# RESEARCH & REVIEWS IN ENGINEERING - II

**SEPTEMBER/2021**

EDITORS

PROF. DR. BANU NERGİS
ASSOC. PROF. DR. SELAHATTİN BARDAK
ASSOC. PROF. DR. MAHMUT KAYAR
DR. ARİF FURKAN MENDİ

gece
kitaplığı

# RESEARCH & REVIEWS IN ENGINEERING - II

## September / 2021

### EDITORS
### PROF. DR. BANU NERGİS
### ASSOC. PROF. DR. SELAHATTİN BARDAK
### ASSOC. PROF. DR. MAHMUT KAYAR
### DR. ARİF FURKAN MENDİ

# CONTENTS

# Chapter 1

## EVALUATION OF SOLID WASTE MANAGEMENT AND SOLID WASTE-INDUCED GREENHOUSE GAS EMISSIONS: TURKEY

*Oylum GÖKKURT BAKİ[1]*

---
1   Assoc. Prof. Dr. Oylum GÖKKURT BAKİ, Sinop University Faculty of Engineering and Architecture, Environmental Engineering Department, Sinop, Turkey (ORCİD Id: 0000-0001-7823-0824) ogbaki@sinop.edu.tr

## Introduction and Objectives

The sustainable management of wastes within the scope of environmental management systems has been on the agenda both around the world and in Turkey for long years. Wastes have recently been viewed as potentially useful raw materials and energy resources. However, the waste management policies adopted and implemented in Turkey have employed a linear economic perspective. In recent years, especially "zero waste"-based applications have come to the forefront as measures for waste reduction and sustainable management-based planning.

Climate change is, today, indisputably the most investigated and discussed environmental issue. However, its waste-induced impact is sadly discussed to a lesser degree despite the significant percentage of the release of greenhouse gas emissions due to unplanned solid waste management and resulting methane ($CH_4$) and carbon dioxide ($CO_2$) release, which worsens global warming. Turkey still encounters problems in some cities especially regarding issues with solid waste management and solid waste-induced problems will continue to be topics of discussion and research. We should keep in mind that solid waste-induced issues and emissions are not local problems. Thus, they should be treated accordingly as global issues. On the other hand, the prevention of problematic and poorly managed solid waste-induced emissions should be planned and resolved using local facilities and considering local environmental factors, since the release of emissions of harmful materials such as solid waste-induced greenhouse gases is among the most powerful culprit of climate change.

The time scales for climate change and waste management are similar. In most developed and developing countries with increasing population, prosperity, and urbanization, it remains a major challenge for municipalities to collect, recycle, treat, and dispose of increasing quantities of solid waste, especially in a changing climate. A cornerstone of sustainable development is the establishment of affordable, effective, and truly sustainable waste management practices in developing countries. It must be further emphasized that multiple public health, safety, and environmental co-benefits accrue from effective waste management practices, which concurrently reduce GHG emissions, improve the quality of life, promote public health, prevent water and soil contamination, conserve natural resources, and provide renewable energy benefits (Christian, 2010).

The 2013 report of ISWA and UNEP states that methane emissions from solid waste management activities, namely landfills, contribute 15% to global anthropogenic methane emissions. The report also mentions that the global warming effect factor of methane gas is higher than carbon dioxide ($CO_2$) and points to the lower reliability of the data on fugitive

CH4 emissions originating from uncontrolled landfills than that of direct data on GHG emissions originating from other sections of the waste sector (ISWA/UNEP, 2013).

Landfill gas emissions increases as waste accumulates. The implication is that emissions are set to increase considerably in line with increase in population and waste generation. In addition, incineration and open dumping and burning of waste are important sources of harmful substances and pollutants (Monni et al., 2006; Sharma et al., 2015; Onyanta, 2016). In general, municipal solid waste (MSW) sector is the fourth largest supplier to global emissions of non-$CO_2$ GHGs that contribute towards global warming and climate change due to their emissions and that approximately contribute 5.5–6.4% towards global methane emissions annually IPCC,2006). The MSW landfills are the third largest anthropogenic source of methane, followed by agricultural and enteric fermentation, accounting for 11% of the global methane emissions as biogas (Kropp & Reckien, 2009; Onyanta, 2016).

Turkey produced 32.2 million tonnes of waste overall per year and it means that waste per capita is approximately 1,16 kg per day and also the average recycling rate of all waste in Turkey is 12.3% according to 2018 records of TurkStat. According to these data, 20.24% of the wastes are sent to the municipal garbage disposals, 67.20% is disposed of by the regular storage method and 38% is recycled by composting. Also, according to the 2016 data, wild storage in Turkey continues at 27% (Gökkurt Baki and Ergun, 2021).

The study evaluates solid waste management and solid waste-induced greenhouse gas emissions in Turkey. By doing so, the study aims to examine the contribution of solid wastes to the impact of global warming and raise awareness of it. Therefore, the importance of the reduction of wastes at their sources, the impact of the reduction of solid waste-induced greenhouse gas emissions, and the importance of increasing the capacity allocated for this purpose were emphasized.

**Solid Waste, Solid Waste Management, and Related Legislation in Turkey**

Valuable materials that are not desired at their original place are called solid wastes (Curi, 1990). In the "Regulation on Control of Solid Wastes" published in the Official Gazette dated March 14, 1991, solid wastes are defined as solid materials and treatment sludge that are intended to be disposed of by the producer and that must be disposed of regularly to protect the public peace and especially the environment. In other words, solid wastes are undesired or discarded wastes that are normally solid and originate from human and animal activities. The term includes all

wastes resulting from industries, settlements, businesses, and agricultural activities (Peavy et al., 1985; Henry & Heinke, 1989; Tchobanoglous et al., 1993). In the "Waste Management Regulation" published in the Official Gazette, dated April 2, 2015, waste is defined as "any substance or material that is thrown or released into the environment or should be disposed of by the producer or the real or legal person who actually holds it". Waste management in the Waste Management Regulation is defined as the "prevention of waste, reduction at source, reuse, separation according to its properties and type, accumulation, collection, temporary storage, transportation, intermediate storage, recycling, recovery including energy recovery, disposal, monitoring, control, and supervision after disposal activities" (EUM, 2015; Figure 1).



**Figure 1.** *Solid waste (original)*

There have been various laws and regulations on the applications regarding solid waste management in Turkey until today. The aims of these regulations in order of years are as follows (Gökkurt Baki, 2021);

-30829 / 2019 Regulation on Zero Waste The sustainable and effective management of raw materials and natural resources

-30283 / 2017 Regulation on the Control of Packaging Waste Prevention of packaging wastes, determination of the principles and policies for the reuse, recycle, and recovery of unavoidable packaging wastes

-29959 / 2017 Regulation on Control of Medical Wastes Determination of all principles, policies, programs, and legal, managerial, and technical principles for the processes from the formation to the disposal of medicinal wastes

-29314 / 2015 Regulation on Waste Management Procedures and principles regarding the management of wastes from their formation to their

disposal without harm to the environment and human health, reduction of waste formation, reduction of the use of natural resources using methods such as the reuse and recovery of wastes

-27533 / 2010 Regulation on the Landfill of Waste Determination of the technical and managerial issues and general rules regarding the disposal of wastes by regular storage

-27721 / 2010 Regulation on Waste Incineration Minimization of the negative effects of waste incineration on the environment and the risks waste incineration poses to human health

-27448 / 2009 Regulation on Control of End-of-Life Vehicles Prevention of the formation of vehicle wastes, reuse of end-of-life vehicles and their parts, reduction of waste using the recycle and recovery processes

-28300 / 2004 Regulation on the Control of Waste Electrical and Electronic Equip Determination of the methods for the reuse, recycle, and recovery of electric and electronic devices to reduce the electric and electronic waste formation and amount of waste to be disposed from their production to disposal

-25569 / 2004 Regulation on the Control of Waste Batteries and Accumulators Determination of the principles, policies, and programs regarding the production of batteries and accumulators that meet certain environmental criteria, terms, and properties and development of a management plan for their recovery and ultimate disposal

-25682 / 2004 Regulation on Recovery of Waste and Control of waste from Ships Determination of the procedures and principles for waste reception vessels to prevent the discharge of load wastes and wastes that are produced by the ships in the maritime jurisdiction of Turkey to seas and protect the marine environment.

In Turkey, certain regulations and applications have been adopted in Turkey, with the "Zero Waste" applications being among them. The Zero Waste Project was launched by the Ministry of Environment and Urbanization in 2017. To date, 18.750 public institutes have joined the project and total of 126-ton paper and 8.7-ton glass were recycled. Moreover, the Ministry of Environment and Urbanization reported that 9.1 tons of organic wastes were composted until May 2019. The purpose of the zero waste application in Turkey involves the recycle and reuse of wastes, sustainable solid waste management, efficient use of natural resources, and reduction of landfill areas. The legislation prohibiting the chargeless use of light plastic bags was enacted in January 2019 and the public positively reacted to the process (Circular Economy Regional Report, 2020; Gökkurt Baki, 2021).

Solid waste management, on the other hand, is the identification, implementation, and development of an efficient and economical organization of services by aiming the storage, collection, transportation, treatment, and removal of all solid wastes and residues under hygienic conditions without harming the air and disturbing animals and plants, natural riches, and ecological balance. As the concept of "Solid Waste Management" has been effectively established and implemented after long efforts in developed countries, the control of basic health and environmental problems in terms of solid waste services has been resolved to an optimum level. Developing countries, on the other hand, suffer from the problems resulting from urbanization and lack the necessary financial resources to provide adequate solutions as a result of their high population density. The scarcity of available resources and the need to improve the quality of the urban environment demonstrate the importance of efficient solid waste management (Sorgun, 1988).

High-quality management is essential to achieve solid waste management and establish an easy-to-operate and efficient organization. The uncontrolled use of limited raw materials and the resulting issues highlight the necessity of the protection of raw material resources. The conservation of raw material resources primarily requires the reduction of waste generation and, then, the development of reuse methods by recovering energy or substances from wastes (ISTAC, 2016). In this manner, the amount of waste generation and the use of raw materials will follow suit and decrease at the same rate. The reduction of waste refers to the elimination of all environmental and public health issues caused by wastes. Sustainability in waste management entails the adoption of an environmentally optimum and efficient, socially acceptable, and economically effective and efficient management approach.

**Figure 2.** *Components of sustainable solid waste management*

As seen in Figure 2, we can begin to discuss sustainable waste management only after reaching the point where the effective management of the three factors intersects. A multidisciplinary effort is needed for sustainable waste management. Solid waste management will be successful in the long term only with the contribution of a multidisciplinary and multifaceted approach. Thus, sustainable waste management can only be efficient when an integrated viewpoint is adopted. Table 1 shows the shareholders and topics of consideration to achieve this goal.

Integrated waste management is defined as the selection and implementation of the appropriate methods, technologies, and management programs for a specific waste management purpose and target (ISTAC, 2016). Integrated solid waste management has also adopted the sustainable management viewpoint. In this manner, the reduction of the use of raw materials or any material, energy, space, and emissions is aimed. The four main strategies to achieve integrated solid waste management in order of priority are:

- Waste reduction (prevention and minimization)
- Material recycle
- Energy recovery
- Sanitary landfills

Ergun (2001) has delineated the hierarchy of integrated solid waste management by considering its impact level (Figure 3).
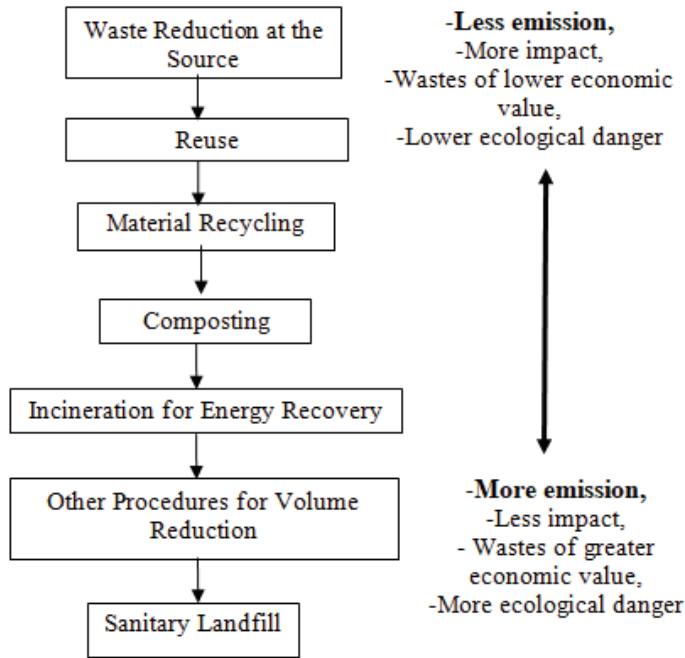


**Figure 3.** *Integrated waste management hierarchy (modified from Ergun, 2001; Stokoe, 1995)*

**Table 1.** *Integrated sustainable waste management framework (adapted from WorldBank, 2012)*

| Stakeholders | It involves the cooperation of local administrators, non-governmental organizations, and the public. |
|---|---|
| Process | It involves the technical aspects of solid waste management and affects all stakeholders and more. The technical, financial, political, and legal processes are effective and prominent in the process. |
| Policies and Impacts | In this section, economic factors, social factors, and political factors are of prominent importance. Sustainable waste management and public health should be planned with all processes in mind. At this stage, more emphasis should be placed on circular economy approaches to waste management. In addition, in an industrial sense, the Life Cycle Assessment (LCA) approach should be adopted at every stage of production. |

### The Role of Wastes in Greenhouse Gas Emissions in Turkey

The physical problem underlying climate change is very simple: dumping carbon dioxide ($CO_2$) and other greenhouses gases into the air raises their concentrations in the atmosphere and causes gradual warming. In the several decades since climate change has been an important international political issue. Carbon dioxide is a waste product; dumping it into the open air is a form of littering. Dumping can be avoided or cleaned up with technological fixes to our current infrastructure. These fixes do not require drastic reductions in energy use, changes in lifestyle, or transformations in energy technologies. Keeping carbon dioxide out of the atmosphere is a waste management problem (Lackner & Jospe, 2017).

According to the Intergovernmental Panel on Climate Change (IPCC) Guidelines (IPCC, 2006), the global warming potential (GWP) for 100 years' time horizon is about 25 times for $CH_4$ and 298 times for $N_2O$ compared with that of $CO_2$ on a weight basis. Methane is mainly produced in strictly anaerobic environments, through the microbial decomposition of easily degradable organic compounds, such as lipids, carbohydrates, organic acids, and proteins present in the organic waste (Husted, 1994; Khan et al., 1997). Nitrous oxide is usually associated with regions of the waste heap where an oxygen ($O_2$) gradient occurs (Beck-Friis et al., 2000) as a result of nitrification-denitrification processes (Pardo et al., 2015).

The carbon dioxide emission rate is higher than that of other gases contributing to the overall gas emission rates in Turkey. The exact solution to the release of carbon dioxide emissions origin of which has been scientifically shown to be wastes can only be achieved through waste reduction. Table 2 shows the statistics for the greenhouse gas emissions in 2019 and their distribution by sector.

**Table 2**. *Greenhouse gas emissions in Turkey in 2019 (TurkStat, 2021)*

| Total | Greenhouse gas emission values in Turkey in 2019 | | | |
|---|---|---|---|---|
| | | $CO_2$ (%) | $CH_4$ (%) | $N_2O$ (%) | F-gases (%) |
| | | 78.9 | 11.9 | 7.9 | 1.2 |
| | Greenhouse gas emissions by sectors in 2019 | | | |
| 506.1 Mt $CO_2$ equivalent | | Energy (%) | Agriculture (%) | Industrial processes and product use (%) | Waste (%) |
| | | 72 | 13.4 | 11.2 | 3.4 |

The data of the Turkish Statistical Institute (TurkStat, 2021) shows $CO_2$ as the cause of 425.3 million tons of the total emissions in Turkey, making up about 80% of the total emission (Table 2). Having become a party to the convention in 2004, Turkey is among the Annex-1 countries of

the United Nations Framework Convention on Climate Change (UNFCCC, 2018). According to the TurkStat data, total greenhouse gas emissions in 2017 were 526.3 million tons (Mt), with energy having the largest share with 72.2%, followed by industrial enterprises and product use with 12.6%, agricultural activities with 11.9%, and waste emissions with 3.3% (Figure 5). In addition, as in 2019, the percentage of carbon dioxide was above 80% compared to other greenhouse gas emissions in 2017, constituting about 425.3 million tons of total emissions.

The $CH_4$ emission data for 2017 show that 62.3% of the emissions were due to agricultural activities, 21.3% were due to wastes, 16.4% were due to energy, and 0.03% were due to industrial processes and product use. The $NO_2$ emission data reveal that 71% of the emission were caused by agricultural activities, 15.1% by waste, 10.7% by energy, and 3.3% by industrial processes and product use. A comparison of the waste emissions in 1990 to those in 2019 reveals that they increased by 55.7%, and decreased by 5% compared to the previous year, thus reaching 17.2 Mt $CO_2$ equivalent (TurkStat, 2021). The data reveal that especially the rates of $CO_2$ and $CH_4$ emissions, which are greenhouse gases that are released by waste storage, are considerably higher than those in 2017. Thus, the emissions due to waste should be reduced urgently.

Figure 4 shows the greenhouse gas emissions with respect to waste disposal methods in 2017 in Turkey in terms of $ktCO_2$/year equivalent.



|  | Total CO2 Equivalent |
|---|---|
| ⊠ solid waste removal | 9079 |
| ⊡ biological treatment of waste | 14 |
| ⊠ incineration | 4 |
| ⊠ wastewater treatment and discharge | 8258 |
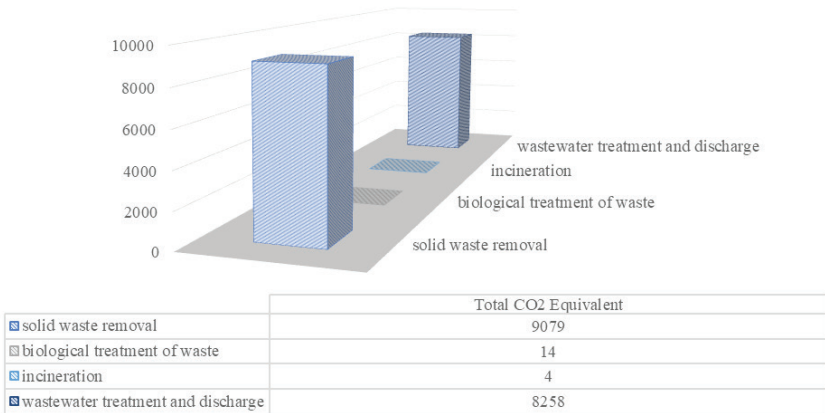
**Figure 4.** *Greenhouse gas emission values by waste disposal methods in Turkey in 2017 (ktCO2/year) (UNFCCC, 2018)*

Figure 5 shows the rate of change in greenhouse gas emissions in Turkey by sectors between 2017 and 2019. As seen in the figure, the

emissions increased by 3% in the waste sector between 2017 and 2019. Moreover, Figure 6 shows the change in the waste-induced emission values between 1990 and 2019 in Turkey.
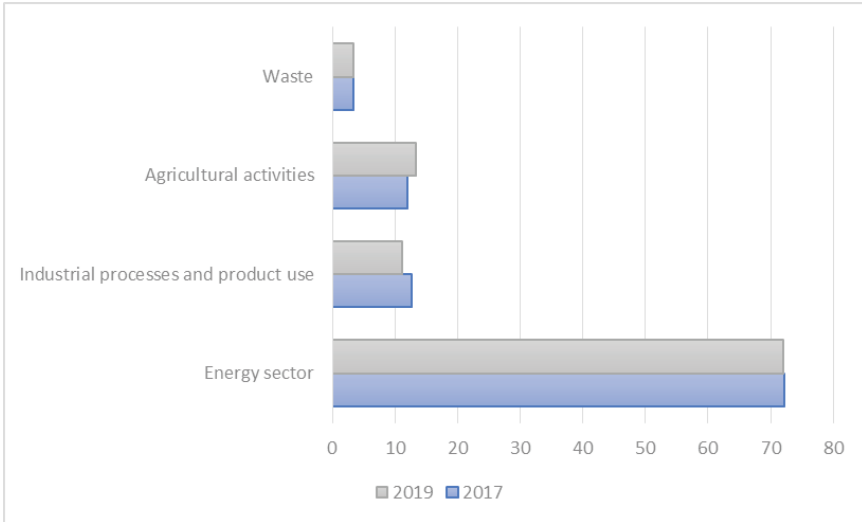


**Figure 5.** *Rate of change in greenhouse gas emissions by sector in Turkey between 2017 and 2019 (TurkStat, 2021)*
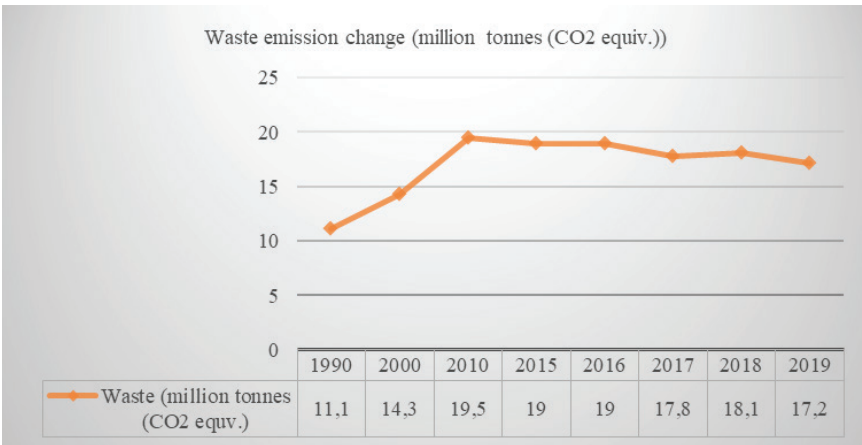


**Figure 6.** *Rate of change in solid waste-induced emissions in Turkey between 1990 and 2019 (TurkStat, 2021)*

As seen in Figure 6, solid waste-induced greenhouse gas emissions were reduced between 1990 and 2019, especially between 2018 and 2019, but even the decreased values have yet to reach the desired levels. The data of the Ministry of Environment and Urbanization show that the change in greenhouse gas emissions in 1995 was 12.95% when compared to the

previous year and was 137.5% in 2018. The share of the waste sector in the emission rate was 18.1 Mt $CO_2$ equivalent (EUM, 2021).

According to the IPCC national greenhouse gas inventory (IPCC, 2006); wastes are evaluated under five categories comprising solid waste disposal, biological treatment, incineration and open waste incineration, wastewater treatment and discharge, and others. Waste-induced emissions are methane and nitrous oxides that are produced by bacterial activities. Methane is the greenhouse gas that is released at the highest rate by sanitary landfills and the European Environmental Agency has pointed to landfills as the greatest cause of waste-induced emission with a rate of 95% (EEA, 2014). Landfills release the majority of the greenhouse gas emissions in waste management. In landfills, biodegradable wastes cause methane emissions, which lead both to climate change and fires and explosions. Therefore, the reduction of methane emissions from especially landfills is important in the minimization of greenhouse gas emissions (Ağaçayak, 2019). Nationwide waste reduction at the source and waste separation are needed to avoid emissions. Furthermore, organic wastes should be controlled at their sources and should be attached importance to just as much as packaging wastes. The studies on the reduction of organic waste storage in landfills should be expedited with this in mind.



**Figure 7.** *Uncontrolled dumping (original)*

The use of fossil fuels in the energy recovery storage facilities brings along different issues. A facility where domestic wastes are incinerated releases greenhouse such as $CO_2$ and $N_2O$, adding to climate change. An incineration facility where air pollutants are not kept can produce pollutants such as NOx, $SO_2$, HCl, particles, and dioxin. Considering waste management from the viewpoint of climate change, the primary thought is the control of the methane released by landfill sites. As mentioned earlier, the control of organic wastes before their transfer to landfill sites is of great importance. The reduction of greenhouse gas emissions is achieved in this manner (Ackerman, 2000). The control of greenhouses gases that are released by the ultimate disposal of solid waste compositions and especially of organic wastes is highly important. The recycling and economic recovery of packaging wastes such as plastics, metals, glass, and paper/cardboard are carried out in Turkey, but the applications for the reuse and recycling of organic wastes are not sufficient. According to the statistics of the waste disposal and recovery facilities in Turkey from 2016, waste recovery facilities constitute 92% of total facilities. However, composting facilities constitute only 1% of the waste recovery facilities. Moreover, 2% of the facilities are co-incineration facilities while 97% of the facilities are recovery facilities that are classified as other and used for the recovery of metals, plastics, paper, minerals, etc. Waste disposal facilities constitute 8% of total waste facilities and comprise 96% sanitary landfills and 4% incineration facilities (TurkStat, 2017). The most important step for the successful application of these methods is the separation of the wastes at their sources and application of the circular economy approach if waste reduction is unattainable and/or waste generation is unavoidable.

**Strategies for the Reduction of Waste-Induced Greenhouse Gas Emissions**

The most important method in the reduction of solid waste-induced emissions is to maximally and efficiently utilize the substances that we have been referring to as "wastes" both during production and after consumption. In other words, this method involves "avoidance of waste generation". To meet the unlimited needs of people using limited resources, developing new ways that are balanced and respectful of ecological balance for using renewable resources with the help of scientific and technological developments is needed.

The advantage of solid waste minimization in the prevention of climate change is undoubtedly the reduction of solid wastes that require storage and incineration. The waste-induced greenhouse gas emissions will decrease with the decreasing amount of waste. The minimization of solid wastes is carried out in three ways (Figure 8).
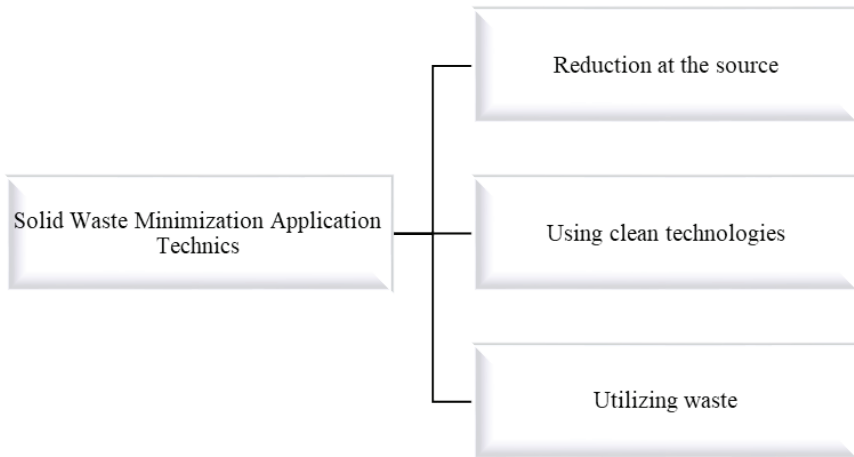
**Figure 8.** *The applications that should be included in solid waste management for the minimization of solid wastes and reduction of emissions*

### Reduction at the Source

Reduction at the source is the first stage and involves the separation of wastes not after their collection but at their source, in other words, moving ahead of waste collection. This also brings along economic gain since the highest-cost step of solid waste management in the local establishments in Turkey is waste collection.

Waste minimization through reduction at the source aims to reduce the production of waste materials to a minimum and collect wastes after their separation at their sources with the help of the public. This can become a highly efficient application when planned correctly and can be employed both by sectors and households. The application will also reduce waste release. Thus, emissions causing global warming can be reduced.

### Use of Clean Technologies

The complete reduction of waste generation is usually not possible. If this is the case, the most effective method in the prevention of pollution is the selection of clean and low-waste-producing technologies and the use of materials produced using these technologies. As is the case in reduction at the source, households, the public, and sectors should undertake certain responsibilities (Ergun, 2001). This will allow the reduction of waste generation and solid waste-induced emissions.

### Waste Recovery

Waste utilization, which constitutes an important part of waste minimization, is carried out in three ways: reuse, use for other purposes, and recycling. Reuse is the first of these methods and the reuse of solid wastes involves the repeated use of materials and products without altering them. The main factor in the reuse of wastes is the appropriate collection of reusable wastes. Thus, even if wastes are generated, no emissions will be released due to their use. Another waste utilization method is the use of wastes for other purposes. In this method, waste materials can be used as raw materials or additives in the production of other materials. The final method is recycling, which involves the collection and appropriate separation of recyclable materials and their conversion into byproducts and/or raw materials through recycling processes (Ergun, 2001).

### Conclusion

Climate change is indisputably the most discussed environmental issue that requires urgent attention. The issues encountered in waste management can be listed among the causes of climate change. The waste-induced greenhouse gas emissions in cities and their role in climate change have been neglected until recently but have long attracted the attention of academic circles.

Anthropogenic activities have changed with developments worldwide and the increasing solid waste generation with each passing year is a reflection of their impact. The supply of lands for the storage of solid wastes becomes harder every year. Thus, the minimization, recovery, and reuse of wastes as solid waste management stages become more important every year.

The open consumption system is based on the buy-produce-use-dispose process (Sapmaz Veral & Yiğitbaşıoğlu, 2018) and is referred to as "linear", "open-ended", or "traditional" (from, Sapmaz Veral 2021; Steffen, 2015). This system is the current economic system and represents the production of raw materials, their processing to obtain end-products, and their conversion into waste after consumption. In this model, production is carried out using resources and emissions occur due to wastes. The recovery of wastes begins after the consumption of products. Recyclable components in solid wastes should be collected at predetermined areas in a regular and cost-effective manner regardless of the purpose and method of their use. This requires detailed and excellent planning. Using wastes that are reintroduced into the economy by adopting the circular economy approach can help achieve the goal of much lower emission release.

The circular economy (CE) approach, which plays a key role in the sustainability of waste management, can be defined as an application where it is possible to reuse products and raw materials, waste is recycled

to the economy, and thus energy and all resources are used efficiently and cleaner production is made in all sectors. In Turkey, if the implementation of the circular economy becomes more efficient, the expenses for local management of waste handling, collection, and disposal will be prevented. Within the context of zero waste, it has been known that waste is very common and substantial in Turkey (Gökkurt Baki and Ergun, 2021).

As the understanding of the potential climate change impacts on waste management is at a very early stage; there is a lack of capacity within the sector in terms of knowledge, information exchange, learning, training, research and response. At present this covers waste management and climate change policy makers, regulators and site operators as well as academia, the professional institutions and other professionals associated with waste management. There are a range of actions that can be taken to address this capacity issue including research, information dissemination and training. The Waste Framework Directive of the European Union requests, however, the application of life cycle assessment (LCA) to identify cases in which it is reasonable to deviate from the classical waste hierarchy (avoid, reuse, recycle, recover, and landfill) (EC, 2008).

Until recently, a linear economy approach has been adopted in Turkey, but it is now being replaced by sustainable solid waste management approaches since its recognition as an unsustainable method. Negative factors such as the inability of the linear economy approach to achieve sustainable waste management, reduce waste generation, and reduce waste-induced emissions have steered the course toward circular economy (CE) approaches.

Life-Cycle Assessment (LCA) evaluates the environmental impact of a service, process, or product from "cradle to grave". The term "cradle to grave" refers to the period beginning from the production of raw materials and ending at waste disposal. LCA was developed to evaluate issues that were not included in other environmental management methods such as environmental impact assessment (Özeler et al., 2006). With regard to its definition, planning should include the beginning of industrial production and the final state of the product, especially for solid wastes. LCA is one of the most appropriate methods for sustainable waste management. If production is made by adopting the LCA approach, its complementation with a circular economy approach during waste management will pave the way for a "cradle to cradle" approach. Thus, the LCA approach is the most suitable application tool of a circular economy. The method and the approach can prevent solid waste generation and protect from the environmental damage caused by emissions.

A healthy and sustainable waste management system requires the

separate collection of recyclable wastes at their sources without mixing them with garbage in addition to carrying out the recycling process within an organized structure. Recycling allows the protection of natural resources and the prevention of the waste of resources. Moreover, the amount of wastes sent to storage areas will be reduced and reusable wastes will be reintroduced into the economy as raw materials. Considering its impact on waste management and climate change, raising the awareness of the public of wastes and separate waste collection and prioritizing capacity enhancement studies are very important for the reduction of wastes at their sources.

The nationwide inclusion of organic wastes to the zero waste applications in Turkey in addition to packaging wastes is surely needed in waste management. A complete zero waste and circular economy approach will remain unattainable where only packaging wastes are collected. The applications will be lacking when only packaging wastes are reintroduced into the economy, which, in turn, will hinder our expectations of reducing solid waste-induced greenhouse gas emissions.

*Recommendations for the reduction of solid waste-induced emissions:*

- Although appearing as local issues, solid waste management and resulting greenhouse gas emissions are global problems, as are all climate change factors. Thus, their handling should be made globally after their local planning and taking the necessary measures.

- The relevant institutions should provide regular data entries to allow carrying out accurate and complete scientific studies, which is an important drawback in Turkey.

- Capacity enhancement should be achieved by the contribution of the public and raising awareness of the impact of solid waste management and greenhouse gas emissions.

- Policymakers and regulators should continue to develop a sustainable solid waste management system.

- A management policy that recognizes current setbacks in solid waste management and involves common policies and applications that can affect climate change is needed.

- Solid waste management can only be sustainable through the use of LCA on an industrial basis, followed by the imminent adoption of circular economy approaches and supported by laws and regulations.

- The "zero waste" applications that are in effect in Turkey and

support the circular economy approach should be adopted around the country and should not be limited to packaging wastes by attaching importance to the inclusion of organic wastes into the cycle.

• Studies on the separate collection of organic wastes at their sources, compost production from this waste composition, and biogas production should be carried out. Complete success in sustainable waste management and the reduction of greenhouse gas emissions are not possible without the planning of policies for the recycling and reuse of organic wastes.

• Consumers should consider the harm caused by the products they purchase as well as the quality of the products and avoid the consumption of dangerous and non-degradable materials. Industries, on the other hand, should use harmless chemicals instead of harmful and toxic chemicals, prefer degradable packaging materials, and prevent the generation of dangerous wastes by changing the technologies and processes they employ. Nationwide capacity enhancement is needed for this purpose.

We can reduce wastes and waste-induced emissions in Turkey by following the circular economy goals. The preparation and thorough application of climate change action plans that clearly delineate waste reduction and adaptation strategies are the most important measures that should be taken within the scope of sustainable waste management in cities to combat climate change. The action plans should most certainly include sustainable waste management (waste minimization-reduction-prevention) policies.

In the study, solid waste management and solid waste-induced greenhouse gas emissions in Turkey were evaluated and the comparison and assessment of especially waste-induced greenhouse gas emissions by years were carried out. The impact and solid waste-induced amounts of greenhouse gas emissions, which can add to global warming, an issue of the utmost importance, were examined. The study emphasizes the importance of the reduction of wastes at their sources, the importance of the circular economy approach in sustainable waste management, and the importance of life-cycle assessment as the most effective tool of the circular economy approach.

## References

Ackerman, F. (2000). *Waste Management and Climate Change*. Local Environment, 5(2): 223- 229.

Ağaçayak, T. (2019). *Climate Change Comprehensive Local Studies in Waste, Wastewater and Air Quality Management in Turkey*, Supporting Joint Efforts in the Field of Climate Change (iklimİN), Ankara

Beck-Friis, B., Pell, M., Sonesson, U. & Kirchmann, H. (2000). *Formation and emission of $N_2O$ and $CH_4$ from compost heaps of organic household waste*. Environmental Monitoring and Assessment, 62, 317–331.

Christian, E.I., **(**2010). *Potential Impacts of Climate change on Solid Waste Management in Nigeria*, Journal of Sustainable Development in Africa (Volume 12, No.8, 2010), *ISSN: 1520-5509*

Circular Economy Regional Report (2020). *Regional Circular Economy Country Specific Report*. Written Contribution for the 2019 Annual Report Issue 1 February 2020 Editor-in-Chief Editorial Manager Co-Authors Rodion Gjoka, Adriatik Olloni Esmerina Hidri, 121 p.

Curi, K. (1998). *Solid Wastes Environment and Human*. Ankara Prime Ministry General Directorate of Environment. Year: 3, Issue:9, 1988.

EC, (2008). European Commission, *Waste framework directive 2008/98/EC of the European parliament and of the council*. European Parliament and Council, Directive European Commission. Brussels: European Commission.

EEA, (2014). *Greenhouse Gas Data Viewer*, 2014. https://www.eea.europa.eu/data-andmaps/data/data-viewers/greenhouse-gasesviewer adresinden erişildi

EUM, (2015). Environment and Urban Ministry, Waste Management Regulation, (date of access: 15.05.2021) https://www.resmigazete.gov.tr/eskiler/2015/04/20150402-2.htm

Ergun, O.N. (2001). *Solid waste lecture notes* (unpublished), 2001. 139 s.

Gökkurt-Baki, O. & Ergun, O.N. (2021). *Municipal Solid Waste Management: Circular Economy Evaluation in Turkey,* Environmental Management and Sustainable Development, ISSN 2164-7682 2021, Vol. 10, No. 2.

Gökkurt Baki, O. (2021). *Integrated Solid Waste Management in Turkey Based on Circular Economy*, 11th International Conference of Ecosystems June 4-6, 2021, Tirana, Albania,

Henry, J.G. & Heinke, G.W. (1989). *Solid wastes*. Environmental Science and Engineering. USA. 1989.

Husted, S. (1994). *Seasonal variation in methane emission from stored slurry and solid manures*. Journal of Environmental Quality, 23, 585–592.

IPCC, (2006). *Guidelines for National Greenhouse Gas Inventories*, (eds Eggleston HS, Buendia L, Miwa K, Ngara T, Tanabe K). Prepared by the National Greenhouse Gas Inventories Programme. IGES, Hayama, Japan.

ISTAC, (2016). *Solid Waste Management and Technologies*, Technical Books Series-7, Editors: Prof. Dr. Ing Oktay Tabasaran, Prof. Dr Ing Martin Kranert, Prof. Dr. İzzet Öztürk, 756 p., 2016 Istanbul, Turkey

ISWA/UNEP, (2013). *Workshop on GHG and SLCP Emission Quantification Methodologies*, September 19-20, 2013 Paris Workshop Report, 21 p.

Khan, R., Muller C. & Sommer, S. (1997). *Micrometeorological mass balance technique for measuring $CH_4$ emission from stored cattle slurry*. Biology and Fertility of Soils, 24, 442–444.

Kropp, J. P. & Reckien, D. (2009). *Cities and climate change: Which options do we have for a safe and sustainable future?*, Fifth Urban Research Symposium 2009 http://siteresources.worldbank.org/INTURBANDEVELOPMENT/ Resources/3363871256566800920/6505269-1268260567624/Kropp.pdf

Lackner, K.S. & Jospe, C. (2017). *Climate Change is a Waste Management Problem*, p. 83

Monni, S., Pipatti, R. Lehtilä, A., Savolainen, I & Syri, S. (2006). *Global climate change mitigation scenarios for solid waste management*. JULKAISIJA-UTGIVAREhttp://www.vtt.fi/inf/pdf/publications/2006/P603.pdf. Accessed July 5 2013

Onyanta, A. (2016). *Cities, municipal solid waste management, and climate change: Perspectives from the South,* Geography Compass 10/12 (2016): 499–513, 10.1111/gec3.12299

Özeler, D., Yetis, U. & Demirer, G. N. (2006). *Life cycle assessment of municipal solid waste management methods: Ankara case study,* Environ. Int. 32, 405-411

Pardo, G., Moral, R., Aguilera E. & Del Prado, A. (2015). *Global Change Biology*, 21, 1313–1327, doi: 10.1111/gcb.12806, Management, Entec UK Limited ISBN: 1-844-32221-1

Peavy, H.S.& Rowe DR. Tchbanoglous G. Solid waste, *Environmental Engineering*, Mc Graw - Hill Book Company. 1985.

Sapmaz-Veral, E. (2021). *The Circular Economy: Barriers, Strategies and Business Models*, Ankara University Journal of Environmental Sciences 8(1), 7-18 (2021)

Sapmaz-Veral, E. & Yiğitbaşıoğlu, H. (2018). *Transition Trends from Waste Management to Resource Management Approach in European Union Waste Policy and Circular Economy Model*. Ankara University Journal of Environmental Sciences 6(1) (2018): 1-19.

Steffen, W., Richardson, K., Rockström, J., Cornell, S. E., Fetzer, I., Bennett, E. M., ... & Folke, C. (2015). *Planetary Boundaries: Guiding human development on a changing planet*. Science, 347(6223), 1259855UN.

Sharma, R., Sharma, M., Sharma, R. & Sharma, V. (2015). *The impact of incinerators on human health and the environment.* Reviews on Environmental Health 28(1), pp. 67–72.

Sorgun, K. (1988). *Regular Sanitary Landfill, Republic of Turkey*, Ministry of Industry and Trade, General Directorate of Training and Development Center, Kuşadası.

Tchobanoglous, G., Theisen, H. & Vigil, S. (1993). *Integrated Solid Waste Management,* McGraw-Hill Inc., New York USA. 1993.

TurkStat, (2017). *News Bulletins-Waste Disposal and Recovery Facilities Statistics*, date of Access: 21 May 2021, <TURKSTAT: http://tuik.gov.tr/...>

TurkStat, 2021. *Turkish Statistical Institute News Bulletin*, No: 37196, 3p.

UNFCCC, (2018). *UN Climate Change Annual Report*, 2018, United Nations Framework Convention on Climate Change, 62 p.

WorldBank, (2012). *What a Waste, A Global Review of Solid Waste Management,* Daniel Hoornweg and Perinaz Bhada-Tata Urban Development Series, March 2012, No. 15, 116 p.

# Chapter 2

## ENERGY EFFICIENCY INVESTIGATION OF AN OFFICE BUILDING

*Doğan Tayyip KARAKELLE*[1]
*Zuhal OKTAY*[2]
*Can COŞKUN*[3]

1  KARAKELLE, D.T., OKTAY, Z., "ENERGY AUDIT AND EFFICIENCY ANALYSIS OF A PUBLIC BUILDING", İzmir 2021, İzmir Demokrasi Üniversity, Graduate School of Applied and Natural Sciences, Departman of Mechanical Engineer, İzmir, Türkiye, dogantayyipkarakelle@hotmail.com
2  İzmir Demokrasi Üniversity, Graduate School of Applied and Natural Sciences, Departman of Mechanical Engineer, İzmir, Türkiye, zuhal.oktay@idu.edu.tr
3  İzmir Demokrasi Üniversity, Graduate School of Applied and Natural Sciences, Departman of Mechanical Engineer, İzmir, Türkiye, can.coskun@idu.edu.tr

# 1. INTRODUCTION

The most important task of the buildings is to provide a high level of user comfort and to offer a minimum level of energy expenditure. Like all over the world, a very important rate of the total energy is used to provide user comfort in the buildings in our country. The rates of using the energy for our country are approxiately 43% in industry, 37% in buildings, and 20% in transport. In the world, the energy used in the buildings has the share of 45-50% of total energy. The energy used in the buildings can increase up to 45-50% of the total energy in the world (Schneider Electric, 2013). This indicates that energy saving in the buildings is very important.

While ensuring energy efficiency, costs need to be kept under control. When efficient buildings are mentioned, especially in our country, automatic control of the mechanical and electrical systems of the building and energy management are understood. At the same time, building design and construction must be energy efficient. However, sub-systems such as architectural design, construction system, carrier system, electrical and mechanical systems are a whole for buildings. Each of these subsystems must match the concept of creating efficiency. Under these circumstances a building is considered an " efficiency building " (Bayraktar and Yılmaz, 2013).

Buildings and all subsystems must be energy efficient from the design stage. Renewable energy sources should be used as much as possible in the energy efficiency of buildings. In addition, passive systems play an important role in energy performance.

Systems used in energy efficiency in buildings are based on average meteorological data. For this reason, it is not possible to achieve the desired performance in buildings where these parameters are ignored during the design phase.

# 2. GENERAL INFORMATION

## 2.1. Energy Efficiency in Buildings

Energy efficiency in buildings ensures that the products we consume in applications for our needs and requirements in our living spaces are consumed at an optimum level. The vast majority of our lives, such as jobs, education, healthcare, services and accommodation, should become efficient in the buildings we consume. For this reason, efficiency studies are carried out in consumption applications. In order to increase

efficiency, applications can be modified by taking various innovations as an example. Therefore, the buildings in which we spend our lifetime should be designed in accordance with these efficiency studies and these modifications. These factors should be taken into consideration while preparing the design and architectural structure. The main purpose of all these studies is to prevent the waste of consumed products. The waste of these products with unproductive adaptations causes an increase in the price of consumption, disrupting the natural balance and environment, reducing the comfort of life and destroying the world of future generations. These efficiency studies are very important for us and our future.

## 2.2. General Properties of Effiency Buildings

Energy efficiency buildings are an integration of comfortable, reliable, environmentally friendly, energy saver systems and energy efficient passive systems. Their main features are:

## 2.3. Location of the Building

The place where the building located is important in terms of solar radiation that affects energy costs, climate elements like air temperature, air movement and relative humidity values, as well as determining the micro-climate conditions that plays a very significant role on buildings energy efficiency. As it is shown in Figure 1, items around the buildings are very important factors that affect the micro-climate surrounding the buildings.

The location of the building, the distance between the other buildings and blocks with the building are one of the important design parameters that affect the amount of solar radiation and air flow type and speed around the building. Hence, location of the building in a site is one of the most important design variables that affect the sun and direct solar radiation utilization rate, and therefore, the total solar energy gain from the sun. (Yılmaz, 2014).
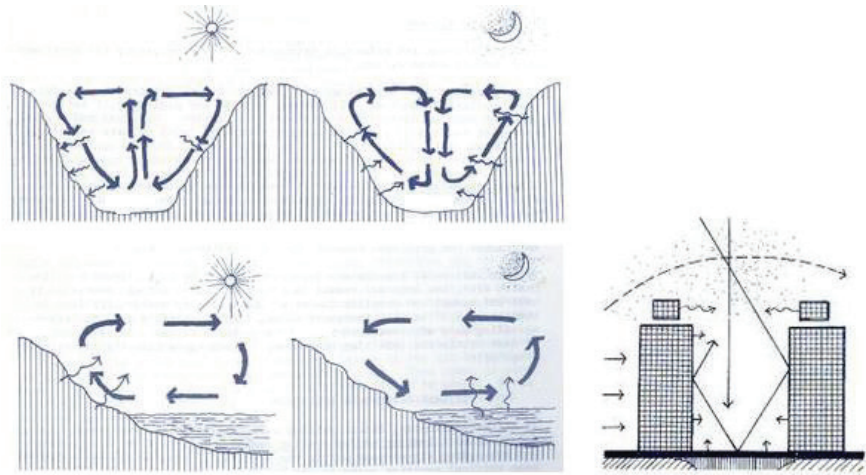
Figure 1. Location relative to other buildings (Yılmaz, 2014)

## 3. ANALYSES

### 3.1. Photovoltaic System Energy Calculation

There are many parameters that affect electricity or heat production from solar energy. The fact that many parameters are taken into account can cause the models used to be complex. In calculations, modeling of different system parameters is usually considered separately. The most important data used at the point of obtaining energy from solar energy systems are meteorological data.

The transposition factor is needed for calculating the radiation coming to the photovoltaic panel. The Transposition Factor (TF) is the ratio of the incident radiation on the plane ($I_{ir}$), to the horizontal irradiation ($I_{hi}$).

$$TF = \frac{I_{ir}}{I_{hi}} \tag{1}$$

The economic life of solar power plants is considered to be over 20 years. The impact of very small changes in the efficiency of a system that will produce energy for 20 years from year to year increases to significant levels. The losses cause the efficiency of these power plants to decrease. The energy produced by a grid-connected PV system depends on many factors such as the nominal characteristics of the components, system configuration, geographical location and structures around the installation point.
.

The reason for the development of such models is that the distribution of solar radiation coming to the world is not homogeneous, and it varies depending on the time and the angle of arrival. If a fixed infrastructure is to be established depending on the location, the system must be installed according to the optimum angle. Therefore, optimum angle determination methods have been developed. Optimum slope angle (β) can be calculated as follows.

$$\beta = \emptyset * 0.9 \tag{2}$$

Optimum slope angle for the winter season can be calculated as follows;

$$\beta = \emptyset + 25° \tag{3}$$

Optimum slope angle for the summer season can be calculated as follows;

$$\beta = \emptyset - 25° \tag{4}$$

The zenith angle is the angle formed between the normal to the horizontal surface and the sun's rays. This angle increases in winter and decreases in summer.
As the angle of inclination increases in solar power plants, the shadow length behind the panel also increases. As the angle of inclination increases in solar power plants, the shadow length behind the panel also increases. If the annual average global horizontal radiation ($I_{gh}$) value is multiplied by the displacement factor (TF), the radiation per unit surface area (I) is obtained.

$$I = I_{gh} \cdot TF \tag{5}$$

In this study, calculations were made on the conversion of global radiation into electrical energy in PV panels. The efficiency of the panel is taken into account in the calculations. The equations used for photovoltaic system modelling are given below.

$$E_{GSL} = E_{elc} + Q_{lost} \tag{6}$$

Here, $E_{GSR}$ is the global solar radiation amount. $E_{elc}$ indicates the converted electricity by the photovoltaic system. In the right-left part of the equation, the energy lost by the photovoltaic system is given as '$Q_{lost}$'. As can be seen in Eq. 6, there is the lost part of the total energy from the sun as well as the part converted into electricity. For the calculation of the lost energy, the following equation can be used.

$$Q_{lost} = U_{PV} \cdot A_{PV} \cdot (T_{cell} - T_{amb.}) \tag{7}$$

Here, $U_{PV}$ is the overall heat transfer coefficient for photovoltaic system. $A_{PV}$ indicates the photovoltaic panel area. $T_{cell}$ and $T_{amb.}$ are cell and ambient temperature. The efficiency of PV systems ($\eta_{PV}$) is found by the ratio of the produced electrical energy to the total global solar radiation.

$$\eta_{PV} = \frac{E_{elc}}{E_{GSR}} \tag{8}$$

In this case, it is possible to express the energy converted into electricity with the following formula.

$$E_{elc} = \eta_{PV} \cdot E_{GSR} \tag{9}$$

Temperature coefficient, cell temperature and reference temperature are the most important factors that directly affect the efficiency of the photovoltaic system. The efficiency of the PV panel can be calculated from Eq. 9 based on the temperature coefficient, cell temperature and reference temperature under standard test conditions.

$$\eta_{PV} = \eta_{PV.ref} \left[ 1 - \beta_{el} \cdot (T_{cell} - T_{ref}) \right] \tag{10}$$

The striking point here is the determination of the cell temperature. While the cell temperature calculation involves a complex calculation, these calculations have been simplified as a result of new studies. Taking wind speed (W), solar radiation (I), ambient temperature ($T_{amb.}$) and some constants into account, the cell temperature ($T_{cell}$)could be calculated as follows.

$$T_{cell} = 0.943 \cdot T_{amb.} + 0.0195 \cdot I - 1.528 \cdot W + 0.35 \tag{11}$$

## 4. Investigated Ofice Building

The Republic of Turkey, Ministry of Transport, maritime affairs and communications, III. Regional Directorate office building was selected for investigation. Information of the selected office building was given in Table 1. In this study, energy wasted resources, poor and incomplete insulation, water loss, fuel leaks and inefficient devices were examined from the perspective of energy efficiency.

In this study, it is aimed to determine the energy saving potential. This study also introduced preventive measures, technical and economic methods.

Table 1 General information about the office building

| Date of establishment | 2006 |
|---|---|
| Indoor area (m²) | 3,125 |
| Heating/cooling system | Fuel-Oil Chiller/Central Fan Coils |
| Isolation status | Not Isolation |
| Building floor / room number | 6 floor and 58 rooms |
| Annual electric consumption per m² (kWh/m² year) | 55.81 |
| Annual fuel-oil consumption per m² (kWh/m² year) | 110.57 |
| Annual energy cost per m² ($/m² year) | 15.1 (10.39 electrity based, 4.71 heating based) |

## 5. RESULTS

The amount of savings that can be achieved after replacing the existing lamps (18W Flourescent Lamps) in the office building with LED lamps (9W Led Lamps) was calculated. The characteristics of the replacement and new lamp are given in Table 2.
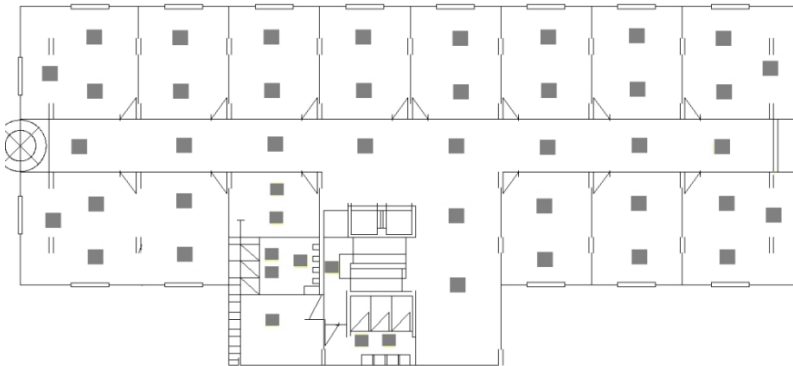


Figure 2. Office building lighting floor plan

Table 2 Features of replacement and new lamp

| Status | Type | Power | Lumen | Color | Live | Price |
|---|---|---|---|---|---|---|
| | | (W) | (Lm) | (K) | (H) | ($) |
| Replaced | 18W Flourescent Lamps | 18 | 1350 | 6500 | 15,000 | 2.10 |
| New | 9W LED | 9 | 1350 | 6500 | 30,000 | 17.00 |

Office building lighting floor plan was given in Figure 2. 1176 fluorescent lamps are used to illuminate the 3,125 m² closed area. There

is an installed lighting power of 6.775 W per m$^2$ for lighting. It was determined that annual 14.09 kWh electricty was utilized per m$^2$ for lighting. This amount equalts to 25.2% of total electricity consuption. It was determined that a 50% saving (22,015.6 kWh) could be achieved with the use of LED lamps.

Considering the suitable roof area of the building (510 m$^2$), it was determined that a PV installation with a total power of 83.325 kW (303 piece) could be provided. It has been calculated that an average of 114,155.25 kWh electricity can be produced on an annual basis. 36.53 kWh/m$^2$ year electricity can be reduced by utilizing the PV system. This amount equals 65.45% of total electricity utilization.

The data of the PV panel planned to be used and the characteristics of the installation are given in Table 3. Monthly basis electricity generation distribution changes between 6,604 kWh and 11,944 kWh for selected location. As expected, electricity generation from PV systems increases in the summer months, but decreases in the winter months.

Table 3. Characteristics of the PV System

| Location (Lat/Lon) | 38.441 / 27.147 |
|---|---|
| Database used: | PVGIS-CMSAF |
| PV technology: | Crystalline silicon |
| PV installed (kWp) | 83.325 |
| Slope angle (Â°): | 35 |
| Annual energy production (kWh/year): | 114,155 |
| Year to year variability (kWh) | 3,206 |
| Angle of incidence (%) | -2.65 |
| Spectral effects (%) | 0.5 |
| Temperature and low irradiance (%): | -12.33 |
| Total loss (%) | -26.24 |

## DISCUSSION

The results obtained as a result of the examination are presented below in articles.

- 7.045 kWh/m$^2$ year electricty can be reduced by lamp replacement. 36.53 kWh/m$^2$ year electricity can be reduced by utilizing the PV system. Total 43.575 kWh/m$^2$ year elecricty can be reduced. This amount equalts to 78.08% of total annual electric consumption per m$^2$ (kWh/m$^2$ year).

- Annual energy cost of the office building was found as 15.1 $/m$^2$ year. After the PV system instalization and lamp replecament, energy cost per m$^2$ could be decreased to 6.989 $/m$^2$ year (46.29% of total energy cost). Annual energy cost saving achives to 25,348 $/year (53.71% of total).
- It has been determined that if the heating with fuel oil is converted to natural gas, there can be a 47.88% reduction in heating costs (7,051 $/year or 2.256 $/m$^2$ year). At the end of the fuel change, together with the other two measures, the energy cost can decrease to 4.733 $/m$^2$ year.

# References

- Bayraktar, M., Yılmaz. Z., (2013). *Bina enerji tasarrufunda pasif akıllılığın önemi*, (retrieved May 16, 2013). http://www.binsimder.org.tr/index_files/TESKON2007__Bildiri_melte m_bayraktar_zerrin%20_yilmaz.pdf
- Schneider Electric, (2013). *Schneider Electric integrated solutions to improve energy efficiency, financial performance and sustainability*, *(retrieved March 4, 2013)*. http://www2.schneider-electric.com/sites/corporate/en/solutions/solutions
- *Yılmaz, Z., (2014). Akıllı binalar ve yenilenebilir enerji, II. Ulusal Tesisat Mühendisliği Kongresi, 387-398 (retrieved May 10, 2014). http://www.mmo.org.tr/resimler/dosya_ekler/101a002877d11b0_ek.pdf*

# Chapter 3

## THEORY, EVALUATION AND OPTIMIZATION OF ORGANIC COATING PREPARATION FOR CONSTRUCTION INDUSTRY

*Nil ACARALI [1]*
*Eda Nur SOYSAL[2]*

1  Yildiz Technical University, Department of Chemical Engineering, Davutpasa Campus, 34210, Esenler-Istanbul, Turkey. nilbaran@gmail.com (https://orcid.org/0000-0003-4618-1540)
2  Yildiz Technical University, Department of Chemical Engineering, Davutpasa Campus, 34210, Esenler-Istanbul, Turkey. eda.n.soysal@gmail.com (https://orcid.org/0000-0002-9667-9702)

## 1. INTRODUCTION

The coating industry was one of the oldest evolving industries with an adventure dating back almost 30,000 years to the present day from people living in caves. The first people, who rapidly evolved to the present day and excavated an existence from the past, wrote their heritage into history by using drawing and paint discovery, decoration, and coating protection methods as the most important communication tools. It was preferred to use the most common materials in nature such as carbon and metal oxides, blood, milk, and plant extrudates as pigments, so the industry gained speed.

The use of materials that provided the incredible durability Egyptian civilization people made paint from sour milk and lime as understood from drawings from 3000-600 BC, along with lime, which acts as a resin in the milk, increased resistance and reached extraordinary durability. From various historical remains, egg whites and yolks were also used as binders. The studies indicated that the Greek and Roman civilizations discovered that paint was not only a visual tool, but also the coating and surface protection feature between 600 BC and 400 AD and used functional paints for this purpose. With the discovery of properties such as surface coating and durability, apart from its aesthetic purpose, perhaps the most important area where the paint was used was realized by Noah, who used pitch to waterproof the ship in the marina area. Ancient Egyptians made great contributions to the development of the paint industry. It was advanced the development of dyes with the production of resins and pigments in dairy products. Throughout the developing industry, it was started to use vegetable oils and fossilized resins instead of raw materials obtained from milk. While the studies on paint were in various places of world, the extract obtained from the sumac tree, which was a common tree in Japan, began to be used in lacquer technology and a unique art was discovered. However, the use of lacquer from the coccid insect Laccifer lacca was proposed in a unique way. Finally, varnishes produced by dissolving in hot vegetable oil called amber were developed and gained an important place in paint development process. Regardless, until the Industrial Revolution, the paint industry, like other industries, did not fully develop due to the difficulty of production. With the advancement of technology because of globalization, companies in the production sector were forced to compete with great competition, it was aimed to increase the market share by putting forward more efficient production and lower cost works. Although this situation put companies more difficult, it played a major role in making progress. The first goal was to produce better quality products by consuming less energy and reducing the workforce. Paint production basically consisted of these stages: dispersion, coloring, filling and packaging. The most important part of these was the dispersion stage in which the colorants and fillers in powder

form were divided into smaller particles and distributed homogeneously in the solvent. Since researchers proved the better the distribution quality, which determined the aesthetic properties such as covering and gloss, as well as the protective properties of the organic coating, the better the protective and aesthetic properties of the organic coating [1].

## 2. THEORY

### 2.1. Definition of Organic Coating

Organic coatings were generally coatings consisting of organic compounds containing carbon and hydrogen, usually oxygen, nitrogen and other elements used for a barrier purpose. Here, the meaning derived from the barrier refers to the part between the substrate and the atmosphere, and its most important and main purpose was to provide protection against corrosion formation. The researchers showed that they were produced with other materials as part of the molecular structure or as additives, fillers, considering its various properties [2]. The stages of organic coating applications were showed in Figure 1.
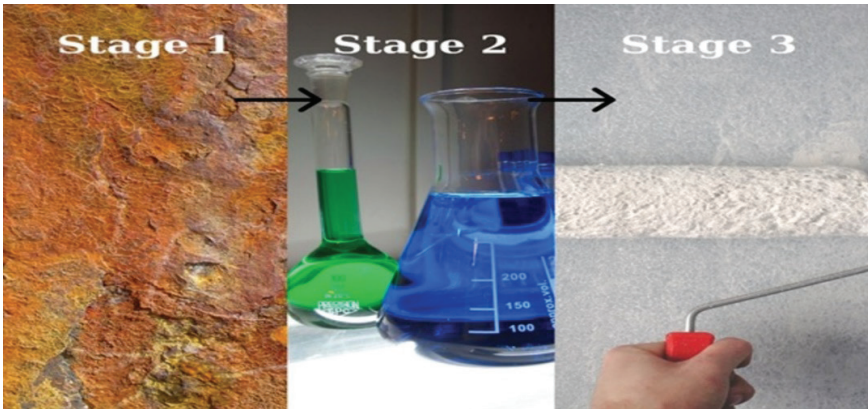


*Figure 1. Organic coating application on surface*

A study demonstrated the organic coating systems classifications as follows according to their usage areas [2]:

a) Decorative Paints

b) Construction Paints

c) Industrial and Corrosion Paints

d) Furniture Paints

e) Auto Paints

f) Marine Paints

g) Packaging Dyes

h) Spray Paints

The organic coating systems formed basically from four components which were as followings:

a.  Resin/binder

b.  Pigments and extenders

c.  Solvents

d.  Additives

## 2.2. Organic Coating Tests

Organic coating tests were required to ensure that the physical and chemical properties of the manufacturer could produce the same dyes and the consumer could choose the dye with the same physical and chemical properties. These factors must be kept constant to obtain reliable results. According to a study [2], the broad definitions of test conditions were made in ASTM, DIN, ISO, TSE. The various test visuals were seen in Figure 2.
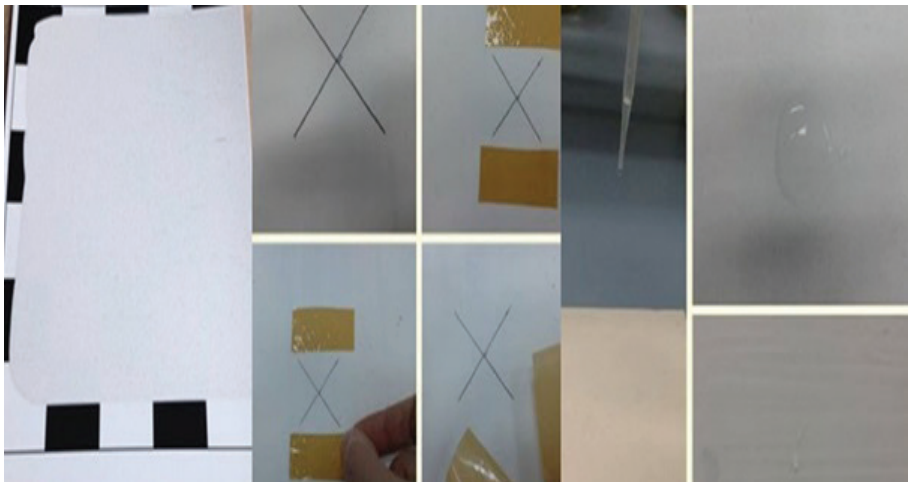


*Figure 2. Various test results*

Organic coating tests were examined in three groups. These were:

i.  Tests on Wet Organic Coating

a) Visual inspection

b) Viscosity

c) Hiding Power

d) Crush Fineness

ii. Tests on Dry Organic Coating

a) Gloss test

b) Impact resistance test

c) Adhesion test

d) Contact angle test

iii. Performance Tests of Organic Coating

a) Moisture chamber test

b) Saltwater mist test

c) Aging test

d) Resistance against chemical substances

## 2.3. Additives in Organic Coatings

In literature, additives were added to the paint in small amounts and provided strengthening of various properties. A research pointed out that for this purpose, while preventing foaming, peeling and pigment precipitation, the other great effect is against dispersion and surface wetting. Additives used as dryers were typically heavy metal derivatives used in alkyd paints to accelerate the oxidation of the paint film. Cobalt and zirconium were among the most widely used. Additives were raw materials that were only 1% in the paint formulation, but had important effects in the production, storage, and application of the paint. By using additives during the production and storage of the paint, the binders and colorants were effective on the surface tension and the rheology of the paint, preventing the paint from collapsing, flowing, and foaming. Dispersion quality was improved with surface wetting agents and dispersants. During the application, creation of a more uniform film is achieved with spreading enhancers, film formers and dryers. Better adhesion of the polymer in the paint to the surface with the adhesion enhancers. In addition, the study specified that there were additives with more special activity. All this significantly affected the performance criteria of the paint [2].

Some of the compounds used as additives were given below:

▪ Waxes

▪ Epoxy Accelerators

▪ Organic Pigments

▪ Defoamers

▪ Amines

▪ Rheology Agents

▪ Polyols

▪ Inorganic Pigments

▪ Cellulose Ether

▪ Water Repellent

Unlike the known additives, the importance of using organic additives, especially plant-based substances were increasing day by day (Figure 3).



*Figure 3. Organic additive examples*

A research synthesized two new acylamino compounds containing gramine groups, based on syntheses to improve antibacterial activity and antifouling activity against Escherichia coli and Staphylococcus aureus, by synthesizing the structures with instrumental analysis methods. As a result of some researches, it was shown that its anti-pollution properties were stronger and more effective than copper oxide and chlorothalonil, which were commonly used additives as pollution inhibitors. These compounds with high antibacterial activity were the best results obtained [3].

## 2.4. Literature Survey

The researchers [4] succeeded in developing lamellar micellar iron oxide paint coatings to extend the life and increase efficiency in structures containing metals that corrode easily under negative conditions. It was very important to be protected from corrosion on metal surfaces especially used in watery areas. The researchers aimed to improve the production by optimizing the dyes with this formula. The biggest deficiency that researchers saw in previous studies is that paint films saturated with moisture were not allowed to escape from the primers, resulting in local adhesion loss. Researchers claimed that improving the permeability property of the film produced at this point to increase the durability, namely the life span, and this could be achieved by examining the dispersion time of the micaceous iron oxides, and it was adopted this principle as the key point and adapted the formulation of the organic coating accordingly. In order to improve the property, it was observed that iron oxide coatings with anti-corrosion micaceous iron had a function suitable for the target.

A study examined the changes that occurred by adding metal ions to dye solutions, which had an important effect on dye agglomeration. Various experiments demonstrated calcium and magnesium ions, which had more interactions than sodium ions. Based on the experimental results, the intended effect of dye hydrophobicity on dye agglomeration did not occur, except in environments with high acidity. The results obtained that it would contribute to the development and understanding of the behavior of multi-purpose reagents used in the silk dyeing industry [5].

Some researchers aimed to coat any potential electrochemical corrosion of organic coatings with another electrochemical reaction that occurred on the film on the metal surface with an electron-conducting polymer in the mild steel layer and enhancing its performance. In the recent period when industrialization increased, generally observing the changes on the electro polymerization of a water-soluble monomer, oxidation method by passivation of the substrate was preferred. The reason why polyaniline material was obtained in nitric acid solution was that the films were large enough despite the fragile and powdery structure on the surface. In fact, those prepared from $Na_2SO_4$ solution were more successful in results with highly conductive films providing good corrosion resistance of the substrate. However, contrary to what was expected in the results, it was observed that adhesion to the substrate was not as sufficient as intended. It was deemed appropriate to carry out future studies and optimization studies on this subject [6].

A research evaluated a cleaner one-step method for robust nano coloration of polyester fibers via hydrophobic coating for the first time.

There were tremendous advantages of this method such as a possible replacement method of conventional textile coloration, eliminating hazardous dyes and auxiliaries, reducing textile effluents via using safe and magnetically recoverable colorant and photocatalytic dye degradation ability. An assorted color from light to dark brown with excellent color fastness ratings against washing, rubbing, acid and alkaline spotting and light was accomplished for fabricated composites. Iron oxide nanoparticles along with fatty acids as safe coating not only improve the safety of the process, additionally enhanced all the color fastness of the prepared samples along with forming very good hydrophobic behavior. The fabricated samples showed proper magnetic feature with saturation magnetization of 1.69 emu/g at room temperature expressed the feasible and simple recovery of the possible released nanoparticles in wastewater from washing. Hence, this method distributed the idea of utilizing magnetic nanoparticles along with fatty acids for cleaner and facile nano coloration of various fibers with multifunctional properties [7].

Another study investigated in detail how the para toluene sulphonic acid-PAni-pTS, which formed polyvinyl butyral coatings by adhering on iron surfaces by using dispersions that were considered to have the effect of preventing the delamination of the volume fraction, illustrating cathodic separation, to form the polyaniline emerald salt. To show the effect of dispersions, chloride electrolyte, which caused coating delamination when faced with the slightest deficiency, was preferred. The atomic force microscope and mass spectroscopy required here were primarily used to measure the time dependent thickness of the oxide layer that might form at the iron interface with the coating. Likewise, a scan was made to calculate how much the delamination rate would affect the surface, where it was found that the cathodic $O_2$ reduction seen in the secondary ion mass spectrometer had an inhibitory nature [8].

Some researchers purposed to fill these functional capsules with a healing agent, linseed oil and corrosion inhibitors. While producing these capsules, which were known as urea-formaldehyde capsules due to the material inside, a method suitable for nano size was chosen and capsules with a small diameter of 30–40 μm were selected and applied to use the mixing technique. This method was proven to be correct by observing that cracks in the paint film, linseed oil and nanoparticles were released from microcapsules broken down under induced mechanical action [9].

In literature, a study group worked on liquid-solid triboelectrification technology that preserves a new way of harvesting hydroelectric energy, the high cost of triboelectric nanogenerator electrode materials and their easily damaged microstructure severely limited the practical application. Improvement the enforcement of solid-liquid triboelectrification and

hydrophobic property was achieved via adding materials with fluorine to acrylic resin as being the friction layer material throughout the manufacturing a different kind of organic coating triboelectric nanogenerator. Thus, the preparation process was aimed to be comprehensive and simple, without microstructure that subsequently contributed to gain good stability and high output achievement as well as simple operation and low cost in coating with power generation function [10].

Some searchers investigated the impact of the quality of being rough on the moisturizing features of ion-coated poly (tetrafluoroethylene) coatings utilizing from analysis techniques. Implementing divergent substrate bias DC voltages to observe the difference of the quality in roughness between different coatings was the main purpose during ion platting part of this study. External part hardness was distinguished by studies performed using an atomic force microscope. For the coarse coatings prepared, high water interacted angles of 150-160° were determined for the coatings with many surfaces' roughness of 6-13 nm. The results provided additional evidence of the wetting property of hydrophobic coatings of the nanometer surface roughness [11].

A study developed coating samples of alloys of various metals that could be used as a barrier to prevent corrosion, suitable for use in the offshore industry, where anticorrosive studies were most focused, and evaluated the optimization of coatings that could be cost effective at the same time increase productivity. Unfortunately, metals were easily affected by the external atmosphere and were susceptible to corrosion and rapid loss of workplace. Here the researcher addressed this issue and put forward the idea of using organic coatings instead of traditional coatings, which were costly. It optimized the performance of organic coatings that could be used especially in sludge areas, areas that were most in contact with water, with the latest developments to provide active protection and coating systems to act as a barrier [12].

Some researchers suggested that impedance measurements aim to evaluate a system under test in terms of formation phenomena that showed insight into the behavior of the coating itself. Although the metals coated in some corrosive environments that were not suitable for the intended purpose were not troublesome, coating with $SrCrO_4$ pigmented styrene acrylic polymer has been presented. This polymer coating was kept in an aqueous NaCl solution for several weeks and as a result impedance measurement were made [13].

Investigators in literature considered dioctyl phosphate doped polyaniline as active pigment. Adhesion strength of the coating was determined to decline marginally after incorporating polyaniline [14].

A study carried out to produce anticorrosive organic coatings that more efficient structures with corrosion resistance could be obtained if used as an alloy of various metals as well as conventional coatings. For this purpose, iron phosphate coating method was preferred, and it was suggested that the data obtained results in quite high efficiency compared to traditional iron phosphating baths and that the changed parameters had enough flexibility. An increase from 10% to 75% was achieved in the corrosion resistance of alloy coatings produced with metal combinations [15].

Researchers made shape analysis and size measurement of microparticles with a high yield application using the method of changing the solvent called reprecipitation by using micro crystals. It was decided to use UV/vis absorption spectroscopy to be the most efficient option, and three fluorescent dyes were preferred, which varied according to the length of the alkyl chain of compounds 1, 2 and 3. At the same time, many values such as optical properties of platelets were analyzed by fluorescence microscopy and transmission electron microscopy [16].

Some explorers investigated some properties with seven coated and uncoated carbon and metal-based NPs, using the knowledge that nanoparticles frequently used in coated surfactants had a great effect on transport and toxicity. These properties were formed in terms of surface adsorption, affinity coefficient and contact. Several experimental setups were prepared for NPs that have the effects of the nuclei of specially designed nanoparticles, and their affinity for octanol and water phases were compared. To evaluate the surface hydrophobicity, especially for nanoparticles, the adsorption of naphthalene on NPs was used, as appropriate characterization methods could not be applied under these conditions. The results showed that Fullerol was the least hydrophobic of the seven NPs [17].

A research group investigated the changes in the activity of polyethylene terephthalate on the fabric by adding various additives to the properties through visible light photocatalytic. The changes were discussed how well they were suitable for self-cleaning and superhydrophobic applications, and a very clear photodegradability was observed as 147° of contact angle. Thanks to its superhydrophobic feature, a stainless fabric was obtained by using water-soluble impurities in these self-cleaning fabrics [18].

## 2.5. Optimization

In an engineering design approach, both analysis and optimization were preferred in first step before any prototype work began. In cases where there was more than one variable that change was desired to be seen before starting an experiment, it was very challenging to reach results by experimenting, considering the limiting nature factors. Maximizing or

minimizing some functions by some clusters often allowed a comparison of different options to determine which one might be "best" as it represented a set of options available in each situation. With all the research carried out within the framework of three effects on the sector, energy, materials and regulations, a study [2] presented that the solutions could be found for all supply and usage problems.

### 2.5.1. Taguchi Method

Industrial applications were particularly showed with the name of an engineer, G. Taguchi. Some researchers [19] studied on the fact for everyone that a lot of time and money was spent on engineering experiments and tests. The Taguchi method obtained design parameters to achieve average values for output parameters. It utilized special orthogonal arrays to examine design parameters with minimal experimentation, and in doing so it is relatively easy to reach the mean value of a target. For example, while L18 (36) to show eighteen studies and six factors with three levels each, the full factorial of the three-level six factor was 729 runs using the Taguchi method. In this way, the study [19] resulted in healthier and more outcomes could be achieved with less effort. An example of Taguchi graph was seen in Figure 4 with Signal/Noise ratio for water-based organic coating. The graph determined the optimum process conditions to prepare the coatings.
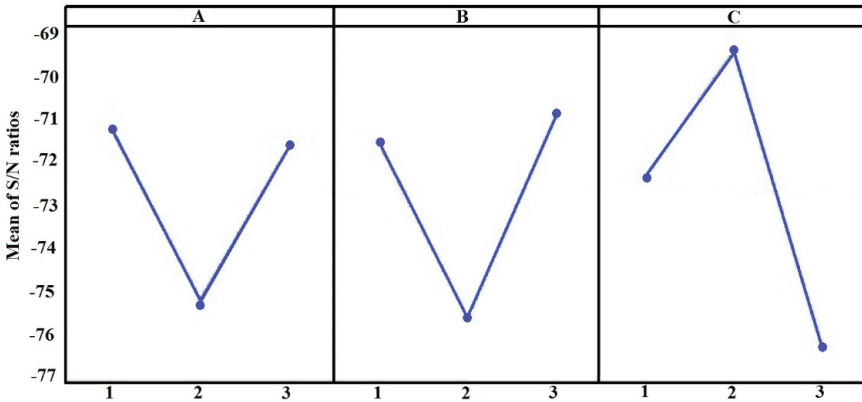


Figure 4. An example for Taguchi results according to the S/N ratio of water-based coating

### 2.5.2. Design Expert Method

Design Expert was a statistical software method to perform the design of experiments (DOE) widely used in optimization applications. Design Expert provided the necessary data and combined designs to minimize

the number of experiments required in real life in applications such as comparison of numerous tests, scanning, characterization, robust parameter design. The program obtained test matrices for screening up to 50 factors. The statistical significance of the parameters was determined by analysis of variance (ANOVA). A study [20] demonstrated that graphical tools were useful to determine the effect of each factor. The ANOVA used in Design Expert tells whether there was any statistical difference between the means of three or more independent groups. As in the same T test, ANOVA was used to find out whether the differences between data groups were statistically significant and work by analyzing the variance levels within the groups with samples taken from each. They [20] also indicated that in cases where there was a lot of variances, it was more likely that the mean of a sample selected from the data was different by chance. An example of Design Expert graph was seen in Figure 5 with parameter interaction for organic coating.
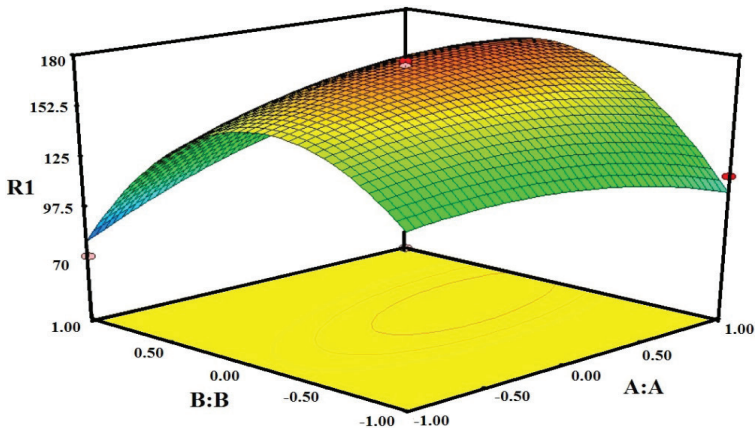


*Figure 5. An example for Design Expert results according to parameter interaction of coating*

## 3. EVALUATION RESULTS

Finally, by examining the optimization within the framework of the literature, some results and methods were oriented to the related subject. By adding fluorinated materials to the acrylic resin, the hydrophobia property was improved and improvement was achieved with lower costs and easier processes. Basically, the consequences of changing the pigment concentration in the coating composition were investigated, and it was concluded that the coating was controlled with additives as an anticorrosive barrier.

Moreover, the studies showed that the anticorrosive properties of the organic coatings could be supplied by adding the $SrCrO_4$ pigmented styrene acrylic polymer. The polymer needed a liquid solvent which was chosen as NaCl and after keeping the solution for many weeks, corrosion mechanism on the surface increased by the time.

There were some developments in the textile industry as the approach for the robust nano coloration of polyester fibers, here due to the hydrophobic coating material was experimentally applied to optimize the options which was flourishingly succeed at the end. As a final example in literature, various effects were tested by using mass spectroscopy and atomic force microscopy, and the use of organic coating in other sectors was investigated. Moreover, the new alternative ways were applicable and that corrosion-resistant organic coatings with high performance could be produced if the selection was made properly in water-based resins and ecological pigments.

## 4. CONCLUSION

In order to achieve better results in this performance evaluation study for long-lasting materials suitable for environments with negative environmental conditions in daily life, the next step should be the evaluation of the organic coatings tests applied in the sector, more resin types used in coatings, feature enhancing additives in coating. For the further investigations, more optimization applications might be given the priorities with the other optimization programming and in the textile and marine industry ought to be in the plans for improving the corrosion protection which it needs to be deal with broader subsumption.

## REFERENCES

[1] Miranda, T.J. (1993). The Future of The Coatings Industry. In: Surface Coatings. Springer, Dordrecht.

[2] Zheng, S.X. (2020). Principles of Organic Coatings and Finishing, Cambridge Scholars Publishing, first edition, UK.

[3] Feng, K., Li, X., Yu, L. (2018). Synthesis, Antibacterial Activity, and Application in the Antifouling Marine Coatings of Novel Acylamino Compounds Containing Gramine Groups, Progress in Organic Coatings, 118, 141–147. https://doi.org/10.1016/j.porgcoat.2017.10.027.

[4] Giudice, C.A., Benitez, J.C. (2000). Optimizing the Corrosion Protective Abilities of Lamellar Micaceous Iron Oxide Containing Primers, Anti-Corrosion Methods and Materials, 47 (4), 226-232. https://doi.org/10.1108/00035590010344321.

[5] Yeung, K.W., Shang, S.M. (1999). The Influence of Metal Ions on the Aggregation and Hydrophobicity of Dyes in Solution, Coloration Technology, 115, 228-232. https://doi.org/10.1111/j.1478-4408.1999.tb00361.x.

[6] Troch-Nagels, G., Winand, R., Weymeersch, R., Renard, L. (1992). Electron Conducting Organic Coating of Mild Steel by Electro Polymerization, Journal of Applied Electrochemistry, 22, 756-764. https://doi.org/10.1007/BF01027506.

[7] Rezaie, A.B., Majid Montazer, M. (2020). A Cleaner and One-Step Approach for Robust Coloration of Polyester Fibers via Hydrophobic Magnetically Recoverable Photocatalyst Fatty Acids/Nano Iron Oxide Coating, Journal of Cleaner Production, 244, 118673. https://doi.org/10.1016/j.jclepro.2019.118673.

[8] Holness, R.J., Williams, G., Worsley, D.A., McMurry, H.N. (2005). Polyaniline Inhibition of Corrosion-Driven Organic Coating Cathodic Delamination on Iron, Journal of the Electrochemical Society, 152 (2), B73-B81. https://doi.org/10.1149/1.1850857.

[9] Selvakumar, N., Jeyasubramanian, K., Sharmila, R. (2012). Smart Coating for Corrosion Protection by Adopting Nano Particles, Progress in Organic Coatings 74, 461– 469. https://doi.org/10.1016/j.porgcoat.2012.01.011.

[10] Wang, B., Wu, Y., Liu, Y., Zheng, Y., Liu, Y., Xu, C., Kong, X., Feng, Y., Zhang, X., Wang, D. (2020). New Hydrophobic Organic Coating Based Triboelectric Nanogenerator for Efficient and Stable Hydropower Harvesting, Applied Materials, and Interfaces, 31351–31359. https://doi.org/10.1021/acsami.0c03843.

[11] Veeramasuneni, S., Drelich, J., Miller, J.D., Yamauchi, G. (1997). Hydrophobicity of Ion-plated PTFE Coatings, Progress in Organic Coatings, 31, 265-270. https://doi.org/10.1016/S0300-9440(97)00085-4.

[12] Olajire, A.A. (2018). Recent Advances on Organic Coating System Technologies for Corrosion Protection of Offshore Metallic Structures, Journal of Molecular Liquids, 269, 572-606. https://doi.org/10.1016/j.molliq.2018.08.053.

[13] Hubrecht, J., Vereecken J. (1984). Corrosion Monitoring of Iron, Protected by an Organic Coating, with the Aid of Impedance Measurements, Electrochemical Science and Technology, 131, 2010-2015. https://doi.org/10.1.1.860.3295.

[14] Samui, A.B., Phadnis, S.M. (2005). Polyaniline–dioctyl Phosphate Salt for Corrosion Protection of Iron, Progress in Organic Coatings 54, 263–267. https://doi.org/10.1016/j.porgcoat.2005.07.002.

[15] Gorecki, G. (1995). Improved Iron Phosphate Corrosion Resistance by Modification with Metal Ions, Metal Finishing, 36-40. https://doi.org/10.5006/1.3315980.

[16] Bertorelle, F., Rodrigues F., Fery-Forgues S. (2006). Dendrimer-Tuned Formation of Fluorescent Organic Microcrystals, Influence of Dye Hydrophobicity and Dendrimer Charge, Langmuir, 8523-8531. https://doi.org/10.1021/la060239e.

[17] Xiao, Y., Wiesner, M.R. (2012). Characterization of Surface Hydrophobicity of Engineered Nanoparticles, Journal of Hazardous Materials 215–216, 146–151. https://doi.org/10.1016/j.jhazmat.2012.02.043.

[18] Min, K.S., Manivannan, R., Son, Y.A. (2019). Porphyrin Dye/TiO2 Imbedded PET to Improve Visible-Light Photocatalytic Activity and Organosilicon Attachment to Enrich Hydrophobicity to Attain an Efficient Self-Cleaning Material, Dyes and Pigments, 162, 8–17. https://doi.org/10.1016/j.dyepig.2018.10.014.

[19] Karazi, S., Moradi, M., Benyounis, K. (2019). Statistical and Numerical Approaches for Modeling and Optimizing Laser, Micromachining Process-Review, Module in Materials Science and Materials Engineering, https://doi.org/10.1016/B978-0-12-803581-8.11650-9.

[20] Dashti, A., Nasab, E.M. (2013). Optimization of the Performance of the Hydrodynamic Parameters on the Flotation Performance of Coarse Coal Particles Using Design Expert (DX8) Software, Fuel, 107, 593–600. https://doi.org/10.1016/j.fuel.2012.11.066.

# Chapter 4

## POWER-VOLTAGE (P-V) CURVES OF ENERGY TRANSMISSION LINES

*Yelda KARATEPE MUMCU[1]*
*Kamil Fırat İrfan GÜNEY[2]*

1   Dr. Öğr. Üyesi Yelda KARATEPE MUMCU, Marmara University, Teknik Bilimler Meslek Yüksekokulu, Elektrik ve Enerji Bölümü, Dr. Mehmet Genç Külliyesi Kampüsü, İstanbul / Türkiye, ykaratepe@marmara.edu.tr, ORCID ID : https://orcid.org/0000-0001-9948-9413

2 Prof. Dr. Kamil Fırat İrfan GÜNEY, Acıbadem Üniversity, Vocational School, Biomedical Equipment Technology, Istanbul / Turkey, irfan.guney@acibadem.edu.tr

## INTRODUCTION

Today, electrical energy consumption is increasing rapidly due to developments in technology and population growth. The large distances between energy production centers and consumption centers necessitate the use of very long high voltage lines in energy transmission. In addition to many factors taken into account in the planning, operation and control of power plants and lines of an electrical power system, another very ixmportant factor is to consider the stability of energy systems.
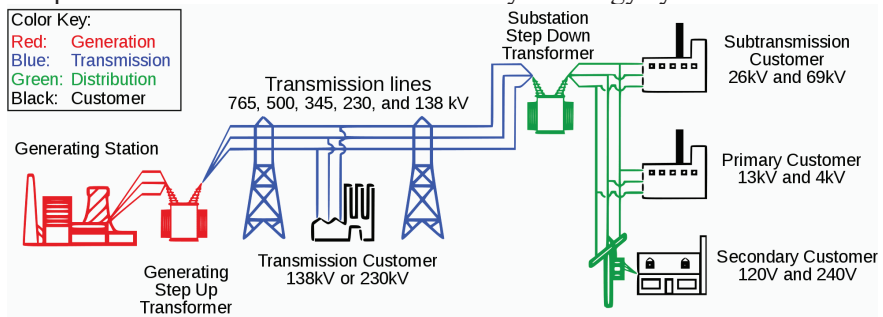


**Figure 1.** *Electric power system (URL 1)*

Stability in energy systems; It is defined as the ability of the system, which is exposed to a disturbance effect (such as short circuits, on-off events), to return to its operating conditions after the disturbance effect (Anderson and Fouad, 1986).

A complete analysis of an energy system after this disturbance is quite complex and if both the disturbance type and time are classified in terms of the stability of the system;

a) Transient Stability

b) Dynamic State Stability

c) Steady State Stability

can be expressed as.

The response of the system to major disturbance effects such as sudden load changes, short circuits in energy transmission lines, where there is a sudden loss of synchronism; The transient stability is the response of the system to the disturbance effect, after a few seconds of transient period, and the response of the system to the disturbance effect in a few minutes when the mechanical regulators are also active, the dynamic state stability and the response of the system to the expected load changes or small disturbance effects is called the steady-state stability(Taylor, 1992).

In order to improve the voltage stability in the system, some measures can be taken during the design and operation phase of the system. As an example to this; Appropriate placement and control of reactive power compensation devices, control of on-load tap changing transformers, performing load shedding in critical situations, and appropriate design of transmission lines can be given. In the classical stability studies, besides the importance of the active power-voltage (P-V) relationship, the reactive power-voltage amplitude (Q-V) relationship comes to the fore in terms of voltage stability (Overby and Dobson, 1994).

## POWER VOLTAGE (P-V) CURVES

With the help of Power-Voltage (P-V) curves, the voltage-power profile at the receiving end of the system can be monitored. Therefore, by looking at the curves; To examine the effects of step change under load and the dynamic behavior of partial loads, as well as various operating conditions such as keeping the source voltage constant at different values, making serial or shunt compensation at different percentages in the transmission line, consisting of a single or multiple line of the transmission line, having loads with different power coefficients at the end of the line. is possible.
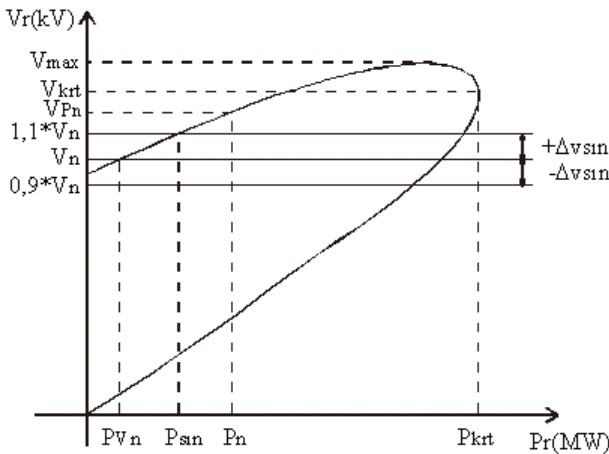


**Figure 2.** *Stable Zone Area*

In Figure 2, two different voltage values such as Vr1 and Vr2 correspond to each constant Pr power. Only at one point a given power (Prcrt) corresponds to a single voltage value (Vcrt). This point shows the maximum power limit that can be carried in terms of voltage stability and the corresponding critical voltage value. The upper part of the P-V curve reflects the normal stable operating region, while the lower region is the region corresponding to voltage instability. For a power demand higher than the critical power, the load voltage will be unstable (Mal, 1992) (Henriet , and Tarkan , N, 1991).

The positive effect in terms of voltage stability is the increase of the critical power value, the decrease of the critical voltage and the critical angle. By looking at the P-V (Power-Voltage) curves, stable operation is determined for both lines between different operating states (Tang, 2021).

## FACTORS AFFECTING P-V (POWER-VOLTAGE) CURVES

While obtaining the P-V (Power-Voltage) curves for a radial transmission line, it is possible to clearly see the effects of some quantities that remain closed in the analytical solution on these curves. These effects;

- Series compensation effects

- Shunt Compensation Effect

- Power factor influence

- Effect of Transmission Line Length

- Line loss factor effect

- Head-of-line voltage effect

- Parallel Power Transmission Line Effect, can be grouped as.

**Series compensation effects**

It is a known feature that the voltage at the ends of the capacitor connected to an alternating current circuit is 90º behind the current passing through it, and the voltage at the ends of the inductive reactance is 90º ahead of the current passing through it. Accordingly, the effect of inductive reactance on the line voltage in a transmission line can be partially or completely compensated by a capacitor to be connected in series. By making use of this feature in transmission systems, the compensation of the serial reactances of the energy transmission lines is realized (Miller, 1982).

With this kind of compensation; The maximum power that can be carried can be increased or the transmission angle ($\delta$) can be reduced for a given power transfer and the natural load of the line increases. Besides, as the line reactance is effectively reduced, the line will now draw less capacitive reactive power at the same system load, so the degree of shunt inductive compensation needed in idling will also be partially reduced (Indulkar and friends, 1987).

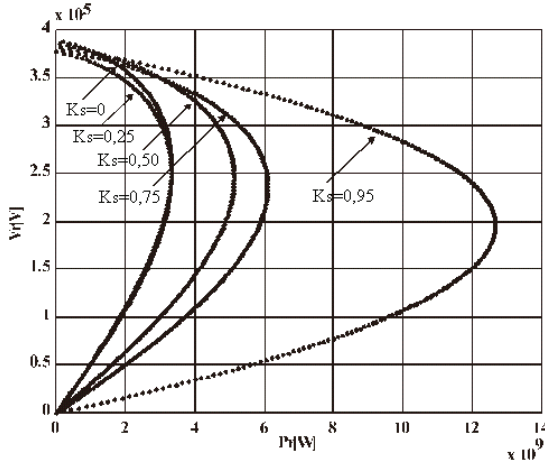$$Ks = \frac{X_c}{\varpi \cdot L}$$

$$(1)$$

**Figure 3.** *P-V curves obtained at different serial compensation percentage values of Atatürk-Göksun-Yeşilhisar transmission line (Mumcu, 2001)*

The series compensation ratio is given in the formula for Ks. Here, the capacitor's reactance to be placed in series is expressed as Xc, and the series inductive reactance of the line is expressed as ω.L.

As seen in Figure 3, with the increase in the series compensation ratio on the line, the critical voltage value and critical angle decreased, while the critical power value increased in all transmission models.

### Shunt compensation effects

Very high transport voltages are chosen so that large amounts of electrical energy can be transported economically over long distances. Thus, line losses are reduced by reducing the carrying current. However, since the capacitive currents that will flow due to the ground capacities of the transmission lines are proportional to the reactive power, the carrying voltage and the line length, it will cause a voltage rise due to the series line impedance, which has an inductive character.

If the end of the transmission line is connected to a characteristic equivalent impedance, the reactive losses in the line inductance are fully compensated by the reactive power produced by the line capacity, in this case the line does not exchange reactive power.

$$Kd = \frac{B_r}{\omega \cdot C} \tag{2}$$

Shunt compensation ratio Kd is the ratio of the reactor susceptance (Br) to be placed shunt on the line to the shunt capacitance (ω.C) of the
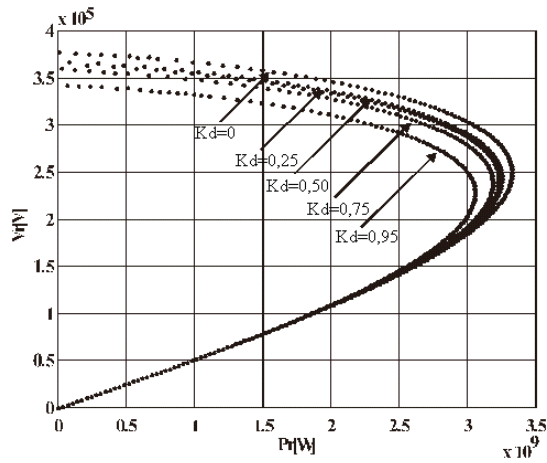
line.



**Figure 4.** *P-V curves of Atatürk-Göksun-Yeşilhisar energy transmission line obtained at different shunt compensation percentage values (Mumcu, 2001)*

As can be seen from Figure 4, clearly with the help of P-V curves, the voltage stability limits get smaller with shunt compensation. The critical voltage value decreased with the increase in the shunt compensation ratio on the line.

### Power factor influence

While the load is drawn from the end of the line with a certain power factor, this power factor is kept constant and the P-V curve is obtained by continuously increasing the value of the drawn power from zero to the critical value. For different power factors, by taking cosφ=1 criteria, three curves are obtained depending on whether the power factor is forward or backward.
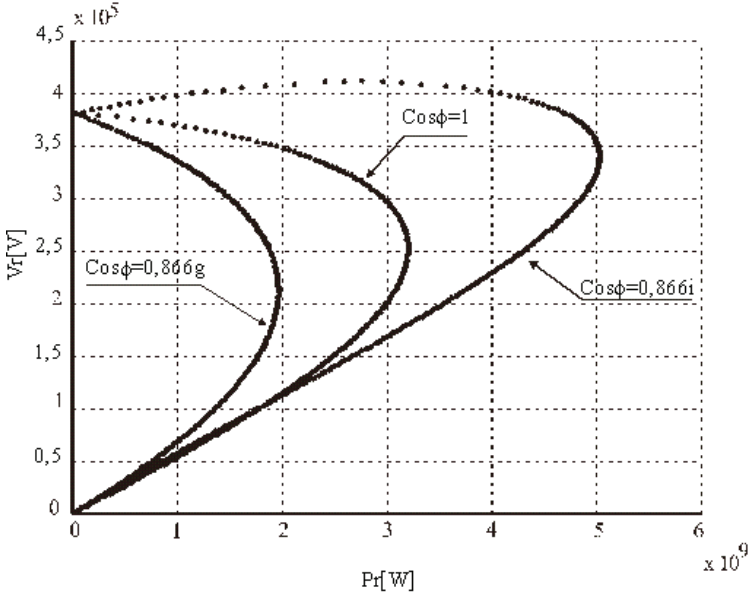
**Figure 5.** *P-V curves for different power coefficients of Atatürk-Göksun-Yeşilhisar power transmission line (Mumcu, 2001)*

It is observed that the critical power limit in terms of voltage stability increases as the power factor changes from inductive (backward) to capacitive (forward). This clearly shows that the power factor must be close to the unit value (cosφ=1) in energy transmission.

### Effect of Transmission Line Length

A, B, C and D circuit constants of the power transmission lines are calculated with the help of the ohmic resistance of the line, r (ohm/km), the reactance of the line, x (ohm/km) and the total capacity of the line, y (1/ohm.km). Since these quantities change depending on the line length, the circuit constants of the line also take different values at different lengths for the same r, x, y unit values. Since line constants are included in the analytical expression used to determine the P-V curves, the curves will differ for different line lengths.
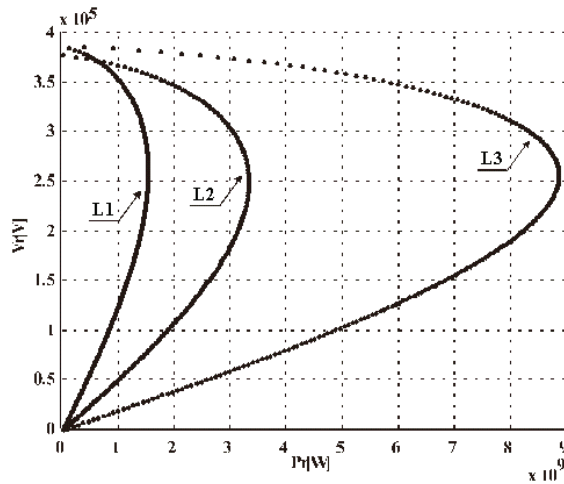
**Figure 6.** *P-V curves by transmission line length (Mumcu, 2001)*

Provided that the r, x, y values of the power transmission line remain the same, the P-V curves of the lines are obtained for three different lengths as L1> L2> L3. It is taken as cosφ=1(Power factor).

**Line loss factor effect**

The loss factor is the ratio between the ohmic resistance r, which arises from the design of the line, and ωL, the inductive reactance, and

$$\alpha = \frac{r}{\omega \cdot L}$$

(3)

is expressed as α.

The loss factor ultimately affects the line selection for transmission and the electrical magnitudes in various aspects. It is seen that the voltage stability is adversely affected by the increase of the loss factor.
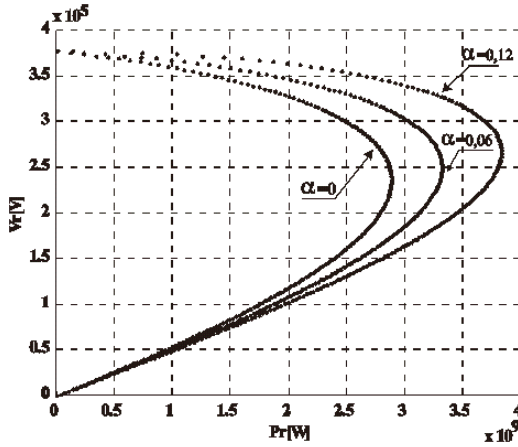
**Figure 7.** *P-V curves for different line loss factors of Atatürk-Göksun-Yeşilhisar power transmission line (cosφ=1) (Mumcu, 2001)*

### Head-of-line voltage effect

By changing the amplitude ($|Vs|$) of the head-of-line voltage, the change in the P-V curves can be monitored. It is seen that the critical values increase with increasing the amplitude of the line head voltage.



**Figure 8**. *P-V curves (cosφ=1) for different line head voltages of Atatürk-Göksun-Yeşilhisar power transmission line. (Mumcu, 2001)*

### Parallel Power Transmission Line Effect

In large amounts of power transmission, it is necessary to establish two or more parallel lines between generation and consumption centers due to various factors such as stability, reliability, operating requirements.

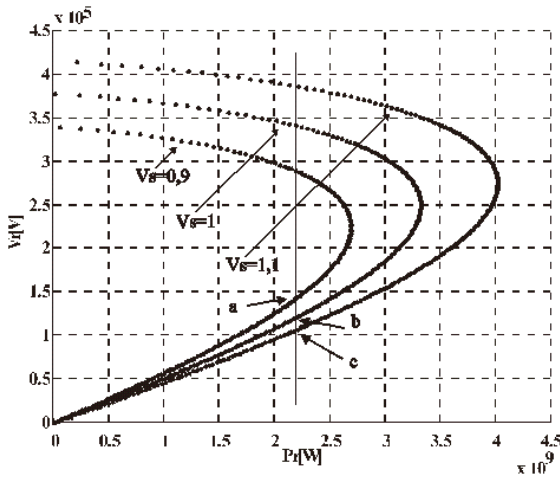In terms of voltage stability, the effect of parallel lines on voltage stability can be clearly seen with the help of P-V curves. Therefore, failure of one of the lines running in parallel may present critical situations in terms of voltage stability.
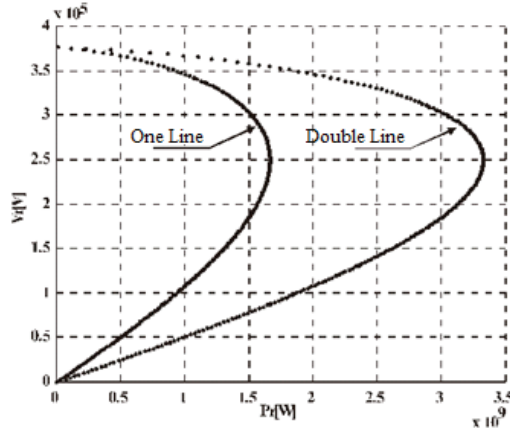


**Figure 9.** *P-V curves of Atatürk-Göksun-Yeşilhisar energy transmission line according to the number of lines (cosφ=1) (Mumcu, 2001)*

When one of the lines is disabled, the total impedance of the system doubles, normally reducing the capacity by about half which improves the power factor. Thus, as a result of increased line losses ($R.I_2$ , $X.I_2$) as well as the decrease in reactive power produced by shunt capacitances, the total voltage drop increases significantly and the possibility of voltage instability arises.

• As the phase angle changes from inductive value to capacitive value, Pkrt, which is the maximum power point on the curve, increases. However, for phase angle φ>85, the critical power value Pkrt increases rapidly towards zero. This is because the power factor cosφ=0 at phase angle φ=90⁰.

• The maximum limit power Psin is obtained around the power factor value cosφ=0.8, corresponding to the capacitive φ phase angle value for each operating state. Capacitive φ phase angle corresponding to power factors ranging from 0.7<cosφ<0.9, two operating regions are obtained for each operating condition. Therefore, care must be taken to allow the system to operate between these two zones.

• For each result obtained as a result of the operating states, it is very important for the system to operate so that the power factor at the receiving end is very close to 1, so that the system operates in the voltage stability region. Although the maximum Psin values are obtained

when the values of the capacitive φ phase angle versus the values of the obtained power factors are between 0.6<cosφ<0.8, care should be taken when determining the maximum power to be transmitted since there are two operating regions in these power factors. In this case, the voltage limit is exceeded. The most suitable rated power value at the receiving end is those with cosφ=1. Because the receiver terminal around this point is within the voltage limits of the voltage values. After the phase angle φ=0, the maximum voltage Vmax value starts to increase regularly.

• The possibility of operating in the region between 0.9*Vn-1.1*Vn, which is the appropriate region in terms of voltage stability, is available for values corresponding to power factors in the range of -90<φ<36. The voltage values obtained for the capacitive phase angle values of φ>36 are outside the operating range.

• Since the operating state IV is the smallest in terms of rated power value compared to other operating conditions, it is the appropriate operating state for situations where the receiver voltage value is required to be at the rated value. The most suitable state to operate at the voltage rated value is the operating state I. Because at capacitive φ phase angles corresponding to 0.5<cosφ<0.9 values, the obtained power values are higher than other operating conditions. The operating state I has the region with the least risk in terms of voltage stability.

• For phase angle φ=300, power value of operating state I with power factor cosφ=0.8090 has maximum limit power Psin2. Thus, the allowable operating zone width to be obtained at constant power factor is larger than in other operating conditions. If the P-V (Power-Voltage) curves obtained in the I, II and III operating states are examined, the values of the limit power Ps to be obtained for the constant transmitter terminal voltage and the same power factor are higher than the values in the operating states IV, V and VI. Therefore, in cases where high power is required at the receiving end, I, II and III. working conditions should be preferred.

• It has been seen from the studies carried out that positive effects in working conditions I., II. and III. It shows very close characteristic values for operating conditions. In terms of dynamic voltage stability, the first three operating states are good, V. and VI. situations are critical, IV. The working condition is seen to be the most critical line model. It is the fourth working position with the narrowest area in terms of working area.

• It has been observed that shunt compensation with two reactors at the beginning and end of the line is superior to the compensation made with a single reactor. Although the choice of reactor depends on other operating conditions, it should also be examined in terms of dynamic

voltage stability, and even the effects of more than two reactors in long lines should be examined.

• Although it has a superiority over other operating conditions in terms of static for the V. operating situation, where the series compensation is in the middle of the line and the shunt compensation is done at the end of the line with a single reactor, it cannot show the same good situation in terms of dynamic effects. Accordingly, it is not sufficient to determine only the static limits of a line operating system in terms of voltage stability. In particular, if compensation is to be made with a single shunt reactor, this examination should be done much more carefully.

• Reactive power generation being as close to consumers as possible, reduces the effects of problems such as parallel line trips, in addition to its well-known feature of reducing losses in generators, lines and transformers(Mumcu, 2001).

*Table 1. Working Situations (Mumcu, 2001)*
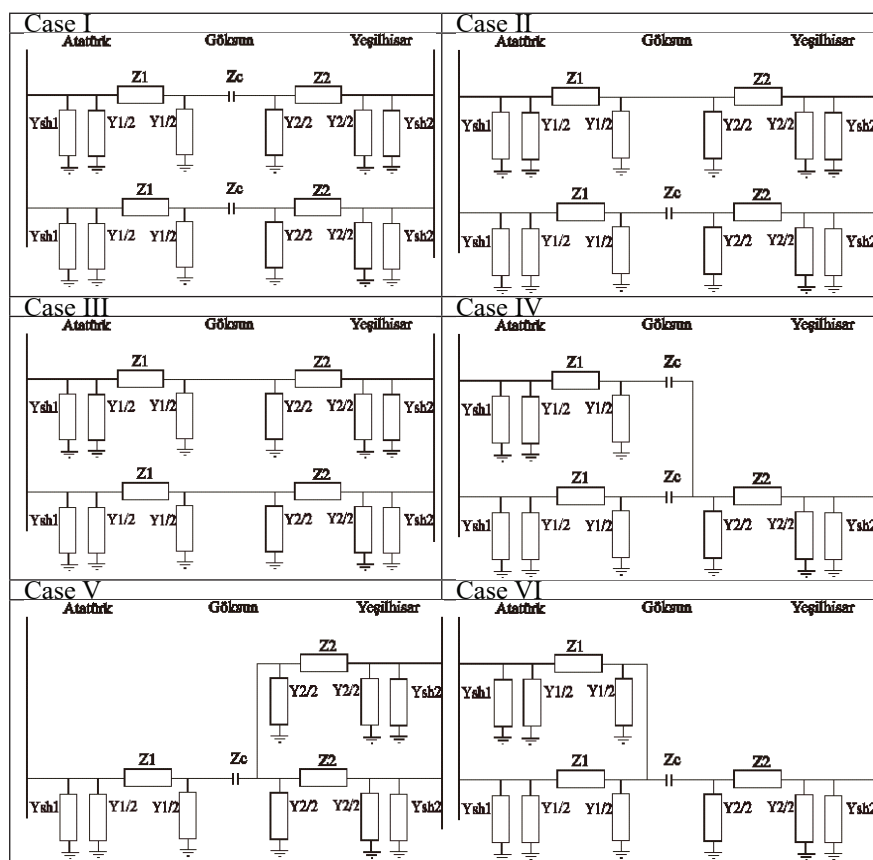
| Not: |
| --- |
| *This book chapter was produced from **Yelda KARATEPE MUMCU**'s **Master's thesis** titled "**Examination of Power-Voltage (P-V) Curves of Energy Transmission Lines**" prepared in 2001 at Marmara University, Institute of Science and Technology Electricity Program.* |

# REFERENCE

Anderson P.M., Fouad A.A., "Power Systems Control and Stability", The Iowa State University Press, 4th printing, 1986.

Henriet , P . , Tarkan , N . , "Elektrik İletim Şebekelerinin İşlemesi ve Korunması", İTÜ Elektrik – Elektronik Fakültesi, Ofset Baskı Atölyesi , İstanbul, 1991

Indulkar C.S., Vıswanathan B., Venkata S.S., "Reactive Power Constrained Loadability Limits of Series CapacitorCompensated EHV Transmission Lines", IEEE Transactions on Power Systems, Vol. PWRS-2, No. 2, pp. 585-591, August 1987.

Miller T.J.E., "Reactive Power Control in Electric Power Systems", John Wiley&Sons, Inc., 1982.

Mumcu K. Y., "Examination of power-voltage (P-V) curves of energy operators", Marmara University Institute of Science, Master Thesis, Istanbul, 2001.

Overby T.J., DOBSON I., "Q-V Curve Interpretations Of Energy Measures For Voltage Security", IEEE Transactions On Power Systems, Vol. 9, No. 1, pp. 331-340, February 1994.

Pal M.K., "Voltage Stability Conditions Considering Load Characteristics", IEEE Transactions On Power Systems, Vol. 7, No. 1, pp. 243-249, February 1992.

Tang Y., "Voltage Stability Analysis of Power System", Power Systems, Springer, Singapore, Science Press 2021.

Taylor C.W., "Power System Voltage Stability", Mc Graw-Hill, Inc., 1992.

URL 1. https://electrical-engineering-portal.com/electric-power-systems (Date of access: 01/07/2021)

# Chapter 5

**BIOCOMPOSITE HYDROGEL BEADS FROM GLUTARALDEHYDE CROSS-LINKED CHITOSAN COATED BIOWASTE APPLICATION FOR REMOVAL OF METHYLENE BLUE**

*Şerife PARLAYICI[1]*

---

1  Department of Chemical Engineering, Konya Technical University, Konya, Turkey

## 1. Introduction

Industrial activities are developing day by day in the world in line with our vital needs. In the light of these developments, it produces large amounts of wastewater containing dyestuffs. Dyes have been used as coloring agents for almost every product since ancient times. It is estimated that 7,105 tons of 10,000 various dyes and pigments are producedare produced world each year[1]. An annual world production of textile is 30 million tons and for this, roughly 10,000 different dyes are produced over an average of 700,000 tons per year, 10% of these dyes are given to the receiving water environments with the industrial treatment plant effluent [2]. The fact that the dyestuffs used in the products in the industry are resistant to various chemicals such as detergents, heat and light makes it difficult to decompose the dyestuffs. Non-degradable dyestuffs are toxic, carcinogenic, mutagenic and teratogenic for aquatic organisms, rapidly depleting the dissolved oxygen of the receiving environment, causing fluctuations in the chemical oxygen demand and biological oxygen demand of wastewater. Dyes can initiate cancer, cause mutations in genetic sequences, and suppress enzyme activities. With dark colors blocking the sun's rays, the balance of photosynthesis and respiration is disturbed, and as a result, it is among the few reasons that are problematic for the ecosystem [3]. In addition, dyestuffs in wastewater can cause aesthetic deterioration in water and the environment. For all these reasons, the dyestuffs in the wastewater must be removed before being discharged to the receiving environment. Because of the toxicological aspects associated with the diversity and complex molecular structure of paints, it is not easy to completely eliminate paint contamination. Many methods are used to eliminate color-related problems from wastewater. Although these methods are very diverse; physical methods (adsorption, membrane filtration, ion exchange, etc.), chemical methods (ozonation, photochemical method, sodium hypochlorite, fenton reagent, electrochemical method, ultrasonic waves, wet air and chemical flocculation and precipitation with peroxide oxidation) methods and biological methods (aerobic biodegradation, anaerobic biodegradation, biosorption, bioaccumulation) [4].

The fact that many of the chemical and biological methods for the removal of dyestuffs are expensive and have many operational difficulties have led researchers to develop inexpensive and effective physical methods. This has encouraged researchers to use a cheap, economical and reliable method such as the adsorption method. Today, the use of animal, plant and food wastes as an adsorbent and the use of biomass (lichen, fungus, algae, etc.) as a biosorbent for color removal from wastewater are of great interest [5-7].

Chitosan is one of the most abundant natural biopolymers on earth. Chitosan is a kind of polysaccharide, which can be obtained by deacetylation of chitin or by extraction from the shells of crustaceans. Chitosan; It is extensively used in cosmetics, drug release, wound healing, implantation, wastewater treatment, agriculture, food, paper industry applications [8]. In recent years, the use of cihatsan in water purification has been investigated. The chelation property of chitosan is utilized in the purification processes from these areas of use. Chitosan chelates with dyes and ensures that these ions are removed from wastewater or polluted streams. Again, by taking advantage of its chelation feature, suspended dyes, amino acids, proteins, organic substances, etc. found in the waste water released from food processing processes. materials can be recovered. Chitosan mostly has primary aliphatic amino groups. The presence of amino groups makes chitosan suitable for chemical modification. The amino and carboxylic groups in the structure of chitosan provide benefits in the increase of wastewater [9].

Thanks to its cationic structure, chitosan can be easily dissolved in some solutions in a pH <6 environment (Figure 1.). On the other hand, the solubility of chitosan in inorganic acids is quite low. Organic acids (formic acid, acetic acid, citric acid, lactic acid) are generally used to dissolve chitosan. Among them, the most used solvent is acetic acid. At the same time, there are many parameters that affect the solubility of chitosan, such as temperature, solvent concentration and particle size. Studies have shown that chitosan should have at least 75-80% degree of deacetylation for a good solubility. In the acidic environment, the $NH_2$ group exists as $-NH_3^+$ and interacts electrostatically with the anionic groups in the environment. In the protonated state, it behaves as a cationic polyelectrolyte, forms viscous solutions and can interact with oppositely charged molecules and surfaces [10]. In addition, the solubility of chitosan is a very important parameter in terms of its chemical modifications, film or fiber formation.
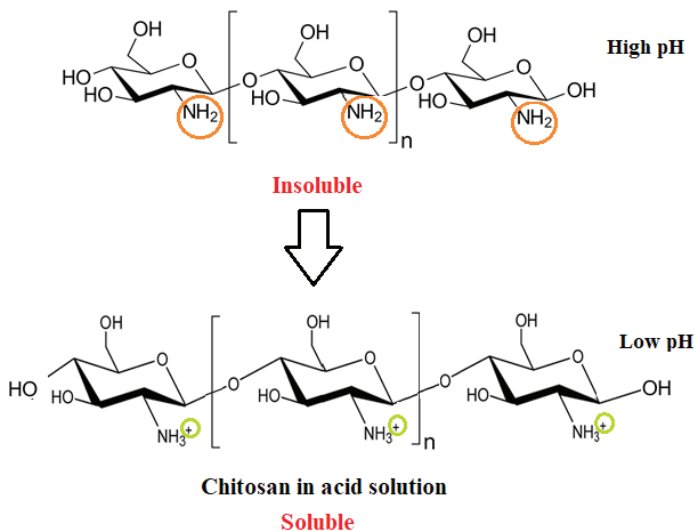
**Fig. 1.** *Transition between protonated (soluble) and deprotonated (insoluble) chitosan.*

However, the low strength of the beads produced by the use of chitosan alone limits their usage areas. Chitosan, which is cationic, allows the formation of electrostatic complexes or multilayer structures with other negatively charged synthetic or natural polymers. Chitosan having functional groups in their chain such as amino (–NH$_2$) and hydroxyl (-OH) can be attached to naturally hydrophilic molecules such as cellulose and starch by specific grafting and cross-linking reactions [11]. Chemical crosslinking is a versatile method in which the mechanical, thermal and chemical stability properties of the material can be increased.

Crosslinkers bind molecules together, increasing molecular weight and generally providing higher mechanical properties and improved stability. Artificial cross-linkers such as formaldehyde, glutaraldehyde, epoxy compounds and carboimide are widely used to modify the material, increase degradation resistance and stabilize it. The crosslinking formation may include two chitosan units, which may or may not belong to the same polymeric chain (Figure 2). Crosslinking using glutaraldehyde is an economical and simple crosslinking method.
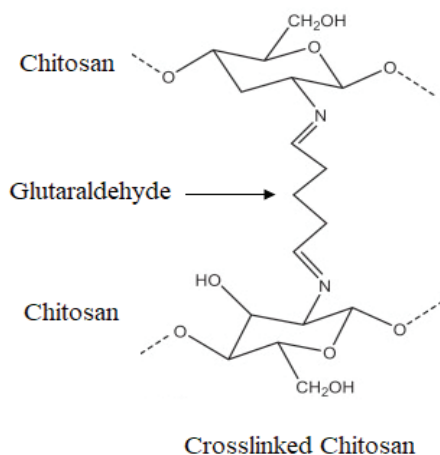
**Fig. 2.** *Chitosan chemically crosslinked with glutaraldehyde.*

The problem of increasing environmental pollution with the developing technology and increasing world population brings to light the need for environmentally friendly, inexhaustible and renewable energy sources. Biomass energy, which is one of renewable energy sources, is used as a raw material in many fields in order not to harm the environment, to be economical, to dispose of wastes and to evaluate them. Biomass can be defined as all natural substances of of including plants, animals, and microorganisms whose main components are carbohydrate components, which can be renewed in a short time. Biomass, especially for developing countries, is one of the widest sources of application, and it is not only renewable, it can be grown everywhere, it provides socio-economic development, and it is counted as a resource such as chemicals. On the other hand, due to the increasing environmental pollution, the efforts to recycle waste have gained importance. In particular, solid wastes are incinerated or disposed of by storage. In both ways, the wastes removed have serious effects on the ecological system. In particular, the use of vegetable wastes as an adsorbent by some pre-treatments or by converting them into composites and thus their recovery research has increased. In addition, the low cost of sorbent production is emphasized as an important advantage. In the literature, there are quite extensive and different studies on this subject [12-15,5,7].

Herein, we report the synthesis of Glutaraldehyde Cross-Linked Chitosan Based cornelian-cherry kernel shells beads copolymer sorbents and an investigation of their equilibrium MB sorption properties in aqueous
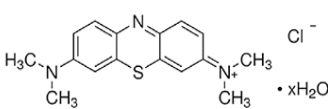
solution. In these study, the optimal set of adsorption parameters were identified first for various combinations of adsorption parameters such as adsorbent dosage, temperature, time, initial concentration of MB solution, initial pH. Kinetic models, isotherm and thermodynamic parameters of adsorption process were determined.

## 2. Materials and Methods

### 2.1. Materials

Cranberry Seed Shell (CSS) used as adsorbent was washed, dried, ground and sieved (125 μm) according to the size range to be used in the experiment. Chitosan (Cht) flakes (degree of deacetylation, DD=75-85%) has been obtained from Sigma-Aldrich. Acetic acid ($CH_3COOH$), glutaraldehyde and ethyl alcohol have been purchased from Merck chemical company. HCl and NaOH (Merck Company) were used to adjust the pH of the prepared solution. The methylene blue used in the experiment was obtained from Acros Organic. 1000 mg/L stock methylene blue solution was prepared. According to the studies, adsorption processes were carried out by making dilutions from this stock solution. Ultrapure water was used to prepare the solutions, dilute, and clean all glass and plastic containers used in the experiments. All properties of MB are given in Table 1. UV-visible spectrophotometer (Schmadzu UV-1700), vibratory sieve shaker (Retsch AS200), oven (Nüve), pH meter (Orion 900S2), Thermostated shaker (GFL 3033 model), analytical balance (Sartorius CP2245), multi magnetic stirrer (IKAMAG) -RO15) was used. The FT-IR spectrum was recorded by a Bruker VERTEX 70 FT-IR spectrometer.

Table 1. *Chemical structures and some other parameters of MB.*

| Commercial name | : | Methylene blue hydrate |
|---|---|---|
| **Structural formula** | : |  |
| **Molecular Formula** | : | C16H18ClN3S. xH2O |
| **Physical Form** | : | Green Fine Crystalline Powder |
| **Molecular Weight (g/mol)** | : | 319.85 |
| **λmax (nm)** | : | 664 |

### 2.2. Preparation of glutaraldehyde cross-linked chitosan based biowaste beads

3 g Cht was dissolved in 300 mL of 1% acetic acid solution with constant mixing until the slurry became was gelled. 3 g of CSS was added

to the Cht gel and stirred for 2 hours to homogeneous solution. The prepared homogeneous mixture was added drop wise using a 5 mL syringe (needle size 0.6 mm) in 200 ml of 5% NaOH and 300 mL of 6% ethyl alcohol solution under constant stirring 100 rpm. Then, the Cht/CSS-glt beads formed and washed until neutral. The beads were reacted with 3 mL of glutaraldehyde at 70 °C for 45 minutes to form a covalent bond with Cts, and the number of functional groups in the structure of adsorbent was increased (Figure 3). Cht/CSS-glt was washed several times to take out excess glutaraldehyde, then dried at ambient temperature.
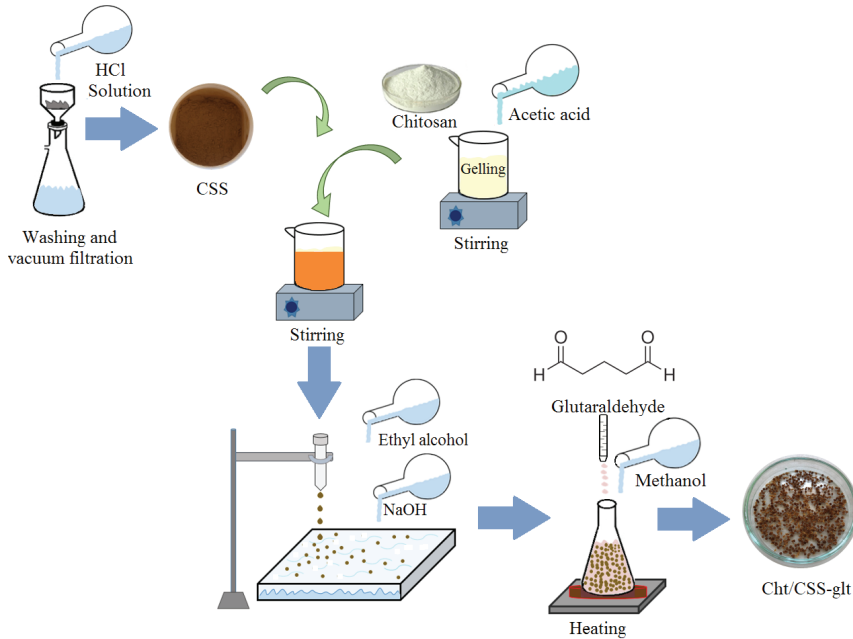


**Fig. 3.** *A schematic diagram presenting the preparation of Cht/CSS-glt*

## 2.3. Batch equilibrium studies

Adsorption of dyes was carried out by mixing 4 g/L Cht/CSS-glt with MB solution in a stoppered glass bottle. All samples were shaken horizontally in a shaker at 200 rpm for optimum time. Then, samples were filtered, UV/Visible, dual beam spectrophotometer was used to analyze the MB concentration of the filtrates at 665 nm. The amount of dye removed from the solution phase was determined as qe (mg/g):

$$\% \text{ Adsorption} = (C_i - C_f)/C_i \ \times 100 \qquad (1)$$

$$\text{Adsorption Capacity} = (C_i - C_f)/m \ \times V \qquad (2)$$

where Ci and Cf are the initial and final MB concentrations, respectively. The adsorption capacity per unit mass of adsorbent (qe) was calculated using Eq. (1 and 2).

## 3.  Results And Discussion

### 3.1.  FTIR

The FT-IR (Bruker VERTEX 70) spectrum of CSS, Cht/CSS and Cht/CSS-glt is given in Figure 4. The band between 3357-3363 cm$^{-1}$ is due to O-H and N-H stretching [12, 16], while at 2958 cm$^{-1}$ aliphatic C-H tension bands are due to the presence of aliphatic structures. The band in the 2000-2300 cm$^{-1}$ range is due to the -C≡C- triple bond of alkynes [17]. Strong peaks of COOH stretching vibrations between 1500-1580 cm$^{-1}$ and -CH, -CH2, -CH3 around 1320-1400 cm$^{-1}$' indicate bending and deformation vibrations [18]. The very intense C-O vibration peak around 1029 cm$^{-1}$ supports the presence of lignin [12, 19]. Aromatic C-H and C-C bonds are moderate and weak in the region between 900-500 cm$^{-1}$. If Cht/CSS beads are in the FTIR spectrum; The peak at 3357 cm$^{-1}$ is due to the –N–H2 tension vibration in the amine functional groups, the COH vibrations of the alcohol groups coming from the chitosan structure, the band at 1374 cm-1 and the bands at 1648 cm$^{-1}$ due to the –N–H2 tension vibration in amide group [18, 20]. The band in the range of 1360-1375 cm$^{-1}$ belongs to -CH vibrations in the alcohol group. The characteristic band of chitosan in the 1151 cm$^{-1}$ range is formed due to (1–4)-glucosidic band in polysaccharide unit [12]. The band around 1025 cm$^{-1}$ belongs to the stretching vibrations of the C-N bonds of amino groups[16]. In the FTIR spectrum of Cht/CSS-glt; It shows evidence of the ether group corresponding to the glutaraldehyde form at 1029 cm$^{-1}$. In some studies, a strong reaction was observed between glutaraldehyde and chitosan, close to 1650 and 1660 cm$^{-1}$ (15, 21).
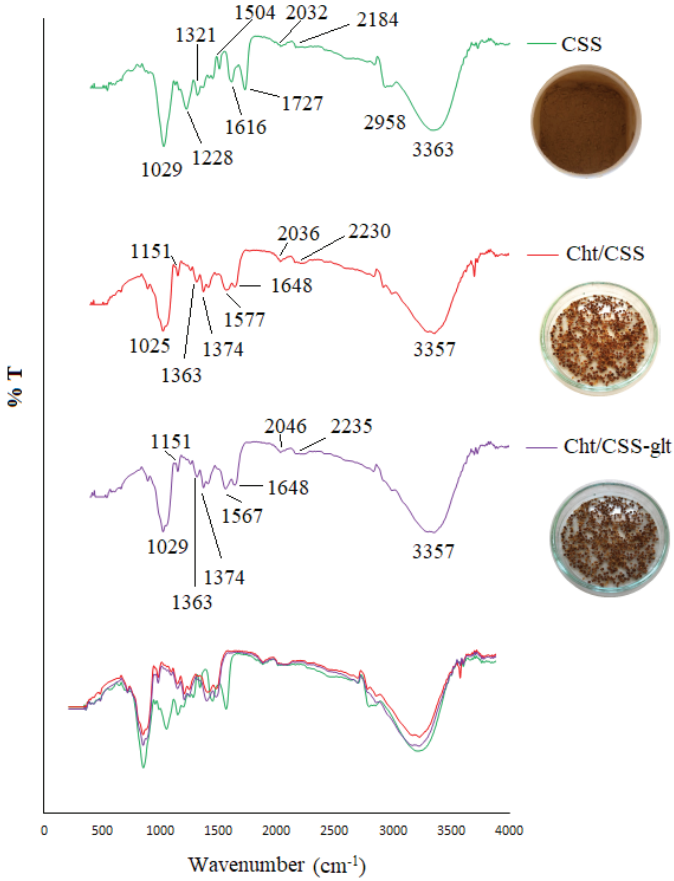
**Fig. 4.** *FTIR spectra of Cht, Cht/CSS and Cht/CSS-glt*

## 3.2. Effect of initial pH

To examine the pH effect on methylene blue adsorption, 0.1M HCl and 0.1M NaOH solutions and MB solutions were pH 2.0; 3.0, 4.0; Adsorption experiments were carried out with fixed parameters set to 5.0 and 6.0. The adsorption capacity and percentages obtained against pH are given in Figure 5. When Figure 5a is examined, it is seen that the amount of methylene blue adsorbed per unit Cht/CSS-glt increases with increasing pH value. At low pH values, it is thought that the adsorption capacity of the chitosan in the Cht/CSS-glt biocomposite decreases as a result of the protonation of the amine groups. Moreover, pH is important parameter for all adsorption processes due to its effects on the surface properties of adsorbents and ionization degrees of functional groups of the adsorbate.

In a solution with a low pH value, the positive charge at the solution interface will increase and the adsorbent surface will behave like a positive charge, resulting in an increase in anionic dye adsorption and a decrease in cationic dye adsorption [22]. This can be interpreted by the protonation of methylene blue at low pH values and the competition of excess H+ ions and dye cations for active sites, as reported by Foo and Hameed [23]. At high pH values, the amount of negatively charged regions that support methylene blue adsorption increases due to electrostatic attraction, while the amount of positively charged regions decreases [24,25]. As a result, an increase in adsorption capacity was observed. The maximum adsorption capacity for MB adsorption was found to be 53.19 mg/g at pH 5.
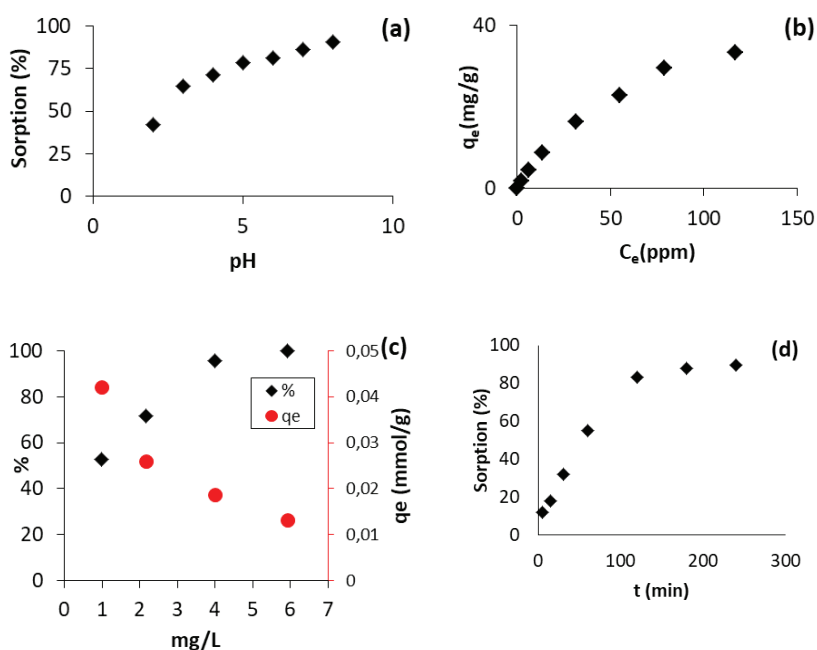


**Fig. 5.** *Effect of pH* **(a)** *Adsorption isotherm* **(b)** *Effect of adsorbent dosage* **(c)** *Effect of contact time* **(d)**

### 3.3. Effect of dye's initial concentration

Adsorption experiments were achieved with a series of samples prepared at different concentrations from the stock MB solution. After equilibrium was attained, the amount of dye remaining in the solution medium was measured with UV-vis. With the increase in the initial concentration, the active areas in the adsorbents are filled with dye molecules. The adsorption capacity is owing to the increase in diffusion on the surface due to the

increase in the concentration difference [26]. Adsorption isotherms provide information about how adsorbate molecules and ions of the adsorbate are distributed between the liquid and solid phase. The different isotherms play a major role in elucidating the mechanism, surface behavior and affinity of the solid towards the adsorbate. The obtained data were applied to Langmuir, Freundlich, Scathard,D-R and Temkin models isotherms, which are frequently used in aqueous solutions (Figure 6). The Kf and n values in Freundlich equation were calculated (Figure 5b). The n value was found to be 2.36, and since this value is between 1-10, adsorption is efficient and $R^2$ is $\geq 0.95$, it fits the Freundlich isotherm. D-R Isotherms adsorption energy value was calculated as 15.43 kJ/mol. Ead value higher than 16 kJ/mol; This strengthens that the adsorption takes place mostly by chemical adsorption. In MB adsorption, the $R^2$ value was found to be 0.99 and the maximum capacity was 71.13 mg/g in the Scatchard isotherm, and it was concluded that it supported the Langmuir isotherm. The qm and b coefficients were calculated and given in Table 2. As a result, when the drawn isotherms are compared, Langmuir isotherm $R^2$ value ($R^2$ ; 0.99) for Cr(VI) being very close to 1 and $R_L$ values of 0.039 showed that it is more suitable for Langmuir isotherm. The Langmuir isotherm indicates that surface is homogeneously distributed. The qm value was calculated as 69.44 mg/g according to the Langmuir isotherm.
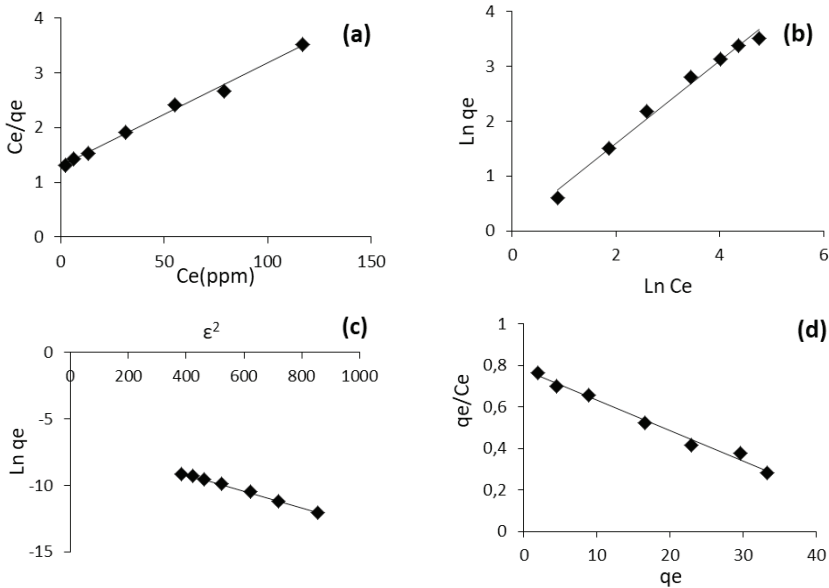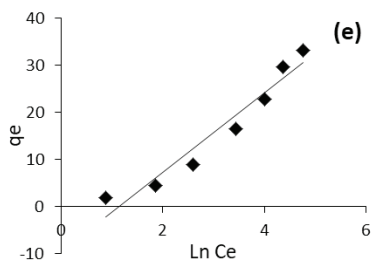
**Fig. 6.** *Langmuir (a), Freundlich (b), D-R (c), Scatchard (d), and Temkin (e) isotherms plot.*

**Table 2.** *Langmuir, Freundlich, Scarthard, D-R and Temkin isotherm parameters.*

| Freundlich | | | Langmuir | | | |
|---|---|---|---|---|---|---|
| $K_f$ | n | $R^2$ | $Q_m$ | b | $R^2$ | $R_L$ |
| 1.084 | 1.33 | 0.988 | 53.19 | 0.015 | 0.995 | 0.733 |
| **Scatchard** | | | **D-R** | | | |
| $Q_s$ | $K_s$ | $R^2$ | $X_m$ | K | E | $R^2$ |
| 53.18 | 0.0146 | 0.989 | 0.0013 | 0.0063 | 8.91 | 0.997 |
| **Temkin** | | | | | | |
| B | | $K_t$ | | $R^2$ | | |
| 8.42 | | 0.319 | | 0.982 | | |

### 3.4. Effect of biocomposite dose

Insufficient amount of adsorbent prevents the adsorption process from taking place effectively, and excess amount means unnecessary economic and experimental burden. For this reason, the most suitable adsorbent amount was determined in the process of adhesion of MB dye on Cht/CSS-glt. The dosage of Cht/CSS-glt used were varied from 0.05, 0.10, 0.15, 0.20, 0.25 g/L for MB while the other parameters (pH, agitation rate, time) are kept constant. Based on Figure 5c, it shows that an increase in the adsorbent's dosage can increase the percentage of MB removal from the solution. In Figure 5c, when the dose of Cht/CSS-glt was increased from 0.01 to 0.1 g, the adsorption percentage of the dye increased from 50.58% to 96.05% and the adsorption capacity was from 379.36 to 72.03 mg. /g decreased. The adsorption percentage increased due to the increase in the amount of adsorbent, the number of available adsorption sites and

the surface area. On the other hand, the adsorption capacity decreases as the active areas on the adsorbent surface are not saturated by reason of overlapping or agglomeration due to the excess of the adsorbent [27]. In the adsorption study with dye, it was observed that the adsorption percentage increased with the increase in the amount of adsorbent. However, 0.1 g adsorbent amount was chosen as the optimum condition to avoid excessive use of adsorbent.

### 3.5. Effect of contact time

The effect of time on the adsorption mechanism was investigated for Cht/CSS-glt. The agitation time was varied from 15 to 1440 min for uptake of MB by Cht/CSS-glt (Figure 5d). Looking at the results, it was seen that the first 30 minutes of adsorption took place very quickly. This is due to the excess of surfactant areas and the loss of the resistance between the liquid and solid due to the mass transfer rate. As the contact time increases, the adsorption decreases as the saturation of the active sites on the surface increases. It was assumed that the adsorption of MB stabilized for Cht/CSS-glt in 60 minutes. When the initial MB concentrations are increased from 50 mg/L to 200 mg/L, capacities at the point where it reaches equilibrium; It increased from 47.5 mg/g to 182.45 mg/g. In other experiments, this time was used as a constant.

### 3.6. Kinetic studies

To investigate the mechanism of the adsorption process, pseudo-first order and pseudo-second order kinetic models were used (Figure 7). Pseudo-first order kinetic and Pseudo-second order kinetic equalities are as stated in Equations (3), (4).

$$In\ (q_e - q_t) = Inq_e - k_1 t \tag{3}$$

$$\frac{1}{q_t} = \frac{1}{k_2\ qe^2} + \frac{1}{q_e t} \tag{4}$$

$qe$ and $qt$ indicate the amount of MB adsorbed at equilibrium and time $t$ (mg/g), respectively. $t$ time (minute), $k1$ pseudo-first-order rate constant (minute-1), and $k2$ pseudo-second-order rate constant. In this study, the experimental results at the initial concentration of 50, 100 and 150 mg/L were used. Parameters calculated from kinetic models for Cht/CSS-glt are given in Table 3. When the correlation coefficients ($R^2$) of both models are compared, the $R2$ value of the pseudo-second-order kinetic model is higher than the pseudo-first-order kinetic model, and the model that best describes the kinetic data is the pseudo-second-order kinetic model.

In this case, it is concluded that adsorption occurs between Cht/CSS-glt and MB via electron sharing or electron exchange. Increasing the initial MB concentration increases the MB dye concentration gradient between the solid and liquid phases, depending on this, the mass transfer of the adsorbate into the adsorbent also increases [28]. This finding supports the chemical interaction accepted by the fictitious second-order kinetic model. In this case, adsorption by chemical interaction takes place between the polar functional groups on the adsorbent surface and MB.
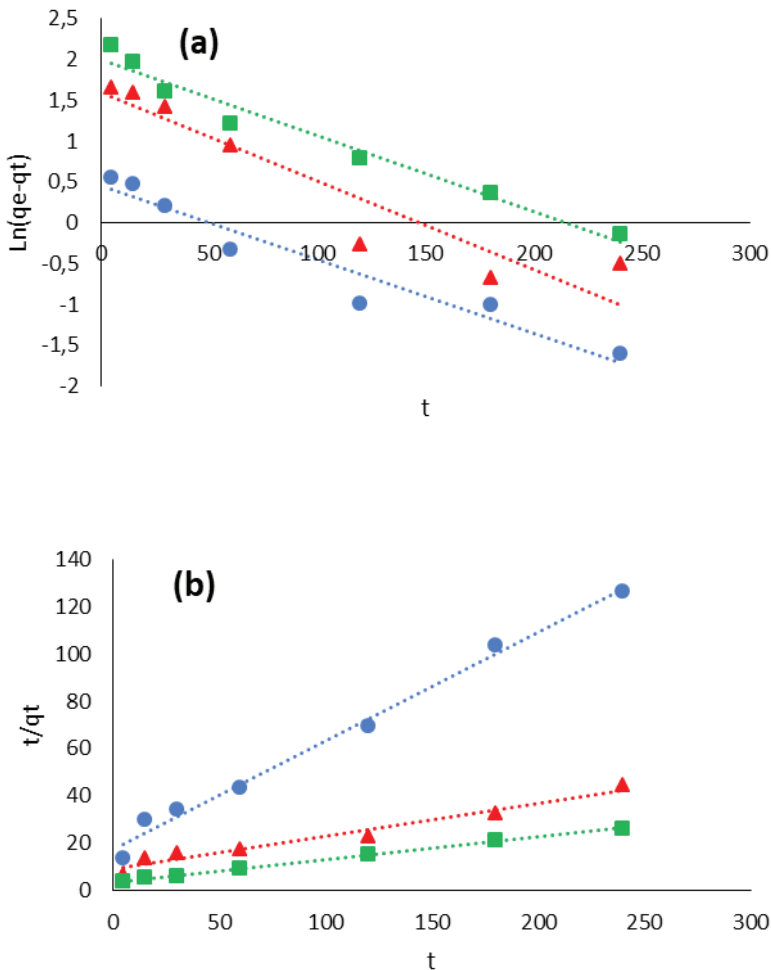


**Fig. 7.** *Kinetic models: (a) pseudo-first-order kinetic model and (b) pseudo-second-order kinetic model.*

**Table 3.** *Kinetic parameters*

| | *Pseudo-first-order* | | | *Pseudo-second-order* | | |
|---|---|---|---|---|---|---|
| $C_o$ | $k_1$ | $qe_{cal}$ | $R^2$ | $k_2$ | $qe_{cal}$ | $R^2$ |
| 10 | 0.009 | 1.56 | 0.937 | 0.0120 | 2.18 | 0.990 |
| 25 | 0.0108 | 4.95 | 0.886 | 0.0022 | 7.14 | 0.965 |
| 50 | 0.0093 | 7.38 | 0.966 | 0.0026 | 10.44 | 0.999 |

## 3.7. Thermodynamic study

The adsorption of methylene blue with Cht/CSS-glt was investigated in terms of thermodynamics. Parameters such as enthalpy, entropy and free energy were calculated with the help of experimental data. The adsorption thermodynamics were determined by working with fixed parameters at 25, 35 and 45°C temperatures. $\Delta H°$, $\Delta S°$ $\Delta G°$ variation parameters were calculated with the help of the graphs, Figure 8, which were drawn using the data obtained from these experiments. Looking at the Table 4., the $\Delta S°$ value was found to be 32.65 kJ/mol.K for Cht/CSS-glt. Positive values of $\Delta S°$ indicate increased irregularity at the solid-liquid interface. Due to the positive $\Delta H°$ values, the adsorption process was endothermic. In addition, it can be thought that chemical adsorption has taken place with the increase of the adsorption capacity as temperature increases [29]. In the literature, similar results have been observed in adsorption studies with different adsorbents in order to remove dyestuffs from aqueous solutions [7, 13]. Whether the adsorption is spontaneous or not depends on the Gibbs free energy, and the negative $\Delta G°$ indicates that the adsorption is spontaneous. In addition, the increase in $\Delta G$ with increasing temperature indicates that MB is more adsorbed at higher temperatures. The $\Delta G°$ for Cht/CSS-glt at temperatures between 298-328 K indicates that adsorption occurs spontaneously [13, 17, 19].
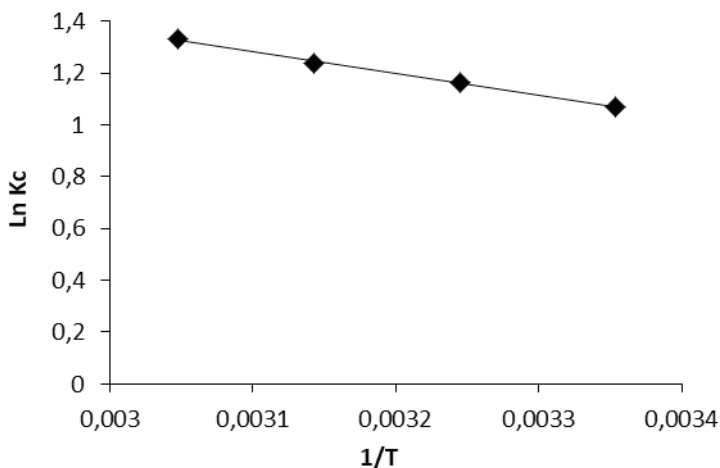
**Fig. 8.** *Effect of temperature.*

**Table 4.** *Thermodynamic parameters.*

| ΔS° | ΔH° | ΔG° (J.mol⁻¹) | | | |
|---|---|---|---|---|---|
| (J.K⁻¹mol⁻¹) | (J.mol⁻¹) | T=298.15 K | T=308.15 K | T=318.15 K | T=328.15 K |
| 32.65 | 7079.67 | -2653.68 | -2980.14 | -3306.60 | -3633.06 |

### 3.8. Comparative evaluation of the adsorption capacities of MB for various adsorbents

There are many studies in the literature on the removal of MB from wastewater and solution media. A comparison of the adsorption capacities of different adsorbent materials in the literature for MB removal is shown in Table 5.

**Table 5.** *Adsorption capacity of the other reported adsorbents for the removal of MB.*

| Adsorbent | qe (mg/g) | References |
|---|---|---|
| H₂SO₄ crosslinked Magnetic chitosan nanocomposite | 20.408 | 20 |
| Chitosan/nano-lignin based composite | 74.07 | 12 |
| Magnetic-modified multi-walled carbon nanotubes | 48.08 | 30 |
| Potassium hydroxide modified apricot kernel shell | 33.67 | 13 |
| Chitosan/Fe₃O₄/graphene oxide nanocomposite | 30.10 | 31 |
| Sulfuric acid-treated coconut shell | 50.6 | 14 |
| Glutaraldehyde Cross-Linked Chitosan Based cornelian-cherry kernel shells | 53.19 | This study |

The maximum adsorption capacity of the prepared Cht/CSS-glt adsorbent for MB adsorption from aqueous solution was found to be 53.19 mg/g. Considering these data given in the table, it is seen that the adsorbent prepared in our study has a high adsorption capacity for MB removal. Therefore, when our study is compared with other studies, it is thought that Cht/CSS-glt is an effective adsorbent in MB removal.

## 4. Conclusions

In this study Cht/CSS-glt was produced by CSS and Cht used as an efficient adsorbent for the removal of MB from aqueous solutions. The adsorption depends on pH and the optimum pH for removal is 5 for MB. The structure in the FTIR spectrum was found to be quite rich in terms of functional groups. The adsorption of methylene blue, which is a cationic dye, is higher due to the increase in basic surface groups with the decrease in acidic surface groups. Since increasing initial concentration provides a driving force, Cht/CSS-glt increased the adsorption capacity, but the removal efficiency decreased with the decrease of accessible areas compared to the adsorbate concentration. MB adsorption is suitable for the Langmuir equation. Adsorption took place on a homogeneous surface and in a single layer in accordance with the Langmuir isotherm. Maximum capacity according to Langmuir isotherm was 53.19 mg/g for MB. The adsorption system conformed to the pseudo-second-order kinetic model, which revealed that chemical mechanisms play a role in the adsorption process. The enthalpy value $\Delta H°=7079.67$ j/mol, the entropy value $\Delta S°=32.65$ j/molK and the free Gibbs Energy value $\Delta G°=-2653.68$, -2980.14, -3306.60 and -3633.06 j/mol for 298.15, 308.15, 318.15 and 328.15 K, respectively, for the adsorption process that took place. These values show us that the adsorption of MB is a voluntary event and a chemical adsorption takes place. Thermodynamic parameters showed that Cht/CSS-glt and MB adsorption took place endothermically. In addition, the negative values of Gibbs free energy changes for Cht/CSS-glt showed that the adsorption process took place spontaneously. It is seen from the results that Cht/CSS-glt beads are effective in removing methylene blue, and it is seen that the removal of dyestuffs in wastewater can be studied as an environmentally friendly, inexpensive and effective method.

# References

1.  Roy, U., Manna, S., Sengupta, S., Das, P., Datta, S., Mukhopadhyay, A., & Bhowal, A. (2018). Dye removal using microbial biosorbents. In Green adsorbents for pollutant removal, Springer, Cham, 253-280.

2.  Rangel, E. M., de Melo, C. C. N., & Machado, F. M. (2019). Ceramic foam decorated with ZnO for photodegradation of Rhodamine B dye. Boletín de la Sociedad Española de Cerámica y Vidrio, 58(3), 134-140.

3.  Priya, E. S., & Selvan, P. S. (2017). Water hyacinth (Eichhornia crassipes)–An efficient and economic adsorbent for textile effluent treatment–A review. Arabian Journal of Chemistry, 10, S3548-S3558.

4.  Parlayıcı, Ş. (2019). Alginate-coated perlite beads for the efficient removal of methylene blue, malachite green, and methyl violet from aqueous solutions: kinetic, thermodynamic, and equilibrium studies. Journal of Analytical Science and Technology, 10(1), 1-15.

5.  Kubra, K. T., Salman, M. S., & Hasan, M. N. (2021). Enhanced toxic dye removal from wastewater using biodegradable polymeric natural adsorbent. Journal of Molecular Liquids, 328, 115468.

6.  Sintakindi, A., & Ankamwar, B. (2021). Fungal biosorption as an alternative for the treatment of dyes in waste waters: a review. Environmental Technology Reviews, 10(1), 26-43.

7.  Parlayıcı, Ş., & Pehlivan, E. (2021). Biosorption of methylene blue and malachite green on biodegradable magnetic Cortaderia selloana flower spikes: modeling and equilibrium study. International Journal of Phytoremediation, 23(1), 26-40.

8.  Morin-Crini, N., Lichtfouse, E., Torri, G., & Crini, G. (2019). Applications of chitosan in food, pharmaceuticals, medicine, cosmetics, agriculture, textiles, pulp and paper, biotechnology, and environmental chemistry. Environmental Chemistry Letters, 17(4), 1667-1692.

9.  Wang, J., & Zhuang, S. (2017). Removal of various pollutants from water and wastewater by modified chitosan adsorbents. Critical Reviews in Environmental Science and Technology, 47(23), 2331-2386.

10. Bhalkaran, S., & Wilson, L. D. (2016). Investigation of self-assembly processes for chitosan-based coagulant-flocculant systems: a mini-review. International journal of molecular sciences, 17(10), 1662.

11. Gulrez, S. K., Al-Assaf, S., & Phillips, G. O. (2011). Hydrogels: methods of preparation, characterisation and applications. Progress in molecular and environmental bioengineering-from analysis and modeling to technology applications, 117-150.

12. Sohni, S., Hashim, R., Nidaullah, H., Lamaming, J., & Sulaiman, O. (2019). Chitosan/nano-lignin based composite as a new sorbent for

enhanced removal of dye pollution from aqueous solutions. International Journal of Biological Macromolecules, 132, 1304-1317.

13. Namal, O. O., & Kalipci, E. (2020). Adsorption kinetics of methylene blue removal from aqueous solutions using potassium hydroxide (KOH) modified apricot kernel shells. International Journal of Environmental Analytical Chemistry, 100(14), 1549-1565.

14. Jawad, A. H., Abdulhameed, A. S., & Mastuli, M. S. (2020). Acid-factionalized biomass material for methylene blue dye removal: a comprehensive adsorption and mechanism study. Journal of Taibah University for Science, 14(1), 305-313.

15. Belho, K., & Ambasht, P. K. (2021). Immobilization of phytase from rice bean (Vigna umbellata Thunb.) on glutaraldehyde activated chitosan microspheres. Journal of Scientific Research, 65, 111-119.

16. Elanchezhiyan, S. S., Preethi, J., Rathinam, K., Njaramba, L. K., & Park, C. M. (2021). Synthesis of magnetic chitosan biopolymeric spheres and their adsorption performances for PFOA and PFOS from aqueous environment. Carbohydrate Polymers, 267, 118165.

17. Kolodynska, D., Bak, J., Koziol, M., Pylypchuk, L.V., 2017. Investigations of Heavy Metal Ion Sorption Using Nanocomposites of Iron-Modified Biochar, Nanoscale Res. Lett. 12.

18. Pap, S., Radonić, J., Trifunović, S., Adamović, D., Mihajlović, I., Miloradov, M. V., & Sekulić, M. T. (2016). Evaluation of the adsorption potential of eco-friendly activated carbon prepared from cherry kernels for the removal of Pb2+, Cd2+ and Ni2+ from aqueous wastes. Journal of Environmental Management, 184, 297-306.

19. Parlayıcı, Ş., & Pehlivan, E. (2019). Comparative study of Cr (VI) removal by bio-waste adsorbents: equilibrium, kinetics, and thermodynamic. Journal of Analytical Science and Technology, 10(1), 1-8.

20. Rahmi; Ismaturrahmi; Mustafa, I. (2019). Methylene blue removal from water using H2SO4 crosslinked Magnetic chitosan nanocomposite beads. Microchemical Journal, 144, 397–402.

21. Gupta, K.C., & Jabrail, F.H. (2006). Glutaraldehyde and glyoxal cross-linked chitosan microspheres for controlled delivery of centchroman. Carbohydrate Research 341, 744-756.

22. Salleh, M. A. M., Mahmoud, D. K., Karim, W. A. W. A., & Idris, A. (2011). Cationic and anionic dye adsorption by agricultural solid wastes: a comprehensive review. Desalination, 280(1-3), 1-13.

23. Foo, K. Y., & Hameed, B. H. (2011). Preparation and characterization of activated carbon from pistachio nut shells via microwave-induced chemical activation. Biomass and Bioenergy, 35(7), 3257-3261.

24. Williams, H. M., & Parkes, G. M. (2008). Activation of a phenolic resin-derived carbon in air using microwave thermogravimetry. Carbon, 46(8), 1169-1172.

25. Rafatullah, M., Sulaiman, O., Hashim, R., & Ahmad, A. (2010). Adsorption of methylene blue on low-cost adsorbents: a review. Journal of Hazardous Materials, 177(1-3), 70-80.

26. Kocaoba, S., Orhan, Y., & Akyüz, T. (2007). Kinetics and equilibrium studies of heavy metal ions removalby use of natural zeolite. Desalination, 214(1-3), 1-10.

27. Ghasemi, R., Sayahi, T., Tourani, S., & Kavianimehr, M. (2016). Modified magnetite nanoparticles for hexavalent chromium removal from water. Journal of Dispersion Science and Technology, 37(9), 1303-1314.

28. Li, W., Zhang, L., Peng J., Li, N., Zhang S. ve Guo, S., 2008. Tobacco stems as a low cost adsorbent for the removal of Pb(II) from wastewater: Equilibrium and kinetic studies, Industrial Crops and Products, 28, 294-302.

29. Cheah W., Hosseini S., Khan M.A., Chuah T.G. and Choong T.S.Y., 2013. Acid modified carbon coated monolith for methyl orange adsorption. Chemical Engineering Journal, 215-216, 747-754

30. Madrakian, T., Afkhami, A., Ahmadi, M., & Bagheri, H. (2011). Removal of some cationic dyes from aqueous solutions using magnetic-modified multi-walled carbon nanotubes. Journal of Hazardous Materials, 196, 109-114.

31. Tran, H. V., Bui, L. T., Dinh, T. T., Le, D. H., Huynh, C. D., & Trinh, A. X. (2017). Graphene oxide/Fe3O4/chitosan nanocomposite: a recoverable and recyclable adsorbent for organic dyes removal. Application to methylene blue. Materials Research Express, 4(3), 035701.

# Chapter 6

## COMPUTATIONALY EFFICENT DESIGN OPTIMIZATION OF MICROSTRIP ANTENNAS

*Peyman MAHOUTI[1]*
*Mehmet Ali BELEN[2]*
*Serdal KARHAN[3]*

1 Associate Professor, İstanbul University - Cerrahpaşa, Department of Electronic and Automation, pmahouti@iuc.edu.tr, ORCID ID: 0000-0002-3351-4433

2 Associate Professor, İskenderun Technical University, Department Electric and Electronic Engineering, mali.belen@iste.edu.tr, ORCID ID: 0000-0001-5588-9407

3 Assistant Professor, İstanbul University - Cerrahpaşa, Department of Electronic and Automation, karaha@iuc.edu.tr, ORCID ID: 0000-0001-6534-1373

90 · Peyman Mahouti, Mehmet Ali Belen, Serdal Karhan

**Introduction**

Thanks to the rapid developments in wireless communication industry, many systems are developed with different standards, GSM, UMTS, WLAN [1]. Each of these requires different operating frequencies and standards where instead of using multiple antenna stages, usage of single multi-band antenna is more convenient [2-5]. For having antenna designs with multiband characteristics many solutions have been proposed in literature such as designs with multi-layer structures [6], stacked designs with novel meta-material [7], planar inverted F antenna (PIFA) elements integrated with two PIN diodes [8], and designs with complex geometrical shape such as star-shaped patch [9]. Although antenna designs with multiband characteristics are efficient solutions for GSM, UMTS, WLAN and other communication protocols, these designs suffers in means of complexity of designs where the designs requires a precise calculation and each of the variables might affect the overall performance of design in case of miscalculations. Thus, design of multiband antenna stages can be considered as a multi- objective, multi- dimension that needs to be solved with high accuracy and stable results. However such design optimization would requires 3D EM simulation results for having precise and stable results which would significantly increase the total simulation time and reduce the computational efficiency. Thus this force designers to make a choice between; (I) using coarse models for design optimization process at the expanse of accuracy, (II) or using fine models at the expanse of having a computationally inefficient design optimization process [10].

A valid solution for solving the mentioned challenging problem is the usage of artificial intelligence algorithms also known as Data-Driven Surrogate models for driving the design optimization process instead of 3D EM simulation methods. Surrogates based models have many applications such as parameter tuning [11], [12], statistical analysis [13]-[15], Multi-objective design [16]-[18]. Although there are series of techniques that can be used for surrogate based modelling such as polynomial regression [19], kriging interpolation [20], radial basis functions [21], support vector regression [22], polynomial chaos expansion [23], Artificial Neural Networks (ANN) [10], [24], and Deep Learning based methods [25-27].

In this work, a study on computationally efficient design optimization of a microstirp patch antenna with dual band characteristics is taken under the study. For this mean, series of Artificial Intelligence based regression algorithms are taken to be used as a regressor tool for creating a mapping between the inputs (Geometrical design variables of the antenna) and the outputs (scattering parameter of the antenna) over the wide frequency range. Firstly, with the use of a 3D EM simulation tool, a data set for training and testing the regression models are generated.

Then selected regression algorithms will be trained and tested using these data sets to determine the most appropriate model to be used in design optimization process. At the next step, meta-Heuristics optimization algorithm assisted by the generated data driven surrogate model will be used for design optimization of a microstrip antenna. In Fig. 1 the flow chart of the proposed Surrogate model assisted design optimization of the microstrip antenna is presented.
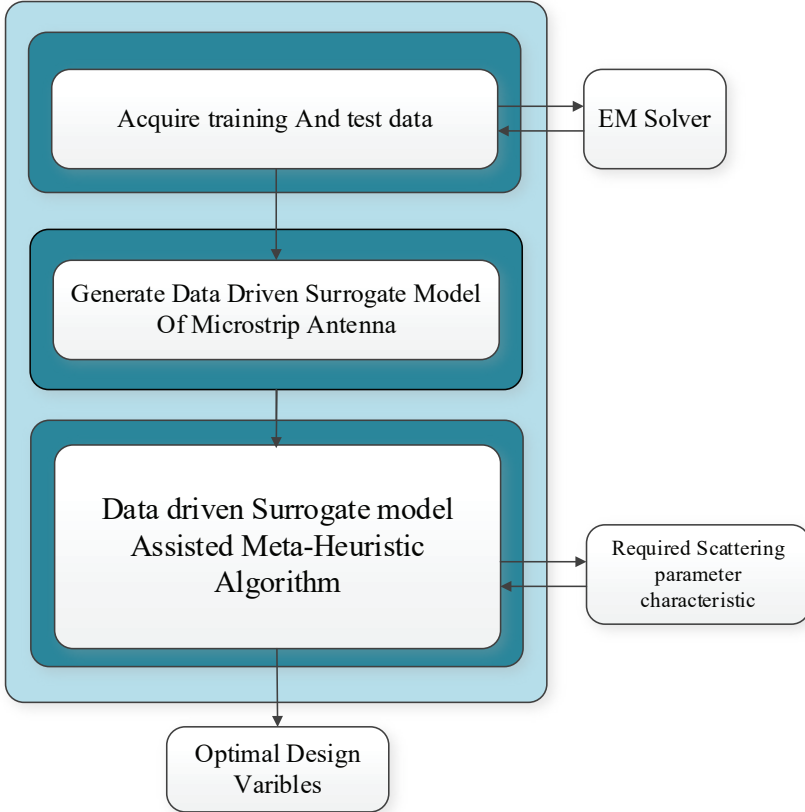


**Figure 1.** *Flow chart of the work.*

## SURROGATE BASED MODELLING OF ANTENNA

### A. Proposed multi-band antenna and its Data set

In this work, a Microstrip antenna (Figs. 2-3) with Inverse C shape and defected ground structure is taken as a suitable design for achieving dual band characteristics. The antenna is consists of a Hexagon radiator with a gap, a 50- microstrip feed line, and a ground plane. A defected

ground structure had been placed in the ground layer to tune the resonance frequency of the design without increasing the overall size of the antenna.
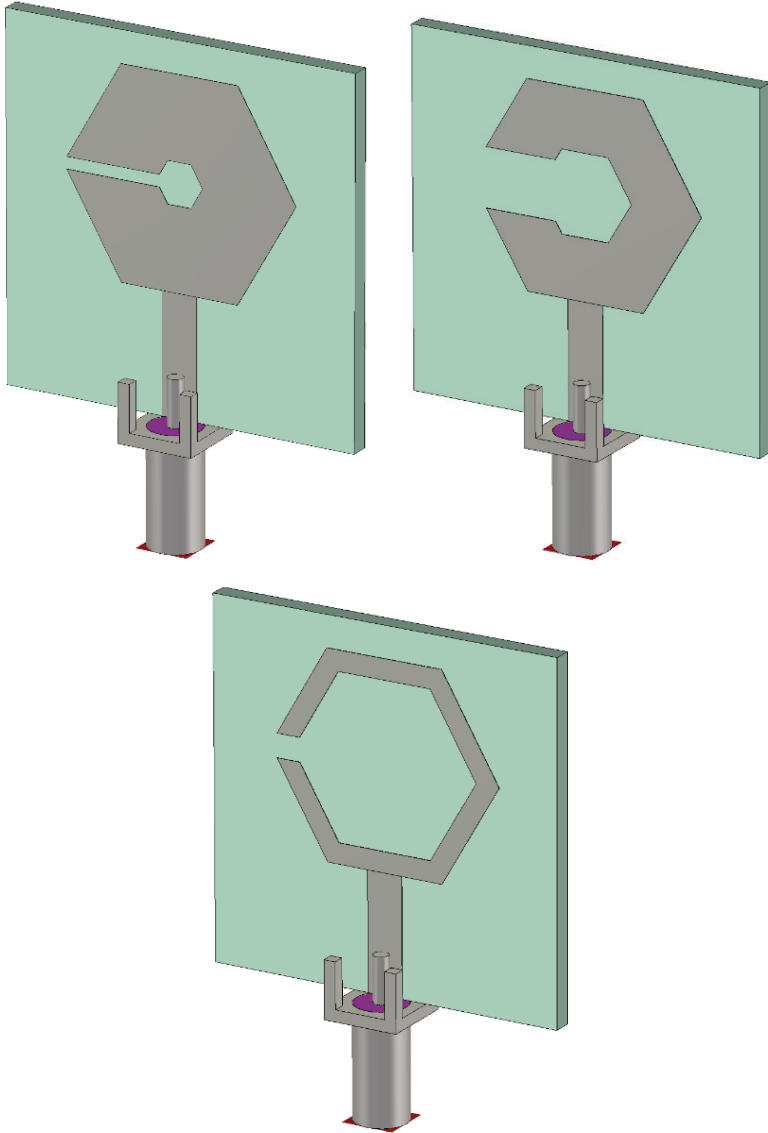


**Figure. 2.** *3D view of the proposed antenna for different design variables values*
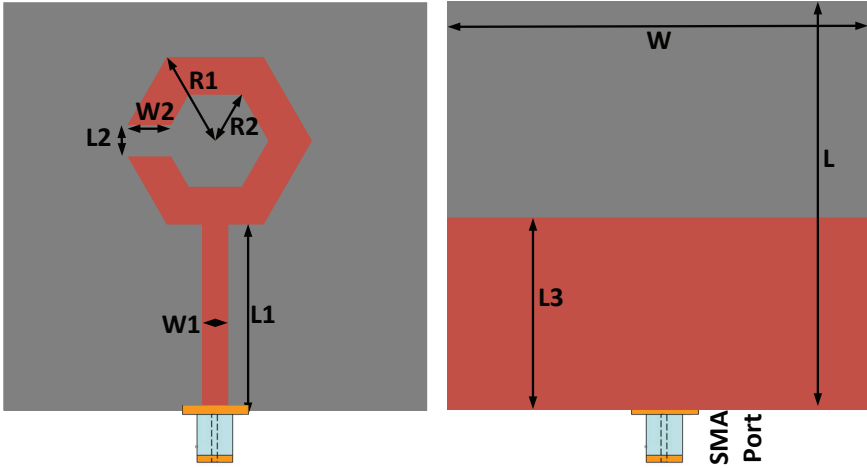
**Figure. 3.** *Schematic of the proposed antenna.*

The upper and lower constraints of the variables of Antenna (Fig. 2) are given in Table 1. Some of the design variables are taken as constant in order to reduce the total number of variables and reducing the dimension size of the optimization problems (L=W=30 [mm], L1=L3, material FR4 H=1.56 [mm] $\varepsilon_r$=4.6).

As it is might known well, high number of training samples usually helps surrogate models to create a better mapping between inputs and outputs of the data set. However, generating such data sets with large samples would also decreases the computational efficiency of the process or even become infeasible. Thus, sampling method that being used for generating the data sets is an important factor for having computational efficient design process. In this work, the antenna has 6 geometrical design variables (Table 1). In case of using traditional liner or grid sampling methods assuming 10 sample point is sufficient for each variables this requires $10^6$ simulation which is almost infeasible, in case of taking sampling point low as 4 sample point for each variables this requires 4096 number of samples which not only is high but also might not be sufficient to generate a good mapping due to the large gaps in variables steps.

Herein, in order to have a comprehensive and computationally efficient data set instead of using traditional linear sampling, Latin-Hyper cube Sampling (LHS) method had been used for sampling design samples from the given ranges in Table 1. By using LHS a training data set with 1000 sample and a test data set with 100 samples had been generated to be used for training the proposed surrogate model. The frequency range of each of

the samples are taken between 0.1-10 GHz with step size of 0.1 GHz. In Fig. 4, the Black-Box representation of the proposed data driven surrogate model has been presented.

**Table 1.** *Design variables and their variation limit.*

| Variable | Min | Max |
|----------|-----|-----|
| W1 | 1 | 10 |
| W2 | 1 | 10 |
| L1 | 5 | 20 |
| L2 | 1 | 5 |
| R1 | 6 | 15 |
| R2 | 1 | 5 |
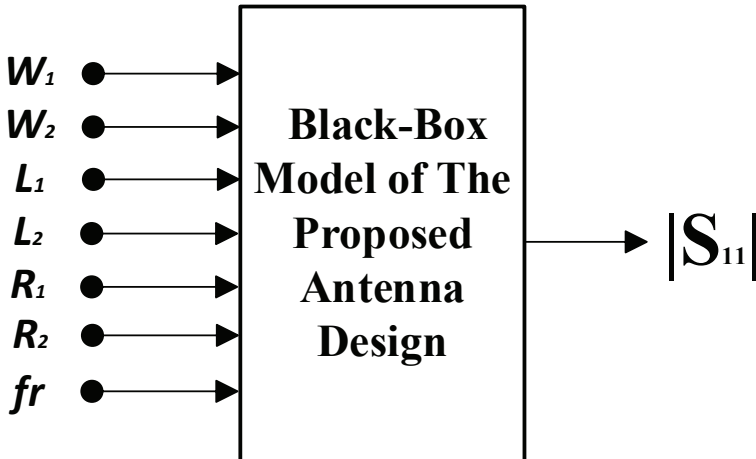


**Figure. 4.** *Black-Box representation of the proposed data driven surrogate model.*

### B.  Data Driven Surrogate Modelling

In this work, the generated data set is being used for training of series of Artificial Intelligence algorithms such as Multi-Layer Perceptron (MLP), Generalized Regression Neural Network (GRNN), Support Vector Regression Machine (SVRM), tree based regression algorithm Gradient Boosted Tree (GBT) and Convolutional Neural Network (CNN) algorithm.

Cross validation method with K=5 had been used for training the

regression models alongside of a hold out data set with 100 samples to further evaluate the performance model with regards to the potential risk of over-fitting. Relative Mean Error (RME) metric (Eq. 1) had been used for calculating the performance of models which are presented in Table 2.

$$(1)$$

$$RME = \frac{1}{N} \sum_{i=1}^{N} \frac{|T_i - P_i|}{|T_i|}$$

Here, $T_i$ is the $i^{th}$ samples targeted value, $P_i$ is the $i^{th}$ samples predicted value, N total number of tested samples over the given operation frequency.

The hyper parameters in the Table 2 are selected via Grid search method based on the models K-fold validation results and the best hyper-parameters setting are taken as final results in Table 3. From Table 3 it can be observed, CNN has the best K-fold and Hold-out performance compared to other regression algorithms. Thus this model is taken as the suitable regression model to assist the meta-heuristic based optimization process in the next section.

**Study Case: Data-Driven Surrogate Assisted Design Optimization**

In this section, a meta-Heuristic Optimization algorithm that is inspired from mating behaviour of Honey Bees is taken as an efficient optimization search tool for the aimed Microstrip antenna design. HBMO is a powerful and efficient algorithm which has the unique capability of hybrid search where a global search is conducted among the hives of a super Bee colony alongside of a local search around the currently known best global solution (Queen of the Colony) using a special nutrition Royal Jelly [28-29]. The main idea in HBMO search is that the known best solution queen would gave birth to new members (usually it is assumed that all new born members are female) of the colony consists of many hives. The new born with better fitness values would compete with their mother to take her crown and become the next Queen of the colony [28]. The cost function that had been used for HBMO search is presented in Eq. (2). The antenna is optimized with respect to the design variables of the problem defined as $x$. By solving

$$x^* = \arg \min_{x} H(x) \qquad (2)$$

where the objective function is defined as

$$H(x) = \max \left\{ f \in [f_{c1}, f_{c2}] : |S_{11}(x, f)| \right\} \qquad (3)$$

The cost function is a minimax task oriented towards minimizing the $S_{11}$ characteristics of the antenna at the aimed operation frequency range of $f_{c1}$ to $f_{c2}$.

**Table 2.** *Hyper-Parameter list of regression algorithms.*

| Model | Hyper-Parameters |
|---|---|
| MLP | Two layers with 15 and 25 neurons |
| GRNN | Spread Parameter 0.78 |
| SVRM | Epsilon SVR, Epsilon=0.25, with Radial basis kernel, |
| GBT | Learning rate of 0.035 8750 number of estimators and depth of 4 |
| CNN | Depth size of 3, Fully Connected Layer Sizes of [64 128 256] |
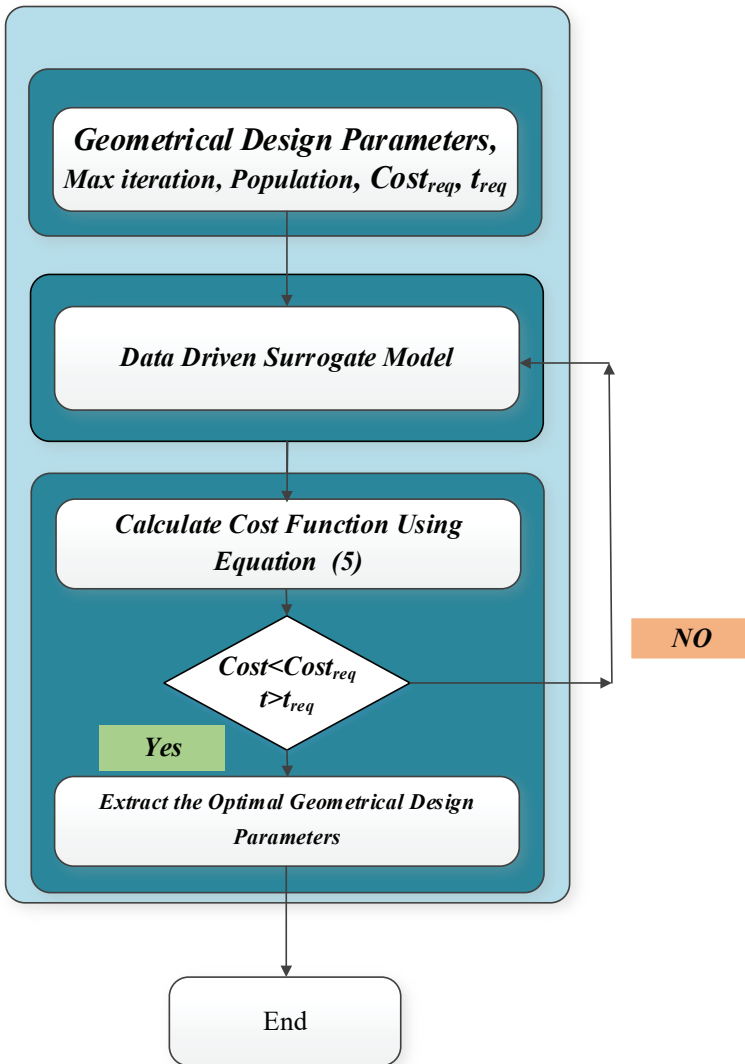


**Figure. 4.** *Schematic of the proposed antenna.*

**Table 2.** *Performance comparison of data driven surrogate model for K-fold and Holdout data sets.*

| Model | K-fold [%] | Holdout [%] |
|-------|-----------|-------------|
| MLP | 9.8 | 10.3 |
| GRNN | 12.5 | 14.6 |
| SVRM | 8.3 | 9.6 |
| GBT | 9.1 | 9.3 |
| CNN | 5.1 | 5.6 |

The cost function is a minimax task oriented towards minimizing the $S_{11}$ characteristics of the antenna at the aimed operation frequency range of $f_{c1}$ to $f_{c2}$. The aimed operation bands of the multi-band antenna are taken as 2.4 and 5.8 GHz for ISM, WLAN applications. The optimally selected geometrical design variables of the antenna are presented in Table 3. As a final validation step, these values had been used for generating a 3D EM model of the antenna which is presented in Fig. 5 (a). The Data driven surrogate model assisted HBMO search for this problem had been used with hyper parameters of [20 iteration, 30 Drone bee, and Royal jelly step size of ±0.05 using HBMO optimization].

**Table 3.** *Optimally selected design values in (mm)*

| W1 | 3 |
|----|---|
| W2 | 6 |
| L1 | 10 |
| L2 | 2 |
| R1 | 10 |
| R2 | 4 |

As a final validation for testing the performance of the proposed technique, the obtained geometrical design values will be used for a 3D EM model (Fig. 5 (a)) and the simulated results of EM simulator will be compared with the results of Data-Driven surrogate model. As it can be seen from Fig. 5 (b), the predicted $S_{11}$ characteristics from the data driven surrogate model is almost same with the simulated result of 3D EM simulation tool which shows that the proposed technique is an accurate and stable solution for computationally efficient design optimization processes.
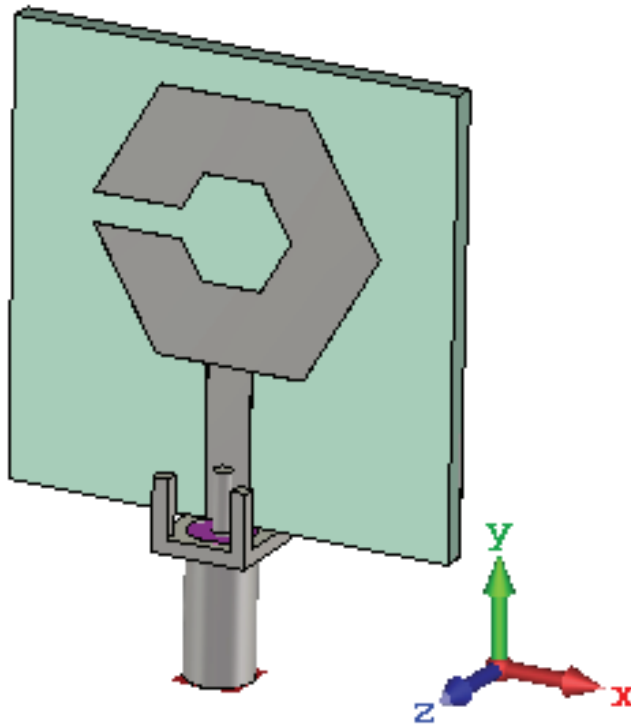
*(a)*
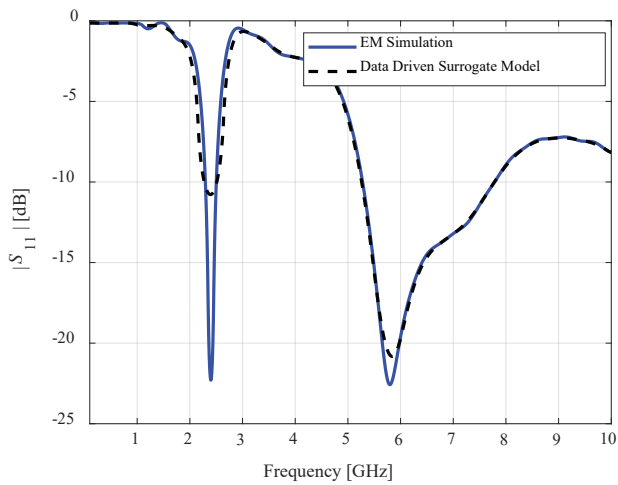


*(b)*

**Figure. 5** *(a) 3D view of the proposed antenna (b) Simulated results of CNN and 3D EM simulator.*

## CONCLUSION

Herein, by means of data driven surrogate model assisted optimization techniques, design optimization of a dual band microstrip antenna had been achieved in a computationally efficient way. CNN one of the commonly used Deep Learning algorithm had been used as an accurate regression tool for creating a mapping between input and outputs of the problem at hand. Then by using this surrogate model, a meta-heuristic search had been assisted for the multi objective multi dimension optimization problem of dual band microstrip antenna. Furthermore, the simulated results of the optimally designed antenna is compared with the predicted results of the CNN based surrogate model. From the comparison it can be said that the proposed technique is an accurate and stable solution for computationally efficient design optimization processes.

## ACKNOWLEDGMENT

# REFERENCES

[1] Mak, A. C. K., Rowell, C. R., Murch, R. D. and Mak, C. (2007). Reconfigurable multiband antenna designs for wireless communication devices. IEEE Trans. on Antennas and Prop., vol. 55, no. 7, pp. 1919-1928.

[2] Ali, T., Prasad, K. D., and Biradar, R. C. (2018). A miniaturized slotted multiband antenna for wireless applications. Journal of Computational Electronics, 17(3), 1056-1070.

[3] Patel, R., Desai, A., Upadhyaya, T., Nguyen, T. K., Kaushal, H., and Dhasarathan, V. (2021). Meandered low profile multiband antenna for wireless communication applications. Wireless Networks, 27(1), 1-12.

[4] Kaur, K., and Sivia, J. S. (2017). A compact hybrid multiband antenna for wireless applications. Wireless Personal Communications, 97(4), 5917-5927.

[5] Hsieh, H., Lee, Y., Tiong, K., and Sun, J. (2009). Design of a Multiband Antenna for Mobile Handset Operations. IEEE Antennas and Wireless Prop. Lett., vol. 8, pp. 200-203.

[6] Ahmed, A. E. and Ali, W. A. E. (2020). A Novel Multiband Antenna with 3D-Printed Multicircular Substrate for Wireless Applications. 2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE), pp. 1-5.

[7] David, R. M., Ali, T., and Kumar, P.A. (2021). Multiband Antenna Stacked with Novel Metamaterial SCSRR and CSSRR for WiMAX/WLAN Applications. Micromachines, 12, pp. 113.

[8] Bharathi, A., Shankar, R., and Gosula, R. (2021). A Novel Compact Multiband Reconfigurable WLAN MIMO Antenna. IETE Journal of Research.

[9] Shaw, M., Chakravorty, D., Islam, S., and Gangopadhyaya, M. (2018, May). A simple Star shaped Microstrip Patch Antenna for penta band application. In 2018 2nd International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech) (pp. 1-3). IEEE.

[10] Mahouti, P. (2019). Design optimization of a pattern reconfigurable microstrip antenna using differential evolution and 3D EM simulation-based neural network model," Int J RF Microw Comput Aided Eng. 29:e21796.

[11] Ullah, U., Koziel, S., and Mabrouk, I. B. (2019). Rapid redesign and Bandwidth/Size tradeoffs for compact wideband circular polarization antennas using inverse surrogates and fast EM-based parameter tuning. IEEE Transactions on Antennas and Propagation, 68(1), 81-89.

[12]   Balaprakash, P., Gramacy, R. B., and Wild, S. M. (2013, September). Active-learning-based surrogate models for empirical performance tuning. In 2013 IEEE International Conference on Cluster Computing (CLUSTER) (pp. 1-8). IEEE.

[13]   Du, J. and Roblin, C. (2017). Statistical modeling of disturbed antennas based on the polynomial chaos expansion. IEEE Ant. Wireless Prop. Lett., vol. 16, p. 1843-1847, 2017.

[14]   Rossi, M., Dierck, A., Rogier, H., and Vande Ginste, D. (2014). A stochastic framework for the variability analysis of textile antennas. IEEE Trans. Ant. Prop., vol. 62, no. 16, pp. 6510-6514.

[15]   Ochoa, J. S. and Cangellaris, A. C. (2013). Random-space dimensionality reduction for expedient yield estimation of passive microwave structures IEEE Trans. Microwave Theory Techn., vol. 61, no. 12, pp. 4313-4321.

[16]   Dong, J., Qin, W., and Wang, M. (2019). Fast multi-objective optimization of multi-parameter antenna structures based on improved BPNN surrogate model. IEEE Access, 7, 77692-77701.

[17]   Easum, J. A., Nagar, J., and Werner, D. H. (2017, July). Multi-objective surrogate-assisted optimization applied to patch antenna design. In 2017 IEEE International Symposium on Antennas and Propagation & USNC/ URSI National Radio Science Meeting (pp. 339-340). IEEE.

[18]   S. Koziel and A.T. Sigurdsson, "Multi-fidelity EM simulations and constrained surrogate modeling for low-cost multi-objective design optimization of antennas," IET Microwaves Ant. Prop., vol. 12, no. 13, pp. 2025-2029, 2018.

[19]   Chávez-Hurtado, J. L. and Rayas-Sánchez, J. E. (2016). Polynomial-based surrogate modeling of RF and microwave circuits in frequency domain exploiting the multinomial theorem. IEEE Trans. Microwave Theory Tech., vol. 64, no. 12, pp. 4371-4381, 2016.

[20]   Queipo, N.V., Haftka, R. T., Shyy, W., Goel, T., Vaidynathan, R., and Tucker, P. K. (2005). Surrogate-based analysis and optimization. Progress in Aerospace Sciences, vol. 41, no. 1, pp. 1-28, 2005.

[21]   Barmuta, P., Ferranti, F., Gibiino, G. P., Lewandowski, A., and Schreurs, D.M.M.P. (2015). Compact behavioral models of nonlinear active devices using response surface methodology. IEEE Trans. Microwave Theory and Tech., vol. 63, no. 1, pp. 56-64.

[22]   Cai, J. King, J., Yu, C., Liu, J. and Sun, L. (2018). Support vector regressionbased behavioral modeling technique for RF power transistors. IEEE Microwave and Wireless Comp. Lett., vol. 28, no. 5, pp. 428-430.

[23]   Petrocchi, A., Kaintura, A., Avolio, G., Spina, D., Dhaene, T., Raffo, A., and Schreurs, D.M.P.-P. (2017). Measurement uncertainty propagation in transistor model parameters via polynomial chaos expansion. IEEE Microwave Wireless Comp. Lett., vol. 27, no. 6, pp. 572-574.

[24] Rayas-Sanchez, J. E. and Gutierrez-Ayala, V. (2006). EM-based statistical analysis and yield estimation using linear-input and neural-output space mapping. IEEE MTT-S Int. Microwave Symp. Digest (IMS), pp. 1597-1600.

[25] Calik, N., Belen, M. A., and Mahouti. P. (2020). Deep learning base modified MLP model for precise scattering parameter prediction of capacitive feed antenna. International Journal of Numerical Modelling: Electronic Networks, Devices and Fields 33, no. 2 (2020): e2682.

[26] Calik, N., Belen, M. A., Mahouti, P., and Koziel, S. (2021). Accurate modeling of frequency selective surfaces using fully-connected regression model with automated architecture determination and parameter selection based on Bayesian optimization. IEEE Access, 9, 38396-38410.

[27] Koziel, S., Mahouti, P., Calik, N., Belen, M. A., and Szczepanski, S. (2021). Improved Modeling of Microwave Structures Using Performance-Driven Fully-Connected Regression Surrogate. IEEE Access 9, 71470-71481.

[28] Güneş, F., Demirel, S., and Mahouti, P., (2016). A simple and efficient honey bee mating optimization approach to performance characterization of a microwave transistor for the maximum power delivery and required noise. Int. J. Numer. Model., vol. 29, pp. 4– 20.

[29] Güneş, F., Demirel, S., and Mahouti, P., (2014). "Design of a front–end amplifier for the maximum power delivery and required noise by HBMO with support vector microstrip model. Radioengineering, vol.23, no.1, pp. 134-143.

# Chapter 7

## BOOSTING - BASED MODELLING OF FREQUENCY SELECTIVE SURFACES

*Hakan KALAYCI[1]*
*Peyman MAHOUTI[2]*
*Mehmet Ali BELEN[3]*
*Umut Engin AYTEN[4]*

1 Phd Student, Yıldız Technical University, Department of Electronic and Communication Engineering, hakan.kalayci@windowslive.com.tr, ORCID ID: 0000-0002-2346-0064

2 Associate Professor, İstanbul University - Cerrahpaşa, Department of Electronic and Automation, pmahouti@iuc.edu.tr, ORCID ID: 0000-0002-3351-4433

3 Associate Professor, İskenderun Technical University, Department of Electric and Electronic Engineering, mali.belen@iste.edu.tr, ORCID ID: 0000-0001-5588-9407

4 Associate Professor, Yıldız Technical University, Department of Electronic and Communication Engineering, ayten@yildiz.edu.tr, ORCID ID: 0000-0003-4174-1799

## 1.  Introduction

Microwave designs that are high-performance and sophisticated have become increasingly important as a result of recent advancements in communication technology to simulate the needed designs, full-wave electromagnetic (EM) analyzers have become a necessity [1, 2]. In the event of recurrent designs or optimization methods that need a large number of model evaluations, these techniques are less efficient [3, 4, 5]. For a computationally efficient process, a rapid and precise model for microwave circuit design optimization issues is essential. Since the last decade, several studies have been done on new techniques of constructing modeling techniques for having highly precise and stable answers that are also computationally efficient in terms of design optimization. As an example, adjoints sensitivity into gradient-based optimization algorithms are an approach [6], accelerating local procedures via sparse sensitivity updates [7], as well as surrogate-assisted procedures involving both data-driven [8], kriging, Gaussian process regression (GPR) [9, 10], neural networks [11, 12],  support vector regression [13], or fuzzy models [14].

 When many models (often called "weak learners") are strategically created and combined to solve a specific computer intelligence challenge, ensemble learning is used. It is primarily used to boost the efficiency of a model (classification, prediction, approximate function approximation, etc.) or to lower the risk of a weak learner group collection. Ensemble learning may also be used to improve model confidence, learner collection of optimal (or nearly ideal) parameters, data merging, progressive learning, incremental learning, and error correction, among other uses [15, 16, 17].
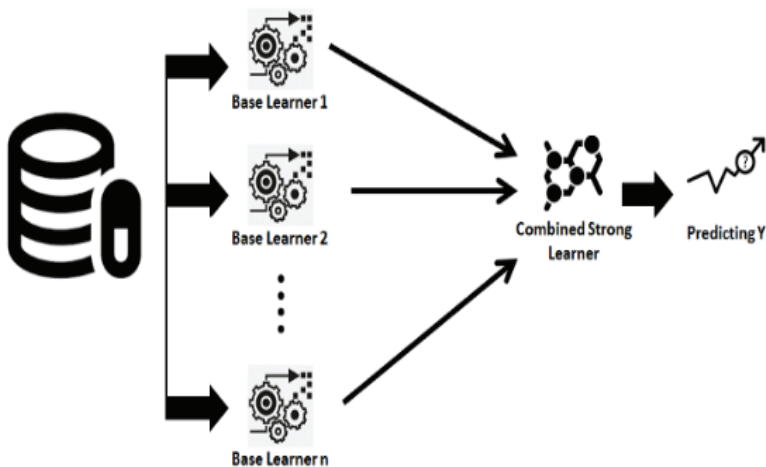


**Figure 1**. Ensemble architecture [18].

The risk of selecting an unsuitable model will be decreased by using learner collections. Ensemble based systems can be beneficial when dealing with enormous volumes of data or a lack of enough data. Partitioning data into smaller sections can be done strategically. Then, each division may be used to train a distinct learner that can be merged with the others. Improved classification/regression decision accuracy can be achieved through the use of data or information fusion. Every learner's decision may be pooled to help determine the optimum overall decision boundary, which lowers the total error. A model that makes varied errors but offers the optimal decision boundary is the goal.

## 2. Boosting Algorithms

Boosting is a technique used in ensemble learning to improve the performance of weak learners in a sequential, iterative manner, and then combine them using a deterministic approach. To create an ensemble of learners, boosters' re-sample the data, which is then combined based on the majority vote. While this is correct, re-sampling is carefully targeted in boosting to generate the most comprehensive training outputs for every subsequent learner.
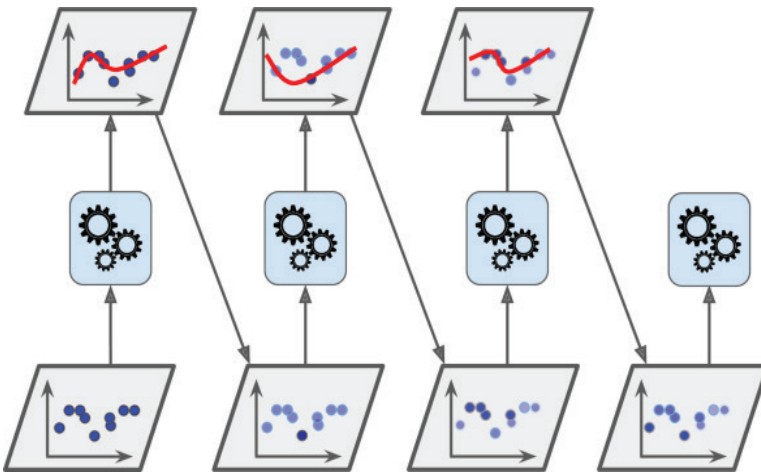


**Figure 2**. Boosting working principle [18].

## A. *Adaptive Boosting (AdaBoost)*

AdaBoost is an adaptive meta-algorithm for classification that can be used in conjunction with a wide range of different types of algorithms for learning. Because it is adaptive, future weak learners are modified in favor of cases misclassified by prior classifiers. As long as the performance of each learner is marginally better than random guessing,

the final model will converge to a strong learner, regardless of how weak the individual learners are [19, 20, 21, 22].

Assume $b(x; r_m)$ is a base learning controlled by parameter $r_m.\beta_m$ controls the addition of weak base learners. M weak base learner's final model $f(x)$ will be as:

Initialize base learner; $f_0$ as $f_0(x) = 0$ and $f(x) = \sum_{m=1}^{M} \beta_m * b(x; r_m)$ $sign(\cdot)$ is the function that converts values to classes. This is the output tree-based structure:

$$G(x) = sign(f(x)) = sign\left(\sum_{m=1}^{M} \beta_m * b(x; r_m)\right) \qquad (1)$$

In the case of D = $\{(x_1, y_1), (x_2, y_2), ..., (x_N, y_N)\}$ with N samples, the global optimal and minimize the below loss function are as follows:

$$(\beta_m, r_m) = Min_{\beta_r, r_m}\left(\sum_{i=1}^{N} Loss(y, \sum_{m=1}^{M} \beta_m * b(x; r_m))\right) \qquad (2)$$

This global optimum can be approximated at each step by finding $\beta$ and r values for the current base learning. Then update the function:

$$f_m(x) = f_{m-1}(x_i) + \beta_m * b(x; r_m) \qquad (3)$$

The final output model $f(x)$ as follows:

$$f(x) = argMin_{\beta_r, r_m}\left(\sum_{i=1}^{N} Loss(y, f_{m-1}(x_i) + \beta_m * b(x; r_m))\right) \qquad (4)$$

Determine the best value of $\beta_m$ and $b_m(x; r_m)$ (simplified as $b_m(x)$) to minimize our exponential loss at each iteration. Assume y belongs to {-1,1}, then the exponential loss function is as below:

$$Loss(y, f(x)) = E(e^{y*f(x)}|x) = P(y = 1|x) * e^{f(x)} + P(y = -1|x) * e^{f(x)} \qquad (5)$$

If we take the derivative and set it to 0, we will find out that when minimizing exponential loss function, we are actually like fitting a logistic regression $y = log\frac{P(y=1|x)}{P(y=-1|x)}$ and we will reach the optimal Bayes error rate:

$$\frac{\partial E(e^{y*f(x)}|x)}{\partial f(x)} = \left[-P(y = 1|x) * e^{-f(x)} + P(y = -1|x) * e^{f(x)}\right] *$$
$$f(x) = 0 \qquad (6)$$

$$\frac{\partial E(e^{y*f(x)}|x)}{\partial f(x)} = 0 \rightarrow f(x) = \frac{1}{2} * log\frac{P(y=1|x)}{P(y=-1|x)} \qquad (7)$$

$$sign(f(x)) = sign\left(\frac{1}{2} * log\frac{P(y=1|x)}{P(y=-1|x)}\right) =$$
$$\begin{cases} 1 & P(y = 1|x) > P(y = -1|x) \\ -1 & P(y = 1|x) < P(y = -1|x) \end{cases} \qquad (8)$$

$$argmax_{y \in \{1,-1\}} P(y|x) = optimal\ Bayes\ Rate \tag{9}$$

$$where;\ sign = \begin{cases} 1 & if\ x > 0 \\ -1 & if\ x < 0 \end{cases} \tag{10}$$

## B. *Gradient Boosting Machine (GBM)*

Gradient boosting [23, 24] is a technique developed by Leo Breiman [25], Jerome H. Friedman [26, 27], Liew Mason, Jonathan Baxter, Peter Bartlett, and Marcus Frean [28] to develop functional gradient-boosting algorithms. For regression, classification and other problems, boosting creates a prediction model in the form of an ensemble of weak prediction models, generally decision trees. In a similar way to previous boosting methods, AdaBoost, it develops the model in a stage-wise fashion, and it generalizes them by providing optimization of an arbitrary differentiable loss function. GBMs that iteratively choose weak hypotheses that lead in the negative gradient direction to optimize a cost function across function space. A number of fields in machine learning and statistics have created boosting approaches that go beyond regression or classification because of this functional gradient perspective. Due to its use of an exponential loss, AdaBoost is more sensitive to outliers. A more robust loss function may be used with GBM since it is continuously differentiable. The Forward Stage-Wise Additive Modeling was discussed in the AdaBoost. Assume iteration *m*:

$$f_m(x) = f_{m-1} + \beta_m * b_m(x) \tag{11}$$

To AdaBoost to decrease loss:

$$Min_{\beta_r,r_m}(\sum_{i=1}^{N} Loss\left(y, f_m(x)\right) \rightarrow Min_{\beta_r,r_m}(\sum_{i=1}^{N} Loss(y, f_{m-1}(x_i) + \beta_m * b(x)) \tag{12}$$

That's only that the situation is different from that of Adaboost in this instance. As the loss function (exponential loss) is known, the exact optimum $b_m(x)$. may be determined in the AdaBoost program. To solve any loss function in GBM, however, the loss function must be a differentiable one. In order to determine the optimal $b_m(x)$, it is necessary to use a method similar to gradient descent by assigning $b_m(x)$ to the direction of the negative gradient.

$$Min_{\beta_r,r_m}(\sum_{i=1}^{N} Loss(y, f_{m-1}(x_i) + \beta_m * b(x) \tag{13}$$

$$b_m(x) = a\ constant * \frac{\partial Loss(y, f_{m-1}(x))}{\partial f_{m-1}(x)} \tag{14}$$

$$f_m(x) = f_{m-1} + a\ constant * \frac{\partial Loss(y, f_{m-1}(x))}{\partial f_{m-1}(x)} \tag{15}$$

$$f_m(x) = f_{m-1} + \eta_m * \frac{\partial Loss(y,,f_{m-1}(x))}{\partial f_{m-1}(x)} \qquad (16)$$

The parameter $\eta_m$ is identical to that of the learning rate, however in this case, it is negative. The definition of GBM regression tree algorithm is presented as follows;

***Model Input***: Dataset D: $D = \{(x_1, y_1), (x_2, y_2),..., (x_N, y_N)\ y_i$ belongs to R\}
***Model Output***: Final regressor: $f_m(x)$
***Loss Function***: Square Loss
***Steps***:
(1) Initialization:$f_0(x) = argmin_\gamma(\sum_{i=1}^N Loss(y, \gamma)$ \qquad (17)

(2) for m in 1,2, 3..., M: calculate the negative gradient:

$$\tilde{y}_i^m = -\frac{\partial Loss(y_i, f_{m-1}(x))}{\partial f_{m-1}(x)} = \frac{\partial[y_i - f_{m-1}(x)]^2}{\partial f_{m-1}(x)} \qquad (18)$$

$$\tilde{y}_i^m = 2 * (y_i - f_{m-1}(x_i)) \qquad (19)$$

(3) Assuming that the CART decision tree splits the area into J different parts $R_{j,m}$, fit a new CART tree by minimizing the square loss:

$$f_0(x) = argmin_{b(x)}(\sum_{i=1}^N[\tilde{y}_i^m - b(x)]^2 \rightarrow \{R_{j,m}\} \qquad (20)$$

$$f_0(x) = argmin_{\{R_{j,m}\}}(\sum_{i=1}^N[\tilde{y}_i^m - b(x_i, \{R_{j,m}\})]^2 \qquad (21)$$

(4) For improved results, instead of performing a linear search to find the most suitable parameter for the entire tree, this could be chosen to find the optimal parameter for each zone inside the tree uniquely:

$$\gamma = argmin_{\eta_j}(\sum_{i=1}^N Loss(y_i - f_{m-1}(x) + \gamma_j * I (x\epsilon\{R_{j,m}\}) \qquad (22)$$

$$\gamma_{j,m} = argmin_{\eta_j}(\sum_{i=1}^N[y_i - f_{m-1}(x) + \gamma_j * I (x\epsilon\{R_{j,m}\})]^2 \qquad (23)$$

(5) Update the function $f_m(x)$:

$$f_m(x) = f_{m-1}(x) + \sum_{j=1}^N[\gamma_j * I (x_i\epsilon\{R_{j,m}\})] \qquad (24)$$

(6) So, we will output our final model $f_M(x)$:

$$f_M(x) = f_0(x) + \sum_{m=1}^M\sum_{j=1}^J \gamma_j * I (x_i\epsilon\{R_{j,m}\}) \qquad (25)$$

## C. eXtreme Gradient Boosting Machine (XGBM)

A second order Taylor's approximation is utilized to connect XGBoost to Newton-Raphson in function space, unlike gradient

boosting, which operates as gradient descent in function space. [29, 30] In GBM, just the first derivative information is used to choose the suitable dimension to minimize loss. In contrast, XGBoost utilizes both the first and second derivatives, resulting in more accurate results than the other two methods. There are four split find methods in XGBoost that can be used to obtain each split at each learner. The definition of XGBM regression tree algorithm is presented as follows;

***Model Input:*** Dataset D: D = $\{(x_1, y_1), (x_2, y_2), ., (x_N, y_N)\ y_i$ belongs to R$\}$

***Model Output:*** Final regressor: $f_m(x)$

***Steps:***

(1) Initialization: $f_0(x)\ argmin_\gamma(\sum_{i=1}^N Loss(y, \gamma)$ $\qquad$ (25)

(2) for m in 1,2, 3..., M:

a. compute

gradient: $g_i = \dfrac{\partial Loss(y_i, f_{m-1}(x))}{\partial f_{m-1}(x)}\ i = 1,2,3,4,..,N$ $\qquad$ (26)

b. compute the second derivative:

$$h_i = \frac{\partial^2 Loss(y_i, f_{m-1}(x))}{\partial^2 f_{m-1}(x)}\ i = 1,2,3,4,..,N$$
$\qquad$ (27)

c. By minimizing the loss function with regularization term, a new decision tree may be fitted:

- New tree; $G_m(x) = \sum_{j=1}^J b_j * I\ (x_i \epsilon R_j)$
  (28)

- Find the best structure;

$\{R_{j,m}\}_{j=1}^J = argmin_{\{R_{j,m}\}_{j=1}^J} (\sum_{j=1}^J (G_j * b_j + \frac{1}{2}(H_j + \lambda) * b_j^2)) + \gamma * J$ $\qquad$ (29)

where;

$G_j = \sum_{x_i \in R_j} g_i\ ,\ H_j = \sum_{x_i \in R_j} h_i$ $\qquad$ (30)

- Best predicted value of tree node; $b_j = -\dfrac{G_j}{H_j + \lambda}$

- Minimal final Loss:

$Loss(y_i, f_m(x)) = \sum_{j=1}^J (-\frac{1}{2} * \frac{G_j}{H_j + \lambda}) + \gamma * J$ $\qquad$ (31)

d.  Update the function $f_m(x)$: $f_m(x) = f_{m-1}(x) + G_m(x)$    (32)

(3) So, output our final model $f_M(x)$:

$$f_M(x) = f_0(x) + \sum_{m=1}^{M} G_m(x)$$
(33)

### i.   Method One: Exactly Greedy Algorithm

As a result of each split of each base learner, the gain is as follows:

**Before Split** $= Loss_{before\ split} = \sum_{j=1}^{J}(-\frac{1}{2} * \frac{(G_L+G_R)^2}{H_L+H_L+\lambda}) + \gamma * J_{no\ split}$   (34)

**After Split** $=Loss\ (y, f_m(x)) = \sum_{j=1}^{J}(-\frac{1}{2} * \left[\frac{(G_L+G_R)^2}{H_L+H_L+\lambda}\right]) + \gamma * (J_{no\ split} + 1)$   (35)

**Gain** $= Loss_{before\ split} - Loss_{after\ split}\ \frac{1}{2} * \left[\frac{G_L^2}{H_L+\lambda} + \frac{G_R^2}{H_R+\lambda} - \frac{(G_L+G_R)^2}{H_L+H_L+\lambda}\right] - \gamma$   (36)

This means that the heart of the exact greedy method is to discover the feature and split point that results in the highest benefit [29].

---

**Exactly Greedy Algorithm for Split Finding**
**INPUT**: $I$, instance set of current nodes, $d$, Feature dimension
**Score** $\leftarrow 0$
$G \leftarrow \sum_{i \in I} g_i$ , $H \leftarrow \sum_{i \in I} h_i$
**for** k = 1 **to m** do
  $G_L \leftarrow 0,$   $H_L \leftarrow 0$
  **for** j in sorted ( $I$, by $X_{jk}$) **do**
    $G_L \leftarrow G_L + g_j, H_L \leftarrow H_L + g_j$
    $G_R \leftarrow G - G_L,$   $H_R \leftarrow H - H_L$
    score $\leftarrow$ max (Score, $\left[\frac{G_L^2}{H_L+\lambda} + \frac{G_R^2}{H_R+\lambda} - \frac{G^2}{H+\lambda}\right]$)
  **end**
**end**
**OUTPUT**: Split with max Score

### ii. Method Two: Approximate Algorithm using Weighted Quantile Sketch

In the case of a large dataset, the greedy method is too slow, and the benefit of each combination can be computed by going over every possible feature and split point. This method is faster and more efficient since the data is divided into preset percentile buckets instead of every potential point. To calculate the percentile bucket, it can be used the second derivative's weight. Assume $D_k$ is a multiset that contains the $k^{th}$ feature values and second order gradient statistics of all training samples. [29]

$$D_k = \{(x_{1k,h_{1k}}),(x_{2k,h_{2k}}),\ldots,(x_{1n,h_{1n}})\}$$
(37)

---

**Approximate Algorithm for Split Finding**
**for** k = 1 **to** m **do**
     Propose $S_k = \{S_{k1}, S_{k2}, \ldots, S_{kl}\}$ by percentiles on feature k
     Proposal can be done per tree (global), or per split (local)
       global: define buckets before building the tree
       local: define buckets before each split
**end**
**for** k = 1 **to** m **do**
    $G_{kv} \leftarrow= \sum_{J\in\{j|S_{k,v}\geq x_{jk}>S_{k,v}-1\}} g_j$
    $H_{kv} \leftarrow= \sum_{J\in\{j|S_{k,v}\geq x_{jk}>S_{k,v}-1\}} h_j$
**end**
Follow same step as in previous section to find max score only among proposed splits

---

Following this is the definition of the rank function $r_k$, which uses weighted second derivative in order to compute percentile:

$$r_k = \frac{1}{\sum_{(x,k)\in D_k} h} \sum_{(x,k)\in D_k, x<Z} h$$
(38)

In order to specify the size of buckets, it is necessary to indicate $\epsilon$ after establishing the rank function. Larger buckets & percentiles are utilized as $\epsilon$ grows.

split point candidate = $\{S_{k1}, S_{k2}, \ldots, S_{kl}\}$

s.t. $\left|r_k(S_{k,j} - r_k(S_{k,j+1})\right| < \epsilon, S_{k1} = min_i x_{ik}, S_{k1} = max_i x_{ik}$
(39)

## iii.    Method Three: Sparsity-aware Split Finding

Missing values are problematic for many tree algorithms before XGBoost. Only non-missing sample data are collected when buckets and split points are created in XGBoost, and then it is checked to see which side of the tree would yield the greatest increase when samples with missing values are placed. [29]

---

**Sparsity-aware for Split Finding**
**INPUT**: $I$, instance set of current nodes, $I_k = \{i \in I \mid x_{ik} \neq missing\}$,
$d$, Feature dimension
also applies to approximate setting, only collect statistics of non-missing entries into buckets
**Score** $\leftarrow 0$
$G \leftarrow \sum_{i \in I} g_i, H \leftarrow \sum_{i \in I} h_i$
**for** k = 1 **to** m **do**
    // enumerate missing value goto right
    $G_L \leftarrow 0, \qquad H_L \leftarrow 0$
    **for** j in sorted ( $I$, ascent order by $X_{jk}$) **do**
      $G_L \leftarrow G_L + g_j, H_L \leftarrow H_L + g_j$
      $G_R \leftarrow G - G_L, \qquad H_R \leftarrow H - H_L$
      score $\leftarrow$ max (Score, $\left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{G^2}{H + \lambda}\right]$)
    **end**
    // enumerate missing value goto left
    $G_R \leftarrow 0, \qquad H_R \leftarrow 0$
    **for** j in sorted ( $I$, ascent order by $X_{jk}$) **do**
      $G_R \leftarrow G_R + g_j, H_R \leftarrow H_R + g_j$
      $G_L \leftarrow G - G_R, \qquad H_L \leftarrow H - H_R$
      score $\leftarrow$ max (Score, $\left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{G^2}{H + \lambda}\right]$)
    **end**
**end**
**OUTPUT**: Split and default directions with max Score

---

## iv.    *System Design of XGboost*

To top it off, XGBoost has a superb system architecture that includes Column Blocks, Cache-aware Access, and Blocks for Outside the Core Computing (OOC). Since Random Forest trees are autonomous, parallel

learning was only possible with Random Forest until XGBoost was developed. Since each subsequent split is reliant on the outcomes of the prior base learners, we cannot create base learners in parallel using standard GBM methods. When using XGBoost, data is stored in CSC format, the gradient & second derivatives of each feature are computed simultaneously and the optimum split points are determined simultaneously.

## D. *Light Gradient Boosting Machine (LGBM)*

LightGBM is a free and open source distributed gradient boosting framework. For ranking, classification and other machine learning tasks it relies on decision tree algorithms. [31] A number of XGBoost's benefits are available in LightGBM, such as sparse optimization and parallel training. What sets GBM apart is that it has to look at all the potential features and split points in order to determine the optimum split point.

---

**The Histogram Algorithm**

**INPUT**: $I$: training data, $d$: max depth

**INPUT**: $m$, Feature dimension

nodeSet $\leftarrow \{0\} \rhd$ *tree nodes in current level*

rowSet $\leftarrow \{\{0,1,2,..\}\} \rhd$ *tree nodes in current level*

**for** i = 1 **to** m **do**

  **for** node in nodeSet **do**

    usedRows $\leftarrow rowSet[node]$

    **for** k= 1 **to** m **do**

      $H \leftarrow new\ Histogram\ ()$

      $\rhd$ ***Build Histogram***

      **for** j in usedRows **do**

        **bin** $\leftarrow I.$f[k][j]. bin

            $H[bin].y \leftarrow H[bin].y\ +\ I.y[j]$

            $H[bin].n \leftarrow H[bin].n\ +\ 1$

    Find the best split on histogram H.

    ....

    ....

    Update rowSet and nodeSet according to the best split point

---

XGboost's Approximate Algorithm Using Weighted Quantile Sketch is quite similar to the Histogram Algorithm [31].

When using LightGBM's Histogram Algorithm, every feature in the dataset is broken down into sets of bins. Samples with A features values

between [0,0.1] will be placed in bin 1, [0.1,1] will be placed in bin 2, and so on.
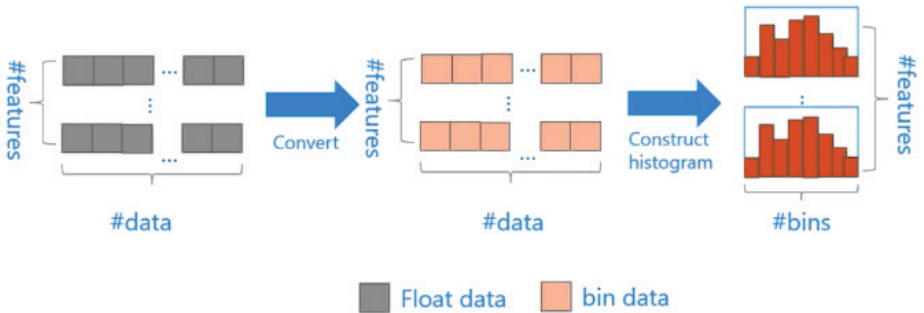


**Figure 3**. Histogram Algorithm.

LightGBM, are Neither categorical nor continuous characteristics problem, official website offers four key advantages, such as; improving efficiency in calculating each split's profit; histogram subtraction may be used to accelerate the process even more; reducing the amount of RAM that is used; and reducing the cost of communication when studying in parallel.

The LightGBM algorithm utilizes two novel techniques called Gradient-Based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB) which allow the algorithm to run faster while maintaining a high level of accuracy.[13]

## i.   *Gradient-Based One-Side Sampling:*

As the GBDT has no unique weight for data instances, Gradient-Based One-Side Sampling (GOSS) takes use of it. A greater gradient will contribute more to the computation of information gain since data instances with various gradients have distinct roles to play. [31] For this reason, GOSS retains instances with high gradients and randomly removes instances with tiny gradients in order to maintain the accuracy of its information collection. [13] Histogram Algorithm is not the only technique proposed by LightGBM. It also recommends the GOSS method to further minimize RAM use and to save computation time for identifying the optimal features and split points. To reduce time while determining the optimum splits, we can exclude samples with low gradients. As a result, in GOSS, we keep examples with high gradients, while also randomly sampling cases with low gradients GOSS provides a constant multiplier for the data instances with tiny gradients while computing the information gain in order to compensate for the effect of only picking a subset of the samples with small gradients.

GOSS uses variance gain $\tilde{V}_j(d)$ on feature j, given split point d, as shown below;

$$\widetilde{V}_j(d) = \frac{1}{n}\left(\frac{\left(\sum_{x_i \in A_l} g_i + \frac{1-a}{a}\sum_{x_i \in B_l} g_i\right)^2}{n_l^j(d)} + \frac{\left(\sum_{x_i \in A_r} g_i + \frac{1-a}{a}\sum_{x_i \in B_r} g_i\right)^2}{n_r^j(d)}\right) \quad (41)$$

The instance set A has high gradients, whereas the instance set B has low gradients, and vice versa gradient of each sample $g_i$ In addition, the following terms are defined: $A_l$, $B_l$, $A_r$, $B_r$, $n_l$, $n_r$

$$A_l = \{x_i \in A : x_{ij} \le d\} \qquad\qquad B_l = \{x_i \in B : x_{ij} \le d\} \ (42)$$

$$A_r = \{x_i \in A : x_{ij} \le d\} \qquad\qquad B_r = \{x_i \in B : x_{ij} \le d\} \ (43)$$

$$n_r^j(d) = \sum \| (x_i \in (A_r \cup B_r)) \quad n_l^j(d) = \sum \| (x_i \in (A_l \cup B_l)) \quad (44)$$

As a result, maximizing $\widetilde{V}_j(d)$ is required to determine the optimal split point d of a feature j. Traditional GBDT has the following loss of variance:

$$\widetilde{V}_j(d) = \frac{1}{n_o}\left(\frac{\sum_{x_i \in O_l} g_i^2}{n_l^j(d)} + \frac{\sum_{x_i \in O_r} g_i^2}{n_r^j(d)}\right) \quad (45)$$

Where O is the current node. $g_i$ is the gradient of each samples. And $O_l$, $O_r$, $n_o$, $n_r$ are define as below:

$$O_l = \{x_i \in O : x_{ij} \le d\} \qquad\qquad n_l^j(d) = \sum \| (x_i \in (O_l)) \ (46)$$

$$O_r = \{x_i \in O : x_{ij} \le d\} \qquad\qquad n_r^j(d) = \sum \| (x_i \in (O_r)) \ (47)$$

$$n_o = \sum \| (x_i \in O) \qquad\qquad\qquad\qquad\qquad\qquad\qquad (48)$$

## ii. *Exclusive Feature Bundling*:

A sparse feature space is characterized by several characteristics that are almost mutually exclusive, meaning they rarely take nonzero values at the same time Features that can only be accessed once are an excellent example of unique features. [31]

---

**Greedy Bundling**
**INPUT**: *F: features, K: max conflict count*
construct graph G
searchOrder ← $G.sortByDegree()$
bundles← {}, $bundlesConflict$ ← {}
**for** i **in** searchOrder **do**

```
        needNew ← True
      for j= 1 to len(bundles) do
        cnt ← ConflictCnt(bundles[j], F[i])
        if cnt + bundlesConflict[i] ≤ K then
           bundles[j]. add(F[i]), needNew ← False
           break
      if needNew then
           Add F[i] as a new bundle to bundles
Output: bundles
```

In order to enhance efficiency, EFB combines various characteristics, decreasing dimensionality while retaining a high degree of accuracy. While identifying the best features, LightGBM also provides a way to minimize feature numbers to save computation time. In fact, since one-hot encoding typically creates a high number of sparse features, this EFB method may be extremely beneficial in many situations.

```
Merge Exclusive Features
INPUT: numData: number of data
INPUT: F: one bundle of exclusive features
binRanges← {0}, totalBin ← {0}
for f in F do
     totalBin + = f. numBin
     binRanges.append(totalBin)
newBin ← new Bin(numData)
for i= 1 to numData do
     newBin [i] ← 0
     for j= 1 to len(F) do
       if F[j].bin[i] ≠ 0 then
         newBin[i]← F[j]. bin[i] + binRanges[j]
Output: newBin,binRanges
```

For example, partitioning features into a small number of distinct bundles was found to be NP-hard in the LightGBM article. The problem is viewed by LightGBM as Graph Color Problem (GCP) and is solved using a greedy technique. For each pair of features that are not mutually exclusive, it creates an instance of the issue by creating a group G= (V; E), then using the features as vertices and adding edges for each pair of features. These edges are weighted according to the total conflicts between two features. Since K is constrained on the maximum number of conflicts in each bundle, the aim becomes a GCP issue. A graph with weighted edges is initially created, whose values correspond to the total

conflicts between features (there are multiple ways to define conflicts). Second, it ranks the features in the graph in descending order based on their degrees. Each feature in the ordered list defined above is checked, and then either assigned to an existing bundle with a minor conflict (regulated by threshold K) or created as a new bundle. In addition, it provides a novel approach of bundling features without the use of graphs. A novel ordering technique was developed by counting the number of nonzero values, which is comparable to ordering by degrees because more nonzero values generally lead to greater disagreement. A critical part of creating a feature bundle is making sure you can identify the original features' values by looking at the feature bundles. This may be done by adding offsets to the initial values of the features, as LightGBM utilizes a histogram-based method.

### iii. *Growing Strategy*

LightGBM and XGBoost are two of the most successful GBDT models before LightGBM that apply a level-wise method for tree growth This is because the model always divides all leaves into two leaves because of the level-wise tree growth approach.
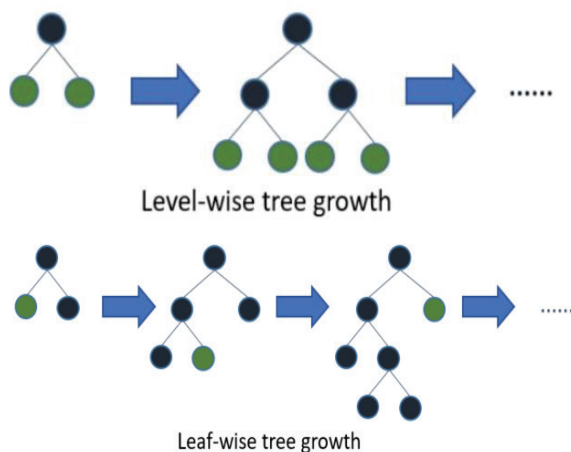


Figure 4. Growing Methods.

LightGBM, on the other hand, utilizes the Leaf-wise tree-growing method. In other words, it will only develop a leaf with the maximum delta loss. According to the official website of LightGBM, leaf-wise algorithms tend to achieve smaller losses than level-wise algorithms when the number of leaves is held constant. Indeed, the Histogram Algorithm, GOSS, and EFB methods in LightGBM not only enhance training speed and RAM use, but also function as the body's natural anti-overfitting mechanism. Because they reduce the model's sensitivity to

small changes in the data collection as a result, the Leaf-wise tree-growing technique in LightGBM works properly.

## 3.  Boosting Algorithms Difference

AdaBoost contains a feature that adds weights to samples that were wrongly categorized. Use the negative gradient as an indication of prior base learners' mistakes in GBM. When samples are successfully categorized, AdaBoost reduces the weights for those samples. As a result of GBM, the next base learner is fitted to approximate the negative gradient of prior base learners. XGBoost Uses both first derivative & second derivative. GBM Only uses the first derivative to find the best base learners at each stage. The initial version of GBM did not include a regularization term in the loss function. Regularization step is added to the loss function by XGBoost. For regression, GBM uses MSE as a scorer to determine the optimal base learners. Better scoring for XGBoost considers over fitness. There is no support for sparse datasets in GBM but XGBoost supports sparse datasets directly. Overfitting is prevented by pre-pruning in GBM in addition to preventing overfit, XGBoost uses post-pruning to avoid under-fit. LightGBM Uses Histogram Algorithm/Gradient-based One-Side Sampling/Exclusive Feature Bundling. XGBoost Uses Level-wise tree growing strategy but Leaf-wise tree-growing technique is used by LightGBM. XGBoost only supports feature parallelism and categorical features as well as direct feature support. There are several ways to use LightGBM simultaneously: Feature Parallel, Data Parallel, Voting Parallel.

## 4.  Case Study: Frequency Selective Surface modelling using Ensemble Learning

Frequency Selective Surfaces (FSS) are one of the sub classes of Meta materials [32] where with respect to their geometrical design variables these designs can alter characteristics of EM waves. Some of the application of FFS can be named as (i) Design of Filtennas, (ii) gain Enhancement of antenna stages, (iii) Reflect array antenna designs, (iv) absorbers.

In this work, modelling of a FSS structure is taken as a study case for testing the performance of Ensemble Learning methods. For this mean a unit element design (Fig. 5) is modelled in 3D EM simulation environment. Then by using Latin Hyper Cube sampling method 1000 samples had been generated through the variable domain defined in Table 1. the aim of the Ensemble algorithms in this case is to create a mapping between input (Geometrical design variables Table 1) and outputs (scattering parameter response of the FSS element). In the next section,

performance results of the Ensemble methods for selected study case is presented.
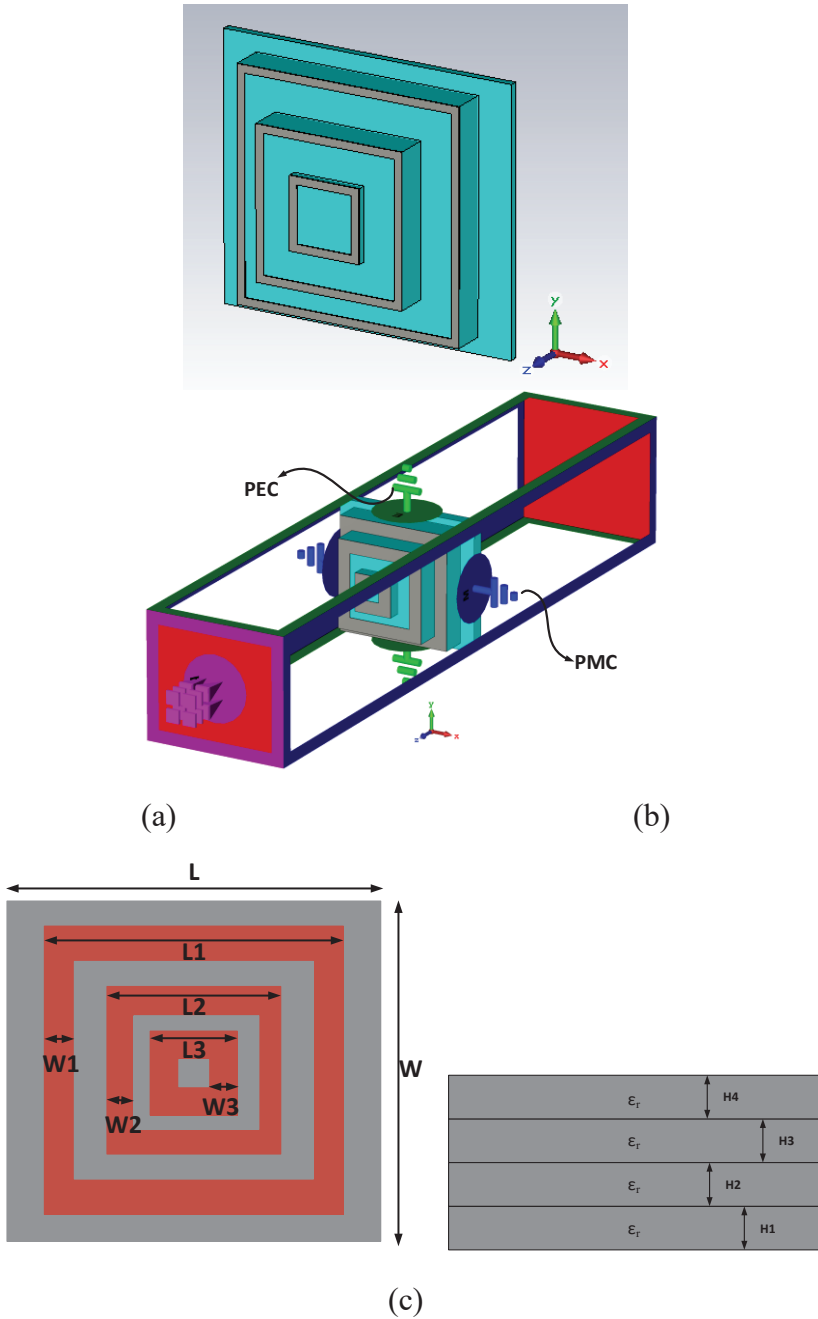


(a)

(b)

(c)

**Figure. 5** (a) 3D view of FSS unit Element, (b) Unit Elements with its boundary condition settings, (c) Schematic of the Unit element.

**Table 1.** Variable space of FSS unit Element.

| Parameter | Lower Boundary | Upper Boundary |
|---|---|---|
| L1 [mm] | 7 | 9 |
| L2 [mm] | 4 | 6 |
| L3 [mm] | 1 | 4 |
| $\varepsilon_r$ | 1.3 | 2.7 |
| **Constant design parameters** | | |
| L [mm] | 11.34 | |
| W [mm] | 5 | |
| H1 [mm] | 1 | |
| H2=H3=H4 [mm] | 2 | |
| W1=W2=W3 [mm] | 1 | |

## A. Verification Case Studies and Benchmarking

In this section the performance results of ensemble learning methods for selected FSS unit element is presented. The data set generated in previous section is used for creating the regression models. The samples in the data set are randomly shuffled, 80% of the data set had been used for training and validation process, while the remaining %20 of the samples are taken as holdout data set for evaluation of over-fitting of the algorithms.

In Table 2, the optimal hyper parameters of each model had been presented. These parameters are obtained via K-fold cross validation method using %80 of the data set. However, in regression models' algorithms tends to fall into over fitting where the memorize the inter-relations of the validation data sets and fails to create a more generic model that also have similar performance for data samples other than validation data. For preventing such case, an additional test process is required named as holdout data set validation. Where a partition of the data set (Here 20%) is taken outside of the modelling process and only presented to the models after finalized optimal hyper parameter determination process. In Table 3, the performance results of the models for both K-fold validation and Holdout data sets are presented. As it can be seen, excluding the ADABOOST method other methods had managed to obtain a high accurate results in means of Mean Absolute Error Metric. It should be noted that between remaining methods of GBM, XGBM, and LGBM although they have similar performance results, their training time is different. If a model is required with fast training time then LGBM can be considered at the expanse of a slightly decrease in the accuracy of the model however if the training time can be neglected then XGBM method is an optimal solution for this type of problems. Thus in this work for furthered justification studies XGBM algorithm will be considered as the proposed data driven surrogate model.

**Table 2.** Hyper parameter list of models.

|  | ADABOOST | GBM | XGBM | LGBM |
|---|---|---|---|---|
| **Learning rate** | 0.07311 | 0.08573 | 0.02243 | 0.0961 |
| **Max depth** | This Algorithm has not this parameter | 9 | 14 | 25 |
| **n estimators** | 509 | 1121 | 1040 | 1927 |
| **Num leaves (*)** | These Algorithms has not this parameter | | | 25 |

(*) Because of LGBM's "2^max_depth > num_leaves" rule, parameter of Num leaves is given only LGBM simulations. Theoretically The reason is that a leaf-wise tree is typically much deeper than a depth-wise tree for a fixed number of leaves

**Table 3.** Performance Results of each model for K-fold and Holdout validation sets.

| Model | K-fold validation ($x10^{-3}$) | Holdout ($x10^{-3}$) | Estimated Time (min) |
|---|---|---|---|
| **ADABOOST** | 100.1 | 199.4 | 38.8 |
| **GBM** | 2.25 | 3.05 | 86.8 |
| **XGBM** | 1.9 | 2.54 | 17.3 |
| **LGBM** | 4.3 | 5.81 | 2.1 |

Furthermore, for justification of the proposed technique, the predicted values form the generated data driven model is compared with the experimental results. A Data-driven surrogate model assisted optimization process is taken into consideration [33] for creating an FSS design operates at X band as a filter for filtering the signals at 10 GHz operation frequency. Then the optimally selected geometrical design values of FSS are used to prototyped a model using 3D printing technology (Fig. 6). In Fig. 7, the measured characteristics of the FSS unit element are compared with the predicted results of the XGBM. Thus, as it can be seen form the results the proposed method is an efficient, accurate and viable technique for modelling of FSS elements.
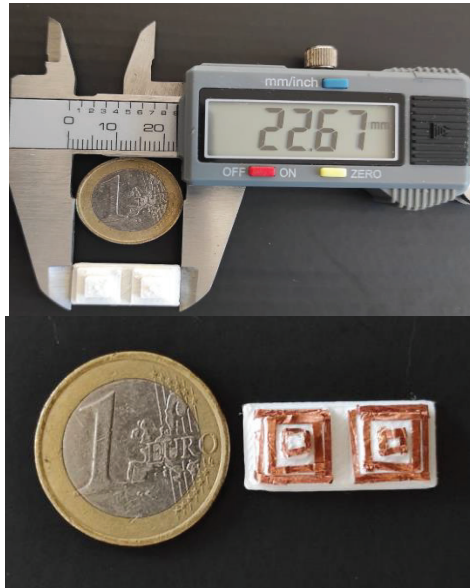
**Figure 6.** Fabricated FSS units using 3D printer (PLA material) with design parameters of [L1=2, L2=6, L3=8, $\varepsilon_r$=2.5].
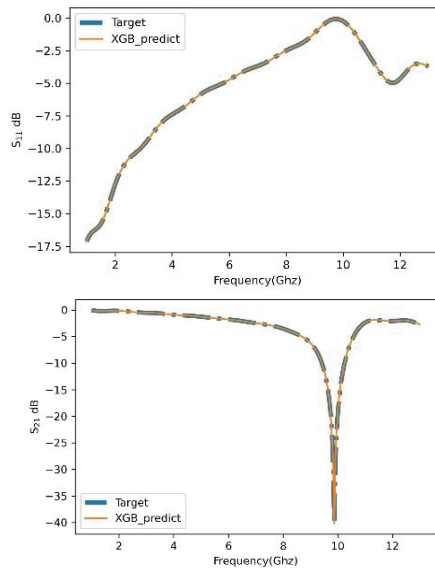


**Figure 7.** Measured and predicted scattering characteristic of FSS design.

## 5. Conclusion

In this work, potential of ensemble learning algorithms for design of microwave elements is taken under the study. As for a study case example, A Frequency Selective Surface (FSS) element is taken under the consideration. A data set for training and validation of Ensemble

algorithms is generated using 3D EM simulation tool and Latin Hyper Cube sampling method. The hyper-parameters of each ensemble model are determined optimally using grid search method. The performance of the models is compared with holdout data sets to check if the models fall in over fitting or not. Finally, the performance results of best model XGBM, are compared with experimental results of a 3D printed FSS design. As a result, the proposed Ensemble based surrogate modelling is an efficient, accurate and viable solution for modelling of data driven surrogate models to be used in design optimization processes.

## 6. Acknowledgment

# 7. References

[1] J. Zhang, F. Feng, W. Na, S. Yan and Q. Zhang, "Parallel space-mapping based yield-driven EM optimization incorporating trust region algorithm and polynomial chaos expansion," *IEEE Access,* vol. 7, p. 143673–143683, 2019.

[2] S. Koziel, P. Mahouti, N. Calik, M. A. Belen and S. Szczepanski, "Improved Modeling of Microwave Structures Using Performance-Driven Fully-Connected Regression Surrogate," IEEE Access, vol. 9, p. 71470–71481, 2021.

[3] S. Koziel and A. Pietrenko-Dabrowska, "Reduced-cost electromagnetic-driven optimisation of antenna structures by means of trust-region gradient-search with sparse Jacobian updates," IET Microwaves, Antennas & Propagation, vol. 13, p. 1646–1652, 2019.

[4] A. Ciccazzo, G. Di Pillo and V. Latorre, "A SVM surrogate model-based method for parametric yield optimization," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems,* vol. 35, p. 1224–1228, 2015.

[5] A. S. Abd El-Hameed, M. G. Wahab, A. Elboushi and M. S. Elpeltagy, "Miniaturized triple band-notched quasi-self complementary fractal antenna with improved characteristics for UWB applications," *AEU-International Journal of Electronics and Communications,* vol. 108, p. 163–171, 2019.

[6] J. Wang, X.-S. Yang and B.-Z. Wang, "Efficient gradient-based optimisation of pixel antenna with large-scale connections," *IET Microwaves, Antennas & Propagation,* vol. 12, p. 385–389, 2018.

[7] S. Kozieł and A. Pietrenko-Dąbrowska, "Reduced-cost design closure of antennas by means of gradient search with restricted sensitivity update," *Metrology and Measurement Systems,* p. 595–605, 2019.

[8] S. Mishra, R. N. Yadav and R. P. Singh, "Directivity estimations for short dipole antenna arrays using radial basis function neural networks," *IEEE Antennas and Wireless Propagation Letters,* vol. 14, p. 1219–1222, 2015.

[9] A.-K. S. O. Hassan, A. S. Etman and E. A. Soliman, "Optimization of a novel nano antenna with two radiation modes using kriging surrogate models," *IEEE Photonics Journal,* vol. 10, p. 1–17, 2018.

[10] J. P. Jacobs, "Efficient resonant frequency modeling for dual-band microstrip antennas by Gaussian process regression," *IEEE Antennas and Wireless Propagation Letters,* vol. 14, p. 337–341, 2014.

[11] J. E. Rayas-Sánchez, "EM-based optimization of microwave circuits using artificial neural networks: The state-of-the-art," *IEEE Transactions on Microwave Theory and Techniques,* vol. 52, p. 420–435, 2004.

[12] F. Güneş, P. Mahouti, S. Demirel, M. A. Belen and A. Uluslu, "Cost-effective GRNN-based modeling of microwave transistors with a reduced

number of measurements," *International journal of numerical modelling: electronic networks, devices and fields,* vol. 30, p. e2089, 2017.

[13] D. R. Prado, J. A. López-Fernández, M. Arrebola and G. Goussetis, "Support vector regression to accelerate design and crosspolar optimization of shaped-beam reflectarray antennas for space applications," *IEEE Transactions on Antennas and Propagation,* vol. 67, p. 1659–1668, 2018.

[14] A. Gehani, P. Agnihotri and D. A. Pujara, "Analysis and synthesis of multiband Sierpinski carpet fractal antenna using hybrid neuro-fuzzy model," *Progress in electromagnetics research letters,* vol. 68, p. 59–65, 2017.

[15] D. Opitz and R. Maclin, "Popular ensemble methods: An empirical study," *Journal of artificial intelligence research,* vol. 11, p. 169–198, 1999.

[16] R. Polikar, "Ensemble based systems in decision making," *IEEE Circuits and systems magazine,* vol. 6, p. 21–45, 2006.

[17] L. Rokach, "Ensemble-based classifiers," *Artificial intelligence review,* vol. 33, p. 1–39, 2010.

[18] Z.-H. Zhou, Ensemble methods: foundations and algorithms, Chapman and Hall/CRC, 2019.

[19] B. Kégl, "The return of AdaBoost. MH: multi-class Hamming trees," *arXiv preprint arXiv:1312.6086,* 2013.

[20] S. J. b. c. w. com, *Joglekar, S,* Retrieved, 2016.

[21] J. Friedman, T. Hastie, R. Tibshirani and others, The elements of statistical learning, vol. 1, Springer series in statistics New York, 2001.

[22] Y. Freund, R. Schapire and N. Abe, "A short introduction to boosting," *Journal-Japanese Society For Artificial Intelligence,* vol. 14, p. 1612, 1999.

[23] S. M. Piryonesi and T. E. El-Diraby, "Data analytics in asset management: Cost-effective prediction of the pavement condition index," *Journal of Infrastructure Systems,* vol. 26, p. 04019036, 2020.

[24] T. Hastie, R. Tibshirani and J. Friedman, "Boosting and additive trees," in *The elements of statistical learning*, Springer, 2009, p. 337–387.

[25] L. Breiman, "Arcing the edge," Technical Report 486, Statistics Department, University of California, 1997.

[26] J. H. Friedman, "Stochastic gradient boosting," *Computational statistics & data analysis,* vol. 38, p. 367–378, 2002.

[27] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of statistics,* pp. 1189--1232, 2001.

[28] L. Mason, J. Baxter, P. Bartlett and M. Frean, "Boosting algorithms as gradient descent in function space," in *Proc. NIPS*, 1999.

[29] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016.

[30] D. Nielsen, "Tree boosting with xgboost-why does xgboost win" every" machine learning competition?," 2016.

[31] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye and T.-Y. Liu, "Lightgbm: A highly efficient gradient boosting decision tree," *Advances in neural information processing systems,* vol. 30, p. 3146–3154, 2017.

[32] N. Calik, M. A. Belen, P. Mahouti and S. Koziel, "Accurate modeling of frequency selective surfaces using fully-connected regression model with automated architecture determination and parameter selection based on Bayesian optimization," *IEEE Access,* vol. 9, p. 38396–38410, 2021.

[33] N. Calik, M. A. Belen and P. Mahouti, "Deep learning base modified MLP model for precise scattering parameter prediction of capacitive feed antenna," *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields,* vol. 33, p. e2682, 2020.

[34] S. Madeh Piryonesi and T. E. El-Diraby, "Using Machine Learning to Examine Impact of Type of Performance Indicator on Flexible Pavement Deterioration Modeling," *Journal of Infrastructure Systems,* vol. 27, p. 04021005, 2021.

[35] S. Joglekar, *Sachin Joglekar's blog. codesachin. wordpress. com,* Retrieved, 2016.

# Chapter 8

## COMPOSTING OF ORGANIC WASTES IN TURKEY AND BIOAEROSOL EMISSIONS FROM COMPOST PLANTS

*Nesli AYDIN[1]*

1  Assist. Prof. Dr. Nesli AYDIN, Tekirdağ Namık Kemal University, Faculty of Engineering, Department of Environmental Engineering, Çorlu, Tekirdağ, Turkey, naydin@nku.edu.tr, https://orcid.org/0000-0002-7561-4280.

## INTRODUCTION

Composting is a process based on the biological degradation of organic wastes under aerobic conditions. Compost, which includes diverse range of nutrients, is produced at the end of this process. It could be used as fertilizer to improve soil quality. Agricultural wastes, which constitute a significant part of organic materials, are produced as a result of various agricultural operations. Many factors affect the composition of them, such as the socio-economic characteristics of the society where the production takes place, nutritional habits, traditions, geographical conditions, climate, distance from the industrial facility and education, etc. With the increase in the world population, it is expected the amount of agricultural production to improve.

During the composting process, there are some important factors affecting the quality (such as pH, temperature and moisture content). The different stages at composting plants, such as screening, turning the pile, sieving, etc., can be the source of pollution involving the generation of odors, noise, dust, and bioaerosols. The emission of bioaerosols during composting activities could dramatically increase the number of microorganisms in the atmosphere. The implications of bioaerosols are crucial for composting plants installed in the open air as these emissions are discharged directly into the atmosphere and conveyed neighborhood surroundings without any pre-treatment. So far bioaerosols have been related to various adverse health effects on the respiratory system causing such as pulmonary allergic disease, chronic asthma and cystic fibrosis, etc. The wider application of composting processes as a sustainable waste management alternative has caused societies to raise concerns about potential health impacts.

Disposal of agricultural and farming wastes poses a serious problem for producers. These wastes either cause environmental pollution by being treated in incinerators or disposed of randomly on the ground. Since these practices are harmful to the environment and potential benefits of these wastes are not utilized, their use in agriculture with composting, which is an alternative disposal method, has gained importance in recent years and has been the subject of many scientific types of research.

Soil protection and development are among the most important parameters for high productivity and sustainability in agricultural activities. It is known that a large part of the soil in Turkey contains very little organic matter and this amount decreases over time. This reduces the quality of any soil physically, chemically and biologically. For sustainable agricultural quality, the application of organic wastes to the soil is a common practice that has a very important effect on the organic

content and nutrients of the soil. Organic fertilizer applications provide a significant increase in soil organic matter and nitrogen content (Özenç ve Şenlikoğlu, 2017). In addition, the application of organic wastes to the soil affects the microbiological activities, structure and air-water balance of the soils positively (Tarakçıoğlu and Yurtseven, 2019).

However, it is known that there are several types of particules being released from different stages of composting processes. For instance, bioaerosols, such as dust, fungi, endotoxin, etc., are emitted from composting plants and they are related with adverse health outcomes (Ciplak and Barton, 2013). For this reason, this study aims to provide an extensive evaluation of organic wastes, particularly produced from agricultural and farming activities by taking into account the different types of composting along with the factors affecting this process. In this study, the possibility of reducing the negative effects of organic wastes from farming (animal waste) and agriculture activities (agricultural waste) on the environment was also evaluated. Composting potential of these wastes was discussed to increase the productivity of any soil with low organic matter content along with bioaerosol emissions.

## AGRICULTURAL ACTIVITIES AND AGRICULTURAL WASTE PRODUCTION

Greenhouse cultivation in Turkey, which started in the province of Antalya in the 1940s, takes an important part in agricultural activities across the country. It is carried out intensively in the Mediterranean and Aegean Regions today as it provides agricultural production that contributes to the economy of Turkey. The total agricultural area in Turkey has a share of 36% (27 575 000 ha) and 68% of this is used for growing field crops, 13% of this is used for growing horticultural crops, whereas 19% is fallowed (Sevgican et al., 2000). While the cereal crop (wheat, coarse grains, rice) production in the world was 2.612.095 million tons in 2017, the total grain production in Turkey was 35.3 million tons (FAO, 2018). Table 1 shows the production of cereal crops in Turkey. Considering the size of production and agricultural areas around the world, it is seen that Turkey's agricultural production is higher than many countries. It is also known that 96% of greenhouse activities, which increase almost every year in Turkey, are mainly based on vegetables, 3% ornamental plants, and 1% fruit growing (Çolakoğlu, 2018).

*Table 1. Production of Cereal Crops in Turkey (FAO, 2019)*

| Nutrients (million tonnes) | Years | | |
|---|---|---|---|
| | 2016/17 | 2017/18 | 2018/19 |
| Wheat | 13.76 | 13.73 | 13.41 |
| Rice (milled) | 0.55 | 0.54 | 0.56 |
| Coarse grains | 20.6 | 21.5 | 20.0 |

## LIVESTOCK ACTIVITIES AND ANIMAL WASTE PRODUCTION

Animal wastes contain high amounts of pathogens and their use in agriculture without any pre-treatment cause environmental health effects. In addition, the irregular storage of animal waste causes serious odor problems. As a result of the anaerobic reaction in animal manure, odorous gases such as hydrogen sulfide are formed. Therefore, the waste needs to be stabilized to remove its hazardous contents and so to prevent odor problem caused by pathogens. Figure 1 gives examples of animal wastes and Table 2 shows the amount of animal waste based on the number of animals in Turkey.



*Figure 1. Animal Waste*

*Table 2. Production of Animal Waste in Turkey (Turkstat, 2017; Çolakoğlu, 2018)*

|  | Number of Animals (million) | Waste (manure) (million tons) |
|---|---|---|
| Poultry | 15.9 | 57.4 |
| Cattle animals | 44.3 | 31.1 |
| Sheep and goats | 1 233.6 | 27.1 |
| Total | 1 293.8 | 115.6 |

## RECYCLING OF ORGANIC WASTES BY COMPOSTING

It takes time for any fertilizer to give effective results when applied to the soil where precipitation and humidity are low and the decomposition time is longer. Compared to the application of fertilizer, compost is a more effective material to increase nutritional value of the soil in a considerably shorter term (Akpınar et al., 2010).

Composting is known as a controlled decomposition process of wastes with the formation of organic materials which can be applied to soil as a conditioner or fertilizer to improve its quality (Taiwo, 2011). In other words, it is a biological process that transforms heterogeneous organic wastes into humus-like materials (Figure 2) by a mixed microbial population under controlled conditions of optimum humidity, temperature and ventilation (Sarangi and Lama, 2013). This process is a widely accepted practice and it is one of the most effective means of recycling organic wastes for agricultural reuse (Hoyos et al., 2002).



*Figure 2. Compost Production*

Çataltaş (2013) indicates that open dumping of animal and vegetable wastes has negative effects on the environment and public health. In this study, it was emphasized that while the accumulation of wastes such as animal manure in rural areas on the land provides a suitable environment for the survival and reproduction of different pathogens, it also causes insects and mosquitoes. This adversely affects the health of society and the environment.

Eskicioğlu (2013) determined that the rapid and irreversible use of resources in agriculture, industry and technology, which develops in parallel with the increasing population. It was stated that recycling, which is the most important component of waste management, is necessary for the protection of the environment. It was underlined that the most foreign currency expenditure in Turkey after oil was made for mineral fertilizers and it was predicted that the yield could be increased by 50% with the right fertilization.

Agricultural wastes include straw, stem, leaf, husk, bark, seed and pruning waste, etc. To prevent environmental pollution, it has become widespread to use compost produced from vegetable wastes in agriculture.

There are many advantages of composting. Some of them are listed below:

• It improves soil structure and ensures easy processing,

• It increases the void volume in the soil and provides aeration of the soil,

• It provides nutrients when plants need,

• It reduces the effect of harmful substances in the soil,

• It regulates the pH balance of the soil,

• It accelerates the growth of the plants and strengthens them,

• It provides a 40%-60% reduction in the waste mass to be stored (Akpınar et al., 2010)

**FACTORS AFFECTING COMPOSTING**

There are several factors affecting compost, such as C/N ratio, moisture, porosity, structure, consistency and particle size. Parameters such as pH, temperature and time particularly affect the quality of compost (EPA, 2021). The quality of compost often depends on the presence of undesirable components in the soil such as soluble salts and weeds. An ideal compost should have a pH of 6.5-8.0, a C/N ratio of 25/1-30/1, moisture of 50-60%, an oxygen content of >5%, a particle size of 0.32-1.27 cm, and a temperature of 54-60$^0$C (EPA, 2021).

### Temperature

An optimum temperature range is required for microorganisms to be effective. Optimum temperatures provide rapid composting and destroy pathogens. Temperature can rise as a result of microbial activity to at least 60° C (EPA, 2021; Wei et al., 2000). If the temperature does not go up, anaerobic conditions could take place.

### Oxygen Flow

Placing a series of pipes under the pile, turning the pile regularly or placing large agents such as chips of wood and papers in the pile could help to aerate the pile. Airflow in the pile allows decomposition of the organic waste and accelerates composting. The provision of too much air could dry out the compost pile.

### Moisture Content

Enough moisture should be provided for the microorganisms in a compost pile to survive. Moisture is the important factor that conveys substances in the pile and causes the nutrients in the waste to homogeneously disperse. Organic waste naturally includes some moisture. However, the moisture also comes because of rain/snowfall.

### Particle Size

To increase the surface area in the pile, grinding, chipping, and shredding could help. Finer particles help compost piles to become more homogeneous and hence improve pile insulation by keeping optimum temperatures.

### Balance of Nutrients

Composting necessitates a proper balance of some organic materials, such as food, grass, and manure, which include large amounts of nitrogen, and some other organic materials, such as dry leaves and branches, which inherits large amounts of carbon but little nitrogen (EPA, 2021).

Külcü and Çaylak (2015) studied compost production in Turkey. In this study, it was determined that although poultry farms produce compost in the Aegean and Marmara Regions locally, off-site compost plants are quite widespread in the Mediterranean Region. It was also revealed in the study that composting does not take an important part in waste management in the Black Sea and Eastern Anatolia Regions in Turkey (Figure 3).
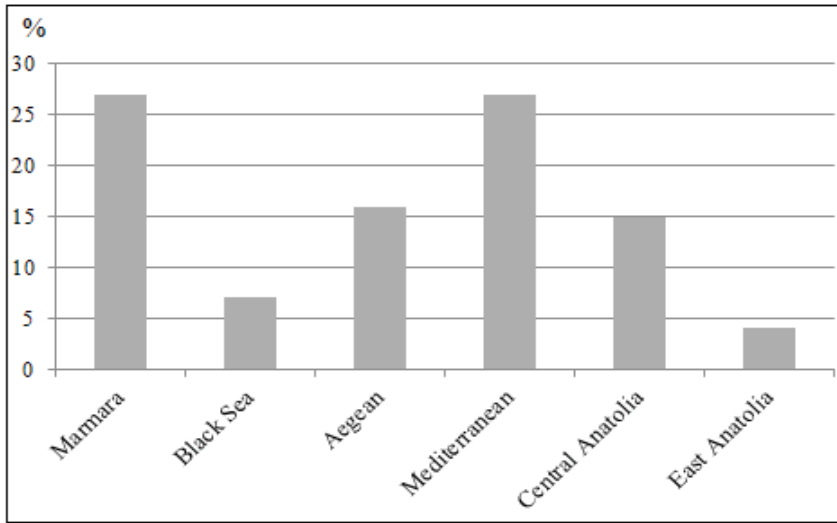
*Figure 3. Distribution of Compost Plants from Agricultural Wastes in Turkey*

If the amount of compost production is examined together with the size of animal husbandry, 236 265 tons of compost (201 445 tons from chicken manure, 31 820 tons from cattle manure, 3 000 tons from other wastes and 2 400 tons from agricultural wastes) is annually produced in Turkey (Figure 4) (Külcü and Çaylak, 2015).

Külcü (2016) determined that it is possible to produce 1 490 451 tons/ year of compost from 2 838 954 tons of organic waste (20% moisture content) in the province of Afyonkarahisar. Likewise, Karaca (2017) indicated that the highest amount of waste in the province of Antalya is produced from tomato cultivation (165.3 thousand tons). This is followed by pepper (27.35 thousand tons) and eggplant cultivation (10 thousand tons).
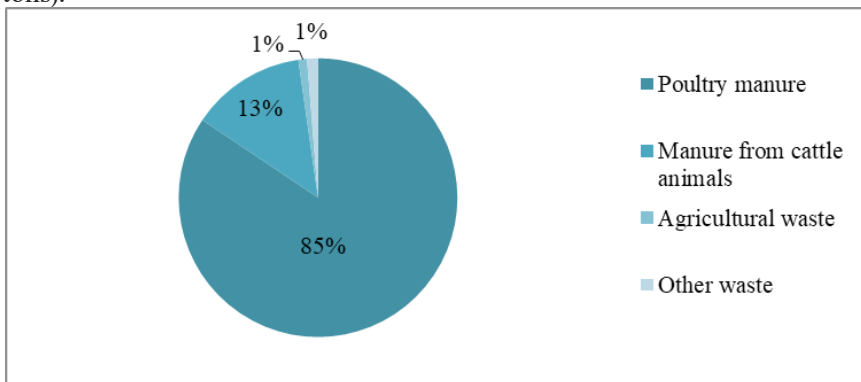


*Figure 4. Distribution of Compost Production in Turkey*

It is also known that the application of compost obtained from the cultivation of high-value vegetables such as tomatoes could be more economical and practical compared to the application of meadows or grasses as compost (Roe and Cornforth, 2000). It was revealed that a 10% yield increase is achieved when compost obtained from greenhouse waste is used instead of sawdust (Çerçioğlu, 2018).

There are different methods and technologies in composting, depending on the characteristics of the waste and the area where the compost is carried out. The commonly used composting methods are windrow composting, aerated static piles and in-vessel composting. While the process success and pasteurization effect are improved in closed systems, the investment and operating costs are quite high. However, closed systems are widely used in composting sewage sludge and municipal waste, where pasteurization effect and odor control are considerably important.

Approximately 3% of the agricultural compost enterprises operating in Turkey use aerated static piles composting, whereas 66% of them are windrow composting and 31% of them are in-vessel composting (Külcü and Çaylak, 2012).

**Windrow Composting**

Windrow composting is an aerobic process for biodegradable organic wastes (Figure 5). During the process, the heat that is produced as a result of biological activity destroys pathogens and a stabilized compost product is produced for use as a soil conditioner. The windrow technique is simple and accomplished easily with standard equipment (Kuhlman, 1990).

In this outdoor process, organic material is stocked into piles to be decomposed and only watering and mechanical turning of pile takes place regularly. This method is commonly used as it is a simple technique with low capital cost. It can also be used for large scale organic waste processing including sewage sludge.

One of the most important factors for the successful operation of windrow composting is keeping oxygen level as desired to maintain decomposition. For this, compost piles should be mixed properly. Temperature is another significant issue which must be high enough to destroy pathogens in the waste without killing the microorganisms.

*Figure 5. Windrow Composting*

## Aerated Static Pile Composting

This sort of composting can be performed both indoors and outdoors (Figure 6). For a quick process, it is more effective to have a uniform incoming waste feed. The typical pile in aerated static pile composting should be kept 3-4 meters high, 6-10 meters wide, and 30- 40 meters long. The pipe system is built under each compost pile and the air is supplied through the pipes. The aeration pipes are usually sheltered with a layer of wood waste for uniform aeration. The temperature of the compost pile requires to be checked regularly to avoid excess heat but to meet the temperature requirements of piles.

The most significant elements to be controlled are oxygen, temperature, and moisture to maintain the activity of microorganisms. Optimum oxygen levels should be provided for the compost system to keep aerobic activity. Composting materials (wastes) should be moist but not wet; otherwise, odor could be a problem.



*Figure 6. Aerated Static Pile Composting*

### In-vessel Composting

It is the process in which compost in vessel channels should have a controlled aeration system (Figure 7). The aeration is provided by aerating machines such as fans providing airflow. The most important advantage of this advanced technology is its requirement for low labor.

Unlike the windrow method, in-vessel systems do not require large space for composting large amounts of wastes. Hence, they provide a solution for the cases where the space is limited but the incoming waste is huge in amount. The produced biogas can be burned for energy to meet the energy demand of the facility. Similar to other methods, such as oxygen and temperature are very important, the composition of incoming waste is less critical for in-vessel systems compared to windrow and aerated static pile systems. This flexibility allows in-vessel compost technology to treat different mixes of wastes. The process takes few weeks to complete, but the compost requires almost a month for the microbial activity to take place and the pile to cool down.



*Figure 7. In-vessel Composting*

## BIOAEROSOLS FROM COMPOSTING

Composting is a microbiological process and during mechanical agitation of composting material, biological agents become airborne and bioaerosols occur. A huge amount of bioaerosols occur in this process

as most of the composting is conducted in open air windrow systems especially in Europe, for example 96% of the organic waste was composted in open-air windrow systems in the UK (The UK Environment Agency, 2002).

The bioaerosol occurrence is a significant issue in composting because it has adverse health effects. Bioaerosols released from composting consist of a range of micro-organisms (bacteria, fungi) and organic constituents of microbial and plant origin. Specifically, the conditions caused by Aspergillus fumigatus are one of the most reported cases as a result of exposure to airborne particles released from composting facilities so far (Sanchez-Monedero and Stentiford, 2003).

## HEALTH AND SAFETY CONCERNS REGARDING COMPOSTING

Bioaerosols are a problematic outcome in composting as they have a potential adverse implication on public health. Importantly, the concerns regarding occupational health and safety are widespread for composting. However, they usually include inhalation of aerosols, and also the potential for bioaerosols to be carried beyond a facility perimeter (Prasad et al., 2004). Bioaerosols include emissions of dust, Aspergillus fumigatus, total fungi and endotoxins.

### Dust

Dust is defined as small solid particles (below 75 μm in diameter), which could remain suspended in the air by the International Standardization Organization (ISO 4225, 1995). Dust generated during composting is technically not a bioaerosol, but it could have microbial agents (Figure 8). The dust produced from a composting facility can include bacteria, insects or fungi.



*Figure 8. Dust Generation from Compost Plants*

Dust at composting plants could be generated during transportation, mixing, sieving, processing and storing of wastes. The reason for dust emission from composting plants is usually the insufficient moisture content of the composting material. Table 3 shows dust concentrations from a variety of activities at composting plants.

*Table 3. Dust Concentrations Recorded at Various Composting Sites*

| Plant | Total Dust Concentration (mg/m³) | Reference |
|---|---|---|
| USA | $3.9 \times 10^{-1}$ - $1.8 \times 10^{0}$ | Hryhorczuk et al. (2001) |
| Granada, Spain | 3.1 | Frostier et al. (2008) |
| USA | $5 \times 10^{-1}$ - $2.47 \times 10^{2}$ | Epstein et al. (2001) |
| North-western Germany | 0.37 | Kampen et al. (2014) |

## Aspergillus Fumigatus

Aspergillus fumigatus is a highly ubiquitous fungus which is related with soil, crop plants, straw, grass, bird droppings, and compost. It is also found on bathroom walls where the conditions for molds to grow could appear.

Aspergillus fumigatus has been proved as an important bioaerosol component in the literature as it is widely monitored. It has significant potential health implications (such as pulmonary allergic disease, chronic asthma and cystic fibrosis) and its emissions were measured at different stages of composting processes (Table 4 - CFU: A colony-forming unit). Figure 9 shows the appearance of Aspergillus fumigatus.

*Table 4. Aspergillus Fumigatus Concentrations Recorded at Various Composting Sites*

| Place | Aspergillus fumigatus concentration (CFU/m³) | Reference |
|---|---|---|
| Sweden | $10^{4}$ (screening/sieving) | Clark et al. (1983) |
| Isfahan, Iran | 50 ( background/upwind/control) <br> 40 (screening/sieving) | Nikaeen et al. (2009) |
| UK | $10^{2}$ (background/upwind/control) <br> $9 \times 10^{2}$-$10^{4}$ (downwind) | Sanchez-Monedero et al. (2005) |
| England and Wales | $6 \times 10^{4}$-$10^{5}$ (shredding) <br> $10^{6}$- $6 \times 10^{6}$ (turning) <br> $55 \times 10^{3}$-$10^{5}$ (screening/sieving) | Taha et al. (2007) |
| Central England | $9 \times 10^{3}$ (background/upwind/control) | Deacon et al. (2009) |

| | | |
|---|---|---|
| France | $5x10^4$ (shredding)<br>$10^4$ (turning)<br>$9x10^4$ (screening/sieving) | Persoons et al. (2010) |
| France | 10 (background/upwind/control)<br>$10^3$ (shredding)<br>$15x10^3$ (turning)<br>$10^4$ (screening/sieving) | Schlosser et al. (2009) |
| Germany | $10^2$-$10^3$ (background/upwind/control)<br>$10^3$-$5x10^6$ (shredding) | Kampen et al. (2014) |
| Scotland and Ireland | $9x10^2$-$11x10^2$ (background/upwind/control) | Drew et al. (2007) |
| Austria | $10^3$-$5x10^3$ (background/upwind/control) | Reinthaler et al. (1999) |
| UK | $5x10^2$-$10^3$ (background/upwind/control)<br>$3x10^2$ (turning) | Stagg et al. (2010) |
| Canada | 60 (background/upwind/control)<br>$10^4$ (turning)<br>60 (screening/sieving) | VanderWerf et al. (1996) |
| USA | $5x10^3$-$10^4$ (turning)<br>$10^3$-$5x10^3$ (screening/sieving) | Epstein et al. (2001) |
| Austria | 10 (background/upwind/control)<br>$10^5$-$5x10^5$ (Indoor composting areas)<br>$9x10^2$-$2x10^3$ (downwind) | Reinthaler et al. (2004) |
| UK | $10^4$ $x10^5$ (Indoor composting areas)<br>$10^4$ (shredding)<br>$10^5$ (turning)<br>$10^4$ (screening/sieving) | Sanchez-Monedero and Stentiford (2003) |
| South East England | $10^4$ (background/upwind/control)<br>$5x10^7$ (turning)<br>$5x10^5$ (screening/sieving)<br>$5x10^4$ (downwind) | Taha et al. (2006) |
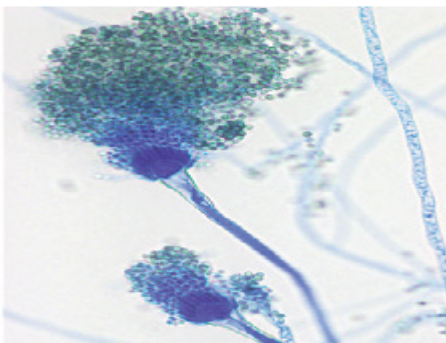


*Figure 9. Aspergillus Fumigatus (URL-1; URL-1)*

## CONCLUSIONS

In today's world, where efficient use of resources and recycling are of great importance, composting of organic wastes is an effective method to evaluate organic wastes for a sustainable environment. The necessity of increasing the production and obtaining more products from the unit area to meet the nutritional requirements of the increasing population has made the composting of organic wastes even more important.

Composting has many benefits such as providing erosion control, easier processing of hard-to-work soils and better use of nutrients. It also destroys the sources of pathogens and weeds in the waste. In sustainable waste management, composting provides both organic matter and nutrient content for the soil, hence less commercial fertilizer is needed.

The climate and seasonal changes do not dramatically affect onsite composting. There are some factors listed below to take into consideration when composting (EPA, 2021; Sanchez-Monedero and Stentiford, 2003; Sanchez-Monedero et al., 2005)

• *Small adjustments should be made when seasonal changes as the warmer or rainy season approach.*

• *Extra care should be taken for organic waste containing food to prevent odors or the attraction of insects.*

• *Education is an important key to success. Local communities (such as municipalities, city agencies, etc.) could help improving the dissemination of knowledge and experience.*

• *Aeration could accelerate the composting process to between three to six months.*

• *Measures should be taken to reduce or completely prevent bioaerosol emissions in compost facilities.*

# REFERENCES

Akpınar, Ç., Ortaş, İ., Demirbaş, A., Şimşek, M., Yüksel, A. (2010). The Effects of Different Compost Processes on Two Different Growing Environments on Clover and Onion Plant's Growing Plant Nutrient Elements Uptake And Mycorrizae Infection Quality, Journal of Agriculture Faculty of Ege University, 659-665

Clark, C.S., Rylander, R., Larsson, L. (1983). Levels of gram-negative bacteria, Aspergillus fumigatus, dust, and endotoxin at compost plants, Appl. Environ. Microbiol., 45, 1501–1505

Çataltaş, A. (2013). Composting of Animal Manure (Master Thesis), Uludağ University, Bursa

Çerçioğlu, M. (2018). Evaluation of Greenhouse Wastes as Compost in Sustainable Waste Management, Journal of Agricultural Faculty of Bursa, 33(1), 167-177

Çıplak, N. and Barton, J.R. (2013). A System Dynamics Approach for the Determination of Adverse Health Impacts of Healthcare Waste Incinerators and Landfill Sites on Employees, Environmental Management and Sustainable Development, 2, 2

Çolakoğlu, B. (2018), Farmers' Approaches on Alternative Uses of Agricultural (Master Thesis), Namık Kemal University, Tekirdağ

Deacon, L., Pankhurst, L., Liu, J., Drew, G.H., Hayes, E.T., Jackson, S., Longhurst, J., Longhurst, P., Pollard, S., Tyrrel, S. (2009). Endotoxin emissions from commercial composting activities, Environ. Health, S9, 8

Drew, G.H., Tamer, V.A., Taha, M.P.M., Tyrrel, S.F., Longhurst, P.J., Pollard, S.J.T. (2007). Project UKPIR12. Edinburgh, Scotland: Scotland and Northern Ireland Forum for Environmental Research and Scotland Environment Protection Agency, Bioaerosol and Odour Monitoring from Three Composting Sites.

EPA, The United States Environmental Protection Agency, 2021, Types of Composting and Understanding the Process, https://www.epa.gov/sustainable-management-food/types-composting-and-understanding-process (Access Date. 08.07.2021).

Epstein, E., Wu, N., Youngberg, C., Crouteau, G. (2001). Dust and Bioaerosols at a Biosolids Composting Facility, Compost Science & Utilization, 9(3), 250-255

Eskicioğlu, A.V. (2013). Vegetable Waste Compost Fertilizer Production System Design (Master Thesis), Namık Kemal University, Tekirdağ

FAO (2018). Food and Agriculture Organization of the United Nations. http://statistics.amis-outlook.org/data/index.html (Access Date. 07.07.2021).

FAO (2019). Food and Agriculture Organization of the United Nations. http://statistics.amis-outlook.org/data/index.html#DOWNLOAD (Access Date. 07.07.2021).

Forestier, D., Thoumelin, P., Efremenko, B., Arfi, C. Edited by Zamorano, M., Brebbia, C.A., Kungolos, A., Popov, V. and Itoh, H. (2008). Occupational Bioaerosol Exposure in Composting Plants. Book Series WIT Transactions on Ecology and the Environment, Waste Management and The Environment IV, doi: 10.2495/WM080921

Hoyos, S.E.G., Juarez,, J.V., Ramonet, C.A., Lopez, J.G., Rios, A.A., Uribe, E.G. (2002). Aer-obic thermophilic composting of waste sludge from gelatin-grenetine industry, Resour. Conserv. Recycl., 34, 161–162

Hryhorczuk, D., Curtis, L., Scheff, P., Chung, J., Rizzo, M., Lewis, C., Keys, N., Moomey, M. (2001). Bioaerosol emissions from a suburban yard waste composting facility, Annals of Agricultural and Environmental Medicine, 8, 177-185

ISO 4225 (1995). Air Quality-Particle Size Fraction Definitions for Health-related Sampling ISO Standard 7708. International Organisation for Standardisation (ISO), Geneva

Kampen, V.V., Sander, I., Liebers, V., Deckert, A., Neumann, H.D., Buxtrup, M., Willer, E., Felten, C., Jäckel, U., Klug, K., Brüning, T., Raulf, M., Bünger, J. (2014). Concentration of Bioaerosols in Composting Plants Using Different Quantification Methods, The Annals of Occupational Hygiene, 58(6), 693–706. doi: 10.1093/annhyg/meu026

Karaca, C. (2017). Antalya'da seracılık biyokütle artıklarının potansiyelinin haritalanması ve enerji üretim amacıyla değerlendirilmesi, Mediterranean Agricultural Sciences, 30(1), 21-25

Kuhlman, L.R. (1990). Windrow composting of agricultural and municipal wastes, Resources, Conservation and Recycling, 4, 1–2, 151-160

Külcü, R. (2016). Evaluation of Agricultural Biomass Potentials of Afyonkarahisar City, Academia Journal of Engineering and Applied Sciences, 1(2, 1-9

Külcü, R. and Çaylak, R. (2012, 17-20 October). Türkiye'de tarımsal atıklardan kompost üretim sektörünün gelişimi, 4. Ulusal Katı Atık Kongresi, Antalya

Külcü, R. and Çaylak, R. (2015). Development of Agricultural Compost Production Sector in Turkey, Academia Journal of Engineering and Applied Sciences, 1(1), 20-25

Nikaeen, M., Mirhendi, H., Hatamzadeh, M., Ghorbani E. (2009). Bioaerosol emissions from composting facilities as a potential health risk for composting workers. In: Bulucea C. A., Mladenov V., Pop E., Leba M., editors; Mastorakis N., editor. In Recent advances in environment, ecosystems and development, 27–29.

Özenç, D.B. and Şenlikoğlu, G. (2017). Effects of compost and nitrogen fertilizer on growth of spinach (Spinacia oleracea L.). Akademik Ziraat Dergisi, 6, 227-234

Persoons, R., Parat, S., Stoklov, M., Perdrix, A., Maitre, A. (2010). Critical working tasks and determinants of exposure to bioaerosols and MVOC at composting facilities, Int. J. Hyg. Environ. Health, 213, 338–347

Prasad, M., van der Werf, P., Brinkmann, A. (2004). Bioaerosols and Composting-A Literature Evaluation, Composting Association of Ireland

Reinthaler, F. F., Haas, D., Feierl, G., Schlacher, R., Pichler-Semmelrock, F. P., Kock M., Wust, G., Feenstra, O., Marth E. (1999). Comparative investigations of airborne culturable microorganisms in selected waste treatment facilities and in neighbouring residential areas, Zentralbl. Hyg. Umweltmed, 202, 1–17.

Reinthaler, F. F., Wust, G., Haas, D., Feierl, G., Ruckenbauer, G., Marth, E. (2004). Simple emission-reducing measures in an open biological waste treatment plant, Aerobiologia., 20, 83–88.

Roe, N.E. and Cornforth, G.C. (2000). Effects of dairy lot scrapings and composted dairy manure on growth, yield, and profit potential of double cropped vegetables, Compost Science and Utilization, 8(4), 320-327

Sanchez-Monedero, M. A. and Stentiford, E. I. (2003). Generation and dispersion of airborne microorganisms from composting facilities, Process Safety Environ. Protect. 81, 166–170.

Sanchez-Monedero, M. A., Stentiford, E. I., Urpilainen, S T. (2005). Bioaerosol generation at large-scale green waste composting plants, J. Air Waste Manage. Assoc., 55, 612–618

Sarangi, S.K. and Lama, T.D. (2013). Straw composting using earthworm (Eudrilus euge-niae) and fungal inoculant (Trichoderma viridae) and its utilization in rice (Oryzasativa)-groundnut (Arachis hypogaea) cropping system, Indian J. Agric. Sci. 83(4), 420

Schlosser, O., Huyard, A., Cartnick, K., Yañez, A., Catalán, V., Quang, Z. D. (2009). Bioaerosol in composting facilities: Occupational health risk assessment, Water Environ. Res., 81, 866–877

Sevgican, A., Tuzel, Y., Gul, A., Eltez, R.Z. (2000). Türkiye'de Örtüalti Yetiştiriciliği, Türkiye Ziraat Müh. V. Teknik Kongresi, Cilt II, 679-707

Stagg, S., Bowry, A., Kelsey, A., & Crook, B. (2015). Bioaerosol emissions from waste composting and the potential for workers' exposure. Health and Safety Laboratory. http://www.hse.gov.uk/research/rrpdf/rr786.pdf (Access Date. 27.09.2021)

Taha, M.P.M., Drew, G.H., Longhurst, P.J., Smith, R., Pollard, S.J.T. (2006). Bioaerosol releases from compost facilities: Evaluating passive and active

source terms at a green waste facility for improved risk assessments, Atmos. Environ., 40, 1159–1169

Taha, M.P.M., Drew, G.H., Tamer, A., Hewings, G., Jordinson, G.M., Longhurst, P.J., Pollard, S.J.T. (2007). Improving bioaerosol exposure assessments of composting facilities—Comparative modelling of emissions from different compost ages and processing activities, Atmos Environ., 41, 4504–4519.

Taiwo, A.M. (2011). Composting as A Sustainable Waste Management Technique in Developing Countries, Journal of Environmental Science and Technology, 4(2). doi: 10.3923/jest.2011.93.102

Tarakçıoğlu, C. and Yurtseven, B. (2019). Effect of Phosphate Rock and Different Organic Fertilizer Applications on Growth And Some Nutrient Contents of Maize Plant (Master Thesis), Ordu University, Ordu

The UK Environment Agency. (2002). Processes and Plant for Waste Composting and other Aerobic Treatment R&D Technical Report P1-311/TR. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/290347/sp1-311-tr-e-e.pdf (Access Date. 28.03.2021)

Turkstat (Turkish statistical Institute), 2017, Agricultural Waste Statistics, https://www.tuik.gov.tr/Home/Index (Access Date. 28.09.2021)

URL-1, Planet Natural Research Center (2021). https://www.planetnatural.com/composting-101/compost-concerns/pathogens/ (Access Date. 07.08.2021).

URL-2, Public Health Expertise and Reference Centre (2021). https://www.inspq.qc.ca/en/moulds/fact-sheets/aspergillus-fumigatus (Access Date. 27.07.2021)

Vander Werf, P. (1996). Bioaerosols at a Canadian composting facility, Biocycle., 37, 78–83.

Wei, Y.S., Fan, Y.B., Wang, M.J., Wang, J.S. (2000). Composting and Compost Application in China, Resources, Conservation and Recycling, 30, 277 – 300

# Chapter 9

## A PRIVATE P2P COLLABORATIVE FILTERING SCHEME WITH CONDENSED VECTORS

*Murat OKKALIOGLU[1]*

1  Yalova University, Faculty of Engineering, Computer Engineering Department, Yalova, Turkey, ORCID ID: 0000-0002-2137-5253, murat.okkalioglu@yalova.edu.tr

# 1. Introduction

Recommendation systems are very popular algorithms that help people decide in on-line avenues. Many popular e-commerce companies such as Amazon, Netflix, and Spotify utilize such approaches to help their customers find the right content. Recommender systems utilize either explicit or implicit user input. Users can express their opinions by providing explicit ratings over a scale. On the other hand, implicit inputs such as browsing history can also be very useful for recommender systems. Recommender systems can be broadly categorized into two approaches: content-based and collaborative filtering. Content-based filtering can use content information of objects while collaborative filtering (CF) relies on user inputs (Bobadilla, Ortega, Hernando, & Gutiérrez, 2013).

The main idea in CF is that users whose preferences were similar in the past will make similar preferences in the future. Thus, CF algorithms usually try to find out neighbors for an active user, who is looking for a prediction. These algorithms usually utilize user ratings made either on a numeric or binary scale. Numeric ratings demonstrate how much a relevant user likes an item, and they can be discrete or continuous. On the other hand, a binary rating indicates whether an item is liked by a relevant user or not. In a traditional CF scenario, ratings collected from users are kept in a large $n \times m$ matrix, for $n$ different users and $m$ different items.

Sparsity is an important issue for CF systems because users are not expected to have an idea for a large set of items. Besides, CF systems rely on enough user participation. While providing any feedback, users might be concerned that their information will be used other than the initial purpose of data collection (Culnan & Amstrong, 1999). Even user data can be sold in case of bankruptcy (Canny 2002). Due to privacy concerns, some users might be reluctant to provide their explicit data (Zhang, Wang, & Jin, 2014). Even if they participate in a CF process, they might show resistance to provide their true opinions. Such a reluctance or resistance might hamper the accuracy of CF systems. CF systems rely on true, quality, and enough user ratings to offer reliable predictions because the more they know about its users, the more reliable relationships between users or items can be extracted. Therefore, users must be assured by CF systems with a protocol that their confidential information will not be disclosed. Privacy measures can help CF systems collect user ratings in these respects. Privacy-preserving collaborative filtering (PPCF) addresses such issues and aims to provide privacy for users without sacrificing prediction accuracy. At this point, privacy and accuracy goals contradict each other (Polat & Du, 2003). An increase in the level of

privacy will eventually cause a decrease in accuracy. Therefore, a balance between these two main parameters must be maintained.

## 1.1. Problem Statement

Privacy is usually maintained by perturbing the original user vectors. Perturbation methods usually add some noise to the original vector. In general, the level of privacy increases with the level of perturbation which inherently leads to a decline in prediction accuracy. Therefore, the primary motivation in the PPCF community is to balance these two metrics. Employing either a central server or distributed servers is highly frequent in the PPCF community (Polat & Du, 2005a,b, 2008; Yargic & Bilge, 2009; Kaleli & Polat, 2007; Guo, Luo, Dong, & Yang, 2019; Kaur, Kumar, & Batra, 2018; Badsha, Yi, & Khalil, 2016; Badsha, Yi, Khalil, & Bertino, 2017). In such systems, predictions are usually produced by utilizing data stored in a central server or distributed servers with two- or multi-party data holders. However, mobile devices can eliminate the need for a server, thanks to their computational power. Our specific motivation in this study is to address the privacy concerns of peers in peer-to-peer (P2P) CF networks. Therefore, the research problem in this paper is *to design a private CF scheme around a P2P, server-less, network that considers both active's and collaborating peers' privacy without sacrificing the prediction accuracy*. An active peer is the one who is looking for a prediction, while a collaborating peer is the one who is joining the prediction process.

## 1.2. Solution and Contribution

We propose a PPCF scheme that operates on a P2P network for numeric ratings. Our scheme allows peers to perform CF operations privately so that they can collaborate in the prediction process without privacy concerns. Our scheme handles both active peer (AP)'s, and collaborating peers' privacy. In terms of AP's privacy,

1. Our scheme lets an AP perturb his/her ratings and rated items by creating a condensed vector from the original one before requesting a prediction. With private condensed queries, other collaborating peers in the P2P prediction network are not able to derive AP's true ratings and rated items. Similarly, the AP is not able to do so.

2. The AP can manipulate his/her rating vector to annihilate the possibility of unrated item discovery.

On collaborating peers' side,

1.  Collaborating peers hide their similarity values from the AP by collaborating with other peers through a communication protocol to calculate interim values required for the prediction.

2.  Although creating condensed vectors prevents the collaborating peers from their item ratings being discovered by the AP, the information whether queried items are rated or not is still at risk. The scheme also considers such consequences and argues that all peers should join the prediction process to stop such a disclosure.

Throughout the paper, we discuss possible privacy disclosure situations and explain our measures to overcome such circumstances. In experiments with three benchmark data sets, our private P2P scheme demonstrates promising accuracy results while providing privacy for its peers.

## 1.3. Organization

The rest of the paper is organized as follows: Section 2 gives the related work in the PPCF literature considering different data distribution scenarios. Section 3 introduces a basic P2P CF scheme without privacy. This scheme will help one understand possible privacy problems. The private PPCF scheme will be given in Section 4 and the related privacy analysis is performed in Section 5. Section 6 lays out the experimental results and the last section gives the conclusion and possible future work.

## 2.  Related Work

PPCF schemes can be categorized into two groups in terms of how data is stored, server-based and decentralized (Casino, Patsakis, Puig, & Solanas, 2013). In server-based schemes, data can be kept in a central or distributed servers with two- or multi-parties. In traditional central server-based PPCF schemes, users perturb their rating vector before submitting it to a central server. The server is not able to know individual original rating vectors due to the perturbation. Such perturbation is called individual privacy (Bilge, Kaleli, Yakut, Gunes, & Polat, 2013). In distributed server-based PPCF schemes, two- or multi-data owners come together for recommendation purposes with privacy. The assumption in distributed PPCF is that each party holds original user rating vectors; however, they want to prevent their corporate data from being disclosed by other parties. Therefore, each party applies privacy measures while exchanging interim results required in the prediction process. Such type of privacy protection is called institutional privacy (Bilge et al., 2013). Other than server-based schemes, users can create a P2P network for CF purposes with privacy. Contrary to server-based schemes, each user (peer) holds their rating vector and participates in the prediction process

collaboratively in a P2P network. Similar to distributed server-based schemes, prediction queries and interim calculations need to be exchanged in a private manner to avoid any disclosure. Therefore, peers' individual privacy should be maintained.

The general research directions to maintain privacy in CF systems are perturbation, anonymity and encryption-based methods. The idea in perturbation is to disguise the original vector and use it for the recommendation purposes so that the service provider cannot tell the true ratings of the users. Anonymization techniques, in general, aim to make it difficult to differentiate a user from a group of users so that identification efforts would be useless. Encryption-based methods, generally, provide secure multi-party computation satisfying better accuracy compared to the aforementioned methods; however, they are computationally more expensive (Wang, Zheng, Jiang, & Ren, 2018).

The randomized perturbation technique (RPT) is utilized by Polat & Du (2003). The authors propose to disguise numerically-rated vectors of users in central server-based PPCF schemes. Each user calculates $z$-scores for items they rate. Then, users generate a random integer, $r_i$, from a range decided by the server and add it to their $z$-scores. The server is not able to derive $z$-scores due to $r_i$. The authors show that aggregate value results via RPT approximate to the original aggregate value results. This method has been utilized in the PPCF community for central and distributed data partitioning scenarios (Polat & Du 2005a, 2005c; Polatidis, Georgiadis, Pimenidis, & Mouratidis, 2017; Bilge & Polat 2013; Yargic & Bilge 2019). Polatidis et al. (2017) utilize RPT with multi-level privacy. They add random values drawn from a scale to the original rating values contrary to the original study (Polat & Du 2003) where random values are added to $z$-scores. The level of privacy determines the range that random values are generated for each user. Yargic & Bilge (2019) propose a private CF approach for multi-criteria ratings. Their approach RPT and random filling methods to disguise rating value and rated items. Besides, they propose a novel entropy-based method to control privacy parameters for users and sub-criteria individually. Bilge & Polat (2013) utilize bisecting $k$-means while they disguise their data by random perturbation. Meng et al. (2019) propose a private quality of service recommendation considering the geographic locations of users. They apply random noise on the quality-of-service data and region aggregation to expand geographic regions to overcome privacy and sparsity problems. Berkovsky, Eytani, Kuflik, & Ricci, (2005) obfuscate user profiles with substitution by choosing random values from uniform or bell-curved distribution or default values. This scheme (Berkovsky et al. 2005) is extended by Berkovsky, Eytani, Kuflik, & Ricci 2006) with peer groups managed by a super-peer. Super

peers manage the neighborhood selection and aggregation calculation that will be transferred to the AP. The impact of data obfuscation is studied by (Berkovsky, Kuflik, & Ricci 2012). The scheme proposed by Boutet, Frey, Guerraoui, Jégou, & Kermarrec (2016) provides obfuscation protocol to hide user vectors and a differentially-private randomized dissemination protocol to forward item vectors. A differentially-private collaborative filtering scheme is proposed by Xiong, Zhang, Zhu, Li, & Zhou (2018). Authors propose a private protocol consisting of five stages, (i) defining a new symbol set for each user, (ii) rating items, (iii) constructing a rating matrix at the recommendation server, (iv-v) generating/publishing recommendations. Privacy is maintained by differential privacy and substitution method while defining a symbol set for ratings. Yang, Li, Sun, & Zhang (2019) propose two methods for privacy preservation. Their methods are differentially private. The first is applied to the input to perturb the original ratings by Laplace noise while their second method perturbs the similarity measurements of prediction by adding Laplace noise. Guo et al. (2019) propose a locally differentially private CF. The novel idea in the paper is that the similarity scores can be recovered from the perturbed data instead of computing similarities from the perturbed data. Besides rating values and rated items, another differentially private CF model keeps the learned model private (Jiang, Li, & Lin 2019), as well. Mazeh & Shmueli (2020) develop a private recommender system based on Open Personal Data Store (openPDS) (de Montjoye, Shmueli, Wang, & Pentland (2014). openPDS allows people to protect their data at a single location and serves the service providers through a module called SafeAnswers. The proposed scheme offers both item- and content-based private CF on openPDS. Healthcare recommendation based on arbitrarily distributed data (ADD) is proposed by Kaur et al. (2018) since patients' preferences can be distributed among different healthcare centers. Unlike homomorphic encryption techniques, this study achieves a computationally efficient PPCF model on ADD with multi-party random masking. Their scheme has two phases: offline and online. In the offline phase, they privately calculate vector dot products and lengths by utilizing randomization and homomorphic encryption, respectively. In the online phase, predictions are produced. When ratings are binary, the randomized response technique (RRT) is frequently used for perturbation. RRT is a survey technique that is employed to discover the percentage of a population with a sensitive attribute by asking a polar question to respondents (Warner 1965). Polat & Du (2006) proposes RRT to perturb binary rating vectors in central server-based systems. This technique either reverses or preserves all item ratings based on a predetermined threshold value by the server. Users generate a random number from a range and they reverse their ratings if the random number is greater than the threshold. However, the drawback is that all item

ratings are disclosed if one of the items is already known by a malicious server. Therefore, the authors propose to divide user vectors into multi-groups and apply RRT for each group separately. In a multi-group approach, item ratings that are in the same group with the disclosed rating are revealed in case of any disclosure. Items in other groups remain private. Besides, RRT is utilized for different data partitioning schemes in PPCF literature. Polat & Du (2005b; 2008) employ RRT when data is distributed between two-parties either horizontally or vertically. Horizontal partitioning means that each data holder holds different users' ratings for the same set of items. Vertical partitioning is the exact opposite: each data holder holds the same set of users' ratings for different items. Scholars also state that their scheme can be extended to multi-party distribution. Similarly, Kaleli & Polat (2007) propose a two-party PPCF scheme with multi-group RRT, and they utilize Naïve Bayes classifier to produce predictions. Although these studies deal with binary rated data, they are designed for server-based PPCF scenarios. On the other hand, Kaleli & Polat (2010) devises a binary P2P PPCF scheme. Similar to other binary PPCF schemes in this section, the AP prepares a query for each by using RRT with a multi-group. The AP also fills her query vector with random values to prevent other peers from disclosing rated items.

A $k$-anonymous approach to privacy is proposed by Casino, Domingo-Ferrer, Patsakis, Puig, & Solanas (2015). First, central value imputation is applied to fill unrated items, then, ratings are standardized with $z$-scores. $k$-most similar users are clustered and ratings in each cluster are replaced by their mean values so that they become indistinguishable from each other. Wei, Tian, & Shen (2018) improve the $k$-anonymous recommender scheme proposed by (Casino et al. 2015). They claim that central value imputation to fill the unrated items leads to inaccuracy; thus, they utilize Latent Factor Model. Besides, they also improve the maximum distance to average vector (MDAV) micro-aggregation algorithm based on importance partitioning (Kokolakis & Fouskakis 2009) to increase homogeneity in each group. Casino, Patsakis, & Solanas (2019) develop a central PPCF architecture with an anonymization server. They utilized variable-group-size microaggreation ($k$-anonymity). According to their default architectural design, users trust the anonymization server; however, if some users do not trust the server, they can route their preferences through other users they trust. Their design requires that all items be filled due to $k$-anonymity. Therefore, they first construct a full matrix by filling the empty cells. Then, variable-size MDAV (V-MDAV) (Solanas, Martinez-Balleste, & Domingo-Ferrer 2006) is applied to microaggregate the data. Besides the private recommendation scheme, the authors also introduce a behavioral precision metric that considers both positive and negative assessments. The rating scale is divided into

sub-ranges, the metric evaluates if the original rating and prediction fall into the same sub-range. A private CF framework is proposed for social communities by Li, Lv, Shang, & Gu (2017). The proposed "you are not alone" (YANA) system organizes the users into groups in a distributed manner by the private $k$-centroid algorithm which utilizes hash vectors for the distance calculation. In this study, pseudo users that are formed from group users' interests are communicated with the recommendation server instead of real user ratings. Hou, Wei, Wang, Cheng, & Qian (2018) propose a PPCF based medical recommendation. Their private scheme offers a neighborhood-based method aiming to protect confidential treatment and demographic information of patients. They propose a $k$-anonymity based neighbor selection and differentially private recommendation with the exponential mechanism.

Badsha et al. (2016) propose a private distributed recommender system with ElGamal encryption. Their method privately calculates item averages and similarity computations required for the prediction. The method also hides rated items. A user-based PPCF based on homomorphic encryption is studied by Badsha et al. (2017). Their scheme recommends an item among a set of items. The scheme has two different service providers, a recommender server and a decryption server, meaning that this is not serverless, as well. The decryption server provides public keys while users interact with the recommender server for recommendations. An unsynchronized secure multi-party computation protocol where a rating value is divided into pieces distributed among users is proposed by Li et al. (2016). The distributed pieces sum to the original rating value. The private recommendation process utilizes the unsynchronized model for the similarity calculation. The private framework also discusses how to update the model when new ratings are available. Badsha et al. (2018) propose a web service recommendation with homomorphic encryption. In their scheme, users do not need a decryption server to share their secret keys. A mobile application offering a privacy-preserving recommendation engine with two frameworks for cloud and social network is proposed by Xu, Zhang, & Yan (2018). The privacy is protected by anonymity, public key and homomorphic encryption in the application.

Our proposed scheme in this paper can be classified as a perturbation method. In general, there are two variants of perturbation-based privacy protection methods. RPT maintains privacy by adding noise to the original numeric vector (Polat & Du 2005a, 2005c; Polatidis et al. 2017; Bilge & Polat 2013; Yargic & Bilge 2019) and obfuscation maintains privacy by substitution (Berkovsky et al. 2005, 2006; Berkovsky et al. 2012). If ratings are binary, RRT is a frequently used technique (Polat & Du 2005b, 2006, 2008; Kaleli & Polat 2007, 2010; Berkovsky et al. 2005,

2006; Berkovsky et al. 2012), which is similar to the substitution by reversing the rating after checking a randomly generated number against a threshold. Our method is different from other perturbation methods on numeric ratings (Polat & Du 2005a, 2005c; Polatidis et al. 2017; Bilge & Polat 2013; Yargic & Bilge 2019). In this work, we propose a novel perturbation method for privacy where a user vector is divided into groups and each group is condensed into a value. Moreover, our method is straightforward and easy to implement. In addition to condensed vectors, we also consider possible disclosures of both AP and collaborating peers' private data. Although our private framework in this paper is built upon a numerically rated data, the main idea of summarizing a group of ratings can be easily extended to binary rated data unlike RPT methods (Polat & Du 2005a, 2005c), which are not applied to binary rated items.

## 3. A scheme for P2P prediction

Before delving into the details of our private P2P scheme, we would like to introduce a numeric P2P CF scheme without privacy concerns. After this section, one can figure out the basic CF scheme and its privacy violations that our PPCF scheme should address.

### 3.1. A simple P2P CF without privacy

A CF prediction process can be summarized in three main steps: (i) calculating similarities between the AP and other users, (ii) picking the neighbors and (iii) producing the prediction. Since our scheme is built on a P2P-network, the AP first probes who is going to join the prediction by sending $q$, the queried item, to other peers in the network. Peers who did not rate $q$ do not need to participate in the prediction process. Peers who rated $q$ become collaborating peers. After the AP learns about the collaborating peers, he/she sends his/her rating vector to them for the similarity calculation. Upon receiving the vector from the AP, collaborating peers calculate their similarities using Pearson correlation in Eq.1, where $sim_p$ is the similarity between a collaborating peer, $p$, and the AP, while $R$ is the set of commonly rated items between $p$, and the AP. $r_{AP,i}$ and $r_{p,i}$ are the ratings given by the AP and $p$ for the related item while $\overline{r}_{AP}$ and $\overline{r}_p$ are the average ratings of the AP and $p$, respectively.

$$sim_p = \frac{\sum_{i=R}(r_{AP,i} - \overline{r}_{AP})(r_{p,i} - \overline{r}_p)}{\sqrt{\sum_{i=R}(r_{AP,i} - \overline{r}_{AP})^2}\sqrt{\sum_{i=R}(r_{p,i} - \overline{r}_P)^2}} \qquad (1)$$

After each peer calculates its similarity with the incoming vector, they are collected by the AP for the neighbor selection. Having received $sim_p$, the AP sorts $sim_p$ values in a descending order to pick the best $k$ neighbors, where $1 \le k \ll n$. The peers with the first $k$ similarity scores are determined as neighbors. Then, the AP notifies the neighbors. Neighbors send their $(r_{p,q} - \overline{r}_p)$ values to the AP in order to produce the prediction by using Eq. 2 by Herlocker, Konstan, Borchers, & Riedl (1999). In this equation, $r_{p,q}$ is the rating of the related peer, $p$, for $q$, and $N$ is the set of neighbors. Since the AP already knows $sim_p$, the neighbors do not need to send them again.

$$r_{AP,q} = \overline{r}_{AP} + \frac{\sum_{p=N} (r_{p,q} - \overline{r}_p) \times sim_p}{\sum_{p=N} sim_p} \tag{2}$$

## 3.2. Privacy problems in the CF scheme

After discussing the CF procedure and the similarity measure that will be utilized in this study, the privacy issues with this CF scheme will be given in this subsection so that an appropriate private P2P prediction scheme can be constructed. Privacy is handled considering both AP and collaborating peers.

Individual privacy describes the privacy needs of users or peers, and it is two-fold Bilge et al. (2013): (i) rating value privacy and (ii) rated/unrated item privacy. To clarify, other peers should not know the item ratings made by any peer. Therefore, ratings must be perturbed so others cannot tell whether the rating they receive is an actual rating value or not if privacy is a concern in a CF prediction. The second is to disguise rated items. In addition to rating values, rated items are also considered private. Some peers might be concerned that they could be profiled over their rated items although their rating values are not known. For example, a peer who would like to rate a sensitive item whose content is controversial might shy away from providing her rating although her rating value will be confidential according to the first aspect of privacy. Because the existence of a rating will reveal that the related peer has accessed such a controversial item, peers might abstain themselves from rating sensitive items. Therefore, rated items should not be differentiated from unrated items. A private CF scheme should consider these two aspects of individual privacy to

encourage and convince people to participate in predictions. First and foremost, the objective of our private scheme is to perturb the AP's vector before sending it to other peers and convince the AP that his/her privacy is preserved in terms of the first and second aspect of privacy.

The CF scheme has privacy problems associated with collaborating peers, as well. Peers in a P2P network participating in a prediction process should protect their confidential data from being disclosed while communicating with the AP. Three pieces of information is dispatched from collaborating peers in this CF setting. The most critical is whether they rated $q$ or not. This can be revealed by monitoring the participation for the prediction, recall that peers rating $q$ can join the prediction process. The second is $sim_p$ value, which is collected by the AP for neighbor selection. The third is $(r_{p,q} - \overline{r}_p)$, which is the deviation from the average rating to calculate the prediction in Eq. 2. Thus, together with the AP, collaborating peers must also be convinced about privacy. We tackle the discussed privacy problems in the following section with our PPCF scheme

## 4. Private P2P predictions with condensed queries

In this section, measures taken to maintain AP's and collaborating peers' privacy will be discussed. First, privacy protection for the AP is given. The general idea in terms of the first aspect of privacy is to perturb his/her rating vector by creating a condensed vector. For the second aspect of privacy, some ratings will be appended into and removed from the condensed vector so that any collaborating peer will not tell which ratings are authentic and fake. Second, we give our approach to maintain collaborating peers' privacy. A communication protocol between peers and the AP is organized to preserve collaborating peers' $sim_p$ and $(r_{p,q} - \overline{r}_p)$ values. Besides, we also discuss that all peers should participate in the prediction to prevent collaborating peers' rated items from being discovered by a malicious AP.

### 4.1. Privacy protection for the AP

As discussed, AP's rating vector should be perturbed so that other peers cannot disclose ratings and rated items. We propose a data perturbation method that addresses these two aspects of privacy for the AP, who is looking for a prediction. Our method is straightforward. The AP creates a condensed rating vector from the original one to mask ratings and rated items. A condensed vector can be considered as a compressed version of the original vector. Items in the original vector are grouped together and an aggregate rating is produced for each group. Creating a condensed vector is a two-step process. First, the AP decides $z$, the number of items

that will be grouped together, then calculates $G = \lfloor m \div z \rfloor$, the required number of groups, where $m$ is the number of items. Notice that the last group could have $z + t$ items where $m \equiv t \pmod{z}$ if $m \not\equiv 0 \pmod{z}$. Then, for each group, the AP finds out which items are rated and calculates an aggregate rating out of the rated items. The aggregated rating can be calculated by using well-known aggregation functions such as *mean()*, *median()*, *max()*, and *min()*. If a group does not have any rated items, then the aggregated rating becomes unrated for that group. Algorithm 1 gives the perturbation method. The $agg\_function\_name$ parameter in the algorithm is a function for the aggregation calculation such as *mean()*, *median()*, *max()*, and *min()*. Fig. 1 gives an illustration of an extraction of a condensed vector from the original one with *mean()* function and $z = 4$.

---

**Algorithm 1: Condense an original vector**

---

| | |
|---|---|
| 1 | ***procedure*** Condense (*originalV, z, agg_function_name*) |
| 2 | $G \leftarrow floor(originalV.Size()/z)$ |
| 3 | $condensedV \leftarrow createEmptyVector(G)$ |
| 4 | $g \leftarrow 0$ |
| 5 | ***while*** $g \neq G$ ***do*** |
| 6 | $lBound \leftarrow g \times z$ |
| 7 | $uBound \leftarrow (g + 1) \times z$ |
| 8 | ***if*** $uBound \geq originalQ.Size()$ ***then*** |
| 9 | $uBound \leftarrow originalQ.Size()$ |
| 10 | $temp \leftarrow originalV[lBound: uBound]$ |
| 11 | $groupRatings \leftarrow temp[temp \neq 0]$ |
| 12 | ***if*** $is\_empty(groupRatings)$ ***then*** |
| 13 | $condensedV[g] \leftarrow 0$ |
| 14 | ***else*** |
| 15 | $condensedV[g] \leftarrow aggregate(groupRatings, agg\_function\_name)$ |
| 16 | $g \leftarrow g + 1$ |
| 17 | ***return*** $condensedV$ |

---

| 1 | 5 | 0 | 4 | 0 | 0 | 0 | 0 | 3 | 2 | 1 | 5 | 5 | 0 | 0 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|

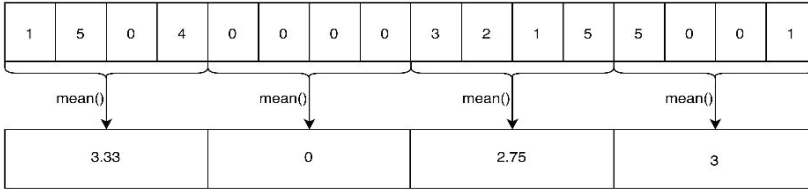| mean() | mean() | mean() | mean() |
|--------|--------|--------|--------|
| 3.33 | 0 | 2.75 | 3 |

**Figure 1.** Extracting a condensed vector from the original one

This method of data perturbation is effective in terms of privacy. When a malicious collaborating peer receives a condensed vector from the AP, he/she can identify neither individual item ratings nor rated items. Item ratings, the first aspect of privacy, are private because the AP generates an aggregate rating of $z$ items for each group. Likewise, the malicious peer cannot differentiate which items are rated among $z$ items in a group by checking the aggregate rating. However, if the related aggregate rating is unrated, the malicious peer figures out that all $z$ items of a group are unrated. Although the disclosure of unrated items is not as important as rated ones, the AP may wish to avoid it, as well. To overcome this problem, the AP can take an extra privacy measure by assigning average aggregate ratings for some of the unrated aggregate ratings. Similarly, the AP can mark some of the aggregate ratings as unrated. In the PPCF community, random filling method is widely used in such circumstances (Polat & Du 2005c, 2007; Kaleli & Polat 2010). The AP determines a $\sigma$ value between [0, 1). Then, he/she calculates $n_{update} = \lceil \sigma \times n_{agg} \rceil$, where $n_{agg}$ is the number of aggregated groups. The AP randomly picks $n_{update}$ number of aggregated groups and converts their rating to unrated. Similarly, she randomly picks $n_{update}$ number of unrated groups and converts their unrated ratings to average ratings. By this protocol, the AP ensures that any collaborating peer does not differentiate whether items of an unrated group are indeed unrated or not. By the way, the AP should realize that larger $\sigma$ values would deteriorate prediction results due to increased randomness. Algorithm 2 gives this process.

We believe the proposed method of condensed vectors will be effective to produce accurate private predictions because the CF data matrices are usually very sparse. Thanks to extreme sparsity in CF, it is highly probable that aggregated groups contain only one rating.

## 4.2. Privacy protection for peers

The previous section addresses the privacy protection for the AP; however, other peers participating in the prediction process might also need the privacy protection measures. The AP starts to disclose whether peers rate an item or not (the second aspect of privacy) as soon as they

become collaborating peers by accepting the prediction process request probed with $q$. The AP learns which peers rate $q$ by looking at which peers become collaborating peers. Moreover, a malicious AP can easily find out which items are rated by whom by simply asking predictions for all items in the data set. Therefore, this scheme needs an update to disguise peers who rate $q$. In order to protect the privacy of peers, we propose that all peers should join the prediction process instead of only letting the peers who rate $q$. However, the problem may arise when a peer who did not rate $q$ is selected as a neighbor because he/she will not be able to calculate $(r_{p,q} - \bar{r}_p)$. Therefore, collaborating peers who did not rate $q$ should fill the rating. At this point, they might choose to fill it with their average rating, $\bar{r}_p$, *median()* or a random rating. Choosing an average rating for $r_{p,q}$ can help a malicious AP figure out which peers did not rate $q$ since it will yield 0 for $(r_{p,q} - \bar{r}_p)$. Other imputation functions such as *median()*, random rating for $q$ will not disclose anything whether a peer rate $q$ or not.

---

**Algorithm 2: Disguise rated and unrated items**

---

| | |
|---|---|
| 1 | **procedure** DisguiseRatedUnrated(*condensedV*, $\sigma$, *G*, *avgrating*) |
| 2 | **if** $\sigma > 0$ and $\sigma \leq 1$ **then** |
| 3 | $nAgg \leftarrow sum(condensedV \neq 0)$ |
| 4 | $nUnrated \leftarrow G - nAgg$ |
| 5 | $nUpdate \leftarrow \lfloor nAgg \times \sigma \rfloor$ |
| 6 | $tempI \leftarrow find(condensedV \neq 0)$ |
| 7 | $sAgg \leftarrow sample(tempI, nUpdate, replace = False)$ |
| 8 | $unratedGs \leftarrow find(condensedV == 0)$ |
| 9 | $condensedV[sAgg] \leftarrow 0$ |
| 10 | **if** $nUpdate > nUnrated$ **then** |
| 11 | $sUnrated \leftarrow sample(unratedGs, nUpdate, replace = True)$ |
| 12 | **else** |
| 13 | $sUnrated \leftarrow sample(unratedGs, nUpdate, replace = False)$ |
| 14 | $condensedV[sUnrated] \leftarrow avgrating$ |
| 15 | **return** $condensedV$ |

---

Since direct transfer between collaborating peers and the AP might cause such disclosures, we propose a protocol that neighbors and the AP should follow. The protocol is about how the numerator part of the Eq. 2 is transferred to the AP. Remember that the prediction is performed by Eq.

2 and the AP needs only $(r_{p,q} - \bar{r}_p)$ from each neighbor to calculate the numerator part of the prediction since $sim_p$ has been already transferred. Also, note that the numerator part of Eq. 2 sums $(r_{p,q} - \bar{r}_p) \times sim_p$ over all neighbors. The AP does not need to know individual values of $(r_{p,q} - \bar{r}_p)$, the summation is enough for the calculation. Therefore, we develop a communication protocol that collaborating peers inform the AP about the final numerator value in Eq. 2. Our protocol updates how collaborating peers are informed about their neighborhood. In our scheme, instead of informing collaborating peers about their neighborhood and asking for $(r_{p,q} - \bar{r}_p)$, the AP now prepares a vector with two cells for each neighbor. The first cell contains a peer id, the second cell is a single bit value that determines the initiator neighbor. Once a collaborating peer receives this vector, she now knows that she is one of the neighbors. The first cell value, which contains a peer-id, identifies the recipient neighbor that the current neighbor will send her $(r_{p,q} - \bar{r}_p) \times sim_p$ value. Once the recipient peer receives $(r_{p,q} - \bar{r}_p) \times sim_p$, she adds her own value to it. This process is repeated by each neighbor and the final neighbor, whose recipient is the AP, will, therefore, have the single value summed over all neighbors and sends it to the AP. The AP cannot tell which values belong to whom. However, he/she now has all the necessary information collected from peers to calculate the prediction for $q$. The second cell in the vector is a single bit that identifies which neighbor peer will initiate the communication between peers. The neighbor peer whose value is 1 initiates the communication among peers. The only question with this algorithm is: "*is the recipient of the initiator able to discover $r_{p,q}$ from $(r_{p,q} - \bar{r}_p) \times sim_p$ ?*". The answer is simply no because the recipient of the initiator does not know if her sender is the initiator. Each neighbor just receives a single value from her sender and cannot differentiate if this value traversed many peers or originating from the initiator. No neighbor knows who the initiator is except the AP. Our algorithm of how the AP selects a random list of a sender-recipient vector with a single initiator bit is given in Algorithm 3.

## 4.3. The private protocol

After discussing both AP and collaborating peers' privacy, our scheme can be summarized as follows:

1.   The AP sends a request to peers in the network to probe their participation status to the prediction process.

**Algorithm 3: Peer privacy, creating sender-recipient vectors**

| | |
|---|---|
| 1 | **procedure** CreateSenderRecVectors($neighborIds$, $AP\_Id$) |
| 2 | $dest \leftarrow emptyDictionary()$ |
| 3 | $sender \leftarrow sample(neighborIds, 1)$ |
| 4 | $neighborIds.remove(sender)$ |
| 5 | **while** $neighborIds$ **do** |
| 6 | $rec \leftarrow sample(neighborIds, 1)$ |
| 7 | **if** $is\_empty(dest)$ **then** |
| 8 | $dest[sender] \leftarrow [rec, 1]$ |
| 9 | **else** |
| 10 | $dest[sender] \leftarrow [rec, 0]$ |
| 11 | $neighborIds.remove(rec)$ |
| 12 | $sender = rec$ |
| 13 | **if** $is\_empty(neighborIdList)$ **then** |
| 14 | $dest[sender] \leftarrow [AP\_Id, 0]$ |
| 15 | **return** $dest$ |

2. Peers inform the AP about their prediction status. Peers who rated $q$ can join the prediction process. As discussed in Section 4.2, the AP and peers can agree on arranging the prediction process without requiring any peers to rate $q$. Therefore, all peers become collaborating peers. Peers who did not rate $q$ fill the item rating with the average, median or random rating.

3. After the AP learns the collaborating peers (Step 2), the AP prepares a condensed query using Algorithm 1 as discussed in Section 4.1. The collaborating peers prepare their corresponding condensed queries, as well, to calculate their similarities with the AP. If the AP wishes to disguise his/her own rated items, the AP might choose to utilize Algorithm 2 to hide own rated items.

4. Collaborating peers calculate their similarities ($sim_p$) to the AP's condensed vector and send them to the AP.

5. After receiving $sim_p$ values from all collaborating peers, the AP sorts them in descending order and picks top $k$ neighbors.

6. Due to the possible privacy problems discussed in Section 4.2, the AP creates a sender-recipient vector using Algorithm 3 for each neighbor and lets them know through this vector.

7.  Each neighbor calculates own $(r_{p,q} - \overline{r}_p) \times sim_p$ value and sends it to their recipient neighbor peer. The recipient neighbor peer now calculates its own $(r_{p,q} - \overline{r}_p) \times sim_p$ and adds it to the incoming value. Each peer repeats this process until the final neighbor peer sends the aggregated sum to the AP.

8.  After receiving aggregated sums from the collaborating peers, the AP can now calculate the final prediction results by using Eq.2. Since the AP has already had $sim_p$ values from all collaborating peers and can compute $\overline{r}_{AP}$ from his/her rating vector, the AP does not need any other value for the prediction.

## 5. Privacy analysis

Privacy analysis of the private binary P2P scheme can be viewed in two aspects. The privacy of the AP and collaborating peers, respectively. First, we will analyze the possibility of recovering the AP's original rating vector from his/her condensed vector in the presence of a malicious collaborating peer.

Assume that the AP applies both Algorithm 1 and 2 to create a condensed vector. A malicious peer must first determine the genuine aggregated groups since Algorithm 2 makes some unrated groups aggregated and aggregated groups unrated. Therefore, a malicious peer needs to choose the correct $n_{agg} - n_{update}$ originally aggregated groups among $n_{agg}$ aggregated groups she received. The related probability is 1 out of $C(n_{agg}, n_{agg} - n_{update})$. After choosing $n_{agg} - n_{update}$ original aggregated groups, there still remains $n_{update}$ aggregated groups to be determined. These groups should be selected among $G - n_{agg}$ unrated groups. The related probability is 1 out of $C(G - n_{agg}, n_{update})$. The final probability to choose correct aggregated groups becomes $\frac{1}{C(n_{agg}, n_{agg} - n_{update})} \times \frac{1}{C(G - n_{agg}, n_{update})}$. After achieving to choose the correct aggregated groups, the second task is to identify each individual ratings in aggregated groups. According to Algorithm 1, each group consists of $z$ items and a single aggregated value represents these $z$ items. Assume that $R_g$ is the set of items that are rated in $g$th group. Since a group consisting of $z$ items can generate $2^z - 1$ subsets (excluding the empty set because we are working on an aggregated group), the possibility of selecting the set $R_g$ is $\frac{1}{2^z - 1}$. Upon determining the rated items, corresponding ratings must be determined. Assume that the rating scale is between 1 and $s$; so, the corresponding probability to discover ratings becomes $(\frac{1}{s})^{|R_g|}$, where $|R_g|$ is the cardinality of the set, $R_g$. After all these assumptions, the final probability of determining ratings

becomes $\left[\frac{1}{C(n_{agg},n_{agg}-n_{update})} \times \frac{1}{C(G-n_{agg},n_{update})} \times (\frac{1}{2^z-1})^{n_{agg}} \times \right.$ $\left. \prod_{g \in G_{agg}} (\frac{1}{s})^{|R_g|} \right]$, where $G_{agg}$ is the set of aggregated group ids.

We discussed the presence of a malicious peer, the AP may also act in a malicious manner. Section 4.2 discusses privacy protection for peers. Remember that an AP peer can easily learn the list of items rated by peers by requesting different participation requests for $q$. Therefore, our scheme proposes that all peers should join the prediction. In this paragraph, we will analyze the probability of recovering the peers who rate the queried item, $q$. Since all peers join the prediction, the first thing is to find out how many of them are the authentic raters of $q$, which is $n_{auth}$. The corresponding probability is $\frac{1}{n}$, where $n$ is the number of total peers. After finding out $n_{auth}$, the task is to pick correct $n_{auth}$ peers among $n$, which can be expressed as $C(n, n_{auth})$. The final probability, $\left[\frac{1}{n \times C(n,n_{auth})}\right]$, is true for only one queried item. For all items in the data set repeatedly, $n_{auth}$ will create a set $N_{auth}$ for each item $q$ in the data set. Then, the probability becomes $\left[(\frac{1}{n})^m \times \frac{1}{\prod_{n_{auth} \in N_{auth}} C(n,n_{auth})}\right]$, where $m$ is the number of items and $N_{auth}$ is the set having the number of authentic peers for each queried item, $q$.

## 6. Experiments
### 6.1. Data set and evaluation criteria

MovieLensMillion (Harper & Konstan, 2015) (MLM), Yahoo Movie (Yahoo! 2018a) and Yahoo Music (Yahoo! 2018b) data sets are utilized in the experiments. Yahoo Movie and Yahoo Music data sets come with their test and train sets. Yahoo Movie data set has 10,136 test samples while Yahoo Music has 54,000 test samples. MLM data set has no test samples; therefore, we randomly removed 20,000 samples from MLM and created test samples out of those removed ones. The remaining details about data sets are given in Table 1.

**Table 1**. Data sets

| Data sets | User | Items | Density | Rating-scale |
|---|---|---|---|---|
| MLM | 6,040 | 3,883 | %4.26 | 1-5 |
| YahooMovie | 7,642 | 11,915 | %0.23 | 1-13 |
| YahooMusic | 15,400 | 1,000 | %2.02 | 1-5 |

*Mean absolute error*, MAE, is selected as the primary evaluation criterion throughout the experiments. It is the average deviation of

the errors. In our case, MAE is the average difference between our predictions and the corresponding original ratings. Eq. 3 displays MAE equation where $r_i$ is the original rating, $\hat{r}_i$ is the prediction and $n$ is the number of predictions.

$$MAE = \frac{\sum_{i=1}^{n} |r_i - \hat{r}_i|}{n} \qquad (3)$$

Besides MAE, we also utilize *coverage*, the percentage of the requests that the system can successfully return a prediction, in the first experiment to show that condensed vectors do not hurt but contribute the coverage.

## 6.2. Experimental results
### 6.2.1. Effects of condensed vectors

The first conducted experiment tests how condensed vectors affect the accuracy of the prediction. Therefore, we only let the AP and collaborating peers form their condensed vectors using Algorithm 1. This experiment does not utilize any other privacy protection mechanism to analyze the pure effects of condensing vectors. In this experiment, we vary $k$, the number of neighbors, between 5 and 80, and $z$, group sizes, between 2 and 64. Besides, four different aggregation functions, *mean()*, *median()*, *max()* and *min()*, are compared in terms of the prediction accuracy. Fig. 2 displays the results with three sub-figures, notice that the x-axis of the figures displays the $z$ parameter. The first value in all figures, CF, displays CF results without any concern of privacy for comparison purposes. Since the use of the aggregation function does not affect the accuracy of the CF setting, it has been run once and the results are repeated for all aggregation cases in the figure.

The main purpose of this experiment is to analyze the degradation of accuracy due to condensed vectors. Our hypothesis is that grouping more items together (larger $z$ values) for condensed vector creations will cause an increase in MAE, which means a decrease in prediction accuracy, because the degree of perturbation increases as more items are aggregated to summarize a long vector into a small one. Before the comparison of the effects of how larger $z$ values affect the prediction accuracy let us gently discuss the accuracy, MAE in our case, in the CF setting. The general and observant trend in Fig. 2 is the obvious improvement when k is varied toward 80. For these data sets, larger neighbor sizes mean better prediction accuracy (lower MAE). When the effects of grouping larger $z$ values are examined for all data sets and aggregation function types, the increase in MAE results (lower accuracy) in MLM is marginal. Although the increasing trend of MAE results for Yahoo Movie and Yahoo Music
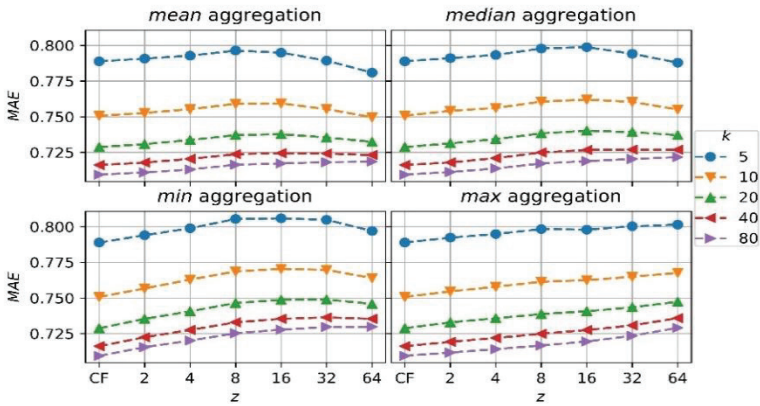
a) MLM
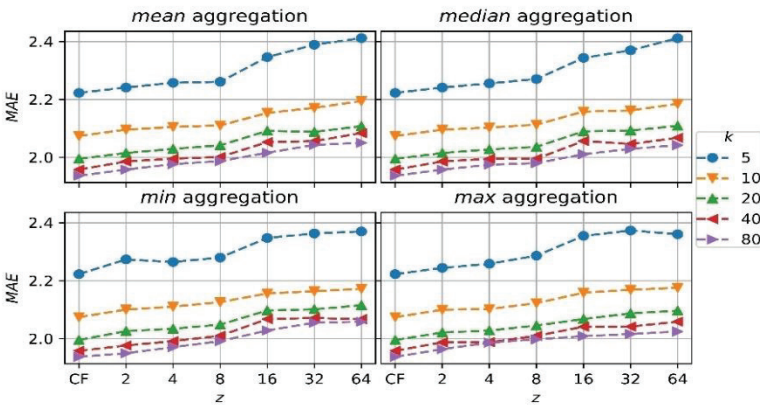


b) YahooMovie
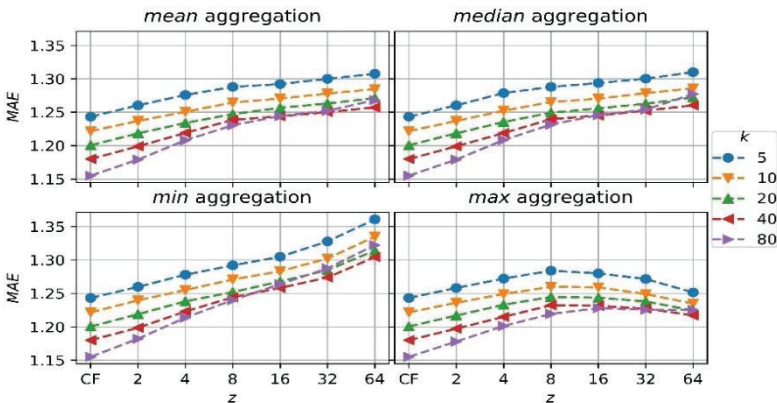


c) YahooMusic

**Figure 2.** Effects of condensed vectors on MAE

data sets is consistent, the slope of increase is not very steep. For example, with *mean*() aggregation function in Yahoo Movie data set for

$k$=20, the CF setting achieves around 2.00 MAE rate. Although the MAE rate increases consistently with increasing $z$ values for this case, it only reaches MAE around 2.11 when $z$=64. The absolute difference is just around 0.11, although the CF setting has no means of privacy while the PPCF scheme with $z$=64 summarizes 64 items together into one rating cell. The perturbation due to privacy is very high; however, the effect of it on accuracy is limited. We would also like to point out the fact that the increase in MAE in all other aggregation function types and data sets is clear; however, their true effect in terms of absolute change in MAE is not dramatic.

Besides MAE, we would like to look into coverage results as well. Fig. 3 displays the results. Generally speaking, as the number of neighbors to select, $k$ in our case, increases, the coverage should decrease because the system may not able to find enough neighbors to pick. This can be verified by the Fig. 3 for different values of $k$ for all data sets. Our method of condensed vectors, however, contributes the coverage although it preserves privacy for all data sets. This is especially clear for sparser data sets, Yahoo Movie and Yahoo Music. The coverage results get better for the same number of neighbors because grouping more items together (larger $z$ values) summarizes vectors and each user/peer is more likely to find enough neighbors to produce a prediction. As the number of $z$ decreases, each vector becomes more unique which makes them to find like-minded peers harder. The reason why sparse data sets, Yahoo Movie and Yahoo Music, record soaring coverage results while MLM follows almost a stable trend is that MLM is able to find enough neighbors regardless of how condensed the vector is due to its density.

To sum up, this experiment shows that privacy preservation via condensed vectors does not hurt accuracy dramatically while contributing to the coverage for sparse data set and retaining the coverage for dense data sets.

## 6.2.2. Effects of disguising rated/unrated items

After discussing the effects of introducing condensed vectors, the second experiment is related to disguising rated items. Note that condensed vectors help AP perturb the original rating vector so that ratings are masked. Any peer receiving the vector cannot tell individual rating values (the first aspect of privacy). On the other hand, disguising rated items in this experiment prevents AP's rated or unrated items from being disclosed. This experiment analyzes how adding or removing aggregated ratings from condensed vectors affect MAE results. Recall that $\sigma$ manages how many aggregated ratings will be canceled and how many unrated aggregated ratings will be filled. In order to measure the effect of $\sigma$, it is varied between 0.05 and 0.25 with 0.05 steps.
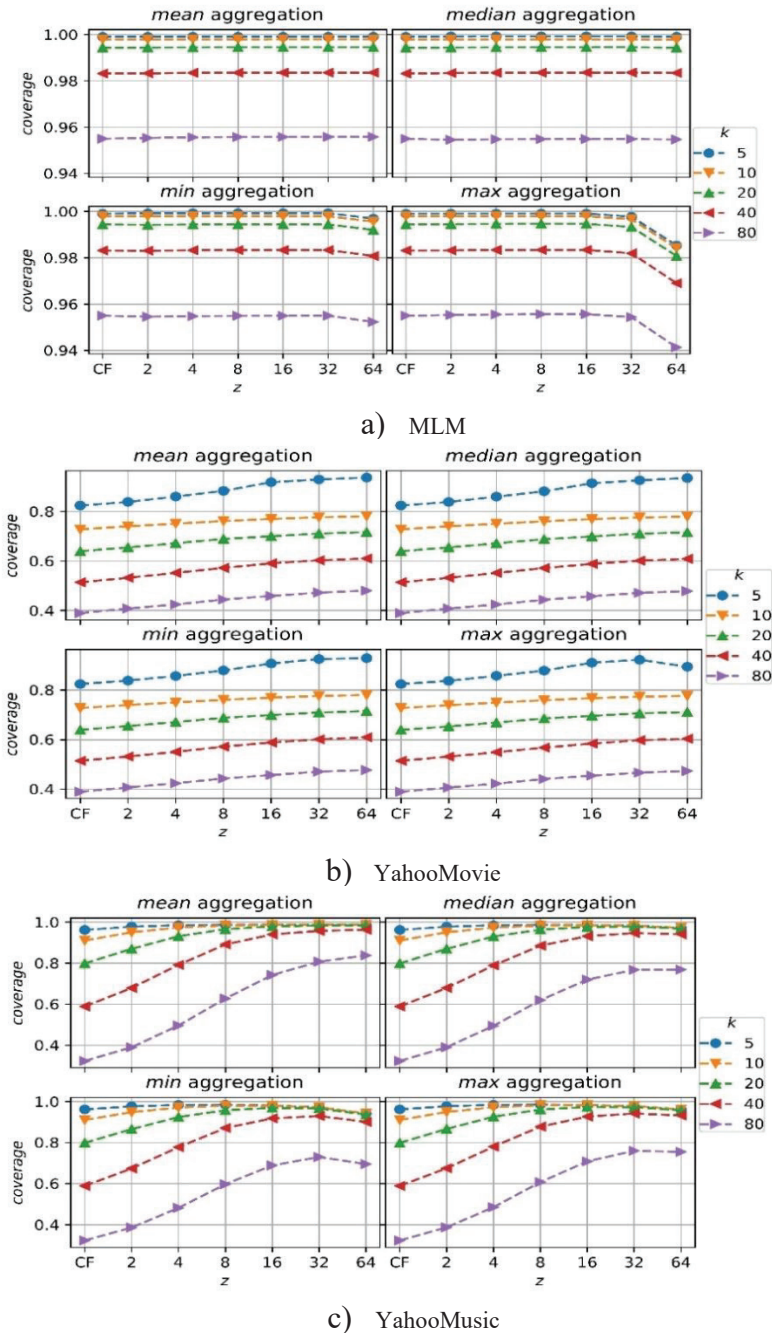
a)  MLM



b)  YahooMovie



c)  YahooMusic

**Figure 3.** Effects of condensed vectors on coverage

Since we already analyzed the effects of condensed vectors and parameters such as aggregation functions, $z$, and $k$, we set the aggregation function *mean()*, $z = 8$ and $k = 40$ for this experiment. The

results are depicted in Fig 4. The orange line in the figure demonstrates the results from the first experiment where $z = 8$ and $k = 40$ ($\sigma = 0$ no disguising) in Fig. 2. Therefore, it is a straight line indicating how much varying $\sigma$ values deviate from it.
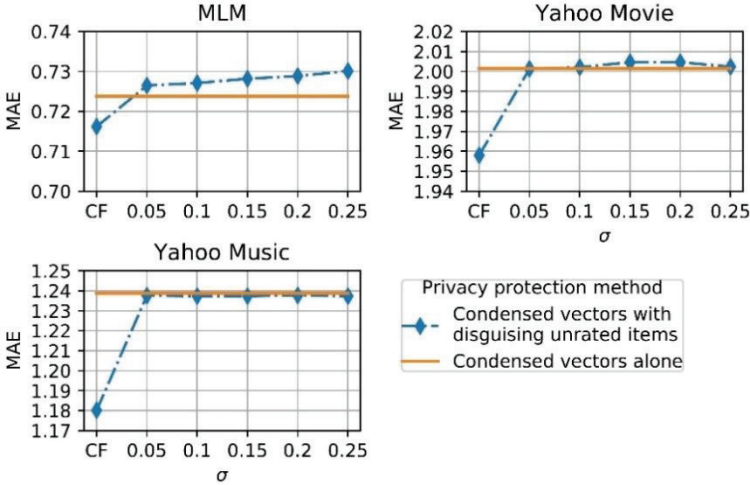


**Figure 4.** Effects of disguising unrated items

The experimental results show that introducing $\sigma$ parameter to disguise rated or unrated items has a deteriorating effect for MLM data set while the graph indicates a constant trend for the other two data sets. For example, the condensed vector approach (straight line) achieves around 0.72 MAE with MLM data set, disguising rated/unrated items approach is around 0.73 even with $\sigma = 0.25$. As a result, the experiment shows that introducing our approach of disguising rated/unrated items by modifying the condensed vector does not dramatically affect MAE results.

### 6.2.3. Effects of peer participation

After discussing the effects of introducing condensed vectors and disguising rated/unrated items for the AP, we now turn our attention to the effects of peer participation privacy on the prediction accuracy. Remember that our privacy-preserving P2P CF scheme with peer participation proposes that all peers in the network should participate in the prediction to prevent an AP from actively discovering which peers rated the queried item, $q$, by asking predictions for different $q$ each time. This mechanism helps the collaborating peers join the prediction process without privacy concerns. Since all peers join the process now, some of the neighbors can be selected among the peers who did not rate $q$. Remember that $(r_{p,q} - \bar{r}_p)$ is essential for the prediction; thus, they need to fill $q$'s rating. Our hypothesis in this experiment is that MAE results

will increase, which means a decrease in the prediction accuracy, due to the fact that ratings of $q$ are filled. In this experiment, similar to the first and second experiments, we compare our proposed scheme, which allows all peers to participate, with the CF setting. Additionally, we also include the private protocol where a condensed vector is applied by disguising rated/unrated items. We set the aggregation function *mean()*, $z = 8$, $k = 40$ and $\sigma = 0.1$ to create condensed vectors and allow all peers to participate in the prediction. This experiment also compares the different functions, $\bar{r}_p$, *median()* or random ratings, to fill the $q$ for the selected neighbors who indeed did not rate $q$.

The results are given in Fig. 5. The orange bar is the condensed vector approach with $z = 8$, $k = 40$ and $\sigma = 0.1$; however, peer participation is not included for comparison purposes. Since peers who did not rate the $q$ can now be selected as neighbors if their similarities are among the best $k = 40$, a decrease in the prediction accuracy can be seen for all filling methods. When compared, the worst MAE results are reported with the random filling because they are non-personalized randomly selected ratings. In addition, allowing all peers to join the prediction records worse MAE results compared to the condensed vectors with $\sigma = 0.1$. Such a case is expected because allowing all peers into the prediction has a direct effect on the prediction when Eq.2 is considered. It is obvious that filling $q$ values directly manipulates the prediction quality.
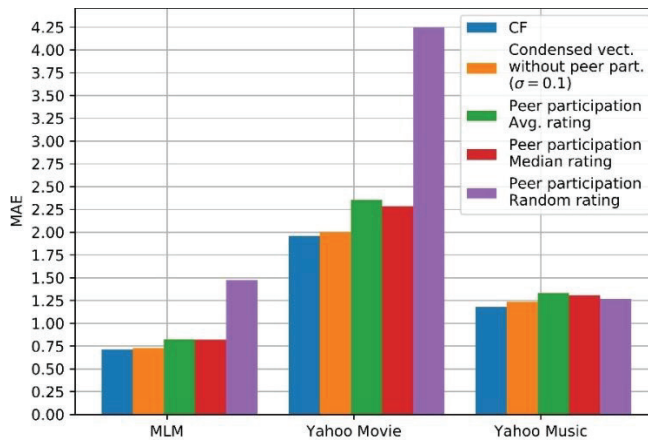


**Figure 5.** Effects of peer participation

## 7. Conclusion and future work

Privacy is an important issue in online avenues. Collaborative filtering systems rely on user ratings to provide accurate predictions. However, users might be reluctant to participate or shy away from providing their true opinions if they are not convinced that their data is confidential. Such

concerns might even worsen the sparsity problem in collaborative filtering systems. In this paper, we aim to provide a private collaborating filtering framework which users act in a peer-to-peer manner without a need for a server.

Our proposed scheme offers privacy measures to active and collaborating peers. An active peer forms a condensed vector by dividing his/her original vector into groups and calculating an aggregate value for each group so that collaborating peers do not disclose individual item ratings. Moreover, an active peer might fill some unaggregated group cells or remove some aggregated rating values to prevent his/her rated/unrated items from being disclosed by collaborating peers. In terms of collaborating peers' privacy, the proposed method requires that all peers join the prediction so that peers who rated the queried item cannot be discovered by the active peer. Second, instead of sending interim values, which is required for the prediction calculation, directly to the active peer, each peer transfers cumulative values through other peers to the active peer so that the active peer cannot learn individual values other than similarities, which are required for the neighbor selection. Our private collaborative filtering scheme has been tested with well-known three different benchmark data sets. Experimental results display promising prediction accuracy results.

In our future work, our scheme can be adapted to other data sharing scenarios other than a peer-to-peer network. Different data holders can co-operate for collaborating filtering purposes. We could also extend our scheme to distributed server-based schemes with two- or multi-party to maintain data holders' privacy. Our scheme can be easily adapted to private prediction schemes with binary ratings. We could also study the robustness of our scheme to analyze whether our scheme is immune to different shilling attacks. In addition, private schemes in the literature generally do not have policies when a user or peer updates the original vector by adding or removes a rating into/from the original vector. Therefore, the policies to apply when there is an update in the original vector remains as future work, as well.

# References

Badsha, S., Yi, X., & Khalil, I. (2016). A Practical Privacy-Preserving Recommender System. *Data Science and Engineering*, *1*(3), 161–177.

Badsha, S., Yi, X., Khalil, I., & Bertino, E. (2017). Privacy Preserving User-Based Recommender System. *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, 1074–1083.

Badsha, S., Yi, X., Khalil, I., Liu, D., Nepal, S., & Lam, K. (2018). Privacy Preserving User Based Web Service Recommendations. *IEEE Access*, *6*, 56647–56657.

Berkovsky, S., Eytani, Y., Kuflik, T., & Ricci, F. (2005). Privacy-enhanced collaborative filtering. *Proc. User Modeling Workshop on Privacy-Enhanced Personalization*.

Berkovsky, S., Eytani, Y., Kuflik, T., & Ricci, F. (2006). Hierarchical neighborhood topology for privacy enhanced collaborative filtering. *Proceedings of PEP06, CHI 2006 Workshop on Privacy-Enhanced Personalization, Montreal, Canada*, 6–13.

Berkovsky, S., Kuflik, T., & Ricci, F. (2012). The impact of data obfuscation on the accuracy of collaborative filtering. *Expert Systems with Applications*, *39*(5), 5033–5042.

Bilge, A., Kaleli, C., Yakut, I., Gunes, I., & Polat, H. (2013). A survey of privacy-preserving collaborative filtering schemes. *International Journal of Software Engineering and Knowledge Engineering*, *23*(08), 1085–1108.

Bilge, A., & Polat, H. (2013). A scalable privacy-preserving recommendation scheme via bisecting k-means clustering. *Information Processing & Management*, *49*(4), 912–927.

Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013). Recommender systems survey. *Knowledge-Based Systems*, *46*(Supplement C), 109–132.

Boutet, A., Frey, D., Guerraoui, R., Jégou, A., & Kermarrec, A.-M. (2016). Privacy-preserving distributed collaborative filtering. *Computing*, *98*(8), 827–846.

Canny, J. (2002). Collaborative filtering with privacy. *Proceedings 2002 IEEE Symposium on Security and Privacy*, 45–57.

Casino, F., Patsakis, C., Puig, D., & Solanas, A. (2013). On Privacy Preserving Collaborative Filtering: Current Trends, Open Problems, and New Issues. *2013 IEEE 10th International Conference on E-Business Engineering*, 244–249.

Casino, F., Domingo-Ferrer, J., Patsakis, C., Puig, D., & Solanas, A. (2015). A k-anonymous approach to privacy preserving collaborative filtering. *Journal of Computer and System Sciences*, *81*(6), 1000–1011.

Casino, F., Patsakis, C., & Solanas, A. (2019). Privacy-preserving collaborative filtering: A new approach based on variable-group-size microaggregation. *Electronic Commerce Research and Applications*, *38*, 100895.

Culnan, M. J., & Armstrong, P. K. (1999). Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation. *Organization Science*, *10*(1), 104–115.

de Montjoye, Y.-A., Shmueli, E., Wang, S. S., & Pentland, A. S. (2014). openPDS: Protecting the Privacy of Metadata through SafeAnswers. *PLOS ONE*, *9*(7), 1–9.

Guo, T., Luo, J., Dong, K., & Yang, M. (2019). Locally differentially private item-based collaborative filtering. *Information Sciences*, *502*, 229–246.

Harper, F. M., & Konstan, J. A. (2015). The MovieLens Datasets: History and Context. *ACM Trans. Interact. Intell. Syst.*, *5*(4), 19:1-19:19.

Herlocker, J. L., Konstan, J. A., Borchers, A., & Riedl, J. (1999). An Algorithmic Framework for Performing Collaborative Filtering. *Proceedings of the 22Nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 230–237. New York, NY, USA: ACM.

Hou, M., Wei, R., Wang, T., Cheng, Y., & Qian, B. (2018). Reliable Medical Recommendation Based on Privacy-Preserving Collaborative Filtering. *Computers, Materials & Continua*, *56*(1), 137–149.

Jiang, J.-Y., Li, C.-T., & Lin, S.-D. (2019). Towards a more reliable privacy-preserving recommender system. *Information Sciences*, *482*, 248–265.

Kaleli, C., & Polat, H. (2007). Providing Na\"ive Bayesian Classifier-Based Private Recommendations on Partitioned Data. In J. N. Kok, J. Koronacki, R. Lopez de Mantaras, S. Matwin, D. Mladenič, & A. Skowron (Eds.), *Knowledge Discovery in Databases: PKDD 2007* (pp. 515–522). Berlin, Heidelberg: Springer Berlin Heidelberg.

Kaleli, C., & Polat, H. (2010). P2P collaborative filtering with privacy. *Turkish Journal of Electrical Engineering & Computer Sciences*, *18*(1), 101–116.

Kaur, H., Kumar, N., & Batra, S. (2018). An efficient multi-party scheme for privacy preserving collaborative filtering for healthcare recommender system. *Future Generation Computer Systems*, *86*, 297–307.

Kokolakis, G., & Fouskakis, D. (2009). Importance partitioning in micro-aggregation. *Computational Statistics & Data Analysis*, *53*(7), 2439–2445.

Li, D., Chen, C., Lv, Q., Shang, L., Zhao, Y., Lu, T., & Gu, N. (2016). An algorithm for efficient privacy-preserving item-based collaborative filtering. *Future Generation Computer Systems*, *55*, 311–320.

Li, D., Lv, Q., Shang, L., & Gu, N. (2017). Efficient privacy-preserving content recommendation for online social communities. *Neurocomputing*, *219*, 440–454.

Mazeh, I., & Shmueli, E. (2020). A personal data store approach for recommender systems: enhancing privacy without sacrificing accuracy. *Expert Systems with Applications*, *139*, 112858.

Meng, S., Qi, L., Li, Q., Lin, W., Xu, X., & Wan, S. (2019). Privacy-preserving and sparsity-aware location-based prediction method for collaborative recommender systems. *Future Generation Computer Systems*, *96*, 324–335.

Polat, H., & Du, W. (2003). Privacy-preserving collaborative filtering using randomized perturbation techniques. *Third IEEE International Conference on Data Mining*, 625–628.

Polat, H., & Du, W. (2005a). Privacy-Preserving Collaborative Filtering on Vertically Partitioned Data. In A. M. Jorge, L. Torgo, P. Brazdil, R. Camacho, & J. Gama (Eds.), *Knowledge Discovery in Databases: PKDD 2005* (pp. 651–658). Berlin, Heidelberg: Springer Berlin Heidelberg.

Polat, H., & Du, W. (2005b). Privacy-Preserving Top-N Recommendation on Horizontally Partitioned Data. *Proceedings of the 2005 IEEE/WIC/ACM International Conference on Web Intelligence*, 725–731. Washington, DC, USA: IEEE Computer Society.

Polat, Huseyin, & Du, W. (2005c). SVD-based Collaborative Filtering with Privacy. *Proceedings of the 2005 ACM Symposium on Applied Computing*, 791–795. New York, NY, USA: ACM.

Polat, H., & Du, W. (2006). Achieving Private Recommendations Using Randomized Response Techniques. In W.-K. Ng, M. Kitsuregawa, J. Li, & K. Chang (Eds.), *Advances in Knowledge Discovery and Data Mining* (pp. 637–646). Berlin, Heidelberg: Springer Berlin Heidelberg.

Polat, H., & Du, W. (2007). Effects of Inconsistently Masked Data Using RPT on CF with Privacy. *Proceedings of the 2007 ACM Symposium on Applied Computing*, 649–653. New York, NY, USA: Association for Computing Machinery.

Polat, H., & Du, W. (2008). Privacy-preserving top-N recommendation on distributed data. *Journal of the American Society for Information Science and Technology*, *59*(7), 1093–1108.

Polatidis, N., Georgiadis, C. K., Pimenidis, E., & Mouratidis, H. (2017). Privacy-preserving collaborative recommendations based on random perturbations. *Expert Systems with Applications*, *71*, 18–25.

Solanas, A., Martinez-Balleste, A., & Domingo-Ferrer, J. (2006). V-MDAV: a multivariate microaggregation with variable group size. *17th COMPSTAT Symposium of the IASC, Rome*, 917–925.

Wang, C., Zheng, Y., Jiang, J., & Ren, K. (2018). Toward Privacy-Preserving Personalized Recommendation Services. *Engineering*, *4*(1), 21–28.

Warner, S. L. (1965). Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias. *Journal of the American Statistical Association*, *60*(309), 63–69.

Wei, R., Tian, H., & Shen, H. (2018). Improving k-anonymity based privacy preservation for collaborative filtering. *Computers & Electrical Engineering*, *67*, 509–519.

Xiong, P., Zhang, L., Zhu, T., Li, G., & Zhou, W. (2018). Private collaborative filtering under untrusted recommender server. *Future Generation Computer Systems*. doi:10.1016/j.future.2018.05.077

Xu, K., Zhang, W., & Yan, Z. (2018). A privacy-preserving mobile application recommender system based on trust evaluation. *Journal of Computational Science*, *26*, 87–107.

Yahoo! (2018a). *Yahoo! Movies User Ratings and Descriptive Content Information, version 1.0*.

Yahoo! (2018b). *Yahoo! Music ratings for User Selected and Randomly Selected songs, version 1.0*.

Yang, J., Li, X., Sun, Z., & Zhang, J. (2019). A differential privacy framework for collaborative filtering. *Mathematical Problems in Engineering*, *2019*.

Yargic, A., & Bilge, A. (2019). Privacy-preserving multi-criteria collaborative filtering. *Information Processing & Management*, *56*(3), 994–1009.

Zhang, B., Wang, N., & Jin, H. (2014). Privacy Concerns in Online Recommender Systems: Influences of Control and User Data Input. *10th Symposium On Usable Privacy and Security (SOUPS 2014)*, 159–173. Menlo Park, CA: USENIX Association.

# Chapter 10

**COMPARISON OF GENERAL PROPERTIES OF HEMP / FLAX NATURAL FIBERS AND GLASS / CARBON SYNTHETIC FIBERS**

*Yalçın BOZTOPRAK[1]*
*Muhammed Ali CAN[2]*

1  Asst. Prof. Dr. Yalçın BOZTOPRAK, Marmara University, Faculty of Technology, Metallurgy and Materials Engineering, Göztepe Campus, Istanbul / Turkey,  yboztoprak@marmara.edu.tr
2 Muhammed Ali CAN, Marmara University, Institute of Pure and Applied Science (Master Degree), Göztepe Campus, Istanbul / Turkey,  canmuhammedali@outlook.com

## INTRODUCTION

Composite materials consist of two or more components. One of these components is called reinforcement material and the other is called matrix. Reinforcing phase materials can be in the form of fibers, flakes, or particles. Matrix phase materials are generally continuous and polymers, ceramics, and metals can be used as matrix phases [1,2].

Material production in many areas for years is moving from single-phase materials to glass fiber or carbon fiber reinforced polymeric composite materials. Natural fibers are shown by scientists as a new alternative for polymer matrix composites. There are debates about whether natural fibers can replace carbon fiber and glass fiber [3]

Biocomposite materials are composite materials containing natural fiber together with resin. These biocomposite materials are biodegradable. The use of all kinds of natural fibers such as wood fibers, hemp, flax, cotton, jute and some cellulose fibers in the production of biocomposite materials is increasing day by day. Biocomposite materials are very useful materials in terms of both their biodegradability and mechanical properties and economic standards. Flax and hemp fibers are also among the fibers used to make biocomposite materials [4,5].
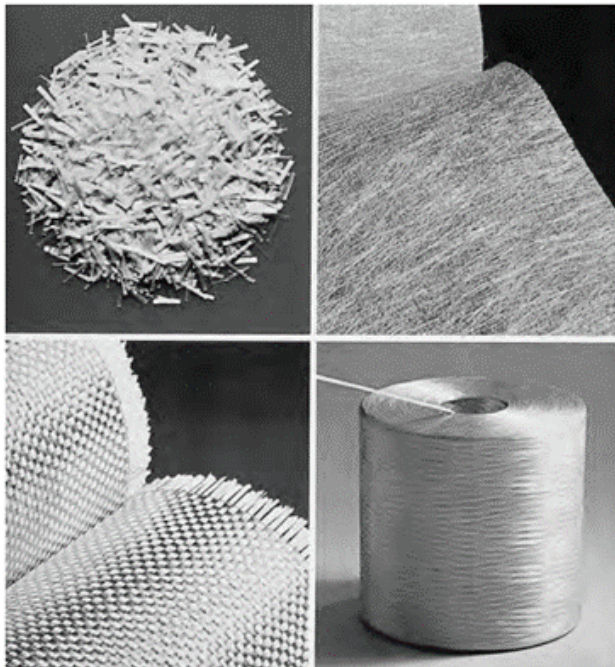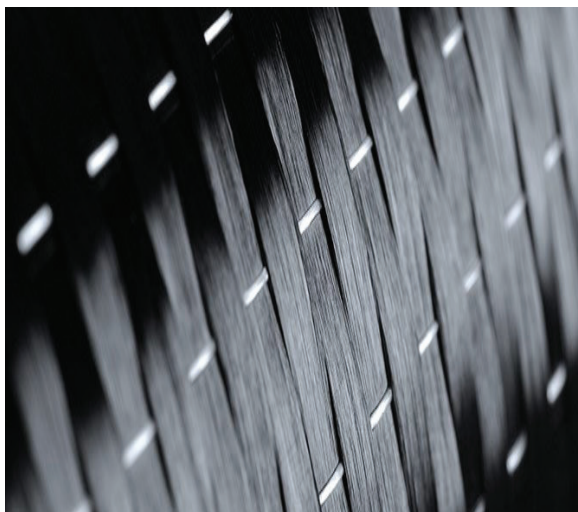
**Figure 1.** *Glass fibre and carbon fibre [6,7]*

Natural fibers such as flax and hemp may have various advantages and disadvantages compared to artificial fibers such as glass and carbon. To list the advantages; low density, corrosion resistance, less health risk, recyclable and biodegradable. In addition to these properties, natural fibers have a very important place compared to artificial fibers in terms of energy consumption. The production cost of natural fibers is lower than glass and carbon fibers [8-10]. Therefore, natural fibers may replace carbon and glass fibers in the future [11].



**Figure 2.** *Flax and hemp fiber [12, 13]*

Natural fibers such as flax and hemp have also some disadvantages. Natural fibers absorb moisture because they contain cellulose. The moisture content of the reinforcement material is undesirable for the production of composite material. On the other hand, these cellulose-containing natural fiber materials have an easily flammable feature.

With the change of thermal conditions, mass loss can be observed in natural fiber materials. If we take into account that the mechanical properties of the fiber will change with this mass loss, flax and hemp natural fibers will have worse properties than artificial fibers such as glass and carbon in terms of stability.

Natural fibers, one of the main materials in biocomposites, have various uses. One of these fibers is hemp fiber. Hemp fiber is one of the strongest and tightest natural fibers, so it has many different uses. Hemp fibers are used in the construction industry as insulation materials, in textile factories, furnitures, in paper making and hemp can also be used in the food and drug industry. Biocomposites produced from hemp are used in the automotive industry, both in the car body and in the doors. Composites made of hemp fiber are used, for example in the German automotive industry such as BMW, Mercedes, Volkswagen Golf, in the French automotive factories such as Peugeot, Renault and in the Italian Alfa Romeo. Hemp fibers reinforced polypropylene are used in automobile door panels and dash boards. Hemp fiber is used in the construction industry for heat insulation and sound insulation. For example, fiber-reinforced concrete has less energy consumption than standard concrete. Fiber-reinforced concrete has reduced energy consumption by 45%, and studies have observed that it reduces humidity [14].

In addition, hemp fiber is used in many places such as aviation industry, building coatings, heat and sound insulation materials, composites, maritime industry, sports and musical instruments.



**Figure 3.** *The use of hemp fiber in automobile door [15]*

**Figure 4.** *The airplane made of hemp fiber [16]*



**Figure 5.** *The car made of hemp fiber [17]*

Flax fiber, another fiber used in the production of biocomposites, is used in many areas considering its various properties, robust and healthy fiber structure. For example, flax fiber is used in various sectors for sound and heat insulation. Like hemp, it is used in the textile industry, automotive industry, aviation, maritime, composite production and construction industry. Especially in the textile industry, flax fibers are much preferred because they are more durable than cotton. The usage areas of natural fibers are parallel to each other [18].

**Figure 6.** *Flax and hemp fibers as reinforcement in composite [19]*

## 1. CHEMICAL COMPOSITION OF HEMP AND FLAX FIBERS

Hemp and flax, which are natural cellulosic fibers, contain cellulose, hemicellulose, lignin and pectin. These chemical contents change the properties of these natural fibers. Figure 7 shows the structure of natural fibers. Flax stalk and fiber structure are shown in Figure 8.
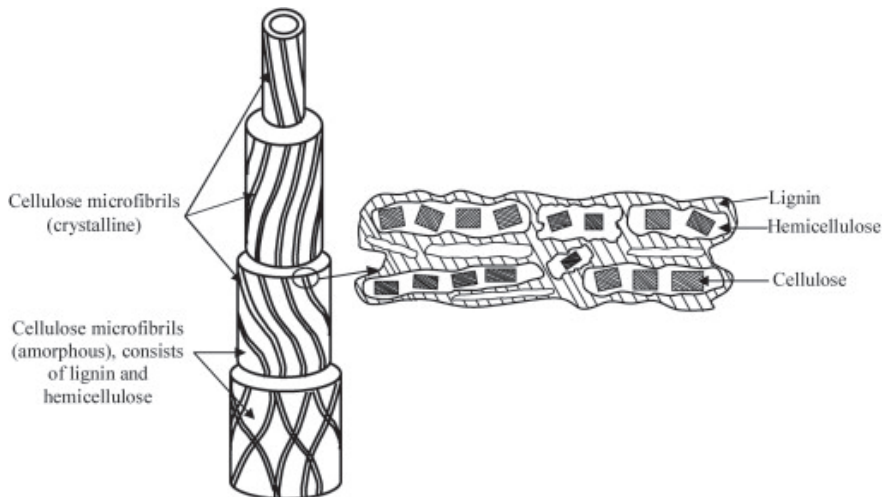


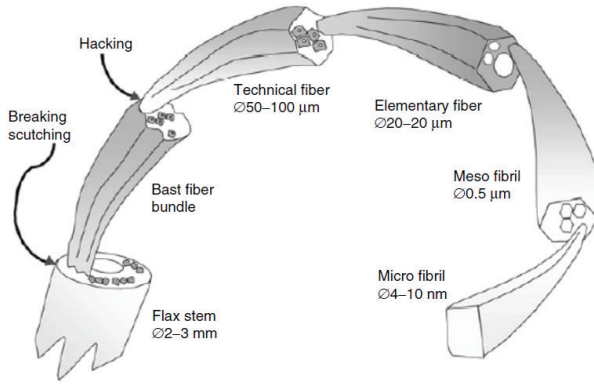**Figure 7.** *Structure of natural fibres [20]*

**Figure 8.** *Schematic of flax stalk and fiber structure [21]*

Some chemical components of flax and hemp fibers are given in Table 1.

**Table 1.** *Chemical properties of some natural fibers [22, 23]*

| Fiber | Cellulose (%) | Hemicellulose (%) | Lignin (%) | Pectin (%) |
|-------|---------------|-------------------|------------|------------|
| Hemp  | 74            | 18                | 4          | 1          |
| Flax  | 81            | 14                | 3          | 4          |

To better understand the physical and chemical properties of these natural fibers more efficiently, it is necessary to examine the materials contained in the fibers. As can be seen from Table 1, flax and hemp fibers are formed by the combination of many elementary fibers. These fibers stay together thanks to pectin, it is an effective structural polysaccharide and has a high water holding capacity. Found in the primary wall and middle lamella. Figure 9 shows a schematic comparison of the stalk cross sections of flax, nettle and hemp.
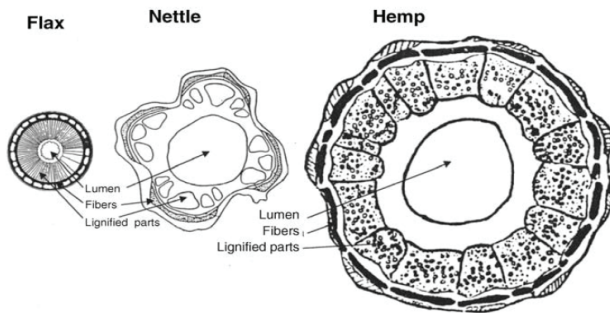


**Figure 9.** *A schematic comparison of the stalk cross sections of flax, nettle and hemp*

Another chemical content of these natural fibers is that the structural content of lignin is the second most abundant biopolymer in the world after cellulose. The most important features that distinguish this biopolymer from cellulose are that it has aromatic and aliphatic groups and its basic component is 4 alkyl catechols. The lignin molecule is difficult to break down and is resistant to enzymes. It also increases the reactivity of the fibers [24].

Another chemical ingredient of natural fibers is hemicellulose. Hemicelluloses are heterogeneous pentose, hexose and sugar-acid polymers. They have a heterogeneous structure. The most important properties that these compounds add to the fibers are air permeability and heat insulation.

## 2. MECHANICAL PROPERTIES OF HEMP AND FLAX FIBERS

The mechanical properties of carbon, glass, flax and hemp fibers are given in Table 1. Here, it is seen that natural fibers and synthetic fibers have different mechanical properties. However, if we compare the superiority of mechanical properties between natural fibers and synthetic fibers, this situation changes according to the properties and usage areas of the material.

Table 1 gives the values for the mechanical properties of flax, hemp, glass and carbon fibers and Figure 10 shows the comparisons of these values. Linen and hemp fibers have lower densities than glass and carbon fibers. When the tensile strength and modulus of elasticity values are examined, it is seen that the strength of glass and carbon fiber is higher. Considering the % elongation amount, it is seen that glass and linen fibers are more flexible than the others. While no moisture absorption is observed in glass and carbon fiber, it is seen that the moisture absorption rate is very high in flax and hemp fiber [25-27].

**Table 1:** *Mechanical properties of fibers*

| Fiber | Density (g/cm³) | Tensile Strength (Mpa) | Specific Tensile Strength (Mpa) | Elastic Modulus (Gpa) | Specific Elastic Modulus (GPa) | Elongation at failure (% | Moisture absorption (%) |
|---|---|---|---|---|---|---|---|
| *Carbon* | 1.7 | 3530 | 2857 | 230-240 | 164-171 | 1.75 | N/A |
| *E-Glass* | 2.6 | 2000-3500 | 800-1400 | 70 | 28 | 4.8 | N/A |
| *Hemp* | 1.4 | 550–900 | 393–643 | 70 | 50 | 1.6 | 6-12 |
| *Flax* | 1.4 | 800–1500 | 571–1071 | 60–80 | 43–57 | 2.7-3.2 | 8-12 |

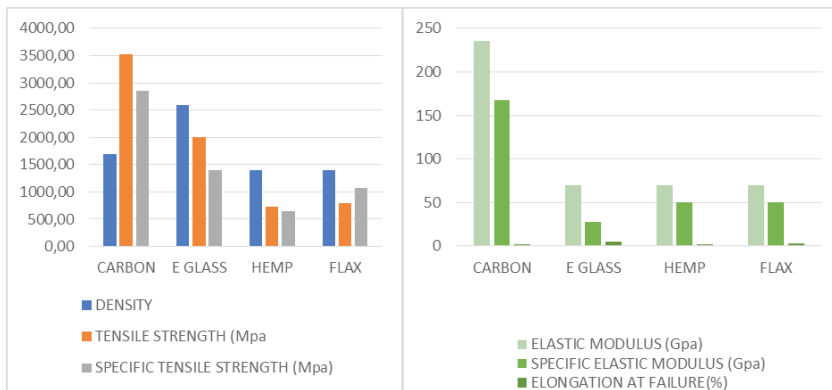* Quoted values may differ in specifically cited sources.

**Figure 10.** *Mechanical properties of fibers*

The tensile strength of natural fibers is an important property in terms of mechanical properties. Prasad and Sain conducted a study on their tensile strength using hemp fibers with diameters ranging from 4 µm to 800 µm [28, 29].

They found that the tensile properties of hemp fibers depend on fiber diameters and gradually decrease with increasing fiber diameter. This is consistent with the general observation that can also be applied to synthetic fibers; As the fiber diameter decreases, the amount of defects in the fibers also decreases, thus causing an increase in the tensile properties of the fibers. The average tensile strength and modulus values of hemp fibers with a diameter of 4 µm are 4200 MPa and 180 GPa, respectively. These values are 250 MPa and 11 GPa for fibers with a diameter of 66 µm, and 10 MPa and 2 GPa for fibers with a diameter of 800 µm.

## 3. THERMAL PROPERTIES OF HEMP AND FLAX FIBERS

Since hemp fiber contains cellulose, it absorbs moisture. All natural fibers also contain moisture like hemp fiber. According to research of A. Shahzad, when hemp fibers are kept in a desiccator and exposed to high temperatures, weight loss behavior appears. We see this weight loss in Asım hahzad's "A Study in Physical and Mechanical Properties of Hemp Fiber" study [29].
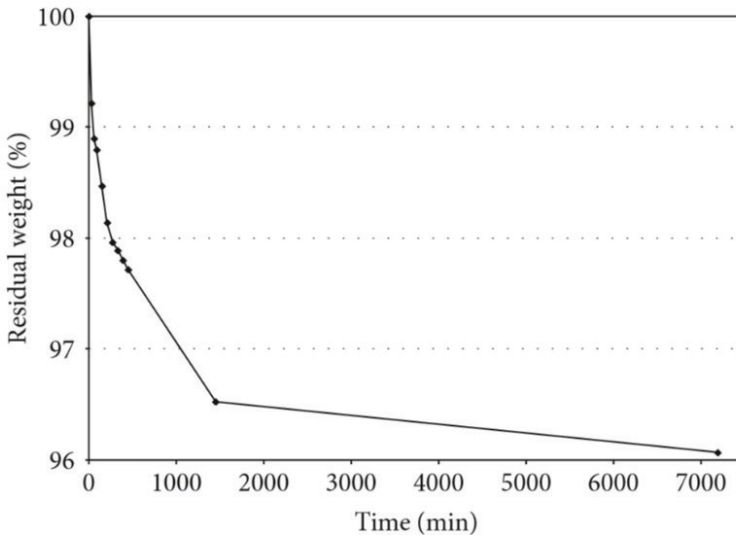
**Figure 11.** *Weight loss of hemp fibers in a desiccator*

In a study, a sample cut from hemp fiber felt conditioned at 23°C and 50% relative humidity was kept in a desiccator containing copper sulfate and weight loss was observed over time. As we can observe in Figure 11, the weight gradually decreases and after about 1500 minutes the weight of the hemp fiber started to stabilize and after 7200 minutes the hemp fiber lost 4% of its original weight.

As we can see from Figure 11, the hemp fiber does not seem to have lost all of its moisture, it is necessary to expose it to higher temperatures to accurately measure its moisture content.

### 3.1. Thermal Insulation Of Hemp And Flax Fibers

It is observed that energy resources are gradually decreasing and energy consumption is increasing nowadays. Comfort and people's needs are increasing day by day, which leads to an increase in energy consumption. Most of the energy consumed in buildings comes from heating and cooling. Thermal insulation is one of the effective methods that can be used to reduce energy consumption in terms of both heating and cooling. In order to better reduce heating and cooling costs in buildings or vehicles, efficiency should be increased and provide more efficient heating with lower energy. This can only be achieved with thermal insulation. Thermal insulation is characterized by many situations. In thermal insulation, we have to consider heat permeability, specific heat, heat conduction resistance, heat conduction coefficient, relative humidity, absolute humidity, specific

humidity and condensation [30].

Granular and fibrous thermal insulation materials can be characterized as cellular reflective composites. In order for these cells to be dense, these cells should be as small as possible. Therefore, materials with a cellular structure can be preferred for thermal insulation. Another preferred material is the fiber ones. The fibrous materials have low density thanks to the width and number of free air channels. The air channels formed between the fibers create resistance to heat transfer that will occur by convection. Therefore, heat transfer by convection is minimal. Granular transfer heat insulation materials exist in the form of particles and there are air gaps between these particles. It provides heat insulation by convection on the same principle as the others. On the other hand, reflective thermal insulation materials are materials that absorb very little heat and reflect a large part of the heat to the outside.

Thermal insulation materials should have many features such as resistance to heat transfer, sufficient compressive and tensile strength, not deteriorating at use temperature, not reacting with the materials used, resistance to fire, resistance to water and moisture, lightness, etc.

Today it is waiting for different tasks from many materials. For example, a material should be able to provide thermal insulation while providing durability. For this reason, natural fiber materials that will provide thermal insulation are preferred in the production of composite materials. Using these natural fiber materials as thermal insulation materials is very important, both in terms of economy and carbon footprint. The carbon footprint of natural materials is very low compared to synthetic materials. In this period of climate and environmental problems, manufacturers should take many things into consideration while producing.

Flax and hemp fibers are among the fibers used in thermal insulation. One of the purposes of use of these fibers is their fiber structure. Flax and hemp consist of bast fiber. Each of the bast fibers are in bundles and are basic fibers consisting of 10 to 40 single cells. For flax and hemp, a single root contains 20 to 50 fiber bundles. These basic fibers are made up of layers and these layers have a porous structure called lumen.
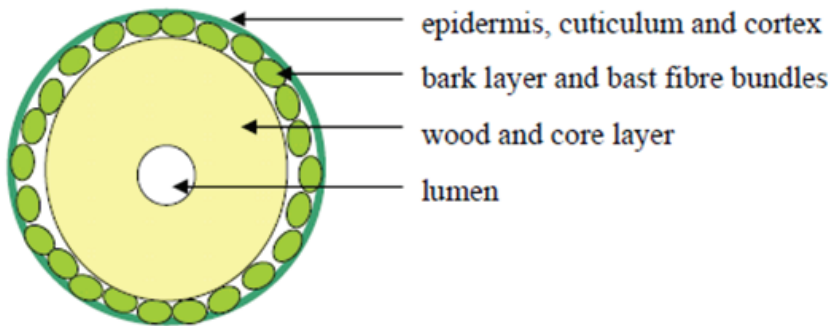
**Figure 12.** *Cross section of the flax*

For these two natural fibers, flax and hemp, their porous structure makes them suitable for thermal insulation. The thermal insulation of the fibers is compatible with traditional methods, but there is variation between the k values of all of them. For example, the relationship between density and k value is linear. We observe this in many studies. However, different k values and density values may occur between different studies. These natural fibers, which fulfill the main function of insulation thanks to bast fibers, cause a large amount of air to be absorbed through their pores.[31]
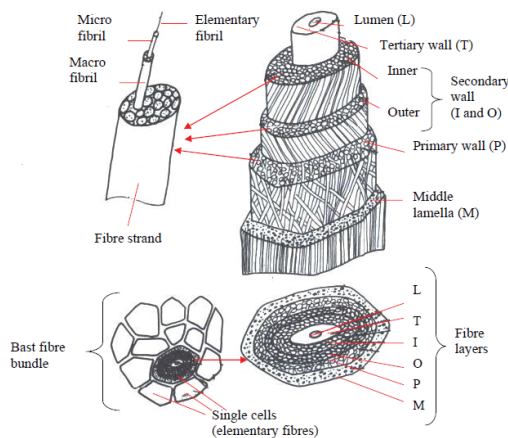


**Figure 13.** *Structure of the bast fibre bundle of flax*

Flax and hemp fibers with cellulosic structure are not very resistant to fire. Because cellulosic based materials are not resistant to fire due to chemical changes. For this reason, flame retardants should be used when using these materials as insulation materials.

In the table below, thermal conductivity values that provide thermal isolation for some jobs are given.

**Table 2.** *Variation ranges of thermal conductivity k (or l) of fibrous thermal insulations at different ranges of densities [31-32]*

| Type of insulation | Fibre raw material | Bulk density (kg/m³) | $k$ (or l) (W/m K) |
|---|---|---|---|
| — | Flax | — | 0.040–0.046 |
| Mat | Flax | 5–50 | 0.038–0.075 |
| — | Flax | 20–100 | 0.035–0.045 |
| — | Flax and hemp | 25–40 | 0.050 |
| Mat | Flax and hemp | 39 | 0.033 |
| Mat | Flax and hemp | 19 | 0.060 |
| Mat | Hemp, retted | 5–50 | 0.040–0.082 |
| Mat | Hemp, green | 5–50 | 0.044–0.094 |
| Loose-fill | Hemp, frost-retted | 25–100 | 0.040–0.049 |
| — | Hemp | 20–45 | 0.040–0.060 |

—       not provided.

If we consider this table, it is possible to obtain different k values at different density values. In other words, these fibers can be preferred as insulation material because they are economical and healthy.



**Figure 14.** *Thermal insulation applications with hemp fiber [33-35]*

## 4. ACOUSTIC PROPERTIES OF HEMP AND FLAX FIBERS

The phenomenon that changes shape in air, water or a similar environment, stimulates the sense of hearing and can be perceived by the ear is called sound. The human ear can hear sounds between 20 and 20000 Hertz. In order for a sound to be heard by the human ear, its intensity must be at a certain level. Sound is the main cause of noise. Noise is defined as the loud sound of many meaningless sounds to the human ear and any hearing aid. Sound insulation is needed to get rid of noise [36, 37].

Sound insulation is any way of reducing sound pressure relative to a particular sound source and receiver. If sound encounters a medium of different density or flexibility than the medium in which it travels in the form of waves, some of the energy is reflected and the other part is absorbed by transforming into heat energy. Non-reflected and non-absorbed sound waves complete their transition. By looking at the structure and design of the insulation materials that make up the outer shell of the building, the passage of external noise into the interior can be prevented or minimized. This measure is done in two ways as sound insulation and absorption [38].
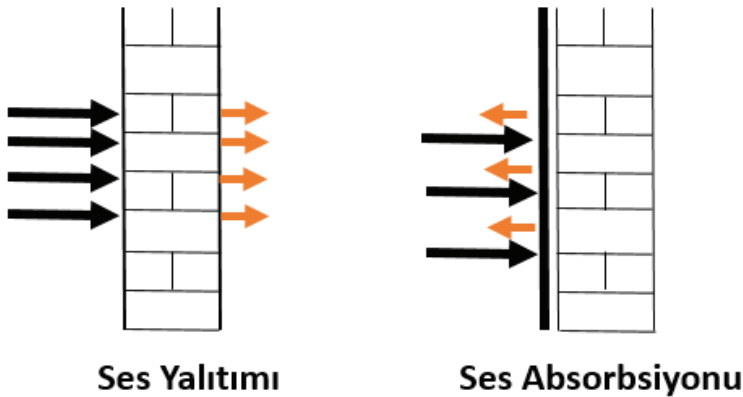


**Figure 15.** *Sound absorption and sound insulation*

The speed of sound depends on the properties of the medium. Sound propagates at different speeds in solid, liquid and gas environments. Since sound is a type of energy that spreads by vibration, it spreads faster in solids and liquids than in gas. Sound encompasses frequencies of vibration. Frequency is the number of vibrations per second of sound waves and its unit is hertz. The time it takes to vibration is called a period and its unit is a second. The distance between the peak and the trough is called the wavelength and its unit is meter. Low frequency sounds are characterized by low tones and high frequency sounds are described as thin sounds. Many

materials are used for sound insulation. The common features of these materials are their hollow structure. These materials can be in a fibrous structure or an open porous structure. In fibrous materials, the sound that hits the surface passes through the fiber and air section, converts some of its energy into heat energy by vibration and is damped. The finer and tighter the fibers, the greater the sound absorption coefficient.

There are many properties that determine the acoustic performance of the material such as phase density, thickness, elasticity, diffusion, surface geometry, porosity and air flow resistance. Having one of the solid, liquid and gas phases of the material will change the distance between atoms as well as change the friction coefficient, which is very important in terms of acoustic. The thicker the material, the higher the time the sound will pass through  and the greater the amount of energy converted into heat energy. The porous nature of the material affects sound absorption and transmission. The open and closed porous structure of the material is among the factors affecting sound transmission.

Hemp fiber is one of the materials that can be used as a sound insulation material. Like all natural fibers, hemp fiber is a natural fiber with a hollow structure and does not contain any risk factors for human health. It has a disadvantage due to its low fire resistance. In addition to having a very high acoustic character, its thermal insulation feature also adds a positive feature. In a study, it is stated that the sound absorption value of hemp fiber with a density of 164 kg/m$^3$ decreases to approximately 0.70 in the frequency range of 600 Hz and 1000 Hz, and reaches the highest sound absorption value of approximately 0.90 at a frequency of 2000 Hz (Figure 16-17) [39].



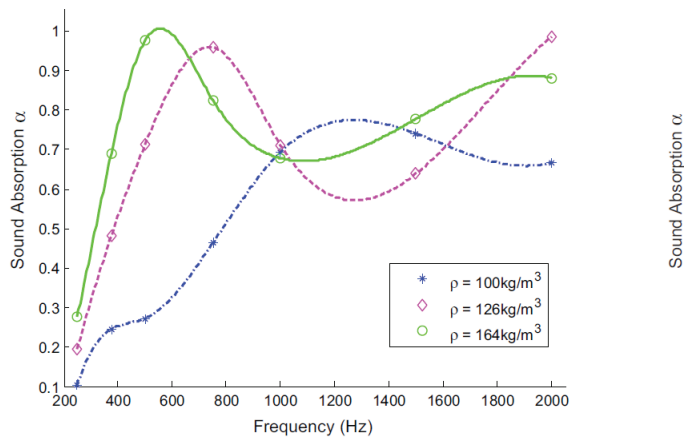**Figure 16.** *Hemp particles of different densities*

**Figure 17.** *Sound absoption characteristics of hemp for different density*

Flax fiber is a non-toxic fiber with similar acoustic properties. It can be used as a sound insulation material such as hemp fiber. Figure 18 shows that the sound absorption coefficient is about 0.80 at a frequency of 6000 Hz in a study for flax [40].
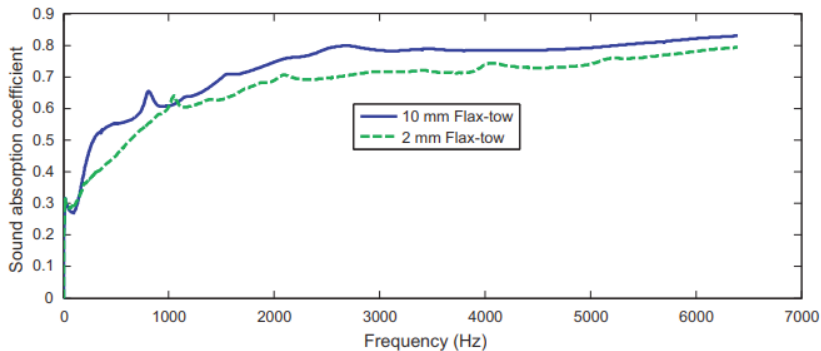


**Figure 18.** *Influence of the flax-tows size grindings on the acoustic absorption coefficient*

## 5. GENERAL COMPARISON BETWEEN NATURAL AND SYNTHETIC FIBERS

Considering the thermal and mechanical properties mentioned above, it will be possible to make a general comment about the fibers. Natural fibers have a lower density than glass and carbon fiber. Natural fibers are renewable, recyclable and require low energy for their production. They emit less carbon, do not pose a health risk and do not pollute the nature like synthetic fibers because they are biodegradable materials.

When hemp and flax fibers are compared with glass and carbon fibers in terms of mechanical properties, it can be said that some properties of flax and hemp are more successful than glass and carbon, although the strength of glass and carbon fiber is high. It has been determined that flax and hemp fibers have usable tensile strength, elastic modulus and density. Natural fibers have some advantages and disadvantages due to their structure. However, in terms of some properties, synthetic fibers are much more durable and useful. Natural fibers constitute an invaluable field of use for composite materials in terms of having sufficient mechanical and thermal properties, not harming the environment and being inexpensive. Table 3 shows comparisons of all these features mentioned.

**Table 3.** *Comparison of natural fibres, glass fibre and carbon fiber*

|  | Natural Fibers (NF) | Glass Fiber | Carbon Fiber |
|---|---|---|---|
| **Density** | Low | Twice that of natural fibers | Higher than natural fibers |
| **Cost** | Low | Low, but higher than NF | Higher than NF |
| **Renewability** | Yes | No | No |
| **Recyclability** | Yes | No | Yes |
| **Energy Consumption** | Low | High | High |
| **Distribution** | Wide | Wide | Wide |
| **CO$_2$ Neutral** | Yes | Yes | Yes |
| **Abrasion To Machines** | No | Yes | Yes |
| **Health Risk When Inhaled** | No | Yes | Yes |
| **Disposal** | Biodegradable | NOT Biodegradable | NOT Biodegradable |

Cellulosic based natural fiber materials differ from carbon and glass fiber in terms of absorbing moisture. The moisture they absorb can cause some defects during composite production.

Natural fibers contain cellulose, hemicellulose, lignin and pectin in their structure. Each of these add different features to the structures in which they are located. While pectin binds the fibers together, it also determines its high water holding capacity. Lignin is resistant to enzymes and allows the plant to stand upright above the soil level. Hemicellulose provides air permeability and heat insulation properties to the fiber.

Table 4 shows the comparison of the properties of flax and hemp fibers [41].

**Table 4.** *Comparison of the properties of flax and hemp*

| Property | Flax | Hemp |
|---|---|---|
| Cellulose content | 65-87%<br>(blanched, up to 98%) | under 80% at technical ripeness |
| Lignin content | small | greater than flax |
| Density | 1460 – 1500 kg.m$^{-3}$ | 1480-1500kg.m$^{-3}$ |
| Length of elementary fibers* | 3 – 60 mm | 4 – 55 mm |
| The shape of the cross-section by elementary fiber | 5- to 7-sided, with sharp peaks | a polygon with rounded peaks |
| Moisture | 12% | 13% |
| Fineness | 0.25 – 0.33 tex | 0.25 – 0.38 tex |
| Breaking length | 52 km | 30 – 50 km |
| Elongation at break | 1 – 2.5% dry<br>2 – 4% wet | 2% dry<br>4% wet |
| Elasticity | slight | slight |
| Shape of lumen | small, less apparent even at the dotted form | broad, dashed (rarely circular) |
| Ends of the elementary fibers | sharp | dull or forked |

* Quoted values may differ in specifically cited sources.

Flax and hemp fibers, which are very successful in terms of health, price and performance, can be used as heat insulation materials. Flax and hemp fibers are also used as acoustic insulation materials, especially in the automotive industry. These natural fibers have very successful acoustic values.

# REFERENCES

1) Callister W., "Materials science and engineering an introduction", 9th Edt., 565 (2014)

2) Kaw A.K., "Mechanics of Composite Materials", 2nd Ed. CRC Press, 3 (2006)

3) Sanjay M.R., Arpitha G.R & Yogeshaa B., "Study on mechanical properties of natural-glass fibre reinforced polymer hybrid composites: a review", Materials Today: Proceedings 2, 2959-2967 (2015)

4) Malkapuram R, Kumar V, Yuvraj S.N., "Recent development in natural fibre reinforced polypropylene composites. journal of reinforced plastics and composites", 1169-1189 (2008)

5) Li X, Tabil L G, Panigrahi S, Crerar W J., "the influence of fiber content on properties of injection molded flax fiber-hdpe biocomposites", Canadian Biosystems Engineering, 1-9 (2009)

6) https://tekstilbilgi.net/karbon-lifi-nedir.html, Date of Access: September 2021

7) http://www.tekniktekstiller.com/articles/cam-lifleri, Date of Access: September 2021

8) Harish S., Peter M.D., Bensely A., Mohan L.D., Rajadurai A., "Mechanical property evaluation of natural fiber coir composite", Materials Characterization, 60, 44-49 (2009)

9) Andrade S.F.,, Toledo F.R.D.,, Melo F.J.A., Moraes Rego F.E., "Physical and mechanical properties of durable sisal fiber-cement composites", Construct Build Material, 777-785 (2010)

10) Ramesh M., Palanikumar K., Reddy K.H., "Mechanical property evaluation of sisal–jute– glass

fiber reinforced polyester composites", Composites Part B: Engineering, 48, 1-9 (2013)

11) Saheb N.D., Jog J.P., "Natural fiber polymer composites: a review", Advanced in Polymer Technology, 18,4, 351-363 (1999)

12) https://www.kenevir.com/post/kenevir-kumas, Date of Access: September 2021

13) https://yetistir.net/keten-yetistiriciligi-ve-kullanim-alanlari/keten-lifi, Date of Access: September 2021

14) Crini G., Lichtfouse E., Chanet G., Crini N.M., "Applications of hemp in textiles, paper industry, insulationand building materials, horticulture, animal nutrition, foodand beverages, nutraceuticals, cosmetics and hygiene, medicine, agrochemistry, energy production and environment: a review", Environmental Chemistry Letters, 18,3, 1451–1476 (2020)

15) Fouad D., Farag M., "Design and manufacturing, Chapter 3: design for sustainability with biodegradable composites, 39-52 (2019)

16)     https://www.outsideonline.com/adventure-travel/advice/hemp-may-be-future-flying, Date of Access: September 2021

17)   https://www.popsci.com/cars/article/2011-02/introducing-first-road-ready-hemp-mobile, Date of Access: September 2021

18) Swarda Satish Radkar, Potential Applications Of Flax Fibers, A Thesis Submitted to the Graduate Faculty of the North Dakota State University of Agriculture and Applied Science Pages 1-77

19)     http://blog.europeanflax.com/flax-hemp-composite-innovation-sustainibility-2, Date of Access: September 2021

20) Kabir M.M., Wang H., Laua K.T., Cardona F., "Effects of chemical treatments on hemp fibre structure", Applied Surface Science, 276, 13-23 (2013)

21) Bismarck A., Mishra S., Lampke T., "Natural fibers, biopolymers, and biocomposites; Chapter: Plant fibers as reinforcement for green composites", 77 (2005)

22) Kaya S., Öner E., "Kenevir liflerinin eldesi, karakteristik özellikleri ve tekstil endüstrisindeki uygulamaları", Mehmet Akif Ersoy Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 11, 1, 108 – 123 (2020)

23) Gedik G., Avinç O.O., Yavaş A., "Kenevir lifinin özellikleri ve tekstil endüstrisinde kullanımıyla sağladığı avantajlar", Tekstil Teknolojileri Elektronik Dergisi, 4, 39-48 (2010)

24) Keller A., Leupin M., Mediavilla V., Wintermantel E., "Influence of the growth stage of industrial hemp on chemical and physical properties of the fibres", Industrial Crops and Products, 13, 1, 35-48 (2001)

25) Lu N., Robert H. Swan J., Ferguson I., "Composition, structure, and mechanical properties of hemp fiber-reinforced composite with recycled high-density polyethylene material", Journal of Composite Materials 46, 1915-1924 (2012)

26) Liu Y., Zwingmann B., Schlaich M., "Carbon fiber reinforced polymer for cable structures a review", Polymers 7 (10), 2078-2099 (2015)

27) Wambua P., Ivens J., Verpoest I., "Natural fibres: can they replace glass in fiber-reinforced plastics?", Composites Science and Technology, 63, 9, 1259-1264 (2003)

28) Prasad B. M., Sain M.M., "Mechanical properties of thermally treated hemp fibers in inert atmosphere for potential composite reinforcement," Materials Research Innovations, 7, 4, 231–238, (2003)

29) Shahzad A., "A study in physical and mechanical properties of hemp fibres, Advances in Materials Science and Engineering, 1-9 (2013)

30) Candan N., "Isı yalıtım sistemleri ve özelliklerinin karşılaştırılması", Sakarya Üniversitesi, Yüksek Lisans Tezi,  1-133 (2007)

31) Kymäläinen H.R., "Quality of Linum Usitatissimum L. (Flax and Linseed) And Cannabis Sativa L. (Fıbre Hemp) Durıng The Production Chain Of Fire Raw Material For Thermal Insulations", University of Helsinki, Doctoral Thesis  (2014)

32) Kymaalainen H.R., Sjoaberg A.M., "Flax and hemp fibres as raw materials for thermal insulations", Building and Environment , 43, 7, 1261-1269 (2008)

33)  https://www.eci.com.tr/productdetail/hemp-wool-insulation-batt,  Date  of Access: (2021)

34)        https://www.isohemp.com/en/building-hemp-blocks-insulating-and-efficient-envelope, Date of Access: (2021)

35) https://www.insulation-info.co.uk/insulation-material/hemp, Date of Access: (2021)

36) Kaya A.İ., Dalgar T., "Ses yalıtımı açısından doğal liflerin akustik özellikleri, Mehmet Akif Ersoy Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 8, Sayı Özel (Special) 1, 25- 37 (2017)

37) Santoni A., Bonfiglio P., Fausti P., Marescotti C., Mazzanti V., Mollica F., Pompoli F., "Improving the sound absorption performance of sustainable thermal insulation materials: Natural hemp fibres", Applied Acoustics, 150, 279-289 (2019)

38) Sagartzazu X., Hervella-Nieto L., Pagalday J.M., "Review in sound absorbing materials", Archives of Computational Methods in Engineering, 15, 311-342 (2008)

39) Kinnane O., Reilly A., Grimes J., Pavia S., Walker R., "Acoustic absorption of hemp-lime construction", Construction and Building Materials, 122, 674-682, (2016)

40)  Hajj N.E., Mboumba-Mamboundoua B., Dheillya R.M., Aboura Z., Benzeggagh M., Queneudec M., "Development of thermal insulating and sound absorbing agro-sourced materials from auto linked flax- tows", Industrial Crops and Products, 34,1, 921-928 (2011)

41) Wiener J., Kovačič V., Dejlová P., "Differences between flax and hemp, AUTEX Research Journal, 3, 2, 58-63 (2003)

# Chapter 11

## AN EXAMINATION OF VGGISH EMBEDDINGS USAGE IN ENVIRONMENTAL SOUND CLASSIFICATION

*Ilker Ali OZKAN*[1]

1    Department of Computer Engineering, Selcuk University, Konya, Turkey. E-mail: ilkerozkan@selcuk.edu.tr,   ORCID: 0000-0002-5715-1040

## Introduction

Environmental sounds usually include daily sound events in their structure. These sound events can occur from live sources such as cats, birds, bee sounds or non-living sources such as wind, rain, engine, siren. It does not consist of musical or speech data and has in its structure daily sound events which are often more diverse and complex. Unlike speech sounds, these sounds, which are made up of many different sources, also naturally have a noise  (Piczak, 2015b; Z. Zhang, Xu, Zhang, Qiao, & Cao, 2019). Analyzing images and videos is often used to describe objects or events in our environment. In addition to these, phonetic information includes a lot of information about our environment (X. Zhang, Zou, & Shi, 2017). By analyzing these environmental sounds, we can obtain information about many events in our environment.

Environmental Sound Classification (ESC) has an increasing interest by researchers (Dogan, Akbal, & Tuncer, 2020). In the classification of the sound, firstly, the process of extracting its features is done. In general, Mel Frequency Cepstrum Coefficients (MFCC) is considered a good manual feature extraction method for the classification of sound signals. Although MFCC is very good for the classification of speech signals, they do not fully reflect the characteristics of non-speech signals (Uzkent, Barkana, & Cevikalp, 2012). Depending on this, researchers continued to study over developing various feature extraction methods as well as MFCC features. For example, Bansal et al. stated that MFCC features are frequently used in the classification of environmental sounds. In their study, they have proposed EMD (empirical mode decomposition) as a new feature extraction approach. In the study that they did with different machine learning methods, they stated that the use of MFCC + EMD features together increases the performance on the classification of environmental sounds (Bansal, Shukla, Goyal, & Kumar, 2021). Chu et al. have proposed several features to fulfill the task of recognizing environmental sounds, including the popular MFCC which describes the spectral shape of sound. They also performed an empirical feature analysis for the characterization of the sound environment and proposed the Matching Pursuit (MP) algorithm to obtain effective time-frequency features. By combining the MP and MFCC features, they achieved an accuracy of 83.9% in distinguishing fourteen classes (Chu, Narayanan, & Kuo, 2009). Wang et al. have obtained the sound properties with spectrum centroid, spectrum spread, and spectrum flatness on a 12-class non-speech database. They achieved 85.10% accuracy in a new sound classification architecture that they obtained by using the hybrid support vector machine (SVM) and k-nearest neighbor (k-NN) delimiters (Jia-Ching Wang, Jhing-Fa Wang, Kuok Wai He, & Cheng-Shu Hsu, 2006).

In addition, deep learning methods have been used in the classification of environmental sounds in recent years. Convolutional neural networks (CNN), with their ability to learn spectro-temporal patterns, are well suited to the environmental sound classification problem (Salamon & Bello, 2017). In a study to evaluate whether convolutional neural networks can be successfully applied to environmental sound classification tasks, they determined that a convolutional model outperforms hand-crafted feature approaches (Piczak, 2015a). As a result of training the CNN model with 5-fold cross-validation on the ESC-50 dataset, it outperformed applications using manually designed features with 64.5% accuracy performance in the best network (Piczak, 2015a). In another study, AlexNet and GoogleNet convolutional neural networks, which are used in image recognition, were used to classify different image representations such as spectrogram obtained from environmental sounds. In the study conducted on three separate environmental datasets, the GoogleNet model provided the best classification performance. Researchers stated that GoogleNet has a deeper network architecture than AlexNet, which affects performance (Boddapati, Petef, Rasmusson, & Lundberg, 2017). Tokozume et al. proposed an end-to-end environmental sound classification system, which they named EnvNet, using CNN in their study. In their study, they compared the performances for different time periods ranging from 0.5 s to 5 s for the input length. They stated that the accuracy was at a high level in the T=1.0~2.5 sec range. In their study, they achieved a 6.5% improvement in classification performance as a result of combining EnvNet with logmel-CNN (Tokozume & Harada, 2017).

VGGish (Hershey et al., 2017) is a CNN-based model trained on audio from 8 million YouTube videos to distinguish 3,000 sound classes. The VGGish model, obtained by modifying VGG16, is widely used in sound recognition. This model creates a sound classification based on log-mel spectrograms and CNN. This deep learning model has six convolutional layers and three fully connected layers, respectively. A 128-dimensional feature can be obtained for each one second segment of the input sound.

In this study, both handcrafted and deep learning methods were used for feature extraction in the classification of environmental sounds. Thus, it is aimed to compare the embeddings obtained by using a previously trained deep learning model, VGGish, with the features obtained by MFCC. The k-NN type classifier, which is mostly used in solving classification problems and has many studies on sound classification, has been used in the classification of environmental sounds with the obtained features.

## Material and Methods

The study includes finding both the event that causes the environmental sound and the source group from which the environmental sound is derived. Therefore, two datasets were used. The first of these includes ten various environmental sounds. The second is categorized into five different groups according to the sounds coming from similar sources. Two methods were used to extract the features of the environmental audio signals and their performances were compared. These methods are MFCC and VGGish.

### *Environmental Sound Classification Dataset (ESC)*

In this study, the publicly available ESC-50 dataset was used (Piczak, 2015b). The ESC-50 dataset is used to compare the environmental sound classification methods. The ESC-50 dataset was created by combining short environmental sound recordings belonging to fifty different classes, and two thousand environmental sound recordings were labeled. All sound recordings were obtained from the freesound.org project, which is an open-access database. The audio clips are 5 seconds long and have a sampling rate of 44100 Hz. There are forty audio clips in each environmental sound class.

The first dataset used in the study is the ESC-10 dataset, which is a smaller subset of ESC-50 and includes ten selected classes. This dataset represents three common sound groups. The first is transient/percussive sounds, including sneezing, dog barking, clock ticking sounds. The second is sound events with harmonic content, including crying baby and crowing rooster sounds. The third is somewhat structural sound events that include the sounds of rain, sea waves, fire crackling, helicopter, chainsaws.

Fifty different classes in the ESC-50 dataset are given in Table 1. These classes are associated with five major categories. This categorically labeled dataset, which was created according to the origin of environmental sounds, is the second dataset used in this study.

**Table 1.** *Categorical representation of environmental sounds*

| Category | Environmental Sounds |
| --- | --- |
| Animals | *Crow, Sheep, Insects(flying), Hen, Cat, Frog, Rooster, Dog, Pig, Cow* |
| Natural soundscapes & water sounds | *Thunderstorm, Toilet flush, Rain, Pouring water, Sea waves, Wind, Crackling fire, Water drops, chirping birds, Crickets* |
| Human, non-speech sounds | *Crying baby, brushing teeth, Sneezing, Snoring Clapping, Breathing, Coughing, Drinking, Footsteps, Sipping, Laughing* |

| | |
|---|---|
| Interior/domestic sounds | *Washing machine, Mouse click, Door, wood creaks, Clock alarm, can opening, Clock tick, Vacuum cleaner, Glass breaking, Door knock, Keyboard typing,* |
| Exterior/urban noises | *Chainsaw, Hand saw, Siren, Engine, Train, Airplane, Church bells, Fireworks, Car horn, Helicopter,* |

The graph of sample environmental sound signals belonging to five major categories is given in Figure 1.
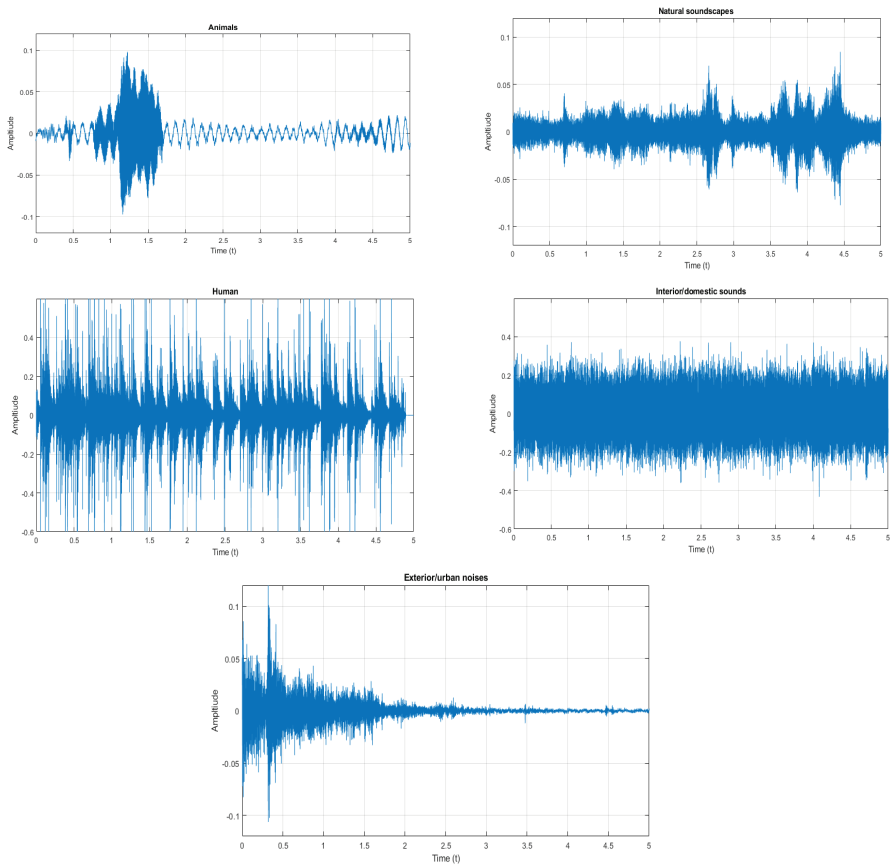


**Figure 1.** *Graphical representation of environmental sounds belonging to five major categories*

### Feature Extraction (MFCC)

Firstly, the Mel Frequency Cepstrum Coefficients (MFCC) method was used to obtain the feature values of the ESC. MFCC is a frequently used and high-performance feature in sound recognition applications. MFCC is based on human hearing perception. Human ear sensitivity is linear up to 1 kHz and logarithmic at higher values. Accordingly, there are

two types of filtering in the system with linear intervals below 1000 Hz and with logarithmic intervals above 1000 Hz. The features in the sound signal are captured as a result of this filtering. (Muda, Begam, & Elamvazuthi, 2010; Tombaloğlu & Erdem, 2016).

The conversion between the Mel scale and the frequency scale is provided by the equation in Equation 1.

$$Mel(f) = 2595x \log \left(1 + \frac{f}{100}\right)$$

(1)

MFCC is defined as the inverse Fourier transform of the logarithm of the time-dependent ESC(n) converted to the Mel scale Fourier transform. In this method, first, the Mel spectrum is obtained by multiplying the power spectrum of x(n) with a filter array arranged according to the Mel scale. Since the Mel spectrum does not contain complex numbers, it is sufficient to take the discrete cosine transform of the logarithm of the Mel spectrum to calculate the MFCC (Fig. 2).
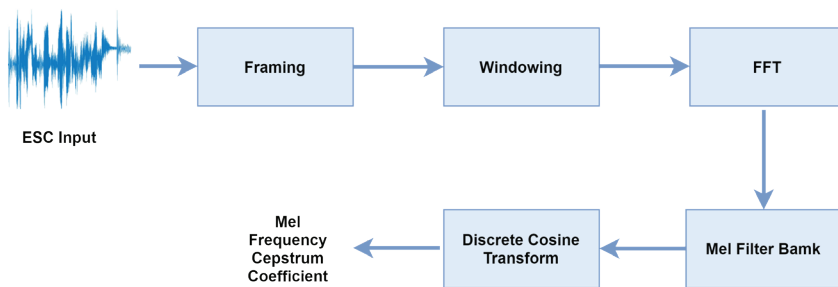


**Figure 2.** *MFCC Block Diagram*

As a result of the transformation, the MFCC are obtained. These coefficients express the acoustic features of the sound.

### Feature Extraction (VGGish Embeddings)

Due to the superior performance of deep networks in the sound classification task, the VGGish model, an 11-layer CNN pre-trained on millions of audio clips from the YouTube 8M dataset, was used in this study. This model has been modified from the VGG16 architecture to include log-Mel spectrogram-based inputs for feature extraction (Hershey et al., 2017).

In this study, firstly, all audio clips were resampled at 16 kHz. Each environmental sound signal is equally divided into five parts, and each has one second. Mel-Spectrograms are considered the input to the

model. A feature matrix of 64 × 96 Mel-Spectrograms is calculated for each one-second audio segment to match the input size of the VGGish model. Here 64 is the Melband number and 96 is the spectrum number in the individual Mel spectrograms. The output of the VGGish model is a 128-dimensional feature embed. VGGish is used to extract features from a 64x96 Mel-spectrogram. The output from VGGish is feature embeddings corresponding to each 0.975 ms audio data frame.

Figures 3-7 show a comparison of Mel-spectrogram and VGGish feature embedding visualizations for the Animals, Natural, Human, Interior, and Exterior major categories.
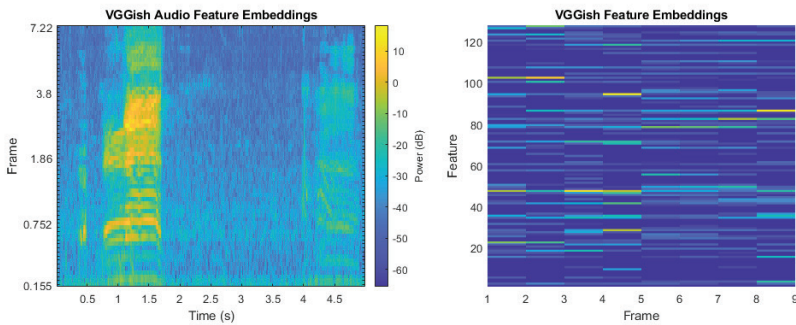


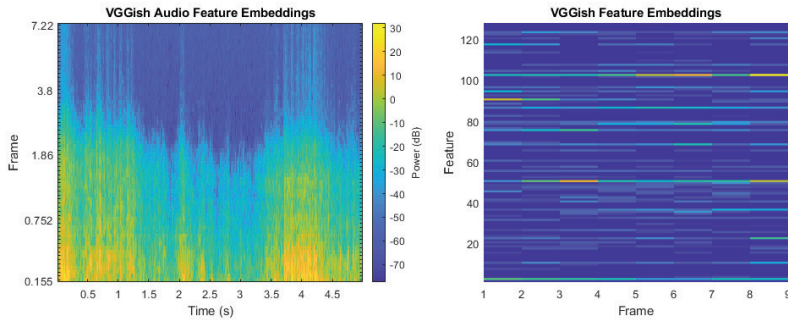**Figure 3.** *Visualizing of the VGGish embeddings and the Mel spectrogram in the Animals category.*



**Figure 4.** *Visualizing of the VGGish embeddings and the Mel spectrogram in the Natural category.*
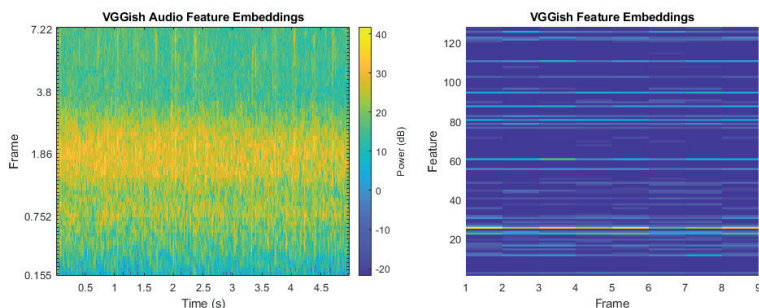
**Figure 5.** *Visualizing of the VGGish embeddings and the Mel spectrogram in the Human category.*
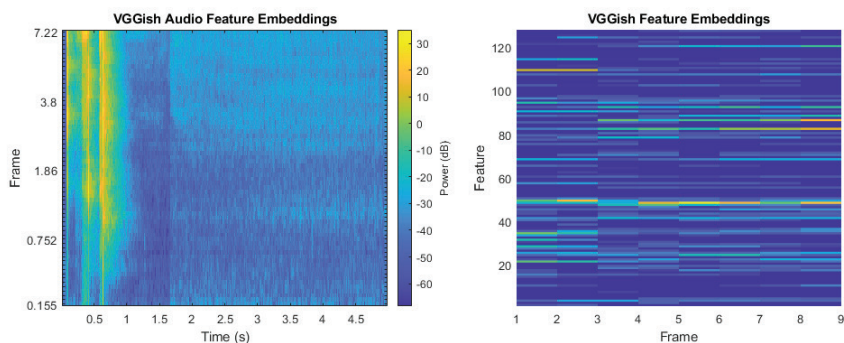


**Figure 6.** *Visualizing of the VGGish embeddings and the Mel spectrogram in the Interior category.*



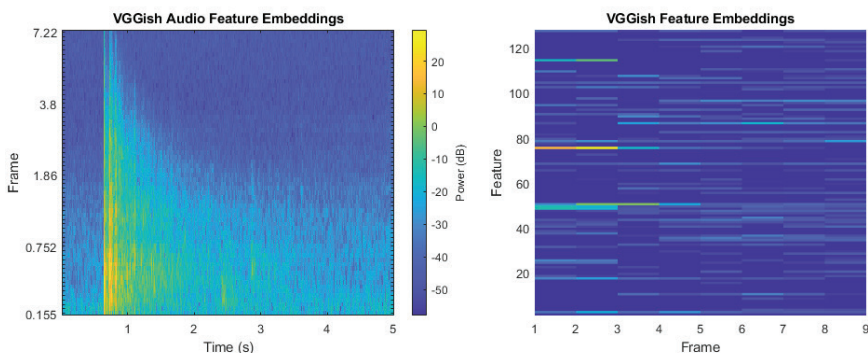**Figure 7.** *Visualizing of the VGGish embeddings and the Mel spectrogram in the Exterior category.*

### *k-NN Algorithm*

The k-nearest neighbor algorithm (k-NN) is one of the most well-known and used algorithms among machine learning algorithms. The output of K-NN classification is class membership. An item is classified by a majority vote of its neighbors. The item is assigned to the class

that is most common among its nearest neighbors (Sutton, 2012). When an unknown item is encountered, the k closest samples are determined from the training set and the class label of the new sample is assigned according to the majority vote of the class labels of its k nearest neighbors (Bhatia, 2010). If k = 1, the item is simply assigned to the class of its nearest neighbor. The distance between the item to be classified and each training items is calculated with the Euclidean equation given in Equation 2 (Wasule & Sonar, 2017).

$$D(Y,X_i) = \sqrt{\sum_{j=1}^{n} (y_j - x_{ij})^2}$$

(2)

In Figure 8, the item u is classified as class C since three of its neighbors are of class C.
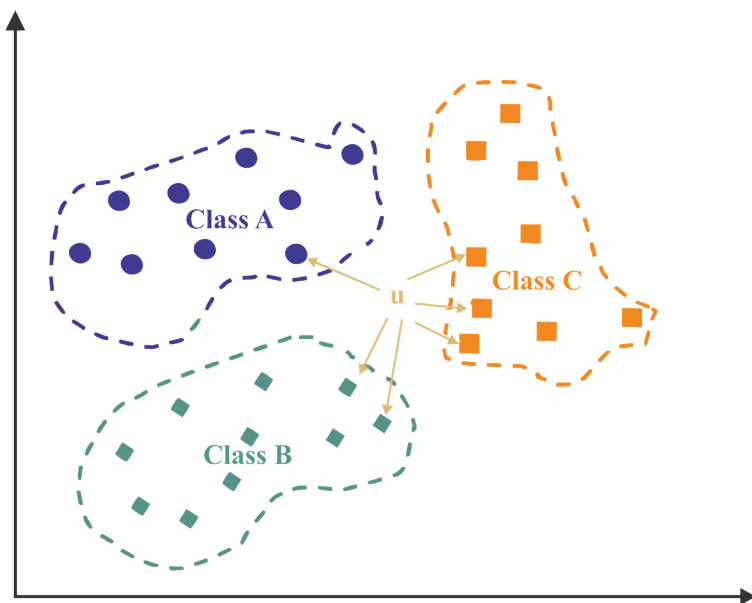


**Figure 8.** *Example of a three-class k-NN classifier*

In this study, trial-and-error method was applied to select the best k value and the k value was chosen as five.

### Results

In this study, the ESC-10 dataset containing audio data from ten different categories on the ESC-50 dataset and the environmental dataset belonging to five major categories were used in the classification of environmental sounds. MFCC feature extraction and the VGGish feature extraction process were applied on both datasets. 5-fold cross-validation

was applied to the obtained features. The accuracy rate of the model was used to measure the model performance. The k-NN classifier was used to classify the obtained features.

The MFCC coefficients obtained from the ESC-10 dataset were trained with the k-NN model, and the overall accuracy rate of the obtained model was obtained as 71.60%. The normalized complexity matrix of this model is given in Figure 9. In addition, Precision, Recall, and F1-Score values for each class of this MFCC feature-based model are given in Table 2.

**Table 2.** *Statistical values for each class in the MFCC-based ESC-10 dataset*

| Class | Precision | Recall | F1-Score |
|---|---|---|---|
| Dog | 0.51 | 0.81 | 0.63 |
| Rain | 0.64 | 0.90 | 0.75 |
| Sea waves | 0.60 | 0.91 | 0.72 |
| Baby cry | 0.90 | 0.61 | 0.73 |
| Clock tick | 0.86 | 0.83 | 0.85 |
| Person sneeze | 0.67 | 0.21 | 0.32 |
| Helicopter | 0.87 | 0.90 | 0.88 |
| Chainsaw | 0.71 | 0.74 | 0.73 |
| Rooster | 0.83 | 0.36 | 0.50 |
| Fire crackling | 0.84 | 0.89 | 0.87 |

**Model 1**

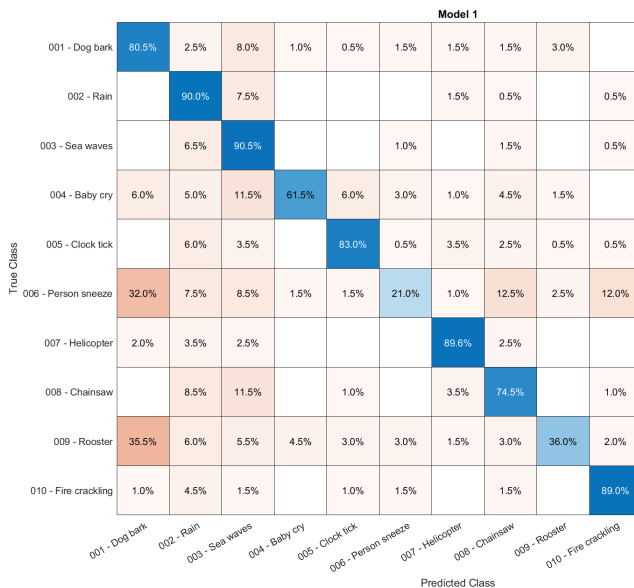| True Class \ Predicted Class | 001 - Dog bark | 002 - Rain | 003 - Sea waves | 004 - Baby cry | 005 - Clock tick | 006 - Person sneeze | 007 - Helicopter | 008 - Chainsaw | 009 - Rooster | 010 - Fire crackling |
|---|---|---|---|---|---|---|---|---|---|---|
| 001 - Dog bark | 80.5% | 2.5% | 8.0% | 1.0% | 0.5% | 1.5% | 1.5% | 1.5% | 3.0% | |
| 002 - Rain | | 90.0% | 7.5% | | | | 1.5% | 0.5% | | 0.5% |
| 003 - Sea waves | | 6.5% | 90.5% | | | 1.0% | | 1.5% | | 0.5% |
| 004 - Baby cry | 6.0% | 5.0% | 11.5% | 61.5% | 6.0% | 3.0% | 1.0% | 4.5% | 1.5% | |
| 005 - Clock tick | | 6.0% | 3.5% | | 83.0% | 0.5% | 3.5% | 2.5% | 0.5% | 0.5% |
| 006 - Person sneeze | 32.0% | 7.5% | 8.5% | 1.5% | 1.5% | 21.0% | 1.0% | 12.5% | 2.5% | 12.0% |
| 007 - Helicopter | 2.0% | 3.5% | 2.5% | | | | 89.6% | 2.5% | | |
| 008 - Chainsaw | | 8.5% | 11.5% | | 1.0% | | 3.5% | 74.5% | | 1.0% |
| 009 - Rooster | 35.5% | 6.0% | 5.5% | 4.5% | 3.0% | 3.0% | 1.5% | 3.0% | 36.0% | 2.0% |
| 010 - Fire crackling | 1.0% | 4.5% | 1.5% | | 1.0% | 1.5% | | 1.5% | | 89.0% |

**Figure 9.** *The complexity matrix obtained for the MFCC-based ESC-10 dataset*

The VGGish features obtained from the ESC-10 dataset were trained with the k-NN model and the overall accuracy rate of the obtained model was obtained as 93.20%. The normalized complexity matrix of this model is given in Figure 10. The Precision, Recall, and F1-Score values of this VGGish feature-based model are also given in Table 3 for each class.

**Table 3.** *Statistical values for each class in VGGish based ESC-10 dataset*

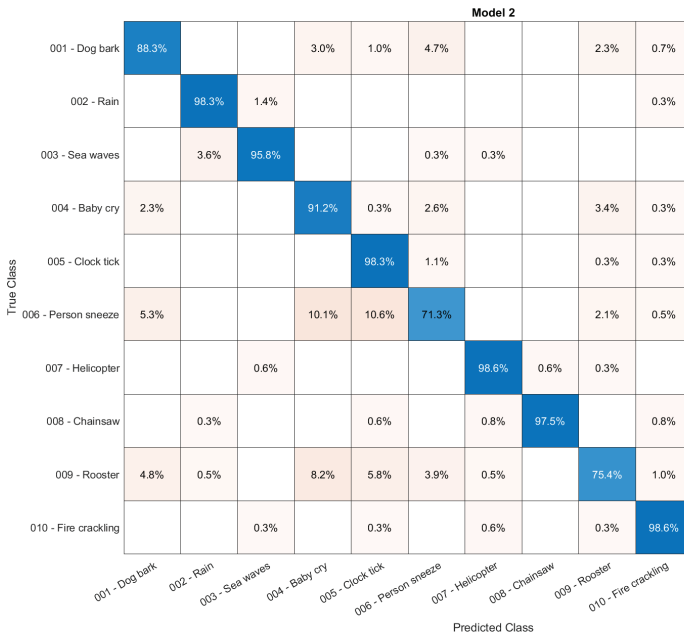| Class | Precision | **Recall** | **F1-Score** |
|---|---|---|---|
| Dog | 0.90 | 0.88 | 0.89 |
| Rain | 0.96 | 0.98 | 0.97 |
| Sea waves | 0.98 | 0.96 | 0.97 |
| Baby cry | 0.88 | 0.91 | 0.89 |
| Clock tick | 0.90 | 0.98 | 0.94 |
| Person sneeze | 0.79 | 0.71 | 0.75 |
| Helicopter | 0.98 | 0.99 | 0.98 |
| Chainsaw | 0.99 | 0.97 | 0.98 |
| Rooster | 0.86 | 0.75 | 0.80 |
| Fire crackling | 0.997 | 0.99 | 0.98 |



**Figure 10.** *The obtained complexity matrix for the VGGish-based ESC-10 dataset*

The MFCC coefficients obtained from the categorized ESC dataset were trained with the k-NN model and the overall accuracy rate of the obtained model was obtained as 68.20%. The normalized complexity matrix of this model is given in Figure 11. In addition, the Precision, Recall, and F1-Score values for each class of this MFCC feature-based model are given in Table 4.

**Table 4.** *Statistical values for each class in the MFCC-based categorized ESC dataset*

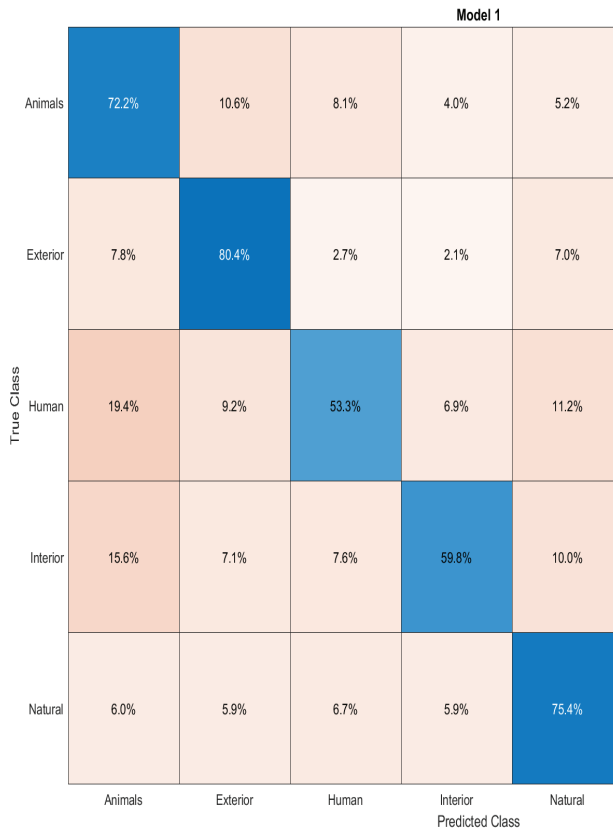| Class | Precision | Recall | F1-Score |
|---|---|---|---|
| Animals | 0.72 | 0.60 | 0.65 |
| Exterior | 0.80 | 0.71 | 0.75 |
| Human | 0.53 | 0.68 | 0.60 |
| Interior | 0.60 | 0.76 | 0.67 |
| Natural | 0.75 | 0.69 | 0.72 |



**Figure 11.** *The complexity matrix obtained for the MFCC-based categorized ESC dataset.*

The VGGish features obtained from the categorized ESC dataset were trained with the k-NN model and the overall accuracy rate of the obtained model was obtained 87.30%. The normalized complexity matrix of this model is given in Figure 12. In addition, Precision, Recall, and F1-Score values for each class of this MFCC feature-based model are given in Table 5.

**Table 5.** *Statistical values for each class in the VGGish-based categorized ESC dataset*

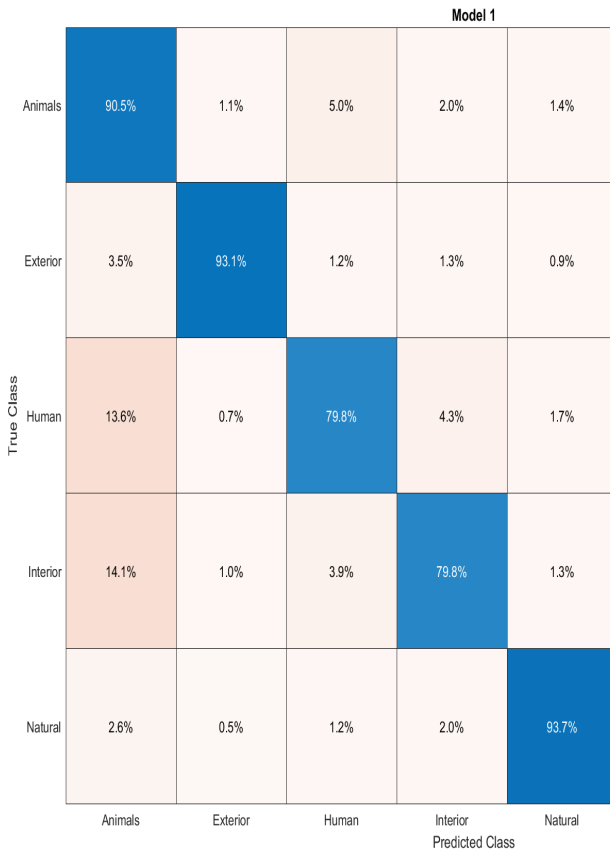| Class | Precision | Recall | F1-Score |
|---|---|---|---|
| Animals | 0.90 | 0.73 | 0.81 |
| Exterior | 0.93 | 0.97 | 0.95 |
| Human | 0.80 | 0.87 | 0.83 |
| Interior | 0.80 | 0.89 | 0.84 |
| Natural | 0.94 | 0.95 | 0.94 |



**Figure 12.** *The complexity matrix obtained for the VGGish-based categorized ESC dataset.*

With VGGish embeddings, a 21.60% better classification accuracy was achieved on the ESC-10 dataset and 19.10% better classification accuracy on the categorized ESC dataset. In environmental sound classification, VGGish embeddings outperforms MFCC feature extraction.

## References

Bansal, R., Shukla, N., Goyal, M., & Kumar, D. (2021). Enhancement and Comparative Analysis of Environmental Sound Classification Using MFCC and Empirical Mode Decomposition. In T. Senjyu, P. N. Mahalle, T. Perumal, & A. Joshi (Eds.), *Smart Innovation, Systems and Technologies* (Vol. 195, pp. 227–235). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-15-7078-0_21

Bhatia, N. (2010). Survey of nearest neighbor techniques. *ArXiv Preprint ArXiv:1007.0085*.

Boddapati, V., Petef, A., Rasmusson, J., & Lundberg, L. (2017). Classifying environmental sounds using image recognition networks. *Procedia Computer Science*, *112*, 2048–2056. https://doi.org/10.1016/j.procs.2017.08.250

Chu, S., Narayanan, S., & Kuo, C. C. J. (2009). Environmental sound recognition with timeFrequency audio features. *IEEE Transactions on Audio, Speech and Language Processing*, *17*(6), 1142–1158. https://doi.org/10.1109/TASL.2009.2017438

Dogan, S., Akbal, E., & Tuncer, T. (2020). A novel ternary and signum kernelled linear hexadecimal pattern and hybrid feature selection based environmental sound classification method. *Measurement: Journal of the International Measurement Confederation*, *166*. https://doi.org/10.1016/j.measurement.2020.108151

Hershey, S., Chaudhuri, S., Ellis, D. P. W., Gemmeke, J. F., Jansen, A., Moore, R. C., … Wilson, K. (2017). CNN architectures for large-scale audio classification. In *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings* (pp. 131–135). https://doi.org/10.1109/ICASSP.2017.7952132

Jia-Ching Wang, Jhing-Fa Wang, Kuok Wai He, & Cheng-Shu Hsu. (2006). Environmental Sound Classification using Hybrid SVM/KNN Classifier and MPEG-7 Audio Low-Level Descriptor. In *The 2006 IEEE International Joint Conference on Neural Network Proceedings* (pp. 1731–1735). IEEE. https://doi.org/10.1109/IJCNN.2006.246644

Muda, L., Begam, M., & Elamvazuthi, I. (2010). Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques. *ArXiv Preprint ArXiv:1003.4083*. Retrieved from http://arxiv.org/abs/1003.4083

Piczak, K. J. (2015a). Environmental sound classification with convolutional neural networks. In *IEEE International Workshop on Machine Learning for Signal Processing, MLSP* (Vol. 2015-Novem, pp. 1–6). https://doi.org/10.1109/MLSP.2015.7324337

Piczak, K. J. (2015b). ESC: Dataset for environmental sound classification. In *MM 2015 - Proceedings of the 2015 ACM Multimedia Conference*. https://doi.org/10.1145/2733373.2806390

Salamon, J., & Bello, J. P. (2017). Deep Convolutional Neural Networks and Data Augmentation for Environmental Sound Classification. *IEEE Signal Processing Letters*, *24*(3), 279–283. https://doi.org/10.1109/LSP.2017.2657381

Sutton, O. (2012). Introduction to k Nearest Neighbour Classification and Condensed Nearest Neighbour Data Reduction. *Introduction to k Nearest Neighbour Classification*, 1–10.

Tokozume, Y., & Harada, T. (2017). Learning environmental sounds with end-to-end convolutional neural network. In *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings* (pp. 2721–2725). https://doi.org/10.1109/ICASSP.2017.7952651

Tombaloğlu, B., & Erdem, H. (2016). Development of a MFCC-SVM Based Turkish Speech Recognition system. *2016 24th Signal Processing and Communication Application Conference, SIU 2016 - Proceedings*, 929–932. https://doi.org/10.1109/SIU.2016.7495893

Uzkent, B., Barkana, B. D., & Cevikalp, H. (2012). Non-speech environmental sound classification using SVMs with a new set of features. *International Journal of Innovative Computing, Information and Control*, *8*(5 B).

Wasule, V., & Sonar, P. (2017). Classification of brain MRI using SVM and KNN classifier. In *2017 Third International Conference on Sensing, Signal Processing and Security (ICSSS)* (pp. 218–223). https://doi.org/10.1109/SSPS.2017.8071594

Zhang, X., Zou, Y., & Shi, W. (2017). Dilated convolution neural network with LeakyReLU for environmental sound classification. In *International Conference on Digital Signal Processing, DSP* (Vol. 2017-Augus). https://doi.org/10.1109/ICDSP.2017.8096153

Zhang, Z., Xu, S., Zhang, S., Qiao, T., & Cao, S. (2019). Learning Attentive Representations for Environmental Sound Classification. *IEEE Access*, *7*, 130327–130339. https://doi.org/10.1109/ACCESS.2019.2939495