT	0.80	0.80	0.78	0.79
LR	0.86	0.83	0.80	0.82

When Table 2 is examined, it is seen that the Word2Vec method algorithms are slightly more successful than the CBOW and SG, Glove and FastText methods. FastText method also performed close to CBOW and SG, however, it was more successful than Glove. In terms of machine learning algorithms, the most successful algorithms were LR with an accuracy rate of 86% and SVM with a performance close to it.

Table 3. *Accuracy of traditional word embedding methods*

	TF-IDF	BoW
K-NN	0.75	0.74
SVM	0.78	0.73
NB	0.78	0.76
DT	0.75	0.71
LR	0.79	0.73

When the performance of the traditional word embedding methods in Table 3 is examined, the TF-IDF method has shown a more successful performance on all algorithms than the BoW (Bag of Words) method. When machine learning algorithms are examined, the most successful algorithms are LR, NB and SVM algorithms that give results close to each other.

CONCLUSION

Together with the increase in the text data produced today, text processing studies have become more popular. However, the number of Turkish text processing studies is still relatively low compared to the other studies conducted in other languages. Therefore, in this study, Turkish data set was created, and modern and traditional word embedding methods approaches were applied.

For the dataset created by collecting news customer reviews about products from several shopping websites. Additionally, K-Nearest Neighbors (KNN), Naive Bayes (NB), Support Vector Machine (SVM), Decision Trees (DT), Logistic Regression (LR) two machine learning algorithms were used for the classification.

Representation type of words is a crucial subject in text processing studies. In this study, both more traditional and more modern approaches were exercised, and the results were analyzed.

Until the neural network-based approaches are in high demand again, word embeddings are examined with the most used method BoW and today's most favorite word embedding method Word2Vec. Moreover, when the results were analyzed, it was realized that Word2Vec performed better results. In other words, modern word embedding methods performed better than traditional word embedding methods in all experiments.

LR and SVM algorithms have been more successful than other algorithms in both traditional word embedding methods and modern word embedding methods experiments.

REFERENCES

Adalı, E. (2012). Doğal Dil İşleme. *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, 5(2).

Altman, N. S. (1992). "An introduction to kernel and nearest-neighbor nonparametric regression". *The American Statistician. **46** (3): 175–185.*

Amasyalı M. F., Balcı S., Mete E., Varlı E. N., (2012)"Türkçe metinlerin sınıflandırılmasında metin temsil yöntemlerinin performans karşılaştırılması" *EMO Bilimsel Dergi*, vol. 2, no. 4.

Angiani, G., Ferrari, L., Fontanini, T., Fornacciari, P., Iotti, E., Magliani, F., & Manicardi, S. (2016, September). A Comparison between Preprocessing Techniques for Sentiment Analysis in Twitter. In *KDWeb*.

Ayata, D., Saraçlar, M., & Özgür, A. (2017, May). Turkish tweet sentiment analysis with word embedding and machine learning. In *2017 25th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.

Balaban, M. E., Kartal, E., (2015). Veri madenciliği ve makine öğrenmesi uygulamaları, Çağlayan *Kitabevi*.

Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2016). Enriching word vectors with subword information. CoRR abs/1607.04606 (2016). *arXiv preprint arXiv:1607.04606*.

Cortes, C., *Vapnik, Vladimir N. (1995*). "Support-vector networks". *Machine Learning. **20** (3): 273–297.*

Çoban, Ö., & Özyer, G. T. (2018, September). Word2vec and Clustering based Twitter Sentiment Analysis. In *2018 International Conference on Artificial Intelligence and Data Processing (IDAP)* (pp. 1-5). IEEE.

D. Kılınç, A. Özçift, F. Bozyigit, P. Yıldırım, F. Yücalar, E. Borandag, (2015) TTC-3600: A new benchmark dataset for Turkish text categorization. *Journal of Information Science*, 43(2), 174-185.

Duwairi, R., & El-Orfali, M. (2014). A study of the effects of preprocessing strategies on sentiment analysis for Arabic text. *Journal of Information Science*, *40*(4), 501-513.

Erdinç, H. Y., & Güran, A. (2019, April). Semi-supervised Turkish Text Categorization with Word2Vec, Doc2Vec and FastText Algorithms. In *2019 27th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.

Joulin, A., Grave, E., Bojanowski, P., & Mikolov, T. (2016). Bag of tricks for efficient text classification. *arXiv preprint arXiv:1607.01759*

Joyce, J. (2003), "Bayes' Theorem", in Zalta, Edward N. (ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2019 ed.), Metaphysics Research Lab, Stanford University.

Kaynar, O., Görmez, Y., Yıldız, M., & Albayrak, A. (2016). Makine öğrenmesi yöntemleri ile Duygu Analizi. *In International Artificial Intelligence and Data Processing Symposium (IDAP'16)* (pp. 17-18).

Krouska, A., Troussas, C., & Virvou, M. (2016, July). The effect of preprocessing techniques on Twitter sentiment analysis. In *2016 7th International Conference on Information, Intelligence, Systems & Applications (IISA)* (pp. 1-5). IEEE.

Liu, B., & Zhang, L. (2012). A survey of opinion mining and sentiment analysis. In *Mining text data* (pp. 415-463). Springer, Boston, MA.

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams engineering journal*, *5*(4), 1093-1113.

Pennington, J., Socher, R., & Manning, C. D. (2014, October). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 1532-1543).

Peter H. L., (1957). "A Statistical Approach to Mechanized Encoding and Searching of Literary Information" . *IBM Journal of Research and Development.* ***1*** *(4): 309–317.*

Q. Le, T. Mikolov, (2014). Distributed representations of sentences and documents. *31th International Conference on Machine Learning*, China.

Safavian, S. R., & Landgrebe, D. (1991). A survey of decision tree classifier methodology. *IEEE transactions on systems, man, and cybernetics*, 21(3), 660-674.

Scott, S., & Matwin, S. (1998). Text classification using WordNet hypernyms. *In Usage of WordNet in Natural Language Processing Systems.*

Seker, S. E. (2016). Duygu Analizi (Sentimental Analysis). *YBS Ansiklopedi*, *3*(3), 21-36.

T. Mikolov, I. Sutskever, G. Corrado, J. Dean, (2013). Distributed Representations of Words and Phrases and their Compositionality, AAAI Spring Symposium AI Technologies for Homeland Security 200591-98, vol. cs.CL, pp. 3111–3119.

T. Mikolov, K. Chen, G. Corrado, J. Dean, (2013). Efficient estimation of word representations in vector space. *Proceedings of Workshop at ICLR*. Scottsdale.

Tolles, J., William J. (2016). "Logistic Regression Relating Patient Characteristics to Outcomes". *JAMA*. **316** (5): 533–4

Ünal, F. (2015). Büyük Veri ve Semantik. *Abaküs Yayınevi*

Wang, Y., Liu, S., Afzal, N., Rastegar-Mojarad, M., Wang, L., Shen, F., ... & Liu, H. (2018). A comparison of word embeddings for the biomedical natural language processing. *Journal of biomedical informatics*, *87*, 12-20.

Weiss, S. M., Indurkhya, N., Zhang, T., & Damerau, F. (2010). *Text mining: predictive methods for analyzing unstructured information*. Springer Science & Business Media.

Chapter 33

EXAMINING THE USE OF FUZZY LOGIC FOR IOT SECURITY

İrfan AYGÜN [1]

1 Research Asst. İrfan AYGÜN, Manisa Celal Bayar University, Software Engineering Department, Turkey, irfan.aygun@cbu.edu.tr

INTRODUCTION

Due to the advancement of technology, the number of internet connected devices used in daily life is increasing rapidly. The majority of these devices were recently used by people directly using computers, tablets or phones. However, as of today, it is observed that almost all electronic devices are connected to the internet with the concept of “smart”. The difference compared to the recent past; these devices can communicate autonomously without the need for instant human control. With these developments in inter-object communication, the concept of IOT (Internet of Things) has taken its place in the literature. The exact number of devices that will be subject to internet of things or inter-object communication is not known. However, researches and studies predict that 26 billion devices will operate within this scope by 2020 (Khan & Salah, 2018). Thanks to the breakthrough of technology in this field, many autonomous systems and smart systems have been implemented. Again, under the leadership of these concepts, many popular working areas such as smart urbanism, smart factories and production facilities, autonomous independent vehicles have been formed. It is vital to ensure the continuity and security of communication between devices in order to perform healthy studies in all these topics. At this point, information security studies and studies on IoT networks are supported.

The importance of cyber security on information security and technological infrastructure is also increasing day by day. Thanks to the use of the Internet, all devices that communicate with each other definitely share information. There may be disruptions due to the scope of the communication or mishaps that will cause massive effects as a result of the changed content. As the number of members in the cyber world increases, the importance of cyber security increases exponentially. Studies on this subject show that cyber attacks increase at a minimum rate of 600% or more each year compared to the previous year (Saravanan & Bama, 2019). These statistics are expected to increase with each passing year at higher rates. All these attacks are not done directly with instant human control. Consequently, it is not possible to resort to instant human control to detect and respond to these attacks. For this reason, many studies are carried out with the topics of automatic attack detection, autonomous reaction in case of anomaly detection, and autonomous control of system operation. Rule-based approaches and artificial intelligence-supported studies have become very popular in order to detect anomalies or damage that may occur in the system. Due to the fact that artificial intelligence offers positive results in many disciplines and offers resource savings, it has been inevitable that it is frequently preferred in studies for autonomous information security. In this way, a wide variety of artificial intelligence approaches have been used to minimize possible losses. These approaches and algorithms can vary according to the type of problem and the desired reaction type. For problems using artificial intelligence approaches; It is possible to present many examples such as the detection of the attack, the detection of the attack type, the instant reaction to the attack, the detection of the attacker type, and the detection of the damage after the attack (Zhao et al., 2020).

In this study, "Fuzzy Logic" model, one of the artificial intelligence approaches to information security, is emphasized. In the following sections; Similar studies conducted in the past, the general use and domain-based use of fuzzy logic, experimental studies and the conclusion of the findings are discussed.

RELATED WORK

When the literature is examined, it is seen that information security is provided by many different strategies and methods. In their study, Ahmad et al. Summarized these strategies under 5 headings: intrusion detection, attack prevention, attack surveillance, attack intervention and deception of the attacker (Ahmad et al., n.d.). Each of these strategies has different important roles in ensuring information security. For this reason, different artificial intelligence solutions are developed for all of these approaches. This section focuses on studies using fuzzy logic as an artificial intelligence method.

In case devices or systems using Internet access are attacked, various solutions based on fuzzy logic are suggested to automatically detect this situation (Acampora et al., 2010; Dovom et al., 2019; Haripriya & Kulothungan, 2019; Sai Satyanarayana Reddy et al., 2019). Zhao et al. Conveyed important studies on this subject in their comprehensive research. (Zhao et al., 2020). In addition to these studies, hybrid approaches and studies using fuzzy logic are very popular. For example; In the study conducted by Novaes et al., DDoS and PortScan attacks are detected by measuring the similarities of the detections made with fuzzy logic (with LSTM) (Novaes et al., 2020). Similarly, in the study conducted by Özyer, the results obtained with fuzzy logic were made more focused and strengthened with the support of genetic algorithm (Tarıyan Özyer et al., n.d.).

After detecting that the system is under attack, it is necessary to determine the type of attack in order to ensure security. Because, if the type of attack is known, the precaution to be taken can be chosen accordingly, the vulnerabilities and weaknesses that the system is exposed to can be determined more easily, and it becomes easier to make inferences for the attacker profiles over time (Uma & Padmavathi, 2013). Fuzzy-based inferences have been used in many studies to detect DDoS attack, which is one of the most popular cyber attack types (Alsirhani et al., 2019; Janarthanan et al., 2020; Sherazi et al., 2019; Uma & Padmavathi, 2013). In these studies, the decision taken between the system's natural density and being attacked is examined. When this decision is evaluated in a mathematical set or binary logic, it can give misleading results. When evaluated with binary logic, the periodic exposure of some systems to abnormal traffic levels will lead to perceive this traffic as an attack continuously. At this point, using a control and percentage and fuzzy logic contributes to the solution of the problem. In their study, Velliangiri and Pandey show that fuzzy logic gives significantly higher results than other algorithms in detecting DDoS attacks (Velliangiri & Pandey, 2020). Apart from DDoS, it is seen that similarly fuzzy logic is used to detect other attack types such as "Sql Injection", "Man in the Middle", "Trojan Sending" (Batista et al., 2019; Dehghani et al., 2020; Farzaneh et al., 2020; Pradeep Mohan Kumar et al., 2019; Scaranti et al., 2020). Most of these studies show that there cannot be a single cluster in predicting the attack type. Computational intelligence techniques are used to determine belonging to appropriate clusters among more than one possibility.

In addition to the studies for the security of individual systems, similar studies are carried out to ensure security in multi-device communications. With the development of IoT and machine-to-machine communication (m2m) concepts, the size of the network to be examined has increased. Various methods have been tried to scan these networks properly and to detect possible anomalies. Computational intelligence solutions specific to fuzzy logic are also frequently used in the security of IoT networks like other artificial intelligence approaches (Almogren et al., 2020; Alshehri & Hussain, 2019; Markovic-Petrovic et al., 2019; Nguyen et al., 2019; Ouechtati et al., 2020). Fuzzy logic-based studies have been conducted to interpret the type of anomaly with the proposal "All anomalies that will occur in large networks may not be due to cyber attack". In the study conducted by Örnek et al., the types of anomalies encountered by autonomous carriers in the smart factory were examined with fuzzy logic (Örnek et al., n.d.). In the study, it is aimed to predict the types of attacks that a large-scale IoT network may encounter by using the topics in the literature.

MATERIAL AND METHOD

The current study is inspired by a dataset that examines the movements of an IoT network and uses a wide range of features (83 Features) (Ullah & Mahmoud, 2020). This data set, which is shared in a reusable manner in the study by Ullah and Mahmood, contains the necessary qualities for detecting network anomalies. As a result of the study, it was aimed to select the relevant features from this data set and evaluate them within the scope of fuzzy logic and reach the result of attack or no attack as the result type.

A description of the method used in line with this goal is presented under subheadings.

The Use of Fuzzy Logic in Cybersecurity and IoT Security

Fuzzy sets theory was developed by Zadeh for representation of uncertainty and representation of mixed distribution between sets. Thanks to this theory, real life problems can go beyond the binary system representation (Zadeh, 1996). Thanks to this approach, which is expressed as fuzzy logic, numerical approaches have become possible to problems that cannot be directly expressed mathematically. There are basically 4 processes in the operation of this method. These processes are as follows.

- Fuzzification
- Rules
- Inference
- Defuzzification

In order for the fuzzy system to go through the above processes, it must have certain input and output data. The data presented as input to the system is expected to turn into output data with fuzzy rules and membership functions. Outputs, which are called final output data, express the result that the system produces in the decision mechanism.

There are many areas suitable for the use of fuzzy logic as an approach. As can be seen from the literature review, cyber security is one of these areas. Studies that make use of fuzzy logic differently for any phase of cyber security was compiled in chapter 'Related Work'. There are many reasons for preferring fuzzy logic instead of logical evaluations in these studies. The main reasons are listed below as examples.

- The idea that a system is attacked may give misleading results according to the classical set theorem. Because there are web systems that are exposed to much more than the usual traffic during certain periods of the year. Contrary to the normal course, these days with heavy traffic are the days when the system needs to work most actively. Just in this process, thinking that it is under cyber attack based on abnormal traffic and band failure will ensure that the system is turned off at the most active use. For this reason, fuzzy logic is preferred due to the approach that other possibilities are taken into consideration together with the possibility of attack.

- Fuzzy logic is used in the problems that are set out with the absolute assumption that a system is under attack. Because the type of attack or the type of the attacker can sometimes be too complex to be separated into 0 and 1. Making this determination constitutes the most important step in determining the reaction to be given to the attack. Therefore, all possibilities should be reviewed at this step. Consequently, using fuzzy logic provides a significant advantage in order not to ignore even small percentage memberships.

- After the attack type, aggressor type and reactions are determined, the current situation is predicted. How much more can the current situation tolerate this attack? How much has the system been damaged in the current situation? In which of the measures to be chosen can the system maintain its stability? It is impossible to rely indifferently on a single result for the answers to such questions. Because these questions, which also include abstract factors, may be answered as "the system is not damaged to a great extent, but it has x% risk of shutdown". In real scenarios, the percentile indicated by x in these expressions can be quite important. For this reason, fuzzy logic is preferred in such problems.

Application Simulation Sample

Within the scope of the study, a sample system simulation was designed. The goal of this system is to detect whether a server serving the distributed network of objects is momentarily exposed to anomaly. In case of anomaly detection, it is planned to determine whether DDoS, which is one of the most frequently used attack types in the literature, or a different anomaly. The input and output values determined for the development of this application using fuzzy logic are as follows.

Inputs:

- bandwidth availability (percent)
- response delay (percent)
- new request rate (not seen in cache)

Outputs:

- Anomaly Condition
  - o There is an anomaly, justification DDoS
  - o There is an anomaly, the reason is different
  - o No anomalies

Fuzzy sets and linguistic expressions in accordance with the input and output values given above were prepared as shown in Figure 1, Figure 2 and Figure 3.

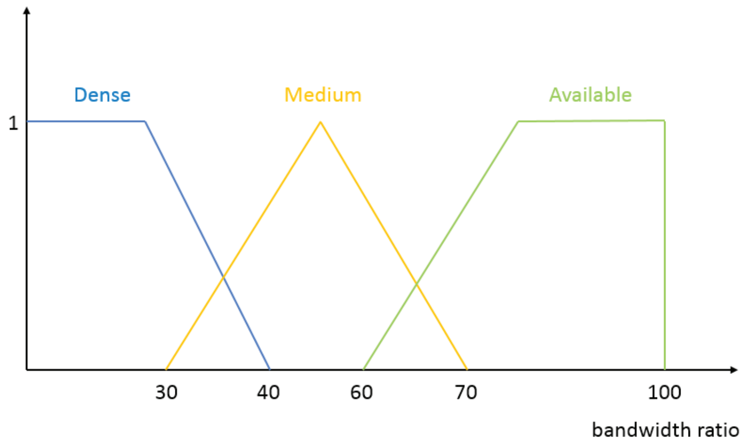


Fig. 1. Input fuzzy set (bandwith ratio)

In the Figure 1, the values are calculated based on the percentage ratio of availability in bandwidth, which is the initial input value. If there is 40% or less available bandwidth, the server is intensive, if there is a gap between 30% and 70%, it is moderately available, if there is a band gap of 70% or more, the server can serve quite comfortably.

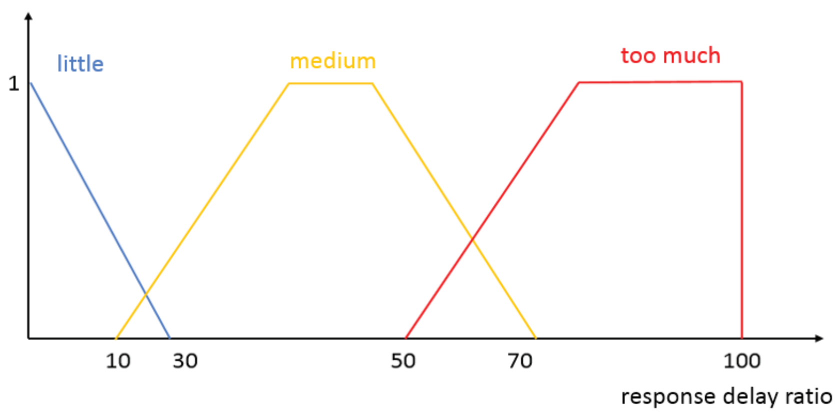


Fig. 2. Input fuzzy set (response delay ratio)

The delay rates seen in returning the requests to the system are shown in Figure 2. If there is 30% or less lag, the 'little' tag is used. If there is a delay between 10% and 70%, the 'medium delay' tag is used. If there is a delay between 50% and 100%, the 'too much' tag is used. As can be seen from this image, if the system delays responding to more than 50% requests, the problem is serious.

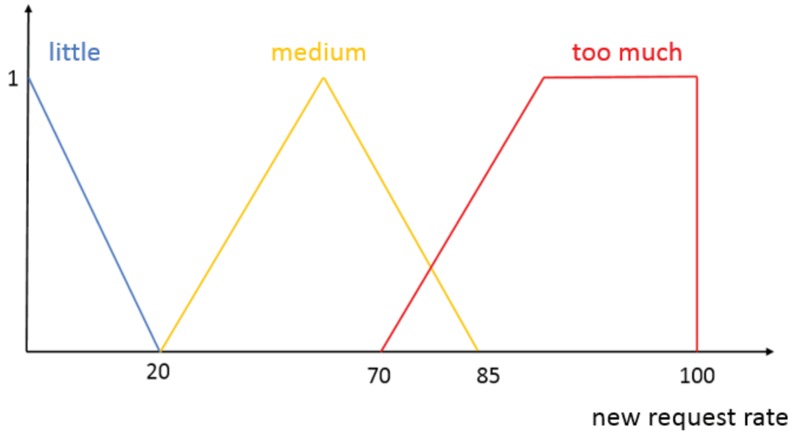


Fig. 3. Input fuzzy set (new request ratio)

Figure 3 shows the rate of requests received by the new users. In a possible DDoS attack, computers that have not used the system before will most likely be used. Based on this assumption, value labels have been assigned as shown in the figure. If there is a delay of 20% or less, the 'small' tag is used. If there are between 20% and 85% new users, the 'medium' tag is used. If there are between 70% and 100% new users, the 'too much' tag is used. As can be seen in Figure 3, less than 20% of the system sees new user requests as harmless. As this ratio rises to 70% and above, the system is expected to go into an alarm state.

The graphical representation shown in the example is also created for all other clusters. Likewise, linguistic expressions were prepared for the output fuzzy set and membership functions were created. The representation of the output fuzzy cluster is shown in Figure 4.

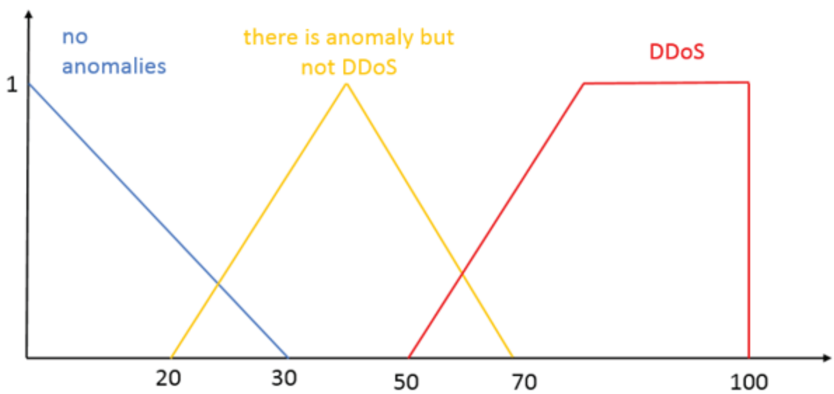


Fig. 4. Representation of Output Fuzzy Sets (anomaly ratio)

Figure 4 shows the fuzzy set that interprets the output values of the system. In this graph, the 'anomaly incidence rate' is shown as a percentage. The system does not have any problems for anomalies experienced at 30% or less. With the assumption that it is resistant to anomaly at this level, 'no anomaly' label is used. The label 'there is an anomaly but not DDoS' is used for the anomaly rate between 20% and 70%. These rates indicate a low level of DDoS attack or an anomaly for a different reason. However, as the rates increase from 50%, it becomes clear that this anomaly is caused by a DDoS attack. The role of fuzzy logic increases in the interpretation of values between 50% and 70% in these outputs. Likewise, the decision whether there is an anomaly or not with a value between 20% and 30% will be made with fuzzy logic.

Using binary logic to interpret this system requires ignoring some possibilities. For example, the exit label caused by an anomaly with a rate of 61% is very close to both values. When making these decisions in real life problems, all serious possibilities are taken into consideration. Considering that the system will react according to the decision, the least possibility must ignore. Fuzzy logic is used in such problems because of its suitability for these reasons. Alternatively, artificial neural networks, deep learning and other similar machine learning methods are frequently used in decision systems.

After the information-based adjustments were completed, rule-based operations were started. A total of 9 linguistic values were used in the inputs of the problem. For the output, it was formed as follows with 27 rule examples using the expressions 'DD' for 3 different verbal value DDoS, 'ND' for the other anomaly type and 'N' for the normal case without anomaly.

Rule 1: If the bandwidth is intense and the response delay is high and the rate of new requests is low, then there is a DDoS type anomaly.

Rule 2: If the bandwidth is medium and the response delay is high and the rate of new requests is low, then there is a different type of anomaly.

Rule 3: If the bandwidth is available and the response delay is high and the rate of new requests is low, then there is a different type of anomaly.

Rule 27: If the bandwidth is available and the response delay is low and the rate of new requests is low, then there is no anomaly.

Rule 4 and later were determined until 27, and a table of rules was created according to the attack characteristics seen in the literature.

For the inference part of the problem, the method of inference proposed by Mamdani and known by its name was used (Mamdani, 1974). According to this approach, the minimum of the expressions connected with ‘and’ and the maximum of the expressions connected with ‘or’ are selected, and the results are expressed. This approach is shown in Figure 5.

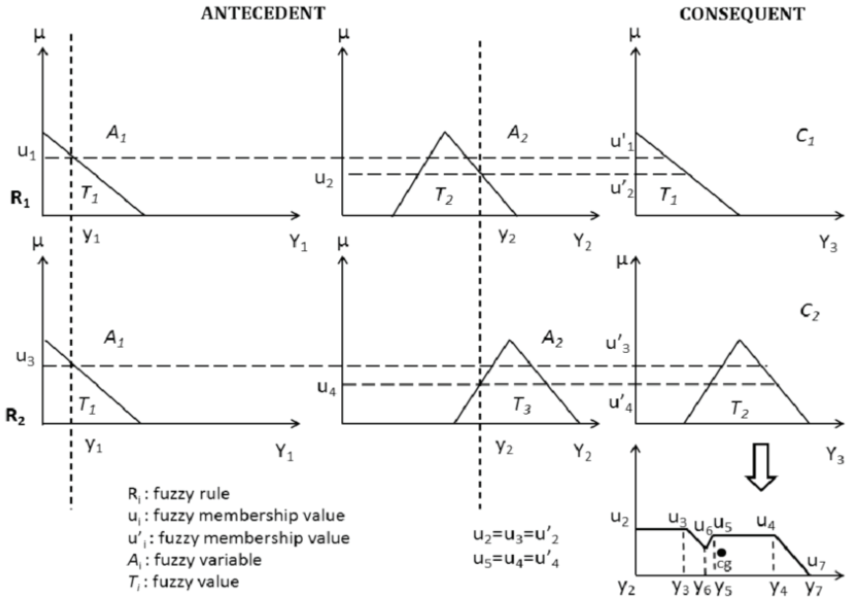


Fig. 5. Mamdani Fuzzy Inference System (Das & Winter, 2016)

A solution is sought for the following problem text in order to examine the working of the developed simulation on a concrete example.

- The server bandwidth is available at 32% and 40 percent of the requests to the server remain unanswered. 95% of the incoming requests consist of new users who are not kept on the server cache yet. In this case, information is needed about whether the system is faced with an anomaly and if it is under an attack, whether it is DDoS type.

After determining the intersection points of the problem given in the example on fuzzy clusters and cluster memberships, inferences were made with the Mamdani method. Before the defuzzification phase, it is seen that there is a 78% probability of being under DDoS attack and a 22% probability of a different type of anomaly. With these values, it is possible to decide that the system is under a DDoS attack after Mamdani extraction. However, it is possible to obtain different results by changing the parameters given in the example and with different inference methods.

Apart from Mamdani, there are many popular inference methods such as Takagi-Sugeno, Tsukamoto and Larsen which are frequently used in the literature. When examining a problem within the scope of fuzzy logic, values obtained from different inference methods are generally interpreted and compared. In this way, it becomes possible to increase the efficiency of the results obtained. The difference that distinguishes these methods from each other is the mathematical method they use in the background or their method of output generation. For example, Mamdani produces a fuzzy output value, but the Sugeno method produces a function as the output value. For this reason, it is very important to check the suitability of the problem to the structure of the problem when choosing the extraction method.

CONCLUTION

The study focuses on the use of fuzzy logic, one of the computational intelligence methods, in cyber security. Within the scope of cyber security, anomalies in IoT networks where communication between objects is provided by large networks, rather than the security of individual devices, have been examined. The research shows that fuzzy logic is very useful in detecting these anomalies, revealing the reasons, responding and predicting the instantaneous situation. As a result of the study, a sample simulation was carried out on the anomaly data sets and inferences used in previous studies. A problem frequently seen in real life was evaluated with sample parameters and made inferences about the current state of the system. Thanks to these inferences, plans to respond or get a reaction according to the type of anomaly will be more successful. There are hybrid approaches where fuzzy logic is strengthened to produce stronger predictions. Considering that IoT networks are growing day by day and problems are diversifying, the need for hybrid approaches becomes clear. Predictions obtained by working with fuzzy values in machine learning or deep learning methods can be made more successful. The use of these methods in future studies is on the agenda to increase the success of existing studies.

REFERENCES

Acampora, G., Gaeta, M., Loia, V., & Vasilakos, A. V. (2010). Interoperable and adaptive fuzzy services for ambient intelligence applications. *ACM Transactions on Autonomous and Adaptive Systems*, *5*(2), 1–26. https://doi.org/10.1145/1740600.1740604

Ahmad, A., Maynard, S. B., & Park, S. (n.d.). *Information Security Strategies: Towards an Organizational Multi-Strategy Perspective*.

Almogren, A., Mohiuddin, I., Din, I. U., Majed, H. Al, & Guizani, N. (2020). FTM-IoMT: Fuzzy-based Trust Management for Preventing Sybil Attacks in Internet of Medical Things. *IEEE Internet of Things Journal*, 1–1. https://doi.org/10.1109/jiot.2020.3027440

Alshehri, M. D., & Hussain, F. K. (2019). A fuzzy security protocol for trust management in the internet of things (Fuzzy-IoT). *Computing*, *101*(7), 791–818. https://doi.org/10.1007/s00607-018-0685-7

Alsirhani, A., Sampalli, S., & Bodorik, P. (2019). DDoS Detection System: Using a Set of Classification Algorithms Controlled by Fuzzy Logic System in Apache Spark. *IEEE Transactions on Network and Service Management*. https://doi.org/10.1109/TNSM.2019.2929425

Batista, L. O., De Silva, G. A., Araujo, V. S., Araujo, V. J. S., Rezende, T. S., Guimaraes, A. J., & De Souza, P. V. C. (2019). Fuzzy neural networks to create an expert system for detecting attacks by SQL Injection. In *arXiv* (Vol. 1, pp. 8–21). arXiv. https://doi.org/10.5769/j201801001

Das, R., & Winter, S. (2016). Detecting Urban Transport Modes Using a Hybrid Knowledge Driven Framework from GPS Trajectory. *ISPRS International Journal of Geo-Information*, *5*(11), 207. https://doi.org/10.3390/ijgi5110207

Dehghani, M., Kavousi-Fard, A., Dabbaghjamanesh, M., & Avatefipour, O. (2020). Deep learning based method for false data injection attack detection in AC smart islands. *IET Generation, Transmission & Distribution*, *14*(24), 5756–5765. https://doi.org/10.1049/iet-gtd.2020.0391

Dovom, E. M., Azmoodeh, A., Dehghantanha, A., Newton, D. E., Parizi, R. M., & Karimipour, H. (2019). Fuzzy pattern tree for edge malware detection and categorization in IoT. *Journal of Systems Architecture*, *97*, 1–7. https://doi.org/10.1016/j.sysarc.2019.01.017

Farzaneh, B., Koosha, M., Boochanpour, E., & Alizadeh, E. (2020). A New Method for Intrusion Detection on RPL Routing Protocol Using Fuzzy Logic. *2020 6th International Conference on Web Research, ICWR 2020*, 245–250. https://doi.org/10.1109/ICWR49608.2020.9122278

Haripriya, A. P., & Kulothungan, K. (2019). Secure-MQTT: an efficient fuzzy logic-based approach to detect DoS attack in MQTT protocol for internet of things. *Eurasip Journal on Wireless Communications and Networking*, *2019*(1). https://doi.org/10.1186/s13638-019-1402-8

Janarthanan, A., Kumar, D., Antony, R. R., & Parvathe, C. B. D. (2020). IADF security: insider attack detection using fuzzy logic in wireless multimedia sensor networks. *Soft Computing*, *24*(18), 13893–13902. https://doi.org/10.1007/s00500-020-04764-4

Khan, M. A., & Salah, K. (2018). IoT security: Review, blockchain solutions, and open challenges. *Future Generation Computer Systems*, *82*, 395–411. https://doi.org/10.1016/j.future.2017.11.022

Mamdani, E. H. (1974). APPLICATION OF FUZZY ALGORITHMS FOR CONTROL OF SIMPLE DYNAMIC PLANT. *Proceedings of the Institution of Electrical Engineers*, *121*(12), 1585–1588. https://doi.org/10.1049/piee.1974.0328

Markovic-Petrovic, J. D., Stojanovic, M. D., & Rakas, S. V. B. (2019). A Fuzzy AHP Approach for Security Risk Assessment in SCADA Networks. *Advances in Electrical and Computer Engineering*, *19*(3), 69–75. https://go.gale.com/ps/i.do?p=AONE&sw=w&issn=15827445&v=2.1&it=r&id=GALE%7CA601908834&sid=googleScholar&linkaccess=fulltext

Nguyen, V. T., Nguyen, T. X., Hoang, T. M., & Vu, N. L. (2019). A new anomaly traffic detection based on fuzzy logic approach in wireless sensor networks. *ACM International Conference Proceeding Series*, 205–209. https://doi.org/10.1145/3368926.3369714

Novaes, M. P., Carvalho, L. F., Lloret, J., & Proenca, M. L. (2020). Long short-term memory and fuzzy logic for anomaly detection and mitigation in software-defined network environment. *IEEE Access*, *8*, 83765–83781. https://doi.org/10.1109/ACCESS.2020.2992044

Örnek, Ö., Gülbandilar, E., Yazici, A., Osmangazi Üniversitesi, E., Mimarlık Fakültesi, M., Meşelik Yerleşkesi, B., & Mühendisliği Bölümü, B. (n.d.). AKILLI FABRİKALARDAKİ OTONOM TAŞIYICILAR İÇİN BULANIK MANTIK TABANLI ANOMALİ TESPİTİ. In *J ESOGU Engin Arch Fac* (Vol. 2020, Issue 1). Retrieved December 8, 2020, from http://orcid.org/0000-0001-5589-2032

Ouechtati, H., Azzouna, N. Ben, & Said, L. Ben. (2020). A Fuzzy Logic Based Trust-ABAC Model for the Internet of Things. *Advances in Intelligent Systems and Computing*, *926*, 1157–1168. https://doi.org/10.1007/978-3-030-15032-7_97

Pradeep Mohan Kumar, K., Saravanan, M., Thenmozhi, M., & Vijayakumar, K. (2019). Intrusion detection system based on GA-fuzzy classifier for detecting malicious attacks. *Concurrency and Computation: Practice and Experience*, e5242. https://doi.org/10.1002/cpe.5242

Sai Satyanarayana Reddy, S., Chatterjee, P., & Mamatha, C. (2019). Intrusion detection in wireless network using fuzzy logic implemented with genetic algorithm. In *Lecture Notes in Networks and Systems* (Vol. 75, pp. 425–432). Springer. https://doi.org/10.1007/978-981-13-7150-9_45

Saravanan, A., & Bama, S. S. (2019). A Review on Cyber Security and the Fifth Generation Cyberattacks. *Oriental Journal of Computer Science and Technology*, *12*(2), 50–56. https://doi.org/10.13005/ojcst12.02.04

Scaranti, G. F., Carvalho, L. F., Barbon, S., & Proenca, M. L. (2020). Artificial Immune Systems and Fuzzy Logic to Detect Flooding Attacks in Software-Defined Networks. *IEEE Access*, *8*, 100172–100184. https://doi.org/10.1109/ACCESS.2020.2997939

Sherazi, H. H. R., Iqbal, R., Ahmad, F., Khan, Z. A., & Chaudary, M. H. (2019). DDoS attack detection: A key enabler for sustainable communication in internet of vehicles. *Sustainable Computing: Informatics and Systems*, *23*, 13–20. https://doi.org/10.1016/j.suscom.2019.05.002

Tarıyan Özyer, S., Özyer, T., Üniversitesi, Ç., Mühendisliği Bölümü, B., & Etü, T. (n.d.). *ÇOK AMAÇLI GENETİK BULANIK SINIFLAMA İLE SALDIRI TESPİT SİSTEMİ*.

Ullah, I., & Mahmoud, Q. H. (2020). A Scheme for Generating a Dataset for Anomalous Activity Detection in IoT Networks. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *12109 LNAI*, 508–520. https://doi.org/10.1007/978-3-030-47358-7_52

Uma, M., & Padmavathi, G. (2013). A Survey on Various Cyber Attacks and Their Classification. *International Journal of Network Security*, *15*(5), 390–396. https://doi.org/10.6633/IJNS.201309.15(5).09

Velliangiri, S., & Pandey, H. M. (2020). Fuzzy-Taylor-elephant herd optimization inspired Deep Belief Network for DDoS attack detection and comparison with state-of-the-arts algorithms. *Future Generation Computer Systems*, *110*, 80–90. https://doi.org/10.1016/j.future.2020.03.049

Zadeh, L. A. (1996). *FUZZY SETS* (pp. 394–432). https://doi.org/10.1142/9789814261302_0021

Zhao, S., Li, S., Qi, L., & Xu, L. Da. (2020). Computational Intelligence Enabled Cybersecurity for the Internet of Things. *IEEE Transactions on Emerging Topics in Computational Intelligence*, *4*(5), 666–674. https://doi.org/10.1109/TETCI.2019.2941757

Chapter 34

COMPARISON OF POPULAR MACHINE LEARNING METHODS IN SOFTWARE DEFECT PREDICTION

Buket KAYA [1]
İrfan AYGÜN[2]

1 Asst. Prof. Dr. Buket KAYA, Fırat University, Dept. of Electronics and Automation, Turkey, bkaya@firat.edu.tr
2 Research Asst. İrfan AYGÜN, Manisa Celal Bayar University, Software Engineering Department, Turkey, irfan.aygun@cbu.edu.tr

INTRODUCTION

In today's technology, the complexity of the software increases considerably due to the complexity of the problems it solves. The errors of the software developed sometimes carry vital risks or cause very high financial losses. Therefore, error testing and post- delivery updates are very important in software. It is inevitable to encounter software errors in the software life cycle. Early detection of these errors is very important in terms of cost. According to research, if the correction of an error is left after the delivery of the software, it is costing 100 times more than the correction of the same error in the software development process (Boehm & Basili, 2001). Therefore, many studies are performed to detect software errors. The aim of the studies is to determine the error in the earliest time and to prevent human source error. The aim of automatic detection of error detection is also aim of machine learning methods too.

It is clear that a software will need maintenance and updates many times in its life cycle. This requires rechecking user feedback each time and manually correcting errors. This situation causes some personnel to use it repeatedly for a software that is completed in its current form. Moreover, each time the process of mastering the project will be experienced again. All these processes cause serious losses for software developers and companies. Thanks to automatic detection of software errors, the amount of error in the software that reaches the user will be minimized. Still, if there are unnoticeable errors, scanning the feedback automatically before a software developer will help you focus on the right point. Even this contribution will be very important and profitable for software projects.

As a result of the work done, it has become very easy to manage large software teams. As a result of these studies, the contact between project employees becomes easier and other problems that may arise in the workflow are resolved early. In addition, the prediction that software will be developed by autonomous devices has become a very strong prediction today. However, in order for devices to develop the right software, they must also be skilled in detecting potential problems. Today's error-detecting algorithms will directly affect the success in building software in the future (Angell et al., 2018).

In the past studies, important steps have been determined for software error detection. When the problem is considered as a whole, it is possible to encounter a complex structure. Therefore, it would be more realistic to use different techniques at different stages of the problem. Because a system is expected to understand the software in order to detect errors in the software. This situation is a major research subject. It would be a mistake to wait for a software step to be performed in one way. While the same process can be done in many different ways, the system's perception of them is a separate study. While accepting one of the possible methods, producing an 'error' output to the other will have undesirable results. The next stage of systems that detect errors will be to offer suggestions for solving these errors. The last phase following this phase is the development of software that automatically generates the solution of the error. There are different solution methods and approaches for each of these stages. In the thesis study conducted by Rana, possible methods that can be used for error detection are shown as in Figure 1.

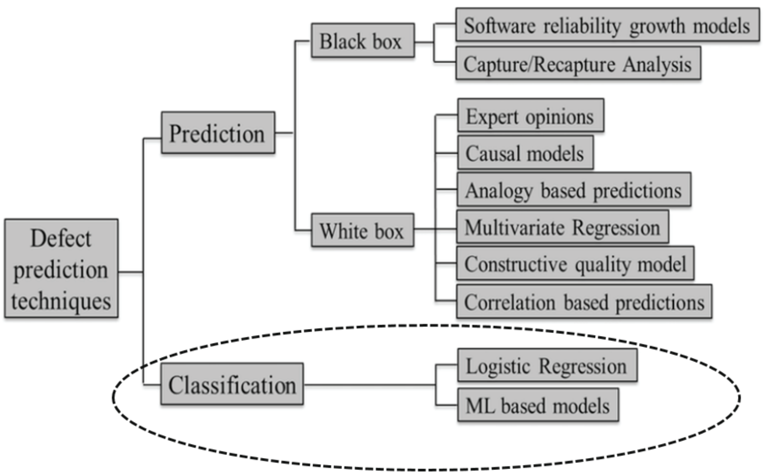


Figure 1. Methods used for software defect detection (Rana, 2015)

As can be seen in the figure, there are many different approaches for software defect prediction. The aim of this study is to find the most commonly used machine learning methods in software error estimation and compare these methods among themselves. According to the findings obtained after the comparison, it was aimed to determine the ideal method or methods.

Litrature Review

In the literature, although there are many methods for software error detection, related studies focusing on classification methods are discussed in this study. With the literature review the aim of this study is to determine the most commonly used classification methods. Studies in this area have recently been supported by deep learning methods or alternative approaches. However, due to the nature of the problem, many studies have sought solutions using classification algorithms. Li et al. (Li et al., 2020) conducted a study which shows 47.7% of the studies on software error estimation and machine learning techniques between 1991 and 2013 were based on Bayes learning technique and 74% of these studies were performed by Naive Bayes algorithm. In the second and third place, Decision Tree and SVM are shown.

Kumar and the others show that in the software error estimation study, SVM could give the most effective results and stated that the performance was improved with LSSVM method (Kumar & Ranjan, 2017). Hidmi O. and others used K-NN and SVM methods for software error estimation (Hidmi et al., n.d.). The reasons for choosing these two models are to compare the results obtained with both parametric and non-parametric techniques.

Turhan B. and others preferred Bayes Classification method for Software Error Estimation. It shows the accuracy of the method by indicating that the obtained performance can be compared with the best results reported in the literature (Hata et al., 2007).

Goyal and others stated that the KNN algorithm provides superior performance over many linear regressions and will provide effective results in Software Error Estimation (Chandra et al., 2014). Hammouri A. and others compared Naive Bayes, Decision Tree and Artificial Neural Networks methods to software error estimation (Hammouri et al., 2018). According to the study, the most effective results are given by the Decision Tree method.

Ranjan et al. Conducted a bug prediction study showing that many open source systems were compared (Bal & Kumar, 2020). In this study, it has become possible to compare extreme machine learning techniques with various algorithms. As a result, a more consistent estimation system is recommended (ELM).

In addition to the studies mentioned above, many studies have been done with hybrid model development and ensemble techniques (Khuat & Le, 2019; Rathore & Kumar, 2019; Rhmann et al., 2020; Sharma & Chandra, 2018; Tumar et al., 2020; Turabieh et al., 2019). In these studies, many machine learning algorithms or deep learning algorithms were used together and more successful results were obtained compared to singular use. The obtained results showed that different algorithms stand out according to the problem type. However, there is an increase in the success percentages of the results obtained with the developing technology.

This study aims to identify the most commonly used and most effective methods in Software Error Estimation, unlike other related issues. After the determination of the methods, these methods were compared according to Accuracy, Recall, F-Measure, ROC, Precision criteria and model building times. In the literature review it was observed that the commonly used methods were KNN, Naive Bayes, Decision Tree and SVM.

METHODS

In this study, four different up-to-date Classification Algorithms used in machine learning were used to determine software fault estimation. These methods are 'decision tree', 'support vector machine', 'k-nearest neighbor' and 'naive bayes' algorithms, respectively. These methods have different categorical characteristics and different mathematical approaches. Each algorithm has its own advantages and disadvantages. As a result of the literature review, it has been observed that these 4 algorithms are mostly used in "software error detection" studies. It has been observed in many studies that these algorithms are used or tested in a hybrid manner with different algorithms. For this reason, it is aimed to measure among the best or most preferred algorithms. These selected algorithms have been tested on a suitable data set and it has become possible to compare their successes thanks to various metrics. The data set and measurement metrics used are detailed in the subtitles.

The following section briefly explains the methods which are used in the study. Then, these methods are compared according to criteria is used and the results are interpreted.

Decision Tree

Decision trees are a learning model in hierarchical supervise. Divide and use approach is follow up. It is a method of performing classification and regression efficiently without requiring parameters (Turney, 1995). Decision trees are consist of nodes. These nodes are called decision and end nodes. Each decision node refers to the test function that will allow the branches to differentiate. When this test function is applied to our data, suitable branch is determined according to the result. This process recursively continues from the root node to the end node. As a result of this process, the class which the input belongs is determined. The working method of the Decision tree algorithm is shown in Figure 2.

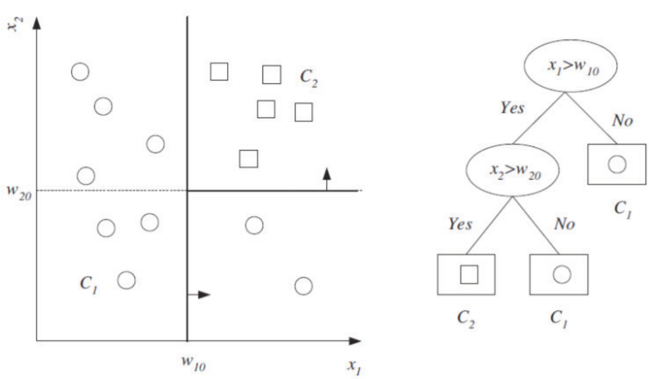


Figure 2. Decision tree construction process on sample (Özdemir,2014)

Decision Tree Regression is preferred due to its ability to explain results easily, not to be parametric and its speed advantages. It continues to give consistent results when faced with missing values or unrelated features. In addition to these advantages, it is preferable because of its low cost. The disadvantage is that it can create 'overfitting' problem.

Support Vector Machines

SVM, finds the appropriate linear separator plane that is farthest away from any input point in a multidimensional space. Implements a useful learning algorithm in the complex datasets of the SVM and in the identification of difficult-to-solve patterns (Shevade et al., 2000) . This algorithm finds that the samples are belong or not belong to a distinguishable class according to former observations. Figure 3 shows the general working principle of an SVM classifier.

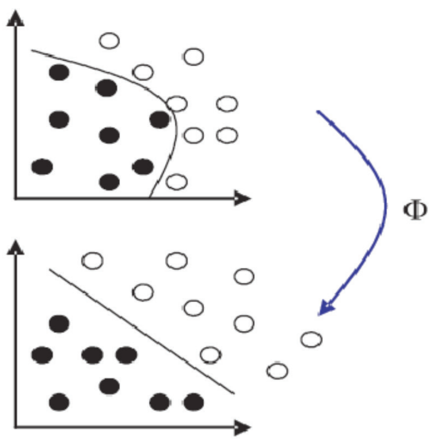


Figure 3. General principle of SVM (Madzarov & Gjorgjevik, 2009)

It is used in two different types to make classification (SVC) and regression (SVR). SVM does not require any unified distribution function information for the data. The learning process is completely independent of distribution. Its main advantages are its ability to work with a large number of independent variables, to be applied to both linearly separated and non-linearly separated data and to provide high accuracy in these studies. According to many other methods, the rate of overfitting is low. Disadvantages are that they cannot produce probabilistic predictions and that the results are in the form of binary classification.

K – Nearest Neighbors (KNN)

KNN is a trained learning algorithm that categorizes a new data on existing learning data. When new data comes to the algorithm, it looks at the closest 'k' neighbor and decides which class this data belongs to.

In order to apply the KNN method, it is necessary to create the attributes feature vectors. The nonparametric KNN classification algorithm has been used for statistical analysis for over 60 years. The correct selection of the k parameter on the accuracy of this algorithm is quite important. In the regulation of the model, the k value is determined by the cross validation method. Another factor that influences the performance of the algorithm is the distance criteria. Different measurement techniques such as Manhattan Distance, Minkowski Distance, Oklid Distance, Chebyschev Distance and Dilca Distance are used as distance measurements. But the most commonly used of these distance measures is the measure of Oklid (Soucy & Mineau, 2001). Figure 4 shows the general working principle of an KNN classifier.

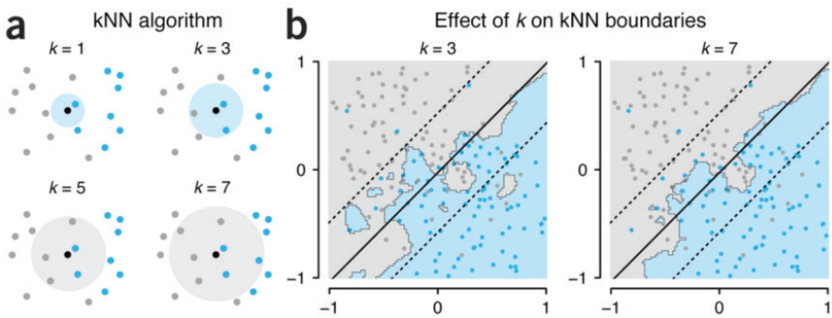


Figure 4. General principle of KNN (Bzdok et al., 2018)

Naïve Bayes Classifier

Based on Bayesian network theory, the Naive Bayes method categorizes using probabilities. The basic logic is to choose the highest probability decision. The use of this algorithm and its mathematical background is quite simple. Excessive amount of data does not create significant impact and cost. The term 'naive' used in nomenclature emphasizes the assumption that the attributes in the data set are independent of each other (Karim & Rahman, 2013). But this assumption will not apply to a significant part of the classification problems encountered in real life. Therefore, it is necessary to make sure that the data set to be studied is compatible with the algorithm to be used. The most important disadvantage is that the input data must be properly prepared for the specified reason. The working principle of the Naive Bayes classifier is shown in Figure 5.

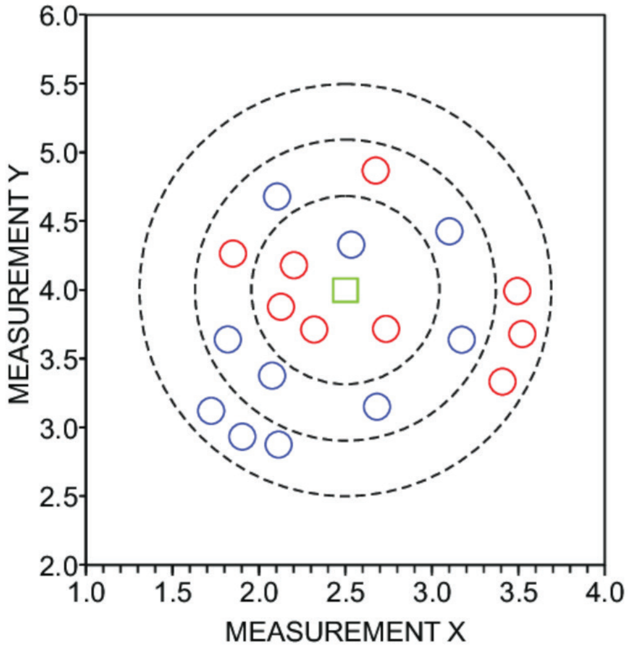


Figure 5. Example of naive bayes classifier setting boundaries according to colors (Silva & Amancio, 2013)

DATA AND COMPARISION CRITERIA

Dataset

For the study, the data set obtained by the related classification and data mining method produced by NASA was used [20]. This data set (pc1) produced by NASA Metrics Data Program contains 10885 samples. 22 attributes are used for each sample. 5 of these attributes show different code criteria, 3 of them are McCabe's metrics, 4 of them are of Halstead-based measure, 8 of them are derived from Halstead's measurement, and 1 of them shows the result. There is no missing data in the data set. The defects attribute is the result that indicates whether there are errors in the software for us.

This field was created by evaluating the other 21 attributes given.

Comparision Criteria

To measure the classification success, the confusion matrix is used first. The confusion matrix consists of rows-columns of up to the number of classes and it holds a different value in each cell. Table 1 shows the confusion matrix annotated.

Table 1. Confusion matrix

	Predicted Positive	**Predicted Negative**
Actual Positive	True Positive	False Positive
Actual Negative	False Negative	True Negative

The rows in the table show the actual classification results, while the columns show the estimates. The following criteria are used for comparison by using the error matrix.

Accuracy : The criterion that gives the ratio of cerrectly classified positive inputs to total positive inputs. Equation (1) shows the formula for calculating the value of accuracy.

$$\frac{TP+FN}{TP+FP+TN+FN} \qquad (1)$$

Precision : The criterion that gives the ratio of positive inputs classified as true to total positive values. Equation (2) shows the formula for calculating the value of precision.

$$\frac{TP}{TP+FP} \qquad (2)$$

Recall : The ratio of positive inputs classified as true to true positive values. Equation (3) shows the formula for calculating the value of recall.

$$\frac{TP}{TP+FN} \qquad (3)$$

F-Measure : It is a measure that is calculated as the harmonic mean of the Precision and Recall values. Equation (4) shows the formula for calculating the value of f-measure.

$$\frac{2 \text{ x } Recall \text{ x } Precision}{Recall+ Precision} \qquad (4)$$

In addition to this criteria which is using the confusion matrix, 'runtime' will be used as a benchmark to evaluate the cost factor in the study as well. Because the selected methods are tested with the same data and the same devices, it is possible to make a cost comparison of the methods by means of working times.

EVALUATION RESULTS

Based on the literature review conducted in this study, 4 of the most frequently used classification methods for software error estimation were selected. The selected algorithms were chosen from different categories attentively. Each method based on Accuracy, Precision, Recall, F-Measure and model creation times as seen in Table 2. The time factor was also included in the evaluation in order to provide the least cost if other evaluation criteria were in close proportion. While the models mentioned in Table 2 were applied, the k variable in the KNN model was determined as 3, the distance criterion was taken Oklid and rhe cross validation value was chosen as 10. SLib library was used for SVM and J48 was used represent as Decision Tree model. The values produced by the algorithms used as a result of the experiments are shown in Table 2.

Table 2. Evaluation results

	Accuracy	Precision	Recall	F-Measure	Time
KNN	79.200	0.843	**0.911**	**0.876**	**0.01 sec.**
SVM	**80.725**	**0.844**	0.807	0.722	3.87 sec.
Naive Bayes	**80.422**	0.765	0.804	0.770	0.15 sec.
Decision Tree	**79.503**	0.759	0.795	0.769	1.25 sec.

When the results are examined, it is seen that there are close Accuracy rates. One of the reasons for this proximity is that there are no faulty or noisy data in the data set used. When the results are examined, it is seen that the highest accuracy ratio is reached in SVM model. However, in this algorithm, the creation time of the model is considerably higher than the others. The KNN algorithm gave the lowest accuracy rate despite the fastest model creation time. The applied models carry such advantages and disadvantages. Therefore, according to the priorities of the project (accuracy ratio, cost), the preferred methods should be changed. The Naive Bayes algorithm yielded the most effective results in the ROC analysis, but it was far below the average in model building time. Therefore, the current software error estimation is used in the majority of studies.

CONCLUTION

In this study, estimation of software errors is considered as a classification problem. In order to solve this problem, the most commonly used methods in the literature were researched and four different methods were compared. The comparison of these methods was based on the model creation time in addition to the criteria related to the error matrix. When the results of the study were examined, it was observed that the Decision Tree method gave worse results than the other methods. The other three methods produced very close results. According to the results of the data to be studied and the cost plan of the study, it was found advantageous to choose one of the SVM, KNN and Naive Bayes methods. According to the literature, the success of Naive Bayes algorithm, which is the most commonly used problem, was confirmed by the study. A suitable future study will be to compare the results by taking the combinations of the algorithms examined. In addition, seeking answers to the question 'Is it possible to minimize the disadvantages of the evaluation methods and achieve higher performance rates with alternative kernels?' is a candidate to be the subject of the next study.

REFERENCES

Angell, R., Johnson, B., Brun, Y., & Meliou, A. (2018). Themis: Automatically testing software for discrimination. *ESEC/FSE 2018 - Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, 871–875. https://doi.org/10.1145/3236024.3264590

Bal, P. R., & Kumar, S. (2020). WR-ELM: Weighted Regularization Extreme Learning Machine for Imbalance Learning in Software Fault Prediction. *IEEE Transactions on Reliability*, 1–21. https://doi.org/10.1109/tr.2020.2996261

Boehm, B., & Basili, V. R. (2001). Software Defect Reduction Top 10 List. In *Computer* (Vol. 34, Issue 1, pp. 135–137). https://doi.org/10.1109/2.962984

Bzdok, D., Krzywinski, M., & Altman, N. (2018). Points of significance: Machine learning: Supervised methods. In *Nature Methods* (Vol. 15, Issue 1, pp. 5–6). Nature Publishing Group. https://doi.org/10.1038/nmeth.4551

Chandra, P., Singh, Y., Mata Vaishno, S., & Goyal, R. (2014). Suitability of KNN Regression in the Development of Interaction Based Software Fault Prediction Models Representing Contexual Relations with Sanskrit Word Embeddings View project Social Media Analytics, like Twitter Sentiment Analysis View project ScienceDirect 2013 International Conference on Future Software Engineering and Multimedia Engineering Suitability of KNN Regression in the Development of Interaction Based Software Fault Prediction Models. *IERI Procedia*, *6*, 15–21. https://doi.org/10.1016/j.ieri.2014.03.004

Hammouri, A., Hammad, M., Alnabhan, M. M., Alnabhan, M., & Alsarayrah, F. (2018). Software Bug Prediction using Machine Learning Approach Network Routing View project E-learning View project Software Bug Prediction using Machine Learning Approach. *Article in International Journal of Advanced Computer Science and Applications*, *9*(2). https://doi.org/10.14569/IJACSA.2018.090212

Hata, Y., İçin, K., Kod, K., Dayalı, Ö., Sınıflandırması, B., Turhan, B., & Üniversitesi, B. (2007). *Predicting Software Defects Across Project View project RESEARCH METHODS USED IN CIVIL ENGINEERING View project*. https://www.researchgate.net/publication/267383488

Hidmi, O., Eng, B. S.-I. J. C. C. I., & 2017, undefined. (n.d.). Software development effort estimation using ensemble machine learning. *Iieng.Org*. Retrieved December 20, 2020, from http://iieng.org/images/proceedings_pdf/26_D_New_formatted_journal.pdf

Karim, M., & Rahman, R. M. (2013). Decision Tree and Naïve Bayes Algorithm for Classification and Generation of Actionable Knowledge for Direct Marketing. *Journal of Software Engineering and Applications*, *06*(04), 196–206. https://doi.org/10.4236/jsea.2013.64025

Khuat, T. T., & Le, M. H. (2019). Ensemble learning for software fault prediction problem with imbalanced data. *International Journal of Electrical and Computer Engineering (IJECE)*, *9*(4), 3241–3246. https://doi.org/10.11591/ijece.v9i4.pp3241-3246

Kumar, S., & Ranjan, P. (2017). A Comprehensive Analysis for Software Fault Detection and Prediction using Computational Intelligence Techniques. In *International Journal of Computational Intelligence Research* (Vol. 13, Issue 1). http://www.ripublication.com

Li, N., Shepperd, M., & Guo, Y. (2020). A systematic review of unsupervised learning techniques for software defect prediction. In *Information and Software Technology* (Vol. 122, p. 106287). Elsevier B.V. https://doi.org/10.1016/j.infsof.2020.106287

Madzarov, G., & Gjorgjevik, D. (2009). Multi-class classification using support vector machines in decision tree architecture. *IEEE EUROCON 2009, EUROCON 2009*, 288–295. https://doi.org/10.1109/EURCON.2009.5167645

Rana, R. (2015). *Software Defect Prediction Techniques in Automotive Domain: Evaluation, Selection and Adoption*.

Rathore, S. S., & Kumar, S. (2019). A study on software fault prediction techniques. *Artificial Intelligence Review*, *51*(2), 255–327. https://doi.org/10.1007/s10462-017-9563-5

Rhmann, W., Pandey, B., Ansari, G., & Pandey, D. K. (2020). Software fault prediction based on change metrics using hybrid algorithms: An empirical study. *Journal of King Saud University - Computer and Information Sciences*, *32*(4), 419–424. https://doi.org/10.1016/j.jksuci.2019.03.006

Sharma, D., & Chandra, P. (2018). Software fault prediction using machine-learning techniques. *Smart Innovation, Systems and Technologies*, *78*, 541–549. https://doi.org/10.1007/978-981-10-5547-8_56

Shevade, S. K., Keerthi, S. S., Bhattacharyya, C., & Murthy, K. R. K. (2000). Improvements to the SMO algorithm for SVM regression. *IEEE Transactions on Neural Networks*, *11*(5), 1188–1193. https://doi.org/10.1109/72.870050

Silva, T. C., & Amancio, D. R. (2013). Discriminating word senses with tourist walks in complex networks. *European Physical Journal B*, *86*(7). https://doi.org/10.1140/epjb/e2013-40025-4

Soucy, P., & Mineau, G. W. (2001). A simple KNN algorithm for text categorization. *Proceedings - IEEE International Conference on Data Mining, ICDM*, 647–648. https://doi.org/10.1109/icdm.2001.989592

Tumar, I., Hassouneh, Y., Turabieh, H., & Thaher, T. (2020). Enhanced binary moth flame optimization as a feature selection algorithm to predict software fault prediction. *IEEE Access*, *8*, 8041–8055. https://doi.org/10.1109/ACCESS.2020.2964321

Turabieh, H., Mafarja, M., & Li, X. (2019). Iterated feature selection algorithms with layered recurrent neural network for software fault prediction. *Expert Systems with Applications*, *122*, 27–42. https://doi.org/10.1016/j.eswa.2018.12.033

Turney, P. D. (1995). Cost-Sensitive Classification: Empirical Evaluation of a Hybrid Genetic Decision Tree Induction Algorithm. *Journal of Artificial Intelligence Research*, *2*, 369–409. https://doi.org/10.1613/jair.120

Chapter 35

SOME THIN-FILM COATING METHODS

Mehmet ÖZKAN[1]

Zehra Nur OZER[2]

1 Assist.Prof.Dr. Mehmet OZKAN, Afyon Kocatepe University, Physics Department.

2 Assoc.Prof.Dr. Zehra Nur OZER, Afyon Kocatepe University, Physics Department.

Thin film coatings are used for many purposes in many different applications such as optoelectronics, semiconductor, environmental applications, energy storage devices, medical fields and so on. Here, we present some of the advance coating methods and principles of thin film deposition that are available.

In order to obtain thin films with good quality, there are many physical and chemical deposition thin film production methods. Physical deposition methods can be classified as evaporation and sputtering methods whereas chemical deposition methods are sol-gel, chemical bath deposition, spray pyrolysis, plating and chemical vapor deposition (CVD) methods. However, thin film coating methods under vacuum are widely used in technological studies in recent years in order to control production parameters.

Here, we will talk about CVD method from chemical methods and thin film coating methods under vacuum; thermal evaporation, electron beam evaporation, RF and DC sputtering, RF and DC magnetron sputtering, molecular beam epitaxy (MBE) and thermionic vacuum arc (TVA) methods.

1. Thermal Evaporation

Thermal evaporation method is a thin film coating method where metal materials are mostly heated and evaporated under high vacuum (10^{-7} mbar), (Tarımcı and Sarı, 2006). The thermal evaporation system is shown schematically in Figure 1. There are many applications of the metal coating process made with this method. Making metal contacts of semiconductor materials, surface metal plating processes are the best examples of processes performed with this method.

In this method, the materials to be evaporated are usually placed in a crucible made of metals such as tungsten (wolfram) with a very high melting temperature. The heating process is provided by a thermal heater. The material heated in the crucible evaporates and spreads in all directions. Thus, the desired thickness can be obtained on the substrate placed in a suitable place in the vacuum chamber (Tarımcı and Sarı, 2006; Akan, 2003).

This coating method has some disadvantages. These;

✓ It is very difficult to evaporate materials with very high evaporation temperatures (eg wolfram (W), tantalum (Ta) and molybdenum (Mo)).

✓ With the heating of the crucible, it is not very useful in the coating of substrates sensitive to high temperatures, as there will be too much heating in the vacuum chamber.

✓ Another disadvantage is that the impurity of the film may increase due to the risk of reaction of the source material to be evaporated with the heated crucible (Tarımcı and Sarı, 2006).

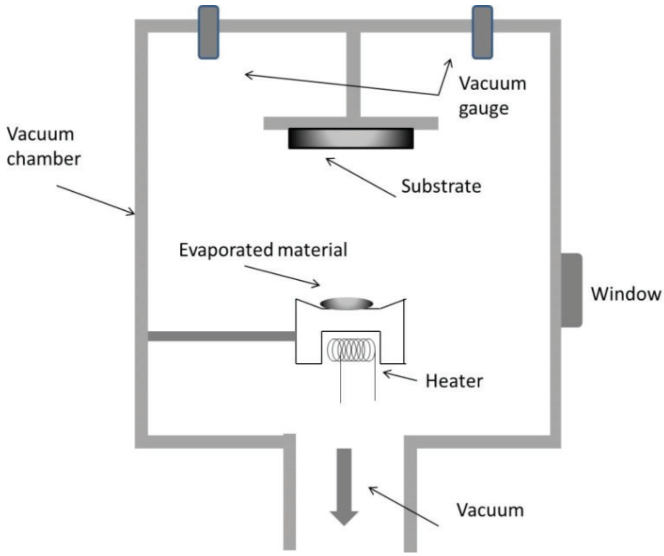


Figure 1.1. *Schematic representation of the thermal evaporation system.*

2. Electron Beam Evaporation

Another method commonly used in the metal coating process is the method using electron beam evaporators. This method in principle has almost the same properties as the thermal evaporation method. The most distinctive difference of this method is that the material to be evaporated is heated directly. In this method, the material to be evaporated is placed in a crucible made of high melting point metal. The material in the crucible is heated and evaporated by bombarding with accelerated and focused energetic electrons emitted by the filament located under the crucible and driven by a magnetic field. In this method, more effective energy transfer is performed to heat the material compared to thermal heaters. In addition, since the crucible can be cooled with water, the evaporation of the metal from which the crucible is made is prevented and the impurity that will form on the film is also improved (Akan, 2003; Tarımcı and Sarı, 2006). Since the material can be heated directly, it is possible to evaporate materials with high evaporation temperatures with this method. Figure 2.1, shows a schematic representation of the electron beam evaporation system.

The positive side of this method is that the impurities emitted from the hot filament used as the electron source are placed in the filament crucible to contaminate the metal of this geometry or the substrate. Although this method is superior to thermal evaporation method, it also has some disadvantages. In this coating method, since high voltage is used to accelerate the electrons, arc and electric discharge may occur during evaporation. In addition, the vacuum level used in this method should be higher than the vacuum level of other similar devices (10^{-8} mbar). High vacuum level is necessary for the electron beam to be directed towards the target (Tarımcı and Sarı, 2006). A quite amount of materials such as amorphous and crystalline semiconductors and molecular materials are also can be prepared by this technique (Jilani et. al., 2017).

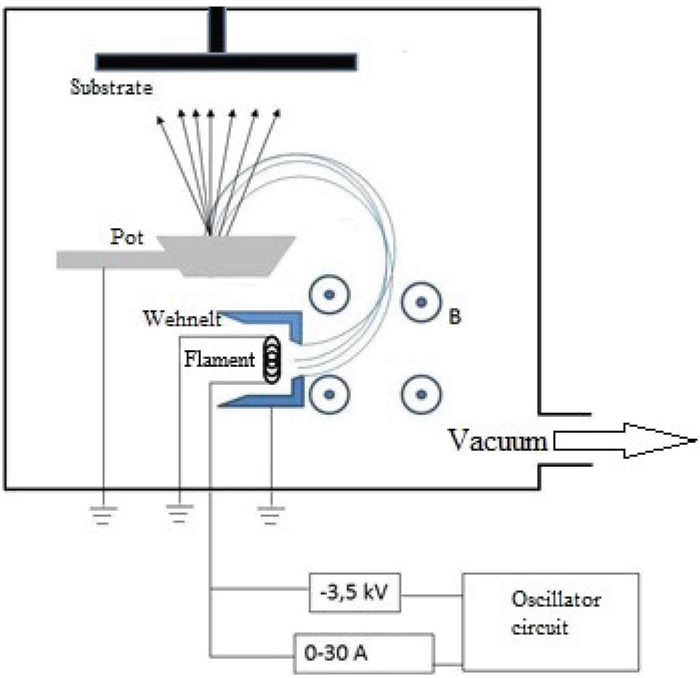


Figure 2.1. *Schematic representation of the electron beam evaporation system*

3. Chemical Vapor Deposition (CVD)

Chemical vapor deposition (storage) can be defined as the storage of solid by a chemical reaction (thermal decomposition, reduction, displacement, etc.) on the heated surface in the vapor phase. CVD technique is the most commonly used in synthesis of nanomaterials (Manawi et. al., 2018). The CVD system mainly consists of three parts (Bunshah, 2001; Pierson, 1992). These parts are; the part that feeds the reaction components in vapor state is the part that discharges the CVD reactor (base heating system) and the exit (exhaust) gases (Geçkinli, 1992). A diagram of this method is shown in Figure 3.1.

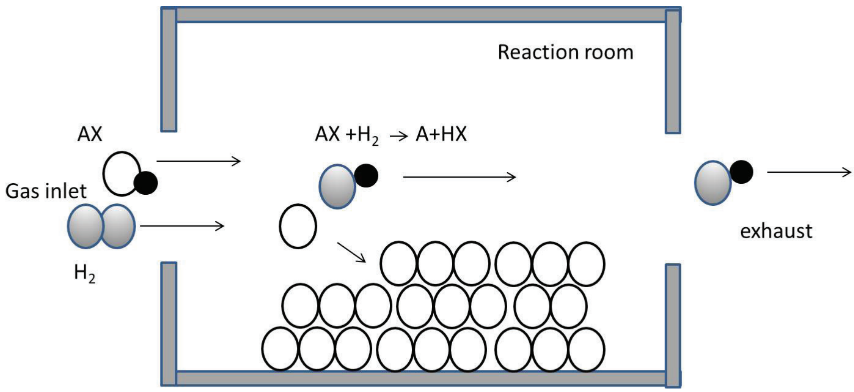


Figure 3.1. *Schematic view of the CVD system.*

There are many advantages of the CVD method:

✓ Generally compatible with substrate, ie these films can be applied to parts with complex shapes.

✓ It allows very high purity storage,

✓ High storage rate (centimeter thick in some cases)

✓ Usually higher vacuum is not required compared to many other methods (Bunshah, 2001; Pierson, 1992).

Disadvantages of CVD method are:

✓ Inability of storage on substrates that are not resistant to high temperatures (suitable for coating substrates at temperatures of 600 ° C and above),

✓ The reactive gases used are often dangerous and expensive,

✓ Some unwanted components formed as a result of the reaction may affect the substrate. (Geçkinli, 1992; Pierson, 1992).

4. Sputtering Method

Sputtering is the name given to the process of mechanically removing atoms from a solid surface and breaking it off (Tarımcı and Sarı , 2006). Spatter method is also the phenomenon of ions hitting a material surface removing atoms of that material surface (Campbell, 1978).

Gases that will not react with materials should be used in this method (eg Argon). In this method, the coating process is carried out in noble gas environment. The pressure of this method is higher than other similar methods (Tarımcı and Sarı, 2006).

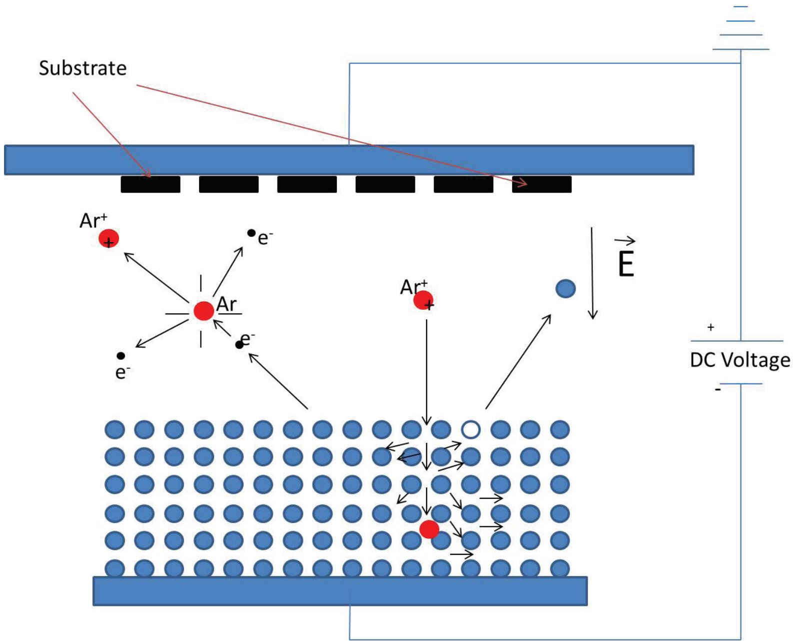


Figure 4.1. *Schematic representation of the sputtering process*

The schematic view of splattering process is shown in Figure 4.1. For the sputtering process, at first an inert gas is filled between the electrodes. By the potential applied between the electrodes, plasma of the inert gas is produced between these electrodes. With the formation of the plasma between the electrodes, the current suddenly increases and potential degreases. The increase in current is caused by the rapid increase in the number of charged particles due to the positive ions those formed in the plasma by hitting the cathode and emitting secondary electrons. Positive ions in the plasma accelerate towards the cathode at this potential degrease and hit the cathode. After a while, neutral material atoms begin to detach from these areas of the cathode surface. These removed neutral atoms are coated on the base surface on the anode located just opposite the cathode (Thornton 1983; Akan 2003).

There are two types of sputtering methods. These are radio frequency (RF) and direct current (DC) sputtering methods. The name of the sputtering method varies according to the type of power source that is used (Tarımcı and Sarı 2006; Johnson 2005; Bunshah 2001). For example, if the voltage applied between the electrodes is obtained from a radio frequency power source, it is called RF Sputter, while the voltage applied between the electrodes is called DC Sputter when the power source that generates a direct current.

Also, to increase the number of positive ions between electrodes, a magnet is used at the back of the cathode. Sputtering process done in this way is called Magnetron Sputtering (Akan, 2003; Johnson, 2005; Bunshah, 2001).

4.1. Radio frequency (RF) Sputtering method

The important aspect of this method is the use of RF power supply in the sputtering process. RF Splatter system consists of four main parts which are vacuum chamber, vacuum pump, radio frequency power supply and matching unit (Grill, 1993; Johnson, 2005).

Vacuum pump is used to reduce the pressure of the vacuum chamber. Vacuum is needed to clean the vacuum chamber from other gases in the air and to obtain high-energy collisions by increasing the average free path of ionized particles (Grill, 1993).

Electromagnetic waves are generated by the radio frequency power supply. These waves create a vibration frequency in the vacuum chamber and enable the target material to be ionized. High frequency makes plasma discharge continuous (Grill, 1993; Johnson, 2005; Bunshah, 2001). The schematic representation of the film formation by the RF Sputtering system is shown in Figure 4.1.1.

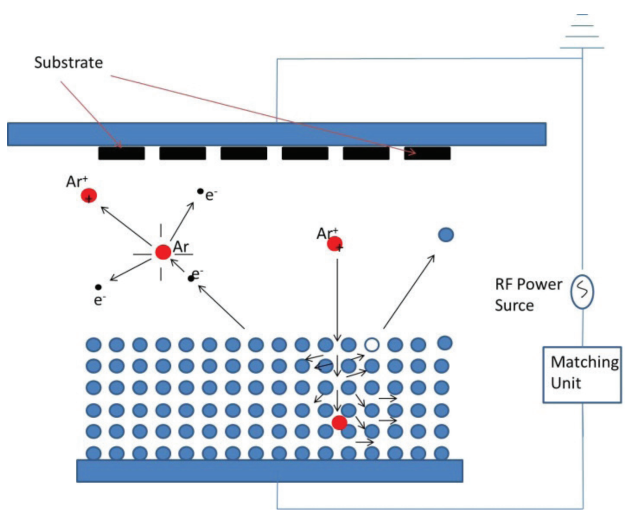


Figure 4.2. *Schematic representation of film formation by RF Sputtering system*

4.2. Direct current (DC) Sputtering method

In this method, the potential difference applied between the electrodes is provided by a direct current power source. It is used for coating metal materials. However, it is not suitable for coating insulating materials. Homogeneous and well-adhering films can be produced. In this coating method, the plasma density formed during the process is low. The lower the plasma density creates lower ion currents at the cathode. For this reason, the storage rate is low (Tüzemen, 2007; Bunshah, 2001).

4.3. Radio frequency (RF) magnetron Sputtering method

In this method, in addition to the electrical field used to accelerate the ionized noble gas (argon) atoms, a magnetic field is applied perpendicular to this field. Usage of magnetic field provides electrons to move along the spiral orbit. Because of this spiral trajectory of electrons their paths lengthen, so they collide with more neutral noble gas atoms along their movement and increase the ion concentration on the target. In this way, the process of atom removal from the target takes place more intensely and at the same time, plasma can be created at lower pressures. With this method, the electrons that are detached from the target are prevented from reaching the substrate by magnetic field. Thus, heating of the film on the substrate is prevented to a certain extent. This method is especially preferred for heat sensitive samples (Tarımcı and Sarı, 2006; Johnson, 2005; Bunshah, 2001). A schematic representation of the film formation with the RF magnetron sputtering system is shown in Figure 4.3.1.

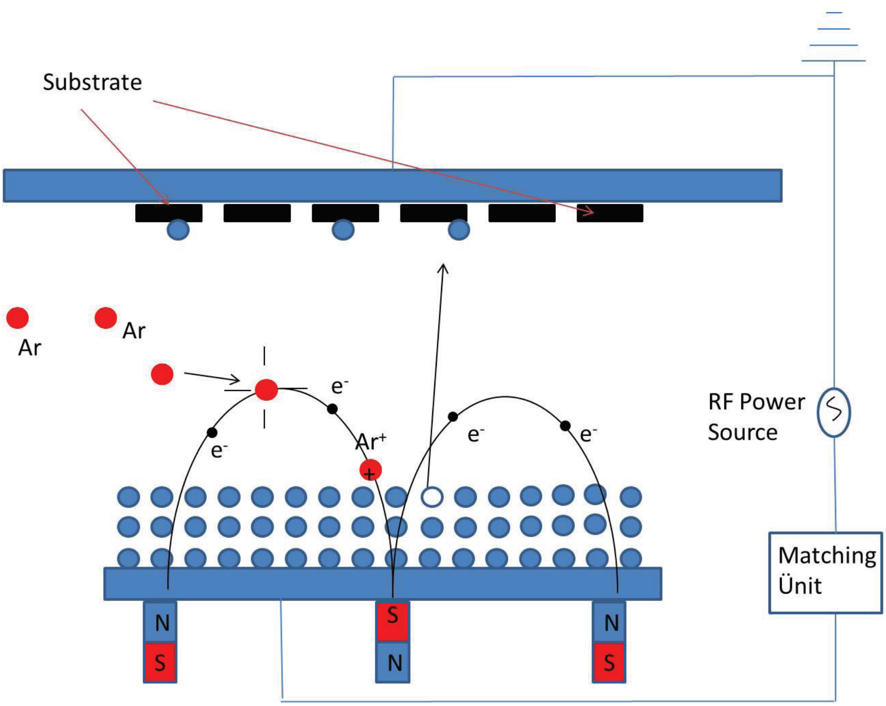


Figure 4.3. *Schematic view of film formation by RF magnetron Sputtering system (Johnson, 2005).*

4.4. Direct current (DC) magnetron Sputtering method

In this method, direct current power supply is used instead of radio frequency. It works on the same principle as the DC Sputtering method. Additionally, magnets are used. Plasma can be created at low pressures (Correia, et. al., 2018). It is a method that can be applied according to the feature of the film to be created (Tuzemen, 2007).

5. Molecular beam Epitaxy Method, MBE

It is usually considered as a refined style vacuum evaporation but it is actually different from simple evaporation. It generally includes multiple sources and coating of a heated material and it takes place in ultrahigh vacuum. Collision free molecular beams are used and it allows in-situ measurements as electron diffraction and mass spectrometry (Asahi and Yoshiji, 2019).

The epitaxial crystal growth method with a beam of molecules is a highly advanced single crystal growth method performed in an advanced high vacuum environment ($<10^{-11}$ mbar), where the film thickness can be controlled very precisely. Epitaxy, refers to the thin film enlargement process carried out in a way to preserve the crystal structure and orientation of the substrate on which it is grown, literally. Figure 5.1 shows a schematic view of an MBE assembly.

In this method, each of the materials containing the film (As, Ga, Al etc.) to be enlarged is placed individually in cylindrical containers known as knudsen cells. By properly heating the materials, the evaporated atoms can reach the substrate and the films with the desired content can be grown. There are flaps in front of the materials that can be opened and closed for a very short time (in the order of milliseconds). During magnification, the flaps in front of the sources that make up the content of the film to be enlarged are opened. Only the materials from this source reach the substrate and start to form the film (Tarımcı and Sarı., 2006).

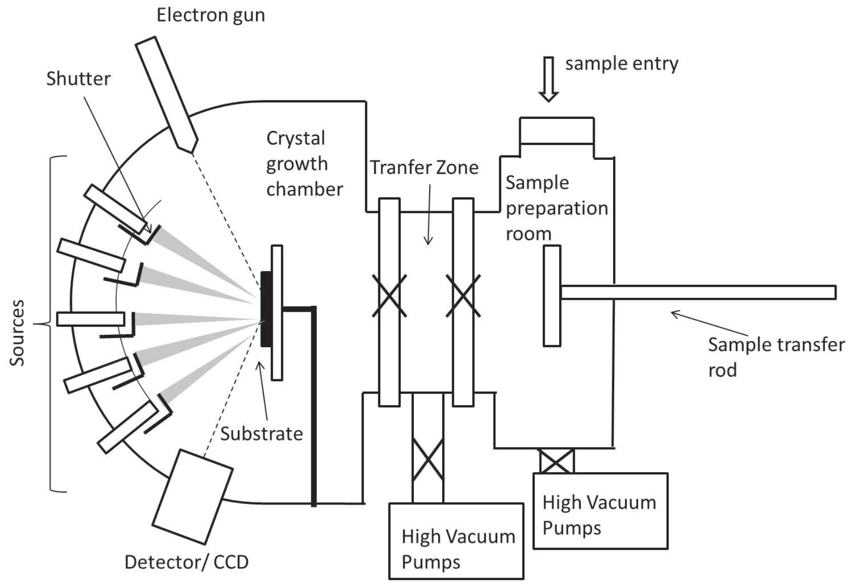


Figure 5.1. *Shematic representation of an MBE assembly (Tarımcı ve Sarı., 2006).*

In molecular epitaxy method, the temperature of the substrate has a great effect on the desired composition of the film to be formed. For this reason, during the crystal growth, depending on the type of film to be grown, the substrate is heated between 450 °C and 700 °C, and allowing the film to grow epitaxially. During the enlargement of the film, the substrate is rotated at a certain speed so that the films grow homogeneously (Tarımcı and Sarı 2006).

The MBE method has many advantages. Some of these can be listed as:

✓ To have film enlargement sensitivity on the order of single-row molecules,

✓ Very precise film thickness control,

✓ Measurement of film thickness during enlargement,

✓ Doping and especially the doping of a very small part of the film (delta doping) can be done.

This method has advantages as well as disadvantages. If the enlargements are not done under a very high vacuum, unwanted atoms can get mixed into the enlarged film. The purity of the materials to be grown must be very high (99.9%). A wide variety of high vacuum pumps used in MBE assemblies must be operated continuously even when no film enlargement is performed in the system. This causes both the installation and operation of the mechanisms to be very costly. Another negativity is that it is not ideal for mass production due to the very slow crystal growth rate (Tarımcı and Sarı., 2006).

6. Thermionic vacuum arc method (TVA)

TVA is an anodic discharge in vapors of the deposited material. The needed energy for the material to be melted is delivered by a focused electron beam. The energy of the ion can be controlled and ions fall on the substrate and they form a well-adhesive layer. The deposited layers are free from droplets that are typical for cathodic arc deposition systems and the substrates thermal stress that is being coated is low (Vladoiu et al., 2020). In this method includes an externally heated cathode mounted with a Wehnelt cylinder and by high voltage electrons are sent to the anode. Electron beam power provides a steady state vapor density. When there is a certain vapor density achieved, a bright discharge occurs between the spaces in the electrodes. Only the evaporating atoms are the maintaining discharged gas and this supply a high purity conditions for thin film deposition (Ozgür et. al., 2019; Pat et. al., 2011; Pat et. al., 2019).

To generate the TVA discharge, the cathode filament can be heated at the desired current with an AC low voltage power supply and thermoelectronic emission from the cathode filament. Here, one end of the cathode is connected to the low voltage power supply while the other end is connected to the vacuum chamber where the electrodes are mounted. The Wehnelt cylinder placed inside the cathode is connected to negative potential. After the electron emission is generated from the cathode filament, high voltage is applied between the anode and the cathode. Thus, electrons emitted from the cathode are focused on the anode by the help of the Wehnelt cylinder with acceleration (Figure 6.1). This is the most important feature that distinguishes the TVA system from other techniques that produce anode material vapor plasma.

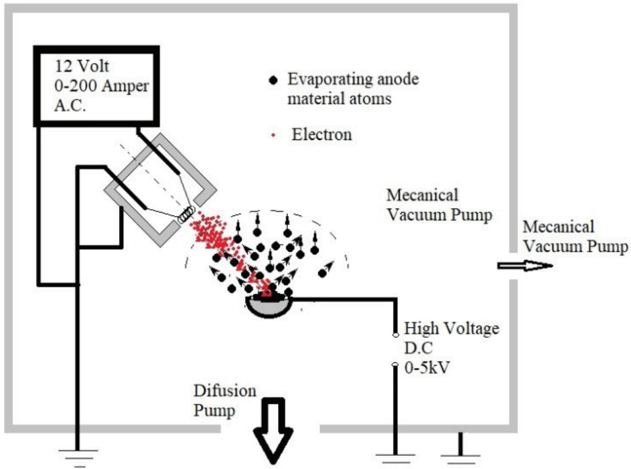


Figure 6.1. *TVA thin film coating system*

With the bombardment of accelerated electrons focused on the anode, energy is transferred into the material inside of the anode and the anode material heats up. By increasing of the applied voltage, the material in the anode is melted and then evaporated (Figure 6.1).

If the applied voltage between the electrodes increases, a bright discharge occurs in the anode material vapor at the appropriate value of the applied voltage in the space between the electrodes. At this moment, the voltage between the electrodes suddenly drops and the current increases. The plasma of the vapor of the anode material formed continuously spreads over the anode towards the vacuum chamber. In this case, substrates or any material placed in a suitable place in the vacuum chamber will be coated in terms of the anode material.

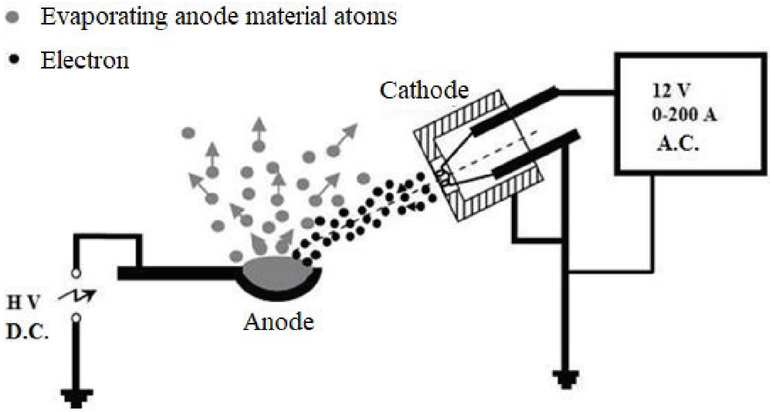
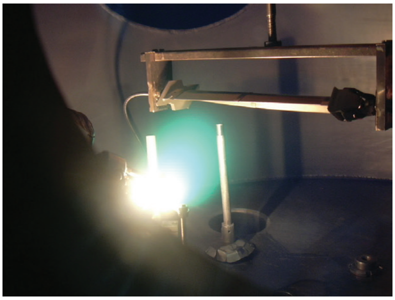
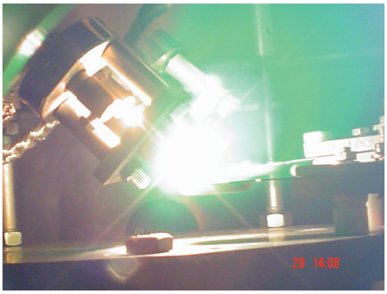


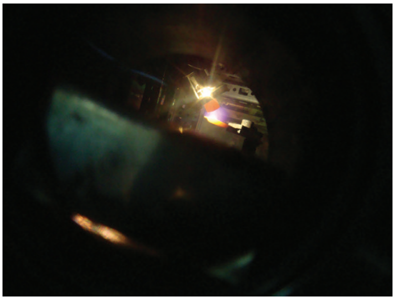
Figure 6.2. *Evaporation of anode material*

Plasma of many materials (such as copper, silver, aluminum, tin, gold, nickel, alumina, zirconium oxide, boron, carbon, rhenium, molybdenum, tantalum, tungsten and beryllium) has been produced with the TVA system and thin films of these materials or thick coatings could be made. In addition, it is possible to produce thin films of alloys, semiconductors or superconductors using TVA system. Figure 6.3 shows the plasma of some elements and compounds formed in TVA system.



(a) (b)





(c) (d)

Figure 6.3. *Some elements and compounds whose plasma is formed in TVA. a) Copper (Cu), b) Boron (B), c) Magnesium (Mg), d) Zinc oxide (ZnO).*

6.1. Scientific Features and Benefits of the TVA Technique

It is possible to work under high vacuum or very high vacuum conditions with the TVA technique. Workings in this level of vacuum increase the quality of the coating and minimize the effects of oxidation or other gases on the film. In some of the other techniques, high vacuum conditions cannot be achieved. A buffer gas is not needed while coating in TVA technique. Thus, it provides high purity of coating the material. Especially in sputter technique, a buffer gas must be used. This may cause the formation of unwanted impurities in the produced films in analysis. In TVA technique, by changing the distance between electrodes, angle, cathode temperature, filament current and voltage, the ionization energies of the material to be coated can be controlled and high ion energies can be obtained. Control of ion energies is very important. This situation increases the adhesion of the ions of the material to be coated to the film and provides the film to be coated with low roughness. In other coating techniques, this adhesion rate (especially in thermal evaporation technique) is very low. Therefore, the surface roughness of thin films produced by thermal evaporation is high (Musa et. al. 1983).

Materials including refractory materials or ceramics with a very high melting point can be easily stored by the TVA technique. It is not possible to store every material in other coating techniques. Simultaneous storage of more than one material with independent power sources can also be performed with TVA technique (Balbağ, 2009).

6.2. Properties of Thin Films Produced with TVA Technique

As a result of the examination of thin film coatings produced with TVA system, the following properties were found:

✓ High purity thin films are produced (Balbağ, 2009).

✓ The surface adhesion of thin films is quite high (Pat et. al., 2019; U. Demirkol, et. al., 2018)

✓ Since the ion energies of TVA plasmas can be controlled, the penetration depth of thin films on the substrate can be changed arbitrarily (Vladoiu et. al., 2003; Pat et. al., 2005; Akan et. al.,2007).

✓ It has been observed that the surface roughness of thin films produced with TVA system is low (surface roughness was found below 10 nm in boron thin films (Ekem et. al., 2008).

✓ It has been observed that thin films produced with TVA are in a very tight structure (Ekem et. al., 2008).

✓ In TEM analysis of thin films produced by TVA technique, it has been observed that the thin films are in nano dimensions (Akan et. al., 2007; Musa et. al., 2006).

✓ Thin films of refractory materials with very high melting temperatures can be easily coated with the TVA system (Surdu et. al., 2007;Musa et. al., 2004; Vladoiu et. al., 2003; Lungu et. al., 2006; Vladoiu et. al., 2006; Lungu et. al., 2005; Musa et. al., 2005; Lungu et. al., 2006).

✓ Thin film storage rate can be adjusted while coating in TVA technique (Musa et. al., 1984; Akan et. al. 2006).

✓ In the TVA system, any type of substrate can be coated. Various substrates such as steel, silicon, glass, chrome thin films, NaCl crystal, copper, aluminum and plastic have been coated up to now (Musa et. al., 1984; Akan et. al 2006).

REFERENCES

Yayıncılık Tarımcı, Ç., Sarı, H. (2006). Vakum tekniği. Seçkin Kitabevi, 1. Baskı, 168 s.

Akan, T., 2003, Investigation of the basic properties of Thermionic Vacuum Arc (TVA), Phd thesis, Osmangazi University, Graduate School of Natural and Applied Sciences.

Jilani, A., Abdel-Wahab, M. S., Hammad, A. H. (2017). Advance deposition techniques for thin film and coating. Modern Technologies for Creating the Thin-film Systems and Coatings, 2, 137-149.

Manawi, Y. M., Samara, A., Al-Ansari, T., Atieh, M. A. (2018). A review of carbon nanomaterials' synthesis via the chemical vapor deposition (CVD) method. Materials, 11(5), 822.

Bunshah, R. F., 2001, Handbook of Hard Coatings, Published in the United Statesof America by Noyes Publications / William Andrew Publishing, LLC Norwich,NewYork, U.S.A. , p.549.

Pierson, H.O., 1992, Handbook of chemical vapor deposition, Principles Technology and Applications, Noyes Publications, Park Ridge, Second Edition, New Jersey, U.S.

Geçkinli, E. (1992). İleri Teknoloji Malzemeleri (1. Baskı). İstanbul Teknik Üniversitesi Basımevi, 288s, İstanbul.

Campbell, D. S., 1978, Active and passive thin films devices, Academic press, London, chapt.2.

Thornton, J. A., 1983, Plasma assisted deposition process: Theory, mechanism and applications, Thin Solid Film, vol.107, 3-19.

Johnson, R. L., 2005, Characterization of piezoelectric ZnO thin films and the fabrication of piezoelectric micro-cantilevers, Master Thesis, Electrical Engineering, Iowa State University, Ames, Iowa, 98 p.

Gril, A., 1993, Cold plasma in materials fabrications (from Fundamentals to applications), , 1993, The institute of electrical and electronics engineers, New York.

Tüzemen, E. S., 2007, ZnO Thin Films and Optimization of Parameters for Device Production, Phd thesis, Çukurova University, Graduate School of Natural and Applied Sciences., 208s.

Asahi, H., Horikoshi, Y. (Eds.). (2019). Molecular Beam Epitaxy: Materials and Applications for Electronics and Optoelectronics. John Wiley & Sons.

Correia, F. C., Salvador, P. B., Ribeiro, J. M., Mendes, A., & Tavares, C. J. (2018). Corrigendum to Effect on the electrical and morphological properties of Bi incorporation into ZnO: Ga and ZnO: Al thin films deposited by confocal magnetron sputtering [Vacuum 152 (2018) 252-260]. Vacuu, 154, 340.

Vladoiu, R., Tichý, M., Mandes, A., Dinca, V., Kudrna, P. (2020). Thermionic Vacuum Arc—A versatile technology for thin film deposition and its applications. Coatings, 10(3), 211.

Özgür, M., Pat, S., Mohammadigharehbagh, R., Musaoğlu, C., Demirkol, U., Elmas, S., Korkmaz, Ş. (2019). Sn doped ZnO thin film deposition using thermionic vacuum arc technique. Journal of Alloys and Compounds, 774, 1017-1023.

Pat, S., Temel, S., Ekem, N., Korkmaz, Ş., Özkan, M., Balbağ, M. Z. (2011). Diamond-like carbon coated on polyethylene terephthalate by thermionic vacuum arc. Journal of Plastic Film & Sheeting, 27(1-2), 127-137.

Pat, S., Korkmaz, Ş., Özen, S., Şenay, V. (2019). Optical, surface and magnetic properties of the Ti-doped GaN nanosheets on glass and PET substrates by thermionic vacuum arc (TVA) method. Particulate Science and Technology, 37(3), 333-338.

Musa, G., Betiu, N., Mustata, I., Baltog, A. and Popescu, A., 1983, Low voltage arc welding in vacuum, Review Roum. Physics, 28, 907.

Balbağ, M., Z., 2009, Production of Magnesium, Boron and Magnesium Boride Thin Films with Thermionic Vacuum Arc (TVA) Technique and Investigation of Some Physical Properties, Phd Thesis, Eskişehir Osmangazi University, Graduate School of Natural and Applied Sciences.

Elmas, S., Pat, S., Mohammadigharehbagh, R., Musaoğlu, C., Özgür, M., Demirkol, U., Korkmaz, Ş. (2019). Determination of physical properties of graphene doped ZnO (ZnO: Gr) nanocomposite thin films deposited by a thermionic vacuum arc technique. Physica B: Condensed Matter, 557, 27-33.

Demirkol, U., Pat, S., Mohammadigharehbagh, R., Musaoğlu, C., Özgür, M., Elmas, S., Korkmaz, Ş. (2018). Investigation of the substrate effect for Zr doped ZnO thin film deposition by thermionic vacuum arc technique. Journal of Materials Science: Materials in Electronics, 29(21), 18098-18104.

Vladoiu, R; Musa, G; Mustata, I, 2003, Thermoionic vacuum arc - A new method of thin film deposition, Journal Of Optoelectronıcs And Advanced Materials Volume:5, Issue: 1 Pages: 325-330.

Pat, S., Ekem N., Akan T., Kusmus O., Demirkol S., Vladoıu R., Lungu C.P., Musa G., 2005, Study on Thermionic Vacuum Arc-A Novel and Advanced Technology For Surface Coating, Journal of Optoelectronics and Advanced Materials Vol.7, No.5, p.2495-2499.

Akan T., Ekem N., Pat S., Issever U.G., Balbag M.Z., Cenik M.I., Vladoiu R., Musa G., 2007, Boron Thin Film Deposition By Using Thermionic Vacuum Arc (TVA) Technology, Materials Letters, 61, p. 23–26.

Ekem, N., Musa G., Pat S., Balbag Z., Cenık I., Vladoıu R., Carbon thin film deposition by Thermionic Vacuum Arc (TVA), 2008, Journal Of Optoelectronics And Advanced Materials, Vol. 10, No. 3, p. 672 – 674.

Musa, G., Vladoiu R., Ciupina V., Lungu C.P., Mustata I., Pat S., Akan T., Ekem N., 2006, Characteristics of boron thin films obtained by TVA technology, Journal of Optoelectronics and Advanced Materials, Vol: 8 (2), p: 617-620.

Surdu-Bob, C; Musa, G; Vladoiu, R, et al., 2007, The synthesis of DLC using a novel cathodic arc technique: Gas-TVA, Journal Of Optoelectronics And Advanced Materials, Volume: 9 Issue: 8 Pages: 2660-2662.

Musa, G; Mustata, I; Ciupina, V, et al., 2004, Diamond-like nanostructured carbon film deposition using thermionic vacuum arc, Diamond And Related Materials, Volume:13, Issue:4-8, p.1398-1401.

Lungu, CP; Mustata, I; Musa, G, et al., 2006, Unstressed carbon-metal films depositedby thermionic vacuum arc method, Journal of Optoelectronics and Advanced Materıals, Volume: 8 Issue: 1 Pages: 74-77.

Lungu, CP; Mustata, I; Lungu, AM, et al., 2005, Influence of Re on the thermo-electron emission from thoriated W cathode during Re deposition by thermionic vacuum arc (TVA) method, Journal Of Optoelectronıcs And Advanced Materıals, Volume: 7 Issue: 5 Pages: 2513-2519.

Musa, G(d)., Mustata I., Lungu C. P., Bob C. S., Lungu A. M., Bursikova V., Vladoiu R., Prodan G., Ciupina V., 2005, Tungsten film deposition by thermionic vacuum arc method, XXVIIth ICPIG, 18-22 July, Eindhoven, the Netherlands.

Vladoiu, R., Ciupina, V., Mustata, I., Lungu, C. P., Musa, G. (2006). Characterization of carbon thin film deposited by thermionic vacuum arc (TVA) method. Romanian Journal of Physics, 51(1/2), 215.

Lungu, CP; Mustata, I; Musa, G, et al., 2006, Unstressed carbon-metal films depositedby thermionic vacuum arc method, Journal Of Optoelectronıcs And Advanced Materıals, Volume: 8 Issue: 1 Pages: 74-77.

Musa, G., Betiu N., Mustata I., Baltog A., Popescu A., 1984, Heated Cathode Arc In Vacuum ArcIn Vacuum Developed In The Vapors of The Heated Anode, Proc. XI th Symp. On Discharges And Insulation In Vacuum, Berlin, p.229.

Akan T., Ekem N., Demirkol S., Pat S., Balbag M.Z., Cenık M.İ., Deligoz H., Musa G., 2006, Studies on Ag-Al2O3 nano-layer composite produced by Thermionic Vacuum Arc Methode (TVA), 6th International Conference of the Balkan Physical Union-August 22-26, Istanbul,Turkey.

Chapter 36

IMPACT BEHAVIOR OF LAMİNATED COMPOSITE MATERIALS EXPOSED TO LOW-VELOCITY IMPACT LOADS

İbrahim YAPICI[1]

1 Bitlis Eren University Technical Sciences Vocational School, 13000, Bitlis, iyapici@beu.edu.tr

1. INTRODUCTION

Today, people's needs increase with the development of technology. New material production is one of these needs. It is seen that the main materials that are limited in nature and their properties are insufficient with the development of technology (Parlakyıldız, 2018 and 2020; Gencer et al, 2017;Efe, 2018). For this reason, efforts to produce materials that are both economical and energy efficient, as well as very light and durable have been intensified. For example, new efficient products in lighting or efficient devices created by optimization studies in power systems are examples (Efe, 2013; Yıldırım et al, 2017; Cengiz et al, 2017). Thus, composite materials obtained by physically combining multiple materials with different properties with special methods have gained great importance. Composite material s better mechanical and thermal properties in order to achieve combine macro level is a made obtained. Composite materials are generally composed of a main structure called matrix and reinforcement material, in a way to show the best properties of the materials that make them up. Therefore, determining the mechanical properties of these materials is important in determining the usability of materials in this class (Kara, 2006).

In engineering applications, especially in mechanical applications, it is desired that the material can give the most appropriate response in order to prevent unexpected results against any impact from outside. Impacts to which the material may be exposed can be in many different forms depending on the application and the purpose of use. Generally, pulses are classified as low-velocity and high velocity. However, a clear transition between low-velocity pulse and high-velocity pulse has not been revealed in the studies conducted so far. Some of the studies on this subject argue that low-velocity impact should be evaluated as velocities varying between 1 and 10 m/s depending on the stiffness of the target, material properties, mass and stiffness of the striking object.

The response to low-velocity impact is determined by the material it self. Namely, damage to metal and metal alloys as a result of impact occurs on the surface exposed to impact. In composite materials, the damage caused by impact may also occur on the surface not exposed to impact, depending on the type of impact. It can be seen as delaminations (separation between layers) in the internal structure. Although the impact response in metals is in the form of rupture or breakage as a result of plastic deformation, composites can be damaged in many different modes and there is no serious change in the structural integrity of the part in these damage modes (Ceyhun & Turan, 2003).

Laminated composite materials are used in many engineering fields and can be hit by foreign bodies. In this case the impact velocities will be small but their impact on the material will be large. Laminated composite structures are more susceptible to impact damage than similar metallic structures. Impact in composite structures creates internal damage that cannot be determined by visual inspection. This internal damage causes a decrease in strength and grows under load. Therefore, it should be understood effects of foreign body impact on composite structures at the design stage and appropriate precautionary er should be taken (Kara, 2006).

Many experimental studies have been done and analytical methods have been developed on the response of laminated composite materials to low-velocity impact. Some of the important work in this area are:

The low-velocity impact response of laminated composites was investigated analytically by (Ramkuar and Chen, 1982) and by (Abatan et al.,1998). An approximate solution for predicting plate response to low-velocity impact is presented by (Gong and Lam, 1999). This solution involves the joint movements of the plate and the reinforcing elements, such as the contact force and the effect of transverse shear deformation. The low-velocity impact response of hybrid laminated composite plates has been investigated using the shear deformation theory (Lee et al., 1997). (Kim and Kang, 2001) developed a new analytical method from dynamic deformation of composite plates subjected to transverse impact to predict impact force.

(Goo and Kim, 1997) conducted dynamic contact analysis of laminated composite plates under low-velocity impact. Simple laws, such as the corrected Hertz contact law, have been used to adapt the composite plates to the dynamic contact situation.

The impact behavior of E-glass/epoxy laminated composite plates has been experimentally studied by Mili and Necip, 2001. They used a weight reduction impact device. They evaluated the effects of impact tip velocities and layering sequence on the behavior of the composite plate. (Aslan et al., 2003), (Aslan and Karakuzu, 2002) evaluated the dynamic behavior of fiber-reinforced laminated composites subjected to low-velocity impact.

(Hosseinzadeh et al., 2006) studied the damage response of fiber-reinforced composite plates in falling weight blows. They made low-velocity pulses to four different fiber reinforced laminated composite plates. They obtained the variation of impact energy-damage diameter by mapping the damage regions in the samples. They reported the changes in different samples.

(Şenel et al., 2009) , in their study, designed and manufactured a unique braking and control system to be mounted on the low-velocity impact test assembly. Low-velocity impact tests were carried out to see the difference of the braking system on sample damage. In the experiments, they used 8-layer symmetrical epoxy-glass fiber composite plates with an orientation angle of [-30/30] 4s as a sample. Experiments were conducted at impact energies of 14.7J, 22.07J, 29.43J, 44.14J. As a result, the braking and control system used in the mechanism retained the mass in the event of the falling mass bouncing, and did not activate in the case of drilling the sample and showed the desired properties. By examining the effect of braking system on damage occurrence, they determined the necessity of this system and its proximity to real modeling.

(Kurşun and Şenel, 2011) investigated the low-velocity impact response of E-glass / epoxy laminated composites. Samples were produced in unidirectional reinforced sheets [0, 90] 2s and tested. Samples of 140 × 140 mm and 2 mm thickness were used in his work. They carried out their impact tests with a specially developed vertical weight drop test device. In their work, they used a striker tip with a 12 mm diameter hemispherical tip geometry. The batter mass is 3.1 kg. In line with the data obtained as a result of this test, they obtained force-time, energy-time and force-displacement graphs during impact on the material and concluded their work.

(Esendemir and Caner, 2018) in a study was made of two different thicknesses and pulse energy inside the glass fiber insert viyel epoxy composite plate behaviors of pulses n is experimentally examined un l av. For this purpose, with dimensions of 100x100 mm; two different thickness of the k (2 and 4 mM), 8, and 16 laminated composite samples vacuum infusion direction using REE Ami Raku rejection that samples i 20J and 60J ‘ luxury two different pulse energy under test they have. In line with the data obtained as a result of this test; They obtained force- collapse curves, impact energy-maximum force energy curves, impact energy heat - failure curves, velocity – time curves, force- time curves and absorbed energy-time curves. As a result of the experimental study, puncture damage did not occur in 8-layer samples exposed to 20J impact energy, while puncture damage occurred in samples exposed to 60J impact energy. They concluded that there was no puncture damage for both impact energies in 16 layer composite samples.

Impact m due to damage lz herein, may occur at the surface exposed to impact or shock may occur depending on the type of internal structure in the form of separation between the layers. In this study, the behavior of laminated composite materials against low-velocity impacts will be discussed.

2. IMPACT BEHAVIOR IN COMPOSITE MATERIALS

Coup; It can be defined as an instantaneous external force applied on a material or structure in a very short time at low, medium or high velocitys (Mili and Necip, 2001).

Damage caused by low-velocity impact can occur during manufacturing, maintenance and service operations. As an example of the impact during use, the impact caused by stones and small particles ejected from the runway after the high velocity of the tires during the take-off and landing of the aircraft can be given. The tools used during production and maintenance can be dropped on the structure (Kara, 2006).

2.1 Tendency to Impact in Composite Materials

The lack of plastic deformation in composite structures means that when a certain stress level is reached, permanent damage occurs that causes local or structural weakening. Unlike a metal that can still maintain its integrity after plastic deformation, composites that exceed a certain stress level, albeit still preserving some of their structural properties, can be permanently damaged. An impact with an energy level of 1 J or less and a velocity of 2 m / s can actually cause irreversible damage to the composite sheets. The causes of damages caused by low-velocity impact can be summarized as follows;

• Low transverse or interlayer shear strength,

• Little or no plastic deformation,

• Laminated structure (effective use of fibers).

In the research conducted by IATA (International Air Transportation Association), the causes of damage that an aircraft may be exposed to during its service period are shown graphically in Figure.1 (Karaca, 2010).

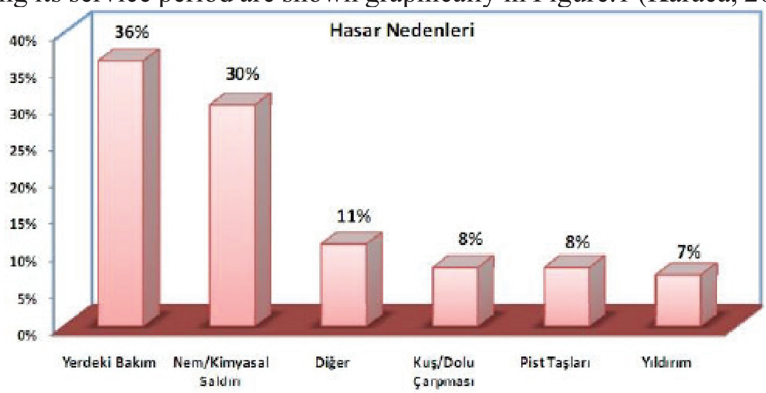


Figure 1. *Percentages of the causes of damage to aircraft during their service life (Karaca, 2010)*

2.2 Impact Tests of Composite Materials

There is no standard test technique or any technique widely accepted among researchers for determining the impact behavior of composite materials. Despite these negativities, the following test methods and devices are widely used today to determine the impact behavior of composite materials. These;

- Weight reduction tests,
- Pendulum tests (Izod, Charpy and Pendulum),
- Built-in beam impact tests,

2.2.1 Weight Drop Tests

The drop weight test method is widely used for impact tests of composite materials and serves different designs. The striker tip, usually attached to a weight group, is released from a certain height. They are equipped with specific mechanical parts to prevent repetitive impact (Adams, 2009).

Some of the drop weight testers measure displacement or acceleration. In this way, the change in load, displacement and acceleration at the moment of impact is recorded. These results can be converted into impact load - time and impact energy - time changes. In this way, properties such as extreme load and absorbed energy can be associated with the fracture process occurring in the material. A typical weight drop test device is as shown in Figure.2. Equipment making up such a device; The sample is supported by platforms, load measuring devices (load cell) placed inside the tube, photoelectric cells used to measure the tube velocity just before the collision and high velocity camera used to view the impact event.

The impact properties of composite materials are determined by the free weight drop impact tester, the characteristics of the striking body (hollow or solid, tip shape and size, etc.), the velocity and mass (or energy) of the striking body, the specimen configuration (size, geometry, specimen and tip at the support points). fixing etc.). Therefore, when it comes to impact properties of composite materials, all these factors should be taken into consideration (Yüce, 2007).

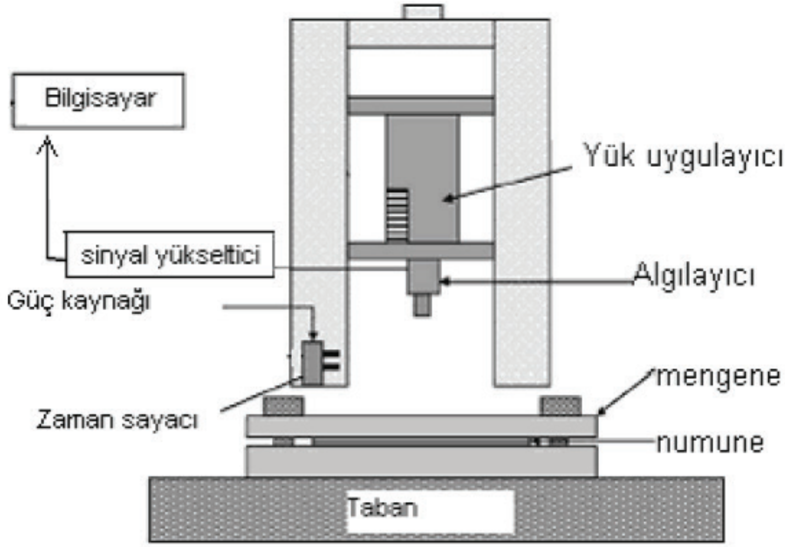


Figure 2. *Weight Reduction Test Setup (Yüce, 2007)*

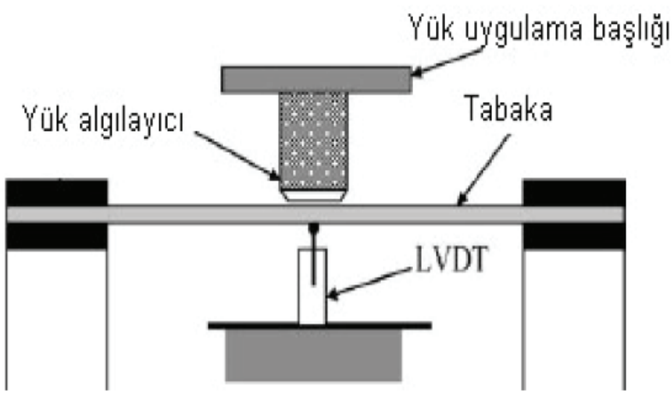


Figure 3. *Detailed View of the Part Where the Weight Contact the Sample in the Drop Weight Test Setup (Yüce, 2007)*

2.2.2 Pendulum Test Applications

Darbe Izod and Charpy test setup schemes of impact test methods are as shown in Figure.4. In the Charpy and Izod impact test methods, a notched test specimen is subjected to impact with a pendulum dropped from a standard height. The height at which the pendulum rises after the impact is determined, and the energy difference in the initial and final positions of the pendulum is measured as the impact energy absorbed by the sample. The lower the height of the pendulum after impact, the higher the impact energy absorbed, hence the impact resistance or toughness of the material. Charpy and Izod impact test methods are very similar, with only minor differences. The most important of these differences is the way the sample is supported and the position of the notch relative to the support and impact points as shown in Figure 4 (Ceyhun & Turan, 2003).

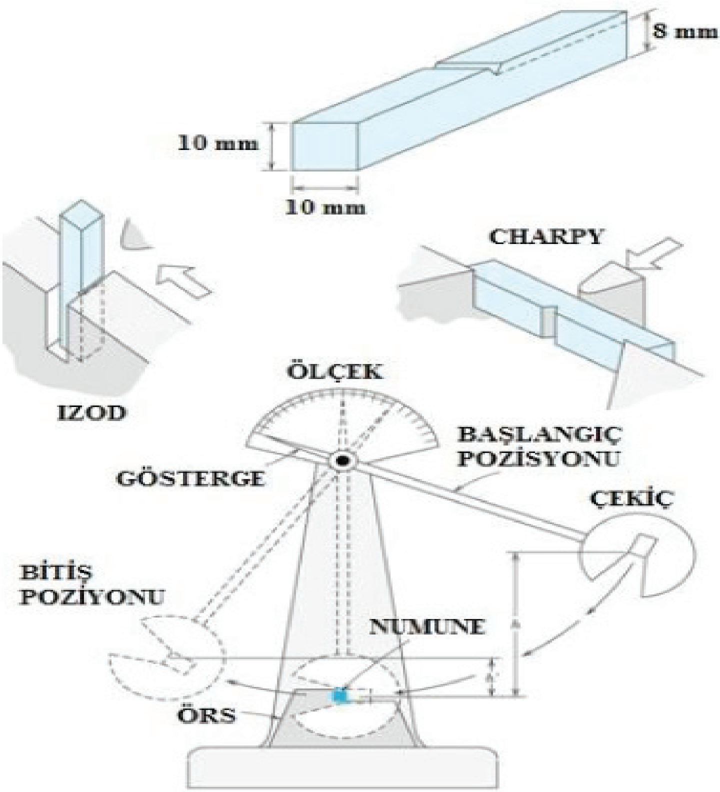


Figure 4. *Sample Image of Charpy and Izod Tests (Karaca, 2010)*

Pendulum pendulum impact test system has some advantages over weight drop impact test system. In this system, it has an advantage in measuring data such as pulse velocity and rebound velocity, as well as being reliable in receiving low impact energy data for the same type of weight. Thus, control is provided for the integration of the equation of motion and the absorbed energy can be accurately measured. In addition, these velocity measurements (when the acceleration of the tube is about zero) are taken under the pendulum pendulum. Thus, it is ensured that more reliable data are obtained from the velocity measurements obtained from the weight reduction impact systems. The schematic picture of the Pendulum impact test system is as shown in Figure 5.

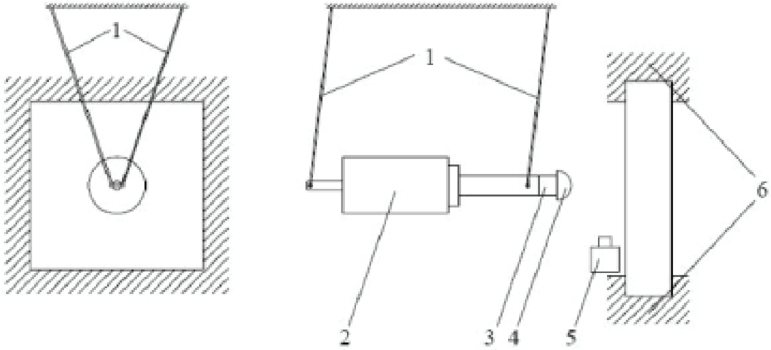


Figure 5. *Schematic picture of Pendulum impact test system (Aktaş, 2007)*

The names of the parts that make up the Pendulum impact test system are as follows; 1) four aramid strings on which the pendulum is hung, 2) a planar weight, 3) a load cell placed at the end of the pendulum to measure the force between weight and mass, 4) a spherical and hardened steel end placed at the end of the load cell, 5) before and after the impact a velocity sensor placed in front of the sample to measure velocity and 6) a rigid stabilizer used to keep the sample stable (Aktaş, 2007 and Herup, 1996).

2.2.3 Built-in Beam Impact Tests

The built-in beam impact test system is schematically shown in Figure 6. The embedded beam impact test is an uncommon low-velocity impact test system.

In this impact test system, when a 1 inch diameter steel sphere mounted on the end of a flexible beam is pulled and released, it hits the sample and the impact is generated (Aktaş, 2007).

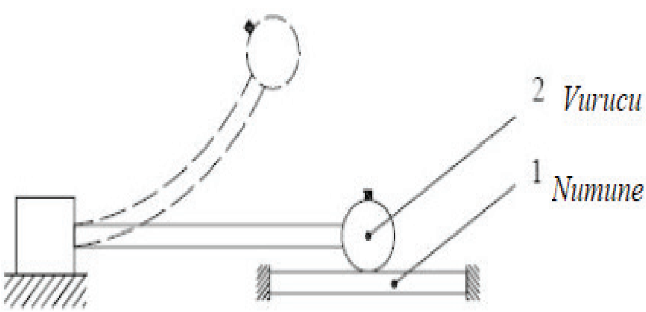


Figure 6. *Schematic view of the embedded beam impact test system (Aktaş, 2007)*

3. DAMAGES DUE TO IMPACT IN COMPOSITE MATERIALS

Damage types and models that occur as a result of impact tests on laminated composite materials are macroscopically examined.

3.1 Types of Macroscopic Damage

- Marking (Crushing Mark): It is the striker making a mark on the composite material and it occurs as a result of matrix crushing.
- Penetration: The striker is stuck in the composite material.
- Perforation: It is the case of the punching the composite material (İçten, 2006).

3.2 Types of Microscopic Damage

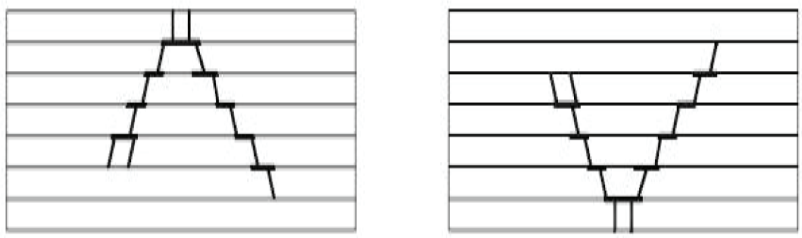
3.2.1 Matrix Breaking

The most common type of damage in laminated composite plates is matrix breakage. First, matrix breaks begin, then delamination (separation between layers) is observed (İçten, 2006).

Matrix breakage is mostly seen in brittle materials such as epoxy. Since the matrix has less deformation than the fiber, impact fracture will start in the matrix. The matrix fracture depends on the sheet fiber angles. The matrix fracture spreads parallel to the fiber angle in the damaged layer and its formation reduces the stiffness of the composite plate (Evci, 2010).

Matrix breakage, in thick composites; It occurs in the first layer struck by the striker due to high and local contact stresses. In this case, the damage is in the form of branches of the tree and progresses from the upper layer of the plate to the lower layer (Fig. 7 (a)). In thin composites, matrix breakage occurs in the lowest layer due to the bending of the plate (Fig. 7 (b)). In this case, the damage proceeds in the form of tree branches. However, the progress of the damage is from the lower layer to the upper layer this time (Karakuzu et al., 2007).

Matrix breakage, in thick composites; It occurs in the first layer struck by the striker due to high and local contact stresses. In this case, the damage is in the form of branches of the tree and progresses from the upper layer of the plate to the lower layer (Fig. 7 (a)). In thin composites, matrix breakage occurs in the lowest layer due to the bending of the plate (Fig. 7 (b)). In this case, the damage proceeds in the form of tree branches.



(a) *(b)*

Figure 7. Progression of matrix fracture damage for (a) thick and (b) thin plates (Karakuzu, 2007).

3.2.2 Delamination

Delaminations are damages that occur due to reduced adhesion between adjacent layers, which significantly reduce the strength of the layer. Experimental studies report that delamination does not occur between layers in the same layer group, but in the matrix-rich region between layers with different fiber orientation. Due to the different fiber orientation between the layers in the laminated composite material, the flexural stiffness of these layers differ. The most important cause of delamination; This bending stiffness difference between the layers is due to the bending shear stresses. (Kara, 2006)

Matrix breaks are a necessary factor in initiation of delamination. There is a tight relationship between matrix breakage and delamination. Delaminations occur in the interface area between the layers and in general the delamination is as seen in Figure 8 and the long part is in the direction of the fiber of the substrate. While the largest delamination occurs at the lowest interface, the delamination decreases as the upper interfaces are reached (Karakuzu et al., 2007).

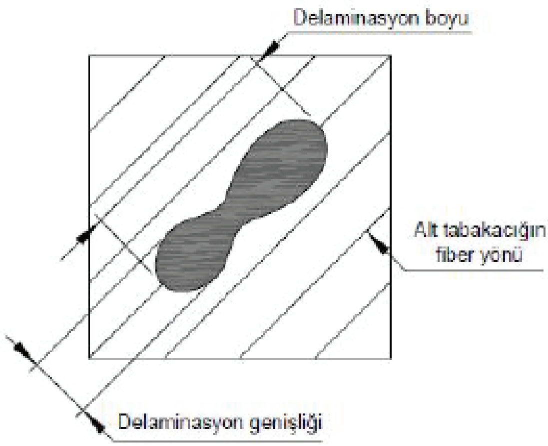


Figure 8. *A typical type of delamination (Karakuzu et al., 2007).*

3.2.3 Fiber Break

Fiber breakage, another type of damage, usually occurs much later than matrix breakage and delamination. The two most important factors causing fiber breakage are:

1) The effects of high local stress and penetration (mainly effects governed by shear forces). This event occurs on the surface just below the striking object.

2) High bending stresses. This event occurs on the surface that is not exposed to impact (Karaca, 2010).

The plate can only be completely damaged if all the fibers are broken. A sharp-tip striker will cause more fiber breakage than a blunt-ended hitter. When the fiber breakage begins, the load-bearing capacity of the plate is not completely exhausted. Because there are still fibers in the other layers of the plate that can carry the load. The complete exhaustion of the load carrying capacity is only possible when all the fibers of the plate are damaged (Evci, 2010).

4. IMPACT GRAPHICS

4.1 Force - displacement (collapse) graph

Force-displacement curves contain very important clues about the impact behavior of a composite material in an impact test. Figure 9 shows a sample graph showing the change of force versus displacement under increasing impact energy. However, there are two types of curves, open and closed curves. The closed curve includes a portion showing the load increment and a return portion containing loading and unloading. In addition, the inclination of the loading section, which increases in the force-displacement curve, expresses the bending stiffness of the composite material under impact loading. Depending on the impact energy, the falling part of the curve after the peak can be in three different ways;

• There may be rebound curve created by the impact tip rebound from the sample.

• There may be rebound after penetration.

• It can also be in the form of an open curve resulting in perforation.

If the material has a completely punctured damage, then the force vs collapse curve will be an open curve. If the impact energy is low, the force - collapse curve is in the form of a closed curve. In other words, force and collapse values decrease as a result of rebound and the curve comes to a point close to the starting point. With the increase of impact energy, the force rises to the maximum point, this point is called the peak force. It is observed that the amount of subsidence increases until the rebound occurs in the graphics in cases with partial damage and therefore partial rebound (Figure 9). As the impact energy increases, the damaged part on the material increases more and the rebound starts to decrease gradually. This situation can be seen in the penetration curve in Figure 9. When the impact energy increases, the curve turns into an open curve, the collapse reaches its maximum and no rebound is observed. This situation is seen in the perforation curve in Figure 9. In this graph, the area under the curves corresponds to the energy absorbed during the pulse duration. Here, after determining the absorbed energy values of each sample, the absorbed energy - time graph of the composite plate can be drawn (Sayer, 2009).

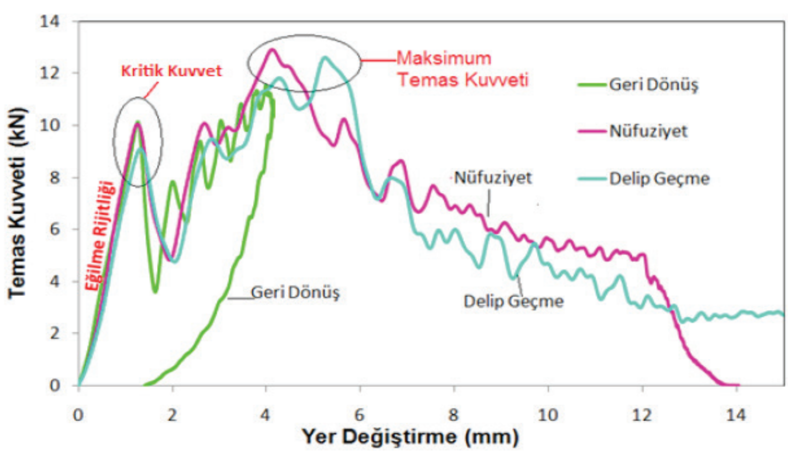


Figure 9. *Force - displacement (collapse) graph (Karaca, 2010).*

4.2 Time relationship of absorbed energy

Absorbed energy is calculated from the area under the force-collapse curve. In the event that the striker bounces back from the sample surface, not all of the energy of the striker can be absorbed by the sample, and in this case the impact energy is used to rebound the striker from the sample surface. In case of the striker stuck in the sample, all the impact energy of the striker is absorbed by the sample and the last part of the curve continues almost horizontally as seen in Figure 10.

In addition, the absorbed energy-time graph of three special cases such as rebounding from the sample surface, the striker tip penetrating the sample and the striker penetrating the sample are as shown in Figure 10. In the event that the batter pierces the specimen, the area under the friction part between the striker and the specimen is added to the amount of absorbed energy calculated by the program of the impact device, and the curve is directed upwards again in Figure 10. However, the area under the friction part must be subtracted from the total amount of absorbed energy obtained. In this case, the energy value obsorbed by the sample is the energy value that the dashed line approximately corresponds to, as seen in the perforation curve (Sayer, 2009).

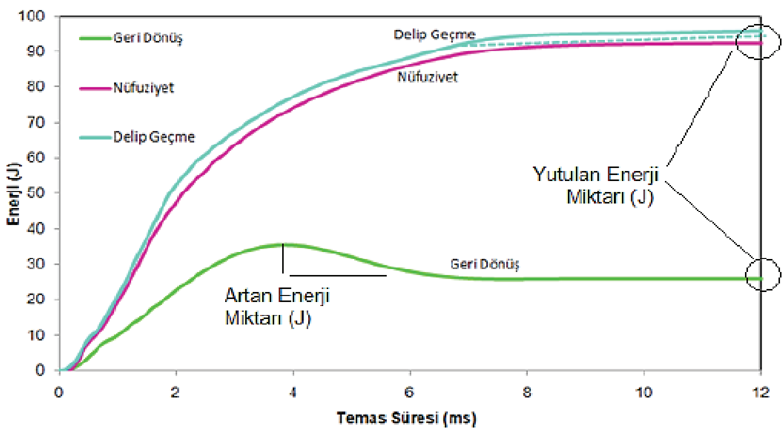


Figure 10. *The relationship between absorbed energy and time graph (Karaca, 2010).*

4.3 Velocity - time graph

When the velocity-time changes are examined, it is observed that the velocity is maximum at the beginning of time and the velocity decreases as the time passes. If the velocity is reset at a certain time (at the time of maximum collapse) and then rebound occurs, it takes a negative value because the velocity of the striker is opposite to the direction of impact (Figure 11). shows. In the case of piercing, the striker continues its movement linearly by piercing the sample since it cannot absorb all of the energy it has after hitting the sample (Ataş & Sayman, 2007).

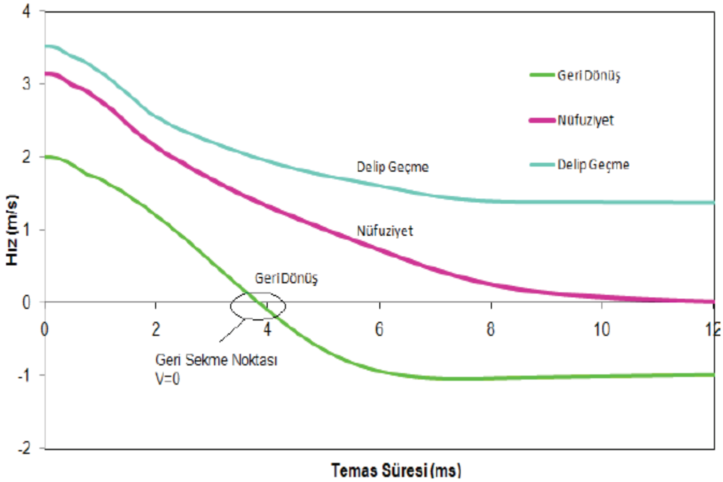


Figure 11. *Velocity - time graph (Karaca, 2010).*

The ratio of rebound velocity to impact velocity decreases as the impact velocity increases. However, as the impact velocity increases, damage to the material will increase and the hitter's rebound velocity will decrease. This ratio varies between 1 and 0 during the stroke. While it is around 1 when the impact velocity is low, it moves towards 0 when the impact velocity increases. If the impact behaves in an elastic behavior, this value becomes 1 (Ataş & Sayman, 2007).

5. CONCLUSION AND EVALUATION

Damage caused by impact in composite materials may occur on the surface not exposed to impact, depending on the type of impact. Low-velocity pulses are impacts with matrix breakage, delamination formation and fiber breakages in terms of damage. A part subjected to low-velocity impact but not pierced may exhibit lower mechanical properties than a part pierced under the same conditions. The reason for this is that most of the damage energy spent on the non-piercing material is spent for delamination, which weakens the internal structure of the material. The force-displacement and absorbed energy-time changes obtained from the impact tests show that the amount of energy absorbed by the sample increases with the increase of impact energy. Bending stiffness is expressed in the force-displacement curve as the slope in the increasing part of the curve. The fluctuations and changes in bending stiffness indicate the damage to the sample. It is seen that the greatest contact force increases with the increase of the impact velocity. As a result of the impact tests, it is seen that the damage on the material increases with the increasing impact velocity, and as a result, the rebound velocity of the striker decreases.

REFERENCES

1. Parlakyıldız S., Gençoğlu M.T., Cengiz M.S. 2018. Development of Rail Systems from Past to Present. International Conference on Multidisciplinary, Science, Engineering and Technology (2018 Dubai, BAE)

2. Parlakyıldız, Ş. Gencoglu, M.T., Cengiz, M.S. Analysıs of Faılure Detectıon and Vısıbılıty Crıterıa In Pantograph-Catenary Interactıon, Light & Engineering 28, 6, pp. 127–135, 2020.

3. Gencer, G. Eren, M. Yildirim, S. Kaynaklı, M. Palta, O. Cengiz, M.S., Cengiz Ç. Numerical Approach to City Road Lighting Standards, Imeset Book of Abstracts, Int. Conf. Mult. Sci. Eng. Tech., 12-14 July 2017 Baku.

4. Efe S. B., UPFC Based Real-Time Optimization of Power Systems for Dynamic Voltage Regulation. Computer Modeling in Engineering & Sciences, 2018. V116, #3, pp. 391–406.

5. Efe S. B., Cebeci M. Power Flow Analysis by Artificial Neural Network. International Journal of Energy and Power Engineering, 2013. V2, #6, pp. 204–208.

6. S Yıldırım, I Yapıcı, S Atiç, M Eren, O Palta, Ç Cengiz, MS Cengiz, Y Yurci. (2017) Numerical Analysis of Productivity and Redemption Periods in Illimunation, Imeset Book of Abstracts, Int. Mult. Sci. Eng. Tech., 1 (1), pp. 12-14.

7. Cengiz, Ç., Kaynaklı, M., Gencer, G., Eren, M., Yapici, İ., Yildirim, S., Cengiz, MS. (2017) Selection Criteria and Economic Analysis, International Conference on Multidisciplinary, Science, Engineering and Technology Bitlis Book of Abstracts,1 (1), pp. 27-29.

8. Abatan, A., Hu, H., & Olowokere, D., ‘‘Impact Resistance Modeling of Hybrid Laminated Composites’’, Journal of Thermoplastic Composite Materials, 11, 249-260, 1998.

9. Abrate, S., “Impact on composite structures”, Cambridge University Press, 135 (1998).

10. Adams, R.D, 2009: *“Adhesive Bonding”*, CRC Press. Woodhead Publishing Limited, 143-187.

11. Aktaş, M., 2007. Hibrit Kompozit Plakaların Darbe Davranışına Sıcaklığın Etkisi. Dokuz Eylül Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi, 136s, İzmir.

12. Aslan, Z. & Karakuzu, R., ‘‘Transient Dynamic Analysis of Laminated Composite Plate Subjected to Low-Velocity Impact’’, Mathematical & Computational Applications, Vol. 7, No.1, 73-82, 2002.

13. Aslan, Z., Karakuzu, R. & Okutan, B., ‘‘The Response of Laminated Composite Plates under Low-Velocity Impact Loading’’, Composite Structures, 59, 119-127, 2003.

14. Atas, C., & Sayman, O., ‘‘An Overall View on İmpact Response of Woven Fabric Composite Plates’’, Composite Structures, 82: 336-345, 2007.

15. Ceyhun, V. ve Turan, M., (2003). Tabakalı Kompozit Malzemelerin Darbe Davranışı, Mühendis ve Makine, 516: 19-25.

16. Esendemir, Ü., & Caner, A.Y.,(2018). Tabakalı Kompozit Malzemelerin Darbe Davranışının Deneysel Olarak İncelenmesi. Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi, vol.22, 207-215.

17. Evci, C., “Seramik Kompozit Zırh Sistemlerinin Darbe ve Balistik Özelliklerinin Belirlenmesi”, Doktora Tezi, Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Ankara, 67 (2010).

18. Gong, S.W., & Lam, K.Y., ‘‘Transient Response of Stiffened Composite Plates Subjected to Low-velocity Impact’’, Composites, Part B, 30, 473-484, 1999.

19. Goo, N.S. & Kim, S.J., ‘‘Dynamic Contact Analysis of Laminated Composite Plates under Low-velocity Impact’’, AIAA Journal, 35, 1518-1521, 1997.

20. Hosseinzadeh, R., Shokrieh, M.M. & Lessard, L., ‘‘Damage Behavior of Fiber Reinforced Composite Plates Subjected to Drop Weight Impacts’’, Composite Science and Technology, 66, 61-68, 2006.

21. İçten, B.M., “Damage in Laminated Composite Plates Subjected to Low Velecity Impact”, Doktora Tezi, Dokuz Eylül Üniversitesi Fen Bilimleri Enstitüsü, İzmir, 12-16, 32 (2006).

22. Kara, M., ‘‘Düşük Hızlı Darbeye Maruz Tabakalı Kompozitlerin Dinamik Cevabı’’, Yüksek Lisans Tezi, Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya, 2006.

23. Karaca, D., 2010: Farklı Kalınlık ve Oryantasyon Yüzdelerindeki Tabakalı Kompozit Levhaların Düşük Hızlı Darbe Enerjileri Altında Davranışlarının Deneysel İncelenmesi, Yüksek Lisans Tezi, Gazi Üniversitesi, Fen Bilimleri Enstitüsü.

24. Karakuzu, R., İçten, B.M., Aktaş, M., Toparlı, M., Ataş, C., Arman, Y., ‘‘Kompozit Plakalarda Sıcaklığın Darbe Davranışına Etkisi’’, TÜBİTAK, Proje No; 104M426, İzmir, 6-8, 23-26, 2007.

25. Kim, J.K. & Kang, K.W., ‘‘An Analysis of Impact Force in Plain-weave Glass-epoxy Composite Plates Subjected to Transverse Impact’’, Composite Science and Technology, 61, 135-143, 2001.

26. Lee, Y.S., Kang, K.H., & Park, O., ‘‘Response of Hybrid Laminated Composite Plates under Low-velocity Impact’’, Computers & Structures, 65, 965-974, 1997.

27. Mili, F., & Necip, B., ‘‘Impact Behavior of Cross-Ply Laminated Composite Plates under Low Velocities’’, Composite Structures, 51, 237-224, 2001.

28. Ramkumar, R.L., & Chen, P.C., ‘‘Low-velocity Impact Response of Laminated Plates’’, AIAA Journal, 21, 1448-1452, 1982.

29. Sayer, M., 2009: “Hibrit kompozitlerin darbe davranışlarının incelenmesi”, Doktora Tezi, Pamukkale Üniversitesi, Fen Bilimleri Enstitüsü.

30. Şenel, M., & Kurşun, A., ‘‘Investigation of the Effect of Low-Velocity Impact on Composite Plates with Preloading’’, Society for Experimental Mechanics, 2011.

31. Şenel, M., Ünal, R., Mumyakmaz, B., Karakoç, F., ‘‘Ön gerilmeli Kompozit Plakların Düşük Hızlı Darbe Yükü Etkisi Altında Davranışlarının İncelenmesi’’, TÜBİTAK, Proje No: 105M195, Kütahya, 2009.

32. Yüce, İ., 2007. Dairesel Kompozit Tabakalarda Düşük Hızlı Darbe Hasarının İncelenmesi. Yüksek Lisans Tezi, Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya.

Chapter 37

PREDICTING GOLD PRICES USING DEEP LEARNING BASED SENTIMENT ANALYSIS IN TWITTER DATA

Buket KAYA [1]

Abdullah GÜNAY [2]

1 Assist Prof. Buket KAYA, Dept. of Electronics and Automation, Fırat University, Elazığ.

2 Instructor Abdullah GÜNAY, Social Sciences Vocational School, Siirt University, Siirt

INTRODUCTION

The rapid development of communication technologies and the use of these technologies by individuals is increasing day by day. Social media tools, in which individuals can easily express their feelings and thoughts in recent years, are also one of the most important mass media tools. Twitter is one of the most used social media platforms to convey direct thoughts. In this platform, where individuals share their feelings and thoughts or by quoting others' messages, interaction generally takes place through daily events. With about 13.4 million according to the research Staista.com have the most Twitter users in the world is Turkey 6. (staista.com, 2020). The number of Twitter users in Turkey, where all events are so intense and developing countries worldwide can be evaluated by users. With the method called hashtag, individuals, institutions or events are tagged and interaction is provided by users over time. This creates a great opportunity to measure the relationship between tagged and shared events with the analysis of individuals' feelings and thoughts. users about the ever-changing gold prices in recent months in Turkey are sharing their thoughts and feelings through social media platforms. In this context, the present study analyzed the price of gold in Turkey with emotion contained in the discarded tweet about it aimed to investigate whether there is a relationship.

On Twitter, which has become one of the most popular social network platforms today, people can talk about their daily lives, different social / national / international issues, etc. they express their opinions about. They share their views in 140-character text and sometimes in audio / video files. These public posts, called Tweets, can be liked, commented on, or re-tweeted by other people. People can follow each other or be friends with each other on Twitter. Unlike most other social networking platforms, Twitter allows one-way connections; this means that one user can follow another user without the second user exchanging communication, and these interactions lead to a communication network. Within the scope of the study, 62146 tweets about "# gold" posted between 1 August 2020 and 31 August 2020 were obtained and the "positive", "neutral" or "negative" emotions contained in these tweets were analyzed.

Emotion Analysis can be defined as computational processing of ideas, emotions and subjectivity in texts framed within the field of natural language processing (NLP) (Pangand Lee, 2008). NLP algorithms include syntactic process creation (for example, symbol generation, sentence detection), information extraction (i.e., converting unstructured text into a structured form), capturing meaning (i.e. assigning a concept to a word or phrase), and defined language rules and related identifies relationships from non-natural language text through the use of domain knowledge (Koleck et al, 2019: 365). Languages that are more natural often make vague definitions, often vague and more inaccurate. Natural language processing includes applications such as machine translation, which involves models reading a sentence in one language and matching an equivalent sentence in another language. Many NLP implementations are based on language models that define probability distribution over words, characters or byte sequences in a natural language (Alpkoçak vd., 2020; Aygün vd., 2020; Müngen vd., 2020).

Sentiment Analysis is an automatic process of analyzing text data and sorting them according to positive, negative or neutral emotions. Using machine learning to do an Emotion Analysis on data from Twitter can help individuals understand how they talk about a topic (Manjappa and Kumar, 2020: 2).

Many academic studies have been conducted in recent years for sentiment analysis of Twitter data. Karimi and Shahrabadi (2019) analyzed deep learning algorithms on Persian texts in their study. In their studies, they observed that BERT works better than grams and LSTM or CNN. Bansal and Srivastava (2019) combined N-gram features to estimate voting shares in the 2017 Uttar Pradesh (UP) legislative election. They also found that the emotion distribution of emoji tweets differed significantly from emoji-free tweets, and they compared the emotion size-based criteria and the polarity of the tweets. The results showed that the proposed method to include N-gram features and emoji emotions significantly reduced the prediction error. Lu et al. (2020), first process the data they obtain through the BERT model. They then simulated linguistic functions in the sentence by combining Chinese grammar rules in the form of constraints with the Bi-gated recurrent neural network (GRU) and standardizing the output of adjacent positions. Use mobile phone review data in e-commerce in the study of Cai, Cao, and Ma (2020) The simulation experiment is performed on set and the experimental results show that the AT-BERT model greatly improves the performance of emotion classification. Chen, Yuan, and Huang (2020) made use of BERT, a pre-trained language model based on designing a unified framework to integrate data from emotion dictionaries. Evaluating using a laptop and restaurant datasets showed that it consistently outperformed the cutting-edge methods across all datasets.

METHODOLOGY

With the development of the Internet, text-based data from the web has grown exponentially where data carries large amounts of valuable information. Directly processing and analyzing text data has become possible with Natural Language Processing (NLP) techniques. Computer programs often operate on specialized languages designed to provide efficient and precise parsing with simple programs. Languages that are more natural often make vague definitions, often vague and more inaccurate. Natural language processing includes applications such as machine translation, which involves models reading a sentence in one language and matching an equivalent sentence in another language. Many NLP implementations are based on language models that define probability distribution over words, characters or byte strings in a natural language. (Goodfellow, Bengio and Courville, 2016).

Deep learning technology makes a significant breakthrough in text classification tasks, and it is possible to say that many researchers use deep learning techniques to solve emotion analysis problems. Deep learning is a branch of machine learning inspired by artificial neural networks that offer ways to learn data representations in a controlled and uncontrolled way with the help of layer hierarchy that allows processing (Habimana et al., 2020: 5). Nonlinear processing units are needed in deep learning for feature extraction and conversation. These processing units consist of multiple layers and their number can be changed as parameters. Sequential layers provide input to each other sequentially. Each layer is fed with the output of the previous layer. Algorithms can have a supervised or unsupervised working structure. In deep learning, there is a structure based on learning more than one feature level or representation of the data. Top-level properties form a hierarchical representation derived from lower-level properties. In this way, the algorithm finds the opportunity to learn multiple representation levels at different abstraction levels. (Şeker vd., 2017:48).

LSTM

LSTM is a special type of RNN designed to solve the loss / detonation problem faced by RNNs. LSTMs, like other RNN types, generate their output according to the input from the current time step and the output of the previous time step and send the current output to the next time step. Each LSTM unit consists of a memory cell and 3 non-linear gates. Memory cells (ct) keep the current state in arbitrary time intervals. Gates are in the form of an input gate (it), a forget gate (ft), an exit gate (ot), respectively. These gates are designed to regulate the flow of information entering and leaving the memory cell (Basiri et al., 2020).

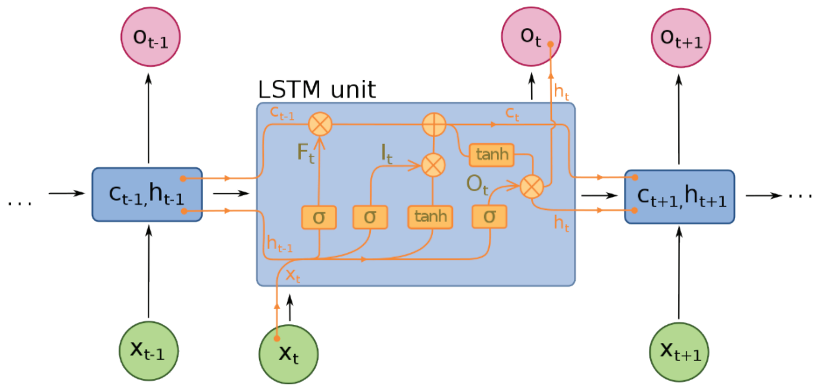


Figure 1. LSTM Unit Courtesy of François Deloche From Wikipedia.

GRU

GRU is an algorithm that aims to improve performance by optimizing the LSTM network structure. It is also called customized LSTM, as it adheres to the LSTM architecture as a structure. Its difference from LSTM is that it uses two gates instead of three. These gates are called update gate and reset gate. Update gate controls how much information from the previous level can be transferred. The reset gate controls the rate at which the previous level information is ignored. Figure 7 shows the working mechanism of the GRU. Using a fixed number of parameters for all models in some datasets, GRU can perform LSTM units in terms of both convergence in CPU time and parameter updates and generalization (Li et al, 2020).

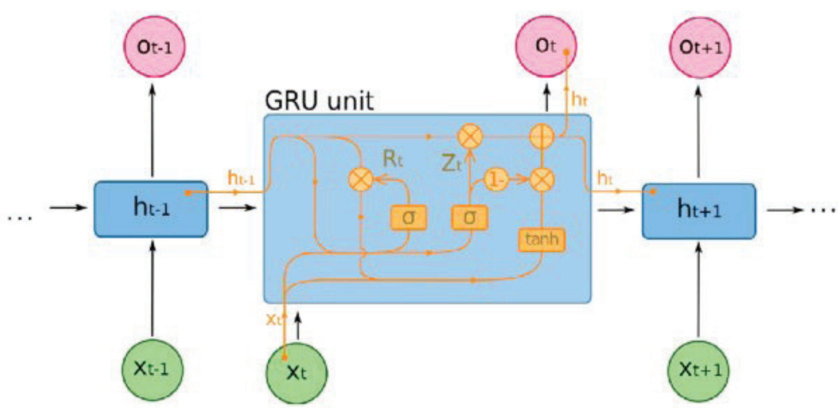


Figure 2. GRU Unit Courtesy Of François Deloche From Wikipedia.

BERT

BERT (Two-Way Encoder Demonstrations from Transformers) is a recently proposed language representation model for achieving text insertions. Pre-trained, deep bi-directional representations on untagged text are provided for BERT, masked word prediction and subsequent sentence prediction tasks. For classification tasks, a special icon is placed at the beginning of the text, and the icon's output vector is designed to correspond to the final text insertion. The pre-trained BERT model has proven very useful for transference learning in multi-class and binary text classification. The fine tuning of the model followed by an additional feed-forward layer and softmax activation function has been shown to be sufficient to provide state-of-the-art results in a downstream task (Yarullin and Serdyukov, 2019:2).

BERT, on the other hand, has a deep bidirectional contextualization that combines representations of both left context and right context models. Its highlight from other algorithms is that it is a multi-layer bidirectional Transformer encoder. Transformer architecture reduces sequential computing with multiple self-interest heads. To compute a representation of an input sequence, the self-attention mechanism correlates different positions of the sequence. Multi-headed attention allows the model to jointly participate in information from different representation subfields at different locations. The original Transformer has an encoder-decoder structure that maps an input sequence used by the decoder to a continuous array of representations to generate an output sequence one element at a time by the decoder. The encoder and decoder each consist of 6 identical layers and contain two sub-layers each of 8 parallel self-attention headers and a fully connected feed forward neural network. In contrast, the BERT model only retains the encoder part of the original model without any decoder (Myagmar et al., 2019:137).

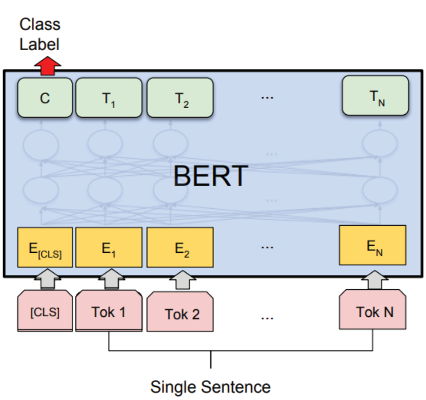


Figure 3. BERT Encoder

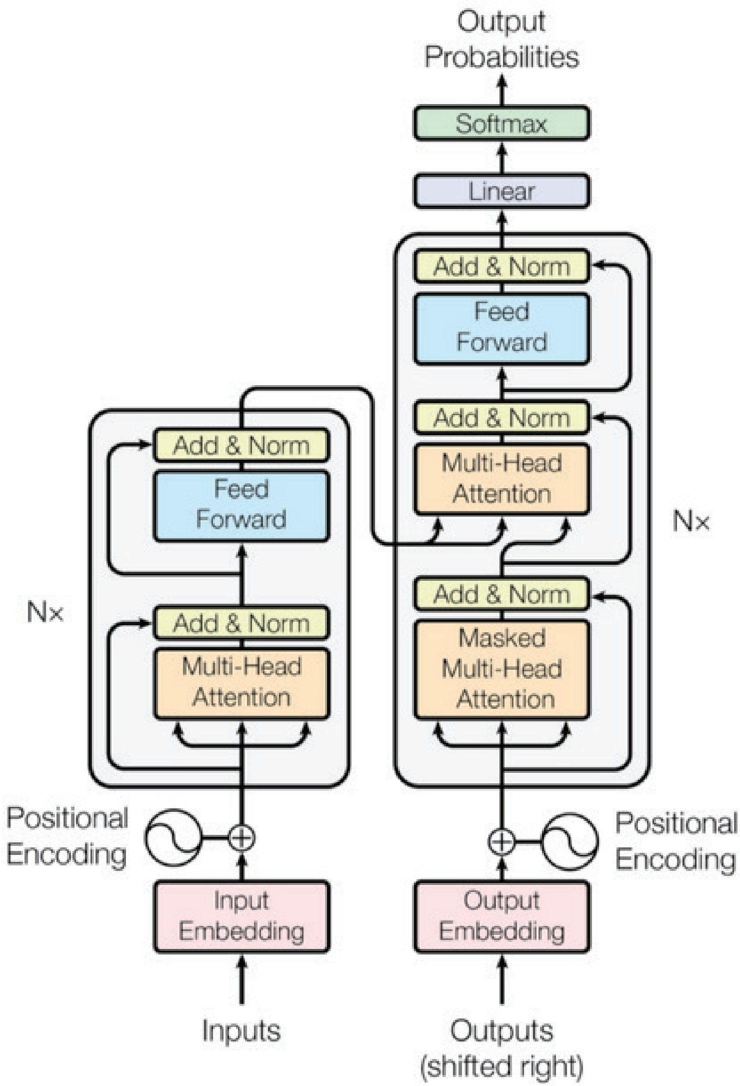


Figure 4. Orginal Transformer Architecture (Vaswani et al., 2017)

EXPERIMENTAL RESULTS

Within the scope of the study, the data obtained from Twitter can be predicted well and has been pretreated to make it into an analyzable form;

- Case converting
- Number filtering
- N Chars filtering

- Stop word filtering
- Tag filtering
- Removing punctuations

Figure 5. Preprocessing Dataset by Using KNIME

Also, word frequencies were calculated by dividing the sentences in the preprocessed data into their items. The most frequently used words in the data set are shown in the word cloud in figure 5.

Figure 6. Word Cloud

Keras neural network library and Tensorflow open source deep learning libraries were called in the program created using Python language in order to analyze the emotional values of the tweets. Afterwards, the training of the machine was tested with the training and test data sets previously created within the scope of the study. The results obtained are given in Table 1.

$$Precision = \frac{TP}{TP+FP} \tag{1}$$

$$Recall = \frac{TP}{TP+FN} \tag{2}$$

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{3}$$

$$F_1\ Score = 2 * \frac{Precision * Recall}{Precision+Recall} \tag{4}$$

Tablo1. Performance Metrics

Performance Metrics	GRU	LSTM	BERT
Accuracy	0,788	0,776	0,789
Precision	0,79	0,78	0,77
Recall	0,93	0,92	0,76
F1 Score	0,85	0,84	0,77
AUC	0,88	0,88	0.88

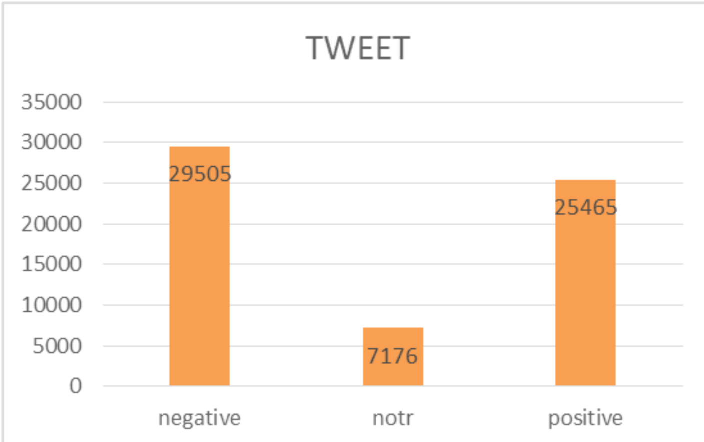


Figure 7. The number of tweets posted according to their mood

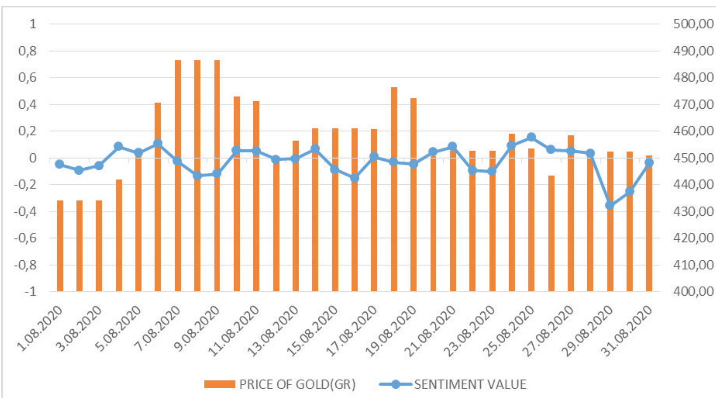


Figure 8. Comparison of Daily Gold Prices and Emotion Scores

As a result of the sentiment analysis, Figure 8 contains the emotion values of the comments made by users on the twitter about the daily gold prices and gold prices. Although there is not a significant relationship, it is seen that the average sentiment of the comments on that day changed in a positive or negative direction during the periods when gold prices suddenly rose and fell.

CONCLUSIONS

Sharing the feelings and thoughts of individuals on social media provides a great source of data for researchers who observe society and individuals. The expressions in the obtained data set were subjected to various text pre-processing techniques, and the emotional value of each tweet was calculated. The relationship between emotions and gold prices was observed using the LSTM, GRU and BERT algorithms. It has been determined that the rise and fall in gold prices are related to the sentiment score. When negative comments are made, gold prices rise, when positive and positive comments are made, they fall or remain stable.

REFERENCES

Alpkoçak, A., Tocoğlu, M. A., Çelikten, A., & Aygün, İ. 2019. Türkçe Metinlerde Duygu Analizi için Farklı Makine Öğrenmesi Yöntemlerinin Karşılaştırılması. *Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi*, *21*(63), 719-725.

Aygün, İ., Kaya, M., & Alhajj, R. 2020. Identifying side effects of commonly used drugs in the treatment of Covid 19. *Scientific Reports*, *10*(1), 1-14.

“Leading countries based on number of Twitter users as of October 2020”, https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/ , Erişim Tarihi : 18.12.2020.

Bansal, B., & Srivastava, S. (2019). Lexicon-based Twitter sentiment analysis for vote share prediction using emoji and N-gram features. *International Journal of Web Based Communities*, *15*(1), 85-99.

Basiri, M. E., Nemati, S., Abdar, M., Cambria, E., & Acharrya, U. R. 2020. ABCDM: An Attention-based Bidirectional CNN-RNN Deep Model for sentiment analysis. Future Generation Computer Systems. doi:10.1016/j.future.2020.08.005

Cai, X., Cao, H., Ma, J. 2020. Sentiment analysis of E-commerce comments based on BERT. International Journal of Science 7(6):239:244.

Chen, F., Yuan, Z., Huang, Y. 2020. Multi-source data fusion for aspect-level sentiment classification. Knowledge-Based Systems 187, 104831. doi : 0.1016/j.knosys.2019.07.002

Goodfellow, I. J., Bengio, Y., Courville, A. C. 2016. Deep Learning (Adaptive Computation and Machine Learning series) Illustrated Edition, MIT Press

Karimi, S., Shahrabadi, F. S. 2019. Sentiment analysis using BERT (pre-training language representations) and Deep Learning on Persian texts. , Iran University of Science and Technology Deep Learning.

Koleck, T. A., Dreisbach, C., Bourne, P. E., & Bakken, S. (2019). Natural language processing of symptoms documented in free-text narratives of electronic health records: a systematic review. Journal of the American Medical Informatics Association, 26(4), 364-379.

Li, W., Wu, H., Zhu, N., Jiang, Y., Tan, J., Guo, Y. 2020. Prediction of Dissolved Oxygen in a Fishery Pond Based on Gated Recurrent Unit (GRU). Information Processing in Agriculture. doi:10.1016/j.inpa.2020.02.002

Lu, Q., Zhu, Z., Xu, F., Zhang, D., Wu, W., Guo, Q. 2020. Bi-GRU sentiment classification for Chinese based on grammar rules and BERT. International Journal of Computational Intelligence Systems, 13(1): 538-548. doi: 10.2991/ijcis.d.200423.001

Manjappa, Rajani Shree and Kumar, Aditya, Twitter Sentiment Analysis (June 11, 2020). Available at SSRN: https://ssrn.com/abstract=3633673 or http://dx.doi.org/10.2139/ssrn.3633673

Müngen, A., Aygün, İ, Kaya, M. 2020. Finding the Relationship Between News and Social Media Users' Emotions in the COVID-19 Process, Sakarya University Journal of Computer and Information Sciences, 3(3).

Pang, B., and Lee, L. 2008 (January). Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval2(1–2): 1–13

Habimana, O., Li, Y., Li, R., Gu, X., Yu, G. (2020). Sentiment analysis using deep learning approaches: an overview. *Science China Information Sciences*, *63*(1), 1-36.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*, 5998-6008.

Şeker, A., Diri, B., & Balık, H. H. (2017). Derin öğrenme yöntemleri ve uygulamaları hakkında bir inceleme. *Gazi Mühendislik Bilimleri Dergisi*, *3*(3), 47-64.

Myagmar, B., Li, J., Kimura, S. (2019). Transferable High-Level Representations of BERT for Cross-Domain Sentiment Classification. In *Proceedings on the International Conference on Artificial Intelligence (ICAI)* (pp. 135-141). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).

Chapter 38

ULTRAVOILET (UV-C) UTILIZATION AND NUMERICAL APPLICATION RESEARCH TO REDUCE THE RISK OF COVID-19 TRANSMISSION

Mehmet Sait CENGIZ[1]

1 Bitlis Eren University Technical Sciences Vocational School, 13000, Bitlis msaitcengiz@gmail.com

1. INTRODUCTION

The quality of the air we breathe, the water we drink, and the cleanliness of our surfaces have a profound effect on our health and well-being. Especially in busy public places such as offices, factories, shops, bars or restaurants, schools, museums, and public transportation, there is a risk of catching and spreading viruses and bacteria. This is explained in the EPA Report, "Strengthening Study for Increased Protection Against Airborne Chemical and Biological Emissions.

UV-C radiation is a disinfectant known to help reduce the risk of infection for air, surfaces, objects, and water. All bacteria and viruses tested to date (hundreds of years over the years, including various coronaviruses) respond to UV-C disinfection. The Flow (UV Dose) Required to Achieve Bacteria, Protozoa, Virus and Algae Incremental Recording Inactivation has been reviewed, updated and expanded by Adel Haji Malayeri, Madjid Mohseni, Bill Cairns and James R. Bolton. Previous contributions were made by Gabriel Chevrefils (2006) and Eric Caron (2006). Supervised by Benoit Barbeau, Harold Wright (1999) and Karl G. Linden.

In laboratory tests, our UV-C light sources are known to inactivate 99% of the SARS-CoV-2 virus on a surface with an exposure time of 6 seconds. UV-C can be crucial in a strategy to protect people. Boston University's laboratory studies through NEIDL confirm the importance of UV-C.

Contamination was achieved by contracting the virus that caused the Covid-19 disease. SARS-CoV-2 (COVID-Signify UV-C light sources were exposed to a dose of 5 mJ/cm^2 UV-C on a surface. The duration of 6 seconds was chosen for this effect. SARS-CoV-2 virus UV-C In this study, it was determined that with a dose of 22 mJ/cm^2. UV-C reduced the SARS-CoV-2 virus on this surface by 99.9999%. In this study, the virus transmission time was 25 seconds.

This makes it important for top air systems that disinfect the air with UV-C lamps in various applications, and open UV-C systems that disinfect rooms and floors during the night (or when not in use).

The coronavirus disease (COVID-19) pandemic continues to seek solutions around the world to control or reduce the spread of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) responsible for the disease.

Covid-19 is usually transmitted from person to person with large respiratory droplets directly or by touching infected surfaces and then physically touching the eyes, nose, or mouth. More importantly, there is increasing evidence of airborne virus transmission (van Doremalen, 2020).

The use of UV-C radiation is an important method of reducing both contact spread and airborne transmission of infectious agents (such as bacteria and viruses). It is within the UV-C range (200 nm-280 nm) and has been used successfully and safely for a long time. However, UV-C should be applied informedly, with proper attention to dosage and safety. Improper application of UV-C can present human health and safety problems and cause inadequate deactivation of infectious agents.

With the development of technology, events that are limited in nature have begun to be imitated by people. For example, for the benefit of humanity, UV-C rays coming from the sun but not reaching the earth were imitated with LED lamps to produce solutions for viral diseases (Gencer et al, 2017; Yıldırım et al, 2017; Cengiz et al, 2017). Power transmission or electric vehicles for energy transfer have emerged as a result of imitation of nature. (Parlakyıldız, et al., 2018 and Parlakyıldız et al., 2020; Gencer et al., 2017; Efe, 2018; Efe and Cebeci, 2013).

2. DIFFERENT DISINFECTION METHODS

There are a wide variety of methods for disinfection, which we can describe as providing a more hygienic environment (water, air, or surface) by removing bacteria, germs, and viruses that are basically harmful to human health, or even permanent damage or death in some cases.

2.1 Disinfection with UVC lamp devices

The method, which we can also call killing bacteria and microbes with ultraviolet light in the wavelength range of 210-280 nanometers, is a reliable and effective method that has been used especially in the sterilization of operating rooms since the early 1900s. Although it is dangerous for humans when it comes into contact with the skin and eyes, it is preferred because of its successful results when necessary safety measures are taken.

It has long been discovered that UVC, the most energetic type of ultraviolet light, deforms the DNA of bacterial cells and the RNA of virus cells, preventing their reproduction. UV-C has been used in the disinfection of operating rooms and surgical instruments since the end of the 1800's.

2.2 Disinfection with Ozone

Ozone Gas, formed by oxygen atoms (O) weakly bonded to oxygen molecules (O_2), also has deadly effects on bacteria and viruses. This method is used extensively in the disinfection of wastewater, more generally in water disinfection. It is preferred because it does not affect the taste of water and its disinfectant effect is very strong.

2.3 Ionization

We wanted to include in this article where we describe all disinfection and sterilization methods, although it is a preferred method in large commercial enterprises. Ionization; is simply the process of removing electrons from atoms of a substance. In this way, even all the DNA of microorganisms is removed from their cells (even other electrons), and microorganisms are killed and neutralized. It is made with the ionizing property of Gamma rays with a wavelength (0.001 nanometers) even shorter than ultraviolet light in the light spectrum. Therefore, although the cost is a bit high, ionization with gamma rays is possible to obtain precise and effective disinfection.

2.4 Disinfection with HEPA Filters

In this method, HEPA filters filter all the particles larger than 0.3 microns in size and thus the air is cleaned. Although the HEPA method is effective in cleaning the air, it cannot provide complete hygiene or complete disinfection. This is because very small, microscopic creatures (smaller than 0.3 microns in size) are not attached to the filter and continue to roam in the air.

In addition, it has been observed that bacteria and viruses close to the filter are cleaned with HEPA filters, and their effect decreases as they move away from the filter. Therefore, HEPA filters may be sufficient for cleaning, but they cannot provide disinfection. At least, they cannot provide effective disinfection.

2.4 Disinfection with Chemical Substances

It is a method of cleaning viruses and bacteria using various chemicals and as you can guess, it is used for disinfection of surfaces, not water and air. The chemicals used for disinfection are various; alcohol, chlorine, formaldehyde, glutaraldehyde, orthophthalaldehyde, hydrogen peroxide, peracetic acid.

Chemical disinfectants used in hospitals due to their strength are formaldehyde, glutaraldehyde, hydrogen peroxide, orthophthalaldehyde, and peracetic acid. All of these chemicals have different potency and their dangers on human health are also variable. Some of them have a very high lethal effect on bacteria and viruses, while others are milder. A major disadvantage of chemical disinfection; leaving residue in the disinfected area and the risk of these residues damaging by contact with human skin. On the other hand, it is one of the least costly and easiest methods to apply.

3. UV RAYS

Ultraviolet radiation has biological effects. These effects are defined according to the spectrums, namely, UV-A, UV-B, UV-C, which are the basic 3 groups.While UV-A is in the wavelength range of 315-400 nm, UV-B emits radiation in the wavelength range of 280-315 nm and UV-C in the wavelength range of 100-280 nm.

The UV-C part of the ultraviolet spectrum carries the highest energy. When the ultraviolet light spectrum carries high energy, it easily damages microorganisms and viruses. So, UV-C is the most effective ultraviolet light. Spectral view for UV-C is shown in Figure1.

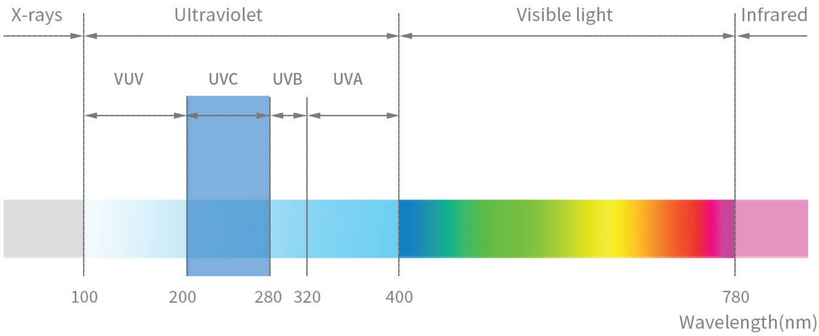


Figure 1. *Spectral view for UV-C*

In disinfection applications, UV-C has been used successfully in water disinfection for years. Moreover, UV-C disinfection is routinely included in air handling units to manage biofilm formation and disinfect the air (CIE, 2003). UV-C disinfection for ventilation systems is shown in Figure-2.

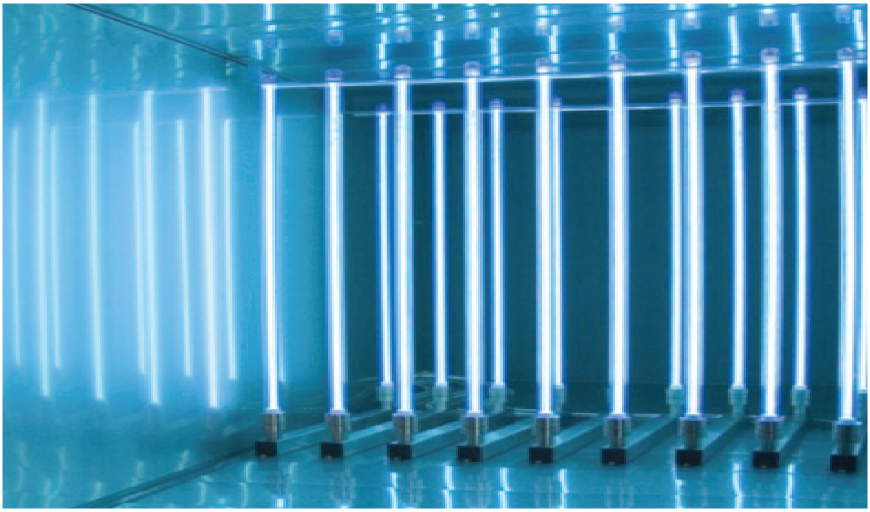


Figure 2. *UV-C disinfection for ventilation systems*

UV-C is used in disinfection processes in many areas today. For example, operating theaters or rooms are exposed to UV-C for disinfection overnight.

Also, UV-C is used to disinfect the air in closed areas. Water is sterilized by exposure to UV-C in water tanks.

Instead of putting a UV-C in a room to destroy viruses on the surface, in addition to this option, UV-C originated mobile robots are used today. There is increasing evidence that the use of UV-C in addition to standard manual cleaning in hospitals can be effective in practice. UV-C disinfection in rooms is shown in Figure 3.

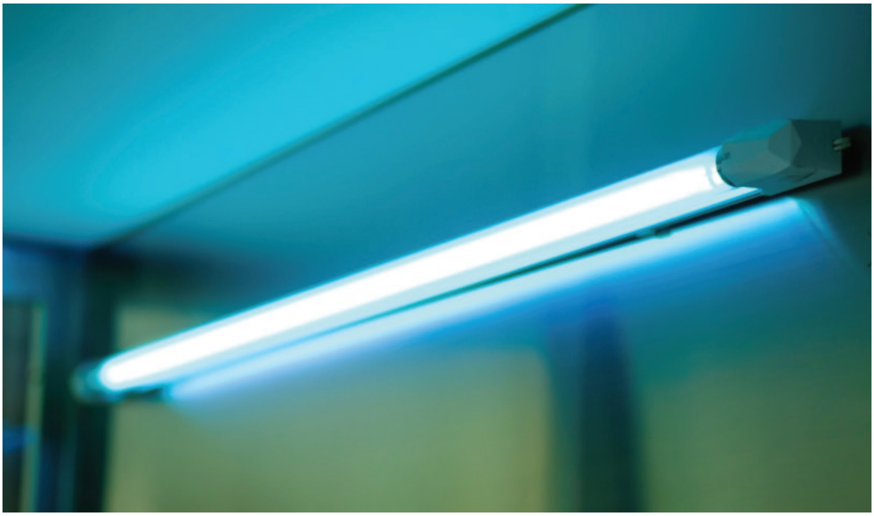


Figure 3. *UV-C disinfection in rooms*

In UV - C disinfection, light sources are located above the relevant areas. In terms of height, it is placed in an area higher than human height for rooms. In this way, the air in the relevant area is disinfected. In scientific research and studies conducted in the past years, UV-C has produced successful results to reduce the possibility of transmission of tuberculosis infectious disease to healthy individuals.

Mphaphlele's 2015 study and Escombe et al's 2009 study proves that the transmission of DHHS-2009 tuberculosis infectious disease with UV-C is reduced. In this context, the World Health Organization declared that UV-C is beneficial and should be used for both preventing tuberculosis infection and controlling the disease (WHO, 2019). UV-C disinfection in hospitals is shown in Figure 4. UV-C disinfection in storage areas can be seen in Figure 5.

Figure 4. *UV-C disinfection in hospitals*

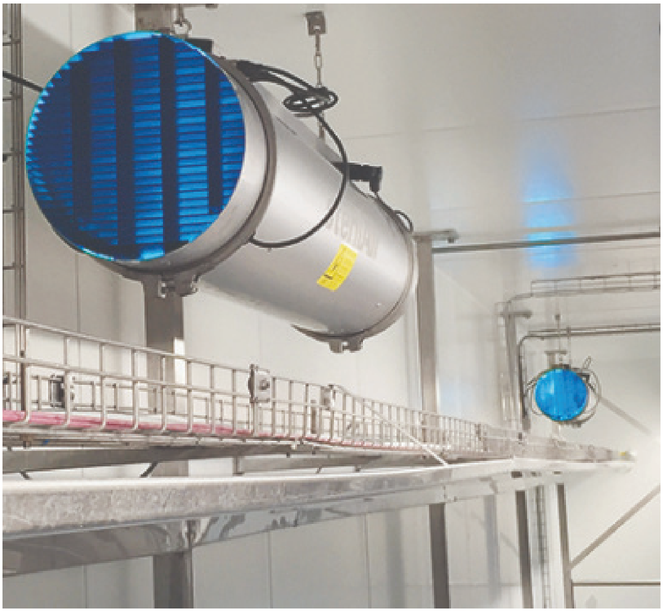


Figure 5. *UV-C disinfection in storage areas*

4. RISKS WHEN USING UV-C

Many people are not exposed to UV-C under normal conditions or are not exposed to UV-C during their lifetime. UV-C rays coming from the sun are absorbed by the atmosphere even at elevated altitudes, and these rays are prevented from reaching the earth. That is, the atmosphere does filtering (Piazena & Häder, 2009). Human exposure to UV-C is through artificially produced technological devices. UV-C effect is physical and has the ability to dry the skin of the human skin, ie the outer layer. Reaching the basal layer of the epidermis UV-C effect. In addition, it cannot affect the inner parts of the eye by remaining in the surface layer of the cornea of a sensitive organ such as the eye. The exposure of the eye to UV-C causes a sand effect on the eye. It causes photokeratitis, which is a very painful condition. A similar practice is seen in the iron welding process. Photokeratitis effects last for about a day, after which the painful process ends.

If human skin is exposed to intense UV-C, skin rashes such as sunburn (ISO / CIE, 2019). Generally, skin burn pain is less than the pain effect of UV-C rays on the eyes.

There are scientific studies that show that exposure of the skin to UV-C adversely affects the body's immune system (Gläser et al., 2009).

Ultraviolet radiation is thought to have a carcinogenic effect (ISO / CIE, 2016). However, there is no definitive evidence that UV-C alone causes cancer in humans. Technical Report CIE187: 2010 (CIE, 2010). The eye and skin effect of UV-C can be seen in Figure 6.

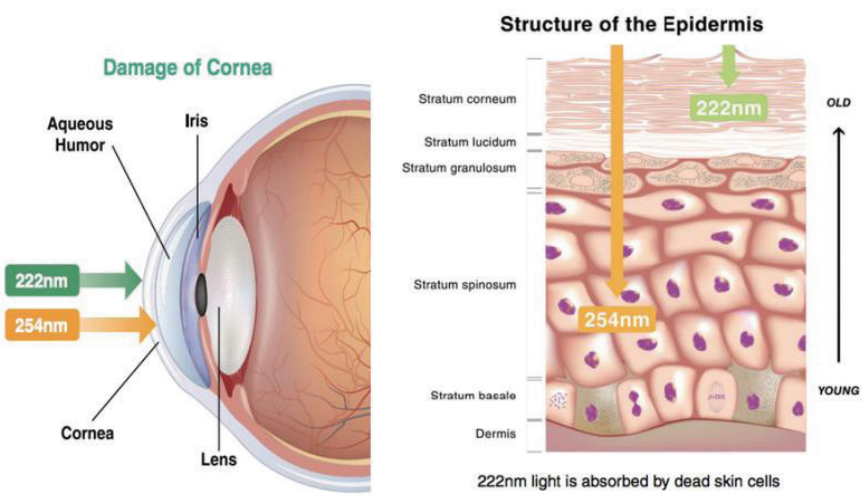


Figure 6. *The eye and skin effect of UV-C*

5. UV-C APPLICATION CALCULATION EXAMPLE

Although there are no clear conclusions about the COVID-19 (SARS-CoV-2) virus, which affects the world, it is known that UVC (Ultraviolet C) light at a dose of 10-20 mJ / cm2 is sufficient to destroy comparable viruses in the same SARS virus family. It is possible to deactivate 99.9 percent of bacteria and viruses with this application. UV-C effect depends on 3 important parameters. These;

• Ultraviolet dose,

• Light intensity

• Application time

depends. The time required to disinfect a surface is directly related to the intensity of the ultraviolet light source per unit area to be used. It is necessary to know the ultraviolet dose and light intensity to determine the application time. In Equation 1, the amount of dose to be applied is calculated.

Dose (mJ/cm^2) = Light intensity (mW/cm^2) x Duration (s) (1)

In Equation 2, the dose administration time is calculated.

Duration (s) = Dose (mJ /cm^2) / Light intensity (mW / cm^2) (2)

As shown in the formula above, the time required to disinfect a surface can be found by dividing the dose by the light intensity.

Light intensity is defined as the power per unit area per surface at a given distance. This value is inversely proportional to the square of the distance from the source to the surface. In other words, the further the light source is from the surface, the less effect it will be. The light intensity can be calculated by dividing the power by the square of the distance. In Equation 3, the amount of light intensity to be applied is calculated.

Light intensity (mW/cm^2) = Power (mW) / Distance^2 (cm^2) (3)

It should be noted that the calculation made with the help of the formula above is an assumption. The actual value may vary depending on various factors such as viewing angle, radiation pattern, and reflection.

Due to these differences, the luminous power value is given in mW instead of light intensity in mW/cm^2 in the data files of ultraviolet LEDs. Ultraviolet LEDs mostly have power in the range of 10-20 mW.

For example; When a UVC light source with a power of 3000 mW is kept at a distance of 300 cm from the surface, a light intensity of approximately 0.3333 mW / cm^2 is reached on the surface. In order to reach the minimum dose requirement of 10 mJ / cm^2, the application time should be at least 300 seconds, ie 5 minutes. In this case, it can be said that the surface is disinfected when the UV-C light is kept at a distance of 300 cm from the surface for 300 seconds (5 minutes).

In order to make correct inferences regarding the disinfection time, the light intensity on the surface must be measured with appropriate equipment. In this way, closer to reality dose rate or duration can be calculated.

6. CONCLUSION

In scientific studies, it has been observed that UV-C produces successful results in disinfection processes. However, the point to be considered here is that the application should be made according to ultraviolet dose, light intensity, and application time. A numerical application exemplifying this situation was made in this study and a solution proposal was made.

It has been concluded that the most effective and cheapest method in disinfection processes is UV-C. UV-Cs are widely preferred today because of their successful results when safety measures are taken, provided that they do not contact the skin and eyes.

Disinfection with UVC Lamp Devices is no longer a choice and has become a necessity. It can be recommended to be used at night when people are absent, especially in areas where people are present, such as schools, kindergartens, hotels, hostels, sports halls, cafes, restaurants, offices. Because UV-C prevents bacteria, viruses, and microbes from reproducing by disrupting the structure of DNA and RNA and causes them to die when they try to reproduce, while it can also harm a living creature.

At the beginning of these dangers; is the occurrence of burns if the beam comes into direct contact with the skin. In case of prolonged exposure to strong UV light, it may cause cancer. UV light also poses a risk to the eyes. If UV-C is in direct contact with the eye for a long time, UV rays will cause damage to the cornea.

REFERENCES

Adel Haji Malayeri, Madjid Mohseni, Bill Cairns ve James R. Bolton - Gabriel Chevrefils (2006) and Eric Caron (2006) - Benoit Barbeau, Harold Wright (1999) ve Karl G. Linden https://iuvanews.com/stories/92816/fluence-required-achieve-incremental-log-inactivation-bacteria -protozoa-viruses-algae.shtml

Arıkan, I. 2020. Ultraviyole web notları. https://www.fiyatimbu.com/blog/uvc-lambali-sterilizasyon-cihazlari-hakkinda-sikca-sorulan-sorular_411

BIPM (2019a) The International System of Units (SI), 9th Edition. https://www.bipm.org/utils/common/pdf/si-brochure/SI-Brochure-9-EN.pdf

BIPM (2019b) The International System of Units (SI), 9th Edition – Appendix 3: Units for photochemical and photobiological quantities. https://www.bipm.org/utils/common/pdf/si-brochure/SI-Brochure-9-App3- EN.pdf, accessed 2020–12–12.

Buonanno, M., Ponnaiya, B., Welch, D., Stanislauskas, M., Randers-Pehrson, G., Smilenov, L., Lowy, F.D., Owens, D.M. and Brenner, D.J. (2017) Germicidal Efficacy and Mammalian Skin Safety of 222-nm UV Light. Radiat Res 187(4): 483–491.

Cengiz, Ç., Kaynaklı, M., Gencer, G., Eren, M., Yapici, İ., Yildirim, S., Cengiz, MS. (2017) Selection Criteria and Economic Analysis, International Conference on Multidisciplinary, Science, Engineering and Technology Bitlis Book of Abstracts,1 (1), pp. 27-29.

CIE (2003) CIE155:2003 Ultraviolet Air Disinfection.

Freely available at http://cie.co.at/news/cie-releases-two-key-publications-uv-disinfection2

CIE (2006) CIE172:2006 UV protection and clothing.

CIE (2007) CIE181:2007 Hand protection by disposable gloves against occupational UV exposure.

CIE (2010) CIE187:2010 UV–C photocarcinogenesis risks from germicidal lamps.

Freely available at http://cie.co.at/news/cie-releases-two-key-publications-uv-disinfection2

CIE (2016) CIE220:2016 Characterization and Calibration Methods of UV Radiometers.

CIE/ICNIRP (2020) CIE/ICNIRP Online Tutorial on the Measurement of Optical Radiation and its Effects on Photobiological Systems, December 2, 2020. http://cie.co.at/news/cieicnirp-online-tutorial-measurement-optical-radiation-and-its-effects-photobiological-systems, accessed 2020–12–12.

DHHS (2009) Environmental Control for Tuberculosis: Basic Upper-Room Ultraviolet Germicidal Irradiation Guidelines for Healthcare Settings, DHHS (NIOSH) Publication Number 2009–105, https://www.cdc.gov/niosh/docs/2009–105/default.html, accessed 2020-12–2.

Efe S. B., UPFC Based Real-Time Optimization of Power Systems for Dynamic Voltage Regulation. Computer Modeling in Engineering & Sciences, 2018. V116, #3, pp. 391–406.

Efe S. B., Cebeci M. Power Flow Analysis by Artificial Neural Network. International Journal of Energy and Power Engineering, 2013. V2, #6, pp. 204–208.

EPA Report, https://www.epa.gov/clean-air-act-overview/progress-cleaning-air-and-improving-peoples-health

Escombe, A.R., Moore, D.A., Gilman, R.H., Navincopa, M., Ticona, E., Mitchell, B., Noakes, C., Martínez, C., Sheen, P., Ramirez, R., Quino, W., Gonzalez, A., Friedland, J.S., Evans, C.A. (2009) Upper-room ultraviolet light and negative air ionization to prevent tuberculosis transmission. PLoS Med. 6(3): e43. DOI: 10.1371/journal.pmed.1000043.

Gencer, G. Eren, M. Yildirim, S. Kaynaklı, M. Palta, O. Cengiz, M.S., Cengiz Ç. Numerical Approach to City Road Lighting Standards, Imeset Book of Abstracts, Int. Conf. Mult. Sci. Eng. Tech., 12-14 July 2017 Baku.

Gläser, R., Navid, F., Schuller, W., Jantschitsch, C., Harder, J., Schröder, J.M., Schwarz, A., Schwarz, T. (2009) UV-B radiation induces the expression of antimicrobial peptides in human keratinocytes in vitro and in vivo. Journal of Allergy and Clinical Immunology 123(5): 1117–1123.

ICNIRP (2004) ICNIRP Guidelines – On limits of exposure to ultraviolet radiation of wavelengths between 180 nm and 400 nm (incoherent optical radiation), Health Physics 87(2):171–186; 2004.

Limited free access until 2020–12–12. Available at http://www.icnirp.org

ICNIRP (2010) ICNIRP Statement – Protection of workers against ultraviolet radiation, Health Physics 99(1):66-87.

Jordan, W.S. (1961) The Mechanism of Spread of Asian Influenza, Am Rev Resp. 83, 2P2, 29–40.

Ko, G., First, M.W., Burge, H.A. (2000) Influence of relative humidity on particle size and UV sensitivity of Serratia marcescens and Mycobacterium bovis BCG aerosols. Tubercle and Lung Disease. Vol. 80, 4–5, 217–228.

Mphaphlele, M. (2015) Institutional Tuberculosis Transmission. Controlled Trial of Upper Room Ultraviolet Air Disinfection: A Basis for New Dosing Guidelines. Am J Respir Crit Care Med. 192(4): 477–84.

Narita, K., Asano, K., Morimoto, Y., Igarashi, T., Hamblin, M.R., Dai, T. and Nakane, A. (2018) Disinfection and healing effects of 222-nm UVC light on methicillin-resistant Staphylococcus aureus infection in mouse wounds. Journal of Photochemistry and Photobiology B: Biology 178: 10–18.

Parlakyıldız S., Gençoğlu M.T., Cengiz M.S. 2018. Development of Rail Systems from Past to Present. International Conference on Multidisciplinary, Science, Engineering and Technology (2018 Dubai, BAE)

Parlakyıldız, Ş. Gencoglu, M.T., Cengiz, M.S. Analysıs of Faılure Detectıon and Vısıbılıty Crıterıa In Pantograph-Catenary Interactıon, Light & Engineering 28, 6, pp. 127–135, 2020

Peccia, J., Werth, H.M., Miller, S., Hernandez, M. (2001) Effects of Relative Humidity on the Ultraviolet Induced Inactivation of Airborne Bacteria, Aerosol Science and Technology. 35, 3.

Philips, 2020 https://www.lighting.philips.com.tr/urunler/uv-c#footnote

Piazena, H. and Häder, D.-P. (2009) Solar UV-B and UV-A irradiance in arid high-mountain regions: Measurements on the island of Tenerife as compared to previous tropical Andes data. Journal of Geophysical Research: Biogeosciences. 114(G4)

Sagripanti, J.-L. and Lytle, C.D. (2011) Sensitivity to ultraviolet radiation of Lassa, vaccinia, and Ebola viruses dried on surfaces. Archives of Virology 156(3): 489–494.

Taylor, W., Camilleri, E., Craft, D.L., Korza, G., Granados, M.R., Peterson, J., Szczpaniak, R., Weller, S.K., Moeller, R., Douki, T., Mok, W.W.K. and Setlow, P. (2020) DNA Damage Kills Bacterial Spores and Cells Exposed to 222-Nanometer UV Radiation. Applied and Environmental Microbiology 86(8): e03039–03019.

Tomas, M.E., Cadnum, J.L., Jencson, A., Donskey, C.J. (2015) The Ebola disinfection booth: evaluation of an enclosed ultraviolet light booth for disinfection of contaminated personal protective equipment prior to removal. Infect Control Hosp Epidemiol. 36(10):1226–1228.

Van Doremalen, N., Bushmaker, T., Morris, D.H., Holbrook, M.G., Gamble, A., Williamson, B.N., Tamin, A., Harcourt, J.L., Thornburg, N.J., Gerber, S.I., Lloyd-Smith, J.O., de Wit, E., Munster, V.J. (2020) Aerosol and Surface Stability of SARS-CoV-2 as Compared with SARS- CoV-1. N Engl J Med. 382: 1564–1567.

Welch, D., Buonanno, M., Grilj, V., Shuryak, I., Crickmore, C., Bigelow, A.W., Randers- Pehrson, G., Johnson, G.W. and Brenner, D.J. (2018) Far-UVC light: A new tool to control the spread of airborne-mediated microbial diseases. Scientific Reports 8(1): 2752.

WHO (2019) WHO guidelines on tuberculosis infection prevention and control. 2019 update. Geneva: World Health Organization.

WHO (2020) https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public/myth-busters, accessed 2020–12–12.

Yamano, N., Kunisada, M., Kaidzu, S., Sugihara, K., Nishiaki-Sawada, A., Ohashi, H., Yoshioka, A., Igarashi, T., Ohira, A., Tanito, M. and Nishigori, C. (2020) Long-term effects of 222 nm ultraviolet radiation C sterilizing lamps on mice susceptible to ultraviolet radiation.

Yıldırım, S., Yapıcı, I., Atiç, S.,Eren, M.,Palta, O., Cengiz, Ç., Cengiz, MS., Yurci., Y. (2017) Numerical Analysis of Productivity and Redemption Periods in Illimunation, Imeset Book of Abstracts, Int. Mult. Sci. Eng. Tech., 1 (1), pp. 12-14.

Chapter 39

ULTRAVIOLET LIGHT TECHNOLOGY

ANALYSIS AND AREA CALCULATIONS

Mehmet Sait CENGIZ[1]

1 Bitlis Eren University Technical Sciences Vocational School, 13000, Bitlis, msaitcengiz@gmail.com

1. INTRODUCTION

With the development of technology, limited events in nature have begun to be imitated by people. The main purpose here is for human comfort. Accordingly, efficient and economical solutions have been produced. For example, UV-C rays coming from the sun but not reaching the world for the benefit of humanity were imitated with LED lamps to produce solutions for viral diseases (Gencer et al., 2017; Yıldırım et al., 2017; Cengiz et al., 2017). Power transmission or electric (rail systems) vehicles for energy transfer have emerged as a result of imitation of nature (Parlakyıldız et al., 2018 and Parlakyıldız et al., 2020; Gencer et al., 2017; Efe, 2018; Efe & Cebeci, 2013).

Ultraviolet (UV) refers to the region in the electromagnetic spectrum with a wavelength between 10 and 400 nanometers. This region is located between visible light and x-ray rays. Ultraviolet rays with a wavelength shorter than visible light cannot be seen by the human eye. It was created by combining the words ultraviolet, ultra, and violet. The Turkish of the English word violet is purple, the color of the highest frequency of visible light. Ultra means extreme, beyond. Ultraviolet rays are classified into three different subtypes according to their wavelength; Ultraviolet A (UVA), Ultraviolet B (UVB), and Ultraviolet C (UVC).

UVA (315-400 nm): These are the UV rays with the longest wavelength and the least energy. This group constitutes approximately 95 percent of the ultraviolet rays originating from sunlight. Short-term exposure can cause tanning, long-term exposure can cause premature aging, wrinkles, and some skin cancers.

UVB (280-315 nm): These are UV rays with a shorter wavelength and medium energy-carrying. It constitutes 5 percent of the UV rays reaching the earth from the sun. Short-term exposure may cause problems such as tanning, sunburn, and blisters on the skin. Long-term exposure increases the risk of skin cancer and premature aging.

UVC (100-280 nm): These are the UV rays with the shortest wavelength and carrying the highest energy. Since it carries high energy, it is the most harmful UV rays. The rays of this wavelength emitted by the sun are absorbed by the ozone layer and the atmosphere before reaching the earth. Artificial UV rays are generally used in food, air, and water disinfection (purification from germs and bacteria). Black lights similar to fluorescent lamps, ultraviolet LEDs, and ultraviolet lasers are examples of artificial ultraviolet light sources.

2. ULTRAVIOLATED BEAM SPECTRUMS

The type of energy emitted by the sun, stars, and various lamps that give light actually depends on the surface temperature of the source that emits that light. And when this light is refracted by a prism, an electromagnetic spectrum is formed. The light that can be seen with the human eye is a very very small part of this spectrum.

"Red" is the light with the lowest energy and "Blue" is the light with the highest energy in the spectrum, which is distributed in colors ranging from red to blue. Lights after blue are invisible to humans, and this part of the spectrum contains Ultraviolet, X-rays, and Gamma rays.

The largest source of ultraviolet light, which has a wavelength of 10 to 400 nanometers, is actually the Sun. And different Ultraviolet rays are emitted from the sun in the form of UV-A, UV-B, and UV-C.

Ultraviolet A rays with the longest wavelength (320 to 400 nanometers) are also the least harmful. Actually, 85-90% of the light that the atmosphere does not absorb and reaches the earth is UV-A light. Sun tanning, sun spots, signs of aging on the skin, and wrinkles are also the works of Ultraviolet A light.

Ultraviolet B rays (between 290 and 320 nanometers) are slightly more harmful, but 90-95% of them are absorbed by the atmosphere before they reach the earth.

Ultraviolet C rays with a wavelength of 100-290 nanometers, the most dangerous of these, are completely absorbed by the atmosphere. Therefore, you cannot be protected from the Covid-19 virus by going out in the sun, because this type of UV light that kills the virus cannot reach the earth.

3. USE OF ULTRAVIOLE LIGHT

Although it contains radiation, UV light is also an extremely useful technology. Uses in daily life;

- To obtain a burnt skin in the solarium,
- distinguishing counterfeit money from real,
- To color the environment with neon lamps in nightclubs
- reading marks on customers' arm with invisible ink at club entrances
- To disinfect the belts where products are produced in the food industry
- Sterilizing surgical instruments after surgery,
- to treat cancer
- hardening dental fillings
- It is also used in low doses for the body to produce vitamin D

4. UV DISINFECTION EFFECT

All living things, regardless of plant, animal, or human, have building blocks called DNA and RNA in the cell nucleus, and these building blocks reproduce by breaking down during cell division. It has long been discovered that UVC, the most energetic type of ultraviolet light, deforms the DNA of bacterial cells and the RNA of virus cells, preventing their reproduction. UVC has been used in the disinfection of operating theaters and surgical instruments since the end of the 1800s.

The disinfection effect of UVC light has been proven in the tests performed on the viruses Sars-CoV-1 and Mers-CoV, which are related to the Covid-19 virus. For this reason, Harvard Medical Research Professor Edward A. Nardell and Ron Hofmann of Toronto University Professors state that UV lamp devices can be used to protect against the Covid-19 virus. Figure 1 shows the UV-C beam breaking down the mRNA or DNA strand.

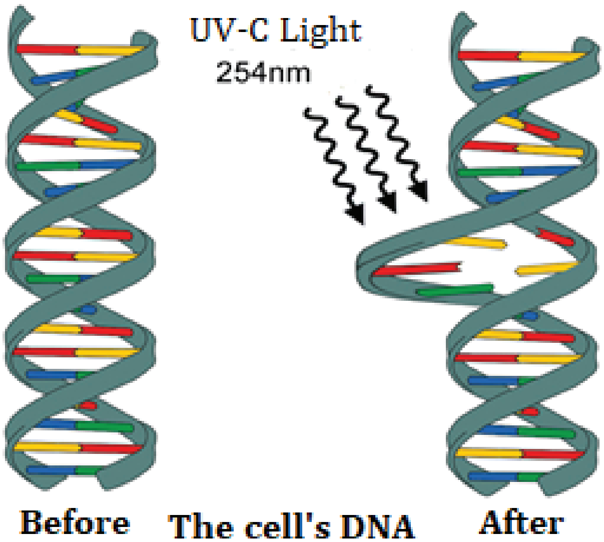


Figure 1. *UV-C beam splitting the mRNA or DNA strand*

This deadly effect of UV light on bacteria and viruses can be used for disinfection in air, water, and surfaces.

- Disinfection with UVC light should be used in conjunction with other hygiene methods. UVC light provides a very high level of protection when used in addition to other disinfection methods in hospitals, food production factories, workshop or restaurant kitchens, stores, cafes, restaurants.

- Since the light can reach even the most hidden surfaces with the right UVC lamp selection, areas that cannot be sterilized sufficiently by other methods can be completely sterilized with UV light.

- With the UVC lamp, disinfection can be applied to the air, surfaces in a room, or water. It is not applied directly or indirectly to human skin.

- UVA light is relatively dangerous as it ages the skin, while UVB light has been discovered to cause skin cancer. UVC light, on the other hand, is much more harmful than these two, and it should never come into contact with humans. However, it is this type of UV light that also kills bacteria and viruses. In fact, UVA and UVB light can also have a fatal effect on bacteria and viruses, but their effects are seen for a long time. UVC light creates this effect in much shorter times.

- Due to the potential dangers of UVC light, the application period should not be human or animal in the area to be applied and the area should be ventilated after the application.

If desired, disinfection with UVC light can be taken for individual use in homes by paying attention to the correct UVC lamp and device selection and of course taking the necessary precautions. Apart from this, such devices can be used in small, medium, and large enterprises for the health of both their employees and customers. Today, there are UV sticks that disinfect drinking water, fixed or mobile UVC lamp devices that disinfect rooms, closed UVC devices that do not harm people because they are closed boxes and disinfect by taking in the air. The representative view of the viruses on the mouth mask is shown in Figure 2.

Figure 2. *Representative view of viruses on mouth mask*

4. LATEST DEVELOPMENTS İN UV-C TECHNOLOGY

Disinfection with UVC light has continued to develop and progress since the beginning of the 1900s when it was used only for the sterilization of operating rooms and surgical instruments. Before the Covid-19 epidemic, UVC lamp devices were used for sterilization in the prevention of seasonal flu diseases and epidemics. Devices with UVC lamps, which are mainly used to disinfect water, air, and surfaces that people come into contact with, are used both industrially and individually. UV devices that sterilize the water are generally in the form of a stick and completely sterilize drinking water by soaking in water for a few minutes. UVC rods used in water are less hazardous to health as they are models with the longest wavelength.

UV lamp devices that disinfect surfaces consist of fixed models suspended on the ceiling and mobile models that can be carried. Since the UVC rays emitted by these devices are dangerous, no people or animals should be present during the application and the area should be ventilated after the application.

Devices with UVC lamps that disinfect the air can be in open or closed forms. Those in closed box form are safer for human health. These devices vacuum the air inside. It disinfects the air in a closed box by exposing it to UVC light and returns the air to the same environment. Today, scientists continue to work on the types of light that do not harm human health rather than how and in what form. Distant UVC Rays are also a result of these studies.

4.1 Far UVC Ray

Distant UVC rays, which can eliminate the carcinogenic and cataractogenic effects of UVC light, can be defined as members of traditional UVC rays. While UVC rays are normally at wavelengths of 100 to 290 nanometers, the longer part of this spectrum, ie UVC rays that are closer to the wavelength of 290 nanometers, is called Far UVC rays.

Distant UVC rays with a wavelength of 210 - 222 nanometers are highly absorbed by microscopic creatures but cannot be absorbed by human skin. The small size of microscopic creatures is the most important parameter here. David Welch and his team, one of the professors at the Center for Radiological Research of Colombia University, also revealed the lethal effect of Far UVC rays on bacteria and viruses with the study they published in 2018. Figure 3 shows the success of the Remote UV-C beam in eliminating viruses.

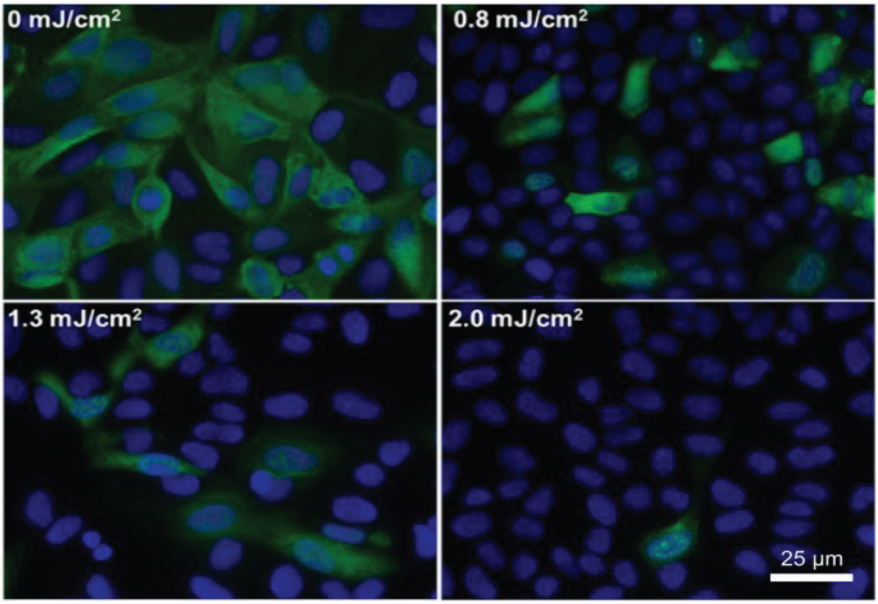


Figure 3. *The success of the distant UV-C beam in destroying viruses*

5. UVC USAGE ADVANTAGES

- Ease of Use

Antibacterial devices with UVC lamps are easy to use compared to methods such as vinegared water and cologne. Therefore, it is ideal for use in places where wet cleaning cannot be performed, such as computers, mobile phones, TV screens. While disinfecting by wiping with liquid cleaners, parts of the objects in a room that is not easily accessible may not be sufficiently disinfected. even in some cases, these places are not accessible at all. As UVC light can reach anywhere that cannot be reached by hand, it performs disinfection with a full performance where it is used. Regardless of the brand and model of UVC lamp disinfection devices, it can be used with an on-off button. It is also possible to make adjustments such as time and power with the help of control panels.

- Reliable Results

The lethal effect of UVC light with a wavelength of 200-280 nanometers on bacteria and viruses has been proven more than 100 years ago and has even been used in the sterilization of operating rooms and surgical equipment in hospitals for a long time. It is one of the safest methods for protection from bacteria and viruses due to its 99% lethal effect on UVC lamp devices.

- UVC lamp disinfection is non-toxic.

Perhaps the most important disadvantage of chemical cleaners is that they are toxic, in other words, toxic to somebodies. Antibacterial UV lamps do not contain any risk in this respect.

- Bacteria and Viruses Do Not Develop Immunity

Some bacteria and viruses are known to develop immunity to various chemical disinfectants and are no longer affected by these chemicals. However, there is no such risk for UV lamp disinfection devices. Because the lethal effect of antibacterial UV lamp devices on viruses and bacteria is not chemical, but physical. Therefore, viruses and bacteria cannot develop any immunity.

6. UV-C VOLUMETRIC AREA CALCULATIONS

A research group from Korea University was able to sterilize the coronavirus (COVID-19) by 99.9 percent in just 30 seconds with the help of UV (Ultraviolet) LEDs. In the study, UVC (Ultraviolet C) light, which is a subtype of UV, was used.

While the coronavirus continues to spread all over the world, LED manufacturers have also concentrated their work on UV LED technology. In this context, some manufacturers allocated resources to show how UV LEDs sterilize the virus.

UV LEDs stand out as an innovative approach to neutralize viruses in the air, sterilize products such as face masks, mobile phones, and minimize the spread of coronavirus.

In the study conducted at Korea University, it was aimed to measure the sterilization effect that occurs when infected surfaces are exposed to UV photons for 30 seconds. In the study, it was also seen that when exposed to UV light for a longer time and closer, the sterilization effect increased and made the surfaces more sterile.

In the experiment performed with UV LEDs, it was proven that 99.9 percent of other harmful bacteria such as escherichia coli, staphylococcus aureus, pseudomonas aeruginosa, klebsiella pneumonlae and salmonella Typhimurium were successfully sterilized in addition to coronavirus.

It was seen that the distance of the exposed surface to the light source and the exposure time were very important for the sterilization to be successful. In the study, a sterilization rate of 1 percent at the end of 10 seconds and 10 percent at the end of 20 seconds was achieved. After 20 seconds, the effect increased exponentially, and after 30 seconds, 99.9 percent sterilization rate was achieved.

NASA's space station in the USA uses UV LEDs to clean the environments where astronauts are located.

Ultraviolet (UV) light effectively prevents and kills the reproduction of viruses and bacteria by breaking the molecular bonds in their DNA. Important parameters in UV sterilization are dose, density, and sterilization rate.

In ultraviolet sterilization, UV dose is the amount of UV energy taken per unit area. It is commonly used in units of mj / cm^2. The UV dose is found by multiplying the UV intensity and time. UV intensity in the equation expresses the energy of ultraviolet energy per unit area and unit time. It is generally used in units of mW / cm^2. In Equation 1, the amount of dose to be applied is calculated.

$$UV\,dose = UV\,intensity\;x\;time \tag{1}$$

Another important parameter in UV sterilization is the inactivation value. Inactivation is the name given to the process of killing or inactivating the microbial agent.

A logarithmic scale is usually used when calculating the change in the number of cells. The logarithmic inactivation value is calculated using the formula below. In the formula, the number of living microorganisms before sterilization is expressed as variable A and the number of organisms after sterilization is expressed as variable B. The amount of logarithmic reduction is seen in equation 2.

$$Logarithmic\,reduction = log10\,(A) - log10\,(B) \tag{2}$$

The logarithmic reduction is translated into the percentage reduction known as the sterilization rate by the formula below. The variable L is the logarithmic reduction value obtained from the formula above. The percentage reduction (decrease) is seen in equation 3.

$$Percent\,reduction = (1\text{-}10^{-L})\;x\;100 \tag{3}$$

A single log reduction is a 90 percent reduction of organisms. The two log reduction is 99 percent reduction of organisms, and the three log reduction is 99.9 percent. The UV exposure dose required for each reduction level is different from each other.

7. CONCLUSION

In scientific studies, it is seen that UV-C is successful in disinfection processes. UV-C disinfection method is the most effective and cheapest method in disinfection processes.UV-Cs must not come into contact with skin and eyes. Since it has harmful effects on human health, it should be used at night and in places where people are not present.

With the increase in technology, the use of UV-C will increase. Suitable UV-C lamps should be selected after engineering calculations in UV-C usage areas.

REFERENCES

Adel Haji Malayeri, Madjid Mohseni, Bill Cairns ve James R. Bolton - Gabriel Chevrefils (2006) and Eric Caron (2006) - Benoit Barbeau, Harold Wright (1999) ve Karl G. Linden https://iuvanews.com/stories/92816/fluence-required-achieve-incremental-log-inactivation-bacteria-protozoa-viruses-algae.shtml

Arıkan, I. 2020. Ultraviyole web notları. https://www.fiyatimbu.com/blog/uvc-lambali-sterilizasyon-cihazlari-hakkinda-sikca-sorulan-sorular_411

Aydınlatma Portalı e-notları: https://www.aydinlatma.org/uv-sterilizasyon-hesaplari-dozu-yogunlugu-ve-sterilizasyon-orani-nasil-yapilir.html

Buonanno, M., Ponnaiya, B., Welch, D., Stanislauskas, M., Randers-Pehrson, G., Smilenov, L., Lowy, F.D., Owens, D.M. and Brenner, D.J. (2017) Germicidal Efficacy and Mammalian Skin Safety of 222-nm UV Light. Radiat Res 187(4): 483–491.

Cengiz, Ç., Kaynaklı, M., Gencer, G., Eren, M., Yapici, İ., Yildirim, S., Cengiz, MS. (2017) Selection Criteria and Economic Analysis, International Conference on Multidisciplinary, Science, Engineering and Technology Bitlis Book of Abstracts,1 (1), pp. 27-29.

Freely available at http://cie.co.at/news/cie-releases-two-key-publications-uv-disinfection2

CIE/ICNIRP (2020) CIE/ICNIRP Online Tutorial on the Measurement of Optical Radiation and its Effects on Photobiological Systems, December 2, 2020. http://cie.co.at/news/cieicnirp-online-tutorial-measurement-optical-radiation-and-its-effects-photobiological-systems, accessed 2020–12–12.

Efe S. B., UPFC Based Real-Time Optimization of Power Systems for Dynamic Voltage Regulation. Computer Modeling in Engineering & Sciences, 2018. V116, #3, pp. 391–406.

Efe S. B., Cebeci M. Power Flow Analysis by Artificial Neural Network. International Journal of Energy and Power Engineering, 2013. V2, #6, pp. 204–208.

Escombe, A.R., Moore, D.A., Gilman, R.H., Navincopa, M., Ticona, E., Mitchell, B., Noakes, C., Martínez, C., Sheen, P., Ramirez, R., Quino, W., Gonzalez, A., Friedland, J.S., Evans, C.A. (2009) Upper-room ultraviolet light and negative air ionization to prevent tuberculosis transmission. PLoS Med. 6(3): e43.

Gencer, G. Eren, M. Yildirim, S. Kaynaklı, M. Palta, O. Cengiz, M.S., Cengiz Ç. Numerical Approach to City Road Lighting Standards, Imeset Book of Abstracts, Int. Conf. Mult. Sci. Eng. Tech., 12-14 July 2017 Baku.

Gläser, R., Navid, F., Schuller, W., Jantschitsch, C., Harder, J., Schröder, J.M., Schwarz, A., Schwarz, T. (2009) UV-B radiation induces the expression of antimicrobial peptides in human keratinocytes in vitro and in vivo. Journal of Allergy and Clinical Immunology 123(5): 1117–1123.

Jordan, W.S. (1961) The Mechanism of Spread of Asian Influenza, Am Rev Resp. 83, 2P2, 29–40.

Ko, G., First, M.W., Burge, H.A. (2000) Influence of relative humidity on particle size and UV sensitivity of Serratia marcescens and Mycobacterium bovis BCG aerosols. Tubercle and Lung Disease. Vol. 80, 4–5, 217–228.

Mphaphlele, M. (2015) Institutional Tuberculosis Transmission. Controlled Trial of Upper Room Ultraviolet Air Disinfection: A Basis for New Dosing Guidelines. Am J Respir Crit Care Med. 192(4): 477–84.

Narita, K., Asano, K., Morimoto, Y., Igarashi, T., Hamblin, M.R., Dai, T. and Nakane, A. (2018) Disinfection and healing effects of 222-nm UVC light on methicillin-resistant Staphylococcus aureus infection in mouse wounds. Journal of Photochemistry and Photobiology B: Biology 178: 10–18.

Parlakyıldız S., Gençoğlu M.T., Cengiz M.S. 2018. Development of Rail Systems from Past to Present. International Conference on Multidisciplinary, Science, Engineering and Technology (2018 Dubai, BAE)

Parlakyıldız, Ş. Gencoglu, M.T., Cengiz, M.S. Analysıs of Faılure Detectıon and Vısıbılıty Crıterıa In Pantograph-Catenary Interactıon, Light & Engineering 28, 6, pp. 127–135, 2020

Peccia, J., Werth, H.M., Miller, S., Hernandez, M. (2001) Effects of Relative Humidity on the Ultraviolet Induced Inactivation of Airborne Bacteria, Aerosol Science and Technology. 35, 3.

Piazena, H. and Häder, D.-P. (2009) Solar UV-B and UV-A irradiance in arid high-mountain regions: Measurements on the island of Tenerife as compared to previous tropical Andes data. Journal of Geophysical Research: Biogeosciences. 114(G4)

Sagripanti, J.-L. and Lytle, C.D. (2011) Sensitivity to ultraviolet radiation of Lassa, vaccinia, and Ebola viruses dried on surfaces. Archives of Virology 156(3): 489–494.

Taylor, W., Camilleri, E., Craft, D.L., Korza, G., Granados, M.R., Peterson, J., Szczpaniak, R., Weller, S.K., Moeller, R., Douki, T., Mok, W.W.K. and Setlow, P. (2020) DNA Damage Kills Bacterial Spores and Cells Exposed to 222-Nanometer UV Radiation. Applied and Environmental Microbiology 86(8): e03039–03019.

Tomas, M.E., Cadnum, J.L., Jencson, A., Donskey, C.J. (2015) The Ebola disinfection booth: evaluation of an enclosed ultraviolet light booth for disinfection of contaminated personal protective equipment prior to removal. Infect Control Hosp Epidemiol. 36(10):1226–1228.

Van Doremalen, N., Bushmaker, T., Morris, D.H., Holbrook, M.G., Gamble, A., Williamson, B.N., Tamin, A., Harcourt, J.L., Thornburg, N.J., Gerber, S.I., Lloyd-Smith, J.O., de Wit, E., Munster, V.J. (2020) Aerosol and Surface Stability of SARS-CoV-2 as Compared with SARS- CoV-1. N Engl J Med. 382: 1564–1567.

Welch, D., Buonanno, M., Grilj, V., Shuryak, I., Crickmore, C., Bigelow, A.W., Randers- Pehrson, G., Johnson, G.W. and Brenner, D.J. (2018) Far-UVC light: A new tool to control the spread of airborne-mediated microbial diseases. Scientific Reports 8(1): 2752.

Yamano, N., Kunisada, M., Kaidzu, S., Sugihara, K., Nishiaki-Sawada, A., Ohashi, H., Yoshioka, A., Igarashi, T., Ohira, A., Tanito, M. and Nishigori, C. (2020) Long-term effects of 222 nm ultraviolet radiation C sterilizing lamps on mice susceptible to ultraviolet radiation.

Yıldırım, S., Yapıcı, I., Atiç, S.,Eren, M.,Palta, O., Cengiz, Ç., Cengiz, MS., Yurci., Y. (2017) Numerical Analysis of Productivity and Redemption Periods in Illimunation, Imeset Book of Abstracts, Int. Mult. Sci. Eng. Tech., 1 (1), pp. 12-14.

Chapter 40

A STUDY ON CLASSIFICATION AND DECISION SUPPORT METHODS IN IMAGE PROCESSING AND MACHINE LEARNING

Olcay PALTA[1]

1 Öğr.Gör. Olcay PALTA, Bitlis Eren Üniversitesi Teknik Bilimler MYO

1. INTRODUCTION

The use of digital imaging and machine learning practices gradually increases and gains importance. Digital imaging can be defined as analyzing results and interpreting these analyzes by affecting the digital data. Image processing process includes determining the distinctive characteristics of the images, which are obtained from various media with different characteristics and under different conditions by making use of a camera or scanner, by cleaning, filtering, subjecting them to certain processes, specifying the needed areas, applying transformation processes. Machine learning can be explained as developing solutions and experiences in computerized systems by making use of digital data. The main function of machine learning is to create programmatic algorithms providing input data, update the algorithm outputs when new data received, and use statistical analyses in order to estimate the next situation. In the present study, the feature extraction, classification, and decision support methods in image processing and machine learning are discussed.

2. FEATURE EXTRACTION, CLASSIFICATION, AND DECISION SUPPORT METHODS

2.1. Convolutional Neural Network

Machine learning is a science that allows a system to create a solution on a subject by making use of a programmatic approach and provides information to computerized systems. Artificial neural network is a method developed based on the bran's methods for processing the information. It can solve problems with mathematics by making use of a neural network artificially constructed using previously developed methods. Deep learning is, however, the use of multilayered deep neuronal networks in order to create systems extracting data clusters trained using large data diversity. Deep learning is a machine learning approach and is an area of study that draws significant attention of computer scientists in recent years. CNN (Convolutional neural network), one of deep learning methods, uses object identification and detection as image classification. These methods offer high accuracy in classification [1]. Local binary patterns (LBP) are the identifiers translating the grey images into digital information and recommended for measuring the pattern from this transformation. LBP includes transformation procedures translating a reference point, which is selected as center, and the neighbor points into digital data in relationship with the reference values.

2.2. Discrete Cosine Transform

It is a transformation of a numeric data and cosines in different frequencies and amplitudes that can be expressed as summative. It is very frequently used for reducing the data size, especially for large data. For reducing the size, the harmonic components having high frequency from the data obtained from discrete cosine transformation process [2].

2.3. Directed Gradient Histogram

Becoming very popular in recent years, directed gradient histogram is widely seen in literature [3]. It is an identifier defining the object with its angular aspects and used in detecting and identifying the object [2]. The use of directed gradient histogram (DGH), which can be defined as a specific image input's pixels' angular direction and amplitude characteristics, draws significant attention in various fields in recent years [4]. In this method, the gradients of image are calculated first. After the directed gradient calculation, every pixel angle of output image is recorded. The image is divided into cells, which are rectangular regions that are independent from each other. Then, the histograms of every region's gradients are calculated. Cells are gathered together and blocks are created. The histograms obtained from the cells in each block are locally normalized. Then, by successively connecting these histograms, DGH features are obtained [5].

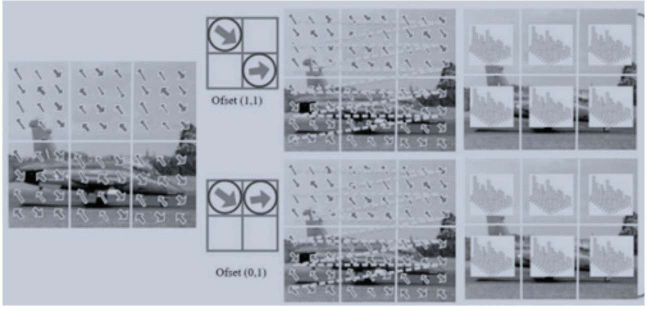


Figure 1. *Sample of directed gradient conformation histogram, one of DGH methods [3]*

2.4. Support Vector Machine

Becoming popular in determining the pattern, support vector machine is a method classifying data into two classes and creating an n-dimensional plane [6]. SVM is a classification method that aims to find the best sub-plane by separating a specific point cluster from a data belonging to two categories [7]. In fact, it can be described as a machine learning method based on vector space creating a decision scale from a random point in training data to the furthest two other classes.

As another definition, by dividing data output set into two with decision line, let's assume that the data in both classes show linear distribution.

Thus, it is aimed to separate them by using decision function obtained from training data.

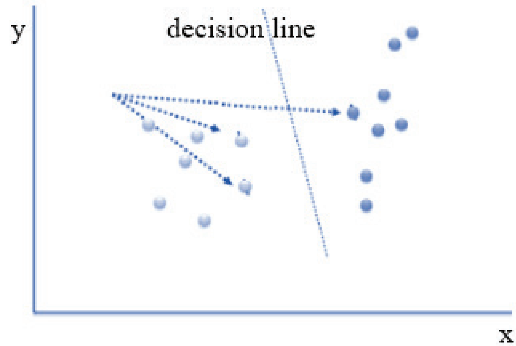


Figure 2. *Illustration of support vector with decision line*

In order for decision line borderline to be robust to new data, it is desired to be close to the border between two groups. The points near to border are known as support points [8].

2.5. Gabor Filters

In interpreting a visual data, human eye uses spectral and spatial data. Spectral data refer to the transition of tones in colors. Spatial data, however, refer to the spatial distribution of changes in transitions of these color tones [9].

Gabor filters, which identify the pattern data, are a color-pattern identification method defining the images. When Gabon filter is applied to an image, the point having the same frequency and angle value in the image will be the strongest point. Thus, the Gabor description of an image is obtained using Gabor filter method [10].

2.6. Canny Edge Detector

It can be described as a feature extraction operator, which detects the edges in an object by making use of multiphase algorithm in detecting the edges. Among edge detection points, there are various algorithms such as Gradient edge detection algorithm and LoG (Laplacian of Gradient) edge detection algorithm.

For an image, edge detection is of a significant importance but, since every user detects the edge differently, Canny operator is used in edge detection. Since the Canny edge detector detects the edges in an image by considering the signal-noise ratio, it is superior to the other edge description detectors [11]. In Canny edge detection algorithm, image is subjected to fuzzyfication, gradient calculation, suppression of non-maximum points, and thresholding processes, respectively.

2.7. Discrete Wavelet Transformation

In continuous wavelet transformation, it is multiplied with main wavelet's iterated and scaled versions and signal's total. There are many calculations in continuous wavelet transformation. In order to reduce the calculations, the discrete wavelet transformation is preferred [12].

In discrete wavelet transformation, two function clusters (scaling function and wavelet function) are used. These correspond to low-pass and high-pass filters [13].

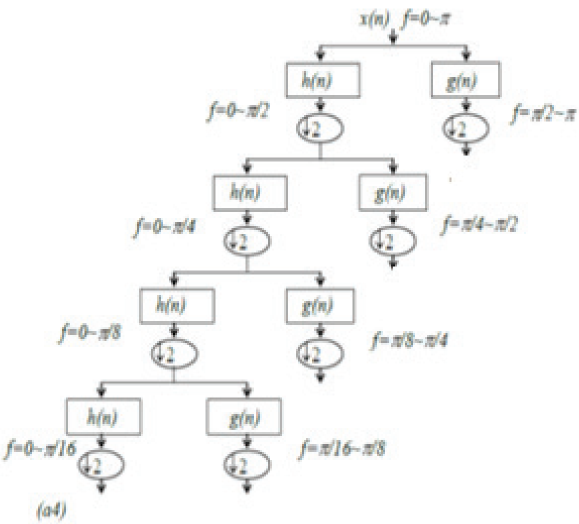


Figure 3. *A sample of 4th grade wavelet [14]*

2.8. Knn Algorithm

It is the most common algorithm in the sample-based learning. The first sample is classified based on the degree of similarity with data in learning set. In KNN algorithm, the data in learning set are specified as n-dimensional.

All the training samples are kept in n-dimensional sample space in the way each sample represents a point in n-dimensional space. When there is a new sample, k units of sample closest to the relevant data in the training dataset are determined and the classification label of new sample is determined according to majority of the classification labels of k's nearest neighbor [15].

2.9. Decision Tree Algorithm

It is a classification-based pattern determination method that is widely used in literature in recent years. The most important reason for the popularization of this method is that the rules sued in tree structure are clear and simplified.

Decision tree is a tree-shaped decision algorithm that is known as induction from the classified sample data. In many classification problems, in which high amount of data are used, decision trees offer useful solutions for the complex data or the data containing error. The decision tree algorithm is a decision algorithm, in which the fundamental decision-making steps are followed and large amounts of data are divided into small groups. In every dividing process, the members of resultant group become more similar to each other [16].

Figure 4 illustrates a sample of decision tree structure. As seen in the figure, the tree consists of three main structures namely nodes, branches, and leaves. Each feature in this tree structure is represented by a node. The farthest parts of tree structure are leaves, whereas the bottom parts are root. The regions between the leaves and the bottom are named branches.

The main objective in creating a decision tree is to ask and answer specific questions about the data, to move in parallel with these answers, and to achieve the conclusion as soon as possible. The answers to the questions are used in setting the decision rules of tree [17].

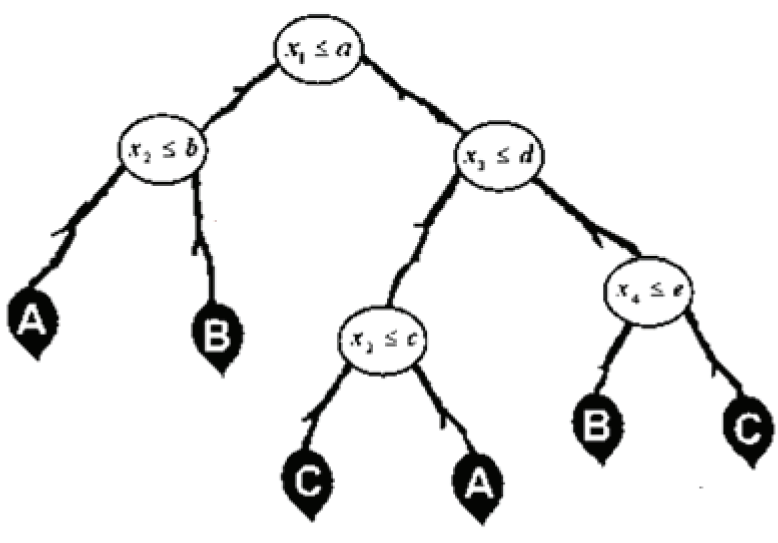


Figure 4. *Structure of a decision tree consisting of 4 dimensions and 3 classes [17]*

2.10. Gauss Mixture Model

It is the optimization of Gauss parameters of sources in the way maximizing the possibility density function of sources by assuming that there are multiple independent sources creating samples using Gauss distribution. As a result, it is assumed that the dataset is created from a single distribution. Thanks to the estimation of necessary distribution parameters, this method provides useful results even when the systems modeling fall short. In Gauss Mixture Model training, all the groups are considered unrelated to each other. For this reason, training is performed separately for each group. It is a mode, in which the training dataset is defined with the mixture of multiple independent Gauss distributions [18].

3. CONCLUSIONS

With image processing and machine learning, various methods were used in determining the object features, identification, result optimization, and maximizing the detection rate. In the present study, among the feature extraction, classification, and decision support methods, Convolutional neural network, Discrete cosine transformation, Gabor filter, Directed gradients histogram, Support vector machine, Canny edge detector, discrete wavelet transformation, Gauss mixture model, KNN algorithm, and Decision tree algorithm are discussed. In conclusion, it was determined that the methods may vary depending on the characteristics of image processing and machine learning study and each method has different application techniques.

REFERENCES

[1] Arı, A., Hanbay, D*, Tumor detection in MR images of regional convolutional neural networks*, Journal of the Faculty of Engineering & Architecture of Gazi University, cilt 34(3), ss.1395-1408,2018.

[2] Abacı, B., Ulucan, E., ve Akgül, T. (2014). Cinsiyet Sınıflandırma için birlesim Tabanlı Öznitelik Vektörü Fusion Based Feature Vector For Gender Classification, *22nd Signal Processing and Communications Applications Conference,* IEEE

[3] Alpaslan, N. (2018). Renkli görüntülerde kenar yönelimlerini almak için iyileştirilmiş yöntem. Türk Doğa ve Fen Dergisi, 7 (1), 29-34. Https://dergipark.org.tr/en/pub/tdfd/issue/37853/364147 adresinden erişildi.

[4] Alpaslan, N., Hanbay, K., Hanbay, D., Talu M.F. (2014). Hessian matrisi ve temel eğriliklere dayanan yeni bir doku sınıflandırma metodu, *22nd Signal Processing and Communications Applications Conference*, IEEE

[5] Akyüz, O., Usanmaz, G., Aydın, M., Gündüz, H., Çevikalp, H. (2013). Konveks sınıf modelleri kullanarak dijital imgelerdeki nesne görüntülerinin konumlarının bulunması, 109E279 nolu proje, TÜBİTAK

[6] Yakut, E., Elmas B. ve Yavuz, S.,(2014). Yapay Sinir Ağları ve Destek Vektör Makineleri Yöntemleriyle Borsa Endeksi Tahmini, *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 19(1):139-157

[7] Sezer, O. G., Erçil, A., Keskinöz, M. (2005). Destek vektör makinesi kullanarak bağımsız bileşen tabanlı 3B nesne tanıma, *Sabancı Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi Dergisi.*

[8] Kaan ülgen, https://medium.com/@k.ulgen90/makine-%C3%B6%C4%9Frenimi-b%C3%B6l%C3%BCm-4-destek-vekt%C3%B6r-makineleri-2f8010824054 ,Erişim: 17 Haziran 2020.

[9] Akar, Ö. (2019). Göktürk-2 ve Worldview-2 Yöntem Görüntülerinin Görüntü Keskinleştirme Değerlendirilmesi. *Erzincan Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 12 (2), 874-885. DOI: 10.18185 / erzifbed.495854

[10] Kılınç, M., Uludağ, U. (2012). Gender identification from face images. *20th Signal Processing and Communications Applications Conference,* ss. 1-4, IEEE

[11] Cam, A, Topan, H . (2018). RASAT ve GÖKTÜRK-2 Görüntülerinin Gerçek Yer Örnekleme Aralığının Belirlenmesi. Karaelmas Fen ve Mühendislik Dergisi, 8 (1), 292-298. Retrieved from https://dergipark.org.tr/en/pub/karaelmasfen/issue/57119/805732

[12] Tepe, C.,Sezgin, H. (2007). EKG Sinyallerinde Gürültü Gidermede Ayrık Dalgacık Dönüşümünde Farklı Ana Dalgacıkların Ve Ayrıştırma

Seviyelerinin Karşılaştırılması. *12. Elektrik, Elektronik, Bilgisayar ve Biyomedikal Mühendisliği Ulusal Kongresi*, Eskişehir

[13] Yılmaz Z., Bozkurt M. R., "Ayrık Dalgacık Dönüşümü Kullanarak Aritmilere Ait Özniteliklerin Çıkarılması", XIV. Akademik Bilişim Konferansı, 23-25 Ocak 2013, Antalya.

[14] Palabaş, T. (2014). Ayrık Dalgacık Dönüşümü Bileşenlerine Ait İstatistiksel Veriler ile Epileptik EEG İşaretlerinin Sınıflandırılması. *Bilişim Teknolojileri Online Dergisi,* cilt 5(15), ss.49-56

[15] Taşcı, E., Onan, A. (2016). K-en yakın komşu algoritması parametrelerinin sınıflandırma performansı üzerine etkisinin incelenmesi, *Akademik Bilişim*, Aydın, Türkiye, ss.1-8

[16] Albayrak, A.S, Yılmaz, Ş. K., "Veri Madenciliği: Karar Ağacı Algoritmaları ve İMKB Verileri üzerine bir Uygulama", *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, C.14, S.1 s.31-52.

[17] Kavzoğlu, T., Çölkesen, İ. (2010). Karar ağaçları ile uydu görüntülerinin sınıflandırılması: Kocaeli örneği. *Harita Teknolojileri Elektronik Dergisi*, cilt 2(1), ss.36-45

[18] Herkiloğlu, K. (2005). *Gauss Karışım Modelleri Kullanılarak Ses İmzalarının Sınıflandırılması,* Yüksek Lisans Tezi, Fırat Üniversitesi, Fen Bilimleri Enstitüsü

[19] Parlakyıldız S., Gençoğlu M.T., Cengiz M.S. 2018. Development of Rail Systems from Past to Present. International Conference on Multidisciplinary, Science, Engineering and Technology (2018 Dubai, BAE)

[20] Parlakyıldız, Ş. Gencoglu, M.T., Cengiz, M.S. ANALYSIS OF FAILURE DETECTION AND VISIBILITY CRITERIA IN PANTOGRAPH-CATENARY INTERACTION, Light & Engineering Vol. 28, No. 6, pp. 127–135, 2020

[21] Gencer, G. Eren, M. Yildirim, S. Kaynaklı, M. Palta, O. Cengiz, M.S., Cengiz Ç. Numerical Approach to City Road Lighting Standards, Imeset Book of Abstracts, Int. Conf. Mult. Sci. Eng. Tech., 12-14 July 2017 Baku.

[22] Cengiz, Ç., Kaynaklı, M., Gencer, G., Eren, M., Yapici, İ., Yildirim, S., Cengiz, MS. (2017) Selection Criteria and Economic Analysis of LED, International Conference on Multidisciplinary, Science, Engineering and Technology Bitlis Book of Abstracts,1 (1), pp. 27-29.

[23] Efe S. B., UPFC Based Real-Time Optimization of Power Systems for Dynamic Voltage Regulation. Computer Modeling in Engineering & Sciences, 2018. V116,#3, pp. 391–406.

[24] Çıbuk M., Cengiz M.S. Determination of Energy Consumption According to Wireless Network Topologies in Grid-Free Lighting Systems. Light & Engineering, 2020. V28, #2, pp. 67–76.

[25] Efe S. B., Cebeci M. Power Flow Analysis by Artificial Neural Network. International Journal of Energy and Power Engineering, 2013. V2, #6, pp. 204–208.

[26] Çıbuk M., Reducing Energy Consumption in Single-Hop and Multi-Hop Topologies of Road Lighting Communication Network, 2020. V28, #4 pp. 91–102.

Chapter 41

CR(VI) REMOVAL FROM AQUEOUS SOLUTIONS BY ELECTROCOAGULATION USING ALUMINUM ANODES

Belgin KARABACAKOĞLU[1]

1 Assoc. Prof. Dr., Eskişehir Osmangazi University, Faculty of Engineering and Architecture, Chemical Engineering Department, Eskişehir, Türkiye, e-mail: bkara@ogu.edu.tr

1. INTRODUCTION

Heavy metals are among the common environmental pollutants and found in wastewater generated by various industrial activities. Chrome, which is among these heavy metals, is released especially in the effluent of electroplating and leather tanning (Babu et al., 2007). Chromium ion can be in +3 and +6 oxidation steps in wastewater. Since Cr^{3+} has low solubility, it can easily precipitate as hydroxides in the appropriate pH range, and is less harmful to the environment. On the other hand, Cr^{6+} is carcinogenic and mutagenic; it spreads rapidly to soil and aquatic environment and is a strong oxidizing agent (Mouedhen et al., 2009;). High levels of Cr(VI) can damage breathing and damage the nose, lungs, stomach and intestine, it was included in the 1st group of carcinogenic substances by the International Agency for Research on Cancer (Öztürk, 2018). According to the water pollution control regulations in Turkey; the discharge limit for Cr(VI) in wastewater from electroplating facilities is 0.5 milligram per liters. Therefore, wastewater containing Cr(VI) cannot be discharged into the receiving environment without treatment. Several treatment methods have been developed to remove chromium from wastewater and water. Chemical coagulation, adsorption, and biological and electrochemical methods can be used to remove Cr (VI) (Jin et al., 2011). Chemical coagulation is a simple and widely used method under suitable chemicals and ambient conditions (Zhao et al., 2011), but this method has number of disadvantages such as requiring coagulant and producing high levels of sludge (Karhu et al., 2012).

The electrochemical methods such as electrocoagulation and electro-Fenton appear to be effective for the treatment of different effluents compared to conventional methods. The electrochemical methods use a clean reagent, the electron. It also has other advantages, such as versatility, ease of automation, being a safe method due to its operation under mild reaction conditions. Therefore, electrochemical technologies for Cr (VI) treatment have attracted great interest (Akhtar et al., 2020; Ahangarnokolaei, 2018; Sharma et al., 2020).

The electrocoagulation method is based on the production of a coagulant in the reactor continuously. Furthermore, there is an opportunity of reduction and oxidation of substances at the cathode and anode, respectively. This method does not require supplementing chemicals externally, so the volume of produced sludge is reduced. Because of these advantages, electrocoagulation eliminates some of the disadvantages of traditional methods and provides a sustainable and economical treatment of wastewater (Heidman and Calmano, 2008; Mollah, 2011; Ahangarnokolaei, 2018). The electrocoagulation provides compact treatment facility, efficient particulate removal, and high treatment rate (Tanyol et al., 2018). The EC process has advantages such as being able to remove very small colloidal particles, produce less sludge, remove non-biodegradable organic materials or make them ready for biological treatment, easy control of coagulant dosage, removal of heavy metal ions by adsorption on hydroxides or flocs. It does not require continuous pH control and requires relatively low investment costs (Akhtar et al., 2020).

In electrocoagulation, coagulant substances are produced in situ by oxidation reactions occurring at the sacrificial metal anode. Iron and aluminum plates are generally used as anode material (Shafaei et al., 2011; Babu et al., 2007).

$M(s) \rightarrow M^{n+} + ne^{-}$ (anodic dissolution) (1)

$2H_2O + 2e^{-} \rightarrow 2OH^{-} + H_2$

(cathodic hydrogen evolution in basic media) (2)

$2H^{+} + 2e^{-} \rightarrow H_2$

(cathodic hydrogen evolution in asidic media) (3)

$M^{+} + 2OH^{-} \leftrightarrow M(OH)_2$ (precipitation in bulk solution) (4)

Hydrogen gas bubbles released at the cathode help the blocked particles to move towards the surface of the wastewater. Electrocoagulation (EC) is a process carried out by dissolving the anode as a result of electrolysis and forming metal hydroxide flocs in bulk solution. It is accepted that EC takes place in 4 consecutive stages (Akhtar et al., 2020). (1) Formation of coagulants (metal cations) via electrolytic oxidation of the anode; (2) Production of hydroxyl ions and hydrogen gas by reduction of water at the cathode, (3) Metal hydroxide formation through reactions between metal and hydroxide ions in bulk solution, (4) Destabilization of pollutants, breakdown of emulsions, aggregation of destabilized phases to form flocs and adsorption of pollutants by metal complexes. Destabilization of pollutants, suspended particles and the breakdown mechanisms of emulsions can be summarized with the following steps (Mollah et al., 2001):

- With the dissolution of the anode, metal ions pass into the solution and interact with the charged species in the wastewater and compress the diffuse double layer around them.

- The ions released by electrochemical dissolution of the anode neutralize the opposite charged ions in the solution. Thus, Vander Vaals attraction force becomes more dominant by decreasing the repulsion force between ions having similar charge in wastewater. This provides a rapid coagulation of colloidal particles and pollutants.

- The flocs formed as a result of coagulation form a sludge blanket that bridges and entraps uncomplexed colloidal particles.

Therefore, compared to conventional coagulation-flocculation applications, it has the potential to remove the smallest colloidal particles due to the presence of its electrical field.

There were many parameters that affect the efficiency of the electrocoagulation process. Electrode material, electrode configuration, surface area of anodes, applied potential, solution pH and conductivity are the main variables that need to be optimized.

The aim of this work is to investigate the performance of electrocoagulation system for removal of Cr(VI) from dilute solutions. Also, the effect of experimental parameters such as applied potential, pH, initial Cr(VI) concentration, and KCl concentration is studied in view of energy consumption and removal efficiency.

2. MATERIALS AND METHOD

2.1. Materials

Cr(VI) solutions was prepared from $K_2Cr_2O_7$. The pH of the solutions was adjusted with 0.1 M HCl and 0.1 M NaOH solutions. The KCl was used as supporting electrolyte for the adjusting of electrolyte conductivity. The concentration of Cr(VI) was determined spectrophotometrically using diphenyl carbazide. The analysis method is based on the formation of a pink colored complex of Cr(VI) and carbazide in a solution acidified with H_2SO_4. Measurements were carried out at 540 nanometers, which is the wavelength at which chromium solutions show maximum absorbance.

2.2. Experimental Procedure

The lab-scale batch experimental setup used for the EC studies is shown in Fig.1. The experimental setup consists of a reactor, DC power supply and magnetic stirrer. 500 mL of plexiglass reactor with the shape of the rectangular was used. All experimental studies were made with a model solution volume of 400 mL. Aluminum plates of 1 mm thickness were used as the anode, and stainless steel plates were used as the cathode. The electrode dimensions were 10 cm x 9 cm x 0,2 cm. The distance between the electrodes was 1.2 cm. The electrodes were connected as monopolar to a digital DC power supply (0 –20 V, 0–5 A). Solution was agitated with magnetic stirrer. All experiments were conducted at room temperature.

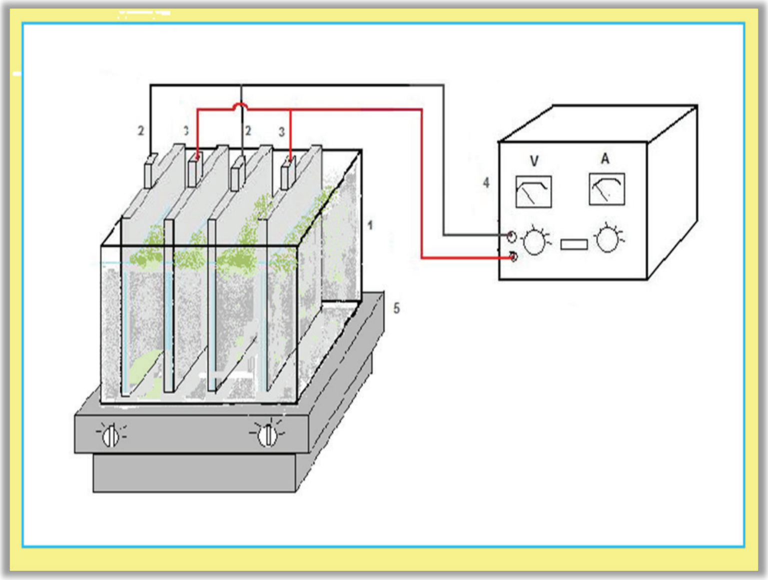


Figure 1. Experimental set-up. 1. electrocoagulation cell; 2. stainless steel cathodes; 3. aluminum anodes; 4. D.C. power supply; 5. magnetic stirrer

In this study, the model wastewater is prepared using potassium dichromate ($K_2Cr_2O_7$) which gives the solution dichromate and potassium ions:

$$K_2Cr_2O_7 \rightarrow Cr_2O_7^{2-} + 2K^+ \qquad (5)$$

When the electric current is applied to the electrocoagulation reactor, the following reactions occur at the anode and cathode (Akbal and Camcı, 2011):

Anode reaction

$$Al\ (s) \rightarrow Al^{3+}(aq) + 3e^- \qquad (6)$$

Cathode reaction

$$2H_2O\ (l) + 2e^- \rightarrow 2H_2\ (g) + 2OH^-\ (aq) \qquad (7)$$

On the other hand the reaction can occur in bulk solution at high pH values:

$$2\ Al + 6H_2O\ (l) + 2OH^- \rightarrow Al(OH)_4^- + 3H_2\ (g) \qquad (8)$$

This and similar monomeric species formed as a result of the reaction of Al released at the anode and OH ions released at the cathode eventually form amorphous aluminum hydroxide according to the following reaction, which adsorbs the pollutants from the solution and removes them from the solution (Akbal and Camcı, 2011):.

$$Al^{3+}(aq) \rightarrow 2OH^{-} + Al(OH)_3 \qquad (9)$$

In addition to these reactions, it is possible to reduce the chromium to the less harmful Cr^{3+} through the following reactions depending on the pH value ((Mouedhen et al., 2009).

$$Cr_2O_7^{2-} + 14\,H^{+} + 6\,e^{-} \rightarrow 2Cr^{3+} + 7\,H_2O \qquad \text{(acidic media)} \quad (10)$$

$$Cr_2O_7^{2-} + 4\,H_2O + 3\,e^{-} \rightarrow Cr^{3+} + 8\,OH^{-} \qquad \text{(basic media)} \quad (11)$$

The parameters studied were solution pH, initial Cr(VI) concentration, applied voltage and supporting electrolyte concentration. In order to determine the effect of a variable on removal efficiency, other parameters were kept constant except for the variable examined. The experimental conditions determined according to this are given in Table 1.

Table 2. *Experimental conditions*

Parameters	Variable range	Values of fixed variables
Applied potential (V)	2.5, 5, 7.5	C_0=50 mg/L; pH=5; KCl=1 g/400 mL
pH	3, 5, 7	C_0=50 mg/L; E=5 V; KCl=1 g/400 mL
Initial Cr(VI) conc. (mg/L)	50, 100, 150	E=5 V; pH=5; KCl=1 g/400 mL
Supporting electrolyte (g/400 mL)	0.5, 1, 1.5	C_0=50 mg/L; E=5 V; pH=5

In this study, the effects of the variables on Cr(VI) removal and energy consumption were investigated in electrocoagulation process. In order to evaluate the results, the percentage of Cr(VI) removal and energy consumption were calculated based on experimental data.

The removal efficiency (RE: %) was calculated using the following equation:

$$RE\,(\%) = 100.(C_0 - C)/C_0 \qquad (12)$$

where C_0 and C were the initial and present concentrations of the Cr(VI) in solution (mg/L), respectively.

Also the electrical energy consumption (EEC: Wh/m^3) was calculated using Eq. 5:

$$\text{Energy Consumption (Wh/L)} = U.I.t/V \qquad (13)$$

where E, is the applied potential (V); I, the current intensity (A), t, the time (h) and V, the volume of treated solution (m^3).

3. RESULTS AND DISCUSSION

3.1. Influence of applied potential

Applied voltage has a significant effect on removal efficiency especially in view of energy consumption. Since the electric current directly affects the dissolution rate of the anode, it determines the amount of coagulant in the solution and the growth rate of the flocs. It also affects the rate of hydrogen generation at the cathode and the size of gas bubbles.

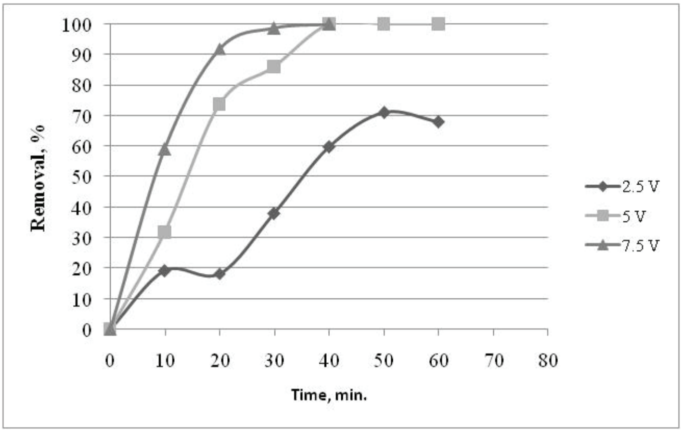


Fig. 2. The effect of applied potential on the removal efficiency

(C_0=50 mg/L; pH=5; KCl=1 g/L; Conductivity= 5.03 mS/cm)

In EC process, the amount of Al^{3+} ions released from the electrodes and hence the formation rate of hydrous oxides is directly proportional to current density. Current density is also proportional to applied potential according to Ohm's law. The effect of potential on the Cr(VI) removal and energy consumption was shown in Fig. 2 and 3. As the current increases, the dissolution of the anode also increases (Fig.2). The removal of Cr(VI) increased regularly with time, but higher voltage induced greater energy consumptions (Fig.3). Therefore, further experiments were carried out at 5 V economical point of view.

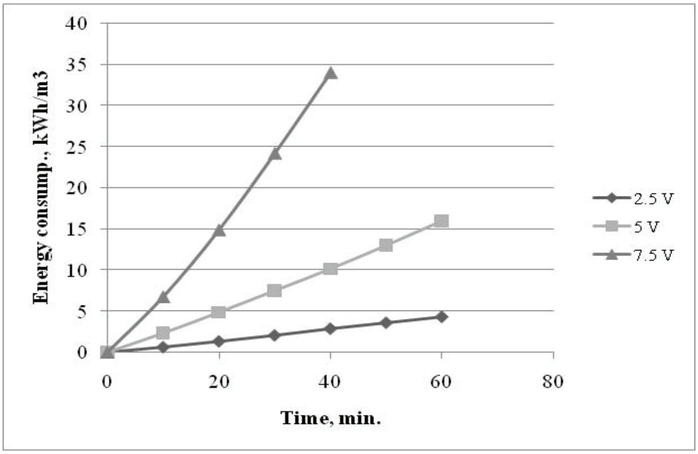


Fig. 3. The effect of applied potential on the energy consumption
(C_0=50 mg/L; pH=5; KCl=1 g/L; Conductivity= 5.03 mS/cm)

3.2. Influence of pH

One of the key parameters affecting the efficiency of the electrocoagulation process is pH. The increase in H^+ concentration accelerates the corrosion of anode. Experiments were carried out with different initial pH (3, 5 and 7) under identical conditions. As shown in Fig. 4, the maximum removal values were obtained at pH 7 and 5 in 50 minutes.

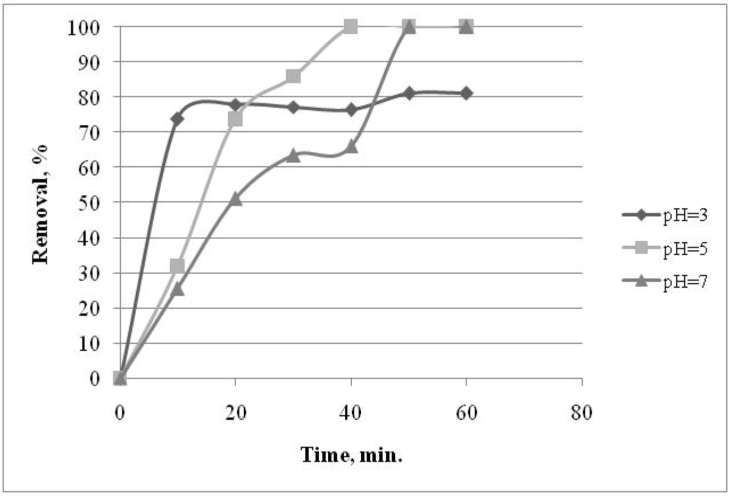


Fig. 4. The effect of pH on the removal efficiency
(E= 5 V; C_0=50 mg/L; KCl=1 g/L; Conductivity= 4,82 mS/cm)

The higher energy consumptions obtained for the pH 3 can be explained by increased conductivity of solution, and hence the higher current density results in higher energy consumption (Fig. 5). Moreover, the increase in the solution pH results in the increase in the side reactions.

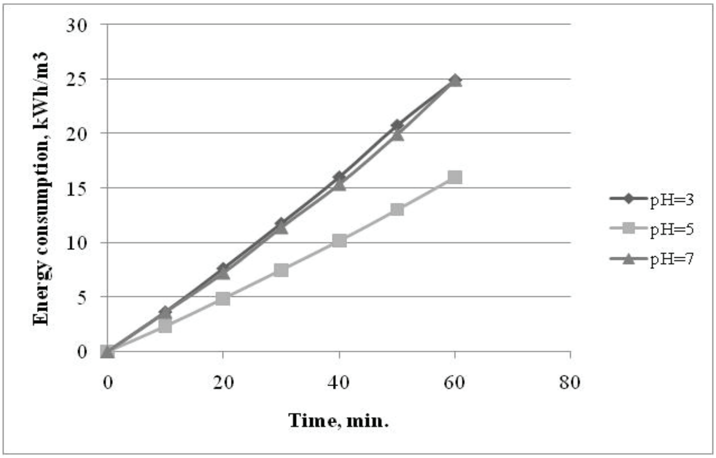


Fig.5. The effect of pH on the energy consump.

(E= 5 V; C_0=50 mg/L; KCl=1 g/L; Conductivity= 4,82 mS/cm)

3.3. Influence of initial Cr(VI) concentration

The results of Fig. 6 and Fig. 7 showed that the performance of the process decrease when the initial Cr(VI) concentration increases. The lower removal efficiency at higher concentrations can be explained by decreased electrode capacity.

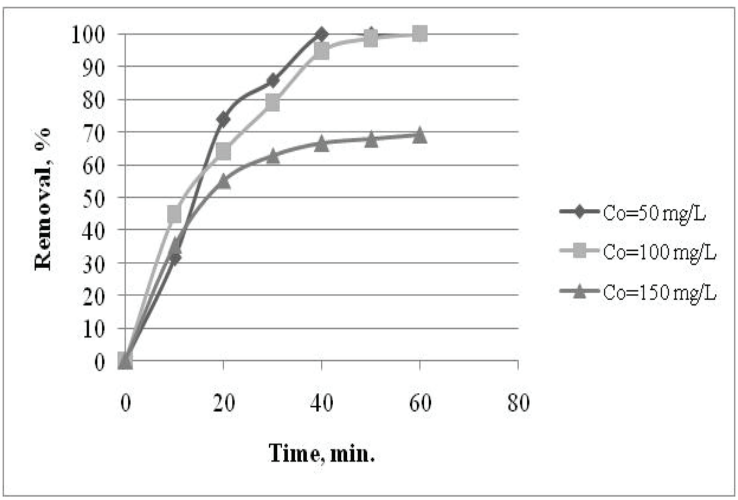


Fig.6. The effect of Cr(VI) conc. on the removal efficiency

(E= 5 V; pH=5; KCl=1 g/L; Conductivity= 4,96 mS/cm)

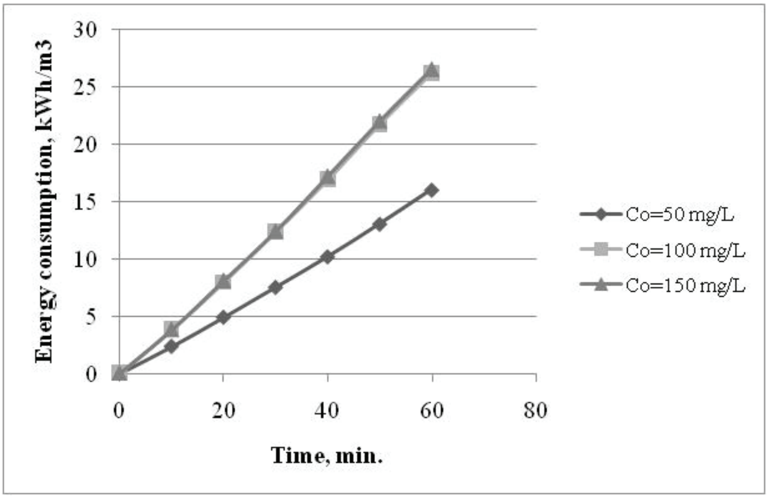


Fig.7. The effect of Cr(VI) conc. on the energy consump.

(E= 5 V; pH=5; KCl=1 g/L; Conductivity= 4,96 mS/cm)

3.4. Influence of KCl concentration

KCl was added to Cr(VI) solution as supporting electrolyte. The performances of electrocoagulation system with various KCl concentrations were analysed in terms of removal efficiency and energy consumption (Fig. 8 and Fig. 9). The removal efficiency increased with the KCl concentration. However, maximum value of energy consumption was obtained in 1,5 g/400 mL KCl concentration as can be seen in the Fig. 9.

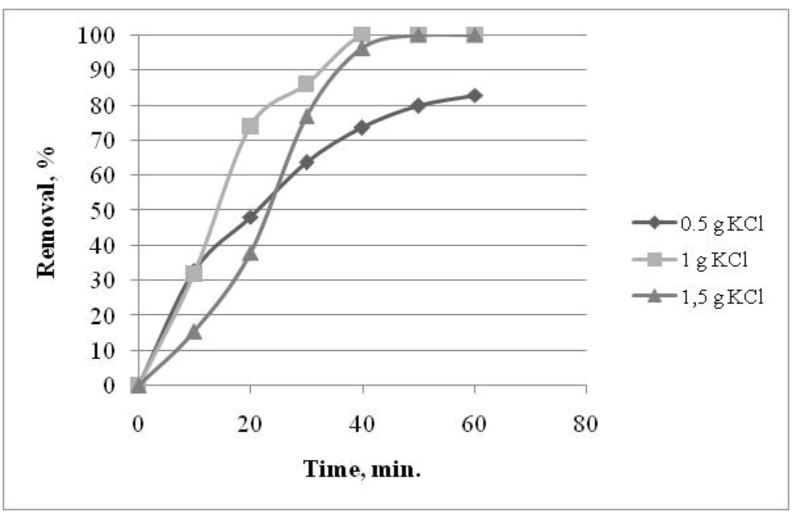


Fig.8. The effect of KCl conc. on the removal efficiency

(E= 5 V; C_0=50 mg/L; pH=5)

Also, increasing of the KCl concentration increases side reactions such as chlorine evaluation at the cathode. The increase in the amount of supporting electrolyte from 1 g to 1,5 g did not have a significant effect on the removal efficiency as seen in Fig.8. When 0.5, 1 and 1.5 g of KCl were added, the conductivity of the solution was measured as 3.84, 4,98 and 6,84, respectively.

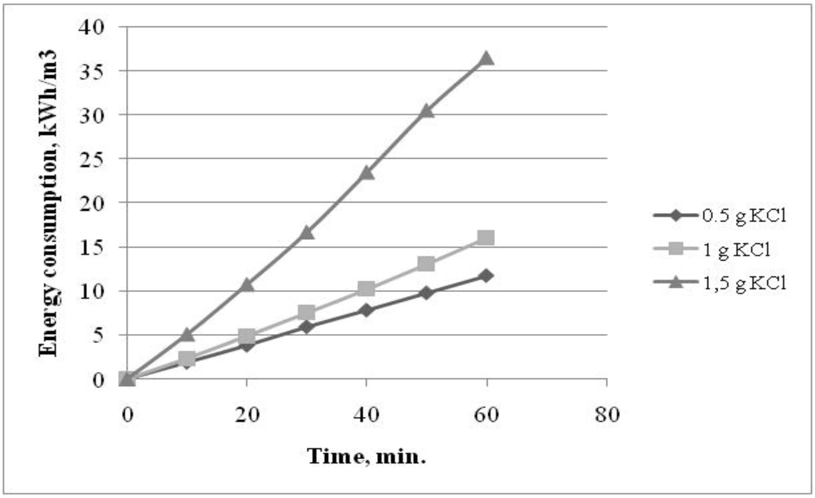


Fig.9. The effect of KCl conc. on the energy consump.

(E= 5 V; C_0=50 mg/L; pH=5)

the addition of supporting electrolyte increases the conductivity by decreasing the solution resistance. The increase in the solution conductivity increases the current density. Increasing the current density also increases the removal efficiency. Chlorine ions also accelerate the dissolution of the anode (Bazrafshan et al., 2008).

In this study, the amount of sludge formed for each parameter were also determined. The highest amount of sludge occurred with the addition of 1.5 grams of potassium chloride. In addition, the increase in applied potential caused an increase in the amount of sludge.

Table 1. The amounts of sludge produced in the electrocoagulation reactor.

Parameters	Applied potential (V)			Cr(VI) concentration (mg/L)			pH			KCl conc. (g/400 mL)		
Levels	2.5	5	7.5	50	100	150	3	5	7	0.5	1	1.5
Sludge amount (g)	0.56	1.2	1.76	1.19	1.94	1.95	1.64	1.19	1.25	0.79	1.19	2.29

4. CONCLUSION

In this study, the removal of Cr(VI) by electrocoagulation method was studied focusing on the influence of process parameters such as applied potential, the initial Cr(VI) concentration, solution conductivity and pH. The electrocoagulation unit consisted of an electrochemical reactor with two aluminum anodes and two stainless steel cathodes in parallel arrangement. Increase in potential enhanced the removal rate, but increased the energy comsumption values. Removal efficiency increased with increase in the value of conductivity and decrease in initial Cr(VI) concentration. The optimum conditions for 100% Cr(VI) removal with energy consumption of 10.17 kwh/m^3 were found as 5 V applied potential, 50 mg/L Cr(VI) concentration, 1 g/400 mL electrolyte (KCl) dose and pH value of 5. Finally, it can be concluded that electrocoagulation method is a reliable, safe, efficient, and cost effective method for removal of chromium from solutions.

REFERENCES

Ahangarnokolaei, M.A., Ganjidoust, H., Ayati, B., (2018) Optimization of parameters of electrocoagulation/ flotation process for removal of Acid Red 14 with mesh stainless steel electrodes. Journal of Water Reuse and Desalination, 8: 278-292.

Akbal F., Camcı, S., (2011) Treatment of Metal Plating Wastewater by Electrocoagulation, Environmental Progress & Sustainable Energy, 31 (3).

Akhtar, A., Aslam, Z., Asghar, A., Bellod, M. M., Raman, A.A., (2020) Electrocoagulation of Congo Red dye-containing wastewater: Optimization of operational parameters and process mechanism, Journal of Environmental Chemical Engineering, 8, 104055.

Babu, R. R., Bhadrinarayana, N.S., Meera, K.M., Begum, S., Anantharaman, N. (2007), Treatment Of Tannery Wastewater By Electrocoagulation, Journal of the University of Chemical Technology and Metallurgy, 42, 2, 2007, 201-206.

Bazrafshan, E., Mahvi, A.H., Naseri, S., Mesdaghinia, A.R., (2008) Performance Evaluation of Electrocoagulation Process for Removal of Chromium (VI) from Synthetic Chromium Solutions Using Iron and Aluminum Electrodes, Turkish J. Eng. Env. Sci.32, 59-66.

Karhu, M., Kuokkanen, V., Kuokkanen, T., Ramo, J. (2012). Bench scale electrocoagulation studies of bio oil-in-water and synthetic oil-in-water emulsions. Separation and Purification Technology, 96: 296–305.

Mollah, M.Y.A., Schennach, R., Parga, J.R., (2001) Electrocoagulation (EC) – Science And Application", Journal Of Hazardous Materials, B84, 29-41,

Mouedhen, G., Feki, M., De Petris-Wery, M., Ayedi, F. (2009) Electrochemical removal of Cr(VI) from aqueous media using iron and aluminum as electrode materials: Towards a better understanding of the involved phenomena, Journal of Hazardous Materials, 168, 983–991.

Öztürk, M. (2018) http://www.cevresehirkutuphanesi.com/assets/files/slider_pdf/PrMlkUEGGwYE.pdf

Shafaei, A., Pajootan, E., Nikazar, M., Arami, M., (2011) Removal of Co (II) from aqueous solution by electrocoagulation process using aluminum electrodes. Desalination, 279:121–126.

Sharma, D., Chaudhari, P. K., Prajapati, A. K., (2020) Removal of chromium (VI) and lead from electroplating effluent using electrocoagulation, Separation Science and Technology, 55:2, 321-331.

Tanyol, M., Ogedey, A., Oğuz, E. (2018) COD removal from leachate by electrocoagulation process: treatment with monopolar electrodes in parallel connection. Water Science & Technology, 77: 177-186.

W. Jin, H. Du, S. Zheng, Y. Zhang, Electrochemical processes for the environmental remediation of toxic Cr(VI): A review, Electrochimica Acta, 191 (2016) 1044–1055.

Zhao, X., Zhang, B., Liu, H., Qu, J. (2011). Simultaneous removal of arsenite and fluoride via an integrated electro-oxidation and electrocoagulation process. Chemosphere, 83: 726–729.

Chapter 42

PERFORMANCE ANALYSIS OF FIXED WINDOW FUNCTIONS FOR APPLICATIONS OF CONTRAST RATIO AND NONRECURSIVE DIGITAL FILTER DESIGN

Kemal AVCI[1]

1 Assoc.Prof.Dr. Kemal AVCI, Izmir Democracy University, Engineerng Faculty, Department of Electrical and Electronics Engineering, Izmir/TURKEY, kemal.avci@idu.edu.tr

INTRODUCTION

A window function (or window for short) defined as a time domain mathematical function which is nonzero for a certain time interval and zero for outside of that interval is used in many electronics engineering applications such as spectrum analysis, nonrecursive filter design, speech processing, image processing, and array processing (Prabhu, 2013; Aljahdali, 2012; Sumamo, 2013; Avci, 2018; Avci and Gumussoy, 2020).

Amplitude characteristic of a typical discrete window in the discrete time domain is generally similar to the one given in Figure-1 (here drawn for 31 length).

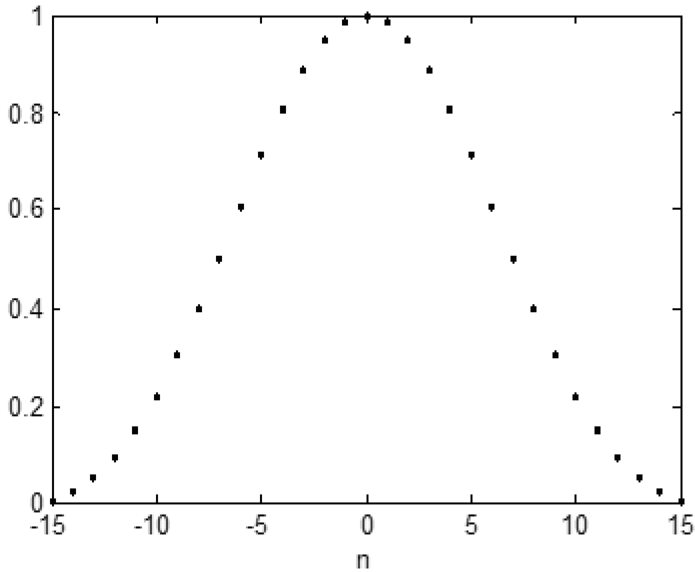


Figure-1: *Amplitude characteristic of a discrete time window function for 31 length*

Performance analysis of a window function is usually carried out in the frequency domain. Figure-2 is the amplitude characteristic of a window function in the frequency domanin. Here, the big lobe in the middle is called mainlobe and the one on the sides is called the sidelobe. Half mainlobe width (wR) and ripple ratio (R) are two important window spectral parameters, and they are used in this study. Therefore, mainlobe width of window can be defined as 2wR. And, the ripple ratio ratio of a window is the difference between the highest sidelobe amplitude value and the mainlobe amplitude in dB, and it corresponds to the parameter denoted as S1 in Figure-2. The parameter ws indicates the sampling frequency in rad/s.

Window functions in literature are classified as fixed or adjustable according to having the number of independent parameters (Saramaki, 1993).

A window function becomes fixed if it has only one independent parameter, namely window length. As for the adjustable windows, they have more than one independent parameters. Kaiser (Kaiser, 1974), dolph-chebyshev (Dolph, 1946), ultraspherical (Bergen and Antoniou, 2005), cosh (Avci and Nacaroglu, 2009), and exponential-hamming (Avci, 2016) windows are some of known adjustable windows.

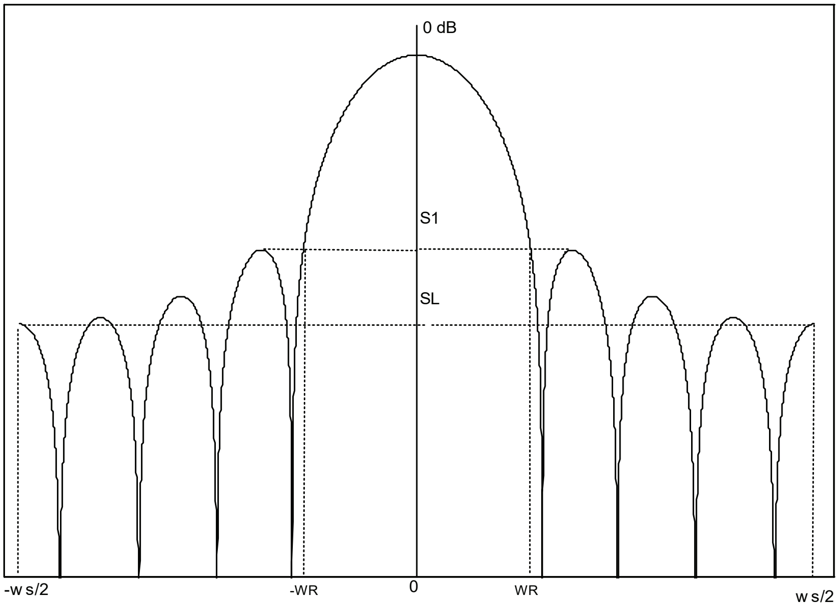


Figure-2: *Normalized amplitude spectrum of a window function*

In literature (especially Harris, 1978; Nuttall, 1981; Antoniou, 2005; Prabhu, 2013), it is observed that there are 20 fixed window functions (Coskun and Avci, 2018). These fixed window functions can be listed alphabetically as Bartlett, Bartlett-Hann, Blackman, Blackman-Harris, Bohman, Cos (x), Cos3 (x), Cos4 (x), Flat-Top, Hamming, Lanczos, Nuttall, Optimized Blackman, Parabolic, Parzen, Rectangular, Riemann, Triangular, Von-Hann, and Welch. The most known fixed window function is Hamming window.

In this study, fixed window functions defined in literature are analysed and compared for contrast ratio analysis and nonrecursive digital filter design application.

MATERIAL AND METHOD

a) Fixed window functions

In literature, it is observed that there are 20 fixed window functions having only one independent parameter, namely window length (N). A discrete window function is denoted as w(n). The time-domain mathematical definitons for 20 fixed window functions are given below in alphabetic order (Coskun and Avci, 2018). All window functions are defined for the discrete time range: $|n| \leq (N-1)/2$

Bartlett Window:

$$w(n) = \begin{cases} \dfrac{n+\frac{N-1}{2}}{\frac{N-1}{2}} & -\dfrac{N-1}{2} \leq n \leq 0 \\ 2-\dfrac{n+\frac{N-1}{2}}{\frac{N-1}{2}} & 0 \leq n \leq \dfrac{N-1}{2} \end{cases} \tag{1}$$

Bartlett-Hann Window:

$$w(n) = 0.62 - 0.48\left|\frac{n+\frac{N-1}{2}}{N-1} - 0.5\right| + 0.38\cos\left(2\pi\left(\frac{n+\frac{N-1}{2}}{N-1} - 0.5\right)\right) \tag{2}$$

Blackman Window:

$$w(n) = 0.42 + 0.5cos\frac{2\pi n}{N-1} + 0.08cos\frac{4\pi n}{N-1} \tag{3}$$

Blackman-Harris Window:

$$w(n) = 0.35875 - 0.48829\cos\left(\frac{2\pi\left(n+\frac{N-1}{2}\right)}{N-1}\right) + 0.14128cos\left(\frac{4\pi\left(n+\frac{N-1}{2}\right)}{N-1}\right) - 0.01168cos\left(\frac{6\pi\left(n+\frac{N-1}{2}\right)}{N-1}\right) \tag{4}$$

Bohman Window:

$$w(n) = \left(1 - \frac{|n|}{(N-1)/2}\right) cos\left(\frac{\pi n}{(N-1)/2}\right) + \frac{1}{\pi}\frac{sin(\pi|n|)}{(N-1)/2} \tag{5}$$

Cos(x) Window:

$$w(n) = cos\left(\frac{\pi n}{N-1}\right) \tag{6}$$

Cos³(x) Window:

$$w(n) = 0.75cos\left(\frac{\pi n}{N-1}\right) + 0.25cos\left(\frac{3\pi n}{N-1}\right) \tag{7}$$

Cos⁴(x) Window:

$$w(n) = 0.375 + 0.5cos\left(\frac{2\pi n}{N-1}\right) + 0.125cos\left(\frac{4\pi n}{N-1}\right) \tag{8}$$

Flat-Top Window:

$$w(n) = 0.21557895 - 0.41663158\cos\left(\frac{2\pi\left(n+\frac{N-1}{2}\right)}{N-1}\right) + 0.277263158cos\left(\frac{4\pi\left(n+\frac{N-1}{2}\right)}{N-1}\right) - 0.083578947cos\left(\frac{6\pi\left(n+\frac{N-1}{2}\right)}{N-1}\right) + 0.006947368cos\left(\frac{8\pi\left(n+\frac{N-1}{2}\right)}{N-1}\right) \tag{9}$$

Hamming Window:

$$w(n) = 0.54 + 0.46cos\frac{2\pi n}{N-1} \tag{10}$$

Lanchos Window:

$$w(n) = sinc\left(\frac{2n}{N-1}\right) \tag{11}$$

Nuttall Window:

$$w(n) = 0.3635819 - 0.4891775\cos\left(\frac{2\pi\left(n+\frac{N-1}{2}\right)}{N-1}\right) + 0.1365995cos\left(\frac{4\pi\left(n+\frac{N-1}{2}\right)}{N-1}\right) - 0.106411cos\left(\frac{6\pi\left(n+\frac{N-1}{2}\right)}{N-1}\right) \tag{12}$$

Optimized Blackman Window:

$$w(n) = 0.412 + 0.5cos\frac{2\pi n}{N-1}s + 0.088cos\frac{4\pi n}{N-1} \tag{13}$$

Parabolic Window:

$$w(n) = 1 - \left(\frac{2n}{N-1}\right)^2 \tag{14}$$

Parzen Window:

$$w(n) = \begin{cases} 1 - 6\left[\frac{n}{(N)/2}\right]^2 + 6\left[\frac{n}{(N)/2}\right]^3 & 0 \le |n| \le \frac{N-1}{4} \\ 2\left[1 - \frac{|n|}{N/2}\right]^3 & \frac{N-1}{4} \le |n| \le \frac{N-1}{2} \end{cases} \tag{15}$$

Rectangular Window:

$$w(n) = 1 \tag{16}$$

Riemann Window:

$$w(n) = \frac{sin\left(\frac{2\pi n}{N-1}\right)}{\frac{2\pi n}{N-1}} \tag{17}$$

Triangular Window:

$$w(n) = 1 - \frac{|2n|}{N} \tag{18}$$

Von-Hann Window:

$$w(n) = 0.5 + 0.5cos\frac{2\pi n}{N-1} \tag{19}$$

Welch Window:

$$w(n) = 1 - \left(\frac{n - \frac{N-1}{2}}{\frac{N-1}{2}}\right)^2 \tag{20}$$

b) Contrast Ratio Computation

One of the most important parameters used to measure the quality of an image obtained from an imaging system is the contrast ratio. Contrast ratio is defined as the difference between the whitest and blackest pixel values in terms of brightness, and a high contrast ratio is desired in images as it always offers better quality (Gonzales and Woods, 2007). There is a direct relationship between the contrast ratio of the images obtained in imaging systems and the characteristics of the window functions used to create these images.

The relationship between a window function and contrast ratio (CR) is defined by Eq. (21)

$$CR = \frac{w^T w}{w^T Q w} \tag{21}$$

Here, CR is the contrast ratio, w is the window function coefficients vector, Q is the matrix with the elements q(n, m) defined by Eq. (22).

$$q(n,m) = \begin{cases} -\frac{w_R}{\pi} sinc[w_R(m-n)] & m \neq n \\ 1 - \frac{w_R}{\pi} & m = n \end{cases} \tag{22}$$

c) Digital filter design by windowing method

Digital filters are very important systems in electronics engineering. They are mainly used for noise reduction and signal enhancement applications. They can be classified as nonrecursive or recursive. Nonrecursive digital filters are attractive for applications because they are always stable and can be designed to have exact linear phase characteristic. Windowing method is the most straightforward way to design nonrecursive digital filters (Antoniou, 2005).

There are four genereal method to design nonrecursive digital filters. These are windowing, frequency sampling, numerical, and optimization methods. Windowing method is the most straightforward one to design the nonrecursive digital filters. This method uses window functions to obtain the practical finite impulse response, h(n), of the digital filter from the infinite impulse response, hid(n), of the ideal prototype digital filter. The method is briefly defined by the following equation (Antoniou, 2005)

$$h(n) = w(n) h_{id}(n) \tag{23}$$

In this study, as a prototype digital filter, lowpass filter type is chosen because other types of filters such as highpass, bandpass, and bandstop can be obtained from the lowpass filter by the related transformation. For a lowpass filter, the ideal impulse response of the digital filter is defined by the equation below.

$$h_{id}(n) = \begin{cases} \dfrac{\omega_c}{\pi} & for \quad n = 0 \\ \dfrac{sin\omega_c n}{n\pi} & for \quad n \neq 0 \end{cases} \tag{24}$$

where ω_c is cut-off frequency of the filter.

A lowpass filter can be expressed by the amplitude response shown in Figure 3 (Avci, 2008).

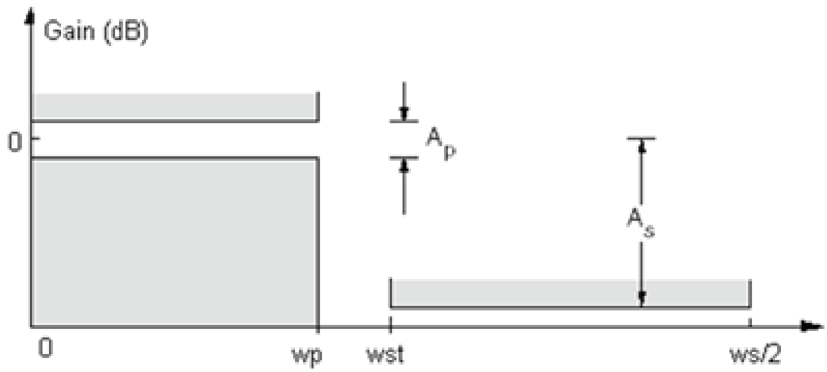


Figure-3. *Lowpass filter amplitude specifications*

The parameters in Figure 3 are

w_p = passband frequency

w_{st} = stopband frequency

w_s = sampling frequency

A_p = maximum allowable attenuation in passband

A_s = minimum allowable attenuation in stopband

Also, the cut-off frequency (w_c) and the transition width (Δw) are defined as

$$w_c = (w_{st} + w_p)/2 \tag{25}$$

$$\Delta w = w_{st} - w_p \tag{26}$$

Since the ripples in the filters designed by the windowing are almost the same in passband and stopband regions, only one parameter is considered in the analysis. In this study, the parameter As is chosen as presenting the ripples.

SIMULATION RESULTS

In this section, the simulation results of contrast ratio analysis and designing nonrecursive lowpass digital filters by using 20 fixed windows for different filter lengths are given. All simulation studies were performed on the Matlab 2017a software platform.

a) Results for Contrast Ratio

In this section, performance analysis of 20 fixed window functions defined in the literature is carried out in terms of contrast ratio. In addition, the analysis of fixed windows in terms of mainlobe width and ripple ratio, which are the basic two window spectral parameters, are also presented and the relationship of these parameters with the contrast ratio is examined in the study.

Contrast ratio results of fixed window functions are given in Table I for N = 21 and 51 lengths, and Table II for N = 81 and 111 lengths. The results showed that the Blackman-Harris window function provided the highest contrast ratio at all lengths, followed by the Nuttall and Flat-Top window functions. The results also showed that the Rectangular window function provided the lowest contrast ratio at all lengths, followed by the Parabolic and Welch window functions.

When Tables I and II are examined, it is seen that window functions with wide half mainlobe width (hence mainlobe width) and low ripple ratio have higher contrast ratio. Similarly, window functions with narrow half mainlobe width and high ripple ratio were observed to have low contrast ratio.

Table I. *Contrast ratio results for different lengths of fixed windows for N = 21 and 51*

Window Function	N = 21			N = 51		
	w_R (rad/s)	R (dB)	CR (dB)	w_R (rad/s)	R (dB)	CR (dB)
Bartlett	0.508	-25.93	24.10	0.204	-26.44	24.83
Bartlett-Hann	0.597	-35.97	34.74	0.238	-35.90	35.03
Blackman	0.888	-58.26	57.44	0.355	-58.13	57.03
Blackman-Harris	1.241	**-91.10**	**90.83**	0.496	-92.11	**88.77**

Bohman	0.850	-45.97	46.39	0.340	-46.01	46.44
Cos(x)	0.424	-22.74	22.22	0.170	-22.97	22.57
$Cos^3(x)$	0.749	-39.29	41.35	0.299	-39.30	41.35
$Cos^4(x)$	0.908	-46.75	49.50	0.363	-46.74	49.49
Flat-Top	0.499	-79.73	78.89	0.199	-86.46	80.66
Hamming	0.620	-40.64	34.10	0.244	-42.33	34.37
Lanczos	0.472	-26.14	26.04	0.189	-26.36	26.41
Nuttall	1.253	-85.46	81.11	0.501	**-92.29**	83.42
Optimized Blackman	0.918	-62.86	61.89	0.367	-62.60	61.24
Parabolic	0.400	-21.05	20.30	0.160	-21.27	20.65
Parzen	0.973	-53.02	52.01	0.401	-53.04	52.12
Rectangular	0.243	-13.20	9.91	0.100	-13.26	9.88
Riemann	0.472	-26.14	25.07	0.189	-26.36	25.21
Triangular	0.462	-26.04	24.24	0.196	-26.44	24.84
Von-Hann	0.588	-31.48	32.61	0.235	-31.48	32.58
Welch	0.381	-21.08	20.33	0.157	-21.26	20.64

Table II. *Contrast ratio results for different lengths of fixed windows for N = 81 and 111*

Window Function	N = 81			N = 111		
	w_R (rad/s)	R (dB)	CR (dB)	w_R (rad/s)	R (dB)	CR (dB)
Bartlett	0.128	-26.50	24.91	0.093	-26.54	24.93
Bartlett-Hann	0.149	-35.91	35.04	0.108	-35.94	35.02
Blackman	0.222	-58.17	57.02	0.161	-58.11	56.99

Blackman-Harris	0.309	-92.14	**88.92**	0.225	-92.05	**89.00**
Bohman	0.212	-46.02	46.43	0.154	-46.01	46.41
Cos(x)	0.106	-22.99	22.60	0.077	-23.08	22.61
$Cos^3(x)$	0.187	-39.34	41.33	0.136	-39.43	41.30
$Cos^4(x)$	0.227	-46.80	49.51	0.165	-46.74	49.47
Flat-Top	0.124	-89.35	81.69	0.090	-90.85	82.10
Hamming	0.151	-42.58	34.38	0.110	-42.66	34.38
Lanczos	0.118	-26.44	26.44	0.086	-26.48	26.50
Nuttall	0.313	**-95.08**	84.39	0.228	**-96.39**	84.77
Optimized Blackman	0.229	-62.61	61.20	0.166	-62.64	61.19
Parabolic	0.100	-21.34	20.67	0.073	-21.29	20.68
Parzen	0.252	-53.05	52.12	0.184	-53.05	52.12
Rectangular	0.063	-13.26	9.87	0.046	-13.28	9.85
Riemann	0.118	-26.44	25.23	0.086	-26.48	25.25
Triangular	0.124	-26.50	24.91	0.091	-26.51	24.94
Von-Hann	0.147	-31.48	32.59	0.106	-31.47	32.48
Welch	0.099	-21.29	20.67	0.072	-21.29	20.68

Since the best CR values are obtained from Blackman-Harris, Nuttall, and Flat-Top, respectively, a detailed comparison among these windows is performed for several characteristics in terms of window length.

Figure 4 shows the half mainlobe width characteristics for these three windows for window lengths from N = 11 to N = 121. It is seen that while Blackman-Harris and Nuttall windows almost have the same characteristics, Flat-Top window has always lower mainlobe widths, i.e. the best values. As expected, while window length increases mainlobe width for all windows becomes narrower.

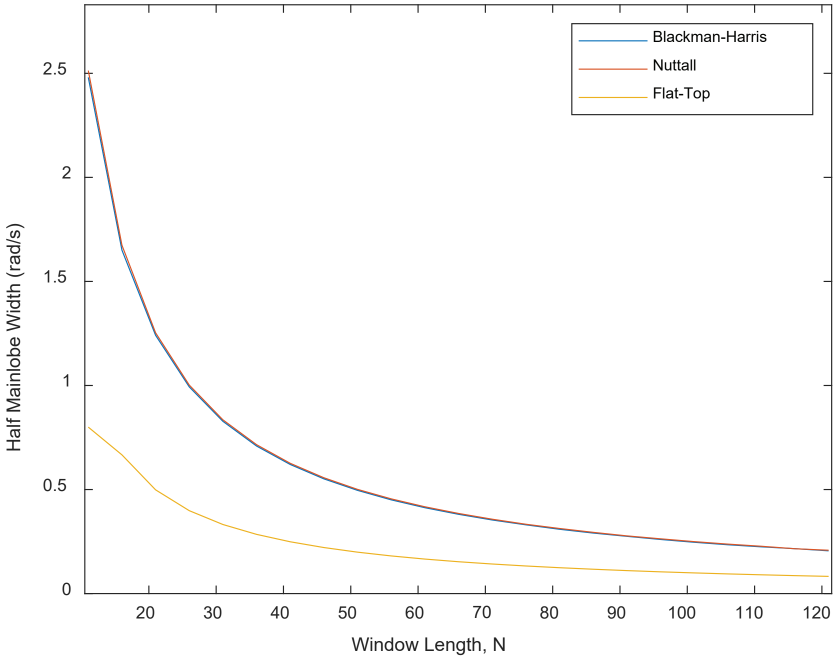


Figure-4. *Half mainlobe width characteristics for three windows for N = 11 to 121.*

Figure 5 shows the half mainlobe width characteristics for these three windows for window lengths from N = 11 to N = 121. It is seen that while Blackman-Harris window has the lowest ripple ratio (i.e., the best one) for N small than 50, Nuttall window has the lowest ones for N larger than 50. Flat Top window has the worst ripple ratio characteristic amog these three windows. expected, while window length increases mainlobe width for all windows becomes narrower.

Figure 6 shows the contrast ratio characteristics for these three windows for window lengths from N = 11 to N = 121. It is seen that Blackman-Harris window has the highest values (i.e., the best ones) for all window length. Nuttall window performs better values than Flat-Top window but worse than Blackman-Harris window.

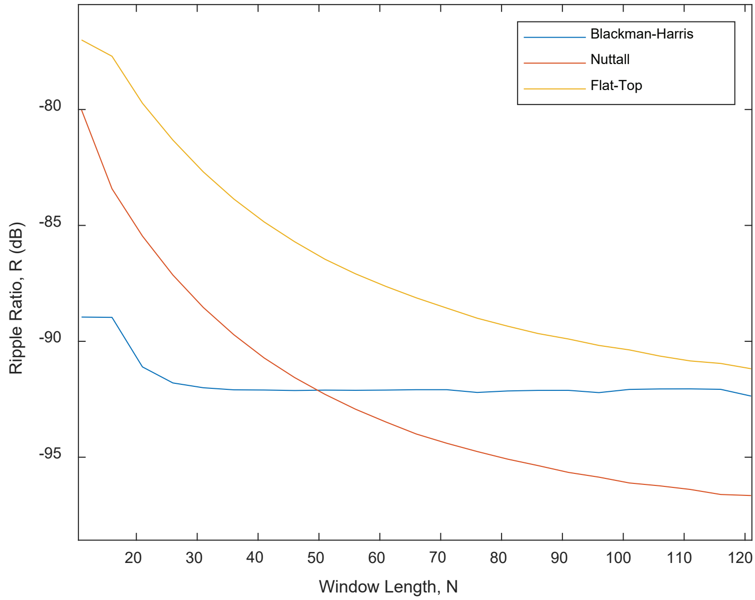


Figure-5. *Ripple ratio characteristics for three windows for N = 11 to 121.*

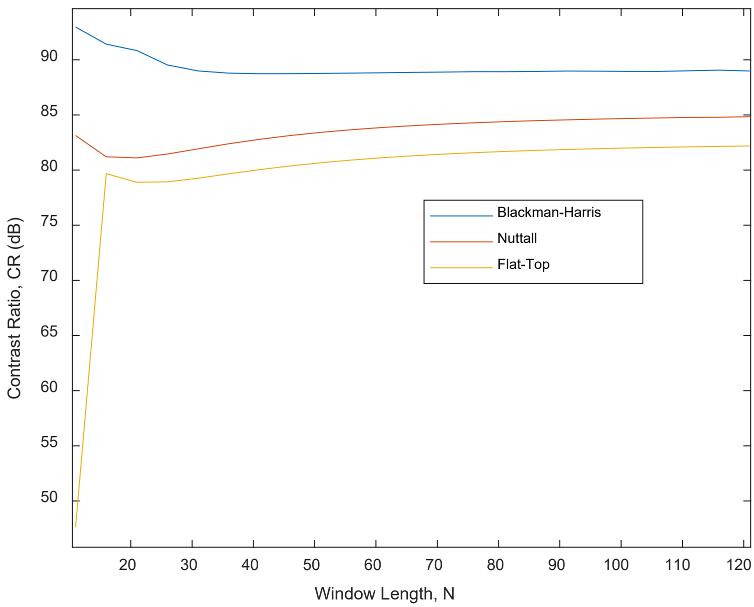


Figure-6. *Contrast ratio characteristics for three windows for N = 11 to 121.*

b) **Results for Nonrecursive Digital Filter Design**

In this section, the simulation results of designing lowpass digital filters by using 20 fixed windows for different filter lengths are given.

As a specific design example, if a 51-length Hamming window shown in Figure 7 is used to design a digital filter having a cut-off frequency $\omega_c = 0.4\pi$ rad/s, then the lowpass digital filter shown in Figure 8 is obtained. From Figure 8, the designed filter has a minimum stopband attenuation, As = 43 dB, and a transition width, $\Delta w = 0.03$ rad/s.

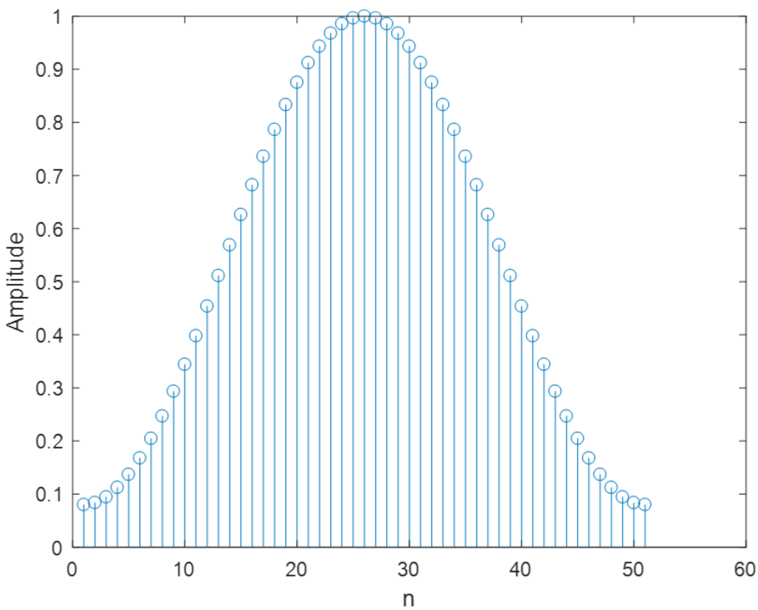


Figure-7. *Plot of Hamming window for N = 51*

Similarly, digital filters are designed by using other fixed window functions for N = 11, 51, 101, and 151. Table III and Table IV show the results in terms of minimum stopband attenuation and transition width, respectively.

From Table III, it is seen that the best As (the largest value) is obtained for Nuttall windows for N = 51, 101, and 151 whereas for N = 11 Optimized Blackman window provides the best As. Also, it is seen from Table III that some windows don't have As for N = 11 because they don't have the ripples in the stopband for this filter length. And, the worst As is presented by the rectangular window for all filter lengths.

From Table IV, it is seen that the best Δw (the smallest value) is obtained for rectngular windows for all filter lengths. Also, it is seen from Table IV that some windows don't have Δw for N = 11 because thay don't have the ripples in the passband and stopband for this filter length. And, the worst Δw (the largest) is presented by the Flat-Top window for lengths N = 51, 101, and 151 whereas for N = 11 Cosx window provides the worst Δw.

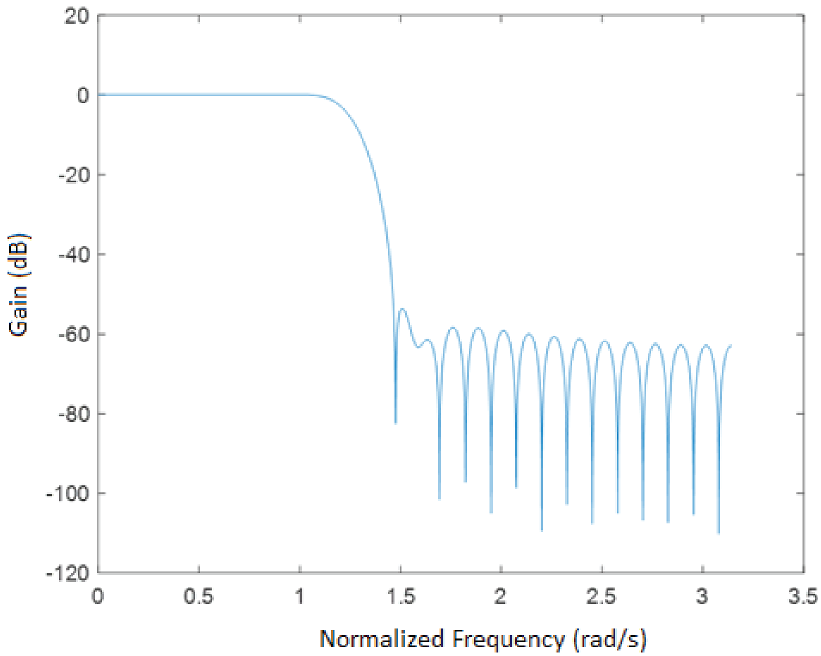


Figure-8. *Spectrum of the lowpass digital filter designed by the Hamming window N = 51*

Table III. *Analysis of designed digital filters by using fixed window functions in terms of minimum stopband attenuation for a cut-off frequency $\omega_c = 0.4\pi$ rad/s.*

Window Function	A_s (dB)			
	N = 11	N = 51	N = 101	N = 151
Bartlett	-	26.38	26.35	26.33
Bartlett-Hann	36.97	39.79	39.71	39.67
Blackman	-	75.30	75.29	75.29
Blackman-Harris	-	109.24	109.28	109.29
Bohman	-	51.95	51.95	51.95

Cos(x)	35.39	33.85	33.80	33.80
$Cos^3(x)$	55.35	52.97	52.97	52.97
$Cos^4(x)$	55.07	61.36	61.36	61.36
Flat-Top	-	114.59	114.45	114.55
Hamming	53.45	53.66	53.86	53.96
Lanczos	40.01	38.57	38.53	38.52
Nuttall	-	**114.87**	**114.79**	**114.66**
Optimized Blackman	**75.20**	79.14	79.12	79.12
Parabolic	32.85	31.44	31.40	31.39
Parzen	-	56.66	56.64	56.64
Rectangular	20.54	20.98	20.98	20.98
Riemann	40.01	38.57	38.53	38.52
Triangular	24.94	26.38	26.35	26.33
Von-Hann	43.41	43.95	43.94	43.95
Welch	30.98	31.36	31.37	31.38

Table IV. *Analysis of designed digital filters by using fixed window functions in terms of transition width for a cut-off frequency* $\omega_c = 0.4\pi$.

Window Function	Δw (rad/s)			
	N = 11	N = 51	N = 101	N = 151
Bartlett	-	0.502	0.253	0.166
Bartlett-Hann	-	0.432	0.216	0.144
Blackman	-	0.725	0.363	0.242
Blackman-Harris	-	0.982	0.492	0.328
Bohman	-	0.755	0.377	0.249
Cos(x)	1.463	0.267	0.134	0.089
$Cos^3(x)$	-	0.534	0.267	0.178
$Cos^4(x)$	-	0.664	0.332	0.221

Flat-Top	-	1.222	0.608	0.405
Hamming	-	0.420	0.210	0.140
Lanczos	-	0.312	0.156	0.104
Nuttall	-	0.980	0.490	0.327
Optimized Blackman	-	0.695	0.347	0.232
Parabolic	1.295	0.244	0.122	0.081
Parzen	-	0.925	0.489	0.333
Rectangular	0.584	0.123	0.062	0.041
Riemann	-	0.312	0.156	0.104
Triangular	1.457	0.448	0.237	0.166
Von-Hann	-	0.402	0.201	0.134
Welch	1.102	0.239	0.121	0.081

CONCLUSION

In this study, comparative analyzes of 20 fixed window functions were presented for contrast ratio and nonrecursive digital filter design applications. Simulation analyzes for different window lengths were performed for these studies on the Matlab 2017a software platform to observe the performances of the fixed windows against the window length.

As a first study, contrast ratio values of fixed window functions for lengths N = 21, 51, 81 and 111 were calculated. In addition, by providing analyzes of fixed window functions in terms of the fundamental two window spectral parameters, namely analob width and mainlobe width, the relationship of these parameters with the contrast ratio was examined in this study. The results of the simulation showed that for all length values, firstly Blackman Harris, then Nuttall and Flat-Top window functions provided the highest contrast ratio values. However, the lowest contrast ratio values were obtained primarily by the Rectangular, then Parabolic and Welch window functions. The simulation results also showed that window functions with wide mainlobe width and low ripple ratio present higher contrast ratio. Also, a detailed comparison for different characteristics among Blackman Harris, then Nuttall, and Flat-Top windows were also presented for N = 11 to 121.

As a second study, nonrecursive lowpass digital filters were obtained for N = 11, 51, 101, and 151. The simulation results demonstrated that best digital filters in terms of As were obtained by using Nuttall, Flat Top, and

Blackman-Harris window functions, respectively. And, the worst filters in terms of As were obtained by using Rectangular, Bartlett, and Triangular windows. Designed filters were also compared in terms of transition width. The results showed that the best window functions in terms of Δw are Rectagular, Welch, and Parabolic window functions. And worst filters in terms of Δw were obtained by using Flat Top, Nuttall, and Blackman Harris window functions.

REFERENCES

Aljahdali S., Nofal M., Albagory Y. (2012). A modified array processing technique based on kaiser window for concentric circular arrays. *International Conference on Multimedia Computing and Systems (ICMCS)*, 458-461, Tangier, Morocco.

Antoniou A. (2005). Digital signal processing: signal, systems, and filters, *McGraw-Hill*, New York, USA.

Avci K. (2008). Design of high-quality low order digital filters by using window functions, *Natural and Applied Institute*, Gaziantep University, PhD Thesis, Turkey.

Avci K. (2014). Kaiser-Hamming window and its performance analysis for nonrecursive digital filter design, *Journal of the Faculty of Engineering and Architecture of Gazi University*, 29(4), 823-833.

Avci K. (2016). Design of FIR filters using exponential-hamming window family. *Turkish Journal of Electrical & Computer Sciences*, 24(4), 2513-2524.

Avci K. (2018). Two dimensional digital filter design using kaiser-hamming window structure and huang transform and image enhancement application. *Journal of the Faculty of Engineering and Architecture of Gazi University*, 33(4), 1459-1474.

Avci K., Gumussoy E. (2020). Design, analysis, and ECG subband processing application of new m-channel cosine modulated uniform filter banks based on exponential window family. *Journal of the Faculty of Engineering and Architecture of Gazi University*, 35(1), 193-211.

Avci K., Coskun O. (2018). Spectral performance analysis of cosh window based new two parameter hybrid windows. *26th Signal processing and Communications Applications Conference (SIU2018),* İzmir, Turkey.

Avci K., Nacaroglu A. (2009). Cosh window family and its application to fir filter design. *AEU-Int J Electron C*, 11, 907-916.

Bergen S. W., Antoniou A. (2005). Design of nonrecursive digital filters using the ultraspherical window function. *Eurasip J Adv Sig Pr*, 12, 1910-1922.

Dolph C. L. (1946). A current distribution for broadside arrays which optimizes the relationship between beamwidth and side-lobe level. *Proc. IRE*, 34, 335-348.

Gonzalez R. C., Woods R. E. (2007). *Digital Image Processing*, 3rd. Edition, Pearson, 2007.

Harris F. J. (1978). On the use of windows for harmonic analysis with the discrete fourier transform. *Proceedings of the IEEE*, 66.

Kaiser, J. F. (1974). Nonrecursive digital filter design using I_0-sinh window function. *Proceedings of IEEE Int Symp Circuits and Systems (ISCAS'74)*, 20-23, San Francisco, Calif, USA.

Nuttall A. H. (1981). Some windows with very good sidelobe behavior. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-29, 84–91.

Prabhu K. M. M. (2014). Window functions and their applications in signal processing, CRC Press, Boca Raton, USA.

Saramaki T. (1993). Finite impulse response filter design. in: Mitra K, Kaiser JF, editors. *Handbook for Digital Signal Processing*. New York, NY, USA: Wiley & Sons.

Sumamo L. (2013). Handwritten word segmentation using kaiser window, *International Conference on Quality in Research (QIR)*, 73-78, Yogyakarta, Endonezya.

Chapter 43

THE EFFECT OF DIFFERENT MATERIALS ON PROPERTIES OF HORASAN MORTAR

Aylin AKYILDIZ[1]

1 TNKÜ, Çorlu Mühendislik Fakültesi, İnşaat Mühendisliği Bölümü, aakyildiz@nku.edu.tr

Introduction

The social legacy within the universe and their conservation for the following era is pivotal. Numerous verifiable structures proceed to offer administrations and verifiable values; and as such, it is vital to get it their auxiliary behavior and collapse component when creating an arrangement for their conservation. Modern advances that can be engaging to authentic structures are rising with current progressions in materials and development procedures [1, 2].

Horasan mortars have been employed with distinctive names in numerous chronicled buildings since antiquated societies. Nowadays, we realize that rebuilding of such chronicled buildings can be accomplished by utilizing materials comparing with the old times. Subsequently, the enhancement of antiquated mortars applied in reclamation of authentic buildings is critical. Mortars containing lime have been promoted broadly since the antiquated Romans and Hassock [3-5]. Horasan mortars have been made used of in Footrest buildings particularly in 15th centuries [6, 7]. In common, mortars were handled in advancement of wells and showers in Roman, Seljuk surviving up to date [8, 9] .

Mortars and plasters of numerous Roman, Seljuk and Footrest period buildings were created by utilizing lime as folio and pulverized bricks altogether. These mortars are called as "horasan" of Turkiye [10]. Horasan mortars have been broadly adopted as waterproof in reservoir conduits, bridges and showers due to their pressure driven properties [11, 12]. Among the chronicled mortars having come to outlive till nowadays, gypsum, lime and lime pozzolan have been benefitted as official materials. As total materials, stream sand, stones, brick pieces and powder have been utilized in conjunction with feed, horse hair and goat hair serving as strands. In our homeland, we experience Horasan mortar with shifting blend proportions in buildings from Byzantion, Seljuk and Hassock periods [13, 14].

Amid the Hassock times, lime was blended with water to create slaked lime and after that warmed and filtered to utilize in mortars [15]. The lime responds with brick tidy to deliver calcium silicate hydrates. The water-absorption of permeable brick clean permits mortars to pick up attachment and quality in dampness, and adaptability amid the drying and settling process. Lime is classified into a few bunches agreeing to the generation prepare, such as quicklime, pressure driven lime, and hydrated lime. Generally, nonhydraulic lime was appropriated in old mortars, though water powered and hydrated mortars are mechanical mortars at the most recent century requiring particular warming, pulverization, and generation strategies [16-21].

In this work, bricks exploited in horasan mortars and plasters are inspected to characterize their features for potential usage at preservation work. The reason for this thinking about is to analyze the improvement of materials which were resorted in chronicled buildings to organize to their properties and to consider the properties at unique materials. This reasoning was carried out to explore the properties of the horasan mortar gotten by appropriating lime and squander brick clean as the most fabric and grain water as blending water.

1. Materials

1.1. Sand

Sand produced in Limak A.S. The grading of sand was shown in Table 1.

Table 1. The Grading of Sand

Sieve (mm)	2.00	1.60	1.00	0.50	0.16	0.08
Percentage (%)	0	7 ± 5	33 ± 5	67 ± 5	87 ± 5	99 ± 1

1.2. Cement

Cement was CEM II 42.5 R. Chemical compositions of cement, mineral components and mechanical properties were given in Table 2, 3 and 4, respectively. Its gravity was about 3.14 g/cm^3.

Table 2. Chemical Properties of cement

Component	CaO	MgO	SiO_2	Al_2O_3	Fe_2O_3	K_2O	Na_2O	SO_3	Loss on ignition
Cement, %	61.66	1.31	21.44	4.96	3.90	0.25	0.43	3.38	0.96

Table 3. Mineral components

C_3S	33.37
C_2S	36.37
C_3A	6.55
C_4AF	11.87

Table 4. Mechanical Properties

Compressive Strength (MPa)	2	days	28.0
	7		41.3
	28		60.5

1.3. Brick powder (BD)

The specific gravity of brick dust was 2.76. The chemical properties of the waste brick dust were seen in Table 3. Picture of brick dust was seen Figure 1.

Table 5. Chemical properties of brick dust

	SiO_2	CaO	Al_2O_3	Fe_2O_3	K_2O	MgO	Na_2O	SO_3	TiO_2	LOI
Brick dust	52.0	4.3	40.0	0.7	0.49	0.11	0.04	0.31	0.67	1.38

Figure 1. Picture of brick dust

1.4. Hydrated Lime (L)

The specific gravity of lime was 2.34 g/cm^3. The bulk density was 480 kg/m^3. Available Lime Index as $Ca(OH)_2$ was 89%.

2. Methods

In this study, the effect of using different materials over the properties of horasan mortar was investigated. For this purpose, lime, brick dust, cement and water were used in horasan mortar. Mix amounts of the mortars are given in Table 6. The fresh mortars were filled into the 40x40x160 mm molds. After 24 h, samples were cured in water until the 28 days.

Table 6. Mix proportions of mortars

	Sand, g (S)	Cement, g (C)	Brick dust, g (BD)	Lime, g (L)	Water, g (W)
Control	1350	450	-	-	225
BD	1350	405	45	-	225
L	1350	405	-	45	225
BD+L	1350	360	45	45	225

2.1. Flexural strength

The test of the composite was **test**ed according to Turkish standard (TS EN 196-1) [22]. Strength was calculated according to Eq. (1);

$$R_f = (1.5 \times F_f \times \ell)/ b^3, \tag{1}$$

R_f is the flexural strength in MPa;

F_f is load in N;

b is the side of the prism in mm;

ℓ is the distance in mm.

2.2. Compressive strength

The strength test was applied with TS EN 196-1 [22]. The one broken parts of the specimen retained after test were used. The tests were applied on cured specimens. Compressive strength was calculated according to Eq. (2);

$$R_c = F_c / 6400, \qquad (2)$$

In the equations above

R_c is the compressive strength in MPa

F_c is the load in N.

6400 is area in mm

2.3. Ultrasonic pulse velocity

In order to assess the quality of the mortar an ultrasonic test was carried out in accordance with TS EN 12504-4 [23]. Ultrasonic pulse velocity test was calculated Eq. (3);

$$v = (L / t) \times 1000, \qquad (3)$$

In the equations above: v is the pulse velocity in m/s;

L is the distance in mm;

t is the time travelled in μs.

3. Results

The strength was the most mechanical properties in the quality of concrete compared with other properties. Figure 2 showed the strength of BD cement mortar at 28 days curing. After 28 days, strength of the samples with BD were in the range 20.00-29.00 MPa. The highest strength of 29 MPa demonstrated by reference samples. After curing, it was observed that the increase of the percentage BD decreased the compressive strength compared to the control mortars, indicating that %10 is the optimum mix proportions for strength.

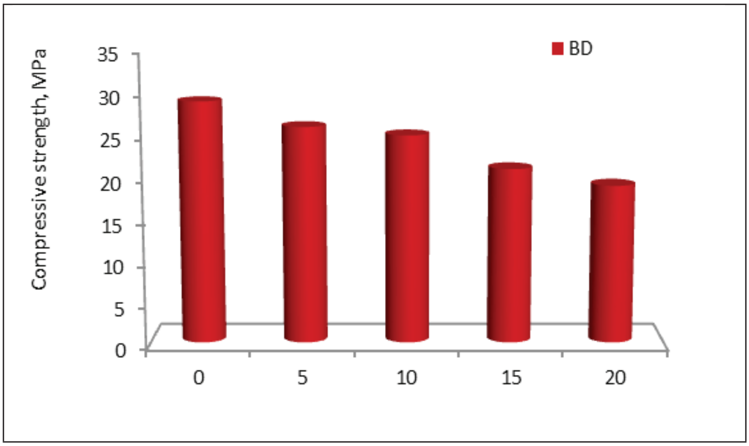
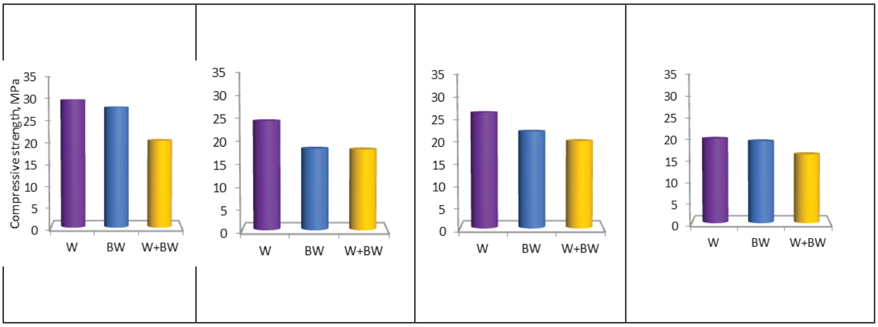


Figure 2. Compressive strength

Mortar is a mixture of the sand and matrix, and it is the interface between these materials. For the strength could be supposed that it is more sensitivity of the transition zone between sand and matrix. The increase in strength of cement based composite containing pozzolan materials play an important role in improving the sand-paste bond and formation of C-S-H. Thus, mortar with BD in water mixing ensuring a better contribution to achieve adequate interfaced to the strength.

The results show that mixing mortar with lime improves the compressive strength of mortar according to brick dust. It can be shown in Figure 3 that the addition of BD+L lowered the strength compared other materials. It can be seen that the mixes with L indicated similar values in contrast with that of the BD and BD+L. The results indicate that the effect of L on compressive strength of mortar was significant compared BD.

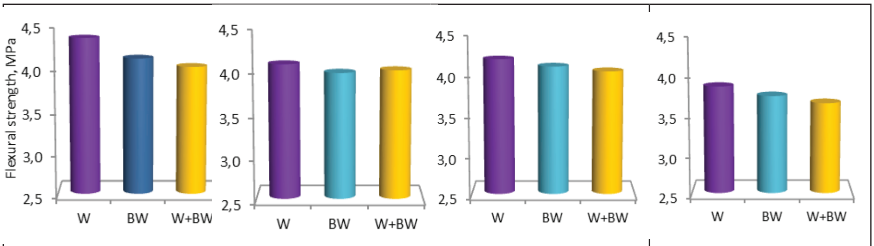


a. **Reference** b. Brick dust c. Lime d. Brick dust + lime

Figure 3. Compressive strength

3.2. Flexural strength

Flexural strength of cement based composite were presented in Figure 4. Figure 4 showed the strength of BD cement mortar at 28 days of curing. After 28 days, strength of the samples with BD were in the range 4.36-3.70 MPa. The highest strength of 4.36 MPa demonstrated by reference samples.



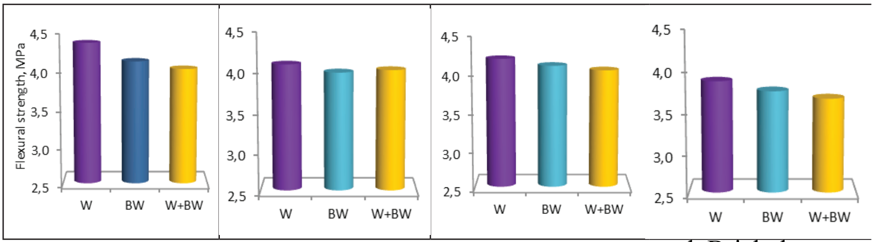
a. Reference b. Brick dust c. Lime d. Brick dust + lime

Figure 4. Flexural strength

The results show that mixing mortar with W improves the flexural strength of mortar. The highest flexural strength demonstrated by reference sample. According to Figure 4b, the average flexural strength of the BD and BD+L mixes was approximately 3.3% and 10% lower than the reference samples, respectively. It can be seen that the L mixes indicated similar values in contrast with that of the BD and BD+L mixes. According to Figure 4c, except for one outlier in which the 28 days flexural strength values were approximately the same, the BD+L samples had flexural strength values at least 10% lower than the other samples. The highest flexural strength of 4.04 MPa demonstrated by L sample.

3.3. Ultrasonic pulse velocity

The ultrasonic pulse velocity results were shown in Figure 5. Figure 5 signifies ultrasonic pulse velocity of composite decreases while the replacement ratio of material increases. During the test, ultrasound transfer inside the material through pores which provides information about the porosity of materials in cement based composite applications. Considering graph, it can be concluded that incorporation of BD and L in cement based composite adversely affects the composite in relation to its ultrasonic pulse velocity.



a. Reference b. Brick dust c. Lime d. Brick dust + lime

Figure 5. Ultrasonic pulse velocity

The quality classification of the concrete is done as follows; a value of above 4.5 km/s points excellent quality, a value varying between 3.5 and 4.5 km/s indicates good concrete; a value between 3.0 and 3.5 km/s indicates medium quality of concrete [24]. As can be seen in Figure 5, the ultrasonic pulse velocity results at samples ranged from 3.0 km/s to 4.57 km/s considered as "good concrete" quality of concrete.

Conclusions

In this study, the effect of using different materials over the properties of horasan mortar was investigated. For this purpose, lime, waste brick dust, as binder cement and water, as mixing water were used in horasan mortar mixture. Following conclusions can be drawn from the experimental studies and analyses mentioned above.

∗ The results show that mixing mortar with lime improves the compressive strength of mortar according to brick dust.

∗ According to Figure 4b, the average flexural strength of the BD and BD+L mixes was approximately 3.3% and 10% lower than the reference samples, respectively. It can be seen that the L mixes indicated similar values in contrast with that of the BD and BD+L mixes.

∗ The ultrasonic pulse velocity results at samples ranged from 3.0 km/s to 4.57 km/s considered as "good concrete" quality of concrete.

References

[1] F. Cakir & H. Uysal, Experimental modal analysis of brick masonry arches strengthened prepreg composites, Journal of Cultural Heritage, 2015, 16 (3), 284-292.

[2] C. Oguz, F. Turker & U. Kockal, Construction Materials Used in the Historical Roman Era Bath in Myra.

Scientific World Journal, 2014, Volume 2014, Article ID 536105.

[3] B. Işıkdağ & İ. B. Topçu, The effect of ground granulated blast-furnace slag on properties of Horasan mortar. Construction and Building Materials, 2013, 40, 448-454.

[4] B. Işıkdağ & İ. B. Topçu, Improvement of Khorasan mortar with fly ash for restoration of historical buildings, Science and Engineering of Composite Materials, 2014, 21(3), 359–367.

[5] A. Tironi, A. M. Trezza, A. N. Scian & E. F. Irassar, Assessment of pozzolanic activity of different calcined clays. Cement and Concrete Composites, 2013, 37(1), 319–327.

[6] H. Böke, S. Akkurt, B. Ipekoglu, & E. Uğurlu, Characteristics of brick used as aggregate in historic brick-lime mortars and plasters, Cement and Concrete Research, 2006, 36(6), 1115-1122.

[7] H. Böke, S. Akkurt & B. Ipekoglu, Investigation of the Pozzolanic Properties of Bricks Used in Horasan Mortars and Plasters in Historic Buildings, Key Engineering Materials, 2004, 264-268, 2399-2402.

[8] S. S. Tunçoku, Ph.D. Thesis, Characterization of Masonry Mortars Used in Some Anatolian Seljuk Monuments in Konya, Beyşehir and Akşehir, Middle East Technical University, Ankara, 2001.

[9] G. Biscontin, M.P. Birelli & E. Zendri, Characterization of binders employed in the manufacture of Venetian Historic Mortars, Journal of Cultural Heritage, 2002, 3(1), 31-37.

[10] A. Moropoulou, A. Bakolas & K. Bisbikou, Investigation of the Technology of Historic Mortars, Journal of Cultural Heritage, 2000, 1, 45-48.

[11] E. Uğurlu, M.Sc. Thesis, Characterization of Horasan Plasters From Some Ottoman Baths In Izmir, İzmir Institute of Technology, 2005.

[12] C. Z. Oğuz, M.Sc. Thesis, Evaluation of Masonry Wall Materials of Byzantine and Early Ottoman Periods in Istanbul, Istanbul Technical University, Institute Of Science and Technology, 2009.

[13] D. Ekşi Akbulut & F. Aköz, Effects of Pozzolanas on Mechanical Properties of Mortars in Historical Buildings, Structural Analysis of Historical Constructions, New Delhi, 2006.

[14] S. Acun & N. Arioğlu, A Method for the Preservation and Restoration of the Stones Used in Historical Buildings, Architectural Science Review, 2006, 49 (2), 143-148

[15] Ö. Çizer, M.Sc. Thesis, Investigation of Lime Mortar Characteristics for the Conservation of the Ottoman Baths in Seferihisar-Urla Region, Izmir Institute of Technology, Izmir, 2004.

[16] A. A.Kasimzade, A.Dushimimana, S.Tuhta, G.Atmaca, F.Günday, O. A. Abrar, Comparative Study on Effectiveness of Using Horasan Mortar as a Pure Friction Sliding Interface Material, 2019, DOI:10.24018/ejers.2019.4.2.1166

[17] M. Canbaz & E. Güler, Effect of Lime Type on the Properties of Khorasan Mortar, Uluslararası

Katılımlı 6. Tarihi Yapıların Korunması ve Güçlendirilmesi Sempozyumu, 2017.

[18] K. Beck & M. Al-Mukhtar Formulation and characterization of an appropriate lime-based mortar for use with a porous limestone, Environmental Geology, 2008, 56 (3–4), 715–727.

[19] Ü. A. Mavioğlu, M.Sc. Thesis, Farklı Puzolanik Katkılar ile Hazırlanan Horasan Harçlarının Değişen Parametrelerinin İncelenmesi Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul, 2011.

[20] C. Yılmaz, M.Sc. Thesis, Portland Çimentosu Kullanımının Horasan Harcı Özelliklerine Etkisi, Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara, 2010.

[21] B. Işıkdağ & İ. B. Topçu, The effect of ground granulated blast-furnace slag on properties of Horasan mortar. Construction and Building Materials, 2013, 40, 448-454

[22] Turkish Standards Institute, TS EN 196-1, 2016, Methods of testing cement - Part 1: Determination of strength, Turkish Standard.

[23] Turkish Standards Institute, TS EN 12504-4, 2012, Testing concrete - Part 4: Determination of ultrasonic pulse velocity, Turkish Standard.

[24] A.M. Neville, Properties of Concrete, 4th Edition, 2000.

Chapter 44

DETERMINATION OF RELATIONSHIP BETWEEN NUTRIENT PARAMETERS USING TROPHIC STATE INDEX AND CLUSTER ANALYSIS

Ayla BILGIN[1]

1 * Artvin Coruh University, Engineering Faculty, Department of Environmental Engineering, 08100, Artvin/ Turkey Tel: +90(466)2151040 /4669 Fax: +90(466)2151057, E-mail:ayla.bilgin@gmail.com, ayla.bilgin@artvin.edu.tr

INTRODUCTION

The quality of a body of water is influenced by high concentrations of nutrients (for example from phosphates and nitrogen), causing changes in the natural balance of water (Yang et al., 2018). The chemical processes and dynamics of ecosystems in the water are negatively influenced by excessive nutrients. The changes occurring in water quality characteristics owing to the chemical and biological processes are known as eutrophication and are frequently connected with human activities (Jiang et al., 2014; Nyairo et al., 2015; Quevedo-Castro et al., 2019).

There are many studies on eutrophication process in lakes and reservoirs. This situation has provided opportunities in the development of numerous eutrophication models and mathematical calculation methods and categorization of waters according to their trophic status. Trophic categories primary productivity values are traditionally related to the biomass values and nutritional values and / or nutrient concentration of primary producers (Chalar et al., 2011). Eutrophication is one of the most significant and permanent water quality problems in freshwater ecosystems (Le Moal et al., 2019; Zou et al., 2020). Eutrophication expresses an ecosystem reaction to the nutritional value, typically nitrogen (N) and phosphorus (P), resulting from the influence of natural factors and anthropogenic activity (Du et al., 2019; Muri et al., 2018).

The reservoirs built for hydroelectric power are large-scaled artificial ecosystems. There are a lot of factors affecting the water quality of these basins (de Oliveira et al., 2020; Tundisi & Tundisi, 2012). It is known that the water quality of river basins deteriorates as the result of the anthropogenic activities in the areas surrounding the reservoirs (Wunderlin, 2018). This activities may also affect the current water quality and volume for electric power production and human resource in some cases (Glibert et al., 2018). Therefore, the water quality at long date and hydroclimatic observations are essential for monitoring and comprehending the systems of ecosystem (Cloern, 2019; Li et al., 2020; Shafiei & McLoughlin, 2017). Physical, chemical and biological components with regard to human effects and possible uses of human are analyzed in order to evaluate water quality. For this reason, poor water quality puts at risk the sustainability and survival of ecosystems and species (Chittoor Viswanathan et al., 2015; Quevedo-Castro et al., 2019; Yu et al., 2016).

Determining the trophic state of a lake has importance in the scientific evaluation of each. This enables us to determine (a) biotic status of abiotic and water bodies, (b) the relationship between chemical and biologic parameters and (3) the status of the lake (Matthews et al., 2002).

The trophic situation of inland waters are divided at three levels: oligotrophic, mezotrophic and eutrophic. Attempts have been made in order to evaluate the trophic status of the inland waters quantitatively by using both univariate and multivariate methods (Wang et al., 2018). Carlson defined a numerical Trophic State Index (TSI), which can be calculated using Secchi depth (SD), chlorophyll-a (Chl-a) or total phosphorus (TP) in 1977.

The purpose of this study is to use the index in order to determine the relationship between nutrients and trophic status of dam water quality and its evaluation was made by using the multivariate statistical analysis (hierarchical analysis).

MATERIALS AND METHODS

Study Area

Borçka dam is Turkey's 59th and Artvin's 3rd largest power plant. The plant is also Turkey's 16th largest Hydroelectric Power Plant (URL-1). The purpose of the construction of the Borcka dam is to protect energy production and flood. The studies with regard to the cage fish farming in the dam lake are continuing. It is observed that there has been an increasing trend in the matter of the production of fishing facilities in Borcka dam lake since their establishment (Sucu & Dinç, 2008).

The data of the study were obtained from Artvin (Turkey) Provincial Directorate of Environment and Urbanization. The samples were taken from the surface, middle and bottom of the Borçka dam lake in 2015. Total phosphorus (TP), total nitrogen (TN), chlorophyll-a (Chl-a) and secchi disk (SD) depths were measured for the purpose of assessing the water quality data. The sample points has been shown in the Figure 1.

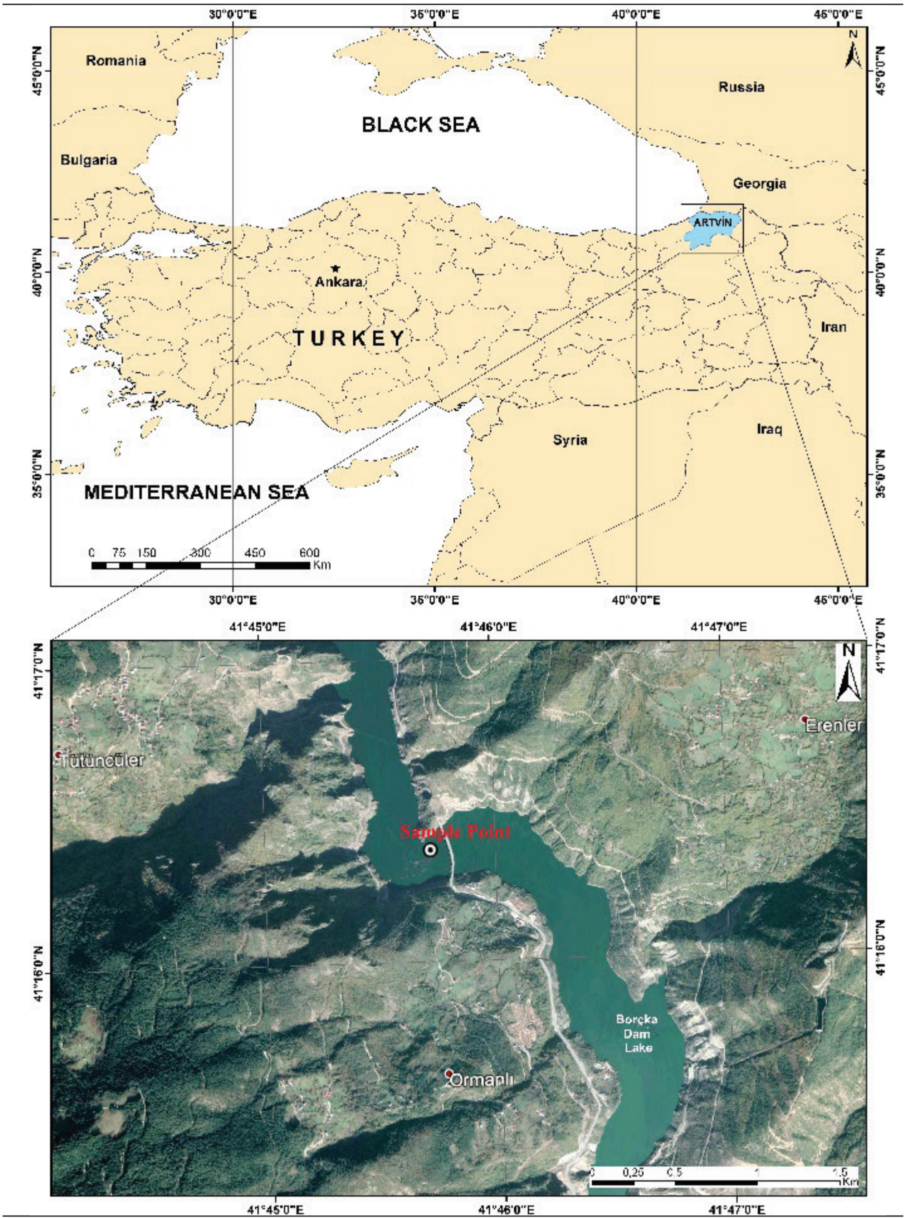


Figure 1.*Monitoring Sites*

Trophic State Index

Carlson (1977) suggested the measurement of the trophic status of a water body in accordance with the Trophic State Index (TSI) value and they can be divided into four basic groups: oligotrophic, mesotrophic, eutrophic and hypereutrophic. TSI has a lot of advantages such as ease of implementation and small data requirements. Carlson (1977) made a calculation by means of using SD, Chl *a* and TP concentrations (Carlson, 1977). Kratzer and Brezonik also included TN concentration in the Carlson type index (Kratzer & Brezonik, 1981). Every 10 units in this system represent a half decrease in Secchi depth, a third increase in chlorophyll concentration and a doubling of the total amount of phosphorus (Dodds et al., 1998). The parameter from which the unit is calculated is shown in parentheses since TSI can be calculated from these four parameters. The categorization of these indexes is shown in the Table 1.

TSI(CHL, µg/L)=10×[6–(2.04–0.68lnCHL)/ln2]

TSI(TP, µg/ L)=10×[6–ln(48/TP)/ln2]

TSI (SD, m) = 10 × [6 – ln SD/ln 2]

TSI (TN, mg/L) = 10 × [6 – ln (1.47/TN) ln 2]

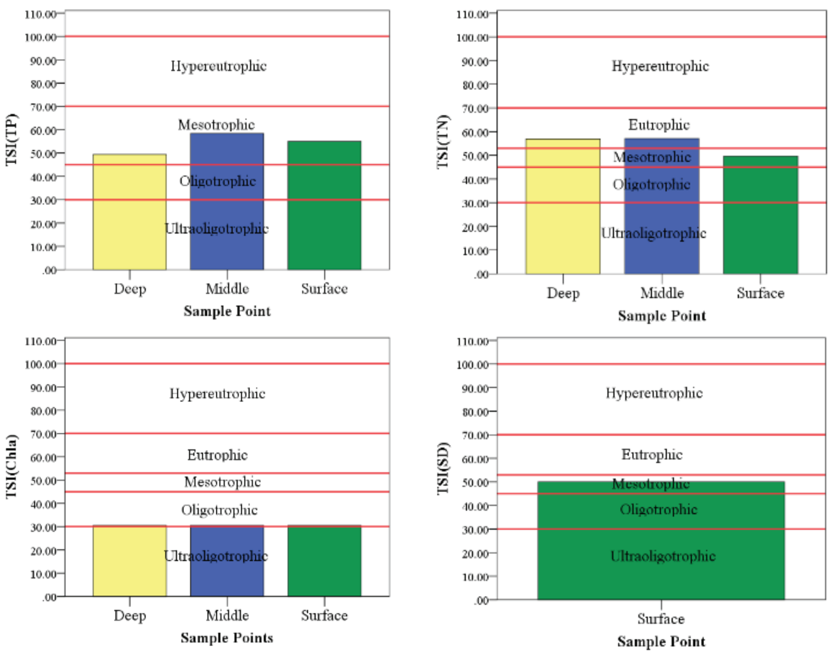
Table 1. Trophic States Index Classifications

	TSI	SD (m)	**TP (µg P/L)**	**Chl-a (µg/L)**	**TN (mg N/L)**
Ultraoligotrophic	0-20	64-16	0.75-3	0.04-0.34	0.02-0.09
Oligotrophic	30-40	8-4	6-12	0.94-2.6	0.18-0.37
Mesotrophic	45-50	2.8-2	17-24	5-6.4	0.52-0.74
Eutrophic	53-60	1.6-1	30-48	10-20	0.92-1.47
Hypereutrophic	70-100	0.5-0.062	96-768	56-1183	2.94-23.6

RESULTS AND DISCUSSION

Summary of Water Quality

The trophic status evaluation of the dam lake water quality was made in line with the "Surface Water Quality Regulation" published in the Official Gazette dated 30.11.2012 and numbered 28483. This regulation was revised twice in 2015 and 2016. The boundary values of "**Lake, Pond and Dam Lakes Eutrophication Criteria**" in the Regulation are given in Table 2. The boundary limits given for the lake, pond and dam lakes were taken as base according to these values (Table 2; Figure 2). The TP parameter is evaluated on the surface (34 µg / L), middle (43 µg / L), and at the bottom (23 µg / L) as mesotrophic in accordance with the boundary values set forth in the regulation (Table 3). TP parameter is determined as mesotrophic on the surface (55) and in the middle (58) and at the bottom (49) as eutrophic according to TSI index. It has been observed that it does not give exactly the same result with TSI index for TP parameter. The TN parameter is evaluated on the surface (716 µg/L) as mesotrophic and in the middle (1197 µg/L) and at the bottom (1185 µg/L) as eutrophic in accordance with the regulation. The TN parameter was determined on the surface (49.63) as mesotrophic and in the middle (57.04) and at the bottom (56.9) as eutrophic according to the TSI index. The water quality is assessed as the same according to the result of TN parameter regulation and TSI index calculation. The Chla parameter is evaluated as oligotrophic on the surface, in the middle and at the bottom pursuant to the regulation and TSI index. The SD parameter is evaluated as mesotrophic on the surface, in the middle and at the bottom pursuant to the regulation and TSI index. The water quality of the lake is determined to be at the mesotrophic level when the average calculation is made in line with the TSI index (Table 4). It was expressed that there may be domestic industrial agricultural activities affecting the water quality in the study carried over in Çoruh basin (Bilgin, 2015). It is determined that the water quality is (Bilgin, 2018) in the studies with relation to the Borçka dam lake water quality and the water quality was found to be at the mesotrophic level in the study of water quality by means of using multivariate statistical analysis (Bilgin, 2015). In the study showing the trophic status of Borcka dam, the data of 2016 and 2013 years were used and this article shows a similar result (Bilgin, 2020).



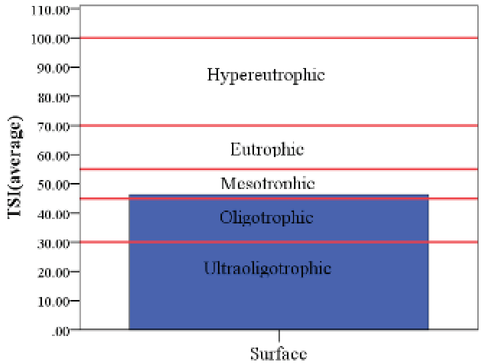


Figure 2. *TSI calculations values calculated at measuring points*

Table 2. The limiting values of "Lake, Pond, and Dam Lakes eutrophication criteria" dam lake trophic states in Turkish regulation

Turkish Water Quality Limited Degree				
Water Quality Degree	**TP (µg/L)**	**TN (µg/L)**	**Chla (µg/L)**	**SD (m)**
Oligotrophic	< 10	< 350	< 3,5	> 4
Mezotrophic	30	650	9	2
	50*	1000*	15*	1,5*
Eutrophic	100	1500	25	1
Hipertrohic	> 100	> 1500	> 25	< 1

* Limit values valid for ponds or dams

Table 3. Evaluation of Water Quality According to Limit Values in Legislation

	TP(µg/L)		TN(µg/L)		Chla (µg/L)		**SD**	
Surface	34	M	716	M	1	O	2	M
Middle	43	M	1197	E	1	O		
Depth	23	M	1185	E	1	O		

Classificaitons; E:Eutrophic M:Mesotrophic, O:Oligotrophic

Table 4. Evaluation of Water Quality According to TSI									
TSI(TP) (µg/L)		TSI(TN) (µg/L)		TSI(Chla) (µg/L)		TSI (SD)		TSI average	
55.0	E	49.63	M	30.6	O	50.01	M	46.3	M
58.4	E	57.04	E	30.6	O				
49.4	M	56.90	E	30.6	O				

Classificaitons; E:Eutrophic M:Mesotrophic, O:Oligotrophic

The cluster analysis ensures for grouping by looking at the similarities of the samples (Shrestha & Kazama, 2007).Wards method was used in order to classify the samples in the hierarchical accumulation clustering analysis (Li et al., 2011; Willet, 1987). Two groups, Cluster -1 (Deep) and Cluster 2 (Surface, Middle), were obtained in accordance with the sampling depth in consequences of the study. This situation demonstrates that the water quality at the bottom differs from the surface and the middle. Two groups such as Cluster -1 (TSI(Chla) and Cluster-2 (TSI(TP), (TSI(TN)) were achieved when it is analyzed pursuant to the TSI (Figure 3).

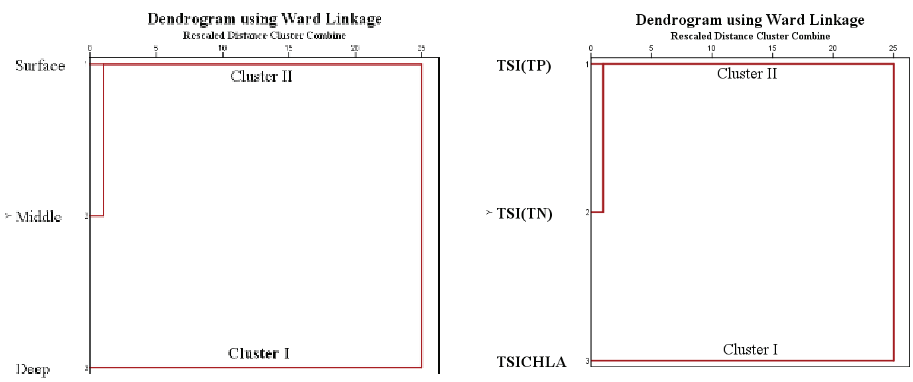


Figure 3. *Cluster analysis*

Determinant of the limiting nutrient

The phosphorus and nitrogen are primary limiting nutrients for aquatic algae production. TN: TP is generally suggested as an index in order to identify N and P limiting nutrients in lakes. (Canfield, 1983; Downing & McCauley, 1992). The threshold value of TN / TP varies in the literature (Huber et al., 1982). The lakes can be categorized according to the TN: TP rate (Table 5). The nitrogen may limit algal growth in the places where TN:TP is low and therefore TN will ensure a better estimate for chlorophyll. On the other hand, the phosphorus is more likely to limit growth in the places where TN:TP is high and TP will be better predictive (Phillips et al., 2008). Nitrogen and phosphorus are primary limiting nutrients for algal production in aquatic systems. However, Dodds & Whiles reported that phosphorus is a more restrictive element than nitrogen in freshwater systems (Dodds & Whiles, 2010). Sakamoto (1966) and Smith (1982) reported that nitrogen is the limiting nutrient when TN / TP <10, phosphorus is a limiting nutrient when TN / TP> 17 and the freshwater environments were considered a balanced system when it was between (10 <TN / TP ≤ 17). The surface is identified as P-limited, the middle and the bottom are identified as co-limited in this study. The fact that it is P limited or co-limited produces the result that the dam lake demonstrates non-algal turbidity. P is most inclined to be limited in lakes turning from oligotrophic into mesotrophic since precipitation are often still oxy and P is relatively permanent owing to the tendency of the insoluble complexes and particles formed with metals to absorb into the substance (Welch & Lindell, 1992). Although there is a little P in rainwater (dust-related on a large scale), It is shown that P has a relatively important contribution in phosphorus-poor basins (Jacoby, 2014).

Two-dimensional graphical approaches frequently used in the explication of the limiting factors that influence the chlorophyll and nutrient concentration, were used. (Carlson & Havens, 2005) It is determined as TSI (Chla) < TSI (SD) ve TSI(TP)>TSI(SD)>TSI(Chla) when it is evaluated pursuant to the TSI index. This situation demonstrates that there is non-algal turbidity. In this circumstances, there is predominantly turbidity generated by sediment or other organic substances rather than chlorophyll. At the same time, it may also occur in the event that the large particles (spirulina) are dominant and if transparency is less affected by particles or zooplankton grazing removes smaller particles and leaves only large forms.

Table 5. Determinant of the limiting nutrient

Sample Point	TN/TP	Evaluation	TSI Evaluation
Surface	11.76	P-limited	a) TSI (Chla) < TSI (SD) 30.6<50.01
Middle	27.84	Co-limited	b)TSI(TP)>TSI(SD)>TSI(Chla)
Deep	51.52	Co-limited	55>50.01>30.6

CONCLUSIONS

TSI and hierarchical analysis were used for the purpose of determining the water quality of Borcka dam lake in this study. The Borçka dam lake is evaluated as mesotrophic in accordance with the TP and SD parameters, eutrophic-mesotrophic pursuant to the TN parameters and oligotrohic according to the Chla parameters in the evaluation conducted in accordance with the legislation. TSI index is determined that it has mesotrophic feature when evaluated according to the average. The quality of bottom water was determined that it was different with respect to the surface and the middle water quality when evaluated pursuant to hierarchical analysis. One of the significant factors in specifying the water quality is the definition of the limiting factor. The limiting factor is designated as P-limited on the surface and as co-limited in the middle ant at the bottom according to the TN/TP rate in this study. At the same time, the fact that the TSI (Chla) value is less than the TSI (SD) value demonstrates that the non-algal turbidity is more dominant in this place. This study shows the usage of index and statistic in the evaluation of the water quality of dam lake. The evaluation of the quality of water only in accordance with the legislation causes incomplete comments. For this reason, the deterioration of water quality can be evaluated more accurately by means of determining nutrient limiting factors.

Acknowledgments

The author is thankful to the Artvin (Turkey) Provincial Directorate of Environment and Urbanization for providing the data. The author is thankful to Sebahat TEMUÇİN KILIÇER for drawing and mapping.

Ethics committee approval: Ethics committee approval is not required.

Conflict of Interests: There are no conflicts of interest to declare.

REFERENCES

Bilgin, A. (2015). An assessment of water quality in the Coruh Basin (Turkey) using multivariate statistical techniques. *Environmental Monitoring and Assessment*, *187*(11), 721.

Bilgin, A. (2018). Evaluation of surface water quality by using Canadian Council of Ministers of the Environment Water Quality Index (CCME WQI) method and discriminant analysis method: a case study Coruh River Basin. *Environmental Monitoring and Assessment*. https://doi.org/10.1007/s10661-018-6927-5

Bilgin, A. (2020). Trophic state and limiting nutrient evaluations using trophic state / level index methods : a case study of Borçka Dam Lake. *Environmental Monitoring and Assessment*, *192*(794).

Canfield, D. E. (1983). PREDICTION OF CHLOROPHYLL A CONCENTRATIONS IN FLORIDA LAKES: THE IMPORTANCE OF PHOSPHORUS AND NITROGEN. *JAWRA Journal of the American Water Resources Association*. https://doi.org/10.1111/j.1752-1688.1983.tb05323.x

Carlson, R. E. (1977). A trophic state index for lakes. *Limnology and Oceanography*. https://doi.org/10.4319/lo.1977.22.2.0361

Carlson, R. E., & Havens, K. E. (2005). Simple graphical methods for the interpretation of relationships between trophic state variables. *Lake and Reservoir Management*. https://doi.org/10.1080/07438140509354418

Chalar, G., Arocena, R., Pacheco, J. P., & Fabián, D. (2011). Trophic assessment of streams in Uruguay: A Trophic State Index for Benthic Invertebrates (TSI-BI). *Ecological Indicators*. https://doi.org/10.1016/j.ecolind.2010.06.004

Chittoor Viswanathan, V., Molson, J., & Schirmer, M. (2015). Does river restoration affect diurnal and seasonal changes to surface water quality? A study along the Thur River, Switzerland. *Science of the Total Environment*. https://doi.org/10.1016/j.scitotenv.2015.05.121

Cloern, J. E. (2019). Patterns, pace, and processes of water-quality variability in a long-studied estuary. *Limnology and Oceanography*. https://doi.org/10.1002/lno.10958

de Oliveira, T. F., de Sousa Brandão, I. L., Mannaerts, C. M., Hauser-Davis, R. A., Ferreira de Oliveira, A. A., Fonseca Saraiva, A. C., de Oliveira, M. A., & Ishihara, J. H. (2020). Using hydrodynamic and water quality variables to assess eutrophication in a tropical hydroelectric reservoir. *Journal of Environmental Management*. https://doi.org/10.1016/j.jenvman.2019.109932

Dodds, W. K., Jones, J. R., & Welch, E. B. (1998). Suggested classification of stream trophic state: Distributions of temperate stream types by chlorophyll, total nitrogen, and phosphorus. *Water Research*. https://doi.org/10.1016/S0043-1354(97)00370-9

Dodds, W. K., & Whiles, M. R. (2010). Freshwater Ecology. In *Freshwater Ecology*. https://doi.org/10.1016/C2009-0-01718-8

Downing, J. A., & McCauley, E. (1992). The nitrogen : phosphorus relationship in lakes. *Limnology and Oceanography*. https://doi.org/10.4319/lo.1992.37.5.0936

Du, H., Chen, Z., Mao, G., Chen, L., Crittenden, J., Li, R. Y. M., & Chai, L. (2019). Evaluation of eutrophication in freshwater lakes: A new non-equilibrium statistical approach. *Ecological Indicators*. https://doi.org/10.1016/j.ecolind.2019.03.032

Glibert, P. M., Al-Azri, A., Icarus Allen, J., Bouwman, A. F., Beusen, A. H. W., Burford, M. A., Harrison, P. J., & Zhou, M. (2018). *Key Questions and Recent Research Advances on Harmful Algal Blooms in Relation to Nutrients and Eutrophication*. https://doi.org/10.1007/978-3-319-70069-4_12

Huber, W. C., Brezonik, P. ., Heaney, J. P., Dickinson, R. E., & Preston, S. D. (1982). *A classification of Florida lakes. Final Report, Florida Department of Environmental Regulation. Tallahassee, FL.*

Jacoby, J. (2014). Pollutant Effects in Freshwater. In *Pollutant Effects in Freshwater*. https://doi.org/10.1201/9781482265118

Jiang, Y. J., He, W., Liu, W. X., Qin, N., Ouyang, H. L., Wang, Q. M., Kong, X. Z., He, Q. S., Yang, C., Yang, B., & Xu, F. L. (2014). The seasonal and spatial variations of phytoplankton community and their correlation with environmental factors in a large eutrophic Chinese lake (Lake Chaohu). *Ecological Indicators*. https://doi.org/10.1016/j.ecolind.2014.01.006

Kratzer, C. R., & Brezonik, P. L. (1981). A CARLSON-TYPE TROPHIC STATE INDEX FOR NITROGEN IN FLORIDA LAKES. *JAWRA Journal of the American Water Resources Association*. https://doi.org/10.1111/j.1752-1688.1981.tb01282.x

Le Moal, M., Gascuel-Odoux, C., Ménesguen, A., Souchon, Y., Étrillard, C., Levain, A., Moatar, F., Pannard, A., Souchu, P., Lefebvre, A., & Pinay, G. (2019). Eutrophication: A new wine in an old bottle? In *Science of the Total Environment*. https://doi.org/10.1016/j.scitotenv.2018.09.139

Li, B., Yang, G., & Wan, R. (2020). Multidecadal water quality deterioration in the largest freshwater lake in China (Poyang Lake): Implications on eutrophication management. *Environmental Pollution*, *260*, 114033. https://doi.org/10.1016/j.envpol.2020.114033

Li, S., Li, J., & Zhang, Q. (2011). Water quality assessment in the rivers along the water conveyance system of the Middle Route of the South to North Water Transfer Project (China) using multivariate statistical techniques and receptor modeling. *Journal of Hazardous Materials*, *195*, 306–317. https://doi.org/10.1016/j.jhazmat.2011.08.043

Matthews, R., Hilles, M., & Pelletier, G. (2002). Determining trophic state in Lake Whatcom, Washington (USA), a soft water lake exhibiting seasonal nitrogen limitation. *Hydrobiologia*. https://doi.org/10.1023/A:1015288519122

Muri, G., Čermelj, B., Jaćimović, R., Ravnikar, T., Šmuc, A., Turšič, J., & Vreča, P. (2018). Factors that contributed to recent eutrophication of two Slovenian mountain lakes. *Journal of Paleolimnology*. https://doi.org/10.1007/s10933-017-9996-5

Nyairo, W. N., Owuor, P. O., & Kengara, F. O. (2015). Effect of anthropogenic activities on the water quality of Amala and Nyangores tributaries of River Mara in Kenya. *Environmental Monitoring and Assessment*. https://doi.org/10.1007/s10661-015-4913-8

Phillips, G., Pietiläinen, O. P., Carvalho, L., Solimini, A., Lyche Solheim, A., & Cardoso, A. C. (2008). Chlorophyll-nutrient relationships of different lake types using a large European dataset. *Aquatic Ecology*. https://doi.org/10.1007/s10452-008-9180-0

Quevedo-Castro, A., Lopez, J., Rangel-Peraza, J., Bandala, E., & Bustos-Terrones, Y. (2019). Study of the Water Quality of a Tropical Reservoir. *Environments*, *6*(1), 7. https://doi.org/10.3390/environments6010007

Shafiei, F., & McLoughlin, P. D. (2017). Change in nutrient loading pattern due to coupled effect of change in concentration and hydroclimatic forces. *Journal of Freshwater Ecology*. https://doi.org/10.1080/02705060.2017.1406873

Shrestha, S., & Kazama, F. (2007). Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. *Environmental Modelling & Software*, *22*(4), 464–475. https://doi.org/10.1016/j.envsoft.2006.02.001

Sucu, S., & Dinç, T. (2008). ÇORUH HAVZASI PROJELERİ. *TMMOB 2. Su Politikaları Kongresi*.

Tundisi, J. G., & Tundisi, T. . (2012). *Limnology*. CRC Press.

URL-1. (n.d.). *https://www.enerjiatlasi.com/hidroelektrik/borcka-baraji.html*.

Wang, S., Li, J., Zhang, B., Spyrakos, E., Tyler, A. N., Shen, Q., Zhang, F., Kuster, T., Lehmann, M. K., Wu, Y., & Peng, D. (2018). Trophic state assessment of global inland waters using a MODIS-derived Forel-Ule index. *Remote Sensing of Environment*. https://doi.org/10.1016/j.rse.2018.08.026

Welch, E. B., & Lindell, T. (1992). Ecological effects of wastewater: applied limnology and pollution effects. In *Ecology*.

Willet, P. (1987). *Similarity and Clustering in Chemical Information Systems. Research Studies Press*. Wiley.

Wunderlin, D. A. (2018). *The Suquía River Basin (Córdoba, Argentina): an Integrated Study on its Hydrology, Pollution, Effects on Native Biota and Models to Evaluate Changes in Water Quality*. Springer.

Yang, Z., Cheng, B., Xu, Y., Liu, D., Ma, J., & Ji, D. (2018). Stable isotopes in water indicate sources of nutrients that drive algal blooms in the tributary bay of a subtropical reservoir. *Science of the Total Environment*. https://doi.org/10.1016/j.scitotenv.2018.03.266

Yu, S., Xu, Z., Wu, W., & Zuo, D. (2016). Effect of land use types on stream water quality under seasonalvariation and topographic characteristics in the Wei River basin, China. *Ecological Indicators*, *60*, 202–212. https://doi.org/https://doi.org/10.1016/j.ecolind.2015.06.029

Zou, W., Zhu, G., Cai, Y., Vilmi, A., Xu, H., Zhu, M., Gong, Z., Zhang, Y., & Qin, B. (2020). Relationships between nutrient, chlorophyll a and Secchi depth in lakes of the Chinese Eastern Plains ecoregion: Implications for eutrophication management. *Journal of Environmental Management*. https://doi.org/10.1016/j.jenvman.2019.109923

Chapter 45

SECTORAL INEQUALITIES IN INFORMATION SYSTEMS: A GINI ANALYSIS OF EUROPEAN COUNTRIES AND TURKEY

Mehmet Erkan YUKSEL

1 Asst. Prof. Dr., Burdur Mehmet Akif Ersoy University, Faculty of Engineering and Architecture, Department of Computer Engineering, Burdur-Turkey. Email: erkanyuksel@mehmetakif.edu.tr, ORCID ID: 0000-0001-8976-9964

INTRODUCTION

Information technologies, which have become one of the fundamental elements of the information society, bring some problems depending on the usage differences. Differences arising from the use of information systems lead to inequalities between people. The inequality issue referred to as the digital gap, digital divide, or digital poverty deepens with the developments in technology and becomes a problem that threatens the information society. Therefore, it is significant to measure the level of inequality caused by the digital divide. Different indexes, parameters, and indicators are employed in many studies on the subject. Data is obtained from the questionnaires and analyzed by the traditional statistical methods. Researches usually focus on demographic variables and income levels in a personal, regional, and global context. However, there are no adequately organizational (company-oriented) and sectoral analyzes. The lack of a standard measurement tool creates methodologic confusion in determining the level of the digital divide and inadequacies in comparing results effectively. Gini coefficients, one of the most suitable instruments to evaluate income inequalities, were rarely used in determining the digital divide levels. Considering that the digital divide is inequality in distribution, the most appropriate approach would be to use Gini coefficients to measure the digital divide level. In this study, we analyzed the digital divide levels between sectors in some European countries and Turkey by using the Gini coefficients. We obtained the data required for calculating the Gini coefficients from the Turkish Statistical Institute (TUIK) and the European Statistical Office (Eurostat) research reports. The results demonstrated that, between sectors, the highest digital divide level is in online sales, and the lowest is in computer ownership.

Digital technology-based communication systems, which started to become widespread after computers were made suitable for personal use, have brought out some differences in social and economic life. After the computers that created the beginning of digital technologies, the second curtain in information systems was opened with the Internet. Internet is a sharing-based system designed for the mutual transfer of information and documents. It formed a worldwide network and enabled users to communicate with each other. It provided a significant contribution to the dissemination of information and eliminated the restriction of information to only one computer. Although limited operations such as plain-text, e-mail, and file sending could be performed in the first years, the Internet expanded the communication channels between the related parties. With the HTML coding technique developed by Tim Barners Lee, the internet users who are introduced to the web concept increased their interactions through web technologies (Melnikas, 2011; Spector, 2001). Image, audio, and video information were transferred over the web, and the interaction environment was supported by languages such as ASP and PHP, which are interpreted by browsers. Thanks to these languages that can be embedded into HTML codes, clients can make queries from databases and instantly access the data they want.

Technological developments bring about changes in business life. The most radical developments that force business processes and organizations to structural changes are seen with information and communication technologies (ICT) (Drucker, 2010). In the early years of the use of computers, information systems, which led to discussions by some scientists led by Solow that they had no contribution to productivity (Solow, 1987), are among the indispensable resources of businesses today. The process of transition to informatics, which started with the idea of not staying behind from competitors, offers companies many advantages such as cost, efficiency, control, flexibility, and speed. The use of ICT increases operational efficiency, organizational flexibility, profitability, competitiveness, and encourages companies to innovate (Akin, 2005; Bayo-Moriones and Lera-Lopez, 2007; Love et al., 2005). In this way, companies gain a dynamic structure where continuous innovation is carried out. ICT that makes the activities of companies more efficient, increases market efficiency. Especially in the years when the internet was opened to social use, it was predicted that information-based problems would disappear with the increase in instant access to information and the speed of information dissemination, thus increasing firm and market activities and creating a perfect market environment (Bakos, 1998; Ellison and Ellison, 2005; Brousseau and Curien, 2007).

There is also a dark side to this comfortable world, which is depicted to be formed by information systems. Authors like Whinston et al. (1997), Lee (1998), Huston and Spencer (2002), Ellison and Ellison (2005) emphasized that information systems did not create the perfect environment as claimed in the 1990s, on the contrary, knowledge-based problems increased during the transition to information technologies. The main problems that researchers draw attention to are incomplete information and the resulting uncertainty. Hoffman et al. (1999), on the other hand, drew attention to the concerns of confidentiality and the capture of private information and stated that these problems negatively affect the use of information systems. The passing of private information to third parties is due to security weakness. According to Egger (2003), recording personal data in databases is the main reason for security weakness.

One of the problems faced with the spread of information technologies is the digital divide. This concept, which is expressed as the difference between those who have information systems and those who do not, is a situation of inequality and is seen as a sociological problem that negatively affects the social structure (Attewell, 2001). In general, it is seen that researches on the subject are conducted at personal, regional, and global levels, and there are not many studies on companies and sectors. The digital divide, which causes inequalities between companies, stands before us as a concept that negatively affects the activities of the companies and the market and economic structure accordingly. Therefore, determining the digital divide level between firms is important in terms of measures to be taken and policies to be implemented.

When the studies on measuring the digital divide are examined, it is seen that two approaches usually stand out. Indexes that contain information technologies variables such as computer ownership, internet access rate, number of hosts, bandwidth are used for comparisons and rankings between countries. These indexes employed by international organizations such as OECD, World Bank, and United Nations contradict each other in terms of results. In another approach, the relationship of information systems usage levels is investigated via demographic variables, and the analyzes such as variance, factor, correlation are used. Digital divide studies using Gini coefficients that have become a standard for calculating inequalities are rarely encountered. In these studies, personal and regional analyzes are performed with data grouped according to demographic variables such as age, gender, and income.

In this study, we carried out a sectoral digital divide analysis of some European countries and Turkey using data related to the computer, Internet, website ownership, and online shopping. We used the Gini method in sectoral analyzes based on the 2013 Eurostat and TUIK data. Data were grouped by NACE Rev.2 system. Considering the results, we demonstrated that Turkey is not in a bad position as it is reported in international indexes and, on the contrary, that Turkey has a lower digital divide level in some information systems usages compared to the other countries.

SECTORAL DISTRIBUTIONS AND INFORMATION TRENDS IN TURKEY AND EUROPE

Sectors consisting of companies operating in a specific production and service area are tools used in grouping enterprises. The characteristic structures of the sectors are determined by the companies they own. Proper classification of companies according to their fields of activity is essential for sectoral analyzes. Many sectoral classifications have been made according to the activity field, and the most common classification between these is the grouping of companies as agriculture, industry, and service sectors. Over the last decade, NACE sectoral classification that Turkey also uses is a system developed since 1970 by Eurostat. In the classification system, in which last changes were made in 2008 and named as NACE Rev.2, it is aimed to give an international dimension to the determination of economic activity areas and to provide comparisons between countries (Eurostat, 2008). TUIK, which carries out sectoral analyzes in Turkey, has begun to employ the NACE system since 2010 (TUIK, 2015). In this context, the sectoral classification given in Table 1 is used for studies on information systems usage in enterprises.

Table 1: *Sectoral classification*

NACE Code	Industrial Sector (Economic Activity)
C	Manufacturing
D	Electricity, gas, steam and air conditioning supply
E	Water supply; sewerage; waste managment and remediation activities
F	Construction
G	Wholesale and retail trade; repair of motor vehicles and motorcycles
H	Transporting and storage
I	Accommodation and food service activities
J	Information and Communication
L	Real estate activities
M	Professional, scientific and technical activities
N	Administrative and support service activities

Source: Eurostat and TUIK

The sectors presented in Table 1 are used jointly by both TUIK and Eurostat. Although there are many sub-sectors, these eleven sectors are typically common in the reports. Since the number of enterprises in sectors D and E is low, these sectors are considered together. According to these sectoral groups, the number of enterprises in some European countries and Turkey is listed in Table 2. Since the latest data on the number of enterprises of TUIK and Eurostat belong to 2013, all tables have been arranged with the data of 2013. Therefore, the analysis carried out only covers 2013.

Table 2: *Number of enterprises in Turkey and some European countries in 2013*

Sector	Denmark	Sweden	Finland	Bulgaria	Romania	Turkey
C	15060	56063	26595	30907	67844	340413
D-E	4390	3855	2826	2926	4854	4705
F	30705	102671	48442	19710	64077	157679
G	42720	140227	58214	141423	274269	1074575
H	11243	31579	22995	19671	59646	421770

I	13670	33865	12905	26532	34907	232738
J	15406	59956	13614	10416	33722	31785
L	27522	56303	25986	21816	15988	24490
M	32336	180814	49595	39544	86172	154456
N	16445	39841	18253	9576	36613	29026
Sum	**209497**	**705174**	**279425**	**322521**	**678092**	**2471637**

Source: Eurostat and TUIK

The percentages of information systems usage, which is common in Eurostat and TUIK data, between enterprises are given by sectors in Table 3. When these values are examined, it is noteworthy that computer and internet usage is higher than web page ownership and online sales. The lowest values are seen in online sales. According to the data in Table 3, it can be said that Denmark is in the best position in information systems usage between countries. It is seen that the highest rate of information systems usage in Sweden, Finland, and Bulgaria is in sector I; in Denmark, it is in D-E, and in Turkey, it is concentrated in the sector J.

Table 3: *The ratio of information systems usage (%)*

Country	Information Systems	All Sectors	Sectors									
			C	D-E	F	G	H	I	J	L	M	N
Denmark	CO	99	99	**100**	**100**	99	98	**100**	**100**	**100**	99	99
	IA	99	99	**100**	**100**	99	98	99	**100**	**100**	99	99
	WO	92	94	**99**	95	89	79	97	96	92	97	96
	OS	27	21	16	12	44	20	**52**	27	19	12	24
Sweden	CO	98	99	**100**	97	99	99	**100**	99	97	**100**	96
	IA	98	98	**100**	96	98	97	**100**	99	97	**100**	96
	WO	89	93	**99**	79	93	73	98	98	97	95	87
	OS	24	28	27	7	35	26	**66**	23	18	14	21
Finland	CO	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
	IA	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
	WO	94	98	92	88	96	82	**100**	98	93	99	90
	OS	16	21	11	4	22	12	**71**	21	15	4	19
Bulgaria	CO	92	91	99	93	92	89	**100**	99	96	96	89
	IA	89	89	99	89	89	88	**100**	99	96	96	86
	WO	47	49	33	46	41	36	**88**	85	51	57	38
	OS	5	5	0	2	6	2	**20**	15	6	3	6

Romania	CO	85	85	96	70	92	88	89	94	**99**	97	84
	IA	83	84	94	69	91	86	84	94	**97**	**97**	78
	WO	42	34	43	37	54	32	51	**80**	43	30	41
	OS	9	6	1	1	15	**28**	12	16	2	3	9
Turkey	CO	92	90	96	91	96	91	85	**98**	94	93	90
	IA	91	88	94	90	96	90	83	**97**	92	91	88
	WO	54	56	62	49	56	44	54	**79**	49	57	45
	OS	10	10	5	4	13	6	19	**20**	4	8	7

CO: Computer Ownership, IA: Internet Access, WO: Website Ownership, OS: Online Sales
Source: Eurostat and TUIK

THE DIGITAL DIVIDE

The digital divide is defined as the differences in the use of ICT. This concept used to describe the differences between individuals, firms, and countries, covers the inequalities in access and effective use of ICT. The digital divide, which was first heard with the ownership of a fixed telephone line and evaluated in this way (Brousseau and Curien, 2007), deepened with the developing technology, and today, it is evaluated with parameters such as computer ownership, internet access, mobile device ownership, and online transactions. The digital divide that is considered as an unbalanced distribution of resources is seen as an inequality (Baker, 2001). Dewan and Riggins (2005) define the digital divide as the distinction between ownership and non-ownership for individuals. OECD (Organisation for Economic Co-operation and Development) describes this concept as the differences in owning and using ICT (OECD, 2001). This definition, which forms the basis of the approach to the digital divide concept in the literature, states that the digital divide is related to the effective use of ICT as well as ICT ownership.

Digital Divide Types

As can be understood from the OECD definition, there are two types of the digital divide: horizontal and vertical. The digital divide between ICT users and non-ICT users is defined as a vertical divide, while the digital divide related to effective use between ICT users is called a horizontal divide (Sedimo, Bwalya, & Plessis, 2011). The vertical divide, which is the first problem encountered with the possession of ICT, is called the first-level digital divide (Attewell, 2001). While analyzing the first-level divide that explains the differences between those who have access to ICT and those who do not have access, parameters such as hardware, software, Internet connection, and Internet connection types should be evaluated (Van Dijk and Hacker, 2003). The horizontal divide is used to express the differences related to the effective use of ICT by those who have access to ICT. Hence, it is referred to as the second-level digital divide (Attewell, 2001). The high complexity of ICT makes them difficult to use, and the different technological competencies of individuals prevent them from taking advantage of the technology at the same level. According to Wei and Hindman (2011), since the differences regarding effective use are more pronounced between social strata, they affect the social structure more. Therefore, academic studies and public policies on the second-level digital divide are required.

Measuring the Digital Divide

Rapid developments in technology give ICT a dynamic structure. This dynamic structure generates difficulties in determining ICT usage levels and revealing the differences in ICT usage. There is no consensus in the literature regarding the measurement of the digital divide due to reasons such as parameter diversity, different approaches of disciplines, diversity in ICT, wide usage areas, and rapid technological developments (Sciadas, 2005). Since an effective methodology has not yet been developed, researchers are forced to create frameworks of their methods. Therefore, different methods, indexes, and parameters are used in different studies.

There are two approaches in studies on measuring the digital divide. In the first approach, indexes calculated from informatics scores with some selected ICT indicators are used. These indexes are usually intended to reveal the digital divide between countries. Some of them are the Digital Opportunity Index (DOI), Digital Access Index (DAI), ICT Development Index (IDI), Networked Readiness Index (NRI), and Digital Evolution Index (DEI). In particular, these studies are promoted by global organizations such as OECD, World Bank, and United Nations. The indicators included in these indexes used to measure informatics inequalities vary according to the indexes, but generally, basic indicators such as computer ownership, internet access, mobile device usage are involved in all indexes (Fidan, 2016). It is aimed to determine the personal usage levels of information systems in the calculated informatics scores, and country rankings are performed. The status of some European countries and Turkey in the ranking of IDI and NRI scores that cover world countries are given in Table 4.

Table 4: *Score values and world rankings of some countries*

Country	IDI (2013)		NRI (2013)	
	Score	Rank	Score	Rank
Denmark	8.86	1	5.58	8
Sweden	8.67	3	5.91	3
Finland	8.31	8	5.98	1
Bulgaria	6.31	49	3.87	71
Romania	5.83	58	3.86	75
Turkey	5.29	68	4.22	45

Source: WEF and ITU reports

When Table 4 is examined, the inconsistency in IDI and NRI rankings draws attention. Denmark ranks first according to IDI, and third according to the NRI. It is also a similar case for Turkey. While IDI shows Turkey in the 68th row, lower than Romania and Bulgaria, NRI shows in the 45th row, above these countries. This discrepancy between indexes poses a question mark on what the scores actually measure.

Another approach to measuring the digital divide is to examine ICT such as computer, Internet, and mobile phone separately instead of compound structure. T-score, variance, factor, and correlation analyzes are used in these studies, which aim to determine the relationship between ICT ownership levels and demographic variables such as age, gender, ethnic origin, income level, and education level. Thus, the factors affecting the digital divide are tried to be determined.

MEASURING INEQUALITES

According to the economic understanding built on the principle that resources are limited, imbalances in the distribution of resources cause inequalities in the economic system. Karoly (1992) defines inequality as the situation in which one of the distributions differs from the other. Several approaches such as the Variance coefficient, Theil index, Atkinson index, Lorenz curve, Gini coefficients, and Robin Hood index can be used to determine the inequality levels. Between these approaches, it is stated that the best method to measure inequalities is Gini (MacLachlan and Sawada, 1997). The Gini coefficient is a value obtained from the Lorenz curve, which is a graphical representation of income inequality.

Gini Coefficients

The Gini method, which is developed by Corrado Gini, is one of the methods used to express inequalities numerically (Ceriani and Verme, 2012). It is based on the Lorenz curve developed by American economist

Max Otto Lorenz in 1905. The Lorenz curve is a graphical representation of how much shares did people take from the total income generated (Kakwani, 1977). In other words, it is a graphical representation of the distribution of income or wealth within a population. Along with its derivative statistics, it is widely used for measuring inequality across a population. The population is divided into segments based on the income levels determined, and the incomes of each group are calculated cumulatively (Chakraborty and Bosman, 2005). Figure 1 depicts a Lorenz curve showing the income distribution inequalities within the population. The inequality is determined by the area A between the diagonal called the absolute equality line and the Lorenz curve. A large area of "$\boldsymbol{A}$" indicates that the inequality in the distribution is high, and a small one indicates that inequality is low. If the total income is collected in a single segment or an individual, the Lorenz curve will consist of diagonals, and thus a situation of complete inequality will occur (Maclachlan and Sawada, 1997).

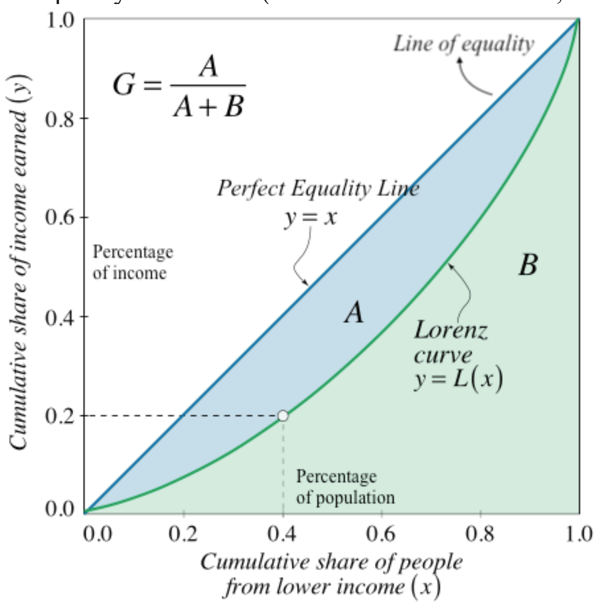


Figure 1: *Lorenz curve*

The Gini approach, which is one of the indexes used to express inequalities numerically, is a method based on the sum of the differences (Goodman and Kruskal, 1959). It is used to determine the numerical levels of the distribution difference in fields such as sociology, economics, geography, biology, health, industrial settlement, and education. The Gini coefficient, which is referred to as the Gini concentration ratio or the Gini index, is the most appropriate method to measure inequalities (Chakraborty and Bosman, 2005). When considering Figure 1, the ratio of the area $(\boldsymbol{A})$ between the Lorenz curve and the absolute equality line to the right triangle

(***A*** + ***B***) under the absolute equality line is called the Gini coefficient and obtained by Equation (1).

$$\mathbf{G} = \frac{\mathbf{A}}{\mathbf{A+B}} \tag{1}$$

$$\mathbf{G} = \frac{1}{n}\left(n + 1 - 2\,\frac{\sum_{i=1}^{n}(n+1-i)y_i}{\sum_{i=1}^{n} y_i}\right) \tag{2}$$

$$\mathbf{G} = \sum_{i=1}^{n} \left| X_i Y_{i+1} - X_{i+1} Y_i \right| \tag{3}$$

The Gini coefficient can be calculated by Equation (2) without using the areas on the Lorenz curve. In the equation, i denotes each of the n groups, Y_i represents the cumulative percentage of income (Shankar and Shah, 2003). The equation, in which the Gini value is calculated without giving the population information, can be used under the assumption that the population is in an equal distribution. In this case, Equation (3) should be used to calculate the Gini coefficients. In the equation, n is the number of regions, X_i denotes the cumulative population ratio in the i^{th} region, and Y_i denotes the cumulative income ratio in the i^{th} region. It is emphasized that before the calculation to be made using the related equation, the income groups have to be sorted in ascending order (Maclachlan and Sawada, 1997). The resulting G value, as a result of the ratio, is between 0 and 1. If the area A becomes larger (the divergence situation of the Lorenz curve from the absolute equality line), the G value approaches 1, and the inequality increases. If the area A becomes smaller (the convergence situation of the Lorenz curve to the absolute equality line), the value of G approaches 0 and the inequality decreases.

Gini Approach in Measuring the Digital Divide

Different approaches, parameters, and methods can be used to measure the digital divide. Although there is an inequality, it is seen that the Gini coefficients, which have become a standard tool for measuring inequalities, are rarely used in measuring the level of the digital divide. Riccardini and Fazio (2002), one of the first researchers to use Gini coefficients in their studies, calculated the Gini coefficients of the digital divide levels between countries, firms, and individuals and revealed the variables that these coefficients are related. Chakraborty and Bosman (2005), on the other hand, determined the digital divide levels between US states according to gender, ethnicity, and income groups by using Gini coefficients. Jin and Cheong (2008), who calculated the Gini values of internet access and internet usage according to age, gender, and education level, found that there was no significant relationship between the digital divide in internet access and internet usage. Howard et al. (2010), who conducted cross-country analysis using the Gini values of the digital divide, compared Canada and the USA according to education and income groups and concluded that there are higher digital divide levels in the USA. Fidan and Sen (2015) calculated the digital divide levels between genders and those living in urban and rural areas in Turkey by using Gini coefficients and indicated that there is a serious digital divide between genders in rural regions of Turkey. They investigated the sectoral inequalities in the digital divide context and determined the sectoral usage differences of companies' information systems (Fidan and Sen, 2017). In another study, Fidan (2018) showed the digital divide levels regionally.

FINDINGS

The calculation method of the Gini coefficients employed in our study is shown in Table 5. The grouping required for calculating the Gini values was carried out based on the NACE Rev.2 system in accordance with the Eurostat and TUIK reports. By using the percentages presented in Table 3, computer, Internet, website ownerships, and the number of enterprises (e.g., dot-com companies), was calculated according to the sectors, and the number of these firms was used in the study.

Table 5: *Calculation of Gini coefficients*

Sectors	Number of enterprises	Number of enterprises using computer	Ratio of computer usage	Ratio of enterprise	Cumulative ratio of computer usage	Cumulative ratio of enterprise	$\lvert X_i Y_{i+1} - X_{i+1} Y_i \rvert$
D-E	4390	4390	0.0210	0.0211	0.0210	0.0211	0.0000
H	11243	11018	0.0537	0.0529	0.0746	0.0740	0.0001
I	13670	13670	0.0653	0.0657	0.1399	0.1397	0.0000
C	15060	14909	0.0719	0.0716	0.2118	0.2113	0.0001
J	15406	15406	0.0735	0.0740	0.2853	0.2853	0.0001
N	16445	16280	0.0785	0.0782	0.3638	0.3635	0.0003
L	27522	27522	0.1314	0.1322	0.4952	0.4956	0.0004
F	30705	30705	0.1466	0.1475	0.6417	0.6431	0.0006
M	32336	32012	0.1544	0.1538	0.7961	0.7969	0.0008
G	42720	42292	0.2039	0.2031	1.0000	1.0000	0.0000
Sum	**209497**	**208204**					0.0024

Table 5, which shows the Gini value calculation, is based on the number of Danish computer usage in 2013. While preparing this table, after determining the number of companies and the number of companies employing computers, their ratio in the total was calculated. After the sectors are ranked in ascending order according to the number of computers, the cumulative values were determined, and the levels of difference were specified by using Equation (3) in order to be weighted depending on the number of companies. The value of the Gini coefficient, which is the sum of these differences, was calculated as 0.0024. The fact that the Gini value is close to 0 indicates the digital divide related to computer ownership in the Danish sectoral structure is quite low. By applying the method in Table 5, Gini values of information systems in other countries are calculated and presented in Table 6.

If the values listed in Table 6 are examined, it is seen that the Gini coefficients in computer and internet usage are at low levels in all countries. Especially in Finland, there is full equality between the sectors in computer and internet use. Similar to the low levels of the digital divide in computer and internet use that Finland, Sweden, and Denmark have, there is also a similar situation for website ownership. The lowest digital divide level in website ownership is in Denmark with 0.0177 and the highest in Romania with 0.1192. It is observed that the highest Gini values between information systems are in online sales. Besides, it is seen that Denmark, Sweden, and Finland have higher Gini values in online sales transactions compared to other countries, in Finland the highest digital divide by 0.3325 and in Turkey with the lowest level of 0.1460.

Table 6: *Gini values of the countries*

Country	Computer	Internet	Website ownership	Online sales	Avarage
Denmark	0.0024	0.0023	0.0177	0.2656	0.0720
Sweden	0.0061	0.0073	0.0292	0.2485	0.0727
Finland	0.0000	0.0000	0.0231	0.3325	0.0889
Bulgaria	0.0116	0.0190	0.1174	0.1717	0.0799
Romania	0.0269	0.0315	0.1192	0.2598	0.1093
Turkey	0.0178	0.0227	0.0365	0.1460	0.0557

When the differences in information systems usage are averaged, it can be said that there is a more balanced distribution in Turkey compared to other countries. According to the average values of Gini coefficients, the lowest digital divide is in Turkey, and the highest is in Romania. The most remarkable finding is that Finland, which is at the top in many studies, ranks after Romania, where the digital divide is the highest. These results contradict both IDI and NRI results. Although the information systems usage levels are high in Denmark, Sweden, and Finland, which are at the top in IDI and NRI scores, the calculated Gini values are higher than in other countries. These values, which show that the sectoral distributions are more unbalanced in information systems usage, reveal that high usage levels do not mean that the digital divide will be low.

CONCLUSION

The digital divide, which is seen as one of the biggest problems facing countries in the information society process, causes the differences between individuals, companies, regions, and countries to deepen. For this reason, measuring the level of the digital divide is crucial in terms of economic and social policies to be implemented. In general, analyzes on firms/enterprises and sectors are rarely encountered when considering studies to measure the levels of the digital divide between individuals, regions, and countries. In these studies, informatics scores are calculated, or the factors affecting the digital divide are determined. Studies using Gini coefficients, one of the methods used to measure inequalities, are not common in the digital divide literature. This study reveals that the Gini approach can be used for determining the digital divide levels between countries.

In our study, we showed the digital divide levels between sectors in European countries and Turkey. According to the average of Gini values calculated for computer, internet, website ownership, and online sales, we determined that the lowest digital divide is in Turkey, and the highest is in Romania. While we obtained similar findings to previous studies in terms of the computer, internet, and website ownership, we derived different results about online sales. Besides, we demonstrated that Finland, which is shown as one of the countries with the lowest digital divide level, has the highest digital divide in online sales. In this context, we can say that the methods and techniques used in studies making comparisons between countries do not yield healthy results and are lacking in determining the digital divide levels. Using the Gini method instead of indexes such as IDI and NRI, which are used in international comparisons, will provide healthier results and create a common ground for efficient comparisons.

In all cases where there is inequality, we can numerically determine the levels of differences in each data set that can be grouped, thanks to the Gini approach. With the data set in our study, we can analyze a particular sector by grouping them according to countries or information systems. The study can also be carried out to cover all European and world countries by increasing the number of countries included in the analysis. Moreover, the methodology presented in our study can be used in digital divide studies in a personal, regional, or national context.

ACKNOWLEDGEMENT

We would like to express our great appreciation to Dr. Huseyin Fidan (Burdur Mehmet Akif Ersoy University Department of Industrial Engineering) for his valuable and constructive suggestions during the planning and development of this research work.

REFERENCES

Akin, H. B. (2005). *Yeni ekonomi*. Konya: Çizgi Kitabevi.

Attewell, P. (2001). Comment: The first and second digital divides. *Sociology of Education*, *74*(3), 252-259.

Baker, P. M. A. (2001). Policy bridges for the digital divide: Assessing the landscape and gauging the dimensions. *First Monday*, *6*(5).

Bakos, Y. (1998). The emerging role of electronic marketplaces on the Internet. *Communications of the ACM*, *41*, 35-42.

Bayo-Moriones, A., & Lera-Lopez, F. (2007). A firm-level analysis of determinants of ICT adoption in Spain. *Technovation*, *27*(6-7), 352-366.

Brousseau, E., & Curien, N. (2007). *Internet and digital economics*. New York: Cambridge University Press.

Ceriani, L., & Verme, P. (2012). The origins of the Gini index: Extracts from variabilità e Mutabilità (1912) by Corrado Gini. *Journal of Economic Inequality*, *10*(3), 421-443.

Chakraborty, J., & Bosman, M. M. (2005). Measuring the digital divide in the United States: Race, income and personal computer ownership. *The Professional Geographer*, *57*(3), 395-410.

Dewan, S., & Riggins F. J., (2005). The digital divide: Current and future research directions. *Journal of the Association for Information Systems, 6* (12), 298-337.

Drucker, P. F. (2010). *Toward the next economics: and other essays*. Boston: Harvard Business Review Press.

Egger, F. N. (2003). *From interactions to transactions: Designing the trust experience for business-to-consumer electronic commerce*. Eindhoven: Technische Universiteit Eindhoven. Access Services (https://doi.org/10.6100/IR572640)

Ellison, G., & Ellison, S. F. (2005). Lessons about markets from the Internet. *Journal of Economic Perspectives*, *19*(2): 139-158.

Eurostat, (2008). *NACE Rev.2: Statistical classification of economic activities in the European Community*. Luxemburg: Eurostat Methodologies and Working Papers. Retrieved from https://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF

Fidan, H. (2016). Measurement of the intersectoral digital divide with the Gini coefficients: Case study Turkey and Lithuania. *Engineering Economics*, *27*(4), 439-451.

Fidan, H., & Sen, H. (2015). Gini approach to measuring digital divide: Digital divide levels in terms of urban, rural and gender in Turkey. *Journal of International Social Research, 8*(39), 1108-1108.

Fidan, H., & Sen, H. (2017). Sectoral inequalities and relations with the digital divide in Turkey. *Sosyoekonomi*, *25*(32), 31-50.

Fidan, H. (2018). Gini approach for determining of regional digital divide in Turkey. *Business and Economics Research Journal*, 8(1), 49-62.

Goodman, L. A., & Kruskal, W. H. (1959). Measures of association for cross classifications. II: Further discussion and references. *Journal of the American Statistical Association*, *54*(285), 123-163.

Hoffman, D. L., Novak, T. P., & Peralta, M. (1999). Building con trust online. *Comminications of the ACM*, *42*(4), 80-85.

Howard, P. N., Busch, L., & Sheets, P. (2010). Comparing digital divides: Internet access and social inequality in Canada and the United States. *Canadian Journal of Communication*, *35*(1), 109-128.

Huston, J. H., & Spencer, R. W. (2002). Quality, uncertainty and the Internet: The market for cyber lemons. *American Economist*, *46*(1), 50-60.

ITU (International Telecommunication Union). (2014). *Measuring the information society reports*. Retrieved from https://www.itu.int/en/ITU-D/Statistics/Pages/default.aspx.

Jin, J., & Cheong, A. W. H. (2008). Measuring digital divide: The exploration in Macao. *Observatorio Journal*, *2*(3), 259-272.

Kakwani, N. C. (1977). Applications of Lorenz curves in economic analysis. *Econometrica*, *45*(3), 719-728.

Karoly, L. A. (1992). Changes in the distribution of individual earnings in the United States: 1967-1986. *The Review of Economics and Statistics*, MIT Press, *74*(1), 107-115.

Lee, H. G. (1998). Do electronic marketplaces lower the price of goods?. *Communications of the ACM*, *41*(1), 73-80.

Love, P. E. D., Irani, Z., Standing, C., Lin, C., & Burn, J. M. (2005). The enigma of evaluation: Benefits, costs and risks of IT in Australian small-medium-sized enterprises. *Information & Management*, *42*(7), 947-964.

MacLachlan, I., & Sawada, R. (1997). Measures of income inequality and social polarization in Canadian metropolitan areas. *Canadian Geographer, 41*(4), 377-397.

Melnikas, B. (2011). Knowledge economy: Synergy effects, interinstitutional interaction and internationalization processes. *Engineering Economics*, *22*(4), 367-379.

OECD. (2001). Understanding the digital divide. *OECD Digital Economy Papers*, OECD Publishing. Retrieved from https://www.oecd-ilibrary.org/science-and-technology/understanding-the-digital-divide_236405667766

Riccardini, F., & Fazio, M. (2002, August). *Measuring the digital divide*. Paper presented at the IAOS Conference on Official Statistics and the New Economy, London.

Sciadas, G. (2005). Infostates across countries and over time: Conceptualization, modeling, and measurements of the digital divide. *Information Technology for Development*, *11*(3), 299-304.

Sedimo, N. C., Bwalya, K. J., & Du Plessis, T. (2011). Conquering the digital divide: Botswana and South Korea digital divide status and interventions. *South African Journal of Information Management, 13*(1), 1-10.

Shankar, R., & Shah, A. (2003). Bridging the economic divide within countries: A scorecard on the performance of regional policies in reducing regional income disparities. *World Development, 31*(8), 1421-1441.

Spector, R. (2002). *Amazon.com: Get Big Fast*. New York: HarperBusiness.

Solow, R. M. (1987). We'd better watch out. *New York Times Book Review*, 36, 36. Retrieved from http://digamo.free.fr/solow87.pdf

TUIK (Turkish Statistical Institute). (2017) *NACE Rev.2-Altılı ekonomik faaliyet sınıflaması*. Retrieved from https://biruni.tuik.gov.tr/DIESS/SiniflamaTurListeAction.do

Van Dijk, J. A. G. M. (2012). The evolution of the digital divide: The digital divide turns to inequality of skills and usage. In J. Bus, M. Crompton, M. Hildebrandt, & G. Metakides (Eds.), *Digital Enlightenment Yearbook* (57-78). Amsterdam: IOS Press.

Van Dijk, J., & Hacker, K. (2003). The digital divide as a complex and dynamic phenomenon. *The Information Society*, *19*(4), 315-326.

WEF (World Economic Form). (2013). Growth and jobs in a hyperconnected world. In B. Bilbao-Osorio, S. Dutta, & B. Lanvin (Eds.) *The Global Information Technology Report 2013*. Retrieved from http://www3.weforum.org/docs/WEF_GITR_Report_2013.pdf

Wei, L., & Hindman, D. B. (2011). Does the digital divide matter more? Comparing the effects of new media and old media use on the education-based knowledge gap, *Mass Communication and Society*, 14(2), 216-235.

Whinston, A. B., Stahl, D. O., & Choi, S. Y. (1997). *Economics of electronic commerce*. Indianapolis: Macmillan Technical Publishing.

Chapter 46

THE STATE AND IMPORTANCE OF FPGAS IN COMPUTATIONAL ELECTROMAGNETICS

Ali Recai CELIK[1]

1 Dr., Dicle University, Engineering Faculty, Department of Electrical and Electronics Engineering, Diyarbakir, Turkey

1. INTRODUCTION

Several electromagnetic (EM) applications like designing and modelling of the antennas, waveguides, communication systems, fiber optic technologies, radar and satellite systems are not analytically computable due to irregular shapes of the devices. Therefore, some numerical techniques are needed to make these calculations.

There are many computational electromagnetics methods used to solve the problems. The most common of these methods are Finite Element Method (FEM), Finite Difference Time Domain (FDTD) and Method of Moment (MoM). While the first two of these methods are used as differential equation solvers, the other is used as integral equation solver. Implementations of the numerical methods are complex and time consuming since millions of equations must be solved simultaneously when making computations. Hence, accelerating the solver part of these methods is very important and necessary. Field Programmable Gate Arrays (FPGA) can overcome this problem thanks to their many features, especially their parallel processing capabilities.

In this study, FEM, FDTD, MoM and FPGA are introduced and some prominent works related with the FPGA implementations in computational methods are mentioned.

1.1. Electromagnetic Computation Methods

EM computations have an important place in the design of various electrical engineering applications. These computations are made by using some numerical techniques. One of these approximations is FEM which is widely used in modelling the EM problems having complex geometries. It has attracted great attention from researchers with the advancement in technological developments [1]. However, it has many linear equations in large sparse systems. Hence, the great part of the solution run-time is consumed when solving these equations.

Another powerful computational technique is FDTD. This technique provides a direct time-domain solution of Maxwell's Equations whose differential form is given in Fig. 1. The reason this method has become a powerful method for solving a wide variety of electromagnetic areas is that it can solve Maxwell's equations with almost any type of medium [2]. Just like in FEM method, the 3D simulation in FDTD method involves millions of computational cell volumes. It is computationally intensive and simulations must run for a long time on multiprocessor supercomputers. A flowchart for the complete process of the FDTD method is given in Fig. 2 as an example [3].

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} - \sigma_m \vec{H} - \vec{M} \quad (1)$$

$$\nabla \times \vec{H} = \frac{\partial \vec{D}}{\partial t} + \sigma_e \vec{E} + \vec{J} \quad (2)$$

$$\nabla \cdot \vec{D} = \rho_e \quad (3)$$

$$\nabla \cdot \vec{B} = \rho_m \quad (4)$$

$$\vec{B} = \mu \vec{H} \quad (5)$$

$$\vec{D} = \epsilon \vec{E} \quad (6)$$

$\vec{E}$: electric field
$\vec{D}$: electric flux density
$\vec{H}$: magnetic field
$\vec{B}$: magnetic flux density
$\vec{J}$: electric current density
$\vec{M}$: equivalent magnetic current density
σ_e: electric conductivity
σ_m: equivalent magnetic loss

Figure 1. Differential form of the Maxwell's Equations and definitions of the symbols

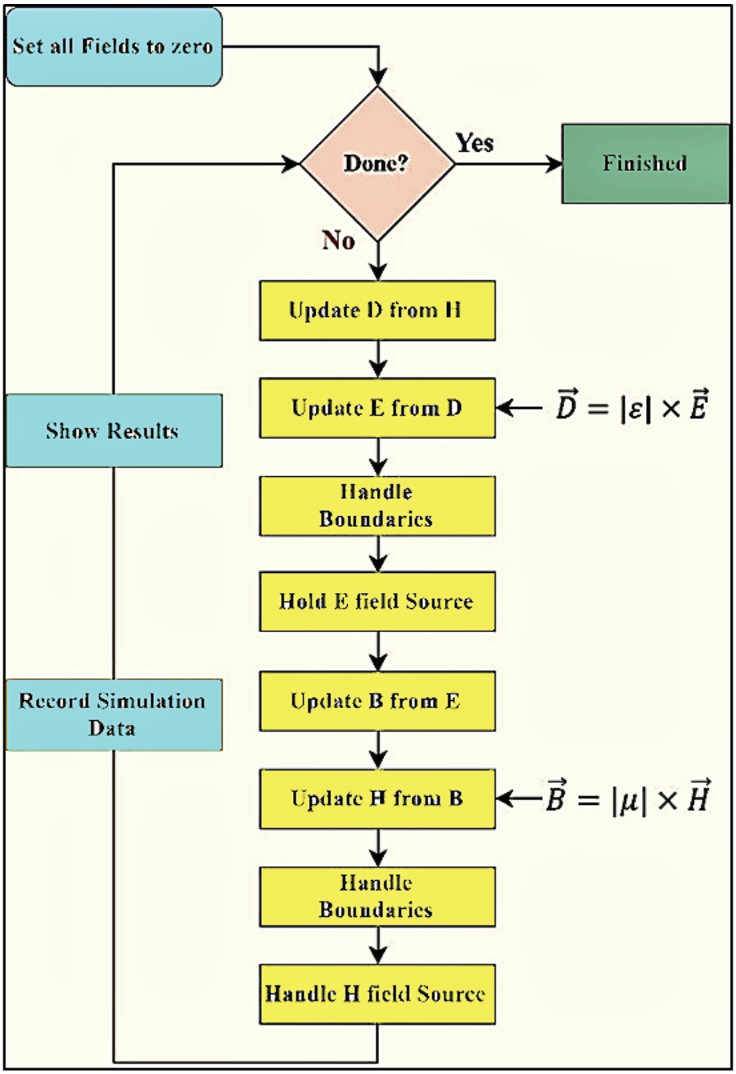


Figure 2. The process of FDTD method [3]

MoM is another basic method used to model the EM problems, which was first proposed by Harrington [4] and is still used in many applications.

MoM is a numerical method based on integral equations in the frequency domain. The main formulation of the method is the integral equations obtained using Green's functions. In this method, the examined structure should be divided into segments. It then needs to be written and resolved as a matrix system. The size of the matrix system depends on the number of segments. As the number of segments increases, the numerical difficulties arise when taking inverse of the impedance matrix. Thus, the calculation volume and its duration increase exponentially in this method.

In summary, implementations of the FEM, FDTD, MoM and other numerical methods are complex and time consuming since millions of equations must be solved simultaneously when making computations. Thus, the solver part of these methods must be very fast.

Software-based solvers are fed into microprocessors and computed serially. Every software instruction must be decoded before starting to execution. The microprocessor must read and decode each instruction sequentially and then execute it. This situation creates an extra overhead for each operation [5]. They don't scale well when the number of equations increase and they are often too slow to be practical for many real-time applications [6].

The usage of Application Specific Integrated Circuit (ASIC) may be another option. However, it has a long development period, costly, and also inflexible. What's more, once the ASIC chip is fabricated, it cannot be modified. If any part of the logic needs to be changed, the chip must be redesigned and refabricated.

FPGAs may be a suitable alternative which can overcome the mentioned problems thanks to their low costs, reconfiguration capabilities, parallel processing features, and speeds enough to work in complex systems. In the next section, the introduction, features and advantages of the FPGA will be mentioned.

1.2. What is FPGA?

FPGA is a general-purpose logic device containing a large number of chips, successfully performs the arithmetic operations and multiplexers can be embedded in. FPGAs are programmable in the field to create logical functions what designers need. They can be defined as integrated circuits whose hardware structures can be changed after production according to the desired function. In other words, FPGAs are the circuits whose transistors are produced independently and freely. Transistors are connected to each other according to the user specified function. Then, the desired application is performed [7]. They can also be reprogrammed run-time, namely reprogramming FPGA partly is possible in case that the system is running on the other part of the chip. These characteristics offer higher elasticity for the FPGAs and appropriateness for several practices. From the outside view, FPGA is a single chip, while the inside structure consists of the adjustable logic blocks, input-output (IO) units around these blocks and programmable connections that connect all these units. The FPGA structure is illustrated in Fig. 3 [8,9].

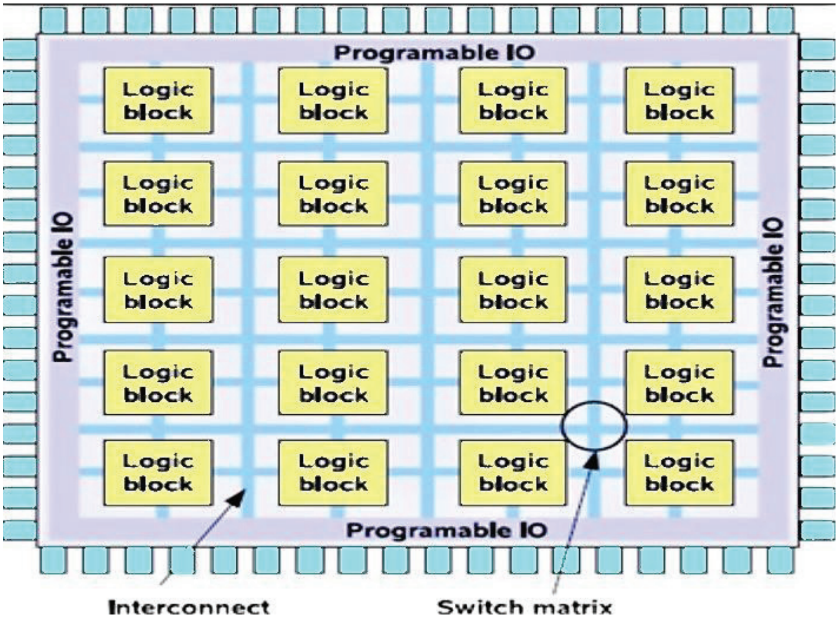


Figure 3. The inside structure of an FPGA

In summary, FPGAs provide the benefits of both ASICs and software-based solvers. Similar to the ASICs, FPGAs could potentially manage a myriad of logic gates, which are optimally designed for a specific function and in a single integrated circuit. Even though FPGA-based designs do not challenge the performance of ASICs, there are few applications that deserve the high cost of very expensive ASICs. Also like software, FPGAs are reprogrammable by designers at any time. Hence, there are flexible and convenient. Although the FPGA-based developments are costly compared to the software, they can show extremely high performance surpassing any programmable solutions [10]. Therefore, the usage of the FPGAs accelerates the solving of numerical methods and has an important role in EM computations. In the next section, some prominent works related with the FPGA implementations in numerical methods will be mentioned.

2. RELATED WORKS

2.1. FPGA Implementations of the FEM

Accelerating the FEM using FPGA for EM computations is one of the most significant works. There are many studies prepared for this purpose. For example, in the study of [11], the researchers introduced a hardware solver going through the most time-inefficient phase in FEM. The device and software of choice were Altera Cyclone II FPGA and Quartus II 9.0 in the study. Pipelining process was fulfilled and the clock speed was recorded nearly at 114.61 MHz. The hardware solver was used to solve the EM boundary value problems. Based on the measurements, the speed was raised by 4 times more than the software performance on a standard computer.

A deeply pipelined FPGA design for efficient sparse matrix-vector multiplication (SpMV), which is a kernel for several iterative numerical techniques to discover sparse linear systems, was developed in [12]. The authors of the paper developed their own pipelinable striping scheme in order to improve the total usability of their own hardware design.

In another study of Zhuo and Prasanna, LU factorization method implemented a direct solver. For the calculations on a Xilinx Virtex- II Pro XC2VP100, a circular linear array of processing elements in double precision became useful. The minimum latency period could be yielded using any design of LU decomposition in that study. The proposed design provided around 4 GFLOPS which was greater compared to the retainable operation of a LU decomposition algorithm exclusively optimized for a 2.2 GHz AMD Opteron processor [13].

Greisen et al. examined many solvent methods and presented a discussion on hardware trade-offs, with demonstration of FPGA architectures of Cholesky direct solver and BiCGSTAB iterative solver. In the study, it was concluded that 32K x 32K matrices were solved at less than 50 fps through the FPGA applications having higher performance than software applications by one or multiple orders of magnitude [14].

In another study related with the FPGA implementation of the FEM, Johnson et al. designed sparse direct LU decomposition and implemented a prototype on FPGA. They monitored the performance with reference to a platform based on universal processor. The results showed that the sparse LU decomposition hardware design application by using FPGA was capable of an order of magnitude speedup relative in comparison with the Pentium 4 based system. These results highlight the usage of FPGAs for high performance sparse direct LU decomposition [15].

2.2. FPGA Implementations of the FDTD

A number of works purposed to speed up the FDTD method for the EM computations. The first use of the FPGA device on this method was proposed in [16]. One-dimensional FDTD cell was implemented on FPGA and the activities were successfully simulated in that work. A pipelined bit-serial arithmetic architecture helped to transfer the algorithm to FPGA-based hardware. The results of the paper showed that the proposed hardware design made the one-dimensional algorithm quicker than the software design.

Kosmas et al. introduced and tested a pseudo-2D FDTD hardware architecture [17]. The study demonstrated that the results of software and hardware implementations. According to the obtained results, the qualified algorithms and the capable resources facilitated the hardware implementation to raise the computational speed up to the peak. The hardware speed was corresponding to that of 24 times the fastest software implementation running on 3.0 GHz PC.

Durbano and Ortiz accelerated the FDTD using a customized accelerator board compatible with the specified configurations (DDR SDRAM ≤ 16 GB, DDR SRAM=36 MB) as well as Xilinx Virtex-II 8000 FPGA and an external PCI controller. Although the accelerator cards used in that study required multiple-PC solutions, such weakness did not obstruct a 23-fold acceleration and an almost 6-fold hike at its zenith of problem size [18].

An FPGA-based architecture using OpenCL was shown in [19]. This architecture was proposed to apply the iteration level parallelism in order to decrease the external memory access. In that study, the processing speed reached up to relatively 4 or more times higher level to the GPU implementation and 13 times higher level to the CPU.

In the study [20], a method using the FPGA which employs a 32-bit operation machine was investigated to short the required time for the FDTD process. For this aim, an analysis about the processing time for an FPGA implementation was made. The authors of the paper focused on decreasing the number of data taken from the DDR memory. It was shown that the used techniques reduced the computation time by 22 times compared to the time required by a computer using commercial software.

2.3. FPGA Implementations of the MoM Studies

There are several studies presenting the FPGA-based accelerator for MoM solutions. Using an FPGA-based hardware implementation, the study of [21] offered parallellization of the iterative matrix solution for multiple RHS vectors in a low-rank compression-based fast solver scheme.

This technique enhanced the capacity of multiple conductors to extract the electrostatic parasitic capacitance in a Ball Grid Array (BGA) package. In another study, the same researchers proposed similar hardware design toward a double-leveled parallelization scheme. The authors used a 2.4 GHz Intel Core i5 processor and observed that speedup was linearly scalable with FPGA resources and the software implementation had 10x higher velocity than equivalent one. Virtex-6 XC6VLX240T FPGA on Xilinx's ML605 board was used in that study. The authors also emphasized that the proposed method could be applied to other forms of EM extractions.

Another study investigated the CPU/FPGA application of the matrix assembly phase of the (MoM), based on the frequency-domain integral equation-based formulation including piecewise linear function and the image methods to appraise the radiation pattern of the antenna system [22]. The code executed on the device ran only one single work-item OpenCL kernel in order to take advantage of the proposed architecture. It was reported that the speedup ratios were computed as relatively approximately 2.08× with reference to the single-core CPU approach and as nearly 1.11× to two-core CPU implementation.

3. CONCLUSIONS

There are several numerical methods used to solve computational EM problems. In this study, FEM, FDTD and MoM, which are the most known and preferred methods, have been briefly introduced. It has been mentioned that implementations of these methods are complex and time consuming since millions of equations must be solved simultaneously in calculation process. It has been emphasized that speed up the solvent part of the methods is very important and necessary. The merits and drawbacks of the sofware-based, ASIC-based and FPGA-based solver systems have been discussed. The properties and structures of the FPGAs have been shown in detail. Several works related with the FPGA implementations have been reviewed. The results of these works indicate that FPGA hardware implementation can achieve significant acceleration for solving numerical method algorithms. Finally, it can be concluded that the usage of FPGAs has an important role in computational electromagnetics.

REFERENCES

[1] Strang, G. & Fix, G. (1973). An Analysis of the Finite Element Method: Englewood Cliffs. *Prentice Hall*, New Jersey, USA.

[2] Kosmas, P., Wang, Y. & Rappaport, C. (2002). "Three-Dimensional FDTD Model for GPR Detection of Objects Buried in Realistic Dispersive Soil," *SPIE Aerosense Conference*, Orlando, pp.330–338.

[3] Memon, N., Naich, M.R., Pathan, A.Z., Mirjat, B.A. & Faiz, M. (2019). "Computational Investigations on CZTS Thin-Film Layers Adopting Grating Structures," *SSRG International Journal of Electronics and Communication Engineering,* 6(8), pp. 38–43.

[4] Harrington, R.F. (1968). "Field Computation by Moment Methods," *Wiley IEEE Press.* ISBN: 978-0-780-31014-8

[5] Compton, K. & Hauck, S. (2002). "Reconfigurable Computing: A Survey of Systems and Software," *ACM Computing Surveys,* 34(2), pp.171–210.

[6] Tarek Ibn Ziad, M., et al. (2015). "Finite Element Emulation-Based Solver For Electromagnetic Computations," *IEEE International Symposium on Circuits and Systems (ISCAS)*, Lisbon, pp. 1434–1437.

[7] Celik, A.R. & Alkan, A. (2015). "Implementation of Basic Image Processing Applications by FPGA Hardware," *International Journal of Scientific and Technological Research*, 1(1), pp. 133–143.

[8] Villasenor, J. & Mangione-Smith, W.H. (1997). "Configurable computing," *Sci. Amer.,* 276(6), pp. 66–71.

[9] Shawahna, A., Sait, S.M. & El-Maleh, A. (2019) "FPGA-Based Accelerators of Deep Learning Networks for Learning and Classification: A Review," *IEEE Access*, vol.7, pp. 7823–7859.

[10] Hauck, S. (1998). "The Roles of FPGAs in Reprogrammable Systems," *Proceedings of the IEEE*, 86(4), pp.615–639.

[11] Zhang, J., Zhang, M., He, H. & Song, Q. (2015). "Accelerating the finite element method using FPGA for electromagnetic field computation," *2015 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER)*, Shenyang, pp. 1763–1768.

[12] El-Kurdi, Y., Giannacopoulos, D. & Gross, W. J. (2007). "Hardware Acceleration for Finite-Element Electromagnetics: Efficient Sparse Matrix Floating-Point Computations With FPGAs," *IEEE Transactions on Magnetics*, 43(4), PP.1525–1528.

[13] Zhuo, L. & Prasanna, V.K. (2006). "High-Performance and Parameterized Matrix Factorization on FPGAs," *International Conference on Field Programmable Logic and Applications*, Madrid.

[14] Greisen, P., Runo, M., Guillet, P., Heinzle, S., Smolic, A., Kaeslin, H. & Gross, M. (2013). “Evaluation and FPGA Implementation of Sparse Linear Solvers for Video Processing Applications,” *IEEE Transactions on Circuits and Systems for Video Technology*, 23(8), pp. 1402–1407.

[15] Johnson, J., Chagnon, T., Vachranukunkiet, P., Nagvajara, P. & Nwankpa, C. (2008). “Sparse LU Decomposition Using FPGA,” *Int. Workshop State-of-the-Art Scientific Parallel Comput. (PARA).*

[16] Schneider, R.N., Turner, L.E. & Okoniewski, M.M. (2002). “Application of FPGA Technology to Accelerate the Finite-Difference Time-Domain (FDTD) Method,” *ACM/SIGDA Tenth International Symposium On Field-Programmable Gate Arrays (FPGA) Association for Computing Machinery*, New York, USA, pp. 97–105.

[17] Kosmas, P., Wang, Y. & Rappaport, C. (2002). “Three-Dimensional FDTD Model for GPR Detection of Objects Buried in Realistic Dispersive Soil,” *SPIE Aerosense Conference*, Orlando, pp.330–338.

[18] Durbano J.P. & Ortiz, F.E. (2004). “FPGA-based Acceleration of the 3D Finite-difference Time-domain Method,” *12th Annual IEEE Symposium on Field-Programmable Custom Computing Machines,* pp.156–163.

[19] Waidyasooriya, H.M., Endo, T., Hariyama, M. & Ohtera, Y. (2017). “OpenCL-Based FPGA Accelerator for 3D FDTD with Periodic and Absorbing Boundary Conditions,” *International Journal of Reconfigurable Computing*, vol. 2017.

[20] Suzuki, H., Takagi, Y., Yamaguchi, R. & Uebayashi, S. (2005). “FPGA Implementation of FDTD Algorithm,” *Asia-Pacific Microwave Conference Proceedings,* Suzhou, pp. 1–4.

[21] Devi, A., Gandhi, M., Varghese, K. & Gope, D. (2016). “Accelerating Method Of Moments Based Package-Board 3D Parasitic Extraction Using FPGA,” *Microw. Opt. Technol. Lett.*, vol. 58, pp. 776–783.

[22] Topa, T. (2020). “Porting Wire-Grid MoM Framework to Reconfigurable Computing Technology,” *IEEE Antennas and Wireless Propagation Letters*, 19(9), pp. 1630–1633.

Chapter 47

AN ONTOLOGY FOR MODELING THE DENTAL PATIENT

***Özgün YILMAZ*[1]**

1 Assoc. Prof. Dr., Ege University, Department of Computer Engineering, İzmir, Turkey, ozgun.yilmaz@ege.edu.tr

1. Introduction

As a biological, a psychological and a social being, a human is a complex organism. The human mouth just like the other body parts interacts with the rest of the body. While considering a dental treatment for a patient, a patient's biological, psychological, and social status should be taken into consideration as a whole. A dentist should be knowledgeable and familiar with these characteristics of the patient. It is essential that this knowledge should be stored systematically for preventing endangering the patient's life due to medical errors and for deciding as well as carrying out the most appropriate dental treatment for the patient. Unfortunately, there is not any formal dental patient model for electronically storing and modeling patient characteristics and behavior currently.

This paper describes the proposed formal dental patient model ontology, Dental Patient Model Ontology (DPMO), which is available online from https://github.com/ozgunyilmaz/DPMO. DPMO is the first formal dental patient model which will make it possible to capture different patient characteristics before any dental procedure. DPMO is encoded in Web Ontology Language (OWL) and is used in a prototype dental information system.

DPMO is used by the prototype system for the knowledge presentation about the dental patient and the related data integration. Although medical knowledge can be represented with a single ontology, this is impractical in a complex domain like medicine which contains vast amount of information about each medical specialty and patients. Decomposing large ontologies into smaller modules provides flexibility and makes medical knowledge and data easier to manage and update. As a result, the ontology modules can be integrated to compose the whole domain (Pathak, Johnson, & Chute, 2009).

By using DPMO and rule-based reasoning, treatment options which are most suitable to the patient characteristics can be presented to the dentist since DPMO makes it possible to achieve a more complex and complete dental patient representation.

In this paper, only DPMO which models the dental patient is described. The subsequent sections of this paper are structured as follows: In Section 2, semantic web and ontology concepts are introduced. In Section 3, factors which are important in modeling dental patient are discussed as well as DPMO and how it is used. Section 4 shows how logic-based reasoning can be used to exploit DPMO. In Section 5, related research is surveyed. Finally, in Section 6, the conclusion is discussed.

2. Ontology and Semantic Web

OWL is a semantic web language designed to represent the rich and complex knowledge describing a domain. The knowledge is represented by using concepts and relations between these concepts. The knowledge expressed in OWL is machine readable and can be exploited by the computer software to verify the consistency of that knowledge or to make the implicit knowledge explicit by using logical inference (Berners-Lee, Hendler, & Lassila, 2001).

Advantages of ontologies are as follows:

- **Knowledge sharing:** Ontologies support knowledge sharing by defining a common set of concepts. Computational entities such as agents and services can interact with each other by using ontologies.

- **Logical inference:** Logical inference can be performed over ontology by using a rule language.

- **Knowledge reuse:** Ontologies support knowledge reuse. An ontology can import and use other ontologies which are developed by other people from different resources.

OWL is modeled through an object-oriented approach and a domain is described by using classes, properties, and axioms. OWL classes are used to represent a concept. There are two types of properties: the object property, which is a complex relation between OWL classes, and the data type property, which is a relation between an OWL class and a primitive data type (string, integer, etc.). There are also individuals, which are instances of OWL classes.

3. DPMO

DPMO is developed using Protégé (available from http://protege.stanford.edu) ontology editor and is encoded in OWL. A general view of DPMO is shown in Figure 1. The figure shows OWL classes, object properties, and datatype properties. The range of datatype properties is not shown in the figure because they are of primitive types, such as string, integer, etc. Also, values of datatype properties are restricted by using *owl:FunctionalProperty* and *owl:one* of built-in OWL properties. Thus the users must choose maximum one value from an enclosed set of values. In DPMO, the age constraints, dental history, expectations, financial status, social status, mental traits and the preliminary examination can be modeled and stored. For flexibility, these concepts are defined using OWL classes.

These concepts can be explained as follows:

- **Preliminary examination:** Understanding the patient and his or her needs are essential for selecting the most appropriate dental treatment.

In addition to the detailed knowledge of the patient's dental condition, it is important to get a thorough insight into his or her emotional makeup. During the preliminary examination which can be conducted at the initial appointment, the patient's psychological profile is examined. A smart clinician can get this kind of valuable information during the contacts with the patient by observing the patient's behavior. In the preliminary examination, the patient's characteristics are modeled by using OWL datatype property. The modeled characteristics under preliminary examination are explained as follows (Grasso & Miller, 1991):

o Neuromuscular coordination: This characteristic is important especially in prosthodontics (prosthetic dentistry). It can be judged from patient's gait or other movements. If a patient's neuromuscular coordination is weak, then the patient may have difficulty using his/her prosthesis.

o Vigorous or phlegmatic: If the patient is phlegmatic, chewing efficiency may be low. This should be considered during treatment planning.

o Tidy or untidy: Untidiness may indicate a low quality of oral hygiene. A patient with poor oral hygiene should be educated by the dentist about the importance of good oral hygiene for the overall well-being of a person as well as how to take care of one's teeth.

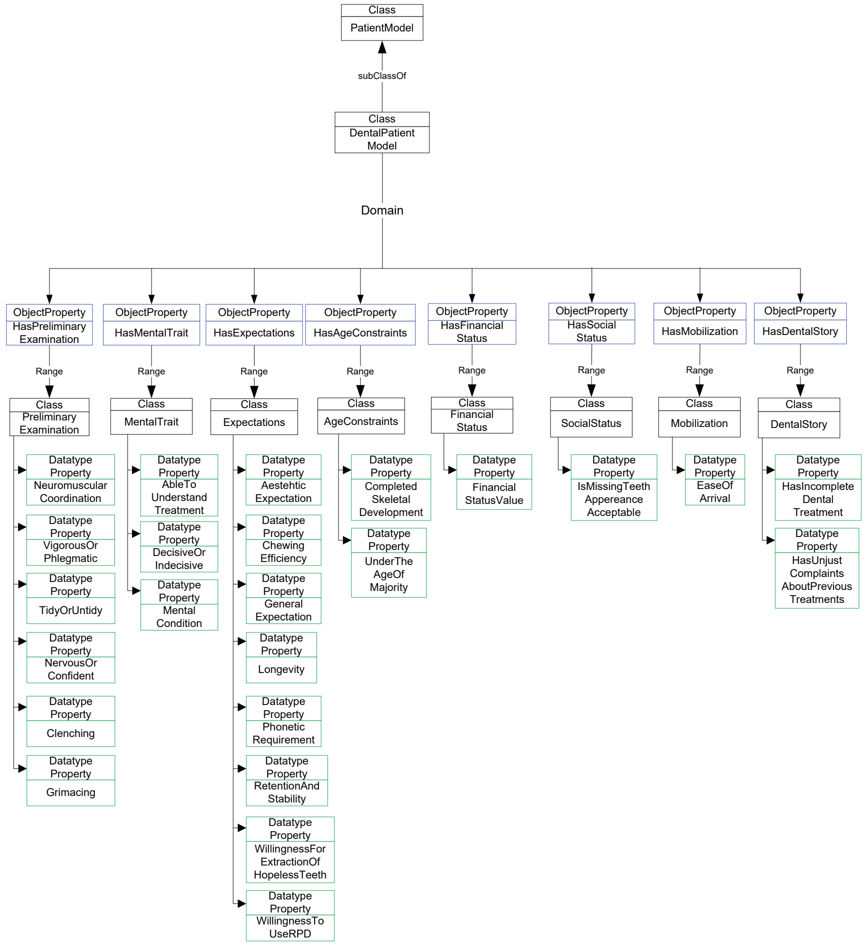


Figure 1. *A general view of DPMO*

o Nervous or confident: This characteristic can be important in treatment planning. A moist, limp hand may indicate nervousness. On the other hand, a firm handshake with a dry hand might suggest confidence.

o Clenching and grimacing: These habits may affect the success of a treatment and indicate that the patient has a lower than average discomfort threshold.

• **Mental trait:** This concept is modeled to see if a patient can perceive reality and/ or can follow through with the treatment. The modeled characteristics under mental trait are explained as follows (Stewart, Rud, & Kuebker, 1983):

o Able to understand treatment: Is the patient able to comprehend his/her dental condition and the treatment plan which needs be carried out?

If the patient is unable to or does not want to understand the treatment plan or if he/she insists on an impossible treatment option, then there is a possibility of a conflict with the dentist.

o Decisive or indecisive: Some indecisive patients may change their minds after the treatment starts. Therefore, these patients should be given extra time for thinking before starting a treatment. Also, there is a possibility of a conflict with the patient.

o Mental condition: If the patient is mentally handicapped or mentally ill, this condition also should be taken into consideration.

• **Expectations:** In this concept, the patient's expectations from the dental therapy are modeled. Expectations may include property values such as the aesthetic expectation, chewing efficiency, general expectations, longevity of the dental work, phonetic requirement, retention and stability as well as willingness for extraction of hopeless teeth and willingness to use removable partial dentures.

• **Age constraints:** In this concept, the age-related patient characteristics are modeled as follows:

o Completed skeletal development: If the patient has not completed his/her skeletal development, then a temporary treatment can be carried out.

o Underage minor: If the patient is an underage minor, then a parent or the legal guardian should be present during examination and treatment process.

• **Financial status:** Patient's financial status is modeled. Treatment options with respect to the patient's financial status can be recommended by using this value.

• **Social status:** The patient's social status is modeled. This includes if a patient's social, professional or psychological status is appropriate for appearing with missing teeth in public as the treatment progresses.

• **Mobilization:** In this topic, the patient's ease of arrival to the dental clinic is modeled. According to this factor, it may be needed to keep number of appointments to a minimum or possibly it may be necessary for the dentist to go to the patient instead.

• **Dental story:** In this topic, the patient's past dental treatments are modeled. If the patient has unjust complaints about his/her past treatments or he/she has had incomplete treatments, then there is a possibility of a conflict with the current dentist also (Grasso & Miller, 1991; Stewart et al., 1983).

As mentioned earlier, DPMO is used in a dental information system. In Figure 2, the dental patient modeling window is shown. This window is created dynamically by parsing DPMO. Every object property (*HasExpectations*, *HasPreliminaryExaminations*, etc.) which is related to the dental patient model is shown in a panel in this window. The dentist can modify property values by clicking buttons. Existing individuals can be added by clicking the button with the text "*", new values can be added by clicking the button with the text "+", and added values can be removed by clicking the button with the text "-" button. For example, when the dentist clicks to add new expectations, the window shown in Figure 3 is shown.

4. Ontology Reasoning

In this section, how DPMO can be utilized by applying ontology reasoning is described. As a result of processing the patient information with the logical reasoning, new facts can be discovered. This derived information can be used to provide recommendations to dentists. For example, if a patient wants a dental treatment with high chewing efficiency, then treatments with low chewing efficiency are discarded or emphasized to the dentist. In Table 1, some rules used by the developed system and their explanations are given. In these rules, treatment recommendations according to the patient model are given.

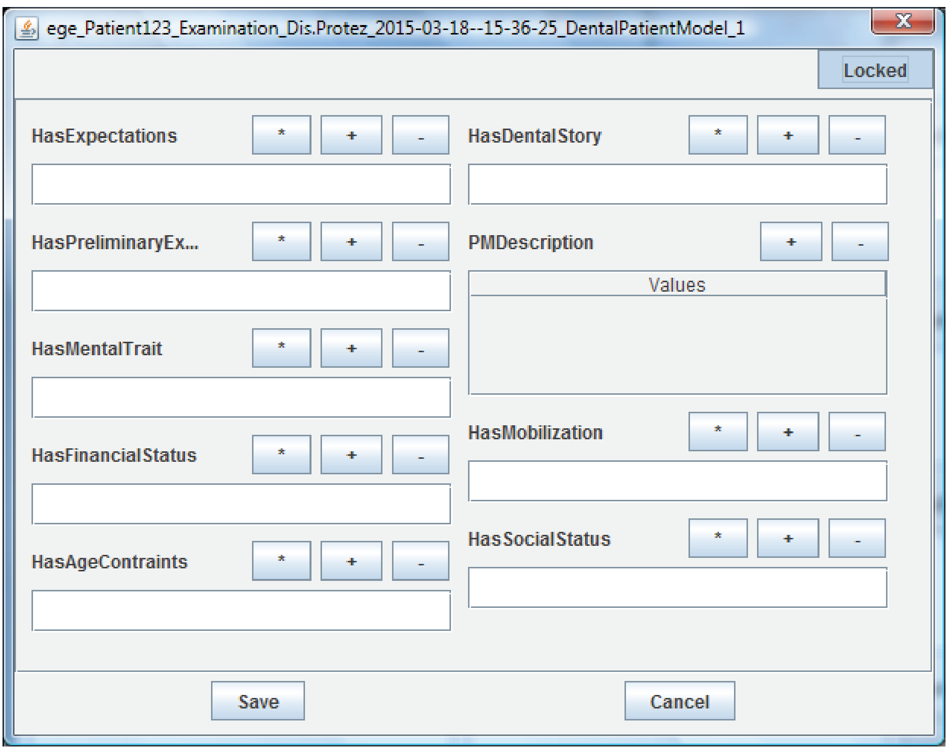


Figure 2. *Dental patient modeling window*

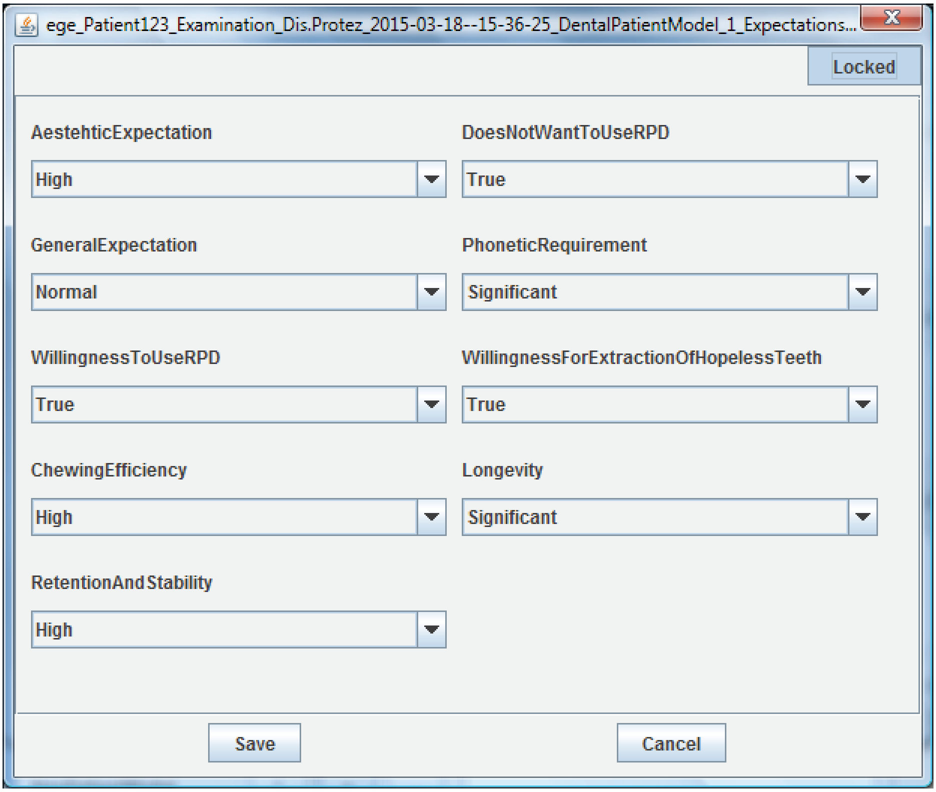


Figure 3. *Dental patient's expectations modeling window*

Table 1. *Voice commands in given order*

Rule	**Description**
(?p HasConsecutiveMissingTeeth 2) ^ (?p HasPatientModel ?pm) ^ (?pm HasExpectation ?exp) ^ (?exp chewingEfficiency "High") => (?p HasTreatmentOption DentalBridge)	If the patient has 2 consecutive missing teeth wants high chewing efficiency, then dental bridge should be applied as a treatment.
(?p HasConsecutiveMissingTeeth 4) ^ (?p HasPatientModel ?pm) ^ (?pm HasFinancialStatus ?fs) ^ (?fs financialStatusValue "BelowAverage") => (?p HasTreatmentOption RemovablePartialDenture)	If the patient has 4 consecutive missing teeth and the patient's financial status is below average, then removable partial denture can be applied.
(?p HasPatientModel ?pm) ^ (?pm HasAgeConstraint ?ac) ^ (?ac underAgeMinor "True") => (?p HasTreatmentWarning "Custodian's approval needed")	If the patient is under the age minor, then an elder custodian's approval is required.

(?p HasPatientModel ?pm) ^ (?pm HasDentalStory ?ds) ^ (?ds HasIncompleteDentalTreatment “True”) ^ (?ds HasUnjustComplaintsAboutPreviousTreatments “True”) => (?p HasTreatmentWarning “Possibility of Conflict”)	If the patient has previous incomplete dental treatments and has unjust complaints about his/her previous treatments, then there is a risk of conflict.

5. Related Work

The systems which have previously used similar ontologies in dentistry have been researched and have been documented in this section. As of now, to the best of our knowledge, there is no dental patient model ontology, other than the one presented in this paper.

Kiani, Francis, Zand-Moghaddam, and Verma (2010) propose a negotiation system which uses semantic web concepts, ontology, and semantic reasoning to facilitate the negotiation process. The system provides a common layer for negotiation by mapping local concepts of each individual to global concepts. The target domain of the system is the negotiation of dentists for the treatment of wisdom teeth (Kiani et al., 2010).

S. Park and Kim (2006) propose a system comprising of a tooth positional ontology and SWRL rules to assist dentists in the decision-making process on ways of replacing a missing tooth. This assistance is provided by reasoning of the use of SWRL rules over the ontology (S. Park & Kim, 2006).

S. G. Park, Lee, Kim, and Kim (2012) propose a shared decision-making system where the doctor and the patient agree on a health care or a treatment decision with its focus on dental restorative treatment planning. The dental patient's preferences about the treatment are taken into consideration during the decision-making process. Treatment options are generated by employing an ontology which captures the clinical knowledge required for treatments (S. G. Park et al., 2012). In this system, the patient preferences are modeled, but biological, psychological, and social factors of the patient are not taken into consideration.

Duncan, Schleyer, and Ruttenberg (2013) refers to The Oral Health and Disease Ontology (OHD) which contains, among others, terms for the diagnosis and the treatment of dental maladies. The authors demonstrate how an intra-coronal tooth restoration is represented in the OHD and discuss some of the ways in which the OHD enables information sharing amongst researchers (Duncan et al., 2013).

All of the previously reviewed work, which somehow use ontologies, don't mention about modeling the dental patient with regards to biological, psychological and social characteristics. Only in the work conducted by S. G. Park et al. (2012), the patient preferences are taken into consideration.

In DPMO, the age constraints, dental history, expectations (preferences), financial status, social status, mental traits, and the preliminary examination are modeled and stored. For flexibility, these concepts are defined using OWL classes and they can be extended by defining related properties.

6. Conclusion

A human being is a complex organism. While considering a dental treatment for a patient, there are many factors such as patient's biological, psychological, and social status which should all be considered by the dentist before making a decision about the treatment. Unfortunately, there is not any formal dental patient model, modeling patient characteristics and behavior currently.

This paper describes the proposed formal dental patient model ontology, DPMO. DPMO is the first formal dental patient model which will make it possible to capture different patient characteristics before any dental procedure; therefore, supporting dentists' decision processes by making appropriate recommendations to them. As a result of this, not only the patient satisfaction can be improved significantly but the most suitable treatment for the patient can be decided upon also.

DPMO is integrated into a prototype dental information system. This system is used as a test bed and if new requirements emerge during the pilot use, then the system will be further improved by making modifications to the system.

References

[1] Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. Sci Am, 284(5), 28-37.

[2] Duncan, W., Schleyer, T., & Ruttenberg, A. (2013, 2013). Representing intracoronal tooth restorations in the ontology for oral health and disease. Paper presented at the Proceedings of the 4th International Conference on Biomedical Ontology.

[3] Grasso, J., & Miller, E. (1991). Removable Partial Prosthodontics: Mosby-Year Book Inc.

[4] Kiani, M., Francis, M., Zand-Moghaddam, Y., & Verma, P. (2010). Ontology-based negotiation of dental therapy options. In M. Joshi, H. Boley, & R. Akerkar (Eds.), Advances in Semantic Computing. TMRF India (pp. 52-78).

[5] Park, S., & Kim, H. G. (2006). Dental decision making on missing tooth represented in an ontology and rules. In R. Mizoguchi, Z. Shi, & F. Giunchiglia (Eds.), The Semantic Web - ASWC 2006. Lecture Notes in Computer Science (Vol. 4185, pp. 322-328). Berlin: Springer.

[6] Park, S. G., Lee, S., Kim, M.-K., & Kim, H.-G. (2012). Shared decision support system on dental restoration. Expert Syst Appl, 39(14), 11775-11781. doi:https://doi.org/10.1016/j.eswa.2012.04.074

[7] Pathak, J., Johnson, T. M., & Chute, C. G. (2009). Survey of modular ontology techniques and their applications in the biomedical domain. Integrated computer-aided engineering, 16(3), 225-242. doi:10.3233/ICA-2009-0315

[8] Stewart, K., Rud, K., & Kuebker, W. (1983). Clinical Removable Partial Prosthodontics: Mosby.