وزارة التعليم العالي والبحث العلمي

BADJI MOKHTAR UNIVERSITY – ANNABA

UNIVERSITE BADJI MOKHTAR – ANNABA

جامعة باجي مختار – عنابة

Faculté : Sciences de l'Ingéniorat
Département : Informatique

THÈSE

Présenté en vue de l'obtention du diplôme de : **Doctorat 3éme cycle**

Intitulé

Système Multi-Agents prédictifs pour la simulation de la pollution due à la production et consommation énergétique

Domaine : Mathématiques et Informatique
Filière : Informatique
Spécialité : STIC (Sciences et Technologies de l'Information et de la Communication)

Par :
Mr. Seif Eddine Bouziane

Devant le jury composé de :

Pr. Mohamed Tarek Khadir	Université d'Annaba	Directeur
Dr. Julie Dugdale	Université de Grenoble-Alpes	Co-Directrice
Pr. Nabiha Azizi	Université d'Annaba	Présidente
Pr. Nadir Farah	Université d'Annaba	Examinateur
Dr. Brahim Farou	Université de Guelma	Examinateur

2021/2022

Dedication

This thesis is wholeheartedly dedicated to my late grandparents who left an empty place in my heart. A special feeling of gratitude to my beloved parents. My aunties, uncles, siblings, and cousins, who have been constant source of support and encouragement. I also dedicate this thesis to my many friends, special mentions to team QLF, team RTOM BO, without forgetting the whole computer science 2017 class and in particular the STIC class, and I dedicate it to anybody who contributed from close or far to this work.

Last but not least, I want to thank me . . .
I want to thank me for believing in me,
I want to thank me for doing all this hard work,
I want to thank me for having no days off,
I want to thank me for never quitting,
And I want to thank me for simply being me at all times.

Acknowledgements

First and foremost, I would like to praise and thank the Almighty God, for providing me this opportunity and granting me the capability to proceed successfully.

I want to thank my advisor Pr. Mohamed Tarek Khadir, for his guidance, patience and encouragement, without him this work wouldn't have been possible.

I must also thank and acknowledge the help and assistance of my co-advisor Dr. Julie Dugdale from the University of Grenoble-Alpes for taking the time to assist me with her precious help.
Many thanks to all my immediate family and friends for their constant encouragement and support.

At the end we thank the National Company for Electricity and Gas (SONELGAZ) for providing the necessary data.

Abstract

ENergy production and consumption are two of the largest sources of greenhouse gases (GHG), along with industry, and are two of the highest causes of global warming. Forecasting the environmental cost of energy production is necessary for better decision making and easing the switch to cleaner energy systems in order to reduce air pollution. This thesis describes a hybrid approach based on Artificial Neural Networks (ANN) and an agent-based architecture for forecasting carbon dioxide (CO_2) issued from different energy sources in the city of Annaba using real data. The system consists of multiple autonomous agents, divided into two types: firstly, forecasting agents, which forecast the production of a particular type of energy using the ANN models; secondly, core agents that perform other essential functionalities such as calculating the equivalent CO_2 emissions and controlling the simulation. The development is based on Algerian gas and electricity data provided by the national energy company. The simulation consists firstly of forecasting energy production using the forecasting agents and calculating the equivalent emitted CO_2. Secondly, a dedicated agent calculates the total CO_2 emitted from all the available sources. It then computes the benefits of using renewable energy sources as an alternative way to meet the electric load in terms of emission mitigation and economizing natural gas consumption. The designed system can be helpful for the environmental agencies and energy planners to examine their energy and air pollution controlling policies.

Keywords: Multi-agent, Neural networks, Carbon dioxide, short-term forecasting, energy.

Résumé

La production et la consommation d'énergie sont deux des plus importantes sources de gaz à effet de serre (GES), avec l'industrie, et deux des principales causes du réchauffement climatique. La prévision du coût environnemental de la production d'énergie est nécessaire pour une meilleure prise de décision et pour faciliter le passage à des systèmes énergétiques plus propres afin de réduire la pollution atmosphérique. Cette thèse décrit une approche hybride basée sur les réseaux de neurones artificiels (RNA) et une architecture multi-agents pour la prévision du dioxyde de carbone (CO_2) issu de différentes sources d'énergie dans la ville d'Annaba à partir de données réelles. Le système se compose de plusieurs agents autonomes, divisés en deux types : premièrement, les agents de prévision, qui prédisent la production d'un type particulier d'énergie à l'aide des modèles RNA ; deuxièmement, les agents de base qui exécutent d'autres fonctionnalités essentielles telles que le calcul des émissions équivalentes de CO_2 et le contrôle de la simulation. Le développement est basé sur les données algériennes du gaz et de l'électricité fournies par la société nationale d'énergie. La simulation consiste tout d'abord à prédire la production d'énergie à l'aide des agents de prévision et à calculer l'équivalent CO_2 émis. Deuxièmement, un agent dédié calcule le CO_2 total émis par toutes les sources disponibles. Il calcule ensuite les avantages de l'utilisation de sources d'énergie renouvelables comme moyen alternatif de répondre à la charge électrique en termes d'atténuation des émissions et d'économie de consommation de gaz naturel. Le système conçu peut aider les agences environnementales et les planificateurs énergétiques à examiner leurs politiques de contrôle de l'énergie et de la pollution atmosphérique.

Mots clés: multi-agents,RNA ,CO_2, prévision à court terme, énergie.

ملخص

يعد إنتاج واستهلاك الطاقة من أهم مصادر غازات الاحتباس الحراري إلى جانب الصناعة، وأحد الأسباب الرئيسية للاحتباس الحراري العالمي. يعد توقع التكلفة البيئية لإنتاج الطاقة أمرًا ضروريًا لاتخاذ قرارات أفضل وتسهيل التحول إلى أنظمة طاقة أنظف للحد من تلوث الهواء. تصف هذه الأطروحة نهجًا هجينًا يعتمد على الشبكات العصبية الاصطناعية وبنية متعددة العوامل للتنبؤ بثاني أكسيد الكربون (CO2) من مصادر مختلفة للطاقة في مدينة عنابة من بيانات حقيقية. يتكون النظام من عدة عوامل مستقلة، مقسمة إلى نوعين: أولاً، عملاء التنبؤ، والتي تتنبأ بإنتاج نوع معين من الطاقة باستخدام نماذج الشبكات العصبية الاصطناعية؛ ثانيًا، العملاء الأساسيون الذين يؤدون وظائف أساسية أخرى مثل حساب انبعاثات ثاني أكسيد الكربون والتحكم في المحاكاة. يعتمد التطوير على بيانات الغاز والكهرباء الجزائرية المقدمة من شركة الطاقة الوطنية. تتكون المحاكاة أولاً وقبل كل شيء من التنبؤ بإنتاج الطاقة باستخدام عوامل التنبؤ وفي حساب مكافئ ثاني أكسيد الكربون المنبعث. ثانيًا، يحسب العميل المخصص إجمالي ثاني أكسيد الكربون المنبعث من جميع المصادر المتاحة. ثم يقوم بحساب فوائد استخدام مصادر الطاقة المتجددة كطريقة بديلة لمواجهة الطلب على الكهرباء من حيث تخفيف الانبعاثات وتوفير استهلاك الغاز الطبيعي. يمكن أن يساعد النظام المصمم الوكالات البيئية ومخططي الطاقة على مراجعة سياسات التحكم في تلوث الهواء والطاقة.

الكلمات المفتاحية: العملاء المتعددين، الشبكات العصبية، ثاني أكسيد الكربون، التنبؤ قصير المدى، الطاقة.

Contents

List of Figures

List of Tables

One of the only ways to get out of a tight box is to invent your way out.

Jeff Bezos

1 Introduction

▷ *This chapter presents a general introduction to the thesis.* ◁

Contents

1.1 General context

Pollution in its various forms affects the quality of life in several cities around the world and has a very considerable economic and social cost. According to the World Health Organization (WHO) air pollution is estimated to be the cause of death of seven million people every year worldwide, and 90 % of people breathe air that contains high levels of pollutants. Air pollution is caused by different sources, both human-made and from natural sources. However, one of the biggest air pollution sources is energy production and consumption. Algeria plays a major role in world energy markets as it has the world's tenth-largest natural gas reserves, and is the sixth-largest exporter of gas and liquefied natural gas. According to the Algerian Ministry of Energy in its annual report for 2018 (MEN 2018) the total energy production is 166.5 Million Ton Equivalent of Petroleum (MTEP), of which 100.8 MTEP was exported in its different forms, with 1.5 MTEP being imported. In terms of production, the primary electricity production saw a big jump from 635 GWh to 783 GWh over the year of 2018, with an estimated increase of 25 %, while the natural gas production had a slight increase of 0.9 % to reach 97.4 Bm^3. On the other hand, the national consumption of energy reached 65 MTEP, with a significant increase of 7.7 % compared to 2017. This was mainly driven by the increase in natural gas consumption (13,4 %), which represents almost two thirds (65 %) of the total consumption and an increase in electricity consumption by 2.9 %. This is in the context of around 90 % of electricity being produced through natural gas-driven power plants (Khraief et al. 2018).

Given the above statistics, it is clear that the energy sector is the center of the Algerian economy, and its growing consumption and production is due to the socio-economic growth of the country. However, this growth comes at a massive cost to the environment, as energy consumption and production are considered two of the biggest air pollution sources. Therefore, finding solutions, including switching to cleaner energy sources, in order to minimize pollution and reduce the environmental cost is crucial. Algeria has immense renewable energy (RE) potential (Bouraiou et al. 2016). This is mainly from solar power due to Algeria's geographical location, which is considered one of the biggest solar energy potentials in the world, with a yearly estimation of 13.9 TWh. This is in addition to possible energy generated by wind and biomass. Therefore, providing energy planners with the necessary tools to forecast energy is essential for better management of energy and its costs.

1.2 Research question and objectives

Short-term load prediction plays an important role in the day-to-day operations of power management systems. it provides data that could help avoid contingencies and provide essential information for decision-making in order to control any technological, economic, or environmental risk. Thus, the new paradigms and the latest advances in electrical systems are based on the use of artificial intelligence, in order to optimize production and minimize losses and environmental costs. In addition, Artificial Neural Networks (ANNs) have shown their efficiency in estimating energy demand and in modeling time series

Once the electric load and natural gas forecasts have been validated, the production forecast, this time, of RE plants in all its forms (Photovoltaics, wind turbines, etc.) will have to be predicted by building suitable models. Knowing that RE production will take an important part in energy production in the short term in Algeria. The forecast of the RE prediction will have to be subtracted from the forecast of the overall demand produced using fossil fuels, natural gas in the Algerian case. The gain induced by RE production in terms of fossil energy must be quantified not only in quantity of fossil fuel (m^3 of natural gas in the Algerian case) but also in terms of impact on the environment and more precisely on-air quality. Therefore, the main objectives of this thesis are:

- Short-term ANN based forecasting models for regional electricity demand and natural gas consumption.
- ANN Forecasting models for the prediction of short-term RE production (photovoltaic and wind energy)
- Integration of the developed forecasting models in a collaborative MAS.
- Integration of an agent in the MAS to calculate and evaluate the contribution of the predicted RE production on pollution and for the quantity of fossil fuel necessary for an equivalent energy production.

1.3 Thesis organisation

The approach taken in carrying out this work is reflected in the organization of this document. This manuscript is structured in 7 chapters including this introductory chapter.

The second chapter presents the concepts related to air pollution, namely, the different types of pollutants, the origins and the effects of this phenomenon. In this chapter, we also present the main methods of forecasting air pollution in addition to a state-of-the-art of the recent research that were conducted in the field of energy related air pollution forecasting.

In the third chapter we present the concept of the artificial neural networks, their origins, architecture and types. Also, we illustrate how the ANN were used in the area of forecasting the energy production and consumption by presenting some recent works that were conducted in this area.

The fourth chapter is dedicated to the multi-agent systems, where we present a definition to the main concepts related to this area, from the agent aspect to the collective dimension of the MAS by presenting its main characteristics. In the end of the chapter we present a review of the MAS models used to simulate pollution related problems.

In the fifth chapter we describe the first part of our contribution, which consists of the forecasting of the energy production and consumption, hence, we'll present the available energy datasets of different energy sources in Algeria and how we used the ANN forecasting models for short-term forecasting for these energy sources.

In the sixth chapter we present our proposed approach for the design of a multi-agent based air pollution system by describing the designed agents and how the forecasting models are integrated into these agents and how these agents interact with each other, and we'll finish by presenting how the CO_2 emission are calculated using the emissions factors of CO_2 and the present some simulation results.

In the end, we conclude this manuscript with a general conclusion by which we summarize the contributions made through this research work, we also present the future perspectives of our work.

Computer Science is no more about computers than astronomy is about telescopes.

Edsger W. Dijkstra

2
Air pollution modelling

▷ *This chapter explores the main concepts related to air pollution.* ◁

Contents

2.1 Introduction

BEfore describing the methods and approaches used for modelling air pollution, some basic concepts mainly related to the environment and air pollution, should first be described. This chapter explores the main concepts related to air pollution. First, some essential definitions as well as the main types of air pollution are presented along with the description of the different pollutants and their sources. Finally, we present a brief state of the art on air pollution modelling with an emphasis on prediction and forecasting methods and models.

2.2 Air pollution

Air is defined as the odourless, colourless and invisible gaseous fluid that makes up the atmosphere and that the living things breathe. It is generally associated with the troposphere, which is the lowest atmospheric layer and closest to the earth (8 to 15 km altitude), however pollutants that have a long lifespan can reach the stratosphere (15 to 50 km altitude). Therefore, the air is sensitive to being influenced by human activities and natural phenomena such as photosynthesis in plants and evaporation from the oceans.

2.2.1 Chemical composition of air

The Earth's atmosphere, or simply air, is a mixture of different gaseous substances and of varying concentrations. The gravitational field of the earth retains these particles around the earth.

The main gases that make up air are molecular nitrogen (N_2) as well as dioxygen (O_2), and traces of rare gases such as krypton, helium, etc. These gases are found in the air with identical concentrations regardless of location and up to an altitude of 100 km, because beyond this height there is hardly more air. The air also contains water vapour in varying concentrations depending on location and weather parameters. The following table illustrates the different constituents of air with their concentrations and molar masses :

Table 2.1: The air constituents.

Constituent gases of dry air	Concentration (%)	Molar masses (O = 16,000)
Azote (N_2)	78,09	28,016
Oxygen (O_2)	20,95	32,000
Argon (A)	0,93	39,944
Carbone Dioxide (CO_2)	0,035	44,010
Neon (Ne)	1,8 10^{-3}	20,183
Helium (He)	5,24 10^{-4}	4,003
Krypton (Kr)	1 10^{-4}	83,07
Hydrogene (H_2)	5 10^{-5}	2,016
Xenon (Xe)	8 10^{-6}	131,3
Ozone (O_3)	1 10^{-6}	48,000
Radon (Rn)	6 10^{-18}	222,00

2.2.2 Structure of the atmosphere

The atmosphere consists of several layers, each layer has its own characteristics, the boundaries of each of its layers have been determined according to the discontinuity in the temperature variations. Therefore, the atmosphere is divided into five main layers:

- **The troposphere:** The atmospheric layer closest to the surface (8 to 15 km altitude), this layer contains 80 to 90% of the total mass of the air and almost all of the water vapour, in this layer meteorological phenomena occur (such as rain, clouds, etc.), and the temperature decreases depending on the altitude.
- **The stratosphere:** up to 50 km, in this layer the temperature increases depending on the altitude up to 0°C, it includes the ozone layer.
- **The mesosphere:** up to 80 kilometres, the temperature drops to -80°C, this is where most meteorites burn off on entering the atmosphere.
- **The thermosphere:** up to 150 Km, it is characterized by the increase in temperature which can reach 1,700°C, and the very low pressure.
- **The exosphere:** it's a thin volume similar to the atmosphere and surrounds the earth, it's where the atmosphere merges with the interplanetary space, it's consists mostly of helium and hydrogen.

2.2.3 Air Pollution

(Barker et al. 1963) describes the atmospheric (or air) pollution as the alteration of the chemical composition of the air, by increasing the concentration of certain substances (chemical or solid), this alteration lasts for a certain period of time to create toxic effects, and which will also have harmful effects on the ecosystem and health. The chemical composition of the air has been constantly evolving for thousands of years, these changes are due to disturbances of natural origin: Geological such as volcanic eruptions, Astronomical such as collisions with meteorites, or even biological (pollen, natural decomposition). Over the past 300 years, human activities have been causing more and more disruption for a relatively short period of time. Pollution can be local (on a regional scale such as: ozone peaks, smog), or global on a planetary scale (the greenhouse effect, destruction of the ozone layer).

Pollution affecting the tropospheric layer is the most important type, as it is in this layer that humans and most species live. Other types of pollution may be considered as less important, except those that influence the climate, such as greenhouse gases, the latter affect the thermosphere layer.

Air pollution is closely linked to other types of pollution, as the clouds of pollutants are transformed and participate in the increase of the phenomenon of acid rain. These rains transform pollutants and spread them in other forms: aquatic pollution, soil contamination and pollution, etc.

2.2.4 Types of pollutants

The main air pollutants are divided into two very distinct large categories: primary pollutants and secondary pollutants.

2.2.4.1 The Primary Pollutants

Primary pollutants are emitted directly from pollution sources such as road traffic, industries activity, heating, energy production and consumption, etc. For example (SEPA 2021):

Sulfur oxides (SO_x)**:**
SO_x is a pollutant produced by many industrials and processed by volcanos, it's mainly found in petroleum and coal, therefore its linked to the production and combustion of these energy sources, its most known for is the Sulfur diox-

ide (SO_2). SO_2 also is caused by some other industrial processes such as steel making and fertiliser manufacturing, and can be caused by natural sources as well, for instance geothermal activity.

This pollutant causes respiratory issues such as bronchitis and asthma attacks, as long as irritating nose, throat and lungs, and it has been linked to cardiovascular disease.

Carbon monoxide (CO)**:**
This toxic pollutant is an odourless and colourless gas, it results from the combustion of coal, natural gas, and wood, in addition to road traffic where studies showed that it is responsible of more than half of the emitted CO into the atmosphere.

CO is a very dangerous gas, as it attaches to haemoglobin's molecules in the blood, therefore it reduces the amount of oxygen that the body receives. The exposure to low amounts of CO makes you feel weak, dizzy, and disoriented, in the other hand, the exposure to high amounts of CO can lead to serious health problems from loss of conscious to death.

Carbon dioxide (CO_2)**:**
CO_2 is known to be the most significant pollutant among the others, as it is speedily produced and emitted by different sources, and it's more harmful to the climate than the others which led it to be called the worst climate pollution. Naturally, (CO_2) exists with and immense quantity in the atmosphere and it's essential for plants, however, recently it's concentrations in the atmosphere increased greatly from 280 parts per million (ppm) in the pre-industrial era to 410 ppm. Its main source is the combustion of fossil fuels.

Nitrogen oxides (NO_x)**:**
Its main form is Nitrogen dioxide (NO_2), it is a chemical toxic gas and it's char-

acterizes by a biting and sharp odour. NO_x results from the of oxygen and nitrogen gases in the air during combustion at high temperatures, therefore it results from the human activities such as combustion of fossil fuels, its natural sources include volcanos, lightning, and bacteria.

NO_x main effect on the human health is on the respiratory system, it can aggravate asthma, and lower the lungs function and decrease their defence which makes them more vulnerable against bacteria and infections.

Ammonia (NH_3)**:**
NH_3 is a compound of hydrogen and nitrogen, it's a colourless gas with a pungent odour. Ammonia is classified as a hazardous pollutant, and its sources include both natural and anthropogenic sources.

Ammonia can cause a decrease in the biodiversity of the ecosystems, and can cause dangerous complications to humans when it's inhaled such as burning the nose and throat, respiratory distress, bronchiolar ... etc.

Particulate matter (PM)**:**
The term PM refers to the liquid or solid particles suspended in air, their source can be either from human activities such as fossil fuel combustion and power plants, or they can be originated from natural sources, such as forest fires, volcanos, dust storms, etc. PM are classified according to their diameter as they vary greatly in size, particles with diameters less than 10 micrometre can penetrate into our noses and lungs and cause a lot of health issues, for instance, $PM_{2.5}$ and PM_1 generally have slower sedimentation rates and can therefore remain suspended in the atmosphere for days or even weeks. Consequently, these particles can be transported over long distances and undergo physico-chemical transformations. Inhaling these particles can aggravate bronchitis and asthma, they can also irritate the eyes, throat, and lungs.

Volatile organic compounds $(VOCs)$**:**

These pollutants are categorized into two categories, methane and non-methane parts. Methane (CH_4) plays a major role in causing global warming, while the non-methane parts $(NMVOCs)$, xylene, benzene, and toluene are suspected to cause cancer such as leukaemia.

2.2.4.2 The secondary pollutants

The secondary pollutants are not emitted directly from pollution sources; instead, they are formed in the atmosphere from different compounds. For example (SEPA 2021):

Ozone (O_3)**:**

O_3 exists in two different parts of the atmosphere and its effects depends on where it is situated, in the higher level of the atmosphere (the stratosphere) it produces naturally and it plays a vital role in that layer, because it protects the earth from the ultraviolet rays coming from the sun which damages the health, thus, this layer is often referred to by the ozone layer. However, the lower layers of the atmosphere and on the earth's surface, O_3 is formed through photochemical reactions between other pollutants such as the reaction between VOC and NO_x, and it can damage our health by causing cardiovascular and respiratory diseases, and affect the performances of lungs.

Peroxyacyl nitrates $(PANs)$**:**

PANs are a secondary pollutant as they form through photochemical reactions, and can cause respiratory problems and diseases. Under cold weather (< -20) PANs can stay in the atmosphere for up to three months, while in the warm weather they persist only for few hours. PANs decomposition can generate several types of other dangerous chemicals including CO_2 and CO.

Sulfuric acid (H_2SO_4)**:**
H_2SO_4 is a very corrosive chemical that can exist as droplets or particles in the atmosphere, and it dissolves when exposed to water. Its main sources in the industrial that produce it or use it during the production such as oil refiners, metal smelters, and phosphate fertilizer producers, and it can be produced from natural sources such as volcanos. This pollutant can severely burn the eyes and skin, it can cause blindness and third degree burns on contact.

Nitric acid (HNO_3)**:**
This pollutant is rarely found in nature in its gas form, however it is produced mainly from traffic emissions, like sulfuric acid, it's a corrosive chemical that can lead to irritating eyes, throats, and lungs, and can cause sever skin burns.

Table 2.2: Sources and effects of some pollutants

Pollutant	sources	Maximum normal concentration in the atmosphere	Effects on the environment	Effects on the health
CO	Traffic, fires, industrial activities	35 ppm (1h period) - 9 ppm (8h period)	contributes to smog formation	Heart diseases, vision problems
NO_x	Traffic, Electricity generation, industrial activities	0.053 ppm (1 year period)	contributes to smog formation, damage to foliage	inflammation and irritation of breathing passages
SO_2	electricity generation, fossil-fuel combustion, industrial activities, traffic	0.03 ppm (1 year period)	Haze, contributes to acid rain formation, reacts to form PM	Heart diseases, breathing problems
O_3	Electrical utilities, gasoline vapours, traffic	0.075 ppm (8h period)	interferes with the respiring ability certain plants, leading to increased susceptibility to other environmental stressors	Lung problems and diseases
PM	Construction sites, fires, reactions between gaseous chemicals	150 $\mu g/m$ (24-hour period for particles <10 μm) and 35 μg/m (24-hour period for particles <2.5 μm)	Haze, acid rain	Heart diseases, lung and breathing problems

Table 2.2 above illustrates the sources and the risks of some of the above-mentioned pollutants (Ghazi 2017).

2.2.5 Air pollution sources

Air pollution has many sources and causes, which can be either natural sources such as forest fires, volcanic eruption, etc., or anthropogenic, i.e., linked to human activity (Mayer 1999). In the second case, pollution is often the direct result of industrial progress in recent centuries, such as the continuous and sometimes careless emission of pollutants associated with combustion processes such as traffic emissions, industrial installations, production of energy by combustion fossil fuels, agriculture, etc.

Human activities play a major role in producing most of the primary pollutants, and mainly due to the energy sector which is by far the biggest source of air pollution emissions that is related to human activity, that's due to fossil energy combustion, coal extraction, bioenergy, oil refining, etc.

Energy generation and consumption in addition to being responsible for the majority of the air pollution from human activities, they are also responsible for a very high proportion of some lead pollutants. Figure 2.1 illustrates that this sector is responsible for more than 85% of the emitted PM, NO_x and SO_2.

2.2.6 The role of energy in air pollution

Global energy markets are still dominated by fossil fuels, despite the increasing attention being given to other sources of energy, the natural gas, coal and oil are still having a remarkably stable share over the last 25 years. For instance,

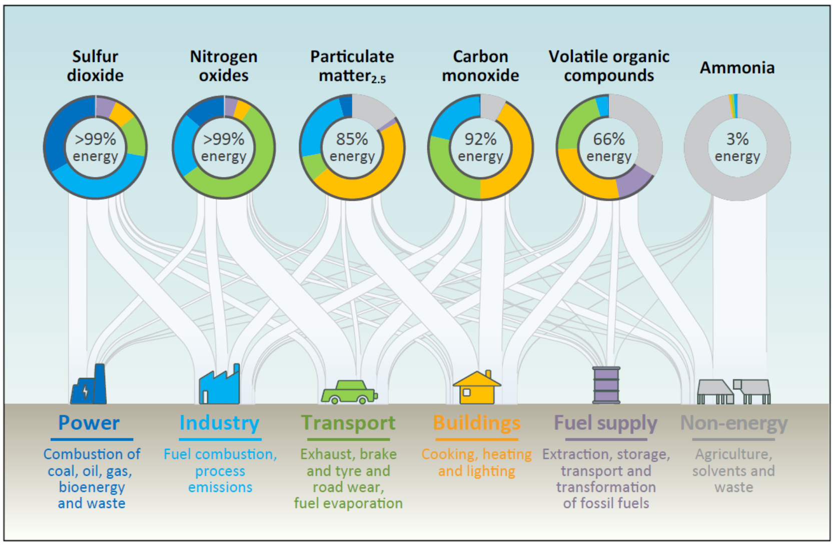


Figure 2.1: The main sources of the primary pollutants (IEA 2016)

in 1989, these fuels share of the total energy use was 81%, and it's the same percentage as 2014.

Similarly, power generation is one of the major sources of air pollution in the world, as the great majority of the world's power generation is still produced using from the combustion of fossil fuels, and mainly the coal-fired power plants which still the global power generation sector by producing 40% of the global electricity. In general, power generation has a big contribution to the global air pollution, for instance, its responsible of one third of the total SO_2 emissions, 14% of the total NO_x, and 5% of the total amount of $PM_{2.5}$. Coal is known to be the largest contributor to air pollution by accounting of 75% the sector's SO_2 emissions, 70% of the NO_x emissions, and more than 90% of $PM_{2.5}$ emissions. Oil-fired power plants are a minor contributor to the global power sector by only 4% of the total power generated, most of it comes from old power plants and in particular in middle east without or with minimal pollution control technology. In the other hand, Natural gas-fired power plants are known to emit less air pollutants than the oil and coal-fired once, however, it still contributes by 20% of total NO_x emitted from global power generation, but barely emitted any $PM_{2.5}$ or SO_2 (IEA 2016).

2.3 Energy related air pollution (state of the art)

Environmental problems in general, and pollution in particular, are major concerns due to their direct impact on our daily life and wellbeing. Air quality prediction has drawn strong attention in recent years and such problem is highly challenging as it is dependent on a variety of complex factors. In this section, we'll present a review of recent work in the field of energy related air pollution prediction.

2.3.1 The characteristics of a prediction model

In the field of air pollution prediction, models can be distinct and categorized according to several criteria:

- The type of the used model: several prediction model types can be used for air pollution prediction, for instance, linear models (ARIMA, ARX, ..., etc), Non-linear like ANN, or other AI models such as genetic algorithms and fuzzy logic, or a hybrid approach that consists of a several types of algorithms.

- The goal of the prediction: some models are used to predict one or more pollutants such as CO_2, O_3, NO_x, ..., etc, and some models are used to predict the air quality index which is an aggregation of several pollutant concentration values.

- The prediction horizon: the goal can be to predict a short, medium or a long-term pollution value.
- The used data: most of the prediction models are data-driven, therefore, the data size and the pre-processing play a major role in their performances.
- The used parameters: choosing which parameters to include, such as meteorological data, traffic data, . . . , etc.
- The studied region: Some models are designed to give predictions to a small region like a city, while the others are designed to give predictions to a whole region or even a whole country.
- Whether or not to include the chemical interactions between the pollutants: some models include these interactions by modelling them using mathematical equations.

Based on table 2.3, it can be seen that different types of methods and algorithms has been used in the field of air pollution prediction, however, ANN emerge as one of the most used methods, as they have become alternatives to the classic statistical methods used in air quality monitoring stations. These methods are based on the analysis of time series of the different pollutants, to detect trends in each series and provide predictions. It would be very effective to use the large amounts of information available from monitoring stations for designing efficient prediction models with the aim of having more effective air pollution monitoring and surveillance systems. Such systems would be very useful in studying the phenomena of air pollution in order to find solutions to reduce it by controlling its main sources.

Table 2.3: State-of-the-art research on energy related air pollution

Approaches	Prediction Goal	Used data	Study area	Methods	Results	Reference
Traditional approaches	Energy related CO_2 emissions	Energy data	Baoding (China)	Gray prediction model	N/A	Wang and Zhang 2013
	Gaseous pollutant concentration	Historical data	N/A	Gaussian Process Mixture model with iterative learning	RMSE: 0.09	Zhou et al. 2019
	CO_2 emissions from Fuel combustion	Energy and socio-economic data	China	Gray multivariable model	MAPE: 2.71	Ding et al. 2017
	GHG emissions	Energy data	India	ARIMA model	RMSE: 0.017, 0.031	Sen et al. 2016
	PM_{10}, CO, SO_2, NO_2, O_3	Historical emission data	India	LR	N/A	Taneja et al. 2016
	Electricity related CO_2 emissions	Energy data	N/A	Kalman Filter	N/A	Lau et al. 2014
	Energy related GHG emissions	Historical emission data	China	Adaptive model with buffered rolling	MAPE: 2.71	Xu et al. 2019
AI approaches	$PM_{2.5}$, NO_2, O_3	Historical data	Anhui (China)	TLS-BLSTM	RMSE: 11.23; 8.15	Ma et al.2019
	SO_2, NO_x, CO_2	Energy and socio-economic data	Lithuania	Fuzzy model	N/A	Oder et al. 1993
	GHG emissions	Historical emission data	Canada	Fuzzy model	RMSE: 5090	Mohsen and Abbassi 2020
	Energy related CO_2 emissions	Historical emission data	Ireland	RNN	MAE: 8.2; 32.8	Mason et al. 2018
	PM_{10}	Meteorological and historical data	N/A	RNN, LSTM, GRU	RMSE: 0.40	Athira et al. 2018
	Energy related CO_2 emissions	Historical emission data	N/A	NARX, EPUNN	RMSE: 33.28	Sheta et al. 2015
	Air quality index	Historical data	Malaysia	ANN, PCA	RMSE: 10.01	Azid et al.2014
	PM_{10}	Meteorological and historical data	Greece	ANN	Index of agreement: 0.82	Moustris et al.2013
	O_3, PM_{10}	Historical data	London	Spiking NN ensemble, clustering	RMSE: 9.33	Maciag et al.2019
	Air quality	Meteorological and historical data	N/A	Deep space-time model	MAE: 22.97; 24.98	Wang and Song 2018
	NO_x	Energy data	N/A	SVM	MAPE: 12.54, 8.32	Duan et al. 2018
	CO_2 tax pricing	Energy data	China	SVR	N/A	An and Zhai 2020
	Coal facility's NO_x and CO emissions	Coal facility related data	USA	NARX, PSO	RMSE: 0.01; 0.05	Safdarnejed et al. 2019
	CO_2 emissions	Historical emission data	China, Japan, USA	Gaussian processes regression, PSO	MAPE: 2.93	Fang et al. 2018
Hybrid approaches	Air quality index, air pollutant concentrations	Meteorological and air quality data	Taiwan	Neuro-fuzzy network	RMSE: 10.92; 5.19	Lin et al.2020
	Air quality index	Historical data	N/A	Secondary decomposition, sample entropy, LSTM, least squares SVM	RMSE: 8.89; 3.8	Wu and Lin 2019
	CO_2 emissions	Energy and socio-economic data	India	MLR, PSO, ANN	RMSE: 11.73; 4.32	Sangeetha and Amudha 2018
	CO_2 emissions	Energy and socio-economic data	China	K-means, Logistic model	R^2: 0.98	Ma et al. 2017
	CO_2 emissions	Energy and socio-economic data	China	Gray relational analyses, PCA, LSTM	MAPE: 0.43	Huang et al. 2019

Several researches were conducted in the field of air pollution prediction, however, in this work, the focus is on energy consumption and production related emissions, different recent researches have been carried out to predict pollutant emissions related to power generation and consumption. These researches can be classified into two categories, the researches based on traditional methods, and the once based on artificial intelligence methods.

2.3.2 Traditional methods

2.3.2.1 Grey model

Grey model concept was first introduced by Ju-Long in 1982 (Ju-Long 1982) as an approach to solve the problems where data is limited and not sufficiently available. Its standard form is noted as GM(1,1), it depends on two parameters, the first parameter (a) represents the developed grey coefficient, while b represents the controlled grey variable. Following a differential equation of GM(1,1), the two parameters are estimated using the least square method, which is a statistical method used in regression analysis. Therefore, the future value to be predicted V (CO_2 concentration for instance) is then predicted using equation 1.

$$V_{(t+1)} = (1 - e^{a})(x(1) - \frac{1}{b})e^{-at} \tag{2.1}$$

In order to model and forecast carbon emissions from production and consumption of electricity from 2011 to 2015 in the city of Baoding in China, (Wang and Zhang 2013) used a grey prediction model to estimate the amount of energy related carbon emissions, the studied energy sources include some primary energy types such as oil, natural gas, and coal.

To estimate the CO_2 emissions from fuel combustion in China, (Ding et al. 2017) used a gray multivariable model to predict the CO_2 emission from 2014 to

2020. The collected data concerns the period from 2005 to 2013 and consists of four indicators, which are the energy use values, CO_2 emissions, GDP, and the urban population.

(Wang and Ye 2017) applied non-linear gray multivariable models to forecast carbon emissions from fossil energy consumption in 30 Chinese provinces by the estimation of the non-linear impact of Chinese economic growth on the carbon emissions from fossil energy consumption.

(Xu et al. 2019) proposed an adaptive grey model with buffered rolling in order to forecast greenhouse gas emissions from energy consumption in China for the period from 2017 to 2025.

2.3.2.2 ARIMA models

Autoregressive integrated moving average (ARIMA) is a first-class time series forecasting technique and it's an alternative of the autoregressive moving average (ARMA) model which was introduced by Box and Jenkins in (Box and Jenkins 1994). ARIMA was introduced also by box and Jenkins in (Box and Jenkins 1976).

ARIMA models are defined by three parameters: p, d, and q, which refer to the number of lag observations included in the model and also called the autoregressive parameters, the number of differencing passes, and the moving average parameters respectively.

(Sen et al. 2016) used ARIMA models to forecast the energy consumption and GHG emissions for an Indian pig iron manufacturing organization in order to implement a better environment policy by observing the current and future trends.

2.3.2.3 Linear regression

Despite a big number of alternatives, linear regression is still one of the most used traditional approach for time series forecasting, the aim of linear regression is to find the relationship between a scalar called dependent and one or more explanatory variables known as independent variables.

In case of one independent variable, it's called simple linear regression, while in the case of more than one independent variable, it's called multiple linear regression (MLR). In a forecasting problem, a MLR model can be formulated as follow:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n + x_n + \varepsilon \tag{2.2}$$

- Where y is dependent variable,
- $x_1, ..., x_n$ are the independent variables
- $\beta_1, ..., \beta_n$ are regression coefficients
- ε is the error term

With the help of data mining techniques and more precisely linear regression, (Taneja et al. 2016) analysed the pollution trends in Delhi to predict the future trends. The predictions showed that there will be a considerable increase in PM by 46 %, in addition to a slight increase in the amount of both CO and NO_2. In the other hand, some pollutants such as SO_2 will show some decrease, mostly due to using non sulphur fuel.

2.3.2.4 Kalman filter

Linear quadratic estimation (LQE), more known as Kalman filter, is an algorithm that uses a series of an observations recorded over time to produce estimations of unknown variables.

This algorithm consists of two steps, the first step is the prediction, it produces estimations of the current variable states and its uncertainties, once the outputs are estimated, these observations are then updated using with a weighted average, with the more certain observations getting the more weights. This algorithm is named after Rudolf Emil Kalman, a Hungarian-American electrical engineer, who was one of the first developers of this theory (Kalman 1960).

(Lau et al. 2014) introduced an adaptive seasonal model based on the hyperbolic tangent function, which was used to identify the daily and seasonal patterns of electricity load demand and used a stochastic model to combine the trends. Then, they employed an Ensemble Kalman Filter to predict the load consumption and CO_2 emissions. The work presented three case studies which are Irish smart grid data, Elexon national electricity production data, and Brunel University photovoltaic production data.

2.3.3 Artificial intelligence-based methods

2.3.3.1 Support vector machine

Support vector machine (SVM) is one of the most used machine learning algorithms, it was first introduced in 1992 by Vapnik (Boser et al. 1992). The idea in SVMs is to fit a hyperplane between two different classes in a multidimensional dataset so that the distance between the marginal levels is as large as possible and there are no data points between the fringe as illustrated in figure 2.2.

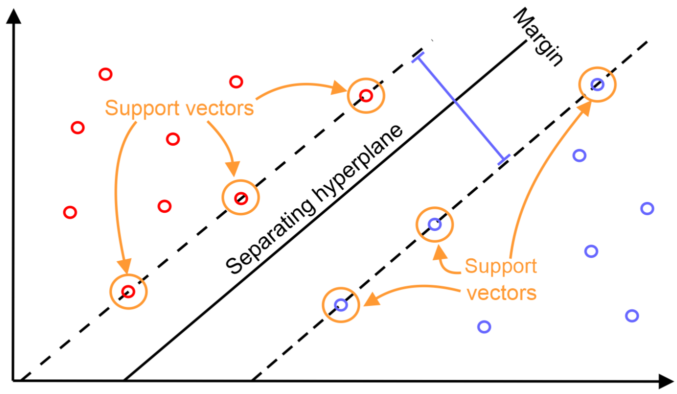


Figure 2.2: Support vector machine (Ghasemlounia and Saghebian 2021).

Another extension to SVM is support vector regression (SVR). SVR has been used successfully to address regression problems on different research areas (Ahmed et al 2020). SVR keeps all the main characteristics and principals of SVM for classification with some minor changes. The main purpose of SVR is to accurately fit a regression function y = f(x) in order to predict the outputs y_i corresponding to a set of input observations x_i.

(Duan et al. 2018) used SVM models to analyse and predict the NO_x emission resulting from a hydrogen enriched compressed natural gas engine, the prediction models achieved results varying from 12.54% to 56.66% according to MAPE metric.

In order to reduce the carbon emissions and lower the energy costs, (An and Zhai 2020) proposed a framework based on SVR model and data envelopment analysis to predict the Chinese thermal power generation and the carbon tax pricing.

2.3.3.2 Fuzzy logic models

Fuzzy logic can be described as a reasoning method that imitates the human reasoning, by simulating the way humans think and take decisions, including all the possibilities between the digital values NO and YES.
Fuzzy logic systems consist of four components:

- **Fuzzification module:** transforms the input values into fuzzy sets
- **Knowledge base:** consists of IF/THEN rules specified by experts
- **Inference Engine:** simulates the human decision making by applying fuzzy inference on the inputs and using the knowledge base.
- **Defuzzification Module:** transforms the obtained fuzzy sets into output values.

(Oder et al. 1993) proposed a fuzzy logic model to assess the decline of the energy related pollutants emissions in regard the decision makers' emission reduction policies in Lithuania, the pollutants that were taken into account are SO_2, NO_x, and CO_2.

(Mohsen and Abbassi 2020) used a fuzzy logic-based model to predict the emitted GHG from large solid waste landfills that is used to generate energy in the Canadian city Ontario. The proposed model includes seven input parameters, and a knowledge base that consists of 19 IF-THEN rules.

2.3.3.3 Gaussian process model

Gaussian process regression (GPR) is a machine learning technique used for regression, it's a nonparametric Bayesian approach, its main advantages are working well on small datasets, and having the ability to give measurements of uncertainty on the predictions.

(Fang et al. 2018) developed a Gaussian processes regression approach based on swarm optimization (PSO) to forecast carbon emissions. The PSO algorithm was used to optimize the hyperparameters of the covariance function. The system was validated using the total carbon emissions data of China, Japan and the US for the period from 1980 to 2012 and was used to predict the emissions values from 2013 to 2020.

(Zhou et al. 2019) applied a Gaussian Process Mixture model adopting iterative learning to predict the gaseous pollutant concentration. The data were collected in real time and concern the hourly concentration of four gaseous pollutants, with a total of 9336 observations. The pollutants that were taken into consideration in this study are Nitrogen oxide (NO_x), Hydro carbon (HC), Ozone (O_3) and Nitrogen dioxide (NO_2).

2.3.3.4 Artificial neural network

ANN are a machine learning technique inspired by the human brain and simulates the way it works; they consist of several interconnected calculation units called artificial neurons. ANN has been largely used to solve problems in multiple areas, such as medicine, statistics, technology, and environment due to its intense ability to adapt to changing inputs and multiple types of data.

(Ma et al. 2020) focused on the data shortage problem in air quality forecasting by proposing a method based on transfer learning and deep learning techniques. The authors presented a stacked bidirectional LSTM neural network based on transfer learning (TLS-BLSTM) to predict the concentration of

$PM_{2.5}$, NO_2, and O_3 in Anhui, China.

(Moustris et al. 2013) developed a PM_{10} prediction model for forecasting the daily concentration levels for 24 h ahead in addition to the number of hours during which PM_{10} level exceeds the recommended rate. Their model consists of a neural network that was trained using data from four cities in Greece, the architecture of the implemented ANN consists of 3 layers, where the input layer contains 23 neurons, the hidden layer contains 14 neurons, and the output layer consists of 2 neurons, one to predict the mean daily PM_{10} concentration and the hours during the day in which the PM_{10} concentration exceeds the selected threshold which in their case was equal to 50 $\mu g/m^3$.

(Azid et al. 2014) presented an ANN model, which uses selected inputs using the principal component analysis (PCA) method. Data from 7 years of measurements concerning air pollution parameters in different cities in Malaysia were used. The most influential parameters that were selected by the PCA are $O_3, NmHC, THC, CH_4$ and PM_{10}, and the aim of the model was to provide predictions of air quality index, and the results identified the pollutants that have most influence on the index.

(Maciag et al. 2019) proposed an ensemble model based on clustering to forecast air pollution using Spiking Neural Networks which is an ANN type where the information processing in the nodes and the neurones communicate through the exchange of spikes. Here, the clustering was used to divide the data into k distinct samples according to their pollution level, and each neural network model is trained on distinct cluster, the pollutants taken into consideration in this work are ozone (O_3) and PM_{10} for the London area.

To forecast (PM_{10}) in the air (Athira et al. 2018) employed three types of ANN, which are RNN, Long Short-Term Memory (LSTM), and Gated Recurrent Unit. The models were trained using the time series dataset AirNet which includes both meteorological and historical air quality data.

(Mason et al. 2018) implemented a recurrent neural network using a covariance matrix adaptation evolutionary strategy for forecasting short term power demand, wind power generation, and CO_2 emissions in Ireland. The data used in this work consists only on historic time series, and did not include any exogenous inputs.

A RNN model was developed by (Safdarnejed et al. 2019) to estimate NO_x and CO emissions issued from a coal fired utility. They used a particle swarm optimization (PSO) algorithm in order to obtain reduced emissions. They also examined the benefits of using dynamic optimization of the combustion process against a steady state approach in reducing the emissions of CO and NO_x, the data was collected from a 480 MW coal fired power plant in the USA, and it consists of 5-min sampled time data from the period of 21/09/2017 to 05/03/2018, and the considered forecasting horizon is 3 hours in advance.

(Sheta et al. 2016) used two types of ANNs, a Nonlinear Auto-Regressive with exogenous Input (NARX) and an Evolutionary Product Unit Neural Network, to predict the global CO_2 emission issued from world energy consumption based on energy attributes dataset, which includes the consumption values of natural gas, oil, and coal from 1980 to 2010.

(Wang and Song 2018) designed an approach to forecast air quality using meteorological data, historical air quality and a deep space-time ensemble model. Their model consists of three components. First, an ensemble method with a partitioning strategy based on weather patterns. Second, a module to discover spatial correlation by evaluating Granger causalities between stations, and producing spatial data relating stations to areas. Finally, a deep LSTM prediction model for both long and short-term air quality prediction. The input data consists of historical air quality time series in addition to meteorological data.

2.3.4 Hybrid approaches

Several studies used a hybrid approaches which proved to be effective in different cases, for instance, (Wu et Lin 2019) proposed a hybrid model to forecast the air quality index. Their model, called SD-SE-LSTM-BA-LSSVM, consists of secondary decomposition, sample entropy, LSTM, and a least squares support vector machine optimized by the Bat algorithm. The forecast result is obtained by aggregating the forecasting values of each component.

(Lin et al. 2020) proposed an approach to predict the air quality index and air pollutant concentrations based on a neuro-fuzzy network trained using the backpropagation and particle swarm algorithms. In order to avoid dimensionality issues, the authors also performed a correlation analysis, to sort out the uncorrelated or weakly correlated variables with the target, and a principal component analysis to reduce the dimensionality of the input variables.

(Ma et al. 2017) introduced a system to forecast the carbon emissions in 30 Chinese provinces between 2014 and 2023 based on data from 2005 and 2013 and by selecting 5 parameters, which include the energy efficiency, urbanization rate, GDP, the proportion of the second industry, and the CO_2 emission intensity. The system consists of two methods: firstly, a K-means clustering algorithm to split the CO_2 emissions into five clusters where each cluster represents a region; secondly, a logistic model to predict carbon emissions in these clusters.

(Huang et al. 2019) employed gray relational analyses to determine the most impactful factors on CO_2 emissions and principal component analyses to extract four principal components. They then used a LSTM recurrent neural network to predict the carbon emissions, in contrast to simple feedforward neural networks. The rationale for this choice was that LSTM has feedback connection and it is able to process entire sequences of data and not only single data points.

(Sangeetha and Amudha 2018) used a multiple linear regression (MLR) model and a PSO algorithm to estimate carbon emissions from several emission sources. An ANN was used to forecast the future projection of CO_2 emissions.

2.4 Conclusion

In this chapter we have presented the different concepts related to air pollution are presented in terms of air constitution and pollutant types. The effects of each type of pollutant on the environment and on the human health are also described as well as the sources of each pollutant. We also presented the tools and approaches used to model or to predict air pollution. These approaches can be categorized into three families: traditional approaches, artificial intelligence-based approaches, and hybrid approaches. We have also presented a state-of-the-art reporting recent studies that address air pollution for each of the mentioned approaches with a table summarizing these studies.

I propose to consider the question, 'Can machines think?

Alan Turing

3

Artificial neural networks as a forecasting tool

▷ *In this chapter we present the concept of the artificial neural networks.* ◁

Contents

3.1 Introduction

IN this chapter we present the concept of the artificial neural networks, their origins, architecture and types, and present how the ANNs were employed in forecasting the energy production and consumption by presenting the recent works that were conducted in this area.

3.2 Machine Learning

Artificial intelligence (AI) is a general term that refers to all techniques and theories that were created to help mimic and simulate the human intelligence by machines or systems. Machine learning and AI are often referred to together as the terms are sometimes used interchangeably, however they do not mean the same thing. Machine learning is a subfield of AI that focuses on designing systems that has the ability to learn to solve a problem automatically based on the data they fed and not using explicit programming.

There are different approaches and techniques that vary depending on the nature of the problem being addressed and the type and volume of the available data. In general, these techniques are divided into three categories, supervised learning, unsupervised learning and reinforcement learning. The main differences between these approaches defined by how each one learns about the data to make predictions.

- **Supervised learning:** The goal of supervised learning is to uncover patterns in data and apply them to an analytical process. These data include characteristics associated with labels that define their meaning. Therefore, the training is guided with inputs and their desired outputs. We're talking about classification or pattern recognition when the data is discrete, or regression if the data is continuous.

- **Unsupervised learning:** Unsupervised learning involves training based on data that has no labels or an accurate and defined output. The goal is to define a function to describe the structure or the hidden model in the inputs.

- **Reinforcement learning:** the inputs are a set of feedbacks from a dynamic environment with which the algorithm is confronted. The goal is to achieve a predetermined goal. It's used to make the intelligent agent learn to maximize their profits and it's different from supervised leaning as it doesn't need labelled inputs and outputs.

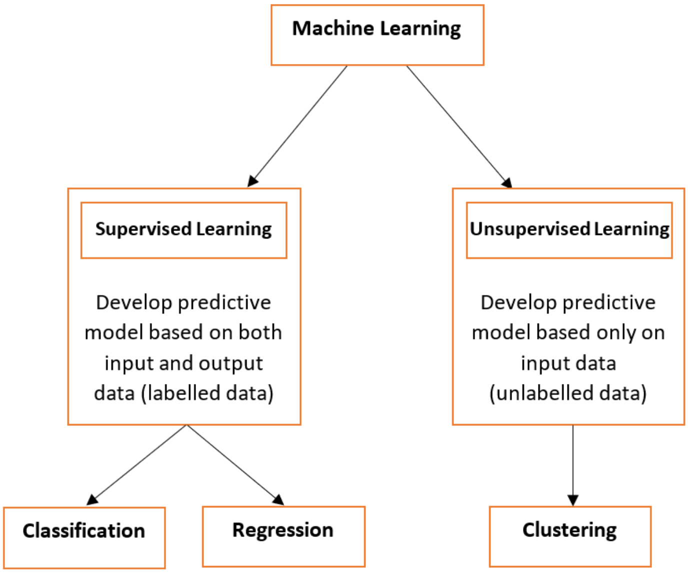


Figure 3.1: Difference between supervised and unsupervised learning

Figure 3.1 above illustrates the main differences between supervised and unsupervised learning.

3.2.1 Supervised learning

As previously described, supervised learning is a machine learning approach that requires a labelled dataset. This dataset should include an appropriate inputs and outputs for the used algorithm. From this set of examples, the program is guided to describe a model capable of predicting or classifying the correct output. At this point, the prediction / classification model should be validated with another labelled dataset independent of the training set, it is only when the validation phase is satisfactory that the algorithm can be considered reliable for use on unknown data. Therefore, considering a supervised problem and the type of data, the learning steps are as follows:

- **Selecting the algorithm:** as it exists a variety of algorithms, therefore selecting the appropriate algorithm to choose is a crucial step, and the choice should be based on the type of the studied problem and the nature of the available date, the available supervised algorithms include artificial neural networks, support vector machines, and decision trees, . . . etc.
- **Training:** This phase is the most important phase in machine learning, as the performance of the constructed model depends on it, in this step a labelled dataset is fed to the algorithm to be trained with, the objective of this phase is to try to build a model capable of adapting to the data provided, i.e. to predict as best as possible the correct output for each input provided, we also should note that the selected data must be as representative and generalized as possible in order to prevent over-fitting and under-fitting.
- **Validation:** In this phase the performances of the model created in the training phase are tested using another dataset called test set, the previ-

ously trained algorithm is used here to predict the input data of the test set. The Major difference between the training and test phase is that during the test phase the output labels are not used to improve the prediction abilities of the model, but they're only used to evaluate its performance.

- **Model deployment:** after training and validating the algorithm, and obtaining satisfactory performances, the model can be saved in order to use it as an automatic system to solve the original problem on the new data. Figure 3.2 illustrates the general process of machine learning.

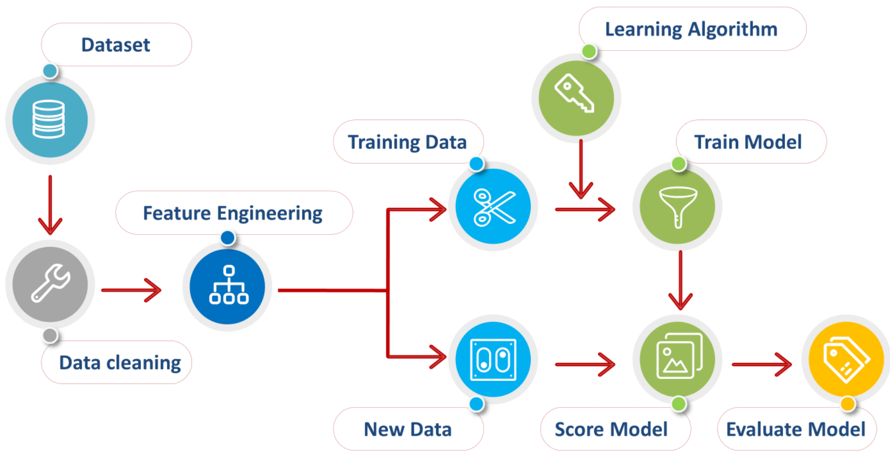


Figure 3.2: The process of machine Learning (Eldho et al. 2020).

3.2.2 ANN basic concept and overview

The concept of artificial neural networks was first introduced in 1943 at the University of Chicago by two researchers: neurophysicist Warren McCullough, and mathematician Walter Pitts. The two researchers presented their theory that the activation of neurons is the basic unit of brain activity (McCulloch and Pitts 1943) and they were the first to show that simple formal ANNs can perform complex logical, arithmetic and symbolic functions. Their model was is a binary neuron, with an output of 0 or 1. To calculate this output, the neuron performs a weighted sum of its inputs then applies a threshold activation function: if the weighted sum exceeds a certain value, the output of the neuron is 1; otherwise, it is equal to 0 (figure 3.3).

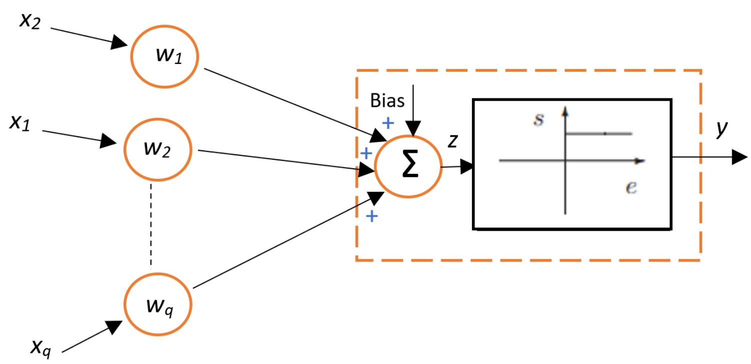


Figure 3.3: The formal neuron

The first ANNs which are Mcculoch-Pitts models lacked procedure for learning, which considered as a pivotal issue in order to be used for artificial intelligence; however more researchers have contributed in evolving the ANNs concept. In 1949, the Canadian psychologist and neurophysiologist Donald Hebb introduced the first learning rule for self-organized learning.

In 1957, Frank Rosenblatt, an American psychologist influenced by the work of Hebb invented the perceptron (figure 3.4) (Rosenblatt 1957), which is the first machine learning algorithm, the perceptron is a supervised linear classifier and it was built around a single neuron, therefore, it was limited to binary classification. Its first implementation was a machine named Mark 1, this machine was designed for image recognition, and included 400 photocells connected to neurons. Synaptic weights were encoded into potentiometers, and weight changes during learning were done by electric motors. This machine is one of the very first artificial neural networks. The algorithm was designed to classify visual inputs, categorize subjects into one of two types and separate groups with a separation line.

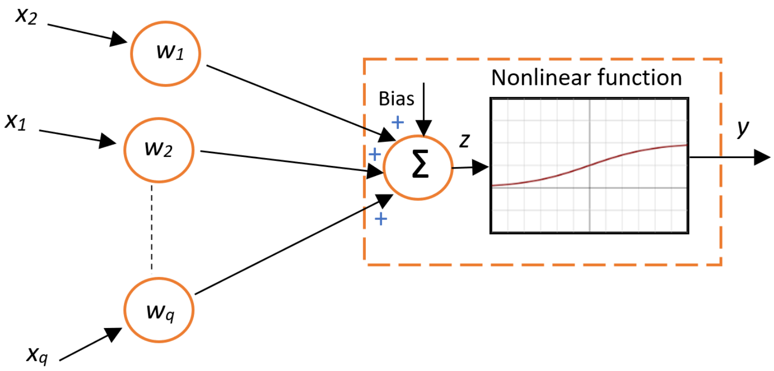


Figure 3.4: The Perceptron

However, Rosanblatt's perceptron suffered from some limitations, as it was analysed and proved by Minsky and Papert (Minsky and Papert 2017) that the perceptron can't learn the classes with nonlinear separability like the XOR function for instance. Moreover, it can be possible to solve such functions if a multilayer network is used, however at that time, there were no existing training algorithm for such architecture, even the Rosanblatt's perceptron learning rule was useless in such cases, as he indicated that the perceptron's learning rule become ineffective in multi-layered architecture. For several years, AI in general and ANN in specific were believed to have reached a dead-end, until 1986 where Rumelhart, Williams and Hinton three independent researchers proposed the delta rule to for learning by backpropagation (Rumelhart et al. 1988) which is until this day one of the most used training algorithms for multi-layered neural networks.

3.2.3 Components of ANN

3.2.3.1 Neurons

ANNs consists of multiple artificial neurons, these neurons are inspired and derived from the biological neurons, each neuron has one or more inputs and just one output, its inputs can be either feature vector of the input data such as texts and images, or they can be other neurons outputs.

3.2.3.2 Connections and weights

The neurons in the network are interconnected through connections, these connections transmit the output of one neuron as an input to another neuron, and each connection includes a weight that represents its relative importance.

3.2.3.3 Activation function

In ANNs an activation function is a mathematical function used to calculate the output of a neuron based on an input signal. It replicates the activation potential found in the human. It can be a simple threshold function that turns the neuron on or off based on a threshold value or it can be a transformation function that maps the inputs into output signal. An activation function can also help to normalize the output of each neuron to a range between -1 and 1 or between 0 and 1. In this section we will present the most used activation functions.

Linear function: Mostly used in the output layer in regression problems, it multiplies the inputs with the weights of each neuron to compute an output signal that is proportional to the input signal as illustrated in figure 3.5. **Binary**

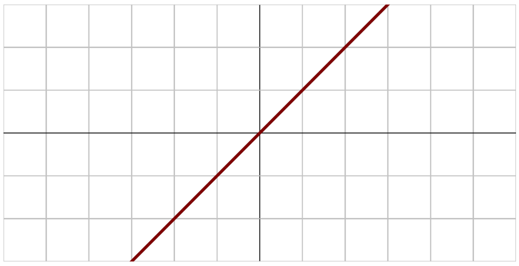


Figure 3.5: Linear function

step function: Used in binary classification, it's a threshold function where the neuron is activated whenever the input value is below or above a certain threshold (figure 3.6).

Sigmoid function: It's the most used function in ANN for decades, it's a differentiable function which is a useful property to train a neural network, one of the common forms of a sigmoid function is the logistic function illustrated in equation 3.1 where α is the slope parameter, the function takes an S shape as shown in figure 3.7.

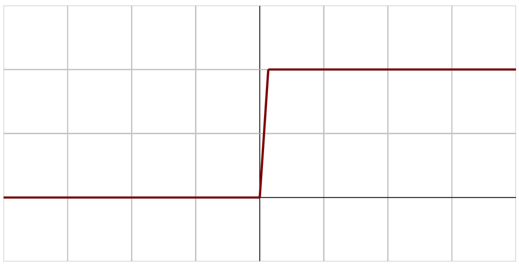


Figure 3.6: threshold function

$$\sigma(x) = \frac{1}{1 + e^{-\alpha x}} \tag{3.1}$$

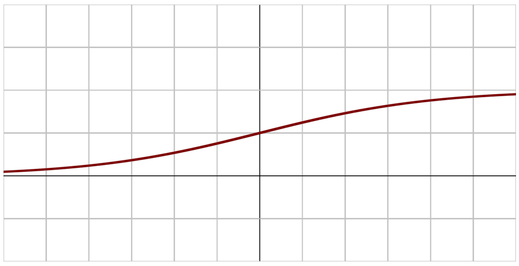


Figure 3.7: Sigmoid function

This function provides several advantages such as smooth gradient as it prevents the jumps in output values, normalizes the output of each neuron by bounding its value in a range of 0 and 1, and helps giving clear predictions by bringing the value of each prediction to the edge of the curve, however it suffers from some limitations as well, for instance, it suffers from vanishing gradient for the very high or very low inputs, the output values are not zero centered, and it's computationally expensive.

Hyperbolic tangent (tanh): Another popular activation function is tanh (equation 3.2), its characteristics are very similar to the characteristics of the logistic function except that it's zero centered as its values are bounded in a range between 0 and -1 which makes it a better choice to model values that have strongly negative or strongly positive values. The equation is illustrated in figure 3.8., we notice that it takes the S shape as well.

$$\sigma(x) = tanh(x) = \frac{sinh(x)}{cosh(x)} \tag{3.2}$$

Rectified Linear Unit (ReLU): This function is simple as it keeps the positive inputs as it is, and truncates the negative inputs to 0, this function allows the ANN to train faster compared to sigmoid and tanh functions, however it suffers from a phenomenon called dying ReLU where the gradient becomes 0 when the inputs are negative or approach 0. To overcome the dying ReLU issue, Leaky

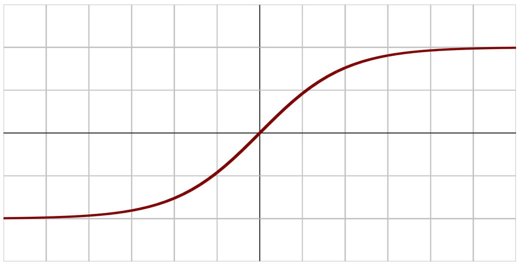
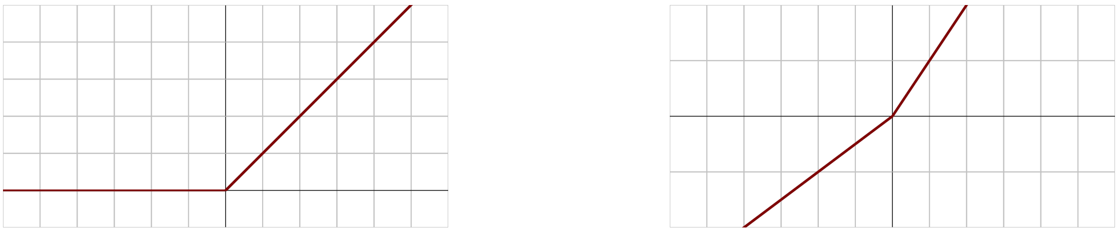


Figure 3.8: Sigmoid function

ReLU was introduced, which is an alternative form of ReLU, this function consists of a small positive slope in the negative area. Figure 3.9 illustrates both functions.



((a)) ReLU ((b)) Leaky ReLU

Figure 3.9: ReLU and Leaky ReLU activation functions

Softmax: Used in multi-class classifications in the output neurons, it normalizes the outputs in a range between 1 and 0 and divides it by their sum in order to give the probability of the input being in a specific class (equation 5).

$$\sigma(z)i = \frac{e^{z_i}}{\sum_{j=1}^{k} e^{z_j}} \tag{3.3}$$

Where i= 1,...,k, and z = $(z_1, \ldots, z_k)$.

3.2.4 Hyperparameters

The terms parameters and hyperparameters are usually used interchangeably, however there is a difference between them, as the parameters is a general term and in machine learning it mostly refers to the parameters that are learnt by the classifier or the model during the training such as the weights in ANNs. In the other hand, hyperparameters are parameters that their values are set before training and cannot be learned. ANNs consists of multiple hyperparameters that need to be set such as:

Number of hidden layers and neurons: Hidden layers are the layers between the output and input layers and adding the optimal number of layers and hidden units is essential to avoid either underfitting or overfitting.

Learning rate: the learning rate controls how fast the network updates the parameters, a large learning rate can speed up the learning but it may lead to the network not converging, in the other hand, a small learning rate can lead to a good convergence but it leads to a slow learning.

Number of epochs: this hyperparameter represents the number of times the whole training dataset is fed to the network during the training phase, a good practice is to keep increasing the number of epochs until the validation error starts increasing while the training error is decreasing which indicates an overfitting.

Batch size: it refers to the number of training samples given to the network after each iteration. Batch size can be selected according to three modes, the first one is to make it equal to the total number of samples and it's called batch mode, this makes the number of epochs and iteration equivalent, the second mode is mini batch mode, where we use a value that can be divided into the total number of samples, and finally the stochastic mode where the batch size is equal to 1, and that makes the parameters update after each given sample.

Hyperparameters optimization:

Hyperparameters defines the performance of the model, therefore they must be tuned carefully and based on the validation error, hyperparameters tuning can be performed in different ways, the most popular methods are as follows:

Manuel Search: In this approach, we follow an iterative process until we find a combination of hyperparameters that works well together using the knowledge acquired on the problem, at each step we try parameters and observe the result. then, we modify the hyperparameters based on these results until the performance become satisfactory.

Grid Search: This approach consists of an exhaustive search through an interval of a manually set of hyperparameters values, the model is then trained using all the combinations of these hyperparameters in order to choose the best combination. After the training is done, is it possible to redo it using a restraint interval of hyperparameters. Grid search is useful as it enables to search for several hyperparameters in parallel, but it demands a big computational capacity.

Random search: This method is similar to grid search; thus, it has the same advantages, the difference between the two methods, is that in the random search approach we choose the value of the hyperparameters randomly, this approach has some inconveniences such as the results are hard to reproduce, and like the grid search it needs a strong computational capacity.

3.2.5 ANN architectures

A neural network can be structured into different forms and shape depending on the purpose it was conceived for and the data it is processing and depending on its complexity and the method of processing the data. Each architecture has its strengths and weaknesses and can be combined with other architectures to optimize results. The choice of architecture is thus crucial, and it is determined mainly by the objective. Neural network architectures can be divided into 3 main families:

Feed-Forward ANN:

Feed-forward simply refers to the procedure for processing data by the neural network. In effect, feed-forwarded simply means that the data crosses the input network to the output without backtracking the information. Typically, in feed-forward neural nets we distinguish two types, single-layer networks, and multi-layer networks. In the single-layer networks the data processing takes place between the input layer and the output layer, which are all interconnected. Hence, the entire network has only one weight matrix. Having a single weight matrix limits the simple perceptron to be a linear classifier allowing to divide the set of information obtained into two distinct categories.

Unlike the single-layer perceptron, the multilayer perceptron has one or more so-called "hidden" layers between the input layer and the output layer. The task of these hidden units is to be part of the analysis of data flowing between the input and output layer. A multilayer perceptron is better suited to process the non-linear functions. Figure 3.10 illustrates a feed-forward MLP architecture with two hidden layers, where each neuron is interconnected with all the neurons in the adjacent layer.

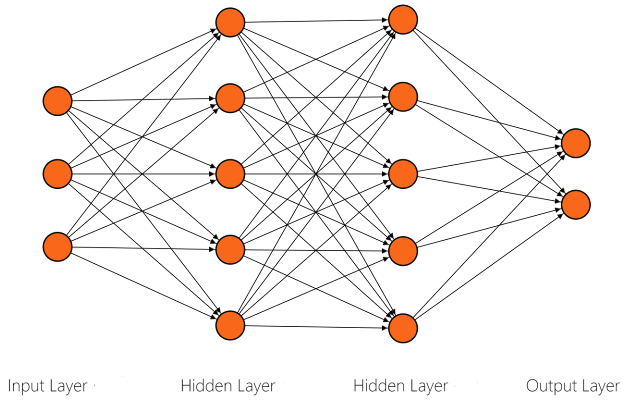


Figure 3.10: Feed-forward MLP

Recurrent Neural Networks (RNN):

Recurrent Neural Networks process information on cycles. These cycles allow the network to process information multiple times, sending it back each time within the network. The strength of RNNs lies in their ability to take into account contextual information following the recurrence of processing of the same information. This dynamic self-sustains the network.

RNNs consist of one or more layers. The Hopfield model is the most well-known single-layer RNN, it is a self-associative memory consists of a single layer that represents both the input to the network and its output. Multi-layered RNNs claim the particularity of having pairs (input / output) such as MLPs between which the data convey both forward and back propagation. Figure 12 illustrates the Hopfield and Elman networks.

Self-organizing neural networks:

Self-organizing neural networks are especially suited for processing spatial information. By unsupervised learning methods, self-organized neural networks are able to study the distribution of data in large spaces, such models are mainly used for clustering problems. The best-known model of this type of neural networks is Kohonen's self-organizing maps (Kohonen 1990) illustrated in figure 3.11.

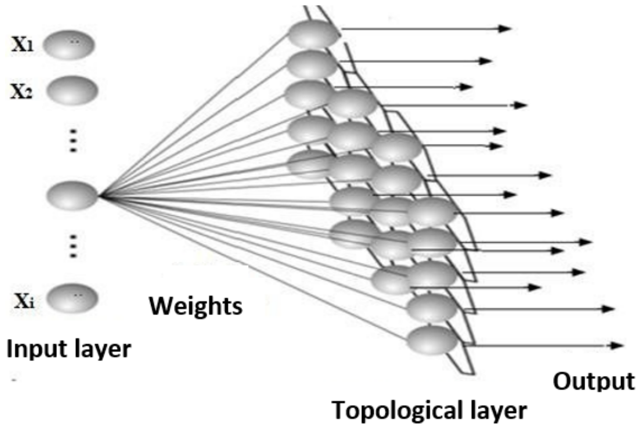


Figure 3.11: Kohonen's self-organizing map

3.2.6 Backpropagation

Backpropagation is the most used method for updating the neural networks weights. This algorithm updates the weights of the network by computing the error's gradient for each neuron of the network from the last layer to the first.

Backpropagation consists of two phases, during the first phase which is the forward propagation phase, each sample of the training set passes through the synaptic weights from one layer to the next, until the data finally comes out in the output neurons while adding the biases as follows:

$$a^{(1)} = x \tag{3.4}$$

$$z^{(2)} = w^{(1)}a^{(1)} \tag{3.5}$$

$$a^{(2)} = g(z^{(2)}) \tag{3.6}$$

Until:

$$z^{(L)} = w^{(L-1)} a^{(L-1)} \tag{3.7}$$

$$a^{(L)} = y = g(z^{(L)}) \tag{3.8}$$

Where:

- x is the input vector
- L is the number of layers in the network
- g is the neuron's activation function
- $a_i^{(j)}$ is the activation function of the neuron i in the layer
- the matrix of weights controlling the mapping of layer j to layer j + 1
- y is the output

The second phase of the backpropagation algorithm is the backward pass, in which we start by the output neurons and pass through all the layers in the network and compute the gradient δ recursively for each neuron in all layers. For each neuron the delta rule is defined as follows:

$$\Delta\varepsilon_j = \alpha(target_j - y_j)g'(z_j)x = \alpha\delta_j x \tag{3.9}$$

Where:

- α is the learning rate
- $target_j$ is the desired output

- z_j is the weighted sum of the neuron inputs

In the output layer, δ simply is the error multiplied by the first derivative of its activation function as in the following equation.

$$E = \sum \frac{1}{2}(target - y_j)^2 \quad (3.10)$$

To compute the weights changes of all connections to the output layer $(o_j = y)$ in the case of using a sigmoid activation function we use equation 3.9. Next, to compute the δ of the hidden layers we use the equation 3.11.

$$\delta_j = \frac{\partial E}{} \partial o_j \frac{\partial o_j}{} \partial Z_j = (\sum_{l \in V} \omega_{jl} \delta_l) o_j (1 - o_j) \quad (3.11)$$

Where V are the neurons receiving the output of the neuron j.

This recursive calculation continues layer by layer by propagating the changes to all the synaptic weights of the network. In summary, the back-propagation algorithm works by passing the data in a forward propagation and a backward propagation of the error.

3.2.7 Deep neural networks

Long short-term memory (LSTM): RNNs are sequence-based models, which are suitable for time series forecasting problems as they're able to learn the temporal dependence between the present and the past information. However, RNNs can suffer from exploding and vanishing gradient problems (Bengino et al. 1994) (Kolen and Kremer 2001), which lead to difficulties in learning long sequences. In order to overcome these limitations, (Hochreiter et al. 1997) introduced the long short-term memory (LSTM), a variant architecture of the traditional RNN which includes a memory cell, and later enhanced by (Gers et al.

2000) by including a forget gate. Therefore, a typical LSTM consists of a cell, an input gate, a forget gate, and an output gate, and its output can be a sequence of a variable length. Figure 3.12 illustrates the structure of an LSTM (Abbasimehr et al. 2020).

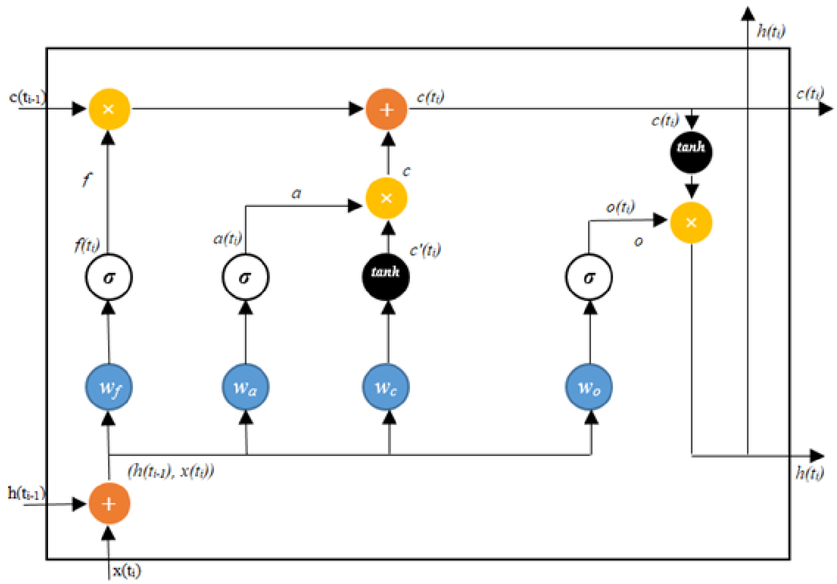


Figure 3.12: LSTM network

- $x(t_i)$ is the input vector
- $h(t_i), h(t_{i+1})$ are the output values at time i and i-1
- $\omega_c, \omega_a, \omega_f, \omega_o$ weight matrixes of the internal state, input, forget, and output gates.
- $\omega_{hc}, \omega_{ha}, \omega_{hf}, \omega_{ho}$ the recurrent weights
- b_c, b_a, b_f, b_o: the biases
- $c(t_i), a(t_i), f(t_i), o(t_i)$: The output values.

Considering the above notations, (Abbasimehr et al. 2020) described the LSTM functioning as follows:

$x(t_i)$ and $h(t_{i-1})$ are used as inputs by the forget gate $f(t_i)$ to calculate the information to be conserved in $c(t_{i-1})$ using a sigmoid activation, the input gate $a(t_i)$ uses the input $x(t_i)$ and $h(t_{i-1})$ to calculate $c(t_i)$, and the output gate $o(t_i)$ regulates the output of an LSTM cell by taking into consideration the cell state $c(t_i)$ and employing tanh and sigmoid layers. The following equations represent the LSTM's forward learning:

$$a(t_i) = \sigma(\omega_a x(t_i) + \omega_{ha} h(t_{t-1}) + b_a \tag{3.12}$$

$$f(t_i) = \sigma(\omega_f x(t_i) + \omega_{hf} h(t_{i-1}) + b_f \tag{3.13}$$

$$c(t_i) = f_t \times c(t_{i-1}) + a_t \times tanh(\omega_c x(t_i) + \omega_h c(h(t_{i-1}) + b_c) \tag{3.14}$$

$$o(t_i) = \sigma(\omega_o x(t_i) + \omega_{ho} h(t_{i-1} + b_o \tag{3.15}$$

$$h(t_i) = o(t_i) \times tanh(c(t_i)) \tag{3.16}$$

Where tanh and σ are activation functions and $\times$ represents the point-wise multiplication. In summary, an LSTM learning process passes by the following steps which will be repeated for a specific number of iterations:

- The forward pass that consists of calculation the output using the equations (3.12-3.16).
- Calculating the error of each layer between the input and the output values.
- Propagate the error backwardly to input gate, cell, and forget gate.
- Using an optimization algorithm to update the weights based on the error.

Convolutional Neural Networks (CNN): CNNs are a special feed-forward networks, they take their name from a mathematical operation called convolution which involves the convolution of different functions. The first part of these networks works as a feature extractor, an image is passed through a succession of filters, creating new images called convolution maps. Some intermediate filters reduce the resolution of these maps by an operation called pooling. Finally, the convolution maps are flattened and concatenated into a feature vector. This feature vector at the output of the convolutional part is then plugged into the input of a second part, made up of fully connected layers (multilayer perceptron). The role of this part is to combine the characteristics of the CNN code to classify the image.

CNNs are best suited for image recognition, video analysis, and natural language processing. Figure 3.13 illustrates the structure of a CNN.

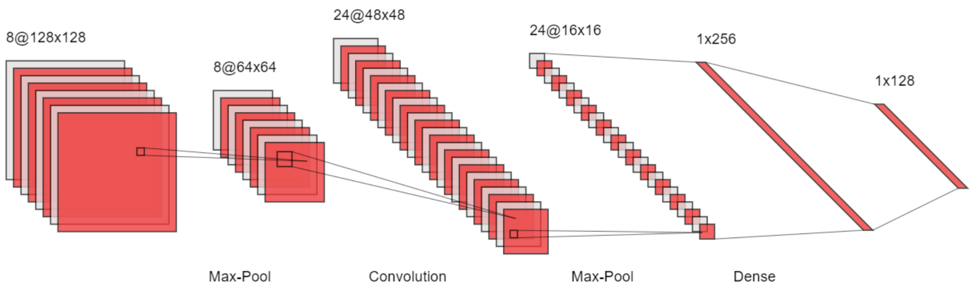


Figure 3.13: The structure of CNN

Autoencoders: The autoencoders are a special type of feed-forward networks, its general aim is to reproduce its inputs as outputs, autoencoder consists of an input layer, a hidden layer, and an output layer that has the same size as the input layer. These networks are trained to compare the produced output of the neurons with their input.

As illustrated in figure 3.14, we refer to the part of the model which calculates the hidden layer h(x) by the encoder as it allows to encode the input, while we refer to the part which calculates the output layer by the decoder. Therefore, the autoencoder firstly encodes the input in a hidden layer and then decodes it to get a reconstruction of that input. Autoencoders are mostly used for anomaly detection and dimensionality reduction.

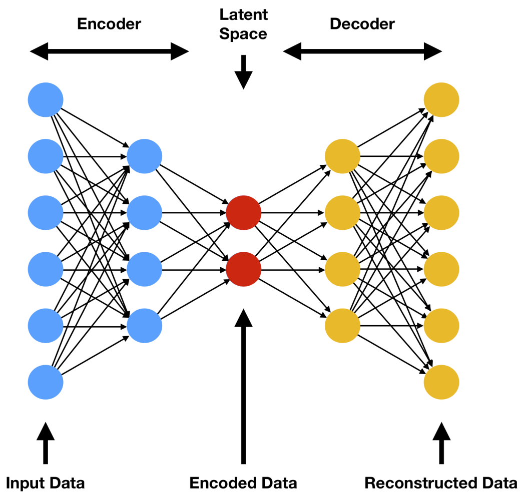


Figure 3.14: Autoencoder architecture (Farfar 2018)

Deep belief networks A Deep belief network (DBN) is deep learning algorithm, i.e., a class of deep neural networks, it's composed of multiple interconnected hidden layers. DBNs can be trained with an unsupervised approach to learn to reproduce its inputs probabilistically, the layers then perform feature extraction, later, DBNs can be further trained with supervised learning approach to perform pattern recognition. DBN can simply be viewed as a composition of multiple unsupervised networks such as autoencoders and restricted Boltzman machines, which give it the possibility to be trained greedily one layer at a time in order to increase its effectiveness.

3.3 ANN for forecasting energy consumption

Forecasting problems consist of predicting future values based on past and present data using dedicated methods, paradigms, and approaches. One of the most commonly used paradigms in machine learning is ANNs.

Using ANNs for forecasting energy consumption and production has been the subject of multiple recent research papers. (Laib et al. 2016) employed multiple MLPs to forecast the Algerian yearly natural gas consumption in different distribution areas. Each MLP was designed specifically to forecast consumption in a specific area, before summing the results of the neural network in order to obtain the total gas consumption.

(Hsu et al. 2018) introduced a methodology based on a two-phase ANN for short term load forecasting. The first stage consisted of forecasting the 24-hour load pattern, while the second phase forecasted the maximum and the minimum loads.

(Jetcheva et al. 2014) proposed a building level neural network-based model for next day electricity load prediction. The method used multiple MLPs where each MLP was trained on a different subset of the data.

Several researches work looked at renewable energy forecasting in the short, medium, and the long-term using the ANNs. For instance, (Wu et al. 2016) used a deep neural network model that consisted of an LSTM recurrent network and a Convolutional neural network for short-term wind power forecasting.

(Peng et al. 2016) used a Multilayer Restricted Boltzmann Machine (MRBM), which is a deep learning neural network with strong feature interpretation ability, to forecast the next 4 hours of wind power production.

(Bhaskar and Singh 2012) proposed a method to forecast wind power that consists of two phases. Phase one forecasts wind speed for the following 30 hours using an adaptive wavelet neural network (AWNN). Phase two used a MLP to map the predicted wind speed into wind power. More works are cited in a survey (Marugan et al. 2018).

(Zhong et al. 2018) used correlation coefficient factors to analyze the correlation between weather factors and solar energy generation. They used a general regression method and backpropagation neural network to predict the generated power.

(Abuella and Chowdhury 2015) used a MLP to forecast solar energy generation for the month ahead in hourly steps using a dataset that consists of 14 weather variables. They then compared the performance of the MLP to multiple linear regression and persistence models.

Table 3.1: Examples of ANN energy forecasting research

Prediction Goal	Used data	Models	Results	Reference
Natural gas	Meteorological and historical data	MLP	MAPE: 0.75	Laib et al. 2016
Electric load	Historical load and temperature	MLP	MAPE: 0.91	Hsu et al. 2018
	Historical load	MLP	MAPE: 2	Jetcheva et al. 2014
Wind power	Historical load and wind speed	LSTM-CNN	MAE: 5.62	Wu et al. 2016
	Historical load and wind speed	MRBM	MAE: 0.12	Peng et al. 2016
	Historical load and wind speed	AWNN-MLP	MAE: 1.92	Bhaskar and Singh 2012
Solar power	Meteorological and historical data	MLP	N/A	Zhong et al. 2018
	Meteorological data	MLP	RMSE: 0.07	Abuella and Chowdhury 2015

Table 3.1 above summarizes the above works, the results highlight the best mentioned results in the papers using some performance evaluation metrics such as root-mean-square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE).

3.4 Conclusion

In this chapter the main concepts of machine learning and supervised learning are presented, putting the focus on ANN theory and its different types and architectures including both shallow and deep architectures.

ANN are algorithms inspired and modeled based on the structure of the biological brain, they consist of several interconnected artificial neurons in a way that this architecture can be used in various fields to solve computational-based problems which cannot be explicitly programmed. ANN was and still one of the most used paradigms in time series forecasting and proved effective in the area of forecasting energy consumption and production where we presented some examples of the state-of-the-art research that were conducted in that field.

The computer was born to solve problems that did not exist before

Bill Gates

4

Multi-Agent Systems

▷ *In this chapter, we will try to give a definition to the main concepts in the domain of multi-agent systems …* ◁

Contents

4.1 Introduction

In this chapter, we will try to give a definition of the main concepts within the domain of multi-agent systems. The notion of "agent" aspect is described by presenting the related characteristics to this basic unit constituting the MAS. Then, we move on to the collective dimension of the MAS by presenting its main characteristics. Approaches that fall into the same category of MAS are also presented along the advantages of MAS motivating their usage. The chapter ends by presenting a review of MAS models used for the simulation of pollution related problems.

4.2 The agent concepts

Despite the maturity of Multi-Agent Systems there is, so far, no consensus on the exact definition of the term agent. For instance:

(Russell and Norvig 1995) give a broad definition which is based on the perception and actions that an agent has on its environment. They define an agent as any entity that perceives its environment via sensors and which acts on this environment via effector components. Consequently, and according to this description, several entities of the real world can according to these arguments be considered as being agents. For example, an automatic gate with a sensor to detect people wishing to enter a building, opening the door each time it is activated, may be considered as an agent.

(Wooldridge and Jennings 1995) defines an agent as a computer system (hardware or software) situated in an environment that can autonomously do actions to achieve the goals of its design. This definition demonstrates clearly the fact that autonomy is among the main properties of an agent. In fact, autonomy is a property that distinguishes agents from other concepts, such as objects.

According to (Ferber and Perrot 1995) ”an agent is an autonomous entity, real or abstract, which is capable of acting on itself and on its environment and which, in a multi-agent universe, can communicate with other agents, and which behavior is the consequence of its observations, knowledge and interactions with other agents”. He also defines it as being ”a physical or virtual entity evolving in an environment of which it has only a partial representation and on which it can act. An agent is able to communicate with other agents and is endowed with autonomous behavior. An agent generally has acquaintances with all the agents with whom he communicates or interacts”.

(Shoham 1993) defines an agent as an entity that functions in continuity and autonomy in an environment in which other processes and agents exist.

(Nwana 1996) defines an agent as a piece of hardware or software that is able to accurately act in order to perform tasks on behalf of its user. Table 4.1 summarizes the properties of the agent based on the different definitions

Table 4.1: Summary of the properties used to define an agent

Definition	Autonomy	Coexistence with other entities	Environment	Perception	Action	Objectif
Russell and Norvig	No	No	Yes	Yes	Yes	No
Wooldridge and Jennings	Yes	No	Yes	No	Yes	Yes
Ferber	No	No	Yes	Yes	Yes	Yes
Shoham	Yes	Processes and agents	Yes	No	No	No
Nwana	No	User	No	No	Yes	Yes

4.3 Agents' typology

Agents are characterized by their architecture and their behavior which remains linked to the point of view of the designer and the way of assembling the different parts of the agent so that the implemented agents can accomplish the goal of their conception. Therefore, the agents are classified into different types based on their architectures and capabilities which qualify them as cognitive, reactive or hybrid agents.

4.3.1 Reactive agents

The category of reactive agents, described by R.A. Brooks (Brooks 1991), includes all agents whose representation of the environment is only sub-symbolic. This means that the representation comes only from the perceptions of the agent, of the "visible" world at the current moment. There is no reasoning strictly speaking in this category of agents since we are in a configuration of the stimulus / action type: stimulus $\rightarrow$ percept $\rightarrow$ action. Thus, as soon as they perceive a change in their environment, they respond with a programmed action. Their quick and unthinking actions are similar to reflexes. These agents react to events in the environment but has no memory and can therefore neither take the past into account, nor foresee beyond the short term.

Reactive agent despite having very limited reasoning skills, they allow the construction of systems composed of many small agents, which are automatons. The interactions of agents with each other allow the emergence of structures from a higher abstraction layer that are potentially observable (Lestel et al. 1994). The main characteristics of reactive agents can be summarized in following points:

- The decision-making of an agent is carried out through a set of modules behavioral corresponding to the task to be performed. Each behavior is implemented as the following rule: Situation $\rightarrow$ action

4.3.2 Cognitive agents

Agents with cognitive abilities come from a metaphor of the human model, unlike reactive agents, the class of cognitive agents includes all agents having an explicit representation of their environment. Cognitive agents are able to reason on their explicit representation, they have a knowledge base which contains the various information related to their areas of expertise and the management of interactions with other agents and their environment. In addition, interaction allows agents to communicate, collaborate and take action. As a result, they are able to make decisions based on the information they have and plan their actions in advance.

There are some characteristics that an agent must have in order to be qualified as a cognitive agent, however an agent does not have to have all these characteristics at the same time, but it is enough that he has some of them in order to be qualified as a cognitive or an intelligent agent, (Nachet 2014) described the following characteristics:

- **Autonomy:** the agent must be able to take initiatives and act on his own and control its own actions as well as its internal state. Autonomy is what distinguishes an agent from another type of software.
- **Situation:** the agent is able to act on his environment from the sensory inputs he receives from this same environment.

- **Interactivity:** The agent must be able to exert actions on his environment and vice versa.
- **Reactivity:** the agent must be able to perceive his environment and develop a reaction in the required time.
- **Proactivity:** a proactive system generates and satisfies goals; he should be able to take initiatives in an appropriate time and not be only driven by the events as he should have an opportunistic behavior.
- **Learning ability:** an agent will have the ability to learn if he knows how to acquire new knowledge, information or habits.
- **Social capacity:** agents must cooperate with each other to accomplish their tasks; therefore, they should have social capacities which give them the possibility to play roles and to interact within the context of these roles.
- **Competition:** the agent is able to compete withing an environment with other agents to reach a common goal in way where only one will succeed while the other agents fail.
- **Mobility:** There are two types of agent mobility: relative mobility and real mobility. In the case of the first type, there is no real displacement of the agent. This initiates a succession of requests to the various servers, while in in the second type, the agent process moves on the network from one server to another, therefore the agent's code is transported and so is its data and it continues to run on the new machine.
- **Delegation:** the agent performs a set of tasks at the request of a user or another agent.

- **Communication:** the agent has the ability to interact either with a user through a user interface or with another agents.

4.3.3 BDI Agents

BDI is an acronym for Beliefs, Desires and Intentions; this type of agents is particular type of cognitive (Bratman 1987), their architecture is built around practical reasoning. These agents are generally represented by a mental state having the following mental attitudes:

- **Beliefs:** Beliefs represent the agent's information about its environment, i.e., its beliefs about the environment and other agents. Using the term belief indicates that what the agent belief may not always be true.
- **Desires:** Desires represent situation and objectives that the agent wants to achieve, hence, desires are the motivational state of the agent.
- **Intentions:** Intentions are the deliberative state of the agent, in another words, intentions are desires that the agent has committed to and chosen to execute.

A BDI agent must therefore update his beliefs with the information that comes from his environment, decide what options are offered to him, filter these options in order to determine new intentions and take his actions in view of his intentions.

In this architecture, the process is divided into two phases: The first phase is called the deliberation phase, where a number of goals are set, while the second phase corresponds to a planning phase and it consists of defining how to achieve these goals. Figure 4.1 illustrates the architecture of a BDI agent.

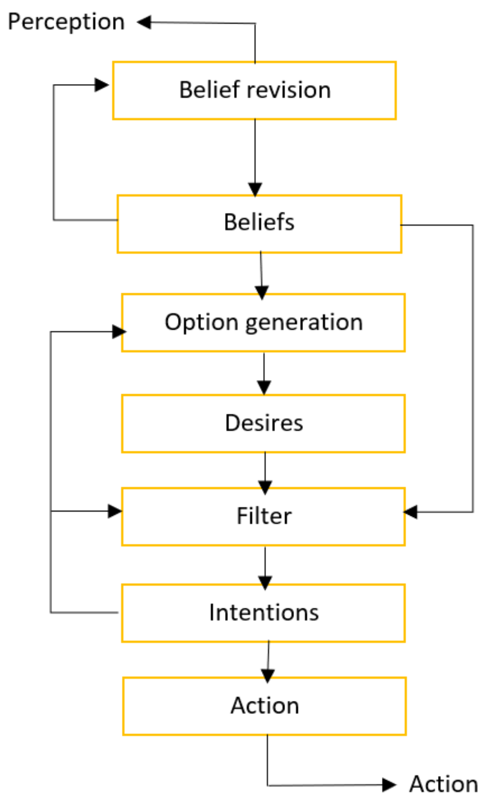


Figure 4.1: BDI agent architecture (Wooldridge 1999)

4.3.4 Hybrid agents

By the early 1990s, it was well known that both reactive and cognitive systems work well for some types of problems but might be less effective for other types Each of these two types of systems is suitable for certain types of problems and less well for others. Therefore, researchers begin to investigate the possibility of combining the two approaches in order to obtain a hybrid architecture (Jarras and chaib-draa 2002).

4.4 The environment

The environment is the universe in which the agent evolves and performs tasks. The environment representation depends on the nature of the modelled system. In some cases, an abstract representation is used. It consists of a Euclidean geometric space with two or three dimensions. The position of each agent is therefore saved among its attributes, for example it can be a geometric coordinate. The notion of proximity in this case is the distance between the agents. In short, defining the environment is a key for defining the problem, which is essential before any attempt at problem solving.

According to (Russell and Norvig 2006) it is possible to categorize environments based on certain properties. These properties play a role in the type of agent that can be designed for certain sets of environments. The properties are as follow:

Partially observable / Fully observable:
On the one hand, if the agent can observe all environment states and be able to use all its sensors, then his environment is fully observable. On the other hand, the environment is partially observable if the agent can only observe his environment within a certain perimeter around it.

Stochastic / determinist:
When the next environment state is determined by the action of an agent and its current state within the environment, the latter is then called deterministic. In cases where we make an exception to the actions of other agents, we say that it is strategic. In all the other cases, we say that it is stochastic, and it compromises the notion of randomness.

Episodic / sequential:

In an episodic environment, the agent's behaviour is divided into atomic episodes. Each episode is associated with a specific action, and all actions are independent, and have no influence on each other. In a sequential environment, next actions depend on the actions previously carried out.

Discreet / continuous:

An environment is discrete when the states of this environment are distinct, even if there is an infinity of them. The way of managing time or of looking at the perception and actions of an agent also determines whether the universe is discrete or continuous.

Static / Dynamic:

If the environment does not change while the agent is thinking or acting, it is a static environment. On the other hand, if the environment can change during the phase of reflection or action of the agent, it is an environment dynamic. It is also possible to define a semi-dynamic environment, for instance in the cases when the environment does not change over time but the performance of an agent is time dependent.

Mono-agent / multi-agent:

The environment is mono-agent when a single agent evolves in it, and it is considered as a multi-agent environment when it includes at least two agents. This environment property is important in the case where it is multi-agent because it will be necessary to specify whether the multi-agent environment is competitive or if it is cooperative. Table 5.2 presents different types of agents along with their environment proprieties.

Table 4.2: Some agents and their environment properties (Grouls 2013)

Environment	Observability	Determinist	Episodic	Static	Discreet	Agents
Chess game	Fully	Strategic	Sequential	Static	Discreet	Multi
Chronometric	chess game	Fully	Strategic	Sequential	Semi-Dynamic	Multi
Pizza deliverer	Partially	Stochastic	Sequential	Dynamic	Continuous	Multi
Cross words	Fully	Determinist	Sequential	Static	Discreet	Mono
Online purchasing advisor	Partially	Stochastic	Episodic	Static	Discreet	Mono

4.5 Multi-agent systems

After giving some definitions to the agent concept and presenting its properties and characteristics, in this section we will focus on the MAS concepts, we will therefore, present the definition of a multi-agent system as well as its components and properties.

4.5.1 Definition

Similarly, to the agent concept, MAS can be defined in different manners; On the one hand, (Macal and North 2010) define a MAS as a set of agents with their behaviours and attributes, a set of relationships between agents, and an environment. The agents interact with their environment as well as with each other. (Jennings 2000) defines a MAS as a set of agents situated in environment as illustrated in figure 4.2.

On the other hand, (Ferber and Perrot 1995) gave a precise definition of MAS: A multi-agent system is a system that has the following components:

- An environment E, which is a space with a metric
- A set of objects O located in E
- A set of Agents A, which are themselves active objects (A included in O).
- A set of relations R, which unites the objects between them.
- A set of operations OP allowing agents of A to perceive, produce and manipulate the objects of O.
- Operators responsible for representing the application, these operations and the reaction of the environment on this attempt at modification called laws of the universe.

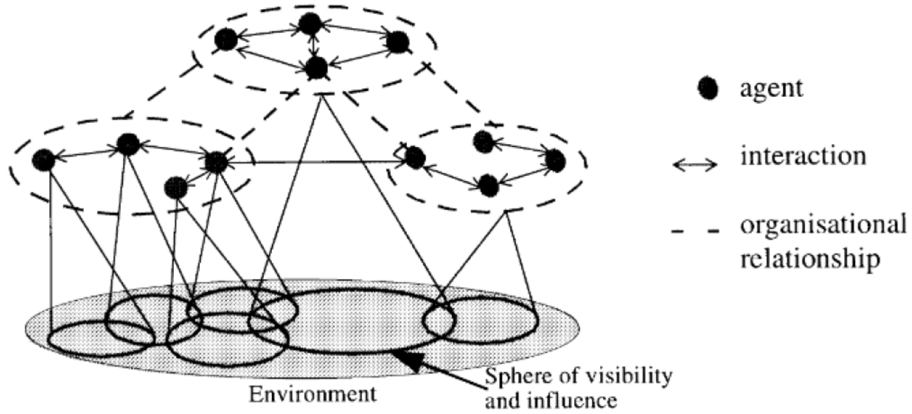


Figure 4.2: MAS representation (Jennings 2000)

Therefore, we can say in general that a MAS consists of multiple entities or individuals called agents, and that these agents are located in an environment and interact with it and interact with other agents. Self-organizing phenomena appears within these systems, allowing the adaptation of a MAS to its environment based on the objectives of these systems.

4.5.2 Interaction

Interaction is a fundamental notion in MAS, and this notion is the cause of many problems as well as many solutions (Ferber and Perrot 1995). The interaction is the reason that the MAS is seen as a whole and not as a set of independent entities. Agents have interaction mechanisms to enable them to accomplish their goals.

An interaction must follow a precise process in order to proceed correctly. Therefore, the interactions are usually structured according to typical patterns called protocols. Interaction protocols allow agents to exchange structured messages and control the exchange of these messages and thus facilitate their coordination.

An interaction protocol specifies rules that must be respected by agents during a conversation, and thus defines for each step the types of messages that can be sent. By following a protocol, an agent interprets the messages in a conversation and changes its own state at each step, along with using the protocol to produce the next message in the conversation.

We generally distinguish different types of interactions that can occur between agents, namely: communication, cooperation, negotiation and coordination.

4.5.2.1 Communication

Communication is one of the key aspects of the multi-agent systems. It is the basic building block of any interaction despite that it can be optional in some cases. Communication is a vector of interaction using a dedicated language that allows the transmission and reception of information or representations. These representations can be representations of the environment, or of a mental state for example. Agents can interact by communicating directly with each other by sending messages or even by establishing structured conversations between them, or by acting on their environment.

A. Communication languages

Inter-agent communication requires mutual understanding of agents among themselves, Hence, a common language for all agents is a necessity. There are many and different languages allowing to formalize and describe the exchanges between the agents.

(Ferber and Perrot 1995) divided the communication languages into five classes based on the aspect of communication that being treated, whether a high-level communication (languages allowing abstraction) or low level (languages allowing realization or implementation) as illustrated in figure 4.3.

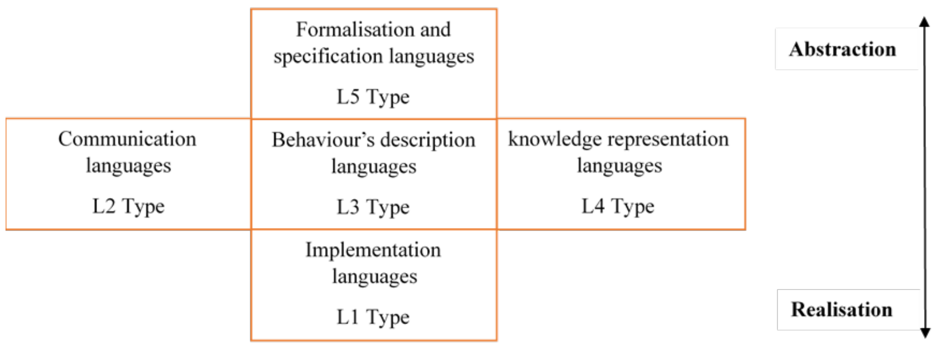


Figure 4.3: MAS conception languages (Ferber 1995)

L1 Type: The L1 type represents the implementation languages used for programming a MAS. The structures needed to implement the environment and the agents such as parallelism mechanisms, the logic of agents, communication protocols . . . etc. Examples of common L1 languages are Java, C++, C♯ . . . etc, must be identified and mastered.

L2 type: L2 represents communication languages between agents. These languages allow agents to interact with each other by means of the transmission of information, knowledge or services. The two most popular multi-agent communication languages widespread are KQML (Knowledge Query and Manipulating Language) (Fini 1995) and FIPA-ACL (Agent Communication Language of FIPA) (FIPA 1999).

L3 type: L3 represents the description languages of behaviours and laws of the environment. This type of language is more abstract and uses a certain formalism so as not to hinder the implementation of behaviours or the implementation of the physical laws of the environment.

L4 type: This type represents the languages of knowledge representation, it is mainly for the aim to implement cognitive or hybrid agents as these agents are required to be able to manipulate the representations that they make to themselves or are transmitted from their environment and mental states.

L5 type: This type is the most abstracted languages type, and represents the formalization and specification languages of MAS. This type of language makes enable the developers to formalize the MAS that they want to develop, and to formalize all the notions that they will need, such as interactions, objectives, as well as development constraints.

B. Agent programming languages

There are no rules regarding programming languages when it comes to the implementation of agents or the implementation of MAS. However, there are some languages that are more suitable, and also languages that have been designed particularly for designing such systems. In this section some of the most common agent languages are presented..

KQML: KQML (Knowledge Query and Manipulating Language) (Fini 1995) is a language that supports inter-agent communication. It is a language based on speech acts that allows to standardize the messages exchanged between agents. KQML message types are for example assertions ('ask', 'tell'), information routing instructions ('forward' and 'broadcast'), persistent commands ('subscribe', 'monitor'), commands that allow consumer agents to ask intermediary agents to find the relevant supplier agents ('dvertise', 'recommend', 'recruit' and 'broker').

ACL-FIPA: FIPA-ACL (Agent Communication Language of FIPA) (FIPA 1999) is an L2-type agent communication language in addition to being a standard for agent communication languages in general. it was developed by FIPA (Foundation for Intelligent Physical Agents), an organization which is responsible of the standardization of communications between agents. FIPA-ACL like KQML is also based on speech act theory and has benefited greatly from KQML research results. It was developed to overcome the limitations and criticisms made to KQML. The specificities of this language are the standardization of language actions (confirm, request, inform, cancel, propose, etc.), the structuring of messages (syntax, protocol, spoken language, etc.) and the management of message transport. (sender, recipient, message size, date, etc.).

AgentO: AgentO is a multi-agent programming language created by Yoav Shoham and presented in (Shoham 1991) that uses the paradigm of agent-oriented programming. Implementations in this language run on a system using LISP. The agents implemented with this language are agents with a great communication capacity and having two types of representation: beliefs and intentions. Agents implemented using AgentO can act in six different ways: do, inform, ask, cancel, repeat, if true then do. Despite being an old language, AgentO laid the foundations for the agent-oriented programming paradigm.

BRIC: BRIC (Block-like Representation of Interactive Components) is a high-level language allowing to design and implement multi-agent systems using a modular approach. It was created by J. Ferber (Ferber 1995). A BRIC system consists of a set of components linked together by communication links; therefore, The BRIC formalism is componential approach in which are used L3 type languages. A BRIC component is defined as a structure that has inputs and outputs, and therefore the interior contains other components such as class and object instances. For BRIC components which are actually made up of several BRIC components, they are said to be structured components and has a specified behaviour (Grouls 2013).

4.5.2.2 Cooperation

Cooperation is the general form of interaction; it is necessary when a goal cannot be achieved by an individual agent and therefore the collaboration of multiple agents becomes necessary. It represents the form of interaction which is concerned with the way of distributing the work, and therefore, allocation of tasks between several agents.

Cooperation can be either statistic by distributing the tasks on the agents when implementing the system, or it can be dynamic by implementing using a centralized agent or a coordinator who distribute the tasks on the other agent based on a demand and offer mechanism. Cooperation can be viewed as an intentional attitude where the agents engage in a collective activity after having identified and adopted a common goal, however, it can also result from an automatic activity of a set of reactive agents (Nachet 2014).

4.5.2.3 Coordination

Coordination is the basis of cooperation, it relates to the coordination of actions, the sharing of resources and the parallelization of actions. Coordination is necessary to improve and maintain consistency in the overall functioning of the system, as in the case of agents' cooperation, where they can share the same resources in order to achieve a goal, therefore they must accomplish the tasks related to the problem to be solved and coordinate their actions. Coordination allows the system to solve problems collectively, save execution time, avoid conflicts between agents and reduce unnecessary which increases system performance (Nachet 2014).

4.5.2.4 Negotiation

Negotiation represents the resolution strategy that is used to reach agreements between agents to resolve conflicting beliefs or goals, this strategy relies on dialogues between agents. These conflicts of beliefs happen due to contradictions between the beliefs of different agents, as the agents may have incomplete or erroneous knowledge, therefore, negotiation is a process of communication between a group of agents in order to reach a mutually accepted agreement and to resolve their conflict by finding a compromise either by sharing resources or by coordinating their actions.

4.5.3 The emergence phenomenon

(Epstein and Axtell 1999) defines the concept of emergence as being the organization of behaviours in the form of recognizable macroscopic patterns, where these patterns can appear during the simulation according to a law, a condition or according to a periodicity. Knowing these conditions and laws is the fundamental goal of modelling and simulation.

This concept is often linked to MAS with reactive agents. It is generally considered to be a phenomenon that characterizes the overall behaviour of a system resulting from local actions and interactions between agents of the same system. Therefore, the MAS allows to reveal so-called emerging behaviours which appear without having been explicitly programmed. This kind of phenomenon emerges from behaviours that take place in micro level. Hence, MAS are considered to be a bottom-up approach, which means to model individuals and their interactions at the low level (micro) in order to see the results at the high level (macro) (Ghazi 2017).

4.5.4 MAS advantages and limitations

MAS disposes of several advantages, (Bonabeau 2002) presents a list of them. The main advantage being their ability to capture emerging phenomena that they offer as described in section 5.3, also, MAS provides a natural description of a system as it is suitable to naturally describe a system composed of behavioural entities by making the model seem close to reality, for example in the cases where we try to describe a traffic jam, voters, or stock market, . . . etc. In summary, MAS approach is convenient when the aim is to describe a system from the perspective of its components' behaviours, i.e., when:

- The individual behaviour is complex, and cannot be modelled using differential equations as its complexity would increase exponentially as the behaviour's complexity increases.

- Describing the system using behaviours presents a more natural way than describing it using processes.
- MAS generally are the most suitable approach to describe and model what's really happening in the real world and enables the experts to calibrate and validate the model which is a crucial phase.
- With MAS it is possible to apply sources of randomness to add stochasticity to agents' behaviours.
- It's not possible to clearly define the behaviour of entities through aggregate transition rates.

Another main advantage of the MAS approach is its flexibility, which can be observed through different aspects, for instance, it's easy to change the number of agents in the model, either by adding more agents or reducing them, it's also possible to change the complexity of the agents by tuning their behaviour, its degree of rationality, change the rules of interactions and their ability to evolve and learn.

Despite the maturity of the MAS approach and its increasing popularity, it still suffers from some limitations and inconveniences, these limitations are classified into three categories, the difficulty of modelling in MAS in contrast to its utility, where the utility and the gain of developed systems are questionable based on the necessary effort and costs for building such models as they suffer from generalization problems, which means that a model designed to represent a phenomenon cannot be reused to model another phenomenon or problem.

Another issue in MAS is the availability of detailed data, as the MAS models requires large amount of detailed data in order for the model to be validated and calibrated as a model designed with a less meaningful dataset will produce results that are far from reality and its utility is up for debate.

The third limitation that MAS approach suffer from is the reproducibility, MAS models are generally very difficult to reproduce, which is due to the sensitivity of these models to initial conditions and the lack of a general framework under which models can be described and communicated (Ghazi 2017).

4.6 MAS for pollution modelling and forecasting (State of the art)

MAS models of environmental pollution are valuable tools. They can help decision-makers put in place environmental management policies in order to preserve the ecosystem and ensure public health, this approach has emerged as a promising approach for modelling environmental problems and specifically those related to pollution (Aulinas et al. 2009). In this section most recent research using MAS approaches to solve environmental and pollution related issues are presented.

(Ghazi et al. 2016) used a MAS system to simulate the control of an air pollution crisis. Their system used a Gaussian Plum Dispersion (GPD) Model combined with an ANN.

In order to control urban air pollution from the activity of power plants, (Dragomir and Oprea 2013) used an agent-based approach applied to the city of Ploiesti in Romania.

(El Fazziki et al. 2017) adopted an agent-based method to model the urban road network infrastructure to control road air quality. The approach used an

ANN to forecast air quality and the Dijkstra algorithm to search for the least polluted path in the road network.

(Corchado et al. 2009) combined a MAS and a case-based reasoning system to detect oil slicks and forecast their evolution and trajectory using meteorological parameters and satellite images.

(Di Lecce et al. 2009) proposed a MAS that uses ANN, fuzzy logic, and sensors to forecast air quality. The system was used to monitor air pollution concentration, caused mainly by road traffic, near a hospital.

(Papaleonidas and Iliadis 2012) proposed a MAS to monitor air pollution in urban centers in Athena (Greece). The system is composed of multiple software agents, ANNs, Fuzzy Rule-based subsystems, and uses reinforcement learning (RL) to improve its learning ability.

(Ahat et al. 2009) employed a MAS to model air pollution in an urban area. Here, a two-dimensional grid representation of the environment was adopted to find the dispersal of air pollution on the grid.

(El Fazziki et al. 2015) proposed an air quality monitoring system for the city of Marrakech (Morocco) based on agent technology and big data. The system uses ANN as a forecasting algorithm and K-means clustering to propose a solution applicable to large scale data. (Hulsmann et al. 2014) presented an approach that combines a multi agent-based transport model (MATSim) with an Operational Street Pollution Model (OSPM) in order to calculate and analyse traffic-related air pollution in Munich, Germany.

(Gao et al. 2018) used a MAS simulation to analyse the correlation between cities in the same river basin assesses the relationship between urban economy and pollution transfer. In this system every agent represents a city, and Each urban behavioural pattern is analysed by considering three aspects for all the agents: network size, network concentration and link intensity. For the implementation they used the Netlogo platform.

(Agarwal and Kickhofer 2015) proposed an agent-based framework to simulate a real-word scenario of the Munich metropolitan area in Germany, this scenario investigates an optimization approach for traffic emissions and congestion, and they used the MATsim framework for implementation.

(Ni and Zhang 2009) developed an intelligent decision support system to control lake water pollution, it consists of an integrated system for acquiring information and processing the lake water environment, in addition to a lake water environment MAS based model, and several decision-making algorithms. The agents in this work represents different entities such as a water plant agent, family agent, company agent, and government agent.

(Li et al. 2012) proposed a simulation and forecasting system of water pollution spread, the proposed system is based on MAS modelling that consists of multiple agents, grid technology, and 3S technique that consists of GIS, RS, and GPS.

(Emery et al. 2020) presented a MAS that aims to simulate a road traffic at the scale of a theoretical and controlled road network.

(Hassan et al. 2022) proposed an air quality index prediction model, the approach is based on MAS architecture and incorporates an ARIMA model.

Table 4.3 summarizes the above described papers, the comparison between the works is not too evident or straightforward, as each system deals with a different phenomenon and uses a different decision-making mechanism. We have tried to analyse the studies according to the elements described below. It should be noted that in some articles, these elements are not always detailed enough and, in some cases, completely ignored.

Table 4.3: Summary of MAS' pollution modelling papers

Prediction Goal	Used data	Methods	Platforms	Reference
Air pollution crisis	Meteorological and historical data	MAS, ANN, GPD	Jade, Encog	(Ghazi et al. 2016)
Power plants related air pollution	Historical data	MAS	Zeus	(Dragomir and Oprea 2013)
Oil slicks detection	Meteorological and sea related data	MAS, CBR, PCA	N/A	(Corchado et al. 2009)
Indoor air quality	Real time pollutants data	MAS, ANN, Fuzzy logic	N/A	(Di Lecce et al. 2009)
Air and water pollution	economic and historical data	MAS	Netlogo	(Gao et al. 2018)
Water pollution	Historical data	MAS	N/A	(Ni and Zhang 2009)
	Historical data	MAS, GIS	N/A	(Li et al. 2012)
	Historical data	MAS, ANN, Fuzzy rule, RL	Java, MySQL	(Papaleonidas and Iliadis 2012)
Air quality index	N/A	MAS, Heat Bugs	Repast	(Ahat et al. 2009)
	Meteorological and historical data	MAS, K-Means, ANN	Prometheus	(El Fazziki et al. 2015)
	Historical data	MAS, ARIMA	Jade	(Hassan et al. 2022)
Road traffic related air pollution	Meteorological and historical data	MAS, ANN, Dijkstra algorithm	Prometheus	(El Fazziki et al. 2017)
	Meteorological and historical data	MAS, OSPM	MATSim	(Hulsmann et al. 2014)
	Meteorological and historical data	MAS	MATSim	(Agarwal and Kickhofer 2015)
	historical data	MAS	N/A	(Emery et al. 2020)

Prediction goal: This represents the objectives of the simulation, these are the problem we seek to solve using the simulation, and depending on the simulation objectives, we can define the variables we are trying to predict or estimate, for example: water pollution, air quality index, . . . etc.

Used data: Most systems use real empirical data collected from: surveillance networks, surveys or existing databases, which gives more credibility to simulation results.

Used methods: It represents the approaches used in each research, including the used decision-making mechanism and algorithms such as ANN, fuzzy logic, CBR, . . . etc.

Implementation platforms: There are several platforms and technologies that can be used to develop simulation systems in general, and MAS specifically. It should be noted that among the systems described in these studies, there are those which have not reached the implementation stage or the implementation details were omitted.

4.7 Conclusion

Modelling means representing a real phenomenon using an abstract model that is easy to manipulate and study. There are several approaches to do this modelling, namely MAS which was the subject of this chapter where we have presented this approach that continues to be applied in many areas. MAS allows to design models with a bottom-up approach, i.e., from the inter-entity interaction level, and then see how the overall behaviour of the system changes. In this chapter we have tried, on the one hand, to explain all the main concepts that are related to MAS. On the other hand, the advantages and the limitations that this approach suffer from are described. The chapter is concluded by presenting the recent state-of-the-art research where MAS approach were used to solve environmental and pollution related issues.

The question isn't, 'What do we want to know about people?', It's, 'What do people want to tell about themselves?'

Mark Zuckerberg

5
Energy modelling and forecasting

▷ *In this chapter we will present a forecasting energy demand, and production case study, starting from the available energy datasets of different energy sources in Algeria and how they are used to construct ANN forecasting models for short-term forecasting of demand and production.* ◁

Contents

5.1 Introduction

ENergy systems are critical and complex infrastructures, starting from the generation and transmission of energy to the distribution circuits. the main concern of these systems is to supply customers with reliable and good quality energy at all times in its different forms. Due to some challenges that the energy sector faces such as the lack of energy storage capacity for instance, the energy companies must forecast the load consumption to ensure a good load balancing between supply and demand instantaneously at any time. Hence, forecasting energy demand and consumption is considered a fundamental process in many energy and power generation and management decisions. It is a tool to help anticipate and determine future demand, a key factor in a process of controlling energy availability and minimizing operating costs. Having an idea about the energy demand, the energy production means using forecasting models, permits the forecast of the amount of CO_2 emission contracted while answering the demand.

Forecasting problems consist of predicting future values based on past and present data using dedicated methods, paradigms, and approaches. One of the most commonly used paradigms in machine learning is artificial neural networks (ANN) which was the used paradigm in this thesis.

In this chapter we will present a forecasting energy demand, and production case study, starting from the available energy datasets of different energy sources in Algeria and how they are used to construct ANN forecasting models for short-term forecasting of demand and production.

5.2 Available data and Context

5.2.1 The region of Annaba

The experiments described in this thesis are based on data from Annaba region in north eastern of Algeria. As illustrated in the geographic map in Figure 5.1, Annaba is one of the largest cities in Algeria and also known as one of the poles of Algerian industry.

Figure 5.1: Geographic map of Annaba (Source: Google)

The topography of the Annaba region favourites air stagnation and the formation of temperature inversions. These situations cause the accumulation of pollutants and the resulting elevation of concentration levels. The effects of sea, land and slope breezes contribute to the transport of clouds of pollutants. Indeed, the clouds of pollutants are carried by the land breeze at night towards the sea. These clouds of pollutants return to the city by sea breeze effect along the mountain of Edough. The clouds revolve over the city in the shape of a circle. Pollutants are deposited slowly by gravity and there is pollution affecting all three receptors (sea, land, air) (Ghazi 2016).

The presented work was validated using datasets provided by the national energy distribution company SONALGAZ and concern the studied energy types. Firstly, the electric load dataset, this dataset comports two years of data in hourly steps, concerning the Algerian city Annaba, from 01/01/2014 to 31/12/2015 with a total of 17544 data observations. Figure 5.2 illustrates the hourly electric load in 2014 and 2015.

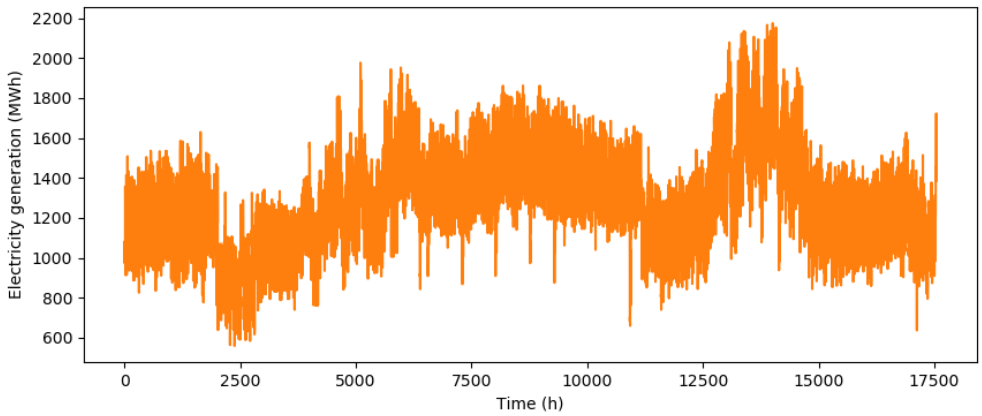


Figure 5.2: Hourly electric load

Secondly, the natural gas consumption dataset, this dataset covers the hourly natural gas consumption data of the city of Annaba for both residential and industrial sectors over one year i.e., 8760 observations as illustrated in figure 5.3.

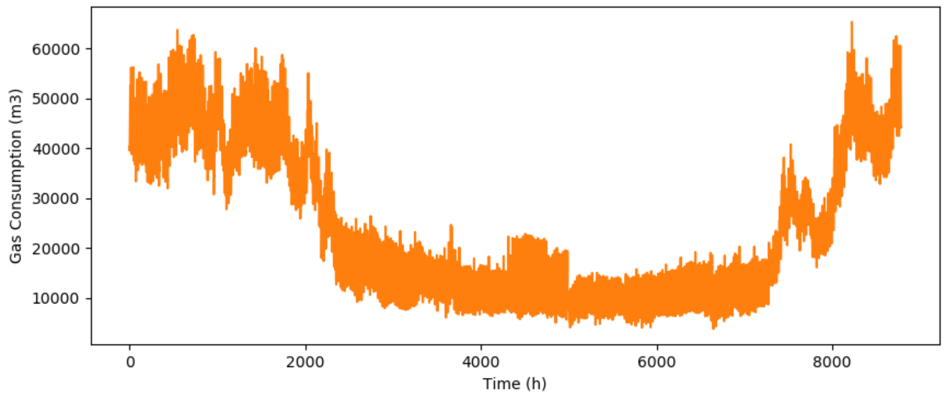


Figure 5.3: Hourly gas consumption

However due to the lack of real hourly temperature data, we used the estimation method presented in (Farfar and Khadir 2018) using the daily minimum T_{mini} and maximum T_{maxi} temperatures recorded by the Algerian meteorological office. The mentioned method is based on two equations; the first equation presumes that the temperature after sunrise follows a sine curve and that the daily temperature reaches its maximum value two hours after noon as follows:

$$T_{ih} = (T_{maxi} - T_{mini}).sin(\frac{\pi.t_a}{DL_i + 4}) + T_{mini} \tag{5.1}$$

where T_{ih} is temperature at hour h and day i, DL_i is day duration from sunrise to sunset (in hours), while t_a is the time in hours counted from the sunrise. Concerning the time after sunset, the temperature is presumed to follow a logarithmic decline, eq (5.2).

$$T_{ih} = T_s - \frac{T_s - T_{mini+1}}{ln(24 - DL_i)}.ln(t_b) \tag{5.2}$$

Where T_s is the temperature at sunset obtained using equation (5.1), and T_{mini+1} is the minimum temperature of the following day i+1, which is replaced by the minimum temperature of the same day for the times before sunrise, while t_b express the time in hours after sunset plus 1 hour.

Based on the previous figures, the gas consumption and the electric load variations over the observed period demonstrates that energy consumption shows a seasonal dependency related mainly to temperature.

5.2.2 The region of Adrar

Currently, the city of Annaba does not possess renewable energy power plants and energy sources, therefore, the simulation presents a hypothetical practical solution to examine the potential contribution of renewable energy on reducing air pollution. In this simulation the case of renewable energy production in Adrar city is used as an example. The region of Adrar is located in the south-west of Algeria, at more than 1400 km from the capital Algiers and falls between the meridians: 2°E and 6°W, and the parallels 20°and 32°North. Its total area is 427,968 Km^2 and Population 431 270 inhabitants, about 18% of the total area of Algeria and almost a fifth of the entire National territory.

Due to the geographical situation, permitting ideal wind and solar conditions, the region has been chosen for the implantation of wind and PV farms where currently around 50 % of its electrical energy comes from renewable sources. Indeed, the frequency of the strong and fast winds is very high throughout the year, especially the sirocco reaching 100 Km/h. Generally, the spring season (March-April), experiences a greater frequency of sand winds (Ainouche and Ainouche 2005).

Concerning the solar potential of the region, (Oudrane et al. 2019) conducted a thorough study where the solar potential for private domestic PV source in the region is modelled. From the results obtained, the authors concluded that the south-facing facade in the summer season is the most optimal for obtaining a very high density of the solar flow. They also, concluded that the best solar gain is recorded in July in the south facade with a density of 245.28 W/m^2, one of the highest ratios in the world. For all that potential, the region was chosen for holding one of the most important Renewable Energy plan and projects. Another reason to choose those regions is that they are not part of the interconnected national electrical grid, and must rely on local isolated ones. The planning is to develop those isolated smaller Grids (preferably RE based) and connect them to the interconnected northern grid once their importance is confirmed.

The used time series datasets were also provided by the national grid company SONELGAZ and concern wind and solar energies. They contain quarterly observations and consist of historical load values and other related exogenous variables mainly meteorological data. To be able to use the datasets in the simulation, resampling the datasets into hourly steps is mandatory, the resampling was done by averaging every four quarter hours' production values into one-hour data point, while doing the same operation with the meteorological data.

5.2.2.1 Kbertene wind farm

The Kabertene (Adrar province) wind farm for electricity production is a successful model for harnessing clean and renewable energy, situated on the territory of the commune of Tissabit (80 km north of Adrar), this project is the fruit of an Algerian-French partnership and the first of its kind on a national scale.

The plant is an experimental station representing a successful model in harnessing wind power for power generation. With a dozen wind turbines, installed on the basis of technical and field studies, taking into account the wind currents that characterize the region, this plant ensures clean and renewable alternative production of 10 Megawatts of electricity, integrated into the local electricity network to strengthen the energy supply capacities of the wilaya of Adrar, see Figure 5.4.

Figure 5.4: Kabertene wind farm (Khadir and Bouziane 2020)

Kabrten's experience has informed the producers on several aspects that must be taken care of by the scientific community, especially the problems related to extreme temperatures, the impact of desert dust on turbines, intermittency and its impact on the network.

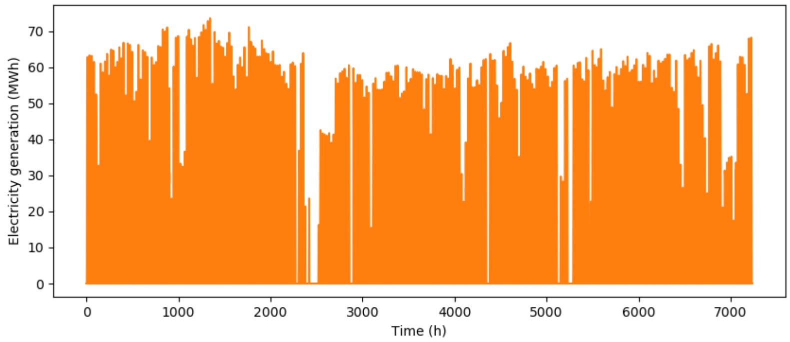
5.2.2.2 Photovoltaic field

A new photovoltaic pilot plant with a capacity of 20 Megawatts has been put on service at the level of the Research unit in renewable energies in Saharan environment (URER-MS) in Adrar. The project aims to test the energy efficiency of this type of installation in the Saharan regions and thus offer a scientific database paving the way for its possible generalization. For now, the RE production (wind and solar) reaches nearly 40% of the 60 to 100 MW produced in the region when meteorological condition are optimum. Figure 5.5 illustrates the PV plant of Adrar.

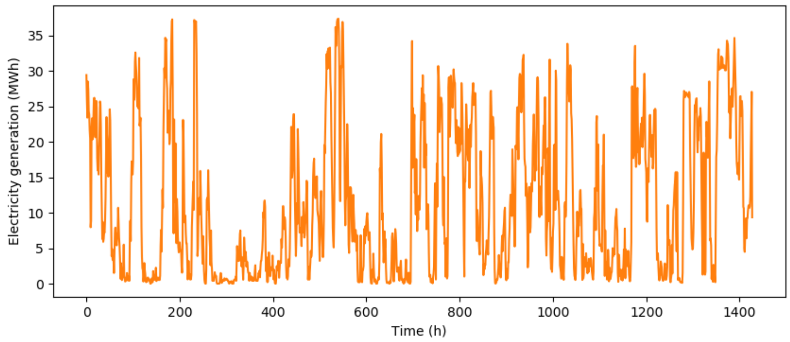
Figure 5.5: Photovoltaic field of Adrar (Khadir and Bouziane 2020)

The available data consists of a set of 5 variables: load produced from PV, load produced from wind turbines, wind speed, temperature, irradiance. The measures are quarterly (every 15 mn) and cover the period from January 2016 to April 2017.

The production of electrical load (both wind and solar) are the main variables, its unit is expressed in watt-hours (Wh). The production values depend mainly on the irradiance for PV, and on wind for Wind Turbines, and range respectively between 0 to 19.4 Wh and 0 to 9 Wh. The total solar irradiance (noted TSI for total solar irradiance) is the quantity of radiative energy coming from the Sun (total intensity of the radiation) received by a surface of 1 m^2 from the top of the Earth's atmosphere. The irradiance of this series varies in an interval between 0 and 1237.8 $Watt/m^2$ approximately. These values are zero at night (between 8 p.m. and 7 a.m. for the summer months and between 5 p.m. and 8 a.m. for the winter months), the maximum values are noted in the interval (12 p.m. - 2 p.m.).



((a)) PV load



((b)) Wind energy load

Figure 5.6: The PV and wind energy generation

For temperatures, celsius degrees (°C) is the unit of the temperature scale, which varies between 0 and 50 (the maximum values are marked on the interval (12 p.m-2p.m) in the months of summer). The wind speed is expressed in km/h, the values vary between 0 and 18 km/h, its values are not correlated with time but with sandstorms which can cover the solar panels and decrease their efficiencies and therefore the production. High wind speeds will affect irradiance and therefore production. For that reason, wind speed must be taken into account as an exogenous variable in PV electrical load production modelling. Figure 5.6 above shows the load data while figure 5.7 illustrates the exogenous inputs.

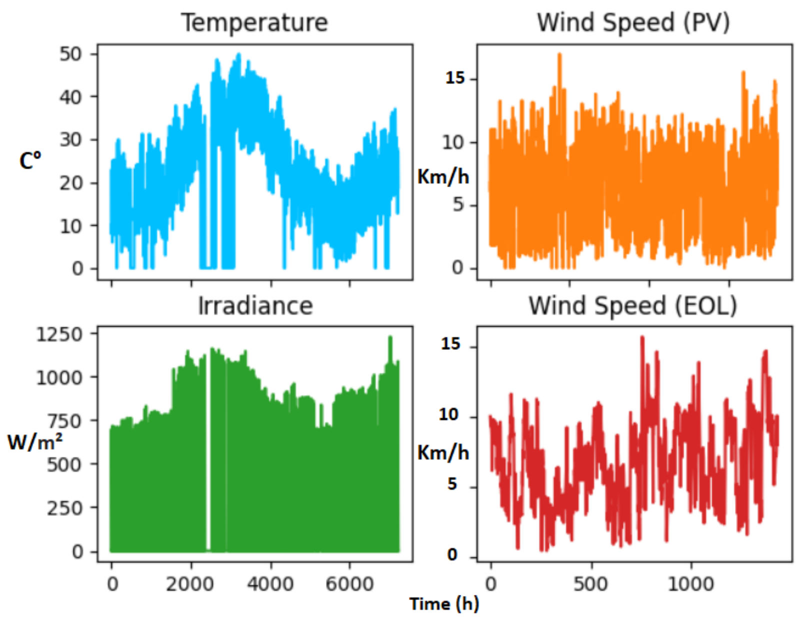


Figure 5.7: The available exogenous inputs

5.3 Shallow ANN models for forecasting energy production and consumption

The first stage of our work focuses on forecasting energy production and consumption, to achieve this goal, we used ANN models, which are a biologically inspired paradigm copying the human and animal nervous system and their information processing capabilities and were briefly presented in chapter 3. One of the major advantages of ANN's is the scalability of the network architecture as well as the values of the weights and biases resulting from a learning process. Once the training process is over and the parameters fixed, the ANN may be able to solve any nonlinear problem and therefore approximate any nonlinear function. As a first step, the developed models are feed forward networks, and more precisely traditional MLP models. The networks topologies and parameters were fixed after multiple trials to choose the best topology in terms of the best validation error.

5.3.1 Gas and electricity consumption models

In order to forecast gas consumption and electricity production, two 3-layer MLP are used; each MLP is dedicated to one of the two tasks. After several training and validation trials, a hidden layer containing 14 neurons was selected along with a backpropagation optimization algorithm. In order to train each model, the available datasets are divided into training sets and test sets. The training sets are used to adjust the weights and the biases of the networks and the test sets are used for validation in order to ensure better generalization.

The next step is the selection of an appropriate input vector to construct the network topology and the training database. As There is a strong correlation of the natural gas load with meteorological elements which is mainly represented by temperature, we used the estimated temperature values that were computed using the approach described in section 5.2 as an exogenous variable, therefore, a 10-dimension vector consisting of past energy consumption and temperature values was constructed. The forecasted energy load E at hour t is therefore explained by the following previous values: $E_{t-1}, E_{t-2}, E_{t-3}, E_{t-24}$ and E_{t-168}, which represents respectively the load in the previous three hours, the previous day and the previous week's load at the same hour of the day. In addition to the past load values, the temperature estimations T at the same selected times, i.e., $T_{t-1}, T_{t-2}, T_{t-3}, T_{t-24}$ and T_{t-168} are used (see figure 5.8).

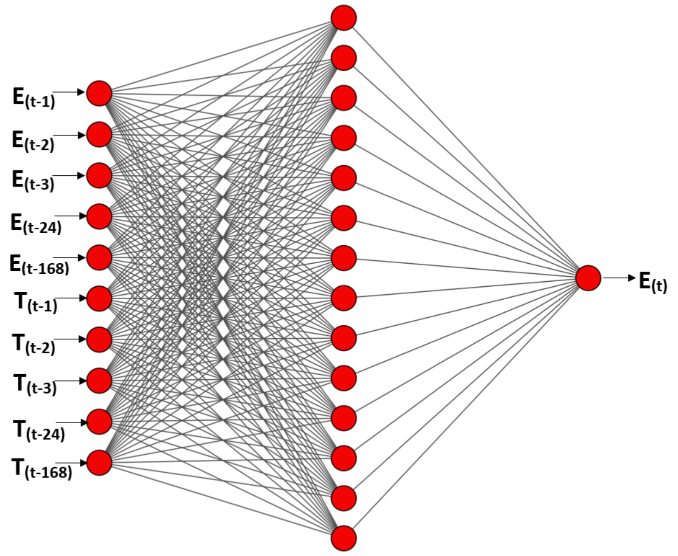


Figure 5.8: The architecture of the proposed MLP models

5.3.2 Renewable energy models

5.3.2.1 PV energy model

Photovoltaic (PV) energy production depends on multiple factors, for instance, the type of PV module, the degradation of the PV module, solar radiation, and temperature, etc., (Vidyanandan 2017). In the case of Adrar's power plant, and in order to forecast the hourly solar energy production, a time series dataset from SONALGAZ was used. The dataset contained 16 months of hourly production in addition to the most influencing factors. The corresponding available exogenous inputs are:

- Irradiance, as variation in solar radiation affects many of the PV parameters.
- Weather temperature, as the PV module is very sensitive to this aspect. The optimum temperature is 25°and generally, the module loses 0.5% of its efficiency per degree increase in temperature.
- Wind speed, which can affect the temperature of the PV panel and the irradiance.
- The energy production of the previous hour.

Finally, and in order to forecast the hourly PV energy production, we used the first 12 months i.e., one year for training and the next 4 months for testing. As a result, for the training, the network is fed with an input vector consisting of four parameters, in addition to the irradiance, the temperature and the wind speed; the previous hour's production value is used, hence, the solar energy PV at the hour t is explained by $Ir_{(t-1)}, T_{(t-1)}, Ws_{(t-1)}$ and $PV_{(t-1)}$ for Irradiance, temperature, wind speed and past load wind production, respectively, equation (5.3).

$$PV(t) = f_{ANN}(PV_{(t-1)}, Ir_{(t-1)}, Ws_{(t-1)}, T_{(t-1)}) \tag{5.3}$$

Note that the input vector is chosen from experimentations and found to be suitable for a satisfactory result. The network topology issued from the selected input vector and experimentation concerning the number of neurons is giving in Figure 5.9.

5.3.2.2 Wind energy model

For the wind energy case, we opted for a MLP, that consists of four neurons in the hidden layer. As stated in the previous Section, the data contains the

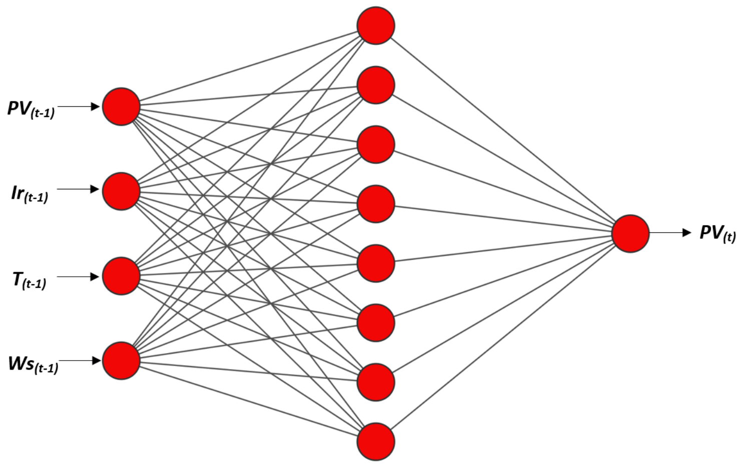


Figure 5.9: The PV ANN model

hourly energy production in addition to the wind speed as an exogenous variable, to forecast wind energy WE at the hour t we fed the network two variables, $WE_{(t-1)}$ and $WS_{(t-1)}$ which represent the energy production at hour t_1 and the wind speed respectively (equation 5.4) and the topology of the used network is illustrated in figure 5.10.

$$WE(t) = f_{ANN}(WE_{(t-1)}, Ws_{(t-1)}) \tag{5.4}$$

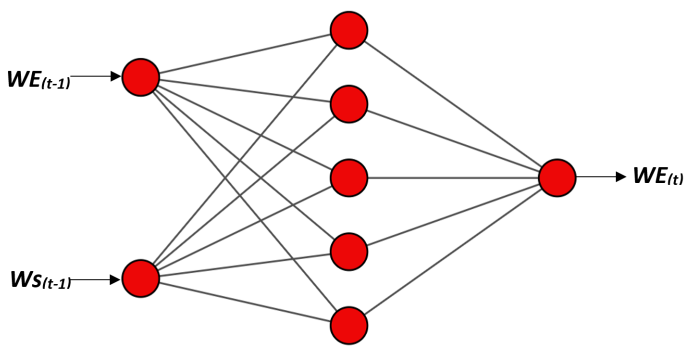


Figure 5.10: The wind energy ANN model

All the forecasting models were trained using the stochastic gradient descent algorithm, which is an optimization algorithm used to minimize the cost function by updating the weights of the neural network. Furthermore, a hyperbolic tangent (tanh) function was used as an activation function for our models. Table 5.1 summarizes the used inputs and parameters of the four forecasting models.

Table 5.1: The used parameters for the MLP models

Models	Hidden layer size	Number of inputs	Inputs	Learning rate	Activation function
Electric load	14	10	Previous load and temperature	0.001	Tanh
Gas consumption	14	10	Previous load and temperature	0.002	Tanh
Solar energy	8	4	Previous load, irradiance, temperature, wind speed	0.001	Tanh
Wind energy	8	2	Previous load, wind speed	0.001	Tanh

5.4 LSTM models for forecasting energy production and consumption

After using traditional feed-forward ANN to forecast the energy consumption and production, and in order to enhance the forecasting performance, a variant of recurrent ANN which are LSTM networks are considered on the same available datasets.

Choosing the best topology of LSTM models is not straightforward, as it requires tuning a set of hyperparameters, these hyperparameters involve mainly: The number of hidden layers and the number of neurones in each layer, learning rate, dropout rate, lag size (number of used past values), batch size, and the number of epochs. Therefore, multiple trials were performed in order to select the best combination of hyperparameters in terms of the best validation errors. Some tested hyperparameters values are illustrated in table 5.2.

Table 5.2: Tested hyperparameters values

Hyperparameter	Values
Number of hidden layers	[1, 2, 3]
Number of neurons in each layer	[10, 20, 40, 80, 160, 320, 500]
Lag size	[1, 2, 3, 5, 7, 12]
Learning rate	[0.0001, 0.001, 0.01]
Dropout rate	[0.1, 0.2, 0.3]
Batch size	[8, 16, 32, 64]
Number of epochs	[500, 1000, 1500, 2000]

As for the optimisation algorithm, the Adam optimizer (Kingma and Ba 2015) is used for all the models. More details about the models are as follows:

Gas and electricity consumption models: In order to forecast the short-term electricity production and natural gas consumption, two LSTM models are trained, where each model is dedicated to one specific task. After multiple experimentation and testing different topologies and hyperparameters combination, two hidden LSTM layers containing 200 and 100 neurons respectively in addition to a hidden dense layer with 100 neurons were selected for each model, along with the hyperbolic tangent (tanh) activation function. In the case of the electricity model, we used 60% of the data for the training, 40% for testing, i.e., 12280 hours for training and 5263 hours as test set. In the other hand, for the gas consumption model we used 70% data for training and 30% for validation and testing. The next step is the selection of an appropriate input vector to train the network by selecting a suitable lag size, selecting an appropriate lag size plays an essential role in the models learning, in our case we chose to train both models using the 5 past values. Therefore, to predict the energy value E at the hour t, we used an input vector consists of $E_{(t-1)}, E_{(t-2)}, E_{(t-3)}, E_{(t-4)}, E_{(t-5)}$ sent the load E at the last five hours in addition to the hourly temperature estimations at the same chosen periods.

The solar energy model: In order to forecast solar energy production, we used 70% of the data for training, and 30% for validation and testing. The constructed LSTM model consists of two hidden layers, with 150 and 80 neurons in the first and second hidden LSTM layer respectively and 50 neurons in the last hidden dense layer, and the chosen lag size is 3 past observations. Therefore, the input vector consists of 3 lagged values of each input variable (previous load, irradiance, temperature, and wind speed).

The wind energy model: Same as solar energy, wind energy generation is influenced by different factors, such as wind speed, air density, pressure, and height of tower. In our case, the available Adrar's dataset contains two variables, the historical load and the wind speed and covers 1429 hours of power generation values. We adopted a two-hidden layer LSTM, with 150 and 80 neurons in the hidden LSTM layers and 50 in the hidden dense layer. Similarly, to the solar energy model, we used 70% of data for training, i.e., 1000 observations, and the rest of the data for validation and testing, and for the inputs, we used 2 lagged values of the two available variables. Table 5.3 below summarizes the used hyperparameters of each LSTM model.

Table 5.3: The used LSTM hyperparameters

Models	Hidden layers size	Lag size	Learning rate	Dropout size	Batch size
Electric load	[200-100-100]	5	0.0001	[0.2-0.1]	32
Gas consumption	[200-100-100]	5	0.001	[0.2-0.1]	32
Solar energy	[150-80-50]	3	0.001	[0.1-0.1]	16
Wind energy	[150-80-50]	2	0.001	[0.1-0.1]	16

5.5 Experiments and results

In order to determine the most suitable models' architecture and hyperparameters, the models are tested through several epochs using RMSE and MAE evaluation metrics, which are calculated using the equations (5.5) and (5.6) respectively, where N represents the number of predicted samples, a_i is the actual value, and b_i is the observed value.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i}^{N}(a_i - p_i)^2} \tag{5.5}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|a_i - p_i| \tag{5.6}$$

Although ANNs could obtain good performances during the training phase, it could experience degradation in and fail to obtain good generalisation on a separate validation and test datasets mainly due to overfitting. Therefore, several were performed using different model architectures and configurations in order to choose the best topology and hyperparameters in terms of the validation error.

Table 5.4 presents the obtained performance of the MLP networks using both RMSE and MAE metrics, while figure 5.11 illustrates 24 hours' forecasting samples for each model for winter season.

Table 5.4: MLP models performance

Models	MAE		RMSE	
	Train	Test	Train	Test
Electric load	0.054	0.068	0.052	0.071
Gas consumption	0.024	0.032	0.023	0.041
Solar energy	0.086	0.124	0.131	0.152
Wind energy	0.065	0.119	0.144	0.214

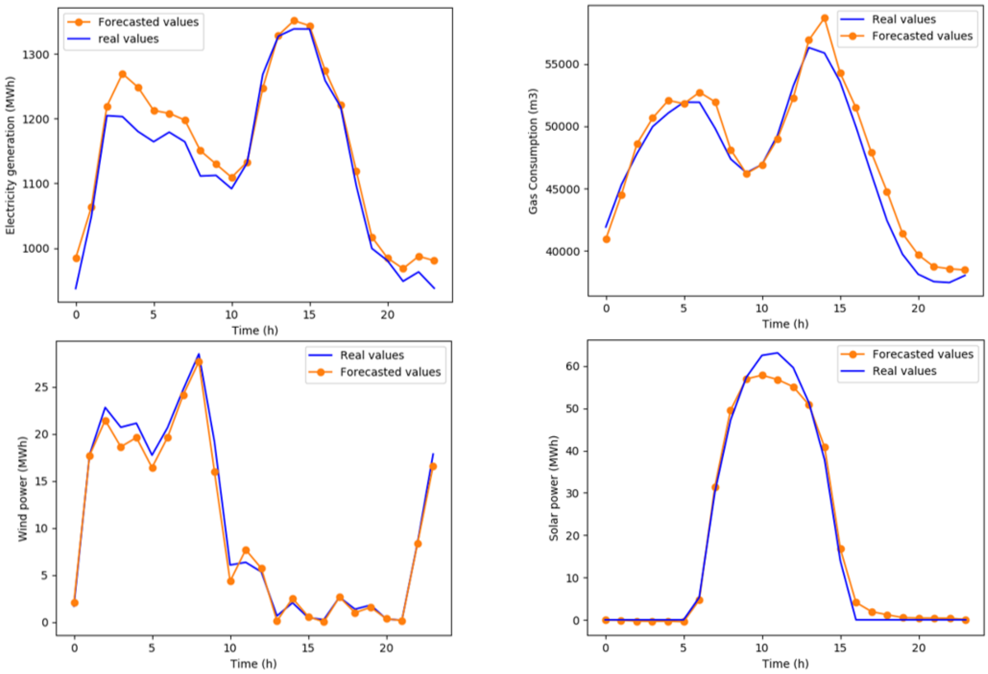


Figure 5.11: 24 hours' prediction samples for each MLP model for winter season

For the case of the LSTM models, the performances according to RMSE and MAE metrics are highlighted in the table 5.5.

Table 5.5: LSTM models performance

Models	MAE		RMSE	
	Train	Test	Train	Test
Electric load	0.058	0.061	0.009	0.012
Gas consumption	0.001	0.002	0.027	0.040
Solar energy	0.021	0.024	0.019	0.026
Wind energy	0.106	0.109	0.105	0.166

Figure 5.12 plots the graph difference between the forecasted and actual values of a 24 hours' samples for all the forecasting models obtained for a random day on a summer season. It can be seen that the models are able to fit the overall

daily trend despite exhibiting some weaknesses in peak's estimation, which is understandable due to the high non-linearity of these phenomena.

It should be noted that for studying purposes, we used weighted values for the renewable energy production, as Adrar's overall demand is considerably low compared to Annaba's which would have made the comparison irrelevant.

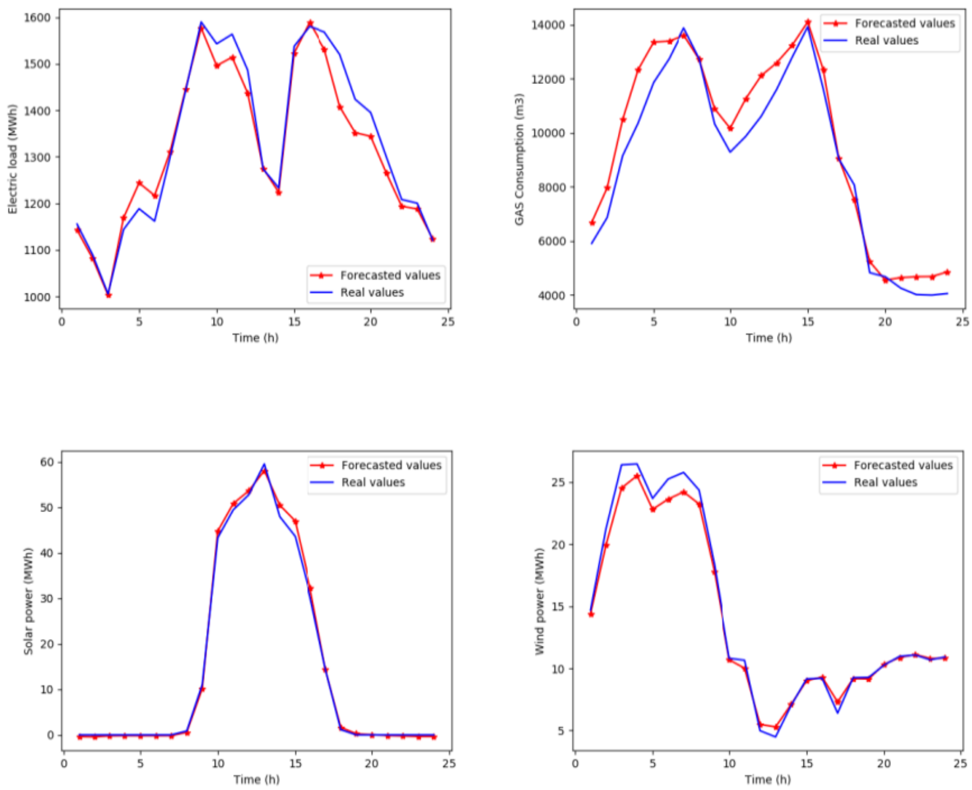


Figure 5.12: 24 hours' prediction samples for each LSTM model for summer season

5.6 Compraison and discussion

In order to validate the performance of the developed models, LSTM and MLP models are compared against each other and against some of the most used state-of-the-art forecasting techniques, such as ARIMA and SVR models.

Here all the tested models received the same training and test data, the performance metrics RMSE and MAE were used in order to compare the performances of the models, and the forecasting error of each model is presented in table 5.6.

Table 5.6: Comparison between the used models and Benchmark models

Metric	Model	LSTM	MLP	SVR	ARIMA
RMSE	Electric load	0.012	0.071	0.218	0.096
	Gas consumption	0.040	0.041	0.361	0.041
	Solar energy	0.026	0.152	0.380	0.217
	Wind energy	0.166	0.214	0.338	0.220
	Average	0.061	0.119	0.324	0.144
MAE	Electric load	0.061	0.068	0.026	0.063
	Gas consumption	0.002	0.032	0.025	0.031
	Solar energy	0.024	0.124	0.073	0.137
	Wind energy	0.109	0.119	0.078	0.149
	Average	0.049	0.085	0.051	0.096

According to the obtained results, all the tested methods achieved good and acceptable performances with a small advantage for the LSTM models which achieved smaller RMSE and MAE errors.

5.7 Conclusion

In this chapter, the first part of our contribution, consisting in short-term forecasting of energy production and consumption is presented where both MLP and LSTM neural network models are considered. Before depicting the design of the ANN models, the studied area and the available datasets that were provided by the national company Sonalgaz and concerns both fossil and renewable energy sources were presented. Its only then when ANN topologies for demand and different Renewable Energy sources (Wind and Solar) are described with an empirical justification of their hyperparameters. Result analysis are then performed using a number of error metrics.

One must be sane to think clearly, but one can think deeply and be quite insane.

Nikola Tesla

6

MAS for forecasting air pollution issued from energy consumption and production

▷ *In this chapter we present our proposed approach for the design of a multi-agent-based air pollution system.* ◁

Contents

6.1 Introduction

In this chapter we present our proposed approach for the design of a multi-agent-based air pollution system. The system is based on the combination of energy consumption and production ANN forecasting models with an autonomous agent in a MAS for forecasting carbon dioxide (CO_2) issued from different energy sources in the city of Annaba.

We'll start by describing the designed agents and how the forecasting models are integrated into these agents, later we'll present how the agents interact with each other, and we'll finish the chapter by presenting the emissions factors of CO_2 and the simulation results.

6.2 The proposed System

6.2.1 Overview

This section describes the architecture of the simulation system which relies on combining the power of MAS collaborative agents and machine learning forecasting by integrating trained ANN models into autonomous agents. The system consists of multiple forecasting models, each of which is specific to a certain energy type. The agents use the models to forecast the energy production and consumption at each simulation time step. These forecasts are sent to a coordination agent, which calculates the total carbon emissions issued from all available sources and calculates the reduced amount of CO_2. The coordination agent also calculates the amount of natural gas saved by using renewable energy.

Figure 6.1 bellow illustrates the general flow and procedure of the system along with the system's components, which is composed of different types of agents.

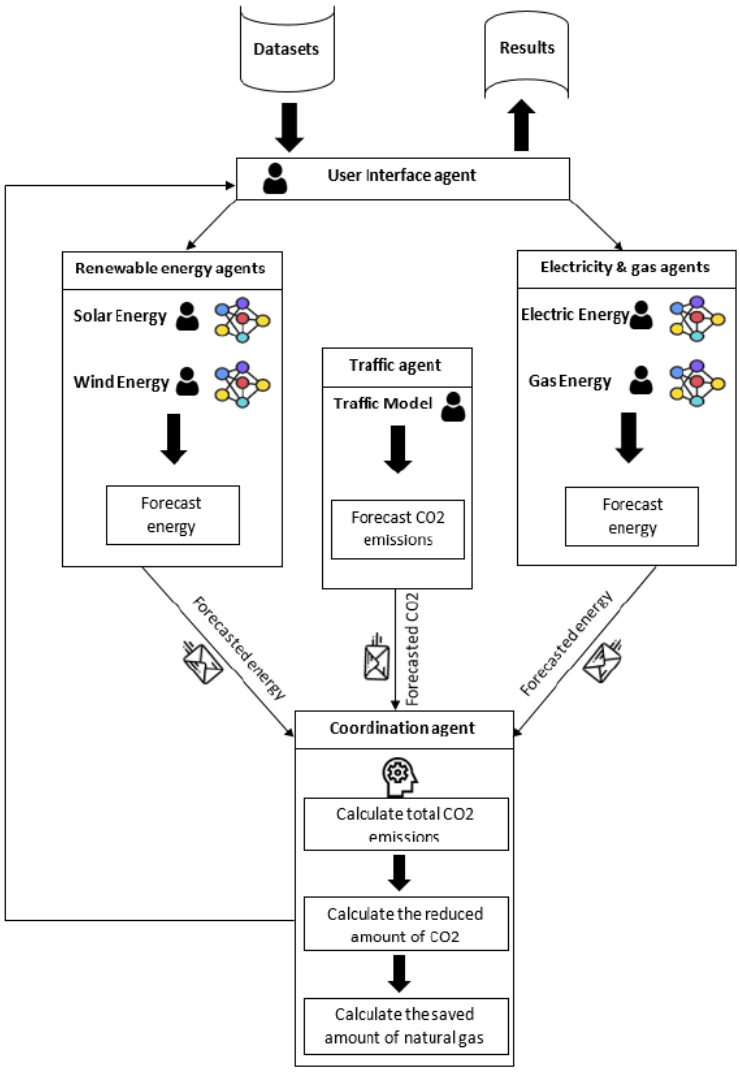


Figure 6.1: The system's general architecture

6.2.2 The MAS Architecture

Considering that our work focuses mainly on short-term forecasting, we used ANN models as they proved to be an effective technique for forecasting problems, and since we deal with multiple subproblems i.e. forecasting multiple energy sources, calculating the carbon emissions, and investigating the effects of renewable energy, we integrated the ANN in a MAS architecture, as it is known that MAS architecture is a convenient solution for complex or distributed problems, by assigning each subproblem to a distinct autonomous agent. Another advantage of using such approach is the flexibility it offers, to cope with further added functionalities, such as adding more energy and CO_2 sources. The proposed architecture incorporates multiple autonomous agents. There are two types of agents. Firstly, forecasting agents, which are responsible for forecasting energy production and consumption. Secondly, core agents that perform other basic tasks for the simulation, e.g., loading the data, starting the simulation, and displaying the results. he system's agents are as follows:

Forecasting agents:

Gas agent: uses an ANN model to forecast the hourly gas consumption.

Electricity agent: uses an ANN model to forecast the hourly electricity production and calculate the equivalent required amount of natural gas to produce such a load.

Renewable energy agents (PV and wind agents): use ANN models to forecast the hourly renewable energy production.

Traffic agent: Traffic emissions are a major source of air pollution, and so to add more depth to our simulator we model traffic emissions. However, due to the lack of datasets and the difficulty of gathering traffic related data, we used a mathematical model based on the equation (6.1) and the emission factors in (EPA 2020). This model gives an estimation of the CO_2 emitted by road traffic at each hour of the day. The amount of CO_2 emissions during rush hours is significantly higher due to the increase in traffic density. Hence, traffic emissions TE at the hour t is calculated as follows:

$$TE_t = E.N(|cos(2\pi(t-8)/24)|) \tag{6.1}$$

Where E is the yearly average carbon emissions per vehicle divided by the number of hours per year and N is the estimated number of vehicles in the city of Annaba, which according to the Algerian National Office of Statistics (ONS 2018) is approximately 200,000 vehicles. The equation can illustrate that the rush hours in Annaba, when there is a peak in traffic emissions, occur at 8 am and 8 pm. Furthermore, this mathematical model is used by an autonomous agent called traffic agent to forecast an hourly estimated value of traffic emissions.

Core agents:

User interface agent: the user interface agent performs some basic yet essential tasks for the simulation, for instance importing data, starting and stopping the simulation, and displaying the results.

Coordination agent: the central agent in our architecture. The agent receives the forecasted values from the forecasting agents and calculates the equivalent carbon emissions from each source. These are then summed to obtain the hourly total emitted CO_2. The agent then computes the effects of using renewable energies in terms of decreasing air pollution and economizing on the use of natural gas.

The used agents are reactive agents, as they react in reasonable time and in an appropriate way to changes in the environment, also, they don't have a symbolic representation of the environment, and their decisions are based on the precepts of their environment (in our case the ANN's input data).

Figure 6.2 illustrates a UML class diagram that describes our system's agents and their architecture, all the agents are inherited from the principal Agent class, and each agent has its own functionalities.

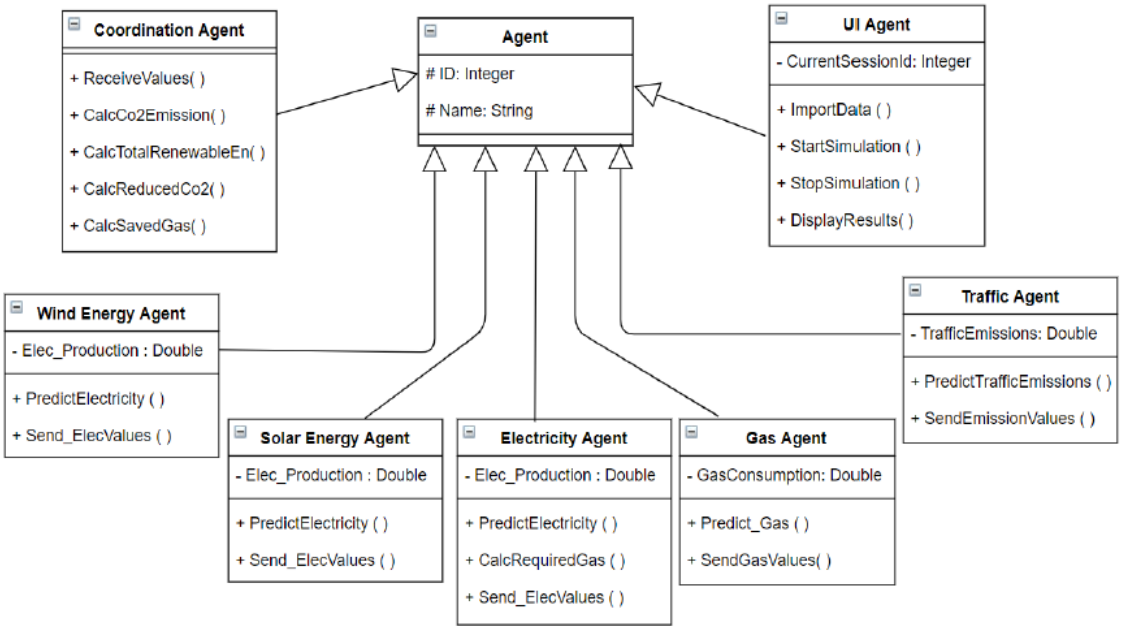


Figure 6.2: Simplified UML diagram of the proposed system

6.2.2.1 Agents' communication

Distributed problem solving is one of the advantages of an agent-based approach. However, it requires cooperation and communications between the agents. The main way of communicating is to exchange messages using an Agent Communication Language (ACL). ACL is a L2 type of language which allows agents to communicate and share their knowledge in order to reach a goal or solve a problem.

An ACL message contains a set of one or several message elements, these elements are needed for establishing an effective agent communication and vary depending on the situation or the scenario. Table 6.1 below presents the full set of FIPA ACL message elements.

Table 6.1: The elements of an ACL message

Category of Elements	Elements
Type of communication	Performative
Participant in communication	Sender Receiver Reply-to
Content of message	Content
Description of content	Language Encoding Ontology
Control of conversation	Protocol Conversation-id Reply-with In-reply-to Reply-by

The only mandatory element in an ACL message is the performative, which can be an inform message, a query, or a propose, . . . , etc, however, it is expected that an ACL message contain also a sender, receiver and content elements.

In our system, the agents coordinate among themselves by exchanging messages according to the ACL standard. For instance, the forecasting agents use the ANN models (described in chapter 4) and share the forecasted energy load with the coordination agent using ACL messages, where the sender is one of the forecasting agents, the receiver is the coordination agent, and the content is the forecasted load with a performative element set as inform.

6.2.2.2 The used technology

The overall system is implemented as a simulation tool using the approach described in the above steps. JADE (Java Agent DEvelopment Framework) (Bellifemine et al. 1999) was chosen to implement the MAS architecture. JADE is an open-source framework written in Java that follows the FIPA (Foundation for Intelligent Physical Agents) specifications and allows the agents to communicate through the ACL Standard. The JADE application consists of one or more containers, and a set of containers composes a platform. Every platform must include a special container called main container which is different from the rest of the containers. The main container includes two special agents: an AMS (Agent Management System) and a DF (Directory Facilitator). The AMS provides the naming service to provide each agent with a unique ID, while the DF enables agents to search and find the agents that provide the service they desire.

Table 6.2 summarizes the charactersitcs of the system used for the implementation and testing of the proposed simulation tool.

Table 6.2: The used system's characteristics

Component	Characteristic
CPU	i5-9300H
RAM	16 Go
GPU	NVIDIA GTX 1660 Ti
VRAM	6 Go
OS	Windows 10

The GPU allows a fast training of the ANNs and in particular the deep neural networks using the Nvidia technology Compute Unified Device Architecture (CUDA).

6.2.2.3 Agents' behaviours

In agent-oriented programming, the various tasks of an agent must obey certain rules and must be written in a form understandable by the platform used. In order for a JADE agent to perform a task, we first need to define these tasks. These tasks are called "Behaviours", in JADE, each behaviour must implement at least two methods: action() which designates the operations to be performed by the agent, and done() which expresses whether the behaviour has finished its execution or not. Jade offers different types of behaviours, varying from simple to composite behaviours, for instance:

One shot behaviour: It has the particularity of performing a task only once and then it ends. The One-Shot-Behaviour class implements the done () method and it always returns true.

Cyclic behaviour: a cyclic behaviour performs its task in a repetitive manner. The CyclicBehaviour class implements the done() method which always returns false.

Waker Behaviour: This behaviour is implemented to execute a method called onWake() after a duration passed as an argument to the constructor. This duration is expressed in milliseconds. The Behaviour ends immediately after executing the onWake () method.

Ticker Behaviour: The Ticker behaviour is implemented to perform its task periodically by the onTick() method. The duration of the period is passed as an argument to the constructor and it′s expressed in milliseconds. Table 6.2 presents the used behaviours in our system's agents.

Table 6.3: The agents' behaviour

Agent	Behaviour
User interface agent	One Shot Behaviour
Coordination agent	Ticker Behaviour
Electricity agent	Ticker Behaviour
Gas agent	Ticker Behaviour
Solar energy agent	Ticker Behaviour
Wind energy agent	Ticker Behaviour
Traffic agent	Ticker Behaviour

6.3 Experiments

For each hourly simulation step, forecasting agents predict energy production and consumption and communicate their values to the coordination agent. The latter has two main tasks. The first is to calculate the equivalent amount of burned natural gas that was used to produce the forecasted electricity production. This is computed according to the U.S Energy Information Administration (EIA 2021) by equation (6.1).

$$AG = HR/HV \tag{6.2}$$

Where AG is the amount of gas used per kWh, HR is the heat rate of the power plant and HV is the heat value of the used fuel. Hence, the amount of natural gas needed to generate one kWh is 0.00786 Mcf which is equivalent to $0.22m^3$. This assumes that the average HR of a natural gas-fired power plant is 8,039 Btu/kWh and the HV of natural gas is 1,023,000 Btu/Mcf.

The second task is to calculate the environmental cost of the produced energy. The agent computes the equivalent carbon emissions in the forecasted hour according to the emissions of natural gas in (EPA 2020). This indicates that the CO_2 equivalent of burning natural gas is 0.0551 metric tons CO_2/Mcf which is equal to 0.0019 metric tons CO_2/m^3. Finally, to simulate the benefits of using cleaner energy sources, the agent calculates the effects of using renewable energies on air pollution, i.e., the reduced amount of carbon emissions and the amount of fossil energy that would be saved.

Figure 6.3 shows a comparison between the carbon emissions coming from electricity generation using natural gas, and carbon emissions using RE to generate a portion of the load. The results show the benefits of using renewable energies as an alternative solution to generate electric power. According to the simulations, using solar and wind energy can help reduce the emissions by up to 33 tons of CO_2 per hour, with an average of 15 tons of CO_2 per hour.

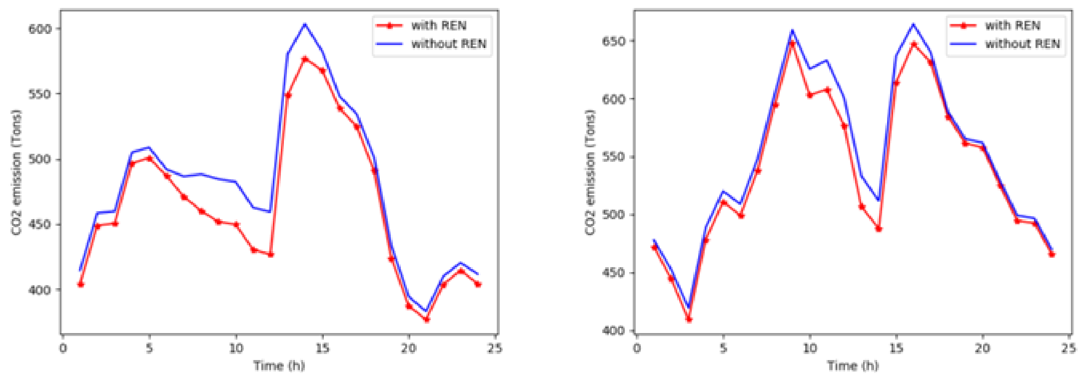


Figure 6.3: The CO_2 emission issued from electricity generation (a) in winter (b) in summer

Since producing electricity consumes a large amount of natural gas, using renewable energies can considerably reduce the amount of gas used. Based on the previous two samples shown in Figure 6.3, Tables 6.4 and 6.5 highlight some simulation results and statistics.

Table 6.4: The obtained CO_2 emissions from electricity generation (in Tons)

Sample	MAX hourly emissions	MIN hourly emissions	AVG hourly emissions	Daily emissions
Winter without RE	603	383	479	11505
Winter with RE	576	377	464	11136
Summer without RE	664	419	551	13232
Summer with RE	647	409	539	12948

Table 6.5: Natural gas (NG) usage for electricity generation (in m^3)

Sample	MAX hourly burned NG	MIN hourly burned NG	AVG hourly burned NG	Daily burned NG
Winter without RE	317640	201740	252303	6055280
Winter with RE	303600	198440	244227	5861460
Summer without RE	349580	220660	290189	6964540
Summer with RE	341000	215380	283955	6814940

To calculate the total carbon emissions, the coordination agent sums the emission values from all of the models' simulation results. Hence, the CO_2 emission at hour t is explained by the following equation:

$$CO2_t = GE_t + EE_t + TE_t \tag{6.3}$$

Where, GE_t, EE_t, TE_t, represent the gas consumption, electricity production, and the traffic emissions respectively at hour t. Although CO_2 emissions in the residential sector may depend on multiple factors, in this work we only consider that it is equal to the emissions from natural gas and natural gas burned to produce electric power. Figure 6.4, illustrates the results obtained from a simulation of the total carbon emissions on a typical winter day.

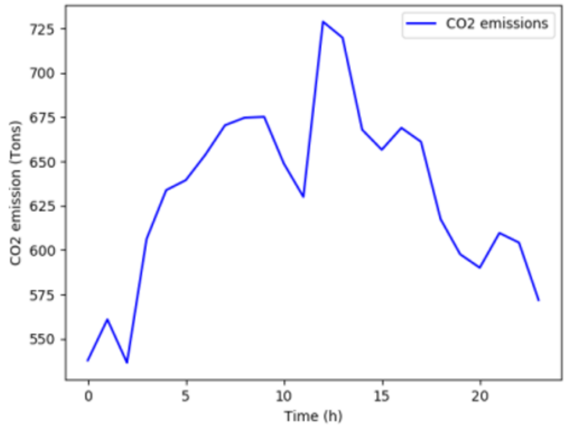


Figure 6.4: Total CO_2 emission

6.4 DZ-SimAP

The system was implemented as a simulation tool named DZ-SimAP, it offers the possibility to configure and start a simulation, visualise the predictions, and retrain the forecasting models if needed to.

Figure 6.5 illustrates the home panel of the tool which is shown at the first execution of the software, it offers access to several functionalities that allow configuration, visualization and saving the simulation scenarios as well as their results.

The simulator is designed in a generic way, it offers the possibility of re-training the prediction models for each energy source using the "Models" tab as shown in Figure 6.6. The data tab enables to configure the simulation by importing the data and choosing the number of simulation time steps, while the simulation tab allows to start and cancel the simulation.

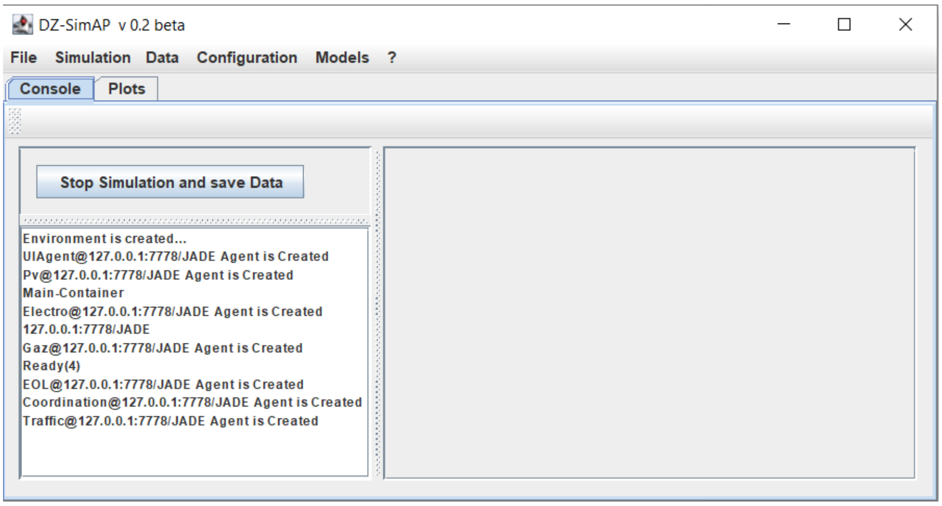


Figure 6.5: Dz-SimAP Home panel

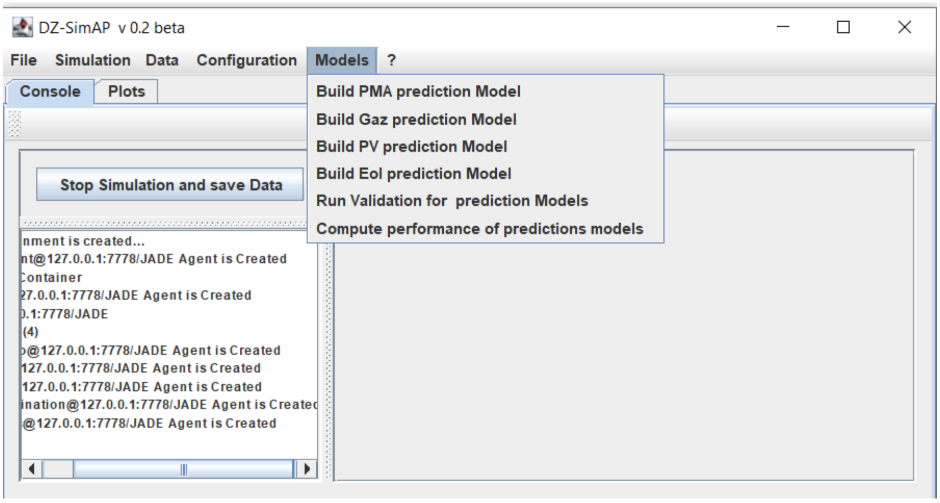


Figure 6.6: ANN Training Tab

Another component in the home panel is the log console, this console displays some information about the system and the simulation, such as the created agents and the imported data.

Also, we have a results console, where the results of the simulation are displayed in real time which allows to monitor the energy values and its equivalent CO_2 emissions as illustrated in figure 6.7.

SimStep	PMA	GAS	PV	CO2	GAS CO2	CO2 ENR
15	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
16	1502.72394..	49066.2674..	1.98417074..	2.8551755...	20.5096997..	2.8514055...
17	1648.48541..	49340.1013..	2.03574369..	3.1321222...	20.6241623..	3.1282543...
18	1688.66803..	50653.7856..	2.04883633..	3.2084692...	21.1732823..	3.2045764...
19	1766.84624..	52543.6776..	1.69621157..	3.3570078...	21.9632572..	3.3537850...
20	1771.93868..	54538.4312..	1.53724606..	3.3666834...	22.7970642..	3.3637627...
21	1731.36189..	54717.0253..	1.77698233..	3.2895875...	22.8717165..	3.2862113...
22	1690.82420..	53986.7920..	0.36114656..	3.2125659...	22.5664790..	3.2118798...
23	1624.68262..	51777.5661..	0.26473126..	3.0868969...	21.6430226..	3.0863939...
24	1403.79602..	50085.5751..	0.29038055..	2.6672124...	20.9357704..	2.6666607...
25	1370.75367..	46536.9672..	0.46156464..	2.6044319...	19.4524523..	2.6035550...
26	1334.94934..	43733.5556..	0.45794656..	2.5364037...	18.2806262..	2.5355336...
27	1340.19967..	42380.6354..	0.13901195..	2.5463793...	17.7151056..	2.5461152...
28	1422.27226..	42361.9955..	0.22489447..	2.7023172...	17.7073141..	2.7018899...
29	1431.02177..	42471.0845..	0.25941775..	2.7189413...	17.7529133..	2.7184484...

Figure 6.7: The results console

6.5 Conclusion

In this chapter we have presented a general framework for modeling air pollution as forecasting system. We have presented how we can combine between prediction models and a multi-agent architecture along with the global architecture of the system and all the developed agents, we also presented the used technologies, from the development framework to the used agents' communication language, in the 3rd and 4th sections of the chapter we presented how the CO_2 emissions are calculated in addition to some simulation results. Lastly, we presented an overview of the implemented tool that englobes the described system.

Truth is ever to be found in simplicity, and not in the multiplicity and confusion of things.

Isaac Newton

7 Conclusion and perspectives

7.1 Conclusion

Pollution in its different forms is a result of a disturbance of an ecological system, these disturbances can be caused from natural sources but they are mostly caused by anthropological activities. Hence, being able to avoid environmental degradation and air quality is the main objective of environmental management and protection organizations. It is crucial for them to be able to measure the effectiveness of their management policies and predict the results of their decisions. Simulation can be used as a decision aid that allows them to better estimate the results of their actions and decisions.

Although the state-of-the-art techniques have proved to be efficient in forecasting CO_2 emissions and solving distributed problems, these works suffer from some limitations, for instance, they focus only on forecasting CO_2 emissions and do not address the problem of forecasting the energy source that is responsible for these emissions. The presented work aimed to fill this gap by jointly forecasting the type of energy source and their equivalent CO_2 emissions. Moreover, another problem that has not been addressed is CO_2 emission reduction policies.

Therefore, in this thesis, how renewable energy can be used to reduce energy related CO_2 and fossil energy consumption is examined. The contribution consists of a collaborative MAS for forecasting hourly CO_2 emissions issued from energy consumption and production, where multiple ANN forecasting models are embedded in various forecasting agents and each model was trained to forecast a specific type of energy; in addition, a mathematical model was used to estimate traffic emissions. The results are promising, as the system was able to achieve satisfying prediction performances, and was able to give some estimations of the benefits of using renewable energies. Therefore, the work has demonstrated that an agent-based approach can be a suitable solution to model complex problems and, in particular pollution-related issues and energy management systems.

7.2 Perspectives

Like any research work, it has some limitations that we will outline in this section, as our work opens the door to possible improvements and topics that can be exploited.

Regarding the short-term forecasting, and despite the good results obtained by our forecasting models, the work can be enhanced by studying and exploring a bigger dataset whether by extending the available datasets or by integrating some auxiliary exogenous information sources (possibly correlated) such as other meteorological parameters. In most cases, adding new factors can improve the prediction accuracy.

Also, the simulator can be improved and extended by adding more pollution sources such as industrial activities, and improving the emissions calculations by adding more influential factors mainly in the gas consumption sector.

We can also extend the simulation to include more pollutants such as Greenhouse gases, $PM_{2.5}$, and even some secondary pollutants.

Finally, we intend to add more agents in our system and include decision making mechanisms to examine how the human decisions affect the air pollution levels and the energy sector.

A

Annex: List of publications

A.1 Journal publication

- Seif Eddine Bouziane, Mohamed Tarek Khadir, Julie Dugdale: A collaborative predictive multi-agent system for forecasting carbon emissions related to energy consumption. Multiagent and Grid Systems, 17, (2021), 39-58.

A.2 International conferences

- Seif Eddine Bouziane, Mohamed Tarek Khadir: Predictive agents for the forecast of CO_2 emissions issued from electrical energy production and gas consumption. 1st International Conference on Embedded Systems and Artificial Intelligence. Fez, Morocco, (2019).
- Khadir Mohamed Tarek, Seif Eddine Bouziane: Artificial Neural Networks modeling of electrical renewable energy both photovoltaic and wind for the region of Adrar Algeria. 4th International Conference on Artificial Intelligence in Renewable Energetic Systems. Tipaza, Algeria, (2020).
- Seif Eddine Bouziane, Julie Dugdale, Mohamed Tarek Khadir: Modeling renewable energy production and CO_2 emissions in the region of Adrar in Algeria using LSTM neural networks. 55th Hawaii International Conference on System Sciences. Hawaii, USA, (2022).

A.3 Book Chapter

- Seif Eddine Bouziane, Mohamed Tarek Khadir: Predictive agents for the forecast of CO_2 emissions issued from electrical energy production and gas consumption. Advances in Intelligent Systems and Computing, Springer. Singapore, 1076, (2020), 183-191.

Bibliography

(Abbasimehr et al. 2020)	H. Abbasimehr, M. Shabani, M. Yousefi: An optimized model using LSTM network for demand forecasting, Computers & industrial engineering,143 (2020).
(Abuella and Chowdhury 2015)	M. Abuella, B. Chowdhury: Solar Power Forecasting Using Artificial Neural Networks, North American Power Symposium (NAPS), Charlotte, USA, (2015).
(Agarwal and Kickhofer 2015)	A. Agarwal, B. Kickhofer: Agent-Based Simultaneous Optimization of Congestion and Air Polution: A Real-World Case Study. 4th International Workshop on Agent-based Mobility, Traffic and Transportation Models, Methodologies and Applications. London, England, (2015).
(Ahat et al. 2009)	M. Ahat, S. Ben Amor, M. Bui, M. Lamure, M. Courel: Pollution Modeling and Simulation with Multi-Agent and Pretopology, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, 4, (2009), 225-231.
(Ahmed et al 2020)	M. Ahmad, S.M. Adnan, S. Zaidi, P. Bhargava: A novel support vector regression (SVR) model for the prediction of splice strength of the unconfined beam specimens. Construction and building materials. Volume 248. 2020.
(Ainouche and Ainouche 2005)	A. Ainouche , H. Ainouche: Promotion of renewable energies in Algeria for a sustainable development and better future for next generations, Contribution of 18th world petroleum congress, (2005), Johannesburg, South Africa.
(An and Zhai 2020)	Y. An, X. Zhai: SVR-DEA model of carbon tax pricing for China's thermal power industry. Science of the total environment, 734 (2020).
(Athira et al. 2018)	V. Athira, P. Geetha, R. Vinayakumar, K.P. Soman, DeepAirNet: Applying Recurrent Networks for Air Quality Prediction, Procedia computer science,132, (2018), 1394-1403.
(Aulinas et al. 2009)	M. Aulinas, C. Turon, M. Sànchez-Marrè, Agents as a Decision Support Tool in Environmental Processes: The State of the Art, Whitestein Series in Software Agent Technologies and Autonomic Computing, (2009), 5–35

(Azid et al. 2014) A. Azid, H. Juahir, M. Toriman, M. Kamarudin, A. Saudi, C. Hasnam, N. Aziz, F. Azaman, M. Latif, S. Zainuddin, M. Osman, M. Yamin: Prediction of the level of air pollution using principal component analysis and artificial neural network techniques: A case study in Malaysia. Water, Air, and Soil Pollution, 225, (2014).

(Ballas et al. 2005) D. Ballas, D. Dorling, B. Thomas, D. Rossiter: Geography matters: simulating the local impacts of national social policies, (2005).

(Barker et al. 1963) K. Barker, F. Cambi, E. Catcott, Chambers, A. Leslie, E. Halliday: La pollution de l'air. 1963.

(Bellifemine et al. 1999) F. Bellifemine, A. Poggi, G. Rimassa: JADE - A FIPA-compliant agent framework, Jade, 99, (1999).

(Bengino et al. 1994) Y. Bengio, P. Simard, P. Frasconi: Learning long-term dependencies with gradient descent is difficult. IEEE Transactions on Neural Networks, 5 (1994), 157-166.

(Bhaskar and Singh 2012) K. Bhaskar, S.N. Singh: AWNN-Assisted Wind Power Forecasting Using Feed-Forward Neural Network. IEEE Transactions on Sustainable Energy, 3, (2012), 306-315.

(Bonabeau 2002) E. Bonabeau: Predicting the unpredictable. Harvard Business Review, 80 (2002), 109-116.

(Boser et al. 1992) B.E. Boser, I.M. Guyon, V.N. Vapnik: A Training Algorithm for Optimal Margin Classifiers. 5th Annual Workshop on Computational Learning Theory, USA (1992). 144-152.

(Bouraiou et al. 2016) A. Bouraiou, A. Necaibia, N. Boutasseta, S. Mekhilef, R. Dabou, A. Ziane, N. Sahouane, I. Attoui, M. Mostefaoui, O. Touaba: Status of Renewable Energy Potential and Utilization in Algeria, Journal of cleaner production, 246, (2016)

(Box and Jenkins 1976) G.E. Box and G.M. Jenkins: Time Series Analysis: Forecasting and Control, Holden-Day, San Francisco, (1976).

(Box and Jenkins 1994) G.E. Box and G.M. Jenkins: Time Series Analysis: Forecasting and Control, Prentice Hall, Englewood Cliffs, 3 (1994).

(Bratman 1987) M.E. Bratman: Intention, Plans, and Practical Reason, (1987) Cambridge.

(Brooks 1986) R. Brooks: A robust layered control system for a mobile robot. IEEE Journal of Robotics and Automation, 2 (1986), 253-262.

(Brooks 1991) R.A. Brooks: Intelligence Without Reason. Proceedings of 12th Int. Joint Conference on Artificial Intelligence. Sydney, Australia, (1991)

(Corchado et al. 2009) M.J. Corchado, A. Mata, S. Rodriguez: A Multi-Agent System for Modeling and Monitoring the Evolution of Oil Slicks in Open Oceans, Advanced Agent-Based Environmental Management Systems, Basel, (2009), 91-117.

(Di Lecce et al. 2009) V. Di Lecce, A. Amato, C. Martines, P. di Bari, R. Dario, Air quality control for health care centres: The application of an intelligent distributed system, IEEE Workshop Environmental, Energy, and Structural Monitoring Systems, (2009).

(Ding et al. 2017) S. Ding, Y.G. Dang, X.M. Li, J.J. Wang, K. Zhao: Forecasting Chinese CO2 emissions from fuel combustion using a novel grey multivariable model, Journal of Cleaner Production, 162, (2017), 1527-1538.

(Dragomir and Oprea 2013) E.G. Dragomir, M. Oprea: A Multi-Agent System for Power Plants Air Pollution Monitoring, IFAC Proceedings Volumes, 46, 2013, 89-94.

(Duan et al. 2018) H. Duan, Y. Huang, R.K. Mehra, P. Song, F. Ma: Study on influencing factors of prediction accuracy of support vector machine (SVM) model for NOx emission of a hydrogen enriched compressed natural gas engine. Fuel, 234 (2018), 954-964.

(EIA 2021) https://www.eia.gov/tools/faqs/

(El Fazziki et al. 2015) A. El Fazziki, A. Sadiq, J. Ouarzazi, M. Sadgal: A Multi-Agent Framework for a Hadoop Based Air Quality Decision Support System. Conference on Advanced Information Systems Engineering. Stockholm, Sweden, (2015).

(El Fazziki et al. 2017) A. El Fazziki, D. Benslimane, A. Sadiq, J. Ouarzazi, M. Sadgal: An Agent Based Traffic Regulation System for the Roadside Air Quality Control, in IEEE Access, 5, (2017), 13192-13201.

(Eldho et al. 2020) B. Eldho, S. Manmadhan, S. Jacob, J.G. Hod: Real Time Dehazing System for Automobiles, International Journal of Computer Trends and Technology, 68 (2020).

(Emery et al. 2020) J. Emery, M. Nicolas, N. Martiny, T. Thevenin: The SCAUP model: multi-agent simulation from Urban sensors for traffic air pollution. Cybergo. 944, (2020).

(EPA 2020) https://www.epa.gov/energy/greenhouse-gases-equivalencies-calculator-calculations-and-references

(Epstein and Axtell 1999) J.M. Epstein, R. Axtell: Agent-Based Computational Models and Generative Social Science, John Wiley & Sons, 4 (1999).

(Fang et al. 2018) D. Fang, X. Zhang, Q. Yu, T.C. Jin, L. Tian: A novel method for carbon dioxide emission forecasting based on improved Gaussian processes regression. Journal of Cleaner Production, 173, (2018), 143-150.

(Farfar 2018) K.E. Farfar: Apprentissage approfondi des réseaux de neurones artificiels pour l'approximation de séries temporelles. PhD Thesis, University of Badji Mokhtar, (2018).

(Farfar and Khadir 2018) K.E. Farfar, M.T. Khadir: A two-stage short-term load forecasting approach using temperature daily profiles estimation. Neural Computing and Applications. 31, (2018).

(Ferber and Perrot 1995) J. Ferber, J. Perrot: Les systèmes multi-agents : vers une intelligence collective. InterEditions, (1995).

(FIPA 1999) http://www.fipa.org/specs/fipa00061/SC00061G.html

(Gao et al. 2018) Y. Gao, G. Liu, M. Casazza, Y. Hao, Y. Zhang, B. Giannetti: Economy-pollution nexus model of cities at river basin scale based on multiagent. Ecological modelling, 379 (2018), 22-38.

(Gers et al. 2000) F.A. Gers, A. Schmidhuber, F.A. Cummins: Learning to Forget: Continual Prediction with LSTM. Neural Computing. 12 (2000), 2451-2471.

(Ghasemlounia and Saghebian 2021) R. Ghasemlounia, S.M. Saghebian: Uncertainty assessment of kernel-based approaches on scour depth modeling in downstream of ski-jump bucket spillways, Water supply, 21 (2021).

(Ghazi 2017) S. Ghazi: Approche Multi-agents pour la simulation de la pollution de l'air. PhD Thesis, University of Badji Mokhtar, (2017).

(Ghazi et al. 2016) S. Ghazi, J. Dugdale, M.T. Khadir: Modelling Air Pollution Crises Using Multi-Agent Simulation, 49th Hawaii International Conference on System Sciences (HICSS), Koloa, HI, (2016), 172-177

(Grouls 2013) A. Grouls : agents et systèmes multi-agents : Vers une synthèse de ces concepts. Thesis, University of Quebec, (2013).

(Hassan et al. 2022) M. Hassan, S. Mostafa, A. Mustapha, M. Saringat, B. Al-rimy, F. Saeed, A. Eljialy, M. Jubair: A New Collaborative Multi-Agent Monte Carlo Simulation Model for Spatial Correlation of Air Pollution Global Risk Assessment. Sustainability, 14 (2022).

(Hochreiter et al. 1997) S. Hochreiter, J. Schmidhuber: Long short-term memory. Neural computation, 9 (1997), 1735-1780.

(Hsu et al. 2018) Y.Y. Hsu, T.T. Tung, H.C. Yeh, C.N. Lu: Two-Stage Artificial Neural Network Model for Short-Term Load Forecasting, IFAC-PapersOnLine, 51, (2018), 678-683.

(Huang et al. 2019) Y. Huang, L. Shen, H. Liu: Grey relational analysis, principal component analysis and forecasting of carbon emissions based on long short-term memory in China. Journal of Cleaner Production, 209, (2019), 415-423.

(Hulsmann et al. 2014) F. Hulsmann, R. Gerike, M. Ketzel: Modelling traffic and air pollution in an integrated approach – the case of Munich, Urban Climate, 10, (2014), 732-744.

(IEA 2016) International Energy Agency: Energy and air pollution. 2016.

(Jarras and chaib-draa 2002) I. Jarras, B. Chaib-draa : Aperçu sur les systèmes multi agents, CIRANO Working Papers, (2002).

(Jennings 2000) N.R. Jennings: On agent-based software engineering. Artificial intelligence, 117 (2000), 277-296.

(Jetcheva et al. 2014) J.G. Jetcheva, M. Majidpour, W.P. Chen: Neural network model ensembles for building-level electricity load forecasts, Energy and Buildings, 84, (2014), 214-223.

(Kalman 1960) R. Kalman: A New Approach to Linear Filtering and Prediction Problems. Journal of Basic Engineering, (1960), 35-45.

(Khadir and Bouziane 2020) M.T. Khadir, S.E. Bouziane: Artificial Neural Networks modelling of electrical renewable energy both photovoltaic and wind for the region of Adrar Algeria. 4th International Conference on Artificial Intelligence in Renewable Energetic Systems. Tipaza, Algeria, (2020).

(Khraief et al. 2018) N. Khraief, M. Shahbaz, H. Mallick, N. Loganathan: Estimation of electricity demand function for Algeria: Revisit of time series analysis. Renewable and Sustainable Energy Reviews, 82 (2018), 4221-4234.

(Kingma and Ba 2015) D.P. Kingma, J. Ba: Adam: A Method for Stochastic Optimization. International Conference on Learning Representations (ICLR). (2015), San Diego, USA.

(Kohonen 1999) T. Kohonen: The self-organizing map, Proceedings of the IEEE, 78, (1990), 1464 – 1480.

(Kolen and Kremer 2001) J.F. Kolen, S. Kremer: Gradient flow in recurrent nets: the difficulty of learning long-term dependencies. A Field Guide to Dynamical Recurrent Networks, (2001).

(Laib et al. 2016) Laib O., Khadir M.T., Chouired L.: Forecasting yearly natural gas consumption using Artificial Neural Network for the Algerian Market. 2016 4th International Conference on Control Engineering & Information Technology (CEIT), Hammamet, Tunisia. 2016, pp. 1-5.

(Lau et al. 2014) E.T. Lau, Q. Yang, A.B. Forbes, P. Wright, V.N. Livina: Modelling carbon emissions in electric systems. Energy Conversion and Management, 80, (2014), 573-581.

(Lestel et al. 1994) D. Lestel, B. Grison, A. Drogoul: Les agents réactifs et le vivant dans une perspective d'évolution coopérative. Intellectica, 19 (1994), 73-90.

(Li et al. 2012) W. Li, J. Xie, J. Li and H. Shen: Three Dimensional Simulation of Basin Water Pollution Incidents Based on Multi-agent and Grid Technology, Fifth International Symposium on Computational Intelligence and Design, Hangzhou, China, (2012), 470-473.

(Lin et al. 2020) Y. Lin, S. Lee, C. Ouyang, C. Wu: Air quality prediction by neuro-fuzzy modeling approach, Applied soft computing, 86, (2020).

(Ma et al. 2017) L. Ma, K. Lin, M. Guan, M. Lin: The prediction of carbon emission in all provinces of China with the K-means cluster based Logistic model, 2017 International Conference on Service Systems and Service Management, Dalian, (2017), 1-6.

(Ma et al. 2020) J. Ma, Z. Li, J. Cheng, Y. Ding, C. Lin, Z. Xu: Air quality prediction at new stations using spatially transferred bidirectional long short-term memory network. Science of the total environment, 705, (2020).

(Macal and North 2010)	C.M. Macal, M.J. North: Tutorial on agent-based modelling and simulation. Journal of simulation, 4(2010), 151-162.
(Maciag et al. 2019)	P.S. Maciag, N. Kasabov, M. Kryszkiewicz, R. Bembenik: Air pollution prediction with clustering-based ensemble of evolving spiking neural networks and a case study for London area, Environmental modelling & software, 118, (2019), 262-280.
(Marugan et al. 2018)	A. Marugan, F. Marquez, J. Perez, D. Hernandez: A survey of artificial neural network in wind energy systems. Applied Energy, 228, (2018), 1822-1836.
(Mason et al. 2018)	K. Mason, J. Duggan, E. Howley: Forecasting Energy Demand, Wind Generation and Carbon Dioxide Emissions in Ireland using Evolutionary Neural Networks, Energy, 155, (2018), 705-720.
(Mayer 1999)	H. Mayer : Air pollution in cities. Atmospheric Environment, 33 (1999), 4029-4037.
(McCulloch and Pitts 1943)	W. S. McCulloch, W. Pitts: A logical calculus of the ideas immanent in nervous activity. The Bulletin of Mathematical Biophysics, 5 (1943), 115-133.
(MEN)	https://www.energy.gov.dz/
(Minsky & Papert 2017)	M. Minsky, S.A. Papert: Perceptrons: An introduction to computational geometry. MIT press, (2017).
(Mohsen and Abbassi 2020)	R. Mohsen, B. Abbassi: Prediction of greenhouse gas emissions from Ontario's solid waste landfills using fuzzy logic-based model. Waste Management, 102 (2020). 743-750.
(Moustris et al. 2013)	K.P. Moustris, I.K. Larissi, P.T. Nastos, K. Koukoulestos: Development and Application of Artificial Neural Network Modeling in Forecasting PM10 Levels in a Mediterranean City, Water, Air, and Soil Pollution, 224, (2013).
(Nachet 2014)	B. Nachet: Modèle multi-agents pour la conception des systèmes d'aide à la décision collective. PhD thesis. University Oran 1, (2014).
(Ni and Zhang 2009)	J. Ni, C. Zhang and L. Ren: An Intelligent Decision Support System of Lake Water Pollution Control Based on Multi-agent Model, International Conference on Computational Intelligence and Security, Beijing, China, (2009), 217-221.

(Nwana 1996) H. Nwana: Software agents: An overview. The knowledge engineering review, 11(1996), 205-244.

(Oder et al. 1993) C. Oder, H. Haasis, O. Rentz: Development of a fuzzy energy-emission model. 2nd IEEE International Conference on Fuzzy Systems. USA. 1993.

(ONS 2018) https://www.ons.dz/. Last accessed: 07/10/2021.

(Oudrane et al. 2019) A. Oudrane, A. Benaoumeur, M. Hamouda : 2019, Étude et Analyse du Gisement Solaire Pour une Région Désertique : Type Adrar, Communication Science & technology, 21, (2019), 683 – 695.

(Papaleonidas and Iliadis 2012) A. Papaleonidas, L. Iliadis: Hybrid and Reinforcement Multi Agent Technology for Real Time Air Pollution Monitoring. Artificial Intelligence Applications and Innovations, 381, (2012), 274-284.

(Peng et al. 2016) X. Peng, L. Xiong, J. Wen, Y. Xu, W. Fan, S. Feng, B. Wang: A very short-term wind power prediction approach based on Multilayer Restricted Boltzmann Machine, IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), Xian, China, (2016).

(Rosenblatt 1957) F. Rosenblatt: The Perceptron, a Perceiving and Recognizing Automaton Project Para, Report: Cornell Aeronautical Laboratory, Cornell Aeronautical Laboratory, (1957).

(Rumelhart et al. 1988) D.E. Rumelhart, G.E. Hinton, R.J. Williams: Learning representations by back-propagating errors. Cognitive modelling, 5 (1988), 1.

(Russell and Norvig 1995) S.J. Russell, P. Norvig: Artificial Intelligence “A modern approach”. Prentice-Hall, 25 (1995).

(Russell et Norvig 2006) S. Russel, P. Norvig : Intelligence artificielle, 2 (2006).

(Safdarnejed et al. 2019) S.M. Safdarnejed, J.F. Tuttle, K.M. Powell, Dynamic modeling and optimization of a coal-fired utility boiler to forecast and minimize NOx and CO emissions simultaneously, Computers & chemical engineering, 124, (2019), 62-79.

(Sangeetha and Amudha 2018) T. Sangeetha, T. Amudha: A Novel Bio-Inspired Framework for CO2 Emission Forecast in India, Procedia computer science, 125, (2018), 367-375.

(Sheta et al. 2016) A.F. Sheta, N. Ghatasheh, H. Faris: Forecasting Global Carbon Dioxide Emission Using Auto-Regressive with eXogenous Input and Evolutionary Product Unit Neural Network Models. 6th International Conference on Information and Communication Systems (ICICS), Amman, Jordan, (2015), 182-187.

(Shoham 1991) Y. Shoham: AgentO: A simple agent and its interpreter. Proceedings of the ninth national conference on artificial intelligence, (1991).

(Shoham 1993) Y. Shoham: Agent-oriented programming. Artificial intelligence, 60(1993), 51-92.

(Taneja et al. 2016) S. Taneja, N. Sharma, K. Oberoi, Y. Navoria: Predicting trends in air pollution in Delhi using data mining. 1st India International Conference on Information Processing. India, (2016).

(Vidyanandan 2017) K.V. Vidyanandan: An Overview of Factors Affecting the Performance of Solar PV Systems. Energy Scan, 27, (2017), 2-8.

(Wang and Song 2018) J. Wang, G. Song: A Deep Spatial-Temporal Ensemble Model for Air Quality Prediction. Neurocomputing, 314, (2018), 198-206.

(Wang and Ye 2017) Z.X. Wang, D.J. Ye: Forecasting Chinese carbon emissions from fossil energy consumption using non-linear grey multivariable models, Journal of Cleaner Production, 142, (2017), 600-612.

(Wang and Zhang 2013) J. Wang, J. Zhang: Research on carbon emissions and its trend forecast of Baoding. Sixth International Conference on Advanced Computational Intelligence (ICACI). Hangzhou, China, (2013), 39-44.

(WHO) www.who.int/health-topics/air-pollution. Last accessed: 27/06/2020.

(Wooldridge 1999) M. Wooldridge: Intelligent Agent. The MIT Press, Cambridge, (1999)

(Wooldridge and Jennings 1995) M. Wooldridge, N.R. Jennings, Intelligent agents: Theory and practice, The knowledge engineering review, 10, (1995), 115-152.

(Wu et al. 2016) W. Wu, K. Chen, Y. Qiao, Z. Lu, Probabilistic short-term wind power forecasting based on deep neural networks, International Conference on Probabilistic Methods Applied to Power Systems (PMAPS). Beijing, China, (2016).

(Wu and Lin 2019)	Q. Wu, H. Lin: A novel optimal-hybrid model for daily air quality index prediction considering air pollutant factors. Science of the total environment, 683, (2019), 808-821.
(Xu et al. 2019)	N. Xu, S. Ding, Y. Gong, J. Bai: Forecasting Chinese greenhouse gas emissions from energy consumption using a novel grey rolling model, Energy, 175, (2019), 218-227.
(Zhong et al. 2018)	J. Zhong, L. Liu, Q. Sun, X. Wang: Prediction of Photovoltaic Power Generation Based on General Regression and Back Propagation Neural Network, Energy Procedia, 152, (2018), 1224-1229.
(Zhou et al. 2019)	Y. Zhou, X. Zhao, K. Lin, C. Wang, L. Li: A Gaussian process mixture model-based hard-cut iterative learning algorithm for air quality prediction, Applied soft computing, 58, (2019).