REPUBLIC OF TURKEY
ADANA ALPARSLAN TÜRKEŞ SCIENCE AND TECHNOLOGY UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING

CARBON DIOXIDE EMISSIONS PREDICTION USING META-HEURISTIC METHODS FOR RENEWABLE AND NON-RENEWABLE SOURCES' APPLICATIONS

İNAYET ÖZGE AKSU
DOCTOR OF PHILOSOPHY

ADANA 2022

REPUBLIC OF TURKEY
ADANA ALPARSLAN TÜRKEŞ SCIENCE AND TECHNOLOGY UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING

CARBON DIOXIDE EMISSIONS PREDICTION USING META-HEURISTIC METHODS FOR RENEWABLE AND NON-RENEWABLE SOURCES' APPLICATIONS

İNAYET ÖZGE AKSU
DOCTOR OF PHILOSOPHY

SUPERVISOR
ASSOC. PROF. DR. TUĞÇE DEMİRDELEN

ADANA 2022

ABSTRACT

CARBON DIOXIDE EMISSIONS PREDICTION USING META-HEURISTIC METHODS FOR RENEWABLE AND NON-RENEWABLE SOURCES' APPLICATIONS

İnayet Özge AKSU
Department of Electrical and Electronic Engineering
Supervisor: Assoc. Prof. Dr. Tuğçe DEMİRDELEN
April 2022, 93 pages

The Shuffled Frog-Leaping Algorithm (SFLA) is a nature-inspired and swarm-based metaheuristic algorithm proposed in 2003. SFLA is an optimization algorithm that deals with the movements of frogs in a given population to reach the maximum amount of food with minimum movement. In this thesis, as a first phase, a hybrid method that increases the speed of reaching the optimum solution of the SFLA and minimizes the possibility of getting stuck in the local minimum is proposed. In the proposed hybrid method, the search capability of the SFLA algorithm is increased at first time with the Levy flight function. Then, the developed method (ILSFLAFA) is performed in a hybrid structure by using the Firefly Algorithm (FA) and the success of this proposed method is proven with the Benchmark functions.

As a second phase, a new estimation method, that involves a multilayer feed forward artificial neural network, is proposed. The artificial neural network is trained with the ILSFLAFA. This newly developed hybrid-swarm-based artificial neural network (HSBNN) estimation method has been used to estimate the CO^2 emission amount of Türkiye. The developed method, which is proposed in the literature as a first time, is compared with the estimation study conducted in 2016 on this issue, and the success and effect of the proposed method are demonstrated.

Then, a new estimation method has been developed for Türkiye's CO^2 emission estimation problem using the proposed HSBNN. At this stage, in order to find the best model, four different forecasting models are developed and the results are examined. The input variables of year, R&D investments, renewable energy ratio of total final energy consumption, population, urbanization, number of motor vehicles, energy consumption and GDP variables are used determining optimal parameters of artificial neural network (ANN). This data set for Türkiye is created by collecting data from different institutions.

In the first prediction model, one layer is used in the hidden layer. The number of neurons in the layers of this method is studied as 8-12-1. In the other three estimation methods, two layers are used in the hidden layer. The number of neurons belonging to the second prediction model is 8-6-4-1. The neuron numbers of the artificial neural network used in the third and fourth prediction models are 8-6-6-1 and 8-8-4-1. In this study, which is conducted with the data of Türkiye's 1990-2018 years, the data set of the years 1990-2012 is used to create the model, while the data set of the years 2013-2018 is used in the testing phase of the model. The results obtained at this stage of the study are given in detail and the best estimation model is selected and the next step is taken.

In the last stage of the thesis, Türkiye's future CO^2 emission estimation is examined in three different scenarios. These scenarios developed; the estimates published by official institutions, the increase rates suggested in the literature studies and the increase rates in Türkiye in recent years have been created. Türkiye's future CO^2 emission estimation is made towards 2030. The results obtained at the end of the study are discussed in detail and the effect of renewable energy use on the amount of CO^2 emissions is emphasized. In addition, the obtained results show that the proposed estimation method can be successfully applied in areas requiring future estimation.

Keywords: renewable energy, CO^2 emission, neural networks, shuffled frog-leaping algorithm, Levy flight, firefly algorithm, hybrid structure, sustainability, green technology

ÖZET

YENİLENEBİLİR VE YENİLENEMEYEN ENERJİ KAYNAKLARININ UYGULAMALARI İÇİN META-SEZGİSEL YÖNTEMLER KULLANILARAK KARBONDİOKSİT EMİSYONLARI TAHMİNİ

İnayet Özge AKSU
Elektrik - Elektronik Mühendisliği Anabilim Dalı
Danışman: Doç. Dr. Tuğçe DEMİRDELEN
Nisan 2022, 93 sayfa

Karışık Kurbağa Sıçrama Algoritması (SFLA), 2003 yılında önerilmiş doğa esinli ve sürü tabanlı bir metasezgisel algoritmadır. SFLA, belli bir popülasyondaki kurbağaların maksimum miktarda besine, minimum hareketle ulaşabilmesi için yaptıkları hareketleri konu alan bir optimizasyon algoritmasıdır. Bu tez çalışmasında ilk olarak, SFLA'nın optimum çözüme ulaşma hızını arttırmak ve lokal minimuma takılma ihtimalini en aza indirmek için hibrit bir yöntem önerilmiştir. Önerilen bu hibrit yöntemde ilk olarak, SFLA algoritmasının arama kabiliyeti Levy uçuş fonksiyonu ile arttırılmıştır. Daha sonra geliştirilen yöntem (ILSFLAFA), Ateşböceği Algoritması (FA) ile hibrit yapıda çalıştırılmıştır ve önerilen bu yöntemin başarısı Benchmark fonksiyonları ile ispatlanmıştır.

Tez çalışmasının ikinci aşamasında ise yeni bir tahmin yöntemi önerilmiştir. Bu tahmin yönteminde çok katmanlı ileri beslemeli yapay sinir ağı kullanılmıştır. Yapay sinir ağı, ILSFLAFA yöntemi ile eğitilmiştir. Geliştirilen bu yeni hibrit-sürü tabanlı yapay sinir ağı (HSBNN) tahmin yöntemi, Türkiye'nin CO^2 emisyon miktarının tahmini için kullanılmıştır. İlk olarak geliştirilen yöntem, 2016 yılında bu konuda yapılan tahmin çalışması ile karşılaştırılmış ve önerilen yöntemin başarısı ve etkisi gösterilmiştir.

Daha sonra, önerilen HSBNN kullanılarak Türkiyenin CO^2 salınımı tahmini problemi için yeni bir tahmin yöntemi geliştirilmiştir. Bu aşamada en iyi modeli bulabilmek için dört farklı tahmin modeli geliştirilmiş ve elde edilen sonuçlar incelenmiştir. Yapay sinir ağının (YSA) optimal parametreleri belirlenirken girdi değişkenleri olarak; yıl, Ar-Ge yatırımları, yenilenebilir enerjinin toplam nihai enerji tüketimi oranı, nüfus, kentleşme, motorlu taşıt

sayısı, enerji tüketimi ve GSYİH kullanılmıştır. Türkiye'ye ait bu veri seti, farklı kurumlardan alınan veriler bir araya getirilerek oluşturulmuştur.

Birinci tahmin modelinde, gizli katmanda bir katman kullanılmıştır. Bu yöntemin katmanlarındaki nöron sayıları 8-12-1 olarak çalışılmıştır. Diğer üç tahmin yönteminde gizli katmanda iki katman kullanılmıştır. İkinci tahmin modeline ait nöron sayıları 8-6-4-1'dir. Üçüncü ve dördüncü tahmin modelinde kullanılan yapay sinir ağına ait nöron sayıları 8-6-6-1 ve 8-8-4-1'dir. Türkiye'nin 1990-2018 yıllarına ait verileri ile yapılan bu çalışmada, 1990-2012 yıllarına ait veri seti modelin oluşturulması için kullanılırken, 2013-2018 yıllarına ait veri seti modelin test aşamasında kullanılmıştır. Çalışmanın bu aşamasında elde edilen sonuçlar detaylı şekilde verilmiş ve en iyi tahmin modeli seçilerek bir sonraki aşamaya geçilmiştir.

Tez çalışmasının son aşamasında ise Türkiye'nin ileriye yönelik CO^2 emisyon tahmini, üç farklı senaryoda incelenmiştir. Geliştirilen bu senaryolar; resmi kurumların yayınladığı tahminler, litaretür çalışmalarında önerilen artış oranları ve Türkiyeye ait son yıllardaki artış oranları göz önünde bulundurularak oluşturulmuştur. Türkiye'nin ileriye yönelik CO^2 emisyon tahmini 2030 yılına kadar yapılmıştır. Çalışma sonunda elde edilen sonuçlar detaylı olarak tartışılmıştır ve yenilenebilir enerji kullanımının CO^2 emisyon miktarındaki etkisi vurgulanmıştır. Ayrıca elde edilen sonuçlar, önerilen tahmin yönteminin ileriye yönelik tahmin gerektiren alanlarda başarılı bir şekilde uygulanabileceğini göstermektedir.

Anahtar Kelimeler: yenilenebilir enerji, CO^2 emisyonu, yapay sinir ağları, karışık kurbağa sıçrama algoritması, Levy uçuş fonksiyonu, ateşböceği algoritması, hibrit yapı, sürdürülebilirlik, yeşil teknoloji

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

ANN	: Artificial Neural Network
CEEFS	: Carbon Emissions Ensemble Forecasting System
CH4	: Methane
CO^2	: Carbon Dioxide
CS	: Cuckoo Search
CSO	: Chicken Swarm Optimization
EU	: European Union
FA	: Firefly Algorithm
FE	: Fitness Evaluation
FLN	: Fast Learning Network
GDP	: Gross Domestic Product
GM	: Gray Model
GRNN	: General Regression Neural Network
GW	: Gigawatt
HEPP	: Hydroelectric Power Plant
HFC	: Hydrofluorocarbons
HSBNN	: Hybrid Swarm-Based MLNN
ILSFLAFA	: Improved Levy Flight Shuffled Frog Leaping Algorithm Based on Firefly Algorithm
IWO	: Invasive Weed Optimization
MLNN	: Multi-Layer Neural Network
MLP	: Multi-Layer Perceptron
Mt	: Million Tons
Mtoe	: Million Tons of Oil Equivalent
MW	: Megawatt
N2O	: Nitrogen Oxide
O3	: Ozone
OECD	: Organisation for Economic Co-operation and Development
PFC	: Perfluorocarbons
PSO	: Particle Swarm Optimization
R&D	: Research and Development

RES	: Renewable Energy Sources
SF6	: Sulfur Hexafluoride
SFLA	: Shuffled Frog Leaping Algorithm
SPP	: Solar Power Plant
UNFCCC	: United Nations Framework Convention on Climate Change
WPP	: Wind Power Plant

1. INTRODUCTION

1.1. Global

Ensuring environmental sustainability is among the important elements of the millennium development goals, which are created at the beginning of the concept of sustainability and entering the third millennium. However, until the adoption of the sustainable development guide (Paris Agreement), there is no research in this area. The Paris Agreement (Paris Agreement) is signed in December 2015 and the efforts made until this year are unsuccessful and the dangers of this are explained, and then in 2019, the European Green Deal is announced and became a law.

The essential element is the spread of a new economic model, that has started in Europe, to the world. The first of the three most important elements of this model are zeroing by reducing emissions, the second is clean, reliable energy that people can buy for purchasing power, and the other is sustainable transportation. “Fit For 55” is announced on 14 July 2021. The EU’s goal of lowering net greenhouse gas emissions by at least 55 percent by 2030 is referred to as “Fit for 55”. According to this program, which is first implemented by the European Green Deal, it is stated that emissions should be reduced by 55% by 2030. Many researchers argue that this ratio is not enough, and the main target should be a reduction of 1.5 degrees. However, for a reduction of 1.5 degrees, there must be a 65% reduction in emissions. For this purpose, carbon pricing systems have come to the fore. With carbon trading, a punishment system has been introduced for those with high emissions.

1.2. Türkiye

The European Green Deal process, which started on 11 December 2019, is formed as a result of many stages and regulations. When the “Fit For 55” package is examined, it is seen that the first stage is emission zero until 2050. As a first step towards this, the EU set the 2030 target and aimed to reduce carbon emissions by 55 percent until this date, and carbon border adjustment mechanism to the fore.

Carbon border adjustment mechanism, which is a package that includes many sub-packages, is the most important issue that will affect both the industrialists in the EU and the companies exporting to the EU and having a trade relationship. The content of this system, whose purpose is to reduce carbon leakage in imports within the EU, is briefly the pricing of the carbon content of the products produced by the industrialists in the EU.

There are different options determined for the carbon border adjustment mechanism, and the certification system has an important place among these options. One of the two possible options to be selected is the greenhouse gas generated during the production process in the company, and the other includes the greenhouse gas emissions generated in the entire life cycle of the product.

An action plan is prepared right after the global studies in Türkiye, but the projects are shelved before they are complied with. However, it is necessary to ensure the permanence of these rapid studies. Since our biggest export route is the European Union, taxation according to emission amounts has become very important for our country.

Türkiye’s efforts towards these regulations are as follows:

- Presidential circular,
- The Green Deal Action Plan arranged by the Ministry of Trade,
- Medium Term Program (between 2022-2024),
- Approval of the Paris Agreement at The Grand National Assembly of Türkiye,
- Name change of the Ministry of Environment and Urbanization.

The studies on “sustainable development” carried out to date reveal how important the issue is. These studies cover two main issues. These are sustainability and pollution. When evaluated on the basis of sustainability, it is seen that one aspect of this concept is actually related to economic development. In the Official Gazette of the Republic of Türkiye regarding the Green Deal, it is aimed to ensure the adaptation of Türkiye’s policies to combat climate change, which has gained momentum on international trade in recent years. In addition, the Green Deal Action Plan, which is a roadmap that will strengthen Türkiye’s competitiveness in exports, has been published.

If Türkiye cannot adapt to this process, it will lose its competitiveness and decrease its market share. Studies suggest that this may affect GDP. This situation has the power to seriously affect exports.

1.3. Energy

Energy is the most important requirement for the infrastructure of industrialization and the survival of people. Since the Industrial Revolution began in the 18th century, the rate of carbon dioxide in the atmosphere has increased by 40% compared to those years. As a result of the industrialization process, an increase in the use of fossil fuels has created problems about global climate change. Climate change, as a result of the greenhouse effect caused by the gases released into the atmosphere, brings along the increase in the average temperatures measured in the world, sea and air throughout the year, in other words, global warming. Reducing the use of fossil fuels and prioritizing renewable energy sources and more efficient production technologies are among the most important steps to be taken in preventing climate change. With the use of renewable energy sources, it is possible to reduce the carbon footprint resulting from global energy use. In the climate change framework convention, Türkiye is counted as group Annex I because of being among OECD member countries. The countries in this group are responsible for limiting greenhouse gas emissions, protecting and improving greenhouse gas sinks, as well as reporting the measures and policies they have taken to prevent climate change. Türkiye is among risky countries in terms of the devastating consequences of global warming. In the long term, the use of renewable energy sources will play a key role in supporting the sustainable economy and reducing environmental impacts. Considering the advantages of Türkiye's geographical location and the fastest-growing energy market among the OECD countries, Türkiye is a pretty rich country in terms of wind and solar energy potential. Therefore, this chapter aims to analyze the role of Türkiye's solar and wind power plants in preventing carbon emissions compared to fossil fuels. Thus, the importance of the use of renewable energy sources in terms of demonstrating the benefits of sustainable development, preventing carbon emissions and reducing the impact of climate changes will be revealed.

Energy is accepted as an indicator of the development of countries against other countries. For this reason, when comparing the development levels of countries, the amount of energy they produce and consume is examined. Today, it can be said that the indicator of the level of welfare that developed societies have reached is the amount of energy that countries have. Energy, which plays an important role in the development and development of countries, has become the most strategic factor in international status. In order to maintain the level of development reached and to provide the demanded energy, countries make various investments and research and development studies. Having energy, which is an indicator of economic superiority for countries, increases the quality of life of societies. Changes in demographic characteristics directly change the energy demand. It has become a necessary consumption item not only for industrialization, but also for social life, as it is necessary for the continuation of the operation of industrial facilities.

As a result of climate change, there are numerous negative effects on food, health, water security and biodiversity. One of the most obvious consequences of climate change is the occurrence of extreme weather events. Today, all these negative effects are clearly observed. It is known that the greenhouse gas emissions that cause climate change are mostly caused by developed countries. However, the countries where the effects of climate change are seen the most are underdeveloped and developing countries. Even if the emission of all greenhouse gases in the world is stopped as of today, the effects of climate change, which is already experienced as an accumulation of old years, will continue to be seen in the next decades. It will be important for countries to come together and create sufficient climate finance to combat the climate crisis in the coming years. As a result, it is necessary to initiate initiatives to combat global warming on a global scale. The frightening level of climate change puts the lives of all living things at risk. As a result of global warming, some plant and animal species are in danger of extinction. Climate change has become a fundamental environmental problem because it threatens the life of all human beings.

The emission of carbon dioxide gas, which is released as a result of individual or institutional use of fossil fuels such as oil, natural gas, coal, into the atmosphere is called carbon emission. Air pollution occurs as a result of the release of this gas into the environment. In previous years, the air pollution caused by carbon dioxide gas is not considered as important by societies as it is now. However, today, with the rapid increase in air and environmental pollution, carbon emission has become a very important issue. Climate change has occurred on a worldwide scale, particularly in recent years. Global warming is the name given to these changes. According to scientific studies, CO^2-based greenhouse gases are the primary cause of global warming. Carbon dioxide in the atmosphere, as is well known, is predominantly created as a result of the use of fossil fuels in many fields and is rapidly rising. Carbon dioxide (CO^2) levels in the atmosphere began to rise rapidly after the Industrial Revolution, as a result of the usage of carbon-based fossil fuels. CO^2 is released into the atmosphere mostly as a result of industrial usage of fossil fuels, and to a lesser extent as a result of the respiration of living things on the ground and microscopic organism decomposition of organic materials. CO^2 contributes 50% of global warming among greenhouse gases. The high amount of CO^2 molecules in the atmosphere, as well as their extremely long life in the atmosphere, are the reasons for this. Therefore, one of the measures to be taken against global warming is to reduce CO^2 emissions. In this direction, extraordinary efforts are being made at the international level. CO^2 gas emission is one of the most common climate change factors on a global scale. This emission value promotes global warming and thus climate change since it heats the air in the atmosphere and elevates the temperature.

Most of the carbon dioxide (CO^2) is produced by just four countries, China, the USA, India, Russia and the European Union. The Paris Agreement, which sets the foundation for the post-2020 climate change regime, is approved at the United Nations Framework Convention on Climate Change's 21st Conference of the Parties in Paris in 2015, and it went into effect in 04.11.2016. The goal of the agreement is to keep the average global temperature rise to 1.5 °C higher than it is before the industrial revolution. Today's temperature has risen by 1.1 degrees Celsius. CO^2 emissions should be decreased by 45 percent in 2030 compared to 2010, and net zero emissions should be achieved by 2050 to avoid exceeding the 1.5°C limit. Different approaches are used by countries and regions in order for international climate policy to achieve their objectives. The EU stands out with the plans and policies it has implemented in the last 5 years. In March 2020, the Climate Law, which will make the net zero emission target legally binding in the EU, was submitted to the EU Parliament by the European Commission and the law proposal was approved by the Commission on 24.06.2021 and entered into force on 09.07.2021. Furthermore, the European Green Deal is implemented with the goal of reducing greenhouse gas emissions, creating new job opportunities, implementing circular economy models, and improving living standards. The European Green Deal is characterized as the EU's new growth strategy, designed and managed by the European Commission, and comprises the core objectives of zeroing net greenhouse gas emissions by 2050, ending economic growth's reliance on resource usage, and leaving no one and no place behind. The agreement is followed with great interest not only by the member states, but also by all third-party countries that have close commercial and political relations with the EU. Many countries, like Türkiye, face both threats and opportunities as a result of climate and environmental change.

With the increase in industrialization and technological developments in Türkiye, CO^2 emissions have increased significantly. When the results of the greenhouse gas inventory are examined, it is seen that, overall greenhouse gas emissions in 2019 are 506.1 million tons (Mt) CO^2 equivalent, down 3.1 percent from the previous year. Total and per capita greenhouse gas emissions in Türkiye, 1990-2019 is shown in Figure 1. The total greenhouse gas emission per capita is 4 tons of CO^2 equivalent in 1990, and 6.4 tons of CO^2 equivalent in 2018. and 6.1 tons of CO^2 equivalent in 2019. is calculated as. When greenhouse gas emission rates are examined according to gases in 2019, it is observed that CO^2 gas has the highest rate with a rate of approximately 79%. In the second place, CH^4 with a rate of 12%, followed by N2O and F-gases, constitute greenhouse gas emissions. Greenhouse gas emissions rates by sector and gases are presented in Figure 2.

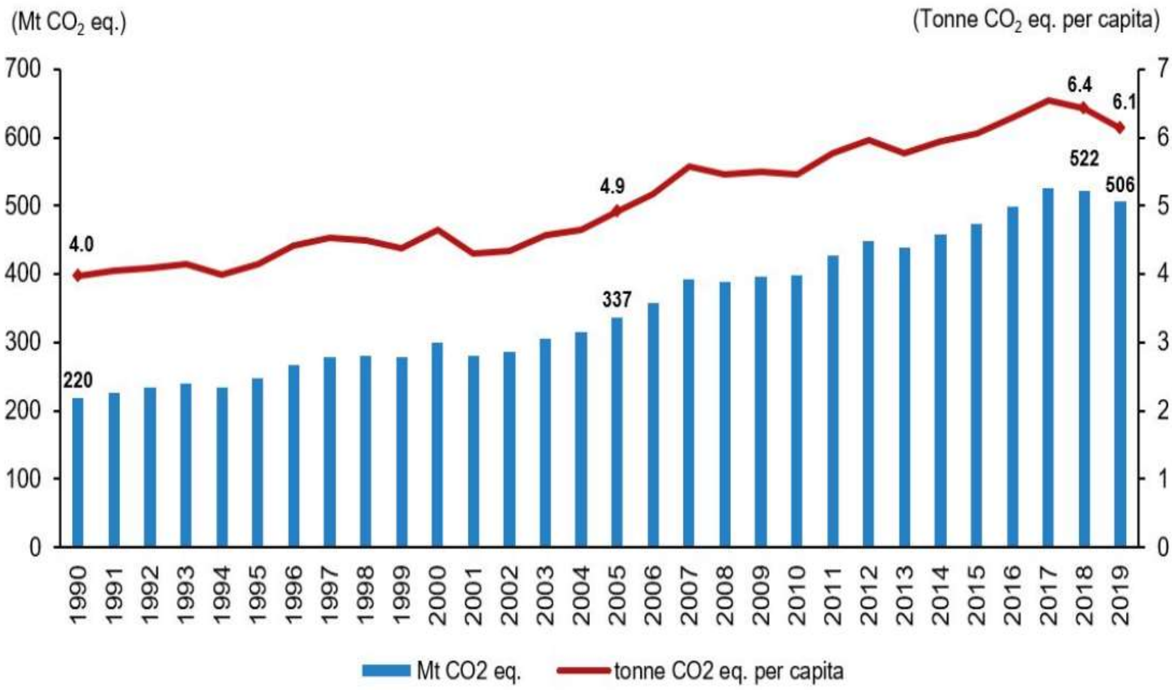
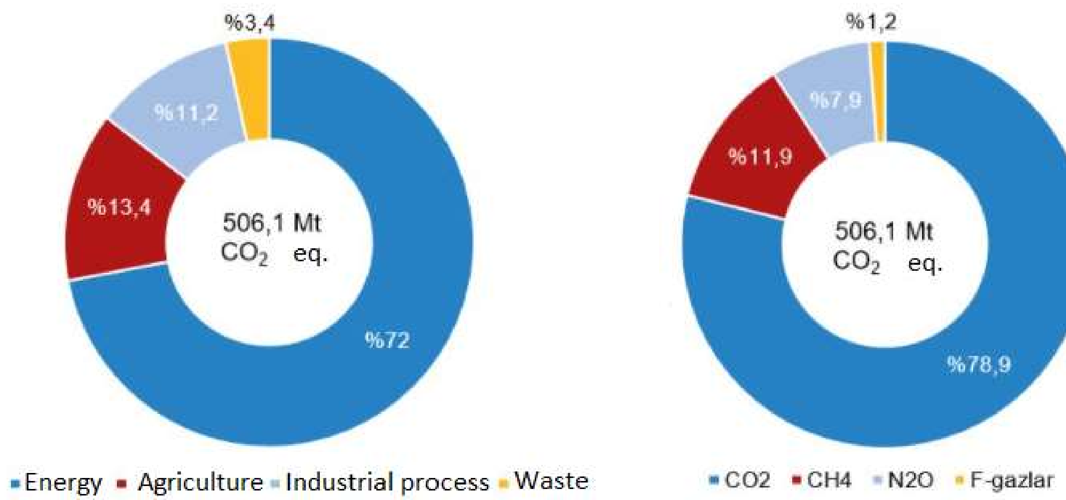


Figure 1.1. Türkiye's total and per capita greenhouse gas emissions in 1990 and 2019 (TURKSTAT, 2022)



(a) (b)

Figure 1.2. (a) Greenhouse gas emissions rates by sector and (b) Greenhouse gas emission rates by gases in Türkiye, 2019 (TURKSTAT, 2022)

The Middle East is an important region because of its geopolitical location, as it contains 70% of the world's oil and natural gas resources. Türkiye is one of the countries that is taken into account when developing energy policies because it is such an essential element of global energy geopolitics. The energy sector's ability to provide sufficient and high-quality energy is seen as a critical determinant in developing countries' economic success. The continuous increase in energy needs, especially in terms of countries, the investment plans made in the energy sector of the countries, and the efforts for the development of energy trade are increasing day by day. For this reason, countries focus on projects to increase their installed power and production values in order to meet their energy demand at a sufficient level.

The sun is the most important energy source for universe. Our country is more advantageous than many other countries in geographical position. The sun emits to the world around 170 million megawatts of energy per second. This amount of energy is more than the amount of energy used by our country. Solar energy is a viable alternative to fossil fuels because it is a clean source of energy. The use of photovoltaic electric energy is constantly increasing in our country as well as worldwide due to reasons such as lack of fuel problems, ease of operation, no mechanical wear, modularity, a clean energy source, etc. Türkiye 36°- 42° south latitude and 26° - 45° is located between the eastern Mediterranean. Accordingly, Southeastern

Anatolia Region of Türkiye is the most sunlight region. After that, the Mediterranean Region is the region with the most sun exposure after Southeastern Anatolia. Eastern Anatolia, Central Anatolia, and the Aegean regions are the next regions. The Black Sea region is Türkiye's least sunny areas. Southeastern Anatolia and the Mediterranean Coast beaches are places that can produce maximum sunlight and solar energy in Türkiye.

Renewable energy sources include wind energy; it is preferred because it is a clean, reliable and low operating cost type of energy. Also; wind turbines are preferred because of their low cost compared to other renewable energies. Although it requires high cost in the initial investment phase, the operating costs of wind turbines are quite low. Wind energy in Türkiye is generally not to meet the energy needs of homes and is used for local needs. In other words, the energy needs required for agri-culture in certain regions are met through wind turbines. Water pumps, grain grinding systems are generally powered by energy from wind turbines. On the other hand, electrical energy required for operations such as logistics and storage is provided through wind power plants. Wind power plants are generally used in lighting systems in residential and residential areas. Wind energy is a viable alternative for disabling fossil fuels, meaning it does not release carbon dioxide or other pollutants. During the three or six months of operation, the wind turbine resets all the oscillation caused by its establishment and in the remaining 20 years it works almost without causing carbon emissions.

Türkiye's energy need is constantly increasing, except for exceptional periods when events that negatively affect demand in economic and social terms. For this reason, studies are carried out to increase production in terms of installed power values in parallel with the increase in energy demand in our country. Türkiye's total installed power, which is around 27.3 gigawatts (GW) as of 2000, reached 99.1 GW at the end of October 2021. With the increase in incentives given from renewable energy sources implemented until 2021, the tendency to increase in the installed power of power plants that generate electricity from Türkiye's total and domestic sources continues. At the end of the 10-month period of 2021, Türkiye's total installed power has reached the level of 99.1 GW. The net installed power increase, which is around 3,160 megawatts in the first 10 months, consisted of power plants generating electricity from renewable resources. The increase in installed power of 485 MW is provided by hydroelectric power plants, 1,420 MW of the total increase came from wind power plants and 991 MW of it came from solar power plants. In this period, the net total installed power of the power plants producing electricity from natural gas and multi-fuel fuels is reduced with the regulations. As the new net installed power increase values between 2016 and 2020 are examined, it is understood that an annual average installed power of 4.549 MW has been activated. The highest net installed power increase among these years is realized in 2017 with 6,702 MW. 49.7% of the power plants commissioned in 2017 consists of wind power plants (WPPs) and solar power plants (SPPs).

In 2018, 62.7% of the total installed power increase of 3,350 MW came from WPPs and SPPs. While the total installed power increased by 4,624 MW in 2020, 1,913 MW WPP and SPP were commissioned in the same period. Renewable energy sources have grown their share of total installed power since 2005. By the end of 2020, the share of power plants producing electricity from renewable energy sources have climbed from 33% in 2005 to 51.7 percent. At the end of 2020, 62.5% of the renewable generation facilities, which reached a total of 49.6 GW, consisted of HEPPs, 17.8% of RES and 13.4% of SPPs. In the eleventh development plan, it is predicted that the total installed power of Türkiye will reach 109.5 GW by 2023.

As shown in Table 1, the EU will experience an increase in renewable energy deployment from around 118 million tons of oil equivalent (Mtoe) now to around 1,000 Mtoe by 2050, representing an increase of more than 88 percent of current renewable energy deployment in just 40 years. On the other hand, all EU member nations are required to raise the percentage of renewable energy sources in their power generation and to establish goals for how many renewable energy sources they wish to develop in relation to total consumption in the coming years (Kaygusuz and Toklu, 2016).

Table 1.1. Contribution renewable energy consumption in the EU region (Mtoe) (Kaygusuz and Toklu, 2016)

Renewable Energy Sources (RES)	**2014**	**2020**	**2030**
Wind	16.2	72	133.5
Hydropower	50.4	34.2	38.5
Photovoltaic	1.2	48	116
Bioenergy	130.6	226	359.1
Geothermal	3.2	35.5	188
Solar Thermal	2.1	70	122
Concentrated solar power (CSP)	0.6	12.1	33.1
Ocean	0.1	1.5	14
TOTAL RES (Mtoe)	**204.4**	**499.3**	**1,0004.2**

The short wavelength solar rays coming to the earth and the atmosphere and the long wavelength ground radiation reflected back must be in balance. Some gases, known as natural greenhouse gases contained in the structure of the atmosphere, are permeable to radiation, but the permeability properties of some of them are much less than for long wavelengths. For this reason, the earth is heating up more with the presence of greenhouse gases. The greenhouse effect is the reflection of short wavelength rays from the sun to the earth by greenhouse gases in the atmosphere, which reach the earth from the atmosphere. Greenhouse gases can be released both naturally and human-induced. Due to the greenhouse effect keeping the temperature of the gases, the average temperatures in the world have increased day by day. Carbon dioxide (CO^2), methane (CH^4), nitrogen oxide (N^2O), and ozone (O^3) are the known natural greenhouse gases in the Earth's atmosphere. Hydrofluorocarbons (HFC's), sulfur hexafluoride (SF6) and perfluorocarbons (PFC's) are examples of greenhouse gases released by human action. Carbon dioxide (CO^2), which has a very important role in the warming of the earth, is a colorless and odorless gas, and it is permeable to the sun's rays and cools the sun's rays when they hit the earth and reflect. Methane gas (CH^4), which has increased so much in the last few centuries, is a natural gas component that has as much greenhouse gas emission effect as carbon dioxide. Greenhouse gases such as CO^2, CH^4 and N^2O are naturally present in the atmosphere. However, human activities affect the concentrations of these gases in the atmosphere. In the last decade, the greenhouse effect has shown a rapid growth.

The use of non-renewable energy sources such as coal, oil and natural gas, which are used as fuel in areas such as shelter, transportation and electricity generation, cause human-induced increases in natural greenhouse gas emissions. Various chemicals and gases used in agricultural activities and in this sector also cause an increase in the greenhouse effect. As the CH^4 and N^2O from agriculture are offset by the CO^2 trapped by forests, they slightly reduce their overall emissions. Industrial greenhouse gases emerge during the production or consumption of products originating from new industries that have started to be developed since the 20th century. In addition, as a result of the destruction of forests and the construction of buildings such as residences, workplaces and factories instead of these areas, the gases that cause the greenhouse gases to be taken out of the atmosphere disappear, and an increase in temperature and naturally an increase in greenhouse gases occur. As a result of this situation, which is called global warming, it has also caused an increase in the earth's surface temperature. Industrial and technological developments have caused an increase in greenhouse gases, as well as population growth has increased greenhouse gas emissions rapidly. Global warming has become a major threat to all living things on Earth, as it causes climate change and therefore natural disasters. Considering such reasons, many countries have made national and global regulations to reduce and control greenhouse gases. In line with these regulations, countries are working on estimation modeling as a forward-looking guide to reducing greenhouse gas emissions.

In line with these regulations, countries are working on estimation modeling as a forward-looking guide to reducing greenhouse gas emissions. Today, 55% of the world's population lives in cities. It is predicted that this rate will increase rapidly as the years progress. Although cities provide economic benefits to people, they cause an increase in energy consumption and greenhouse gas emissions. However, smart cities are expected to come to the fore in the near future within the framework of the climate struggle. "Smart cities", defined by the European Commission as making traditional networks and services more efficient for residents and businesses with digital solutions, offer solutions in areas such as transport, water supply, and heating. Basically, smart city solutions based on 5G and IoT technologies primarily require a strong telecommunication infrastructure. The integration of renewable energy systems and various artificial intelligence technologies into smart cities shows that smart cities can play a role in the fight against climate change.

About half of Türkiye's greenhouse gas emissions are caused by the use of coal, cars and the construction sector. Almost one-fifth of these greenhouse gases, which are released by the presence of forests in Türkiye, are removed by trees. Türkiye accounts for approximately 1% of the total CO^2 emissions in the world and ranks 16th among the countries emitting CO^2 [42] (Union of Concerned Scientists). Like all countries in the world, Türkiye is also seriously affected by the negative consequences of climate change. In order to minimize the negative effects of climate change, Türkiye has signed global agreements on reducing greenhouse gas emissions.

Türkiye has joined the Paris Agreement, which aims to keep the average global temperature rise below 2 °C, enters into force from 2020 and offers a roadmap in the fight against climate change. On April 22, 2016, Türkiye signed the Paris Agreement with officials from 175 countries at a signing ceremony in New York. The Paris Agreement is also accepted by Russia during the United Nations (UN) summit. As a fast-developing economy, Türkiye has planned to reduce its estimated greenhouse gas emissions by more than 21% by 2030. Considering the long-term targets of the Paris Convention, which is realized with the participation of many countries, it is aimed to keep the temperature rises below 2°C until 2100 compared to the time before industrialization. In addition to reducing fossil fuel consumption, energy consumption and similar consumptions, this target will be achieved by using renewable energy sources and increasing green areas.

In parallel with the global outlook, the largest share in Türkiye's emissions belongs to energy-related emissions (72%). Emissions originating from energy are followed by agricultural activities with 13.4%, industrial processes and product use with 11.2%, and iste sector with 3.4%. In 2004, Türkiye ratified the United Nations Framework Convention on Climate Change and then, Türkiye ratified to the Kyoto Protocol in 2009 (Mills and House, 2014). The Protocol separates countries into different categories based on historical responsibility for emissions and financial responsibility for providing financial support to developing countries. Legislative infrastructure and policies for the energy sector, which are supportive of climate policies in Türkiye, have also started to be developed since the early 2000s, and within this framework, the Law on the Use of Renewable Energy Resources for the Purpose of Electricity Generation is adopted in 2005 for the support and development of renewable energy resources. As in certain geographies of the world, the share of renewable energy sources in both installed power and production has started to increase since 2009 in parallel with climate change policies in Türkiye. Similarly, the Energy Efficiency Law is enacted in 2007 for the effective implementation of energy efficiency. In 2012, the National Energy Efficiency Strategy Document is issued. In 2014, the Regulation on the Monitoring of Greenhouse Gas Emissions, which is a direct regulation on climate change, and communiqués regarding the Regulation are published in the following years. At the beginning of 2018, the National Energy Efficiency Action Plan is announced.

About half of Türkiye's greenhouse gas emissions are caused by the use of coal, cars and the construction sector. Almost one-fifth of these greenhouse gases, which are released by the presence of forests in Türkiye, are removed by trees. Türkiye accounts for approximately 1% of the total CO^2 emissions in the world and ranks 16th among the countries emitting CO^2. Like all countries in the world, Türkiye is also seriously affected by the negative consequences of climate change. In order to minimize the negative effects of climate change, Türkiye has signed global agreements on reducing greenhouse gas emissions.

2. LITERATURE REVIEW

Many countries in the world are developing forecasting models for the future in order to control the CO^2 emission and to minimize the emission, and continue to study towards this goal. Today, it is emerging that the estimation of carbon emissions with high accuracy is a very important issue for societies from an economic and environmental point of view. There are many factors that affect the performance of models developed to predict carbon emissions. In order to minimize the negative effects of lag factors on forecast accuracy, a new carbon emission forecasting model with the time-lag driving term and the linear correction term is developed. The effective performance of the developed forecasting model is presented by comparing it with the existing gray forecasting model, machine learning model and statistical forecasting approach in detail (Ye et al., 2022).

The negative effects of global warming and climate change threaten human health and the importance of the studies carried out to control carbon dioxide emissions has increased. The realization of the carbon dioxide emission forecasting guides the plans developed on this issue. A hybrid algorithm consisting of lion swarm optimizer and the genetic algorithm has been developed by using carbon dioxide emission data of countries with different development levels between 1965 and 2017. The performance of eight different algorithms is compared to reveal the performance superiority of the proposed hybrid algorithm. Thanks to the hybrid model, it is observed that the mean absolute percentage error decreased by 0.726% - 1.788% (Qiao et al., 2020).

When the sectors that cause total greenhouse gas emissions in the world are examined, the sectors with the highest rate are electricity and heat production sectors. Studies to reduce greenhouse gas emissions should focus on these sectors. Greenhouse gas emission estimation based on artificial neural network, support vector machine and deep learning algorithms from machine learning algorithms has been realized by using data from the electricity generation sector in Türkiye between 1990-2018. Five statistical performance evaluation criteria are calculated to compare the effectiveness of the prediction models used in the GHG emission estimation results. According to the estimation results obtained, it has been observed that the R2 value for carbon dioxide emissions varies between 0.861 and 0.998. In addition, it is seen that the forecasting results based on the algorithms used are less than 10% in terms of rRMSE values (Bakay and Ağbulut, 2021).

A hybrid Fast Learning Network-Chicken Swarm Optimization (CSO-FLN) model is proposed to predict the carbon emissions of Guangdong province between 2020-2060. The superiority of the developed carbon emission estimation model over other estimation models has been examined in terms of three different error indicators such as MAE, MAPE and RMSE. According to the results, it is concluded that the maximum carbon emission rate for Guangdong province will be reached approximately in 2030 or before. In addition, plans and recommendations for reducing carbon emissions are given in detail (Ren and Long, 2021).

European Union countries carry out many research and development studies to reduce greenhouse gas emissions. Greenhouse gas emissions forecasting based on artificial neural networks has been performed by using the current economic and industrial indicators as input data for 28 European countries between 2004 and 2010. Multiple statistical performance indicators are calculated to present the performances of the proposed artificial neural network models. It has been observed that the greenhouse gas estimation model for 2011 has a MAPE value of 3.60% (Antanasijević et al., 2014).

Improvement of models based on artificial neural networks has been carried out to predict greenhouse gas emissions in Serbia. In addition, in the emission forecasting for European countries, the R^2 value is found to be 0.9125. According to the estimation results obtained, it has been revealed that the artificial neural network can be applied to predict greenhouse gas emissions in order to realize the sustainable development of Serbia (Guo et al., 2021).

Economic activities and the use of energy resources directly affect greenhouse gas emissions and global warming. The data that will help countries to control their greenhouse gas emissions are the data to be obtained from the national level carbon dioxide emission forecasting studies. Forecasting of carbon dioxide emissions based on artificial neural networks has been performed for the countries of Kuwait, Iran, Saudi Arabia, United Arab Emirates and Qatar. In order to examine the performance efficiency of the developed model, the average absolute relative error and the R-squared values are calculated and 2.3% and 0.9998 values are found, respectively (Javed and Cudjoe, 2022).

The countries of China and India contribute significantly to greenhouse gas emissions in the Transportation, Construction, and Manufacturing sectors. Greenhouse gas emissions from these sectors until 2028 are estimated using the improved discrete gray forecast model. The estimation results obtained show that the proposed model predicts greenhouse gas emissions with an accuracy of over 95%. In addition, various suggestions have been made to reduce greenhouse emissions (Ahmadi et al., 2019).

In order to observe the carbon dioxide emissions of Russia, Brazil, South, Africa, China and India, a new forecasting model has been created by improving the existing gray model. According to the estimation results carried out using the carbon emission data of these countries, it has been revealed that the emissions in Brazil and Russia have a decreasing trend. Moreover, it has been observed that the carbon dioxide emissions of other countries have an increasing trend (Radojević et al., 2013).

Different data and methods have been used to estimate China's carbon emissions, but the contradiction between the results could not be resolved. In this study, the association rule algorithm has been used to make an accurate prediction and to correctly identify the factors affecting the formation of carbon emissions. It has been confirmed that urbanization, energy consumption, economic growth, foreign direct investment and industrial structure are the main factors causing the increase in carbon emissions. A gray model optimized by the firefly algorithm (FA-GM) model has been developed to realize the carbon emission estimation. This model is the best performing model among 8 different models and its MAPE value is 0.0499 (Ma et al., 2020).

A new information priority generalized accumulative gray model (NIPGAGM (1,1, k)) has been proposed to estimate greenhouse gas emissions from eight member states of the Shanghai Cooperation Organization. GHG emissions are estimated and compared using five gray models (NHGM, GM, NIPGM, DGM, and NIPGAGM) and linear regression models. NIPGAGM, which has the best forecast performance, has forecast MAPES for each country (Russia Kazakhstan, Kyrgyzstan, Tajikistan, Uzbekistan, Pakistan, India) are 0.42%, 0.11%, 6.36%, 0.06%, 0.48%, 0.52% and 0.84%, respectively (Li et al., 2022).

Optimized metabolic gray model OMGM (1,1) and optimized nonlinear metabolic gray model (ONMGM (1,1)) are applied for the GHG prediction from Türkiye's energy sector. ONMGM (1,1) gave better results than other models and its MAPE value is 5.19%. The annual growth in the energy sector is estimated at 0.49% and the carbon emission in the energy sector in 2025 is estimated as 585.2Mt of CO^2 equivalent (Şahin, 2019).

In order to contribute to the accurate planning of carbon emission reduction studies, a new forecasting method called carbon emissions ensemble forecasting system (CEEFs) has been developed. This method made a prediction consisting of 2 data sets using daily data from America and China, and used 659 data in each data set. The average MAPE value is 1.1102% in the first data set and 1.1382% in the second data set. According to other models, the authors achieved a successful result (Liu et al., 2022).

In another study, an artificial neural network consisting of 11-3-2 structure has been used to calculate energy-based production efficiency and gas emission in wheat production in Iran. In wheat production, energy input is calculated as 80.1 and output as 38 GJ ha-1. Electricity, chemical fertilizer and irrigation water are the most effective factors in energy consumption with 39.5, 23.3 and 6.17 GJ ha-1, respectively. The coefficients of determination (R^2) of the best topology in this study are 0.99 for wheat yield and 0.998 for GHG emissions (Khoshnevisan et al., 2013).

A GRNN (general regression neural network) applied to predict energy consumption and greenhouse gas intensity of energy consumption using historical data for the period 2004-2012 for 26 European countries has been developed. This developed model is first used in the estimation of energy consumption, and it is also used in the estimation of greenhouse gas emissions due to the high accuracy of the results. Energy consumption and greenhouse gas emission MAPE values are 3.6% and 6.4%, respectively (Antanasijević et al., 2015).

By using nine parameter input variables, an artificial neural network model is used for the 5 countries with the highest greenhouse gas emissions, using 3-month data between 1980 and 2015. The estimated and actual values are nearly identical, with higher coefficients of determination (R^2) of 0.80 for Australia, 0.91 for Brazil, 0.95 for China, 0.99 for India, and 0.87 for the United States of America (Acheampong and Boateng, 2019).

China’s transportation sector is responsible for a major portion of the country’s carbon dioxide emissions. Forecasting the carbon emissions caused by the aviation industry is very important for societies to plan a safe future. The forecasting of carbon emissions and fuel consumption caused by passenger flights in Shanghai has been performed using the Autoregressive Integrated Moving Average model. It is concluded that the carbon dioxide emissions for the Shanghai region will increase by 6.41% compared to the same period of the previous year (Yang and O’Connell, 2020).

An effective carbon emission forecasting method is a guide to lead way countries’ attempts to reach low carbon emission levels. In addition, the integration of more renewable energy sources into microgrids also plays an important role in minimizing carbon dioxide emissions. Long-term carbon dioxide emission forecasting in Iran, Canada and Italy is carried out by using the Generalized Regression Neural Network and Gray Wolf Optimization. When the results obtained are compared in detail, the superiority of the proposed method is presented (Heydari et al., 2019).

Forecasting models based on Artificial Neural Network, Holt-Winters Exponential Smoothing (H-W) and Autoregressive Integrated Moving Average models have been proposed to create a low carbon emission economic development plan in the country of Saudi Arabia. Various error criteria have been calculated to demonstrate the accuracy of the prediction models. According to the results obtained, it has been revealed that the ARIMA (2,1,2) model has high accuracy for the forecasting of the carbon emission of the Saudi Arabian country (Alam and AlArjani, 2021).

Gray forecasting models are widely used in studies forecasting carbon dioxide emissions of developing and developed countries. Gray forecast models have been used to estimate the carbon dioxide emissions in the years 2010-2019 for the country of Vietnam. According to the results obtained, it has been observed that the carbon dioxide emissions of Vietnam will increase by 3%. In addition, solutions that try to improve energy efficiency and provide economic growth have been proposed (Ho, 2018).

The development of a high-accuracy short-term carbon dioxide emission prediction model is a very important study to contain climate change. Carbon dioxide emission forecasting based on a multi-layer artificial neural network model has been performed for the 17 countries emitting the most carbon dioxide emissions. It is presented that the proposed forecasting model is superior to the current forecasting results. According to the carbon dioxide emission estimation results based on a multi-layered artificial neural network model, it is seen that some countries will increase their carbon emissions over the years and some countries will reduce their carbon emissions (Jena et al., 2021).

Carbon dioxide emission has become an important environmental problem for the countries of China and Russia. Carbon dioxide emission estimation data have been a guide in determining environmentally friendly strategies and policies. The Gray Verhulst model has been improved to predict with high accuracy the carbon dioxide emissions in China and Russia. It is presented with the estimation results that the improved Gray Verhulst model is more efficient than the existing Gray Verhulst model and the ARIMA model (Tong et al., 2021).

In order to effectively control carbon dioxide emissions in the world, it is necessary to estimate carbon dioxide emissions from the building sector. A forecasting model based on metaheuristic algorithms has been developed to predict carbon dioxide emissions from the building industry in South Korea. Various error criteria are calculated to evaluate the performance of the proposed forecasting model. The estimation results obtained contain data that will be of great help in estimating national CO^2 emissions in 2030 (Hong et al., 2018).

The EU-28, which reduced CO^2 emissions by 25.7 percent in 2014 compared to 1990, is aiming for a 40 percent reduction in CO^2 emissions by 2030, and in this study, it aims to investigate the optimum model for estimating the CO^2 emissions in the EU-28. An autoregressive integrated moving average (ARIMA) (1,1,1)- autoregressive conditional heteroscedasticity (ARCH) (1) model is proposed to achieve this aim. It is estimated by applying both a dynamic and a static process using the carbon emission data of the European Union countries between 1960 and 2015. The RMSE and MAE values of the dynamic model are 0.243838 and 0.194225, respectively. The RMSE and MAE values of the static model are 0.236567 and 0.185870, respectively (Dritsaki and Dritsaki, 2020).

In another study, Iran's CO^2 emissions in 2030 using Multiple linear regression (MLR) and multiple polynomial regression (MPR) analysis under the assumptions of two scenarios, business-as-usual (BAU) and the Sixth development plan (SDP), which includes targets from the Paris Agreement, forecast its emissions. For BAU scenario, it offers 730 594 kt in MLR model and 876 263 kt in MPR model, while in SDP scenario it is 211 096 kt and 315 738 kt for MPR and MLR models, respectively. According to the results, it is seen that Iran is far from its carbon emission target in 2030 (Hosseini et al., 2019).

Using the data obtained from Hebei Province of China for the period 1995-2015, the new model is used to estimate carbon dioxide emissions, and the proposed model performed well compared to other compared models, showing the accuracy of the CO^2 emission estimation and potential for improvement. For carbon dioxide emission forecasting, a hybrid model integrating the random forest and extreme learning machine is developed. MAPE and RMSE values are found to be 0.155328 and 0.001949, respectively (Wei et al., 2018).

The authors performed a two-stage forecasting method for Taiwan using a multivariable gray forecasting model and genetic programming. The multivariable gray forecasting model makes predictions by adding four factors and more, increasing its accuracy compared to other models. By using genetic programming, it reduces the prediction error. The suggested forecast approach offers more accuracy than earlier approaches, according to experimental comparisons on various combinations of multiple factors. The MAPE value between 2000 and 2013 is calculated as 2.14% and between 2014 and 2015 as 0.91% (Lin et al., 2018).

The new model is used to estimate carbon dioxide emissions in China, and the proposed model performed well compared to other compared models, showing the accuracy of the CO^2 emission estimation and potential for improvement. For carbon dioxide emission forecasting, a weighted Adaboost-ENN Model is developed. $R^{2,}$ MAPE and RMSE values are found to be 0.9977, 0.37 and 0.0059, respectively (Zhou et al., 2019).

Artificial neural network (ANN), fuzzy linear regression (FLR), and conventional regression (CR) applied to predict energy consumption and greenhouse gas emission of energy consumption for Brazil, Canada, France, Japan, India, UK and US countries has been developed. This developed model is first used in the estimation of energy consumption, and it is also used in the estimation of greenhouse gas emissions due to the high accuracy of the results. greenhouse gas emission MAPE values is less than 0.89 (Azadeh et al., 2015).

In this paper, a novel quantum harmony search (QHS) algorithm-based discounted mean square forecast error (DMSFE) combination model aims to find the best mean absolute percent error (MAPE) by providing a way to search for optimal β values. For the countries China, United States, India, Russia and Japan, this estimate is made and the best MAPE value is found to be 2.9949 (Chang et al., 2013).

Particle Swarm Optimization (PSO) and Multiple Linear Regression (MLR) models are used to estimate India's carbon emissions. The RMSE value is 11.73 and 4.32 for PSO and MLR values, respectively. The MAPE value is calculated as 23.76 and 4.07 for the PSO and MLR values, respectively (Sangeetha and Amudha, 2018). A comprehensive review of previous research on the different forecasting model for greenhouse gases is given Table 2. A comprehensive review of previous research on the different forecasting model for greenhouse gases is given Table 2.1.

Table 2.1. A comprehensive review of previous research on the different forecasting model for greenhouse gases

Ref.s	Proposed Model	Location	Error Criteria
(Ye et al., 2022)	Dynamic time-delay discrete grey forecasting model	China	MAPE <1.5% RMSE 0.011
(Qiao et al., 2020)	Improved lion swarm optimizer	China, India, Canada, Brazil, Iran, United States, South Africa, Japan, German, Türkiye and Saudi Arabia	MAPE decreased by 0.726%-1.878%
(Bakay and Ağbulut, 2021)	Deep learning, support vector machine and artificial neural network	Türkiye	R^2 varies from 0.861 to 0.998 rRMSE values < 10%.
(Ren and Long, 2021)	Optimized Fast Learning Network, Chicken Swarm Optimization	Guangdong in China	RMSE 27.755 MAPE 1.656% MAE 11.609
(Antanasijević et al., 2014)	Optimized artificial neural network	28 European countries	MAPE 3.6%
(Guo et al., 2021)	Artificial Neural Networks	Serbia	R^2 0.9125
(Javed and Cudjoe, 2022)	Artificial neural networks	Kuwait, Iran, Saudi Arabia, United Arab Emirates and Qatar	Average absolute relative error 2.3% R^2 0.9998

Table 2.1

(Ahmadi et al., 2019)	A novel grey forecasting model	China and India	MAPE <11.24% RMSE <299.9
(Radojević et al., 2013)	Exponential cumulative grey model	Brazil, Russia, India, China and South Africa	MAPE <18.44% MAE <0.384 RMSE <0.479
(Ma et al., 2020)	Gray model optimized by the firefly algorithm	China	MAPE 0.0499
(Li et al., 2022)	New information priority generalized accumulative gray model	Russia Kazakhstan, Kyrgyzstan, Tajikistan, Uzbekistan, Pakistan, India	MAPE <6.36
(Şahin, 2019)	Optimized metabolic gray model and optimized nonlinear metabolic gray model	Türkiye	MAPE 5.19
(Liu et al., 2022)	Multi-objective tangent search algorithm	America and China	MAPE (Dataset1) 1.1102% MAPE (Dataset2) 1.1382%
(Khoshnevisan et al., 2013)	Artificial Neural Networks	Iran	R^2 0.99
(Antanasijević et al., 2015)	General regression neural network	26 European countries	MAPE 6.4%
(Acheampong and Boateng, 2019)	Artificial neural network	Australia, Brazil, China, India, and USA	R^2 <0.99
(Yang and O'Connell, 2020)	Autoregressive Integrated Moving Average linear model	Shanghai in China	RMSE 63.612 MAE 18.610 MAPE 2.634
(Heydari et al., 2019)	Generalized Regression Neural Network and Grey Wolf Optimization	Iran, Canada and Italy	RMSE 0.0111 MAE 0.0094 MAPE 2.2194 (The best values)
(Alam and AlArjani, 2021)	Autoregressive Integrated Moving Average, Holt-Winters Exponential Smoothing and Artificial Neural Network	Saudi Arabia's	RMSE 1.44 MAE 1.12 MAPE 0.08 (The best values)
(Ho, 2018)	Grey Models	Vietnam	Mean relative error 1.21% (Model 1) Mean relative error 4.83% (Model 2)

Table 2.1

(ena et al., 2021)	Multilayer Artificial Neural Network	17 key Carbon dioxide emitting countries	MAPE < 9.6898% MAE < 0.0653 RMSE < 0.1077
(Tong et al., 2021)	Action Verhulst Grey Model	China and Russia	MAPE < 11.3849%
(Hong et al., 2018)	Optimized Gene Expression Programming Model (Using the Metaheuristic Algorithms)	South Korean	MAE < 0.0188 RMSE < 0.0229 MAPE < 8.0%
(Dritsaki and Dritsaki, 2020)	An autoregressive integrated moving average (ARIMA) (1,1,1)-autoregressive conditional heteroscedasticity (ARCH) (1) model	28 European countries	RMSE 0.243838 and MAE 0.194225 for the dynamic model. RMSE 0.236567 and MAE 0.185870 for the static model.
(Hosseini et al., 2019)	Multiple linear regression (MLR) and multiple polynomial regression (MPR)	Iran	MAPE 0.05
(Wei et al., 2018)	Random forest and Extreme learning machine	Hebei in China	MAPE 0.155328 RMSE 0.001949
(Lin et al., 2018)	Multivariable gray forecasting and Genetic programming	Taiwan	MAPE < 2.14%
(Zhou et al., 2019)	Weighted Adaboost-ENN Model	China	R^2 0.9977 MAPE (%) 0.37 RMSE 0.0059
(Azadeh et al., 2015)	Artificial neural network (ANN), fuzzy linear regression (FLR), and conventional regression (CR)	Brazil, Canada, France, Japan, India, UK and US	MAPE (%) < 0.89
(Chang et al., 2013)	Quantum harmony search (QHS) algorithm-based discounted mean square forecast error (DMSFE)	China, United States, India, Russia and Japan	MAPE (%) < 2.9949
(eetha and Amudha, 2018)	Particle Swarm Optimization (PSO) and Multiple Linear Regression (MLR).	India	RMSE 11.73 and 4.32 MAPE 23.76 and 4.07

3. MATERIALS AND METHODS

Optimization is a term that can be thought of as word-for-best. As we think of it mathematically, optimization is the determination of the set of variables that will minimize/maximize an objective function in accordance with the structure of the problem. Until today, many different optimization methods have been proposed for the solution of problems of different nature.

Since the most real-time problems have complex structures, the mathematical methods have become insufficient at solution such complicated problems over time. Therefore, new methods have been proposed for the solution of optimization problems. One of the important topics of these new solution proposals is a metaheuristic method approach (Koopialipoor and Noorbakhsh, 2000).

The metaheuristic approach includes methods that search for a global optimum solution in the search space. These methods, which use various operators to enable them to search different regions of the search space during these searches, produce statistically better solutions (Yetkin, 2016). The basis of these methods is mostly modeling of the behaviors of living creatures in nature to reach food and protect themselves from dangers. Genetic Algorithms, Particle Swarm Optimization, Cuckoo Search, Ant Colony Optimization, Shuffled Frog Leaping and Firefly Algorithms are examples of this approach.

3.1. Shuffled Frog Leaping Algorithm

In optimization studies, the first step is to determine the objective function to fit the nature of the problem. Optimization methods can be grouped mainly in terms of the number of objective functions. If the purpose determined according to the structure of the problem examined is one, the problem is single-purpose. Otherwise, the problem is multi-purpose. In single-objective problems, the first step is adjusting the parameters of the problem to give the optimum value of the objective function. The Shuffled Frog Leaping Algorithm is also one of the recommended methods for solving single-purpose problems.

The Shuffled Frog Leaping Algorithm (SFLA) is first used in 2003 by Eusuff and Lansey to optimize the water distribution network problem (Eusuff and Lansey, 2003). This is a memetic algorithm that seeks to mathematically model the behavior of swarming frogs, such as reaching for food and escaping danger. The main purpose of the SFLA algorithm is to search for food in groups with minimal movement and share their food-related information with each other.

The basis of this algorithm is the meme known as the cultural transmission unit. This structure is similar to the gene in genetics. Traits belonging to an individual are passed on to the next generation through genes. By using same way of thinking, the skills and behaviors of the individual in social life are transferred to the next generation through memes. The algorithm is logically similar to PSO (Kennedy and Eberhart, 1995) and genetic algorithm. Just as the gene that is better among individuals in the genetic algorithm is transferred to the next generation, a better behavior in the SFLA algorithm is transferred to the next generation (Eusuff et al., 2006).

Evolutionary algorithms, which are a subset of metaheuristic methods, work with population structure. The algorithm initially creates a population with randomly generated individuals, then searches for the global best solution in the search space with the methods of the algorithm structure. SFLA is also an evolutionary based algorithm. It is based on the logic of the transfer of memes between generations and the exchange of information between individuals in the same generation. In SFLA, the population is formed into subsets called memeplex. Each memeplex has a different culture amongst itself and makes a local search.

In the first step of the SFLA algorithm, the population is created with random individuals. Each frog represents a solution to the problem. The population is formed by the combination of frogs. The fitness value of each individual is expressed by the distance from the food. After that, the fitness values of these individuals are calculated and this fitness value represents the degree of proximity of each individual to the food. Then, all individuals in the population are ranked according to their fitness values and divided into memeplexes according to a certain rule. After that, each memeplex is evaluated on its own. The next step is to assess each memeplex separately. Individuals in each memeplex go through memetic evolution in order to improve the quality of their memes. The goal in memetic evolution is to bring each individual closer to food. At this stage, the individual with the best position on the basis of memeplex (local best) or the best position on the basis of population (global best) is used. After these stages are completed in all memeplexes, all individuals belonging to the population are reassembled to provide information sharing globally. According to these new results, the population is divided into new groups. Thus, the memetic knowledge of frogs is shared globally with the contribution of frogs in different locations, and it is aimed to reach the optimum solution in the fastest way possible.

The algorithm consists of five steps.

Step 1. Parameter setting and population initialization
In this step, firstly, the parameters of the algorithm are determined and then the initial values are assigned. Afterwards, the first population (P) is created by generating random individuals in line with the determined limits.

Step 2. Creating memeplexes
In this step, the individuals in the population are first sorted according to their fitness values and divided into m memeplexes according to a certain rule (n individuals in each memeplex so P=mxn). In the process of dividing the population into memeplexes, the first frog goes to the first memeplex. The second frog goes to the second memeplex and the third to the third memeplex. After, frog m goes to mth memeplex, frog (m+1) goes to first one, and so forth. The steps for the division of the population into memeplexes is illustrated in the Figure 4.1. In the next step, each memeplex is evaluated on its own.

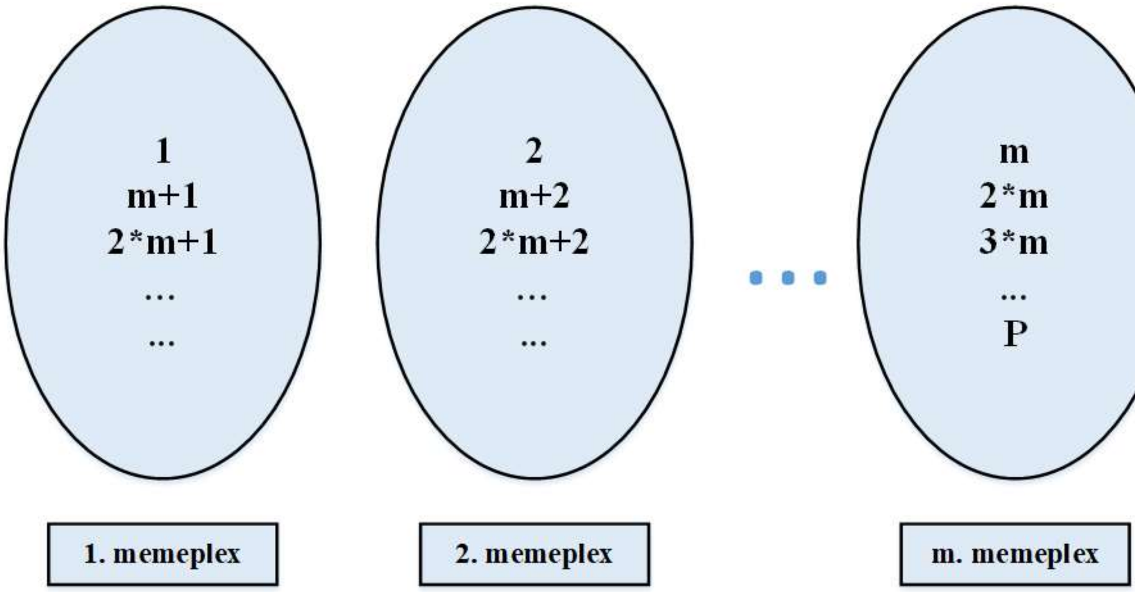


Figure 3.1. Structure of the division of the population into memeplexes

Step 3. Creating sub-memeplexes

In this step, the sub-memeplex is first created. With the triangular probability distribution, higher selection rates are given to individuals with high fitness values in the memeplex, increasing their chances of entering the sub-memeplex. In addition, frogs with low fitness values are not ignored and they are given a chance to be selected in proportion to their low fitness values.

The triangular probability distribution formula is:

$$O_i = \frac{2 \times (n + 1 - i)}{n \times (n + 1)} \tag{1}$$

where n is the number of individuals in the memeplex and O_i is the probability that the ith individual will be selected.

Step 4. Memetic evolution

After that, it is aimed that the individual with the worst fitness value in the sub-memeplex will jump to a better solution point. At this point, the individual with the best and worst positions on the basis of memeplex (local best - X_{best}, local worst - X_{worst}) and the best position on the basis of population (global best - X_{gbest}) is used. First, the local best individual will be used for the location update. If the desired conditions are not met, the location update will be tried according to the global best individual. If the result obtained is still not a suitable value, the position of this individual will be updated randomly.

As S is specified as the individual's location update amount, step size and new position is calculated according to the local best as follows:

$$S = rand \times (X_{best} - X_{worst}) \tag{2}$$

$$X'_{worst} = X_{worst} + S, \qquad S_{min} \leq S \leq S_{max} \tag{3}$$

where, S is the step size and rand is a uniformly distributed random number in (0, 1). S_{min} and S_{max} minimum and maximum values of step sizes allowed to frog, respectively. X'_{worst} is the updated position of the individual with the worst fitness value.
If the X'_{worst} position is not better than the previous position, the update is made based on the global best individual. In this case, the location update formulas are as follows:

$$S = rand \times \left(X_{gbest} - X_{worst}\right) \tag{4}$$

$$X'_{worst} = X_{worst} + S, \qquad S_{min} \leq S \leq S_{max} \tag{5}$$

Here, X_{gbest} is the location of the global best individual.

If this new location information is not better for the individual with the worst fitness value, a new individual is created randomly and replaced with the worst individual.
Interaction between subgroups continues until these processes are done in all subgroups. Then, it is proceeded to the next step.

Step 5. Global interaction
In this step, all frogs in the population are brought together. The population is again ranked according to fitness values. Then, it is divided into memeplexes and the local operations specified in the previous steps are performed. These processes continue until the termination criterion is met.

The pseudo code of the SFLA algorithm is given in Figure 3.2.

1. **Begin**
   define the fitness function f(x)
   set the initial value of the parameters
   generate the initial population P= (1, 2,..., p)
   calculate the fitness value of each frog
2. **while** *(t<maximum_generation_number)*
   Creating memeplexes
   for *(i=1:m) % for all memeplexes*
       for *(j=1: z) % defined iteration number for local search*
           Creating sub-memeplexes
           Determine the X_b, X_w *and* X_g
           Update new position with Eqs. 2-5
       end *for j*
   end *for i*
   combine all memeplexes and rank the population according to their fitness value
   end_while
3. *find the frog with the highest fitness and display it as optimal solution*

Figure 3.2. The pseudo code of the SFLA algorithm

In addition, the detailed flow diagram of the algorithm is given in Figure 3.3.

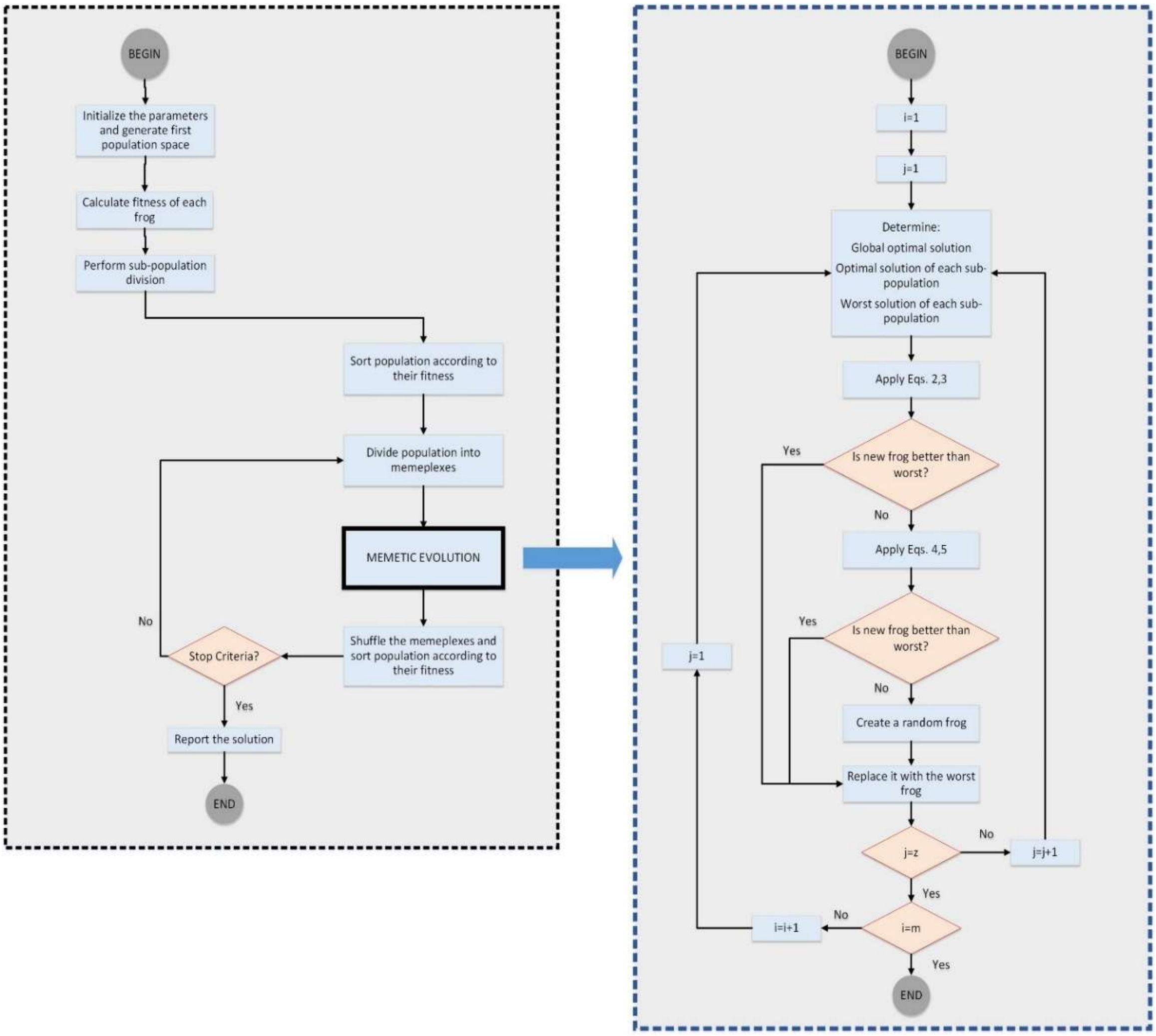


Figure 3.3. Flow diagram of the SFLA algorithm

3.2. Firefly Algorithm

The Firefly Algorithm (FA) is one of the meta-heuristic optimization methods that mimics the swarm behavior of animal/insect species. This algorithm is developed by modeling the behavior of fireflies and their communication by glowing. This is a swarm-based heuristic optimization algorithm developed by Xin-She Yang in 2008 (Yang, 2008). There are three basic rules in the firefly algorithm method:

(1) All fireflies can communicate with each other regardless of sex,

(2) The attraction of any firefly is proportional to the glow brighter. Therefore, the one that emits lighter than the two fireflies attracts more attention and tends towards it,

(3) The brightness of the firefly is determined by the nature of the objective function.

Each firefly is attracted by the one with a brighter glow from among the neighboring fireflies. As the distance between two fireflies increases, the amount of attraction decreases. If there is no brighter firefly around the firefly, this firefly will move randomly. The two most important parameters in the firefly algorithm (FA) method are the variation of light intensity and the attractiveness of the firefly (β). The attraction of the firefly is proportional to the light it emits. As the light intensity decreases, the attractiveness of the firefly decreases.

There are two important terms in this algorithm: light intensity and attractiveness.
Light intensity: In Firefly Algorithm, the light intensity of each firefly is the distance of that firefly from the food. In other words, the light intensity value of each firefly is the fitness value of that firefly and it is defined as:

$$I_i = f(x_i) \tag{6}$$

where f(x), is the fitness function of the problem. I_i, is the light intensity of the ith firefly (Yang, 2008), (Senthilnath et al., 2011).

Attractiveness: Fireflies move towards fireflies, which are among their neighbors and are brighter than themselves. If there is no individual brighter than themselves, they move randomly. The attractiveness value, which affects the movement of fireflies, is associated with neighboring fireflies. In FA, attractiveness value of each individual is proportional to the light intensity.

The distance between two fireflies (r_{ij}) at points i and j is calculated as:

$$r_{ij} = \left\| x_i - x_j \right\| \tag{7}$$

here, x_i and x_j are positions of the i^{th} and j^{th} fireflies, respectively. The attractiveness value (β) of the firefly is calculated as follows:

$$\beta(r) = \beta_0 \times e^{-\gamma \times r^2} \tag{8}$$

here, γ is the light absorption coefficient. $\beta(r)$ is the attractiveness at r and β_0 is the attractiveness at r=0.

The movement of firefly i towards the more attractive firefly j is calculated as:

$$x_i(t+1) = x_i(t) + \beta_0 \times e^{-\gamma \times r^2}\left(x_j - x_i\right) + \alpha \times (rand - 0.5) \qquad (9)$$

where α is the randomization parameter. The first part of this equation provides the effect of the current position of the individual on the next generation. The second part provides attraction between individuals, while the third part provides randomization. In general, the Firefly Algorithm can be examined in four steps.

Step 1. Initialization
Parameters and limitations of the algorithm are determined. Initial value assignments are made to the parameters. The locations of the fireflies are randomly specified at the boundaries of the search space.

Step 2. Determining light intensity
In this step, a fitness function is determined to be suitable for the problem structure. This fitness function is used to determine the light intensity of each firefly.

Step 3. Movement of fireflies in search space
Fireflies move depending on the attractiveness of their neighbors and the distance between them and their neighbors. In line with this movement, they navigate the search space. In each fitness evaluation, they move towards the more attractive individual and approach the optimum solution. The attractiveness value of individuals is calculated with the fitness function.

Step 4. Checking termination condition
In this step, the termination criterion specified at the beginning is checked. If the criterion is not met, go to Step 3. The algorithm will continue to run and search for the optimum result in

the search space until the termination criterion is met (Wang et al., 2017). The pseudo code of the algorithm is given in Figure 3.4.

```
1.  Begin
    set the initial value of the parameters
    generate the initial population of the fireflies
    define the fitness function f(x)
    define the light intensity (I_i) associated with the fitness function
2.  while (t<maximum_generation_number)
    for (i=1: n) % for all fireflies
            for (j=1: n) % for all fireflies
                    if (f(x_j)<f(x_i))
                            move firefly i towards j
                    end if
                    update the attractiveness
                    evaluate the new solutions and update the light intensity
            end for j
    end for i
    rank the population according to their fitness value and find the current optimal solution
    end_while
3.  find the firefly with the highest fitness and display it as optimal solution
```

Figure 3.4. The pseudo code of the FA algorithm

In addition, the flow chart of the algorithm is given in Figure 3.5.

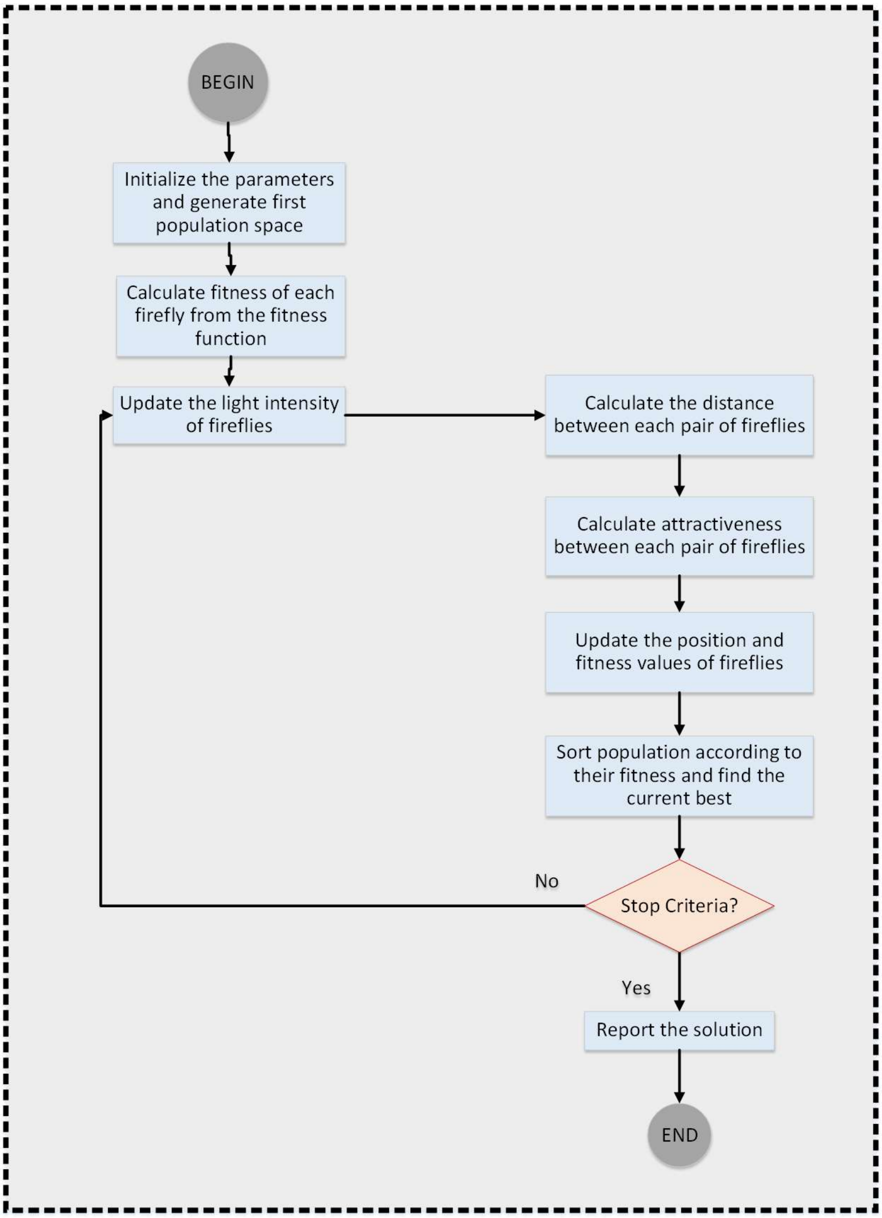


Figure 3.5. Flow diagram of the firefly algorithm

3.3. Artificial Neural Networks

Artificial Neural Networks (ANN) have emerged as a result of mathematical modeling of functions such as the ability to generate new data from the data collected by the human brain through learning, remembering and generalization. In general, computer software developed by imitating the structure of biological neural networks in the brain is called artificial neural networks. Inspired by the human brain, the learning ability is transferred to the machines by applying the learning ability to the machines. Using this method, it is possible to solve problems that are difficult to model linearly in many areas.

An artificial neuron created by software is defined by being inspired by the basic behavior of a biological neuron. Biological nerve cells process the signal they receive from other nerve cells and transmit it to the next nerve cells. This working principle is applied to the artificial nerve cells. The structure of an artificial neuron is as illustrated in the Figure 3.6.

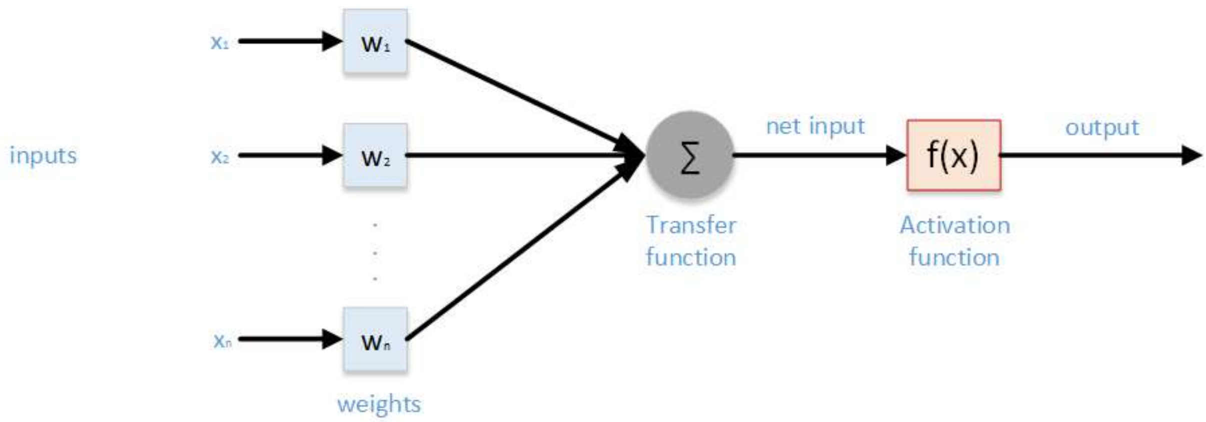


Figure 3.6. Structure of an artificial neuron

As shown in Figure 3.6, each artificial neuron consists of 5 basic components: inputs, weights, transfer function, activation function and outputs (Sharma et al., 2012). Each neuron in the artificial neural network receives n signals. This signal is multiplied by the network coefficient of each neuron and collected in the transfer function and net input is obtained. Finally, the net input signal is passed through the activation function and the output signal is obtained as a result (Lee et al., 2005).

ANNs are used in many different fields such as estimation, classification and pattern recognition in the world of science, due to their simple construction and easy application (Rostami and Jafarian, 2018). Artificial neural networks are divided into two according to the flow direction of the signals and the connection structures between the nerves: Feedback and Feed Forward Network architecture (Sharma et al., 2012). In feedback networks, at least one signal is returned to the previous neuron. This feedback can also occur between layers. In addition, it is necessary to develop an architectural structure that generally uses a delay signal in feedback networks.

Feed forward networks have one-way signal flow. The signal flow starts with the input of the signal received from the physical environment from the input layer. This information is then processed in the hidden layer(s) and the resulting signal is given to the outside world over the output layer. Multilayer neural networks are frequently used to solve nonlinear complex problems. The fact that different learning algorithms can be used easily during the training of the network is one of the reasons for the frequent use of this network structure. The most widely known of feed-forward networks are Multi-Layer Perceptron (MLP).

A multi-layer perceptron is thought to be one of the most efficient structures suggested in modeling (Rostami and Jafarian, 2018). MLPs consist of an input layer, one or more hidden layers, and an output layer. The signal received from the outside is transmitted from the input layer to the hidden layer without any changes. The signal processed in the hidden and output layer is finally given from the output layer. The number of neurons in the input and output layers is determined according to the problem structure. The number of hidden layers and the number of neurons in the hidden layers are decided by trial and error during the solution of the problem. The neurons in the layers are in connection with all the neurons in the previous layer. In this structure, the flow of information is unidirectional and forward (Marugán et al., 2018), (Marugán et al., 2017). The MLP structure is demonstrated in the Figure 3.7.

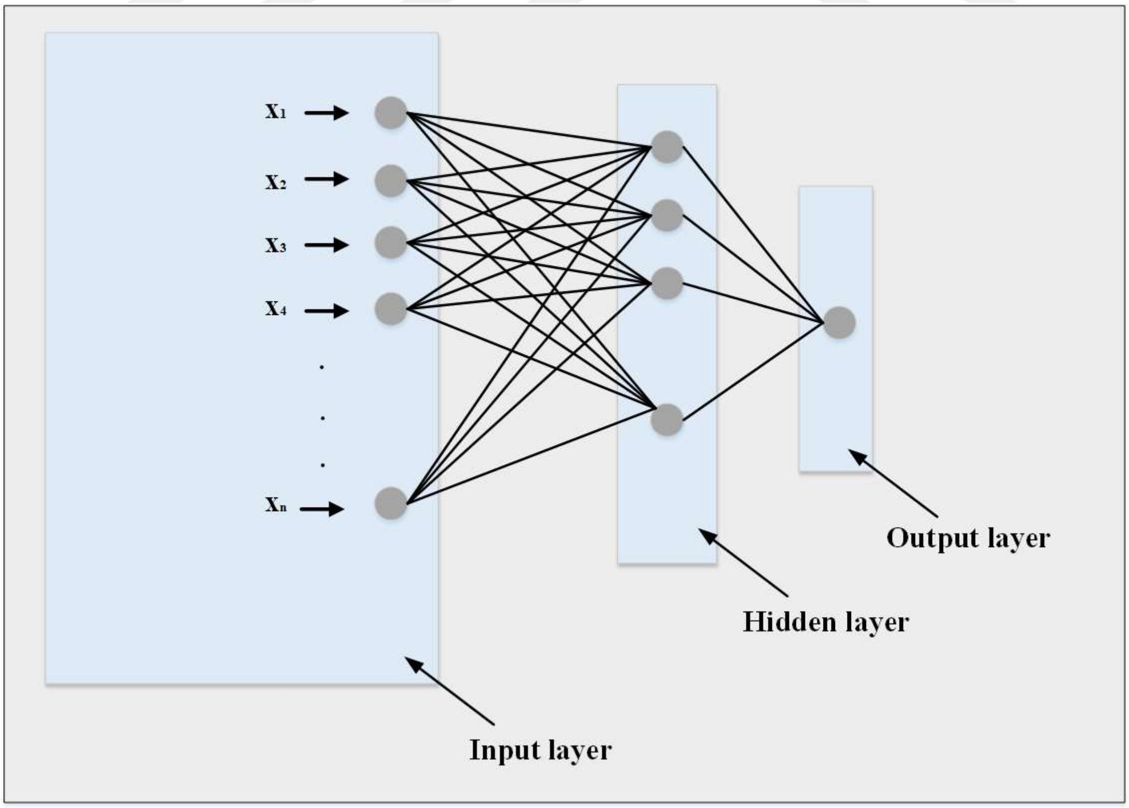


Figure 3.7. The MLP neural network architecture

3.4. Levy Flight

Levy flight is a distribution strategy which provides a non-Gaussian random walk. This distribution is expressed using the power law formula. Levy flight expands the search space, allowing the algorithm to continue searching without trap into the local minimum.

As the movements of animals in nature are examined, randomness draws attention. Therefore, in optimization methods, this randomness is of great importance. In the literature, it is shown that the use of Levy flight in optimization methods affects the results positively. It is seen that Levy flight has been added as an improvement to the nature inspired optimization methods that are frequently used today. Yang and Deb used the Cuckoo Search algorithm together with the Levy flight method. At the end of the study, it is tested with Benchmark Functions and the results are compared with the Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) (Yang and Deb, 2009). Senthilnath et al. used the Cuckoo Search with Levy Flight method for clustering (Senthilnath et al., 2013). Baskan used the Cuckoo Search algorithm together with the Levy Flight method for determining optimal link capacity expansions in road networks (Baskan, 2013). Levy flight is also used in the proposed PSO algorithm by simulating the movement of birds in nature. Levy flight is preferred in the PSO method used in the parameter adjustment step of the study for heart disease and breast cancer detection (Lan et al., 2020). Hariya et al. examined the Levy Flight on PSO in detail in their work. At the end of the study, they supported the proposed method with numerical simulations (Hariya et al., 2015). While Hassanzadeh et al. integrated the Levy Flight method into the Firefly Algorithm for maximum entropy-based image segmentation (Hassanzadeh et al., 2011), Liu and Yang used the Levy Flight with Firefly Algorithm for multilevel thresholding image segmentation (Liu and Wang, 2021).

The reason of the SFLA algorithm is inefficient during search is the absence of acceleration terms in the position update equation (Rajpoot et al., 2018). In traditional SFLA, the position equation of the X_{best} (local best), X_{gbest} (global best) and X_{worst} (local worst) frogs is updated by using the rand value, which is not sufficient to escape from the local minimum. Levy Flight function is used to enhance the search capability of the SFLA method. There are two rules used for searching in the SFLA; local search and global search (Tang et al., 2016). In both rules, Levy Flight function is used to improve search capability.

In general, Levy Flight is a random search method with a step length selected from the Levy distribution which has a long-tailed probability distribution for the step length. In this method, short-distance steps enable individuals to search in their close surroundings, while long-distance steps enable individuals to search far places by jumping in the search space. The mathematical formula for the Levy distribution is as follows (Jensi and Jiji, 2016):

$$Levy(s) \sim |s|^{-(1+\beta)}, \quad 0 < \beta \leq 2 \tag{10}$$

where s is the step size and β is an index of o stability.

Mantegna's algorithm can be used to compute the step length s for a random walk.

$$s = \frac{u}{|v|^{1/\beta}} \tag{11}$$

where u and v are drawn from normal distributions with an average of 0 and a standard deviation of 1. u and v are defined as:

$$u \sim N(0, \sigma_u^2) \tag{12}$$

$$v \sim N(0, \sigma_v^2) \tag{13}$$

where

$$\sigma_u = \left\{ \frac{\Gamma(1+\beta) \times sin\left(\pi\beta/2\right)}{\Gamma\left[(1+\beta)/2\right] \times \beta \times 2^{(\beta-1)/2}} \right\}^{1/\beta} \tag{14}$$

$$\sigma_v = 1 \tag{15}$$

and in the last step s is calculated as

$$s = 0.01 \times s \tag{16}$$

3.5. Improved Levy Flight Shuffled Frog Leaping Algorithm based on Firefly Algorithm

The goal of an optimization problem is to find the best of all potential solutions. In more formal terms, locate a solution in the possible solutions that has the objective function's minimum/maximum value. Optimization problems have some limitations depending on the nature of the problem. These constraints need to be determined before proceeding to the solution of the problem. Also, a fitness function must be determined to measure the success of the solution found. In the line of this information, the stages of creating optimization problems can be specified as follows:

- Determine the parameters that need to be adjusted to solve the problem

$$X = [x_1, x_2, \ldots, x_n] \tag{17}$$

- Determine the fitness function and whether to maximize or minimize

$$f(x), \;\; maximize\ or\ minimize \tag{18}$$

- Identify the constraints of the problem. These may be some constraints on the parameter sought for the solution, depending on the nature of the problem.
- Determine the maximum/minimum values of the searched parameters.

$$lb_i < x_i < ub_i, \qquad i = 1,2, \ldots, n \tag{19}$$

here, lb and ub are lower-bound and upper-bound of the parameters, respectively.

Optimization problems are structures that are constantly looking for the best possible answer. Today, solutions to optimization problems in various structures are sought in a variety of fields. When we look at the literature, we can find that different techniques for solving these challenges have been offered. Meta-heuristic approaches are used in the majority of the solution proposals.

One of the main aims of this study is to propose a new meta-heuristic approach. While creating this method, the efficiency of the SFLA method in the search space is first developed with the Levy Flight method. Then, a new estimation method is developed by using the SFLA method and the FA method in hybrid structure. The purpose of developing hybrid method is to avoid the disadvantages while using the advantages of the methods together. Another goal to propose a hybrid method is to find better results in a shorter time.

In the proposed method, Levy-Flight distribution strategy is integrated into the Shuffled Frog Leaping Algorithm method (LSFLA) to improve global search. By modifying the local and global search equations of the SFLA, the new individual produced with the Levy-Flight distribution strategy is replaced with the worst individual as follows:

$$X'_{worst} = \tau \times Levy \oplus (X_{best} - X_{worst}) \tag{20}$$

here the operator $\oplus$ denotes the entrywise multiplication. τ is a random number. X_{best} and X_{worst} are the individual with the best and worst positions on the basis of memeplex, respectively.

If the position of the X'_{worst} is not better than the previous one, the update is based on the global best individual. The following are the location updating formulas in this case:

$$X'_{worst} = \tau \times Levy \oplus \left(X_{gbest} - X_{worst}\right) \tag{21}$$

here, X_{gbest} is the best position on the basis of population. If there is still no improvement for the individual with the worst fitness value, a new individual is randomly generated and replaced with the worst individual.

The SFLA is an optimization method with a strong local search direction due to its mementic structure. The division of the population into memeplexes allows individuals to search the solution space in parallel. In addition, as frogs in each memeplex group are affected by each other, they also interact with frogs in other memeplexes. Also, the use of the triangle probability approach at the stage of memetic evolution ensures that strong frogs are more efficient during evolution than weak frogs. This aspect of the algorithm plays a role in making the next generation stronger. After the local search performed by the memeplexes within themselves, the entire population is mixed and the search continues by creating new memeplexes. This allows frogs from different regions of the search space to come together and thus interact with individuals with different memes (Eusuff et al., 2006).

The use of different optimization methods in the hybrid structure provides an effective and successful approach in finding the optimum result. At this step of the study, a new hybrid meta-heuristic optimization algorithm is proposed. In the proposed algorithm, the LSFLA and FA methods are used together in the hybrid structure (ILSFLAFA). This proposed hybrid algorithm can find the global solution with fast search feature without trapping into in the local optimum by balancing the SFLA and FA methods appropriately. A common population is used in this hybrid system. The steps of the proposed hybrid optimization method (ILSFLAFA) are as follows:

1. The search space is first constructed with randomly assigned individuals, like in other metaheuristic approaches.
2. The parameters of the LSFLA and FA algorithms are determined and initial values are assigned to these parameters.
3. The fitness function is determined and the fitness value of each individual is calculated between the specified range.
4. The population is sorted by fitness values.
5. The population is divided into a specified number memeplexes.
6. Memetic evolution is performed using the Levy Flight method.
7. All memeplexes are combined and the population is reordered.
8. The population is divided into memeplexes.
9. Individuals in the last one third of the memeplexes in the ranking are sent to the FA algorithm.

10. Individuals from the FA algorithm are reintegrated into the population and the population is ranked according to their fitness values.

11. If the termination criterion is met, the most successful individual is presented. If the termination criterion is not met, it is returned to step 5.

The detailed flow diagram of the proposed method is as shown in the Figure 3.8.

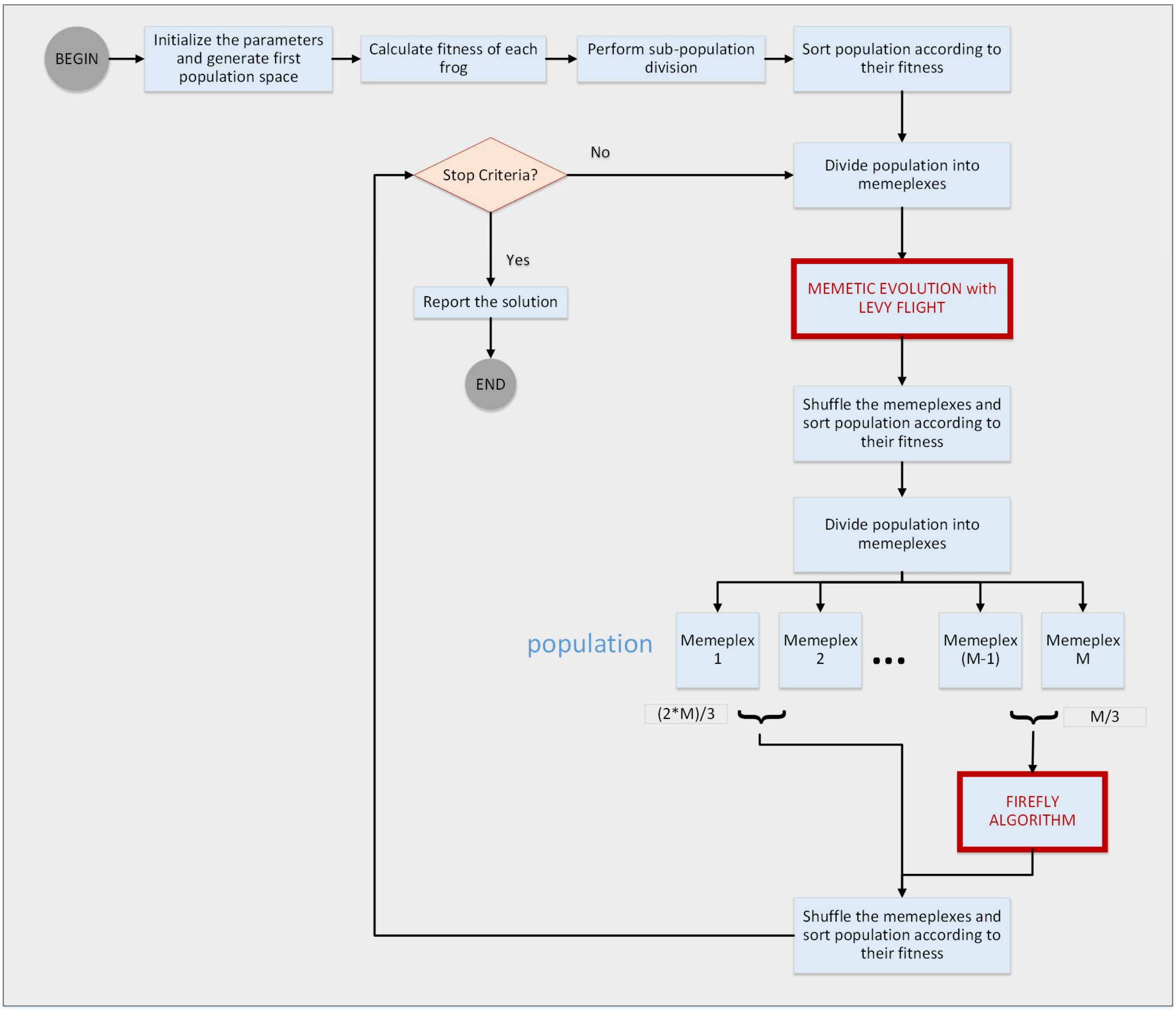


Figure 3.8. General flow diagram of the hybrid optimization procedure

As seen in the detailed flow diagram given Figure 3.8, two improvements are applied to the algorithm. The first improvement is the Levy Flight function, which is applied to the SFLA algorithm in order to reduce the probability of the trapping into the local minimum. The second improvement is the use of the FA algorithm in a hybrid structure in order to integrate the advantage of the fast solution direction into the proposed method.

3.6. Proposed HSBNN

The main purpose of this study is to forecast the CO^2 emission in Türkiye with the minimum error. For this purpose, a new estimation method is proposed in this study and it is developed by optimizing artificial neural networks using the metaheuristic optimization algorithms. In order to get the best result from an artificial neural network, the network coefficients should be optimally adjusted. In the proposed estimation method, the coefficients (weights and biases) of the artificial neural network are adjusted with improved Levy Flight Shuffled Frog Leaping Algorithm based on Firefly Algorithm (ILSFLAFA), which is a method developed by using meta-heuristics optimization algorithms in a hybrid structure.

In the estimation step, a multilayer feedforward neural network is used. The training of the network coefficients is provided by the proposed hybrid method. Meta-heuristic methods in hybrid structure are used to;

- improve forecasting performance,
- increase the ability to train,
- improve the speed of convergence and
- find the optimum result in the search space.

Heuristic methods are utilized in the decision-making phase to optimize the bias and weight values of the feedforward multilayer neural network. Each individual in the heuristic methods consists of the weight and bias values of the network structure. Initially, each individual in the search space is created. The fitness value of each individual in the search space is the error value of the network structure obtained by using the values held by that individual in the network. In this step, Normalized Mean Squared Error (NMSE) is used as the error criterion. MSE is calculated as:

$$MSE = \frac{\sum_{i=1}^{n}\left(x_{actual} - x_{predicted}\right)^2}{n} \tag{22}$$

here, x_{actual} and $x_{predicted}$ are represents the measured and predicted output values, respectively. n is the number of samples. The NMSE is obtained by dividing the MSE by the mean variance of the target rows. A detailed flowchart of the proposed Hybrid Swarm-Based MLNN (HSBNN) as a new estimation method in this thesis is presented in Figure 3.9.

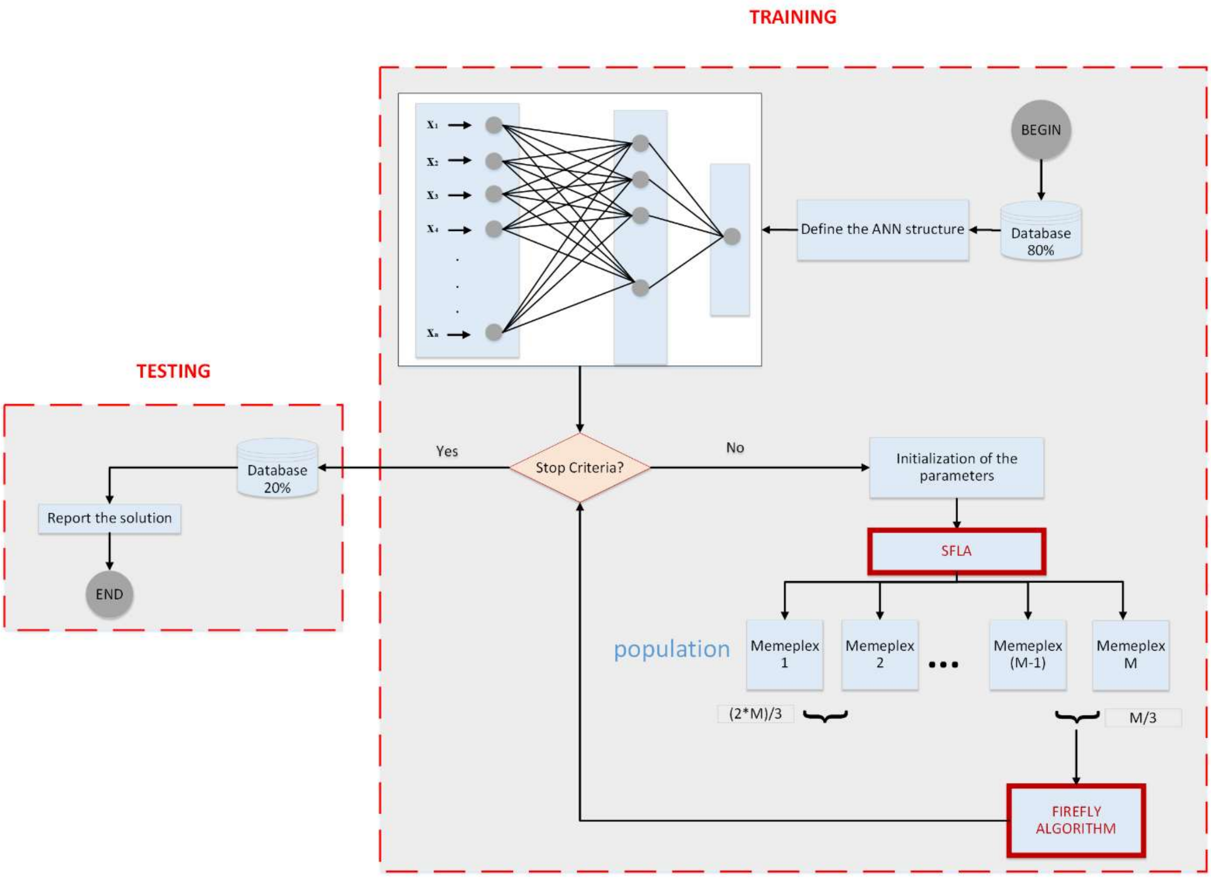


Figure 3.9. Flowchart of the proposed Hybrid Swarm-Based ANN method

The steps of the proposed Hybrid Swarm-Based MLNN (HSBNN) method, as detailed in the Figure 3.9, are as follows:

- The data set to be used is initially prepared in this procedure.
- Normalize the data set and separate it into a training set and a testing set.
- Then, the multilayer feed forward network structure that will be employed to solve the problem is decided. The number of neurons in the input layer network is determined in this stage based on the data set. The number of neurons in the output layer is decided according to the problem. (The number of neurons in the hidden layer are decided by trial and error.)

  Training phase

- In the next step, the training of the network is started. In this step, the parameters of the metaheuristic optimization methods to be used are determined and the initial population is created.
- The solution space is searched using the SFLA.
- The solution space is searched using the FA (with one third of the memeplexes).
- The training phase of the network continues until the termination criterion is met.

  Testing phase
- After the training of the network is provided, the testing phase begins. Prediction is performed by applying test data to the trained network.
- Denormalize the obtained data set and present the results.

Normalization is a pre-processing stage of any form of problem statement. In the fields of machine learning and artificial intelligence, normalization is very significant. There are several different types of normalization procedures, including Min-Max normalization, Z-score normalization, and Decimal scaling normalization (Eusuff et al., 2021). When the data has different units, the training and testing data need to be normalized to reduce the errors caused by this issue. The Z-score normalization method is used for normalization in this study (Eusuff et al., 2021), (Patro and Sahu, 2015):

$$x_{scaled} = \frac{x - mean(x)}{std\_dev(x)} \tag{23}$$

where x is the original value. $mean(x)$ and $std\_dev(x)$ are the mean and standard deviation of the x, respectively. Mean value of x is calculated as:

$$mean(x) = \frac{\sum_{i=1}^{n} x_i}{n} \tag{24}$$

where n is the number of data.

4. RESULTS AND DISCUSSIONS

At this stage of the thesis, firstly, a new optimization method in hybrid structure is proposed. The success of this proposed method has been tested with Benchmark functions. Then, using this proposed method, a new artificial neural network-based estimation method is developed. This developed estimation method has been used for the estimation of the amount of CO^2 emission in Türkiye, taking into account the amount of use of renewable and non- renewable energy sources, and the results are given in detail. In the last stage, the future CO^2 emissions of Türkiye has been forecasted until 2030.

4.1. Results of the Proposed ILSFLAFA

Today, optimization methods in different structures are proposed to solve many real-life engineering problems. Benchmark functions are used to test the convergence rate, parameter sensitivity and robustness of these methods against uncertainties. Benchmark functions are test functions with different properties that are used to test the performance of optimization methods from different aspects. These functions are generally accepted in terms of comparison of optimization methods.

In this thesis, benchmark functions are used to show the success and convergence speed of the proposed method as in most optimization studies in the literature. Many benchmark functions have been proposed by different researchers until today. These functions can be grouped according to the number of local optima and the frequency of finding local optima: unimodal and multimodal.

Here, the most frequently used 12 Benchmark Functions are preferred and equation, type, limit values and optimum value information of the functions are given in Table 4.1.

Table 4.1. Benchmark Functions

Func.	Benchmark problem	Formulation	Search range	Initial range	Category*	Global min.
F1	Sphere	$f(x) = \sum_{i=1}^{n} x_i^2$	[-100, 100]	[-100, 50]	U	0
F2	Schwefel 2.22	$f(x) = \sum_{i=1}^{n} \lvert x_i \rvert + \prod_{i=1}^{n} \lvert x_i \rvert$	[-10, 10]	[-10, 5]	U	0
F3	Sum Square	$f(x) = \sum_{i=1}^{n} i x_i^2$	[-10, 10]	[-10, 10]	U	0
F4	Penalized 1	$f(x) = \frac{\pi}{n}\{10\, sin^2(\pi y_1) + \sum_{i=1}^{n-1} (y_i - 1)^2 [1 + 10\, sin^2(\pi y_{i+1})] + (y_n - 1)^2\} + \sum_{i=1}^{n} u(x_i, 10, 100, 4)$ $y_i = 1 + \frac{1}{4}(x_i + 1) u_{x_i,a,k,m} = \begin{cases} k(x_i - a)^m x_i > a \\ 0 - a \leq x_i \leq a \\ k(x_i - a)^m x_i < -a \end{cases}$	[-50, 50]	[-50, 25]	M	0
F5	Penalized 2	$f(x) = \frac{1}{10}\{sin^2(\pi x_1) + \sum_{i=1}^{n-1} (x_i - 1)^2 [1 + sin^2(3\pi x_{i+1})] + (x_n - 1)^2[1 + sin^2(2\pi x_{i+1})]\} + \sum_{i=1}^{n} u(x_i, 5, 100, 4)$	[-50, 50]	[-50, 25]	M	0
F6	Cigar	$f(x) = x_i^2 + 10^6 \sum_{i=2}^{n} x_i^2$	[-100, 100]	[-100, 100]	U	0
F7	Schwefel 1.2	$f(x) = \sum_{i=1}^{n} \left(\sum_{j=1}^{i} x_j \right)^2$	[-100, 100]	[-100, 100]	U	0
F8	Zakharov	$f(x) = \sum_{i=1}^{n} x_i^2 + \left(\sum_{i=1}^{n} 0.5 i x_i \right)^2 + \left(\sum_{i=1}^{n} 0.5 i x_i \right)^4$	[-5, 10]	[-5, 10]	M	0
F9	Alpine	$f(x) = \sum_{i=1}^{n} \lvert x_i sin(x_i) + 0.1 x_i \rvert$	[-10, 10]	[-10, 10]	M	0
F10	Schaffer	$f(x) = 0.5 + \frac{sin^2\left(\sqrt{\sum_{i=1}^{n} x_i^2}\right) - 0.5}{\left(1 + 0.001\left(\sum_{i=1}^{n} x_i^2\right)\right)}$	[-100, 100]	[-100, 100]	U	0
F11	Chung Reynolds	$f(x) = \left(\sum_{i=1}^{n} x_i^2 \right)^2$	[-100, 100]	[-100, 100]	M	0
F12	Generalized Griewank	$f(x) = \frac{1}{4000} \sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	[-600, 600]	[-600, 200]	U	0

* M - Multimodal functions and U - Unimodal functions

Benchmark functions may show different characteristics and can be classified as unimodal, multimodal and rotated multimodal functions as indicated in the table. While unimodal functions such as Sphere, Schwefel 2.22, Sum Square, Cigar, Schwefel 1.2, Schaffer and Generalized Griewank functions have a single optimal solution, multimodal functions such as Penalized 1, Penalized 2, Zakharov, Alpine and Chung Reynolds have more than one local minimum. Different benchmark functions are used to examine the trapping into the local minima of the methods used and proposed in this study. In the table, it is indicated to which category (column 6) the specified function belongs.

The proposed method (ILSFLAFA) in this thesis is obtained by hybridizing the FA and SFLA algorithms and additionally making some improvements. Therefore, in order to demonstrate the success of the proposed method, the results obtained using the Benchmark functions are compared with the results obtained by the FA and SFLA methods. The parameters of the optimization methods are given in Tables 4.2 - 4.3.

Table 4.2. Optimization parameters for the FA

Algorithm	**Parameters**	**Value**
FA	Population size	24
	Fitness evaluations	400
	Number of decision variables	30
	Light absorption coefficient	2.0
	Attraction coefficient	1.8

Table 4.3. Optimization parameters for the SFLA

Algorithm	**Parameters**	**Value**
SFLA	Population size	24
	Fitness evaluations	400
	Number of decision variables	30
	Number of Memeplexes	6
	Memeplex Size	4
	Number of Fitness evaluations for local search	4

Table 4.4. Optimization parameters for the proposed algorithm

Algorithm	**Parameters**	**Value**
ILSFLAFA	Population size	24
	Fitness evaluations	400 (100x4)
	Number of decision variables	30
	Light absorption coefficient	2.0
	Attraction coefficient	1.8
	Number of Memeplexes	6
	Memeplex Size	4
	Number of Fitness evaluations for local search	4
	Levy index	2.0

In Table 4.5, the results obtained with 400 fitness evaluations for 30-dimensional Benchmark Functions are given. The best optimum values, mean values and standard deviations of the results obtained over 20 runs obtained with the FA, SFLA and the proposed hybrid optimization method are given in Table 4.5. In each function, the best results are indicated by boldface.

Table 4.5. Comparison of the FA, SFLA and ILSFLAFA on unimodal and multimodal Benchmark functions

		FA	**SFLA**	**ILSFLAFA (proposed)**
F1	best	5,37E-18	8,38E-09	**0,00E+00**
	worst	3,30E+00	8,06E-04	**0,00E+00**
	std	7,37E-01	2,93E-04	**0,00E+00**
F2	best	2,74E-16	2,44E-05	**0,00E+00**
	worst	7,50E-03	7,03E-02	**0,00E+00**
	std	2,27E-03	2,14E-02	**0,00E+00**
F3	best	5,42E-201	3,11E-282	**0,00E+00**
	worst	1,64E-198	1,69E-274	**0,00E+00**
	std	0,00E+00	0,00E+00	**0,00E+00**
F4	best	4,71E-31	**4,71E-31**	2,15E-19
	worst	1,24E+01	**4,71E-31**	6,91E-16
	std	4,76E+00	**0,00E+00**	1,61E-16

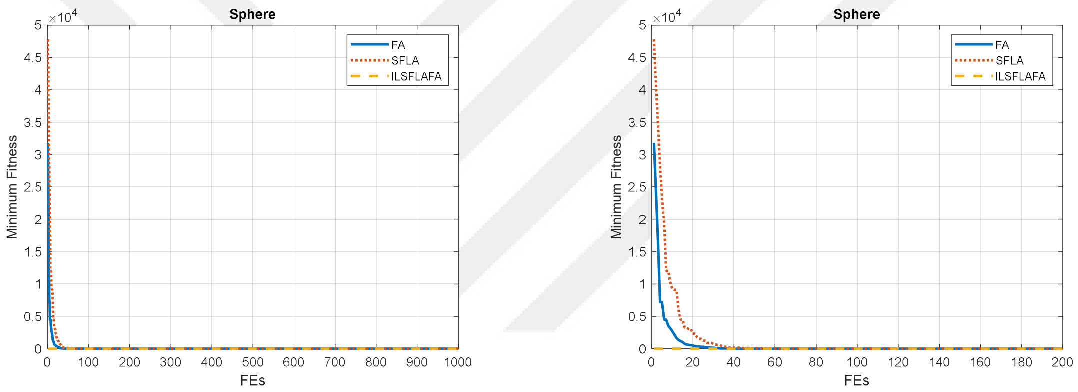
Table 4.5

F5	best	**1,35E-32**	**1,35E-32**	6,00E-14
	worst	**1,35E-32**	**1,35E-32**	5,15E-10
	std	**0,00E+00**	**0,00E+00**	1,36E-10
F6	best	2,85E-203	3,90E-280	**0,00E+00**
	worst	6,87E-199	5,41E-267	**0,00E+00**
	std	0,00E+00	0,00E+00	**0,00E+00**
F7	best	3,88E-202	1,98E-281	**0,00E+00**
	worst	1,06E-198	2,16E-273	**0,00E+00**
	std	0,00E+00	0,00E+00	**0,00E+00**
F8	best	6,46E-202	9,97E-284	**0,00E+00**
	worst	1,40E-198	9,50E-273	**0,00E+00**
	std	0,00E+00	0,00E+00	**0,00E+00**
F9	best	0,00E+00	0,00E+00	**0,00E+00**
	worst	3,34E-101	8,17E-137	**0,00E+00**
	std	7,72E-102	1,99E-137	**0,00E+00**
F10	best	1,78E-01	3,72E-02	**0,00E+00**
	worst	3,96E-01	1,78E-01	**0,00E+00**
	std	6,05E-02	3,20E-02	**0,00E+00**
F11	best	1,57E-22	9,61E-12	**0,00E+00**
	worst	3,50E-11	6,15E-04	**0,00E+00**
	std	8,37E-12	1,45E-04	**0,00E+00**
F12	best	2,56E-13	1,91E-07	**0,00E+00**
	worst	1,58E-01	4,83E-01	**0,00E+00**
	std	3,63E-02	1,06E-01	**0,00E+00**

As the results obtained from the benchmark functions are examined, it is seen that the proposed method is generally successful compared to the other two methods. By analyzing the result of unimodal functions Sphere, Schwefel 2.22, Sum Square, Cigar, Schwefel 1.2, Schaffer, and Generalized Griewank are analyzed, it is clear that the suggested method is more successful than other algorithms in achieving the global optimal solution.

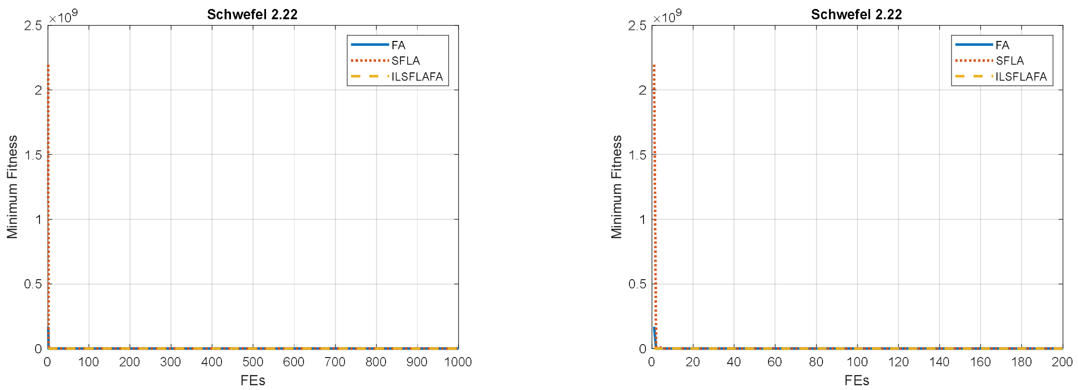
In the review of the results, it is seen that for functions Zakharov, Alpine and Chung Reynolds, which have a multimodal structure, the proposed method has found the global optimum solution without getting stuck on the local optimum results. Besides, the SFLA found the best results in function Penalized 1, while the FA and SFLA found the best results in function Penalized 2. As comparing the ILSFLAFA approach to the other two swarm-based algorithms, it has been seen that the ILSFLAFA method achieves more effective and successful results in majority of the benchmark functions.

In optimization methods, the convergence speed of that method to finding the optimal solution is also important. The convergence speed is of great importance in performance comparison. Convergence graphs are shown in Figures 4.1 - 4.12 to compare the convergence rates of the FA, SFLA, and proposed method. The results are also given by zooming in order to see the convergence more clearly.



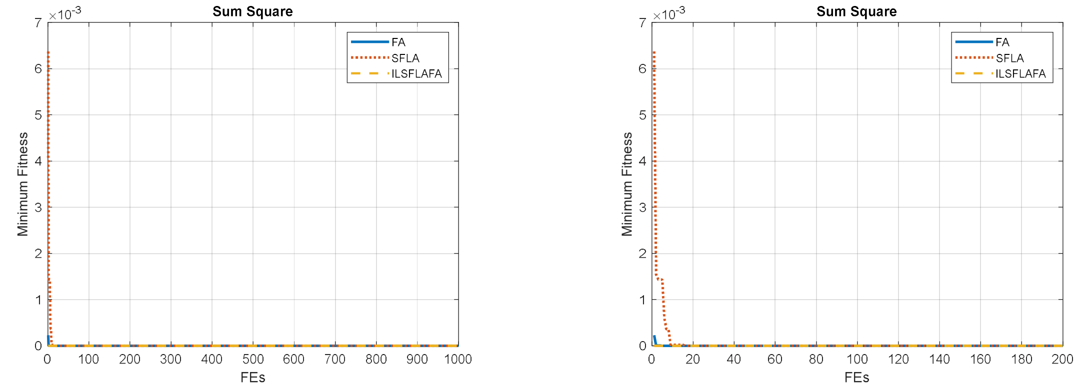
(a) (b)

Figure 4.1. Convergence graphs for Sphere function (a) for 1000 iterations (b) zoomed part



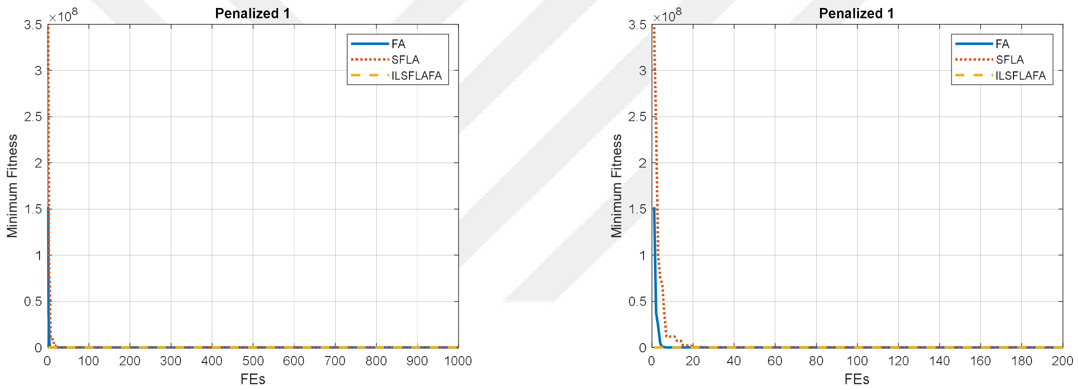
(a) (b)

Figure 4.2. Convergence graphs of Schwefel 2.22 function (a) for 1000 iterations (b) zoomed part



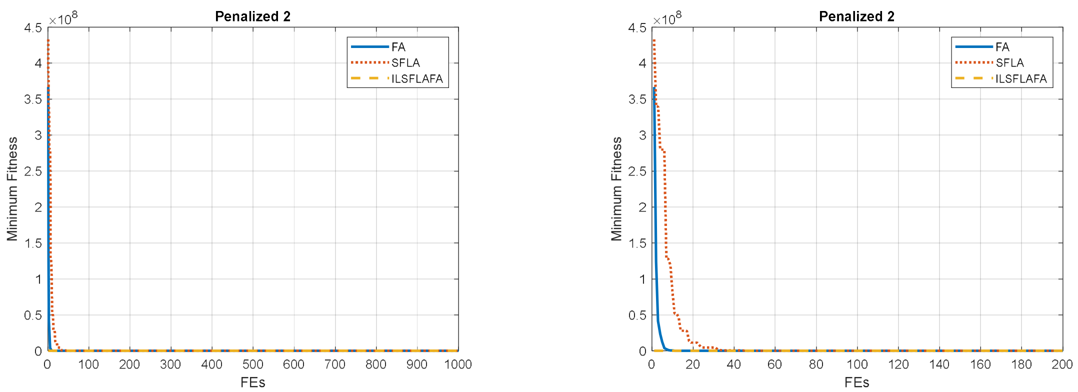
(a) (b)

Figure 4.3. Convergence graphs of Sum Square function (a) for 1000 iterations (b) zoomed part



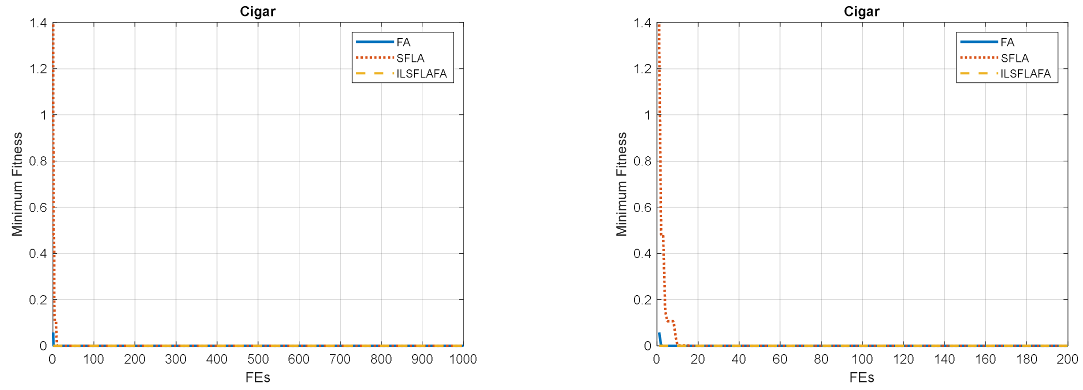
(a) (b)

Figure 4.4. Convergence graphs of Penalized 1 function (a) for 1000 iterations (b) zoomed part



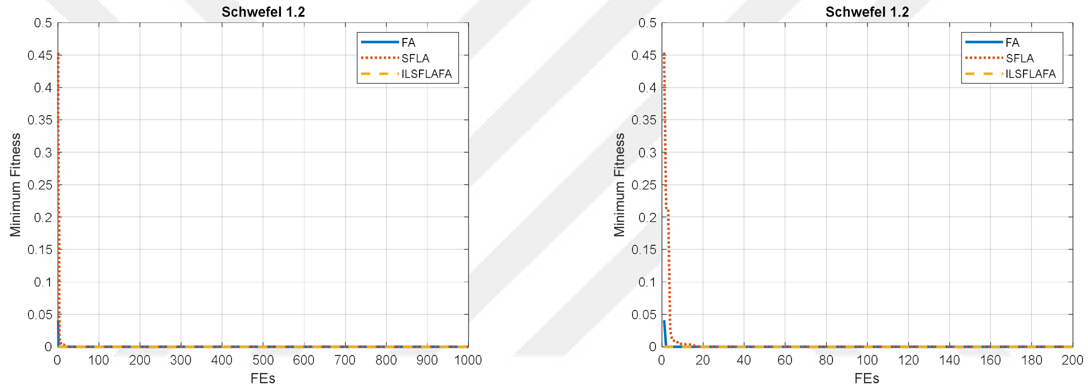
(a) (b)

Figure 4.5. Convergence graphs of Penalized 2 function (a) for 1000 iterations (b) zoomed part



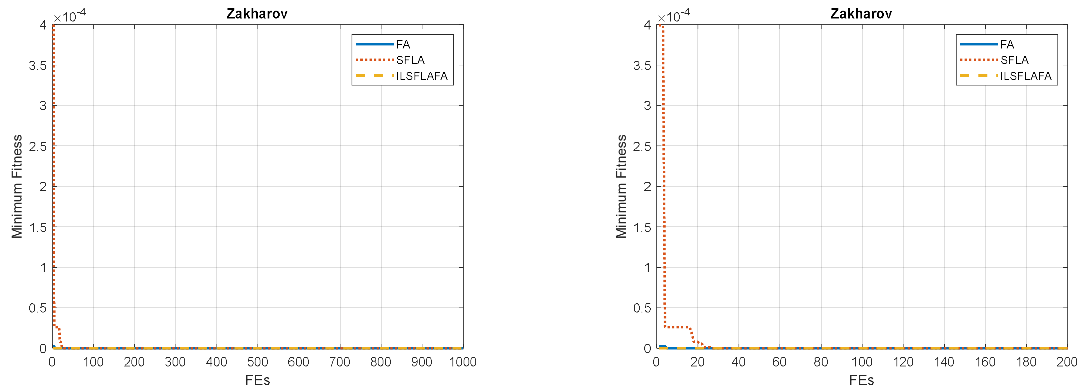
(a) (b)

Figure 4.6. Convergence graphs of Cigar function (a) for 1000 iterations (b) zoomed part



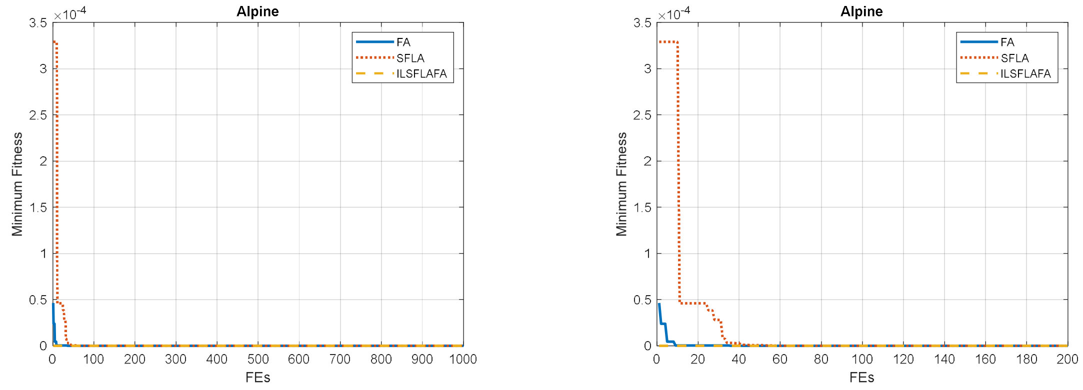
(a) (b)

Figure 4.7. Convergence graphs of Schwefel 1.2 function (a) for 1000 iterations (b) zoomed part



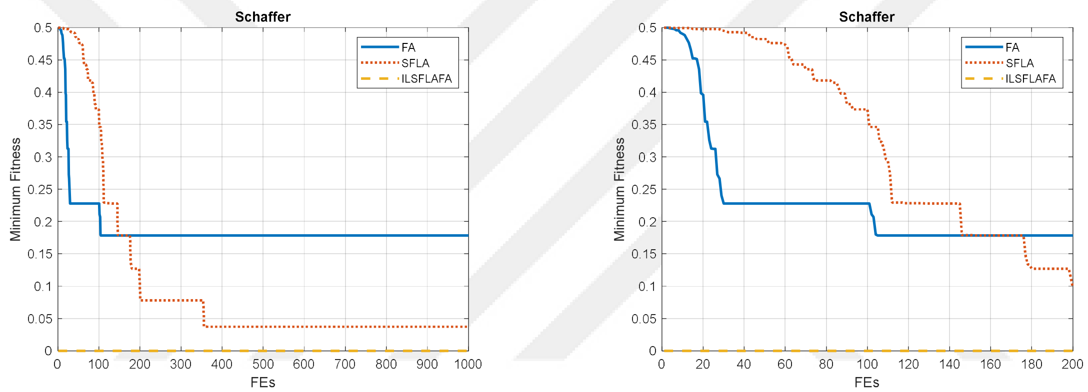
(a) (b)

Figure 4.8. Convergence graphs of Zakharov function (a) for 1000 iterations (b) zoomed part



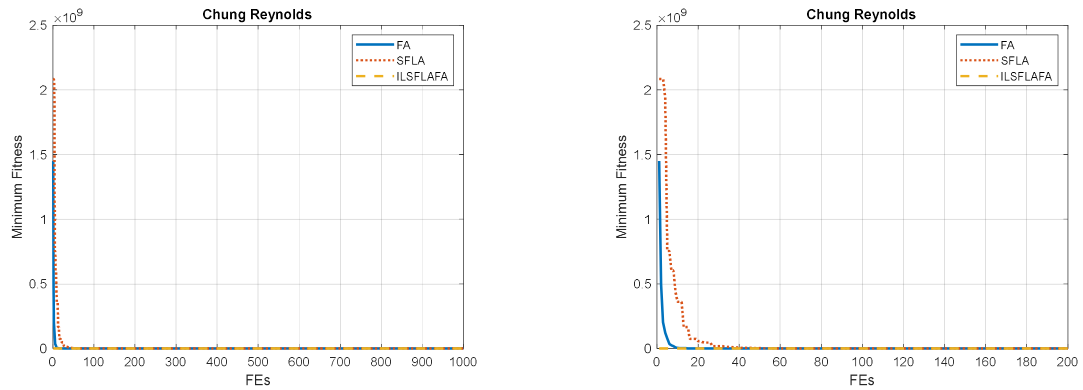
(a) (b)

Figure 4.9. Convergence graphs of Alpine function (a) for 1000 iterations (b) zoomed part



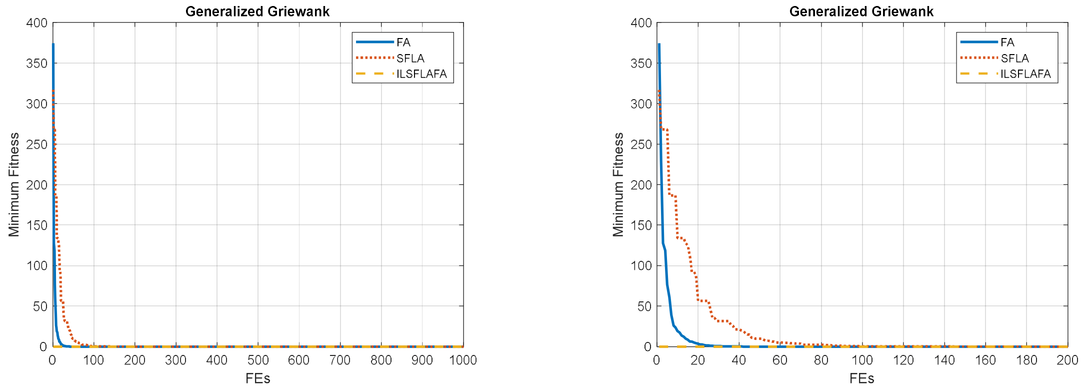
(a) (b)

Figure 4.10. Convergence graphs of Schaffer function (a) for 1000 iterations (b) zoomed part



(a) (b)

Figure 4.11. Convergence graphs of Chung Reynolds function (a) for 1000 iterations (b) zoomed part



(a) (b)

Figure 4.12. Convergence graphs of Generalized Griewank function (a) for 1000 iterations (b) zoomed part

As the numerical results obtained as a result of the application of the FA, SFLA and ILSFLAFA methods to 12 different Benchmark Functions are examined, the success of the proposed method is clearly seen. In addition, when the graphs in Figures 4.1 - 4.12 are examined to compare the converge speeds of the optimization methods, it is seen that the proposed method reaches the optimum solution in the first 100 fitness evaluations.

As a result, numerical and graphical results of the three approaches show that the proposed method mostly finds the optimum result with fast convergence, making it the method with the best performance.

4.2. Results of the Proposed HSBNN

This work aims to contribute to the studies on this subject by developing an estimation method for Türkiye's CO^2 emissions in the future. There are different studies carried out to date for the estimation of CO^2 emissions in Türkiye. One of these studies is carried out by Özceylan in 2016 [22] (Özceylan, 2016). CO^2 estimation method is modeled by Özceylan using the linear, exponential and quadratic equation forms. Particle Swarm Optimization (PSO) and Artificial Bee Colony (ABC) approaches are employed in the development of these models. Energy consumption, population, gross domestic product (GDP), and number of motor vehicles data of Türkiye's for the years 1980-2008 are used in this estimation model as socio-economic indicators. At the end of the article, emission forecasts until 2030 are made using different scenarios.

In this thesis, an artificial neural network-based prediction method has been developed for the estimation of the CO^2 emissions. In this developed method, a multilayer feedforward artificial neural network is used. The proposed method is first applied to the data of Özceylan's article and the results are compared. The data set in Özceylan's paper is as given in Table 4.6.

Table 4.6. Socio-economic data for CO^2 emission (Özceylan, 2016), (Indexmundi, 2013)

Year	Energy consumption (mtoe)	Population (10^6)	GDP ($\$10^9$)	Number of motors vehicles (10^6)	CO^2 emission (mt)
1980	31.970	44.439	68.789	1.431	75.702
1981	32.050	45.540	71.040	1.513	79.809
1982	34.390	46.688	64.546	1.626	86.917
1983	35.700	47.864	61.678	1.761	90.468
1984	37.430	49.070	59.990	1.927	95.718
1985	39.400	50.307	67.235	2.098	106.630
1986	42.470	51.480	75.728	2.275	116.786
1987	46.880	52.370	87.173	2.484	129.801
1988	47.910	53.268	90.853	2.705	126.206
1989	50.710	54.192	107.143	2.895	139.203
1990	52.980	55.120	150.676	3.228	150.667
1991	54.270	56.055	151.041	3.548	151.675
1992	56.680	56.986	159.095	3.974	154.368
1993	60.260	57.913	180.422	4.568	164.125
1994	59.120	58.837	130.690	5.047	161.527
1995	63.680	59.756	169.486	5.355	176.561
1996	69.860	60.671	181.476	5.753	192.342
1997	73.780	61.582	189.835	6.282	202.722
1998	74.710	62.464	269.287	6.923	205.254
1999	76.770	63.364	249.751	7.377	196.607
2000	80.500	64.252	266.568	7.966	215.971
2001	75.400	65.133	196.005	8.453	194.379
2002	78.330	66.008	232.535	8.612	205.510

Tablo 4.6

2003	83.840	66.873	303.005	8.882	218.330
2004	87.820	67.723	392.166	9.686	225.222
2005	91.580	68.566	482.980	10.666	237.174
2006	100.580	69.395	530.900	11.945	261.357
2007	109.060	70.215	647.155	12.708	288.445
2008	113.850	71.517	730.337	13.512	297.120

The first 25 years of data (1980-2004) are used to validate the models and set the weighting parameters, while the last five years of data (2005-2008) are used to test the models. At the end of the study, relative error is used as the error criterion. The equation for the relative error criterion is as follows:

$$Relative\ Error = \frac{absolute\ error}{actual\ value} = \frac{\left(x_{predicted} - x_{actual}\right)}{x_{actual}} \qquad (25)$$

here, x_{actual} and $x_{predicted}$ are represents the measured and predicted output values, respectively. Özceylan used 40 individuals in the search spaces of swarm-based methods, and the methods ran 10000 fitness evaluations during the creation of the models.

The feedforward artificial neural network utilized in the study for this data set has four neurons in the input layer, four neurons in the hidden layer, and one neuron in the output layer, according to the method provided in the thesis study. The network structure used is as shown in Figure 4.13. 25 data sets from 1980 to 2004 are utilized to train the network. The parameters whose optimum values are sought during network training are:

- 24 parameters for weights between input and hidden layers,
- 6 parameters for weights between hidden and output layers,
- 6 parameters for biases at hidden layer and
- 1 parameter for biases at output layer.

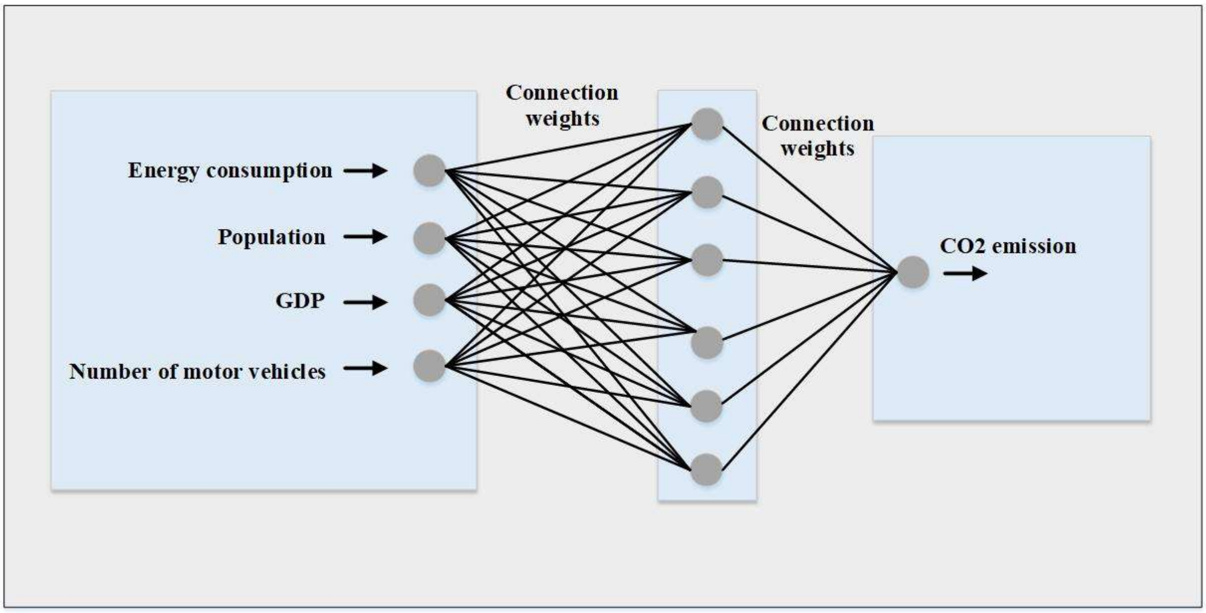


Figure 4.13. Architecture of the Multi-Layer Neural Network (MLNN) with the structure of 4-6-1 (four inputs, one output and six neurons in the hidden layer)

The results given at the end of Özceylan's article, the results obtained with the HSBNN method and the relative errors are as shown in the Table 4.7.

Table 4.7. The results obtained with the HSBNN method and the relative errors (Özceylan, 2016)

	2005	**2006**	**2007**	**2008**
Actual data (Indexmundi, 2013)	237.174	261.357	288.445	297.120
HSBNN	**239.3081**	**259.8382**	**286.4580**	**299.8035**
Relative Error %	**0.8998**	**-0.5811**	**-0.6889**	**0.9032**
PSO^{CO2} linear	231.399	251.480	272.824	282.906
Relative Error %	-2.435	-3.779	-5.416	-4.784
PSO^{CO2} exponential	236.319	254.021	270.648	279.787
Relative Error %	-0.361	-2.807	-6.170	-5.834
PSO^{CO2} quadratic	227.135	240.122	273.136	289.881
Relative Error %	-4.233	-8.125	-5.307	-2.436
ABC^{CO2} linear	231.399	251.468	272.806	282.888
Relative Error %	-2.435	-3.784	-5.422	-4.790

Tablo 4.7.

ABC^{CO2} exponential	226.430	242.060	258.865	263.508
Relative Error %	-4.530	-7.383	-10.255	-11.313
ABC^{CO2} quadratic	234.139	245.389	285.749	312.542
Relative Error %	-1.280	-6.110	-0.935	5.191

In the HSBNN method, 40 individuals are used in the study and the training of the network is performed in 1000 fitness evaluations. Optimization parameters of the HSBNN method are given in Table 4.8.

Table 4.8. Optimization parameters of the proposed prediction model

Parameters	**Value**
Number of input neurons	4
Number of hidden neurons	6
Number of output neurons	1
Fitness evaluations	1000 (100x10)
Population size	24
Number of Memeplexes	10
Memeplex Size	4
Number of Fitness evaluations for local search	5
Levy index	2.0
Light absorption coefficient	1.0
Attraction coefficient	2.0

In the test phase, data from 2005 to 2008 are used. The average relative errors of the proposed models are as shown in Table 4.9.

Table 4.9. Average relative errors of proposed models (Özceylan, 2016)

Prediction Models	**Average Relative Errors**
HSBNN	**0.7682**
PSO^{CO2} linear	4.104
PSO^{CO2} exponential	3.793

Table 4.9

PSO^{CO2} quadratic	5.025
ABC^{CO2} linear	4.108
ABC^{CO2} exponential	8.37
ABC^{CO2} quadratic	3.379

By examining Table 4.7, it can be seen that the estimation results achieved using various methods are closer to the real values in some cases. The mean relative errors of the results are given in Table 4.9 so that the method successes may be seen more clearly. The comparison of the results obtained with the HSBNN and methods used in the article is graphically shown in the Figure 4.14.

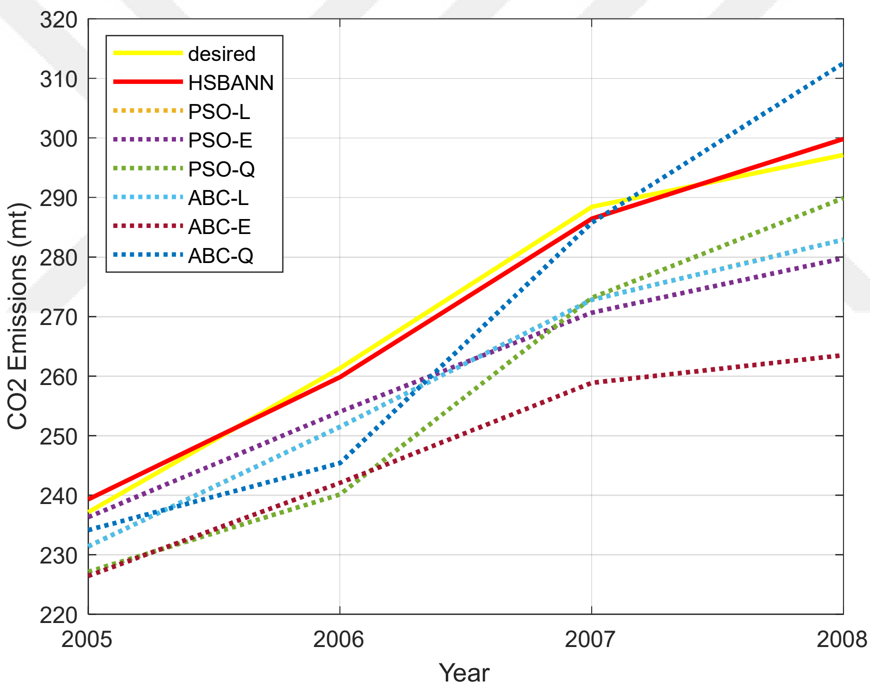


Figure 4.14. Comparison of the desired data and the proposed models (Özceylan, 2016)

In the reviewed article, CO^2 estimation is made with six different methods. The HSBNN method proposed in this thesis is also applied to the same data set. The desired CO^2 emission values and the results obtained from these methods are given in the Table 4.7 and Table 4.9 and Figure 4.14. In the estimation results for the years 2005-2008, ABC^{CO2} quadratic model obtained the best result from the methods used in the article. While the average relative error of the prediction result obtained with this model is 3.379, the average relative error value obtained with the HSBNN is 0.7682. When these results are examined in detail, it is clear that the proposed HSBNN method produces more successful and effective results.

In this thesis, hybrid swarm-based metaheuristic optimization method is used in four different structures and is used in the training phase of the multilayer feedforward neural network. According to the literature, it is seen that socio-economic factors are used for the estimation of CO^2 emissions. In addition, in this study, the renewable energy factor is also used to estimate the amount of emissions. The number of neurons in the input layer is determined by the data set used in studies using artificial neural network-based approaches. The number of neurons of output layer is determined by the problem structure. In this study, the inputs of the neural network are:

- Year,
- R&D investments (TurkStat, 2020),
- Renewable energy ratio of total final energy consumption (World Bank),
- Population (Indexmundi),
- Urbanization (Indexmundi),
- Number of motor vehicles (TurkStat, 2020),
- Energy consumption (US Energy Information Administration (EIA)),
- GDP (World Bank).

In addition, the output of the network is amount of CO^2 emissions (World Bank).

The data set are obtained from different institutions. The historical values for the data titles that are desired to be used are limited. The provided data covers the years 1990 to 2018. At the beginning of the study, the data is divided into two separate sub-set. First set, that involves data from 1990 to 2012, are used for training the proposed forecasting model and while second set, that involves data from 2013-2018, are employed to test the accuracy of the forecasting model. To estimate the CO^2 emission amount, the ANN structure in HSBNN is used in four different ways. In the first proposed method, one hidden layer is used. In the other three methods, there are two layers in the hidden layer. In these four methods, four different estimation methods have been developed by using different neuron numbers in the hidden layers to obtain the best prediction model by examining the obtained results.

First prediction model (HSBNN[1])

In the first model developed for the estimation process, the feed-forward Multi-Layer Neural Network (MLNN) is formed with one hidden layer. MLNN has 8 neurons in the input layer, 12 neurons in the hidden layer, and 1 neuron in the output layer. In the proposed prediction model, the training of the coefficients of the MLNN is provided by ILSFLAFA, in which the improved SFLA and FA optimization methods are used in a hybrid structure. The diagram of the MLNN used in the prediction model is as given in the Figure 4.15.

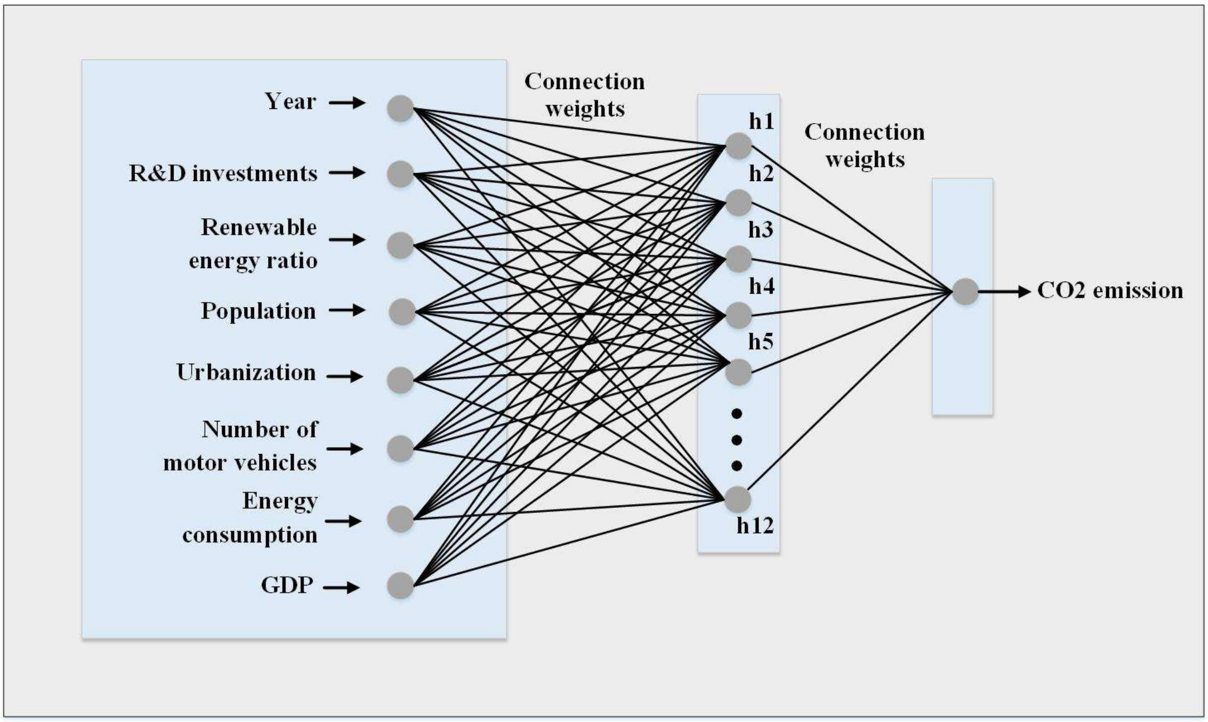


Figure 4.15. The diagram of the artificial neural network used in the prediction model (HSBNN[1])

At the beginning of the study, the initial population of the hybrid optimization method is formed with 24 agents and these agents search the solution space to find the optimum solution. The parameters of the SFLA and FA methods are given in the Table 4.10. The work of the hybrid optimization method consists of two stages. After each SFLA step, the FA method works. In this model, the SFLA method is run in 100 fitness evaluations, and the FA method is run in 10 fitness evaluations.

Table 4.10. Optimization parameters of the HSBNN[1]

Parameters	**Value**
Number of input neurons	8
Number of hidden neurons	12
Number of output neurons	1
Fitness evaluations	1000 (100x10)
Population size	24
Number of Memeplexes	6
Memeplex Size	4
Number of Fitness evaluations for local search	5
Levy index	2.0
Light absorption coefficient	1.0
Attraction coefficient	2.0

The function of the input layer in artificial neural networks is to transmit the incoming signal to the next layer. Therefore, there is no bias value in the input layer. In order to obtain the best prediction result, the parameters that need to be trained for the feed forward multi layered ANN are as follows:

- 96 parameters for weights between input and hidden layers,
- 12 parameters for weights between hidden and output layers,
- 12 parameters for biases at hidden layer and
- 1 parameter for biases at output layer.

Second prediction model (HSBNN[2])

In addition, the prediction model is also built up by using 2 layers in the hidden layer. In this prediction model, 6 neurons are used in the first hidden layer and 4 neurons are used in the second hidden layer. The network coefficients are trained with the ILSFLAFA method. 24 agents are used during the training of the network. The training algorithm is run for 1000 fitness evaluations (100 SFLA x 10 FA). The parameters of the SFLA and FA methods are given in the Table 4.11.

Table 4.11. Optimization parameters of the HSBNN[2]

Parameters	Value
Number of input neurons	8
Number of neurons in the first hidden layer	6
Number of neurons in the second hidden layer	4
Number of output neurons	1
Fitness evaluations	1000 (100x10)
Population size	24
Number of Memeplexes	6
Memeplex Size	4
Number of Fitness evaluations for local search	5
Levy index	2.0
Light absorption coefficient	1.0
Attraction coefficient	2.0

The input parameters of the network are year, R&D investments, renewable energy ratio of total final energy consumption, population, urbanization, number of motor vehicles, energy consumption and GDP. The output layer of the network is the CO^2 emission amount. Signals received from the input layer are transmitted to the output layer through the hidden layers. It is calculated as the error value of the network for that solution for the fitness function required during the adjustment of the optimum values of the network coefficients. The structure of the MLNN with an input, two hidden layers and an output layer are presented in the Figure 4.16.

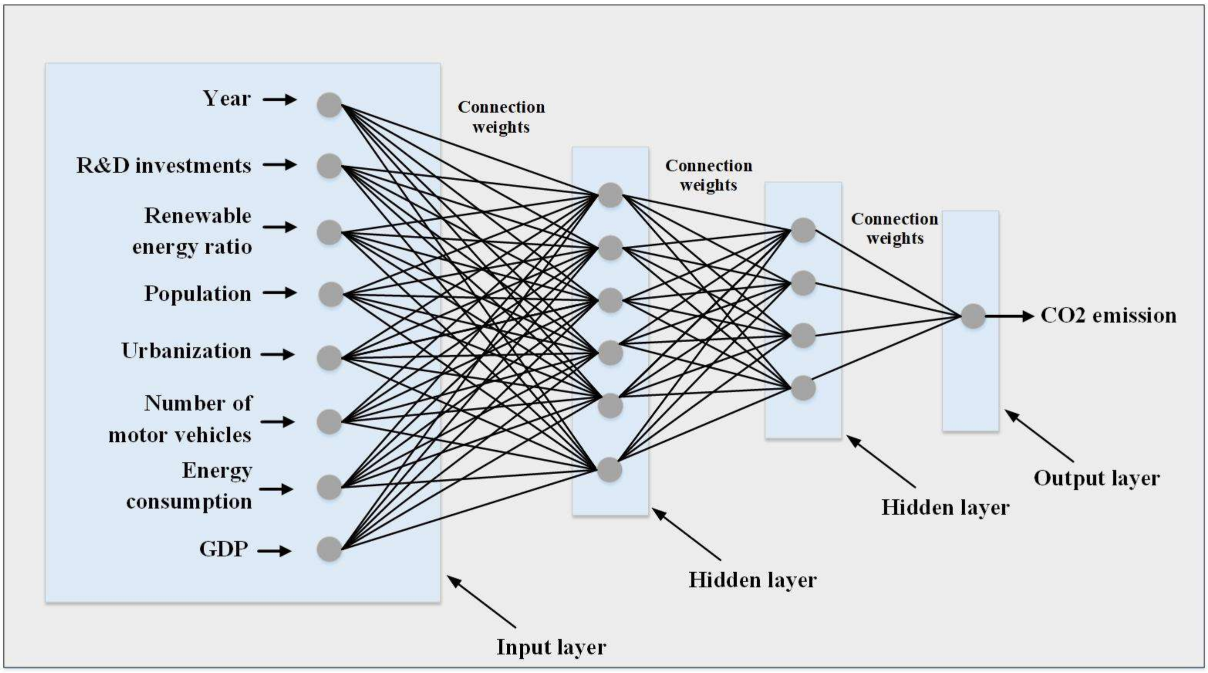


Figure 4.16. The diagram of the artificial neural network with an input, two hidden layers and an output layer used in the prediction model ($HSBNN^2$)

The parameters that need to be trained for the feed forward multi-layered ANN in order to produce the best prediction result are as follows:

- 48 parameters for weights between input and first hidden layers,
- 6 parameters for biases at first hidden layer,
- 24 parameters for weights between first hidden and second hidden layers,
- 4 parameters for biases at second hidden layer,
- 4 parameters for weights between second hidden and output layers and
- 1 parameter for biases at output layer.

Third prediction model ($HSBNN^3$)

In the third estimation method, the artificial neural network is formed with two hidden layers. The number of neurons in the first and second hidden layers are 6. In this estimation method, a search is made in the population space with 24 agents. The training algorithm is run for 1000 fitness evaluations (100 SFLA x 10 FA). The parameters of the hybrid optimization method are given in the Table 4.12.

Table 4.12. Optimization parameters of the HSBNN^{3}

Parameters	**Value**
Number of input neurons	8
Number of neurons in the first hidden layer	6
Number of neurons in the second hidden layer	6
Number of output neurons	1
Fitness evaluations	1000 (100x10)
Population size	24
Number of Memeplexes	6
Memeplex Size	4
Number of Fitness evaluations for local search	5
Levy index	2.0
Light absorption coefficient	1.0
Attraction coefficient	2.0

The input parameters of the network are same as the HSBNN^{1} and HSBNN^{2} methods. The output layer of the neural network is the CO^{2} emission amount. The structure of the artificial neural network with an input, two hidden layers and an output layer is as shown in the Figure 4.17.

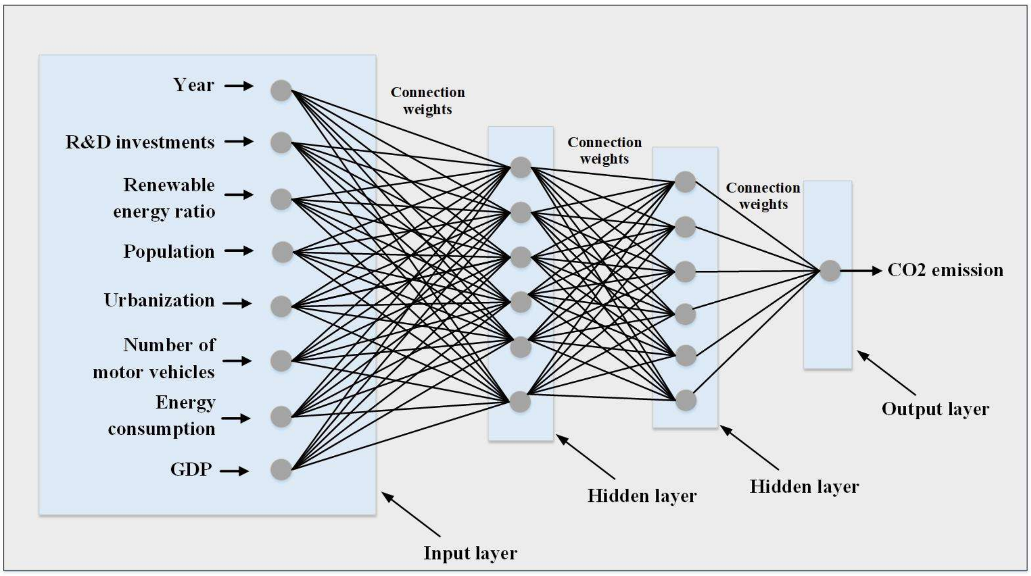


Figure 4.17. The diagram of the artificial neural network with an input, two hidden layers and an output layer used in the prediction model (HSBNN^{3})

The parameters that trained for the feed forward multi-layered neural network in order to produce the best prediction result are as follows:

- 48 parameters for weights between input and first hidden layers,
- 6 parameters for biases at first hidden layer,
- 36 parameters for weights between first hidden and second hidden layers,
- 6 parameters for biases at second hidden layer,
- 6 parameters for weights between second hidden and output layers and
- 1 parameter for biases at output layer.

Fourth prediction model ($HSBNN^4$)

In the fourth estimation method, the artificial neural network is formed with two hidden layers. The number of neurons in the first and second hidden layers are 8 and 4, respectively. In this estimation method, a search is made in the population space with 24 agents. The training algorithm is run for 1000 fitness evaluations (100 SFLA x 10 FA). The parameters of the hybrid optimization method are given in the Table 4.13.

Table 4.13. Optimization parameters of the $HSBNN^4$

Parameters	**Value**
Number of input neurons	8
Number of neurons in the first hidden layer	8
Number of neurons in the second hidden layer	4
Number of output neurons	1
Fitness evaluations	1000 (100x10)
Population size	24
Number of Memeplexes	6
Memeplex Size	4
Number of Fitness evaluations for local search	5
Levy index	2.0
Light absorption coefficient	1.0
Attraction coefficient	2.0

The input parameters are the same as the other three methods. The output layer of the neural network is the CO^2 emission amount. The structure of the artificial neural network with an input, two hidden layers and an output layer is as shown in the Figure 4.18.

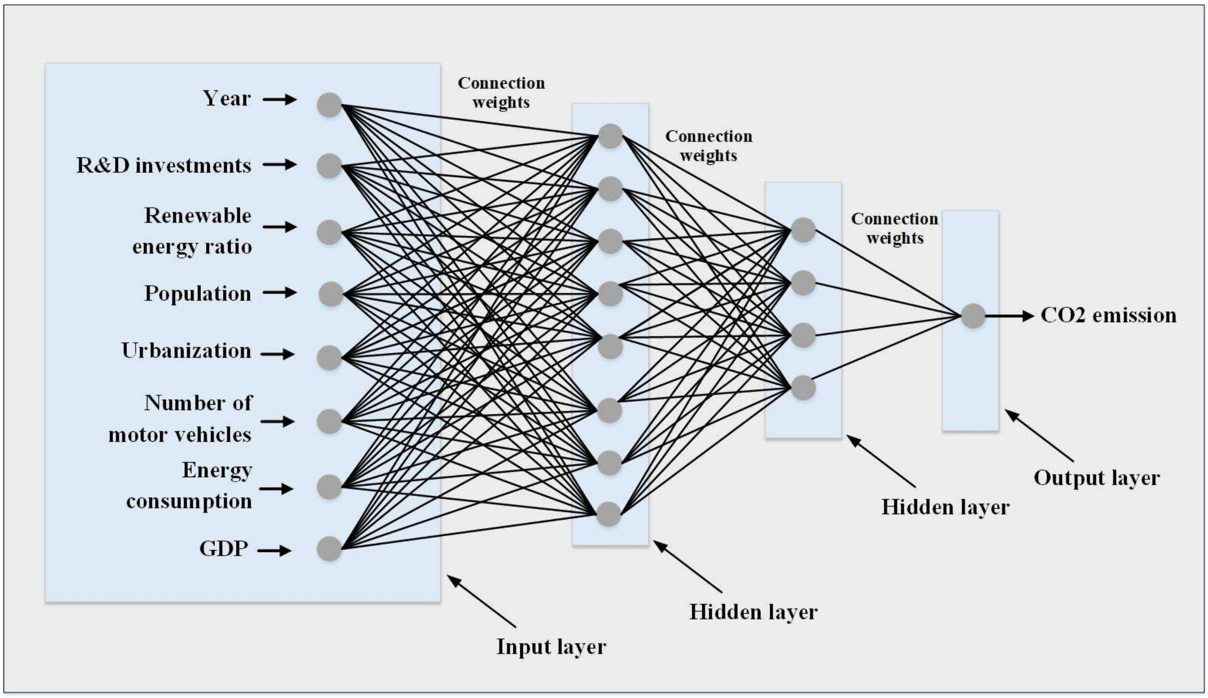


Figure 4.18. The diagram of the artificial neural network with an input, two hidden layers and an output layer used in the prediction model (HSBNN[4])

The parameters that trained for the feed forward MLNN in order to provide the best prediction result as follows:

- 64 parameters for weights between input and first hidden layers,
- 8 parameters for biases at first hidden layer,
- 32 parameters for weights between first hidden and second hidden layers,
- 4 parameters for biases at second hidden layer,
- 4 parameters for weights between second hidden and output layers and
- 1 parameter for biases at output layer.

The choice of fitness function is important in optimization methods. The determined fitness function is used to calculate the closeness of the solution to the global best. That is, the degree of compatibility of the solution evaluated in that step with the solution of the problem is calculated using the fitness function.

In the methods in which the artificial neural network model is used, the best setting of the parameters of the neural network affects the success of the model. Therefore, various methods are used to optimally adjust the neural network's coefficients. In this study, swarm-based methods are used in hybrid structure for this purpose. The fitness value of each solution in the search space is calculated according to the Mean Squared Error (MSE) criterion. The calculation formula for the MSE criterion is as follows:

$$MSE = \frac{\sum_{i=1}^{n}\left(x_{actual} - x_{predicted}\right)^2}{n} \tag{26}$$

here, x_{actual} and $x_{predicted}$ are represents the measured and predicted output values, respectively. n is the number of samples.

At the end of the study, the model results are compared according to various error criteria. The formulas for the used Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) and Theil's inequality coefficient (TIC) error criterias are as follows:

Mean absolute error (MAE)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|x_{actual} - x_{predicted}\right| \tag{27}$$

Root mean square error (RMSE)

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_{actual} - x_{predicted}\right)^2} \tag{28}$$

Mean absolute percentage error (MAPE)

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_{actual} - x_{predicted}}{x_{predicted}}\right| \tag{29}$$

Theil's inequality coefficient (TIC)

$$TIC = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_{actual} - x_{predicted}\right)^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_{actual})^2} + \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_{predicted}\right)^2}} \tag{30}$$

here, x_{actual} and $x_{predicted}$ are represents the measured and predicted output values, respectively. n is the number of samples.

The main purpose of this thesis is to estimate the CO^2 emission amount of Türkiye with minimum error. Therefore, a new artificial neural network-based estimation method has been developed. For the training of the coefficient and bias values of the artificial neural network, a new training method has been proposed by using nature inspired meta-heuristics optimization methods in a hybrid structure.

The multi-layer artificial neural network in the proposed prediction model is used in four different structures in the CO^2 emission estimation phase. In the first model, the ANN is designed with 12 neurons in 1 layer in the hidden layer. The structure of the ANN of this model is 8-12-1. In the second model, 2 layers are used in the hidden layer. The number of neurons in these layers is 6 and 4, respectively. So, the number of neurons in the neural network structure is 8-6-4-1. The HSBNN[3] model is built with a neural network with 6 neurons in both hidden layers. Hence, the number of neurons in the network structure is 8-6-6-1. Finally, the hidden layer of the artificial neural network in the HSBNN[4] model is created as two layers. There are eight neurons in the first hidden layer and 4 neurons in the second hidden layer. The structure of the neural network of this model is 8-8-4-1. The estimation results obtained with these four different structure methods are as shown in the Table 4.14.

Table 4.14. The relative errors obtained with the proposed prediction methods for CO^2 emission amounts of Türkiye

	2013	2014	2015	2016	2017	2018
Actual data	318.1700	340.64	351.59	374.59	415.9	412.97
HSBNN^1	321.1266	338.1796	357.8222	369.8425	408.152	408.4414
Relative Error %	0.9293	-0.7223	1.7726	-1.2674	-1.8629	-1.0966
HSBNN^2	320.1293	348.1403	354.4147	360.0696	409.4185	393.4433
Relative Error %	0.6158	2.2018	0.8034	-3.8763	-1.5584	-4.7284
HSBNN^3	324.0743	335.5070	354.6640	366.2286	409.0119	405.7487
Relative Error %	1.8557	-1.5069	0.8743	-2.2321	-1.6562	-1.7486
HSBNN^4	324.8164	312.0843	354.3832	370.1855	410.7208	416.4065
Relative Error %	2.0889	-8.3830	0.7945	-1.1758	-1.2453	0.8321

In order to determine the best prediction model, studies are carried out using different neuron numbers and layer numbers in the hidden layer. Each method is run 20 times with the data set and the results of the optimum solution obtained are given in Figure 4.19.

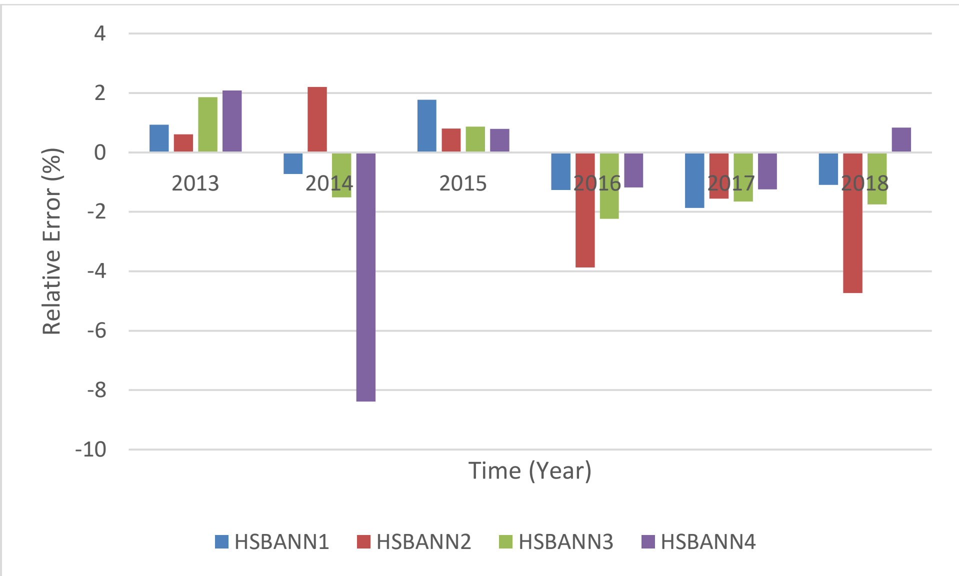


Figure 4.19. Comparison chart of relative errors of different prediction methods

The number of hidden layers of artificial neural networks used in different prediction methods at this stage of the study and information on the number of neurons in these layers are given in the Table 4.15. Moreover, the numerical values of the best results obtained at the end of these studies according to different error criteria are given in this table.

Table 4.15. Training and testing errors for four different prediction models

	No of neurons in 1st hidden layer	**No of neurons in 2nd hidden layer**	**Error for training data set**	**Error for testing data set**			
			NMSE	**MAPE (%)**	**RMSE**	**MAE**	**TIC**
HSBNN^1	12	-	0.000873	1.2752	5.1107	4.7789	0.0138
HSBNN^2	6	4	0.000774	2.2974	10.8183	8.8022	0.0295
HSBNN^3	6	6	0.000469	1.6456	6.3267	6.0970	0.0172
HSBNN^4	8	4	0.001321	2.4199	12.4193	8.5026	0.0338

In this thesis, a multi-layer artificial neural network is used in the proposed estimation method. In the proposed estimation method, the NMSE value is used as the fitness function during the parameter training of the neural network and the graph of the curves of the obtained NMSE values is shown in Figure 4.20. It can be seen that, the prediction model that found the best result at the end of the training phase is HSBNN^3.

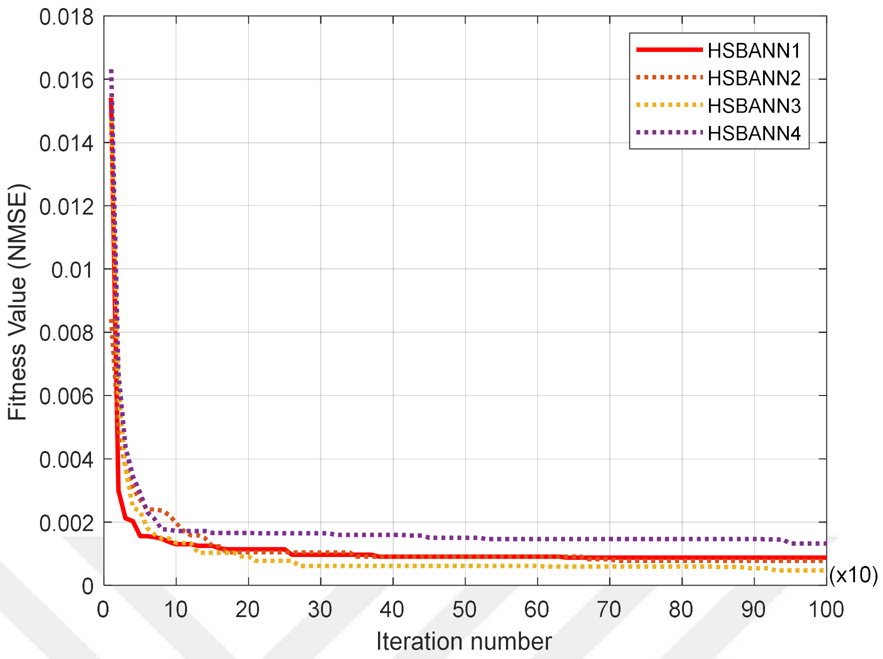


Figure 4.20. The fitness function curves of the proposed prediction algorithms

Figure 4.21 shows the comparison of the actual data and proposed four prediction models results of training and testing data set. The accuracies of the chosen models are difficult to distinguish, but the numbers reveal that the training part is more accurate than the test part.

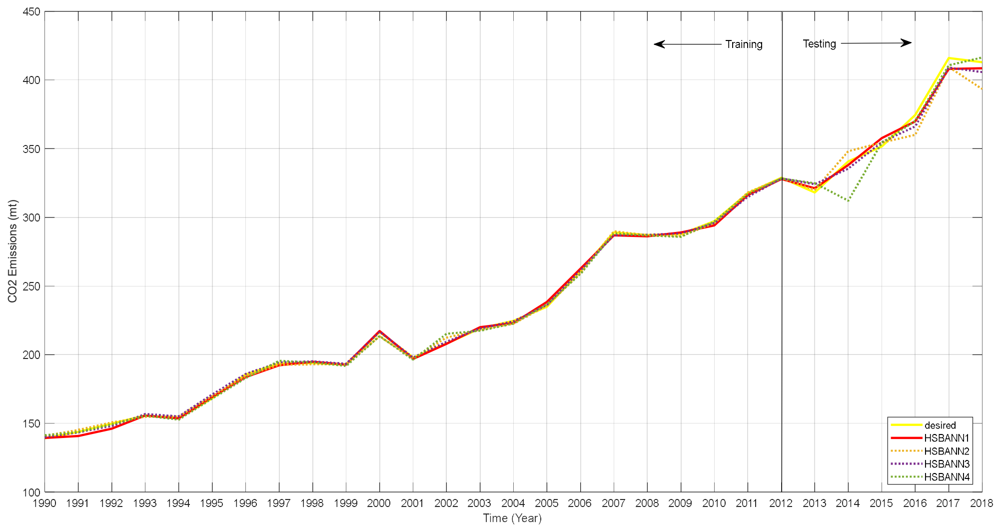


Figure 4.21. Comparison of the actual data and proposed prediction models results (for training and testing data set)

Comparisons of the performances of prediction methods for the testing set are presented in Figure 4.22.

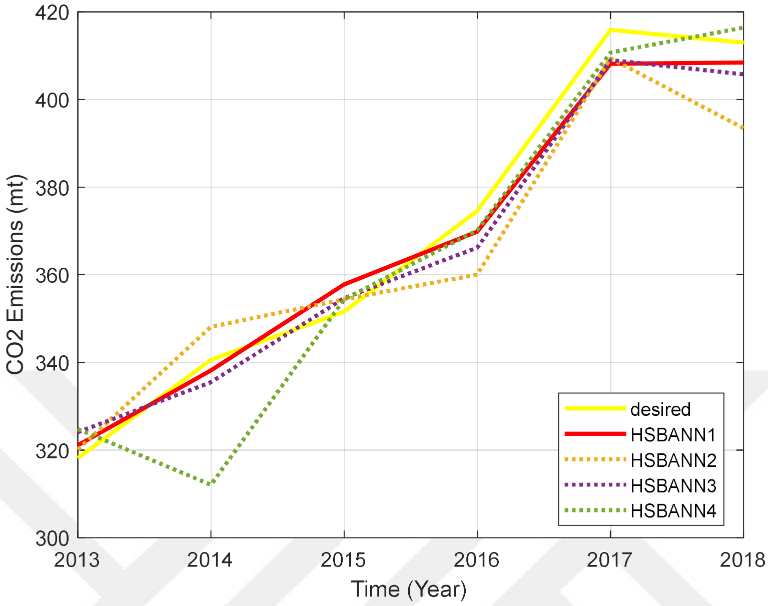


Figure 4.22. Comparison of the actual data and proposed prediction models results (for testing data set)

Linear regression of the actual CO^2 emission measurement value that is taken from the institution relative to results from prediction models is shown in the Figures 4.23-4.26.

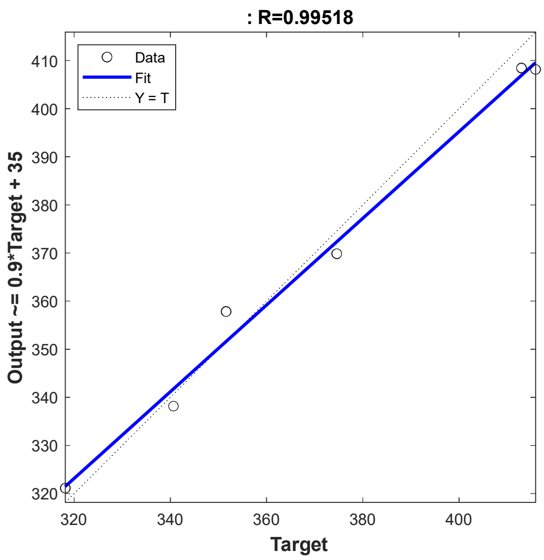


Figure 4.23. Linear regression graphs of the HSBNN[1] method

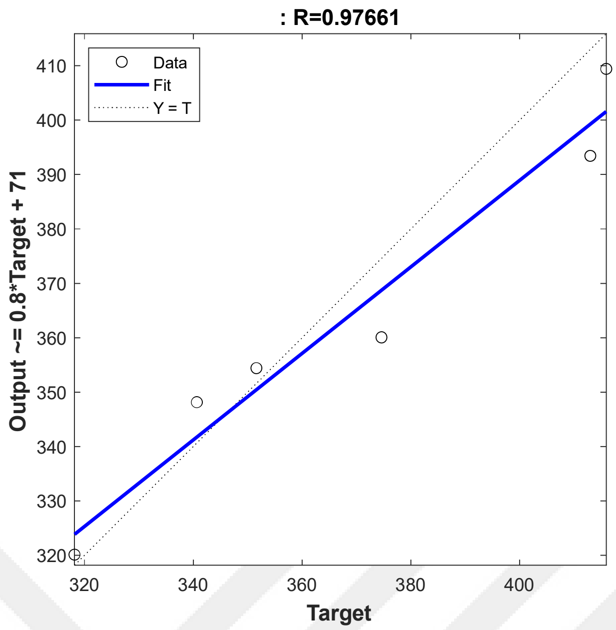


Figure 4.24. Linear regression graphs of the HSBNN^2 method

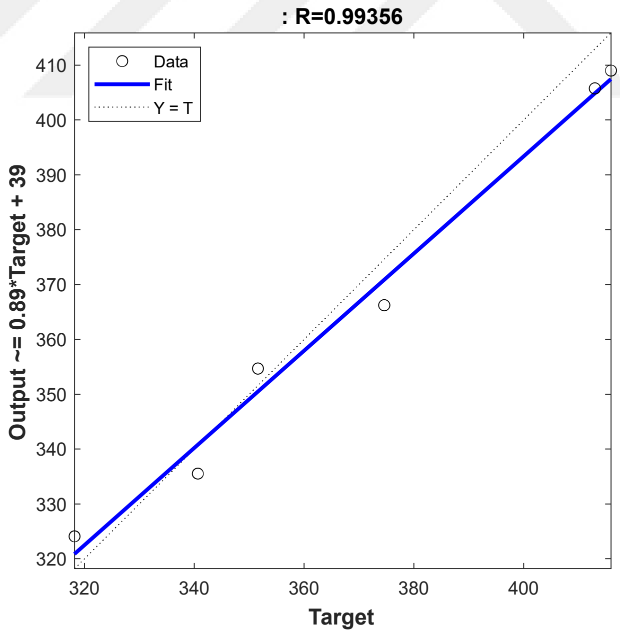


Figure 4.25. Linear regression graphs of the HSBNN^3 method

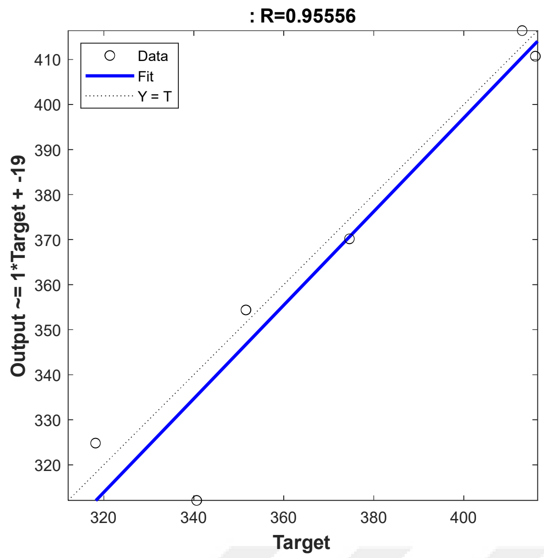


Figure 4.26. Linear regression graphs of the HSBNN^4 method

The evaluation of the prediction results according to the criteria of lineer regression, R^2 and mean relative error are given in the Table 4.16.

Table 4.16. Comparison table of Lineer Regression, R2 and Relative Error of prediction results

	Regression	R^2	Relative error (%)
HSBNN^1	0.99518	0.9904	1.2752
HSBNN^2	0.97661	0.9538	2.2974
HSBNN^3	0.99356	0.9872	1.6456
HSBNN^4	0.95556	0.9131	2.4199

This thesis introduces the effectiveness and applicability of the developed prediction models, HSBNN^1, HSBNN^2, HSBNN^3, and HSBNN^4, to estimate carbon dioxide emissions, with the results provided in Table 4.15 and Figure 4.21. Table 4.15 and Figure 4.21 show that the HSBNN^1 model's results are compatible with the trend of actual carbon dioxide emissions from 2013 to 2018, allowing for a lower error value prediction. The results generated with HSBNN^4 have a significant error value, and the HSBNN^2 and HSBNN^3 methods produce results that are relatively better than this method. In Table 4. 14 and Figure 4.19, the estimation values by years and the relative error of this estimation are given. According to the year-based forecast results, it can be said that, $HSBNN^1$ is outperformed the $HSBNN^2$, $HSBNN^3$, and $HSBNN^4$ prediction methods, respectively.

As it can be understood from the numerical results given in the Figure 4.19 and Table 4.15, the prediction model that found the best result at the end of the training phase is $HSBNN^3$. $HSBNN^1$ and $HSBNN^2$ methods have completed the training phase by finding error values close to each other. At the end of the training phase, the $HSBNN^4$ method achieved the highest error with an error value of 0.0013 according to the NMSE criterion.

As can be seen in Table 4.14, relative error of $HSBN^1$ is relatively stable fluctuating between --2% and 1% and the maximum value is 1.8629%. The $HSBN^2$ and $HSBN^3$ methods gave better results in the first three test points. It is seen that the degree of deviation from the actual value is much higher than the other methods of the $HSBNN^4$ method. Additionally, forecast results according to the criteria of linear regression, R^2 and mean relative error are given in the Table 4.16. Linear regression is the method used to determine the relationship between two variables. According to the results given in Figures 4.22-4.25 and Table 4-16, it is seen that the results obtained with the $HSBNN^1$ method are the most related to the value of the measured CO^2 emission amounts.

4.3. Future CO^2 Emissions in Türkiye

At this stage of the thesis, proposed prediction method is used for 3 different scenarios to future forecast of CO^2 emissions in Türkiye with usage renewable and non-renewable sources. The data to be used in this step of the study is set as follows:

R&D investments

The average increase of the last 2 years (according to TurkStat data (TurkStat, 2020)) is preferred for Scenario 1.

The average increase of the last 3 years (according to TurkStat data (TurkStat, 2020)) is preferred for Scenario 2.

The average increase of the last 5 years (according to TurkStat data (TurkStat, 2020)) is preferred for Scenario 3.

Renewable energy ratio

Average rate of increase in the last 3 years (according to Worldbank data (World Bank)) is preferred for Scenario 1.

Average rate of increase in the last 5 years (according to Worldbank data (World Bank)) is preferred for Scenario 2.

Average rate of increase in the last 8 years (according to Worldbank data (World Bank)) is preferred for Scenario 3.

Population

Estimated population data from TURKSTAT is used (TURKSTAT, 2018).

Urbanization

Average values are determined according to the rate of increase in recent years (according to Indexmundi data (Indexmundi)).

Additionally, the rate of increase in the number of motor vehicles and energy consumption data is taken from the Özceylan's article (Özceylan, 2016). In addition, the increase rate of GDP data is taken from the Uzlu's article (Uzlu, 2021). The details of the scenarios are presented in Table 4.17.

Table 4.17. Scenarios for Türkiye's CO^2 emission prediction (approximate growth rates)

	Scenario 1	**Scenario 2**	**Scenario 3**
R&D investments (TurkStat, 2020)	19.42 %	22.64 %	21.72 %
Renewable energy ratio (World Bank)	12.16 %	12.27 %	12.60 %
Population (TURKSTAT, 2018)	TurkStat population forecasts	TurkStat population forecasts	TurkStat population forecasts
Urbanization	1.0 %	1.05 %	1.08 %
Number of motor vehicles (Özceylan, 2016)	3.0 %	3.5 %	4.0 %
Energy consumption (Özceylan, 2016)	2.0 %	2.5 %	3.0 %
GDP (Uzlu, 2021)	4.87 %	6.37 %	7.87 %

The results of future projections for different scenarios are shown in Table 4.18. According to the table, the amount of CO^2 emissions in Türkiye varies between average 500 Mt and 515 Mt. While the lowest emission amount is obtained with Scenario 1, the highest emission amount is obtained with Scenario 3. When the emission values in 2030 are examined, the lowest and highest CO^2 emission amounts are 504.6270 Mt and 510.0260 Mt, respectively.

Table 4.18. Future projections of CO^2 emission amount according to scenarios with HSBNN[1] method

Year	Scenario 1	Scenario 2	Scenario 3
2018 (World Bank)	412.9700	412.9700	412.9700
2019	419.7500	418.8049	419.2450
2020	427.1719	427.1067	426.2700
2021	434.7241	436.2368	433.4406
2022	442.1495	445.6232	440.4107
2023	449.3373	454.8216	447.0232
2024	455.8839	463.3066	452.8473
2025	460.7198	470.1485	457.0039
2026	464.1629	476.0640	467.5716
2027	469.7298	484.0587	473.9727
2028	478.7903	494.4381	483.6537
2029	490.7800	506.6813	496.0707
2030	504.6270	519.2656	510.0260

The graphical results of the future estimation of the amount of CO^2 emission given numerically in Table 4.18 is presented in the Figure 4.27.

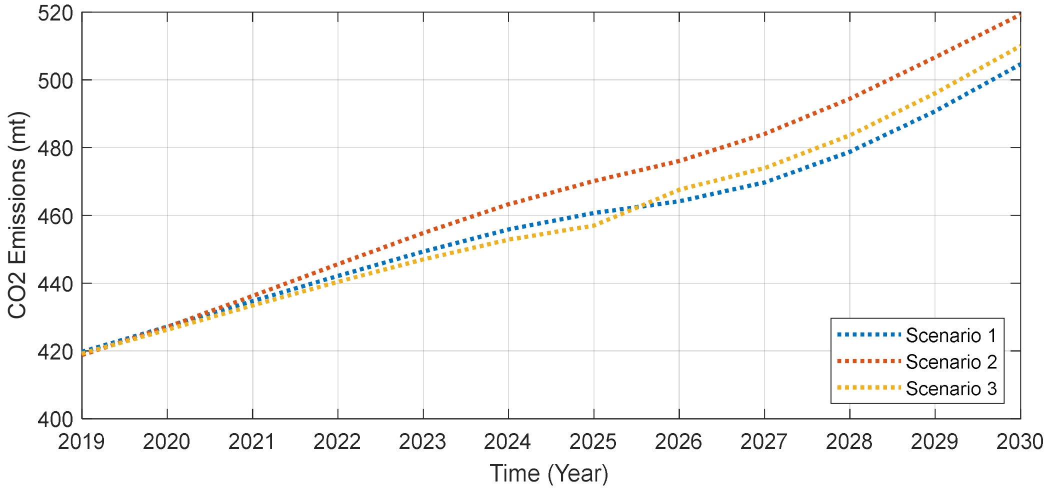


Figure 4.27. Graphical results of future CO^2 emission estimates according to different three scenarios with HSBNN[1] method

Türkiye's average CO^2 emission amount in 2030, according to the future CO^2 emission estimation generated with the approach given in this study, is 510 Mt. According to the Paris Agreement, it is aimed to reduce greenhouse gas emissions by 50% by 2030. In this direction, countries need to accurately estimate the amount of CO^2 emissions, which has the most effective ratio in the amount of greenhouse gas emissions, and take precautions accordingly.

5. CONCLUSIONS

Optimization is a method that aims to find the best solution to a specific problem in fastest way. However, one of the most critical disadvantages of these methods is that they are stuck at the local minimum and cannot reach the global best. In the literature, different methods have been suggested in order to increase the performance of optimization methods, such as using functions to improve search capabilities, using more than one method in a hybrid structure.

Researchers propose different nature-inspired optimization methods by modeling the movements and intelligences of different types of animals, such as insects and birds in nature to reach their goal and communicate with each other. These methods are used successfully in solving real-time physical problems. In this thesis, a new hybrid method is proposed as a first time in the literature. Then, this hybrid method is used together with ANN to develop a new estimation method for CO^2 emission estimation.

Step 1:

In this study, The Shuffled Frog-Leaping Algorithm (SFLA), that is recently proposed nature inspired metaheuristic method is used. To increase the performance and speed of finding the optimum solution in SFLA's search space:

- Levy flight function has been used in the mementic search phase (LSFLA),
- Another nature inspired metaheuristic optimization method, FA, is used in a hybrid structure with SFLA.

Benchmark functions, which are optimization approaches of real-world problems, are used to prove the success of the proposed **improved Levy Flight SFLA based on FA (ILSFLAFA)** method, that is developed by making these improvements. The best, worst and standard deviation values of the results obtained with the 12 most frequently used Benchmark functions are given in the table. The results of this method are compared with the results obtained by the SFLA and FA methods. In addition, the fitness function curves of three different methods are given graphically. As the results are examined, it has been revealed that the performance of the proposed method is better than the other methods in most of the tested functions/problems.

Step 2:

In the 20th century, technological developments have started to increased exponentially, and it has become easier for people to access goods and services. Developments such as the establishment of new factories, population growth and the development of the construction sector and the increase in the number of households have led to an increase in the amount of CO^2 emissions worldwide. This increase worries scientists and government officials. Many countries have started various studies to reduce the amount of greenhouse gas emissions of their countries until 2030, within the scope of The European Green Deal.

In line with the reduction of greenhouse gas emissions by countries, the estimation of CO^2 emissions with the least error gains importance and has addressed in the many researches in recent years. The scope of this thesis is developing a new estimation method by using a multilayer feedforward artificial neural network. The coefficients of the Multi-Layer Neural Network (MLNN) that is used in this thesis are trained by the hybrid optimization method developed in first step/phase of the thesis. As a first time, the **developed hybrid swarm-based MLNN (HSBNN)** is applied to the data of a study conducted in 2016 to estimate the amount of carbon emissions in Türkiye. The results obtained are compared with the results given at the end of the article, and the effect and success of the proposed estimation method are demonstrated. While the lowest relative error value obtained with the methods suggested in the article is 3.379 in the CO^2 emission estimation study conducted in Türkiye for the years 2005-2008 in 2016, the error value obtained with the estimation method proposed in this thesis study is 0.7682.

The aim of this thesis is to use data from renewable energy sources in the estimation of CO^2 emissions. For this reason, the HSBNN method is employed to estimate the amount of CO^2 emissions in Türkiye by considering the increased use of renewable energy sources. The estimation model is developed by giving the year, R&D investments, renewable energy ratio of total final energy consumption, population, urbanization, number of motor vehicles, energy consumption and GDP parameters to the input layer of the Artificial Neural Network. Data collecting between the years 1990-2018 from different institutions are used at the phase of development and testing of the estimation model, Türkiye's. The numerical and graphical demonstration of the results are given in detail. These results prove the success of the proposed method and its usability in studies in this field.

In this study, the MLNN-based prediction method has been tested with artificial neural networks with four different structures. In the first estimation method (HSBNN^1), estimation is made using one hidden layer with twelve neurons. In the other three methods, two layers are used in the hidden layer. The number of neurons in the hidden layers are 6-4, 6-6 and 8-4, in the second method (HSBNN^2), third method (HSBNN^3) and the fourth method (HSBNN^4), respectively. At the end of the estimation of the CO^2 emission values for the years 2013-2018 in Türkiye, the RMSE values obtained from the four proposed methods are 5,1107, 10,8183, 6.3267 and 12,4193, respectively. Additionally, the MAPE values are 1.2752 %, 2.2974 %, 1.6456 % and 2.4199 %, respectively. According to these results, it is recommended to use the **HSBNN^1 method** for forecasting of Türkiye's CO^2 emissions.

Finally, the **proposed estimation method (HSBANN^1)** is used for Türkiye's prospective CO^2 emission estimation. At this point, three different scenarios are considered as Scenario 1, Scenario 2 and Scenario 3. The numerical and graphical results for each scenario are given in detail. By examining the results, it can be concluded that the increase in the amount of use of renewable energy sources (scenario 1) slows down the increase in the amount of CO^2 emissions. As a conclusion, increasing the use of renewable energy sources and supporting the projects that is developed for this purpose will be beneficial for reducing the amount of CO^2 emissions in our country.

6. RECOMMENDATIONS

Suggestions to improve the scope of the study and contribute to the study are as follows:

- Tests can be made with different Benchmark functions to demonstrate the success of the proposed method for training the network. Furthermore, CEC functions can be used to see the method's ability to solve real-world situations.

- By working with Artificial Neural Networks with different structures and features, the estimation ability of the proposed method can be improved.

- Different improvements can be made to improve the search speed of the SFLA algorithm and reduce the possibility of getting stuck in the local minimum. In addition, a new optimization method can be proposed to train the multilayered artificial neural network by using the SFLA algorithm in a hybrid structure with different swarm-based optimization method.

- Since there are not enough studies on the amount of CO^2 emissions in Türkiye in the literature, the results obtained could not be compared with different methods. Therefore, the data set in this study can be compared by working with different methods.

- The diversity of the data used in the emission estimation phase in the study can be increased. Renewable energy sources have been used under a single heading. The results obtained by using them under separate headings (such as the amount of wind energy use, the amount of solar energy use) can be examined.

7. REFERENCES

Acheampong, A. O., & Boateng, E. B. (2019). Modelling carbon emission intensity: Application of artificial neural network. *Journal of Cleaner Production, 225,* 833-856. https://doi.org/10.1016/j.jclepro.2019.03.352

Ahmadi, M. H., Jashnani, H., Chau, K.-W., Kumar, R., & Rosen, M. A. (2019). Carbon dioxide emissions prediction of five Middle Eastern countries using artificial neural networks. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects,* 1-13. https://doi.org/10.1080/15567036.2019.1679914

Alam, T., & AlArjani, A. (2021). Forecasting CO 2 Emissions in Saudi Arabia Using Artificial Neural Network, Holt-Winters Exponential Smoothing, and Autoregressive Integrated Moving Average Models. *2021 International Conference on Technology and Policy in Energy and Electric Power (ICT-PEP),* 125-129. https://doi.org/10.1109/ICT-PEP53949.2021.9601031

Antanasijević, D. Z., Ristić, M. Đ., Perić-Grujić, A. A., & Pocajt, V. V. (2014). Forecasting GHG emissions using an optimized artificial neural network model based on correlation and principal component analysis. *International Journal of Greenhouse Gas Control, 20,* 244-253. https://doi.org/10.1016/j.ijggc.2013.11.011

Antanasijević, D., Pocajt, V., Ristić, M., & Perić-Grujić, A. (2015). Modeling of energy consumption and related GHG (greenhouse gas) intensity and emissions in Europe using general regression neural networks. *Energy, 84,* 816-824.

Azadeh, A., Sheikhalishahi, M., & Hasumi, M. (2015). A hybrid intelligent algorithm for optimum forecasting of CO2 emission in complex environments: The cases of Brazil, Canada, France, Japan, India, UK and US. *World Journal of Engineering, 12(3)*, 237-246. https://doi.org/10.1260/1708-5284.12.3.237.

Bakay, M. S., & Ağbulut, Ü. (2021). Electricity production based forecasting of greenhouse gas emissions in Turkey with deep learning, support vector machine and artificial neural network algorithms. *Journal of Cleaner Production, 285,* 125324. https://doi.org/10.1016/j.jclepro.2020.125324

Baskan, O. (2013). Determining optimal link capacity expansions in road networks using cuckoo search algorithm with lévy flights. *Journal of Applied Mathematics,* 1-11. https://doi.org/10.1155/2013/718015

Chang, H., Sun, W., & Gu, X. (2013). Forecasting Energy CO2 Emissions Using a Quantum Harmony Search Algorithm-Based DMSFE Combination Model. *Energies, 6(3),* 1456-1477. https://doi.org/10.3390/en6031456.

Dritsaki, M., & Dritsaki, C. (2020). Forecastıng European Unıon CO2 Emissions Using Autoregressive Integrated Moving Average-Autoregressive Conditional Heteroscedasticity Models. *International Journal of Energy Economics and Policy, 10(4),* 411-423. https://doi.org/10.32479/ijeep.9186.

Eusuff, M. M., & Lansey, K. E. (2003). Optimization of water distribution network design using the shuffled frog leaping algorithm. *Journal of Water Resources planning and management, 129(3),* 210-225. https://doi.org/10.1061/(ASCE)0733-9496(2003)129:3(210).

Eusuff, M., Lansey, K., & Pasha, F. (2006). Shuffled frog-leaping algorithm: a memetic meta-heuristic for discrete optimization. *Engineering optimization, 38(2),* 129-154. https://doi.org/10.1080/03052150500384759.

Guo, J., Liu, W., Tu, L., & Chen, Y. (2021). Forecasting carbon dioxide emissions in BRICS countries by exponential cumulative grey model. *Energy Reports, 7,* 7238-7250. https://doi.org/10.1016/j.egyr.2021.10.075.

Hariya, Y., Kurihara, T., Shindo, T., & Jin'no, K. (2015). Lévy flight PSO. *In 2015 IEEE congress on evolutionary computation (CEC),* 2678-2684. *IEEE.* https://doi.org/10.1109/CEC.2015.7257220.

Hassanzadeh, T., Vojodi, H., & Moghadam, A. M. E. (2011). A multilevel thresholding approach based on levy-flight firefly algorithm. *In 2011 7th Iranian Conference on Machine Vision and Image Processing, 1-5. IEEE*. https://doi.org/10.1109/IranianMVIP.2011.6121552.

Heydari, A., Garcia, D. A., Keynia, F., Bisegna, F., & Santoli, L. D. (2019). Renewable Energies Generation and Carbon Dioxide Emission Forecasting in Microgrids and National Grids using GRNN-GWO Methodology. *Energy Procedia, 159,* 154-159. https://doi.org/10.1016/j.egypro.2018.12.044.

Ho, H. X. T. (2018). Forecasting of CO2 Emissions, Renewable Energy Consumption and Economic Growth in Vietnam Using Grey Models. *2018 4th International Conference on Green Technology and Sustainable Development (GTSD),* 452-455. https://doi.org/10.1109/GTSD.2018.8595679.

Hong, T., Jeong, K., & Koo, C. (2018). An optimized gene expression programming model for forecasting the national CO2 emissions in 2030 using the metaheuristic algorithms. *Applied Energy, 228,* 808-820. https://doi.org/10.1016/j.apenergy.2018.06.106.

Hosseini, S. M., Saifoddin, A., Shirmohammadi, R., & Aslani, A. (2019). Forecasting of CO 2 emissions in Iran based on time series and regression analysis. *Energy Reports, 5,* 619-631. https://doi.org/10.1016/j.egyr.2019.05.004.

Indexmundi, http://www.indexmundi.com/facts/indicators/EN.ATM.CO2E.KT, (Accessed 23 October 2014).

Indexmundi, https://www.indexmundi.com/facts/Turkey/urban-population (Accessed 02 March 2022).

Indexmundi, https://www.indexmundi.com/Turkey/population.html (Accessed 02 March 2022).

Javed, S. A., & Cudjoe, D. (2022). A novel grey forecasting of greenhouse gas emissions from four industries of China and India. *Sustainable Production and Consumption, 29,* 777-790. https://doi.org/10.1016/j.spc.2021.11.017.

Jena, P. R., Managi, S., & Majhi, B. (2021). Forecasting the CO2 Emissions at the Global Level: A Multilayer Artificial Neural Network Modelling. *Energies, 14(19)*, 6336. https://doi.org/10.3390/en14196336.

Jensi, R., & Jiji, G. W. (2016). An enhanced particle swarm optimization with levy flight for global optimization. *Applied Soft Computing, 43,* 248-261. https://doi.org/10.1016/j.asoc.2016.02.018.

Kaygusuz, K., & Toklu, E. (2016). The increase of exploitability of renewable energy sources in Turkey. *Journal of Engineering Research and Applied Science, 5(1),* 352-358.

Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. In Proceedings of ICNN'95-international conference on neural networks, *4, 1942-1948. IEEE*. https://doi.org/10.1109/ICNN.1995.488968.

Khoshnevisan, B., Rafiee, S., Omid, M., Yousefi, M., & Movahedi, M. (2013). Modeling of energy consumption and GHG (greenhouse gas) emissions in wheat production in Esfahan province of Iran using artificial neural networks. *Energy, 52,* 333-338. https://doi.org/10.1016/j.energy.2013.01.028.

Koopialipoor, M., & Noorbakhsh, A. (2020). Applications of artificial intelligence techniques in optimizing drilling. *Emerging Trends in Mechatronics,* 89-118.

Lan, K., Liu, L., Li, T., Chen, Y., Fong, S., Marques, J. A. L., ... & Tang, R. (2020). Multi-view convolutional neural network with leader and long-tail particle swarm optimizer for enhancing heart disease and breast cancer detection. *Neural Computing and Applications, 32(19),* 15469-15488.

Lee, K., Booth, D., & Alam, P. (2005). A comparison of supervised and unsupervised neural networks in predicting bankruptcy of Korean firms. *Expert Systems with Applications, 29(1),* 1-16. https://doi.org/10.1016/j.eswa.2005.01.004

Li, K., Xiong, P., Wu, Y., & Dong, Y. (2022). Forecasting greenhouse gas emissions with the new information priority generalized accumulative grey model. *Science of The Total Environment, 807(2),* 150859. https://doi.org/10.1016/j.scitotenv.2021.150859.

Lin, C.-C., He, R.-X., & Liu, W.-Y. (2018). Considering Multiple Factors to Forecast CO2 Emissions: A Hybrid Multivariable Grey Forecasting and Genetic Programming Approach. *Energies, 11(12),* 3432. https://doi.org/10.3390/en11123432.

Liu, S., & Wang, Y. (2021). A lévy flight based firefly algorithm for multilevel thresholding image segmentation. *In Journal of Physics: Conference Series, 1865(4),* 1-6. *IOP Publishing.*

Liu, Z., Jiang, P., Wang, J., & Zhang, L. (2022). Ensemble system for short term carbon dioxide emissions forecasting based on multi-objective tangent search algorithm. *Journal of Environmental Management, 302,* 113951. https://doi.org/10.1016/j.jenvman.2021.113951.

Ma, X., Jiang, P., & Jiang, Q. (2020). Research and application of association rule algorithm and an optimized grey model in carbon emissions forecasting. *Technological Forecasting and Social Change, 158,* 120159. https://doi.org/10.1016/j.techfore.2020.120159.

Marugán, A. P., Márquez, F. P. G., & Papaelias, M. (2017). Multivariable analysis for advanced analytics of wind turbine management. *In Proceedings of the tenth international conference on management science and engineering management, Springer, Singapore,* 319-328.

Marugán, A. P., Márquez, F. P. G., Perez, J. M. P., & Ruiz-Hernández, D. (2018). A survey of artificial neural network in wind energy systems. *Applied Energy, 228,* 1822-1836. https://doi.org/10.1016/j.apenergy.2018.07.084.

Mills, S., & House, P. (2014). Prospects for coal and clean coal technologies in Turkey. *IEA Clean Coal Centre.*

Özceylan, E. (2016). Forecasting CO2 emission of Turkey: swarm intelligence approaches. *International Journal of Global Warming, 9(3),* 337-361.

Patro, S., & Sahu, K. K. (2015). Normalization: A preprocessing stage. *arXiv preprint arXiv:1503.06462*, 1-4. https://doi.org/10.48550/arXiv.1503.06462.

Qiao, W., Lu, H., Zhou, G., Azimi, M., Yang, Q., & Tian, W. (2020). A hybrid algorithm for carbon dioxide emissions forecasting based on improved lion swarm optimizer. *Journal of Cleaner Production, 244,* 118612. https://doi.org/10.1016/j.jclepro.2019.118612.

Radojević, D., Pocajt, V., Popović, I., Perić-Grujić, A., & Ristić, M. (2013). Forecasting of Greenhouse Gas Emissions in Serbia Using Artificial Neural Networks. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 35(8),* 733-740.

Rajpoot, V., Chitrakoot, M. G. C. G. V., Tiwari, A., & Mishra, B. (2018). AMSFLO: Optimization Based Efficient Approach For Assosiation Rule Mining. *International Journal of Computer Science and Information Security (IJCSIS), 16(3),* 147-154.

Ren, F., & Long, D. (2021). Carbon emission forecasting and scenario analysis in Guangdong Province based on optimized Fast Learning Network. *Journal of Cleaner Production, 317, 128408.* https://doi.org/10.1016/j.jclepro.2021.128408.

Rostami, F., & Jafarian, A. (2018). A new artificial neural network structure for solving high-order linear fractional differential equations. *International Journal of Computer Mathematics, 95(3),* 528-539. https://doi.org/10.1080/00207160.2017.1291932.

Şahin, U. (2019). Forecasting of Turkey's greenhouse gas emissions using linear and nonlinear rolling metabolic grey model based on optimization. *Journal of Cleaner Production, 239,* 118079. https://doi.org/10.1016/j.jclepro.2019.118079.

Sangeetha, A., & Amudha, T. (2018). A Novel Bio-Inspired Framework for CO 2 Emission Forecast in India. *Procedia Computer Science, 125,* 367-375. https://doi.org/10.1016/j.procs.2017.12.048.

Senthilnath, J., Das, V., Omkar, S. N., & Mani, V. (2013). Clustering using levy flight cuckoo search. *In Proceedings of Seventh International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA 2012), Springer, India,* 65-75.

Senthilnath, J., Omkar, S. N., & Mani, V. (2011). Clustering using firefly algorithm: performance study. Swarm and Evolutionary Computation, *1(3),* 164-171. https://doi.org/10.1016/j.swevo.2011.06.003.

Sharma, V., Rai, S., & Dev, A. (2012). A comprehensive study of artificial neural networks. *International Journal of Advanced research in computer science and software engineering, 2(10),* 278-284.

Sumanto, M., Martoprawiro, M. A., & Ivansyah, A. L. (2021, November). The prediction of molecule atomization energy using neural network and extreme gradient boosting. *In Journal of Physics: Conference Series, 2072(1),* 1-9. *IOP Publishing.*

Tang, D., Yang, J., Dong, S., & Liu, Z. (2016). A lévy flight-based shuffled frog-leaping algorithm and its applications for continuous optimization problems. *Applied Soft Computing, 49,* 641-662. https://doi.org/10.1016/j.asoc.2016.09.002.

Tong, M., Duan, H., & He, L. (2021). A novel Grey Verhulst model and its application in forecasting CO2 emissions. *Environmental Science and Pollution Research, 28(24),* 31370-31379. https://doi.org/10.1007/s11356-020-12137-5

Turkish Statistical Institute (TURKSTAT). https://data.tuik.gov.tr/Bulten/Index?p=Nufus-Projeksiyonlari-2018-2080-30567 (Accessed 02 March 2022).

Turkish Statistical Institute (TURKSTAT). Greenhouse Gas Emissions Statistics 1990-2019, https://data.tuik.gov.tr/Bulten/Index?p=Greenhouse-Gas-Emissions-Statistics-1990-2019-37196, Accessed: 02 March 2022.

Turkish Statistical Institute (TURKSTAT). https://data.tuik.gov.tr/Bulten/Index?p=Arastirma-Gelistirme-Faaliyetleri-Arastirmasi-2020-37439 (Accessed 02 March 2022).
Turkish Statistical Institute (TURKSTAT). https://data.tuik.gov.tr/Bulten/Index?p=Motorlu-Kara-Tasitlari-Aralik-2020-37410#:~:text=T%C3%9C%C4%B0K%20Kurumsal&text=T%C3%BCrkiye'de%202020%20y%C4%B1l%C4%B1nda%20bir,bin%20577%20adet%20art%C4%B1%C5%9F%20ger%C3%A7ekle%C5%9Fti (Accessed 02 March 2022).

UNFCCC (2022), Paris Agreement, https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement, Union of Concerned Scientists, Each Country's Share of CO2 Emissions, https://www.ucsusa.org/resources/each-countrys-share-co2-emissions, Accessed: 02 March 2022.

US Energy Information Administration (EIA), https://www.eia.gov/international/data/country/ TUR/total-energy/total-energy-consumption?pd=44&p=000000001000000000000000000 000000000000000000000000000u06&u=0&f=A&v=mapbubble&a=-&i=none&vo=value &&t=C&g=none&l=249--230&s=315532800000&e=1546300800000 (Accessed 02 March 2022).

Uzlu, E. (2021). Estimates of greenhouse gas emission in Turkey with grey wolf optimizer algorithm-optimized artificial neural networks. *Neural Computing and Applications, 33(20),* 13567-13585. https://doi.org/10.1007/s00521-021-05980-1.

Wang, H., Wang, W., Zhou, X., Sun, H., Zhao, J., Yu, X., & Cui, Z. (2017). Firefly algorithm with neighborhood attraction. *Information Sciences, 382,* 374-387. https://doi.org/10.1016/j.ins.2016.12.024.

Wei, S., Yuwei, W., & Chongchong, Z. (2018). Forecasting CO2 emissions in Hebei, China, through moth-flame optimization based on the random forest and extreme learning machine.

Environmental Science and Pollution Research, 25(29), 28985-28997. https://doi.org/10.1007/s11356-018-2738-z.

World Bank, https://data.worldbank.org/indicator/EN.ATM.CO2E.KT?locations=TR (Accessed 02 March 2022).

World Bank, https://data.worldbank.org/indicator/EG.FEC.RNEW.ZS?locations=TR (Accessed 02 March 2022).

World Bank, https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?locations=TR (Accessed 02 March 2022).

X.S. Yang, Nature-Inspired Metaheuristic Algorithms, Luniver Press, 2008.

Yang, H., & O'Connell, J. F. (2020). Short-term carbon emissions forecast for aviation industry in Shanghai. *Journal of Cleaner Production, 275,* 122734. https://doi.org/10.1016/j.jclepro.2020.122734.

Yang, X. S., & Deb, S. (2009, December). Cuckoo search via Lévy flights. In 2009 World congress on nature & biologically inspired computing (NaBIC), 210-214. *Ieee.* https://doi.org/10.1109/NABIC.2009.5393690.

Ye, L., Yang, D., Dang, Y., & Wang, J. (2022). An enhanced multivariable dynamic time-delay discrete grey forecasting model for predicting China's carbon emissions. *Energy, 249,* 123681. https://doi.org/10.1016/j.energy.2022.123681.

Yetkin, M. (2016). Metasezgisel Algoritmaların Jeodezi'de Kullanımı. Geomatik*, 1(1),* 8-13. https://doi.org/10.29128/geomatik.294070.

Zhou, J., Xu, X., Li, W., Guang, F., Yu, X., & Jin, B. (2019). Forecasting CO 2 Emissions in China's Construction Industry Based on the Weighted Adaboost-ENN Model and Scenario Analysis. *Journal of Energy, 2019,* 1-12. https://doi.org/10.1155/2019/8275491.