REPUBLIQUE ALGERIENNE DÉMOCRATIQUE ET POPULAIRE

MINISTÈRE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

جامعة باجي مختار عنابة

UNIVERSITE BADJI MOKHTAR –ANNABA

BADJI MOKHTAR UNIVERSITY –ANNABA

Faculté des Sciences de la Terre

Département de Géologie

Laboratoire des Ressource en Eaux et Développement Durable

THÈSE

Présentée en vue de l'obtention du diplôme de Doctorat LMD en Hydrogéologie

Option : Hydrogéologie

UTILISATION DES TECHNIQUES NUMERIQUES POUR L'OPTIMISATION DE LA GESTION INTEGREE DE L'EAU

Présentée par :

ZEROUAL Menal

Devant le jury :

Président :	DJABRI Larbi	Pr. Université Badji Mokhtar - Annaba
Directeur de thèse :	HANI Azzedine	Pr. Université Badji Mokhtar - Annaba
Examinateurs :	CHAFFAI Hicham	Pr. Université Badji Mokhtar - Annaba
	HADJI Riheb	M.C.A Université Ferhat Abbas - Sétif
	SHAHROUR Isam	Pr. Université de Lille - France
	MAOUI Ammar	Pr. Université 8 mai 1945- Guelma

Année : 2020

Remerciements

Avant tout je tiens à remercier le bon dieu qui m'a aidé à l'élaboration de ce travail.

Je tiens à remercier particulièrement professeur HANI Azzedine, mon directeur de thèse, pour ses conseils. Pour m'avoir fait profiter de son expérience et pour m'avoir prodigué des conseils sans les quels je n'aurais pu éviter des égarements qui auraient nui au cheminement de cette étude ; pour l'honneur qu'il me fait en acceptant de juger ce travail. Soyez assuré, Monsieur, de toute mon estime et de mon profond respect.

Je tiens tout spécialement à remercier Mr BOUSTILA Amir, de son soutien et son aide inestimable qui m'ont accompagné lors des moments difficiles et sans lequel ce travail n'aurait jamais vu le jour.

Un grand merci aussi à tout le personnel d'ADE de Souk Ahras pour m'avoir lors de l'élaboration de cette étude.

Mes remerciements vont également à l'ensemble des personnes que j'ai croisé au cours des années passées au sein du département de géologie.

Enfin, je voudrais adresser toute ma tendresse à mes parents et mon mari dont l'amour inconditionnel m'a permis de remonter la pente lors des moments difficiles.

Dédicace

À mes chers parents qui ont toujours été là pour moi.

À mon époux qui m'a toujours supporté et mon fils Lokmane.

À mes frères et sœurs pour leurs encouragements, je vous dis Merci.

À mes vrais amis.

Menal ZEROUAL

ABSTRACT

The water resources and demand are inversely related. Such link endangers human lives and represents an important concern in cities planning. The effect of human water usage behavior is clear and already investigated on global scale, but, detailed studies dealing with particularities of household water consumption are still lacking.

The water scarcity is a worldwide issue but its consequences vary dramatically from nation to another especially in semi-arid areas and Algeria no exception .For that raison, Sedrata city, Souk-Ahras, North-east of Algeria is subjected to current study. The parameters governing water consumption are chosen basing on literature review and obtained by a public inquiry for five years period 2012-2017.

In order to define the main determinants of water use, the present thesis adopted two main paths, a conventional statistical analysis dealing with data preparation and preprocessing, and the second dealing with applicability and results of both Artificial Neural Networks and Adaptive Neuro Fuzzy Inference System.

The current thesis and with help of ANNs and ANFIS highlighted twelve explanatory variables, of which nine are socio-economic parameters and three physical characteristics of building units. Another major finding is that the artificial neural networks can be trained to predict water consumption with the correlation coefficient for training, testing and validation phases are about 0,99. In addition, results demonstrate that the combination of socio-economic parameters with physical characteristics of building units enhances the assessment of household water consumption.

Key words:

Artificial neural networks, adaptive neuro fuzzy inference system, Determinants of household water use, Sedrata city, Water consumption

RÉSUMÉ

Les ressources et la demande en eau sont inversement liées. Tel lien met en danger des vies humaines et représente une préoccupation importante dans la planification des villes. L'effet du comportement d'utilisation humain de l'eau est clair et déjà étudié à l'échelle mondiale, mais des études détaillées traitant des particularités de la consommation d'eau des ménages sont encore manquantes.

La pénurie d'eau est un problème mondial mais ses conséquences varient considérablement d'une nation à l'autre surtout dans les zones semi-arides et l'Algérie ne fait pas exception. Pour cette raison, la ville de Sedrata, Souk-Ahras, au nord-est de l'Algérie fait l'objet de l'étude. Les paramètres affectant la consommation d'eau de ménage sont choisis sur la base d'une recherche bibliographique et récoltés par une enquête durant la période 2012-2017.

Afin de définir les principaux déterminants de l'utilisation de l'eau de ménage, la présente thèse a suivis deux chemins principaux, une analyse statistique conventionnelle de la préparation et du prétraitement des données, et le second traitant l'applicabilité et les résultats des réseaux de neurones artificiels et du système d'inférence neuro-fuzzy adaptatif.

La présente thèse et avec l'aide des RNA et de l'ANFIS ont mis en évidence douze variables explicatives, dont neuf sont des paramètres socio-économiques et trois caractéristiques physiques des logements. Les résultats démontrent la capacité d'apprentissage des réseaux de neurones artificiels pour prédire la consommation d'eau avec un coefficient de corrélation pour les phases de calculs d'environ 0,99. De plus, la combinaison des paramètres socio-économiques avec les caractéristiques physiques des logements améliore l'évaluation de la consommation d'eau des ménages.

Mots clé

Réseaux de neurones artificiels, system d'inférence adaptive neuro floue, déterminants d'eau de ménage, Sedrata, consommation d'eau

ملخص

يعد توفير الحاجات الكافية من المياه العذبة من أكبر المشاكل التي تواجه البشرية خاصة في تخطيط المدن، خاصة مع شح الموارد المائية و التغير المناخي. يوجد العديد من الدراسات التي تطرقت لتأثير البشر السلبي بشكل عالمي، لكن هذه الدراسات لم تتناول سلوكيات البشر الاستهلاكية بالتفصيل.

على الرغم من عالمية مشكل ندرة المياه إلا أن أثاره متفاوتة النتائج من منطقة إلى أخرى خاصة المناطق شبه الجافة و التي تعاني أصلا . لهذه الأسباب تم اختيار مدينة سدراتة بولاية سوق أهراس الواقعة في الشمال الشرقي للجزائر كنموذج للدراسة الحالية. لغرض الاحاطة بكامل متغيرات استهلاك الماء ، تم التطرق لكامل الدراسات السابقة بالموضوع و اعتمادا على الاستبيان قمنا بجمع كل المعلومات الخاصة بسلوكيات المستهلكين خلال المدة 2012-2017 .

من أجل تحديد العوامل المتحكمة في استخدام المياه ، اعتمدنا في الأطروحة مسارين رئيسيين ،الأول تحليل إحصائي تقليدي يتعامل مع إعداد البيانات ومعالجتها المسبقة ، والثاني يتعامل مع قابلية تطبيق ونتائج كل من الشبكات العصبونية الاصطناعية ونظام الاستدلال العصبي الضبابي.

أبرزت الأطروحة الحالية وبمساعدة كلا المقاربتين اثني عشر متغيرًا توضيحيًا، منها تسعة معايير اجتماعية اقتصادية وثلاث خصائص فيزيائية للمساكن. نتيجة رئيسية أخرى هي أنه يمكن تدريب الشبكات العصبية العصبونية على التنبؤ باستهلاك المياه مع معامل احصائي مرتفع يصل إلى حوالي 0،99. بالإضافة إلى ذلك، توضح النتائج أن الجمع بين المعايير الاجتماعية والاقتصادية مع الخصائص الفيزيائية للمساكن يعزز دراسات تقييم استهلاك المياه في المنزل.

المصطلحات المفتاحية

الشبكات العصبونية الاصطناعية، الأنظمة العصبية الضبابية ،عوامل استهلاك الماء، مدينة سدراتة ، استهلاك الماء.

TABLE OF CONTENTS

List of Abbreviations

List of Tables

List of Figures

General Introduction

Chapter 01: Literature Review

Chapter 02: Study area

Chapter 03: Data Collection

Chapter 04: Analysis Methodology and Tools

Chapter 05: Results and discussion

Part I: Regression Analysis, ANOVAs Test, Principal Component Analysis, Cluster Analysis and Factor Analysis

Part II: Application of Artificial Neural Network (ANNs), Adaptive Neuro Fuzzy Inference System (ANFIS) to predict water consumption

LIST OF ABBREVIATIONS

ADE	Algérienne des eaux
AG1	Category of age 1 (Under 8 years old)
AG2	Category of age 2 (Between 9 to 15 years old)
AG3	Category of age 3 (Between 15 to 35 years old)
AG4	Category of age 4 (Older than 35 years old)
ANFIS	Adaptive Neuro Fuzzy Inference System
ANNs	artificial neural networks
ANOVAs	analysis of variance
BAR	Building Area
CA	Cluster Analysis
CARN	Number of Cars
CLF	Climatic Factors
e (t)	forecast error
FA	Factor Analysis
FEM	Number of Female
FAO	Food and Agriculture Organization
FIS	Fuzzy Inference System
FL	Fuzzy logic
FSHW	Female takes shower
GAR	Garden Area
GEN	Gender
GIS	geographic information system
GRNN	General Regression Neural Network
GWAT	Number of Garden Watering
HGS	High School
HOUS	Household size
INC	Household Income
INH	Indoor Habits
IWRM	Integrated Water Resources Management
MAE	Mean Absolute Error
MAL	Number of Male
MLP	Multilayer Perceptron Neural Networks
MDS	Medium School
ME	Mean Error
MF	membership functions
MSE	Mean square error
MSHW	Male takes shower
NWSAS	North-Western Sahara Aquifer System
PCA	Principal Component Analysis
PHC	Physical Characteristics of Buildings Units
R	Correlation coefficient
PHC	Physical Characteristics
PRE	Precipitation
PRS	Primary School
RMSE	Root Mean Square
ROMN	Number of Rooms
SEP	Socio-economic parameters
SPSS	Statistical Package of Social Sciences
TAR	Total Area
TEM	Temperature
UNIV	University
UTLT	Using Toilets (day)
WCAR	Number of washing Car
WCL	Washing Clothes
WDISH	Washing Dish
WCP	Water Consumption

LIST OF TABLES

LIST OF FIGURES

General Introduction

1. Research context

Africa in general is highly threatened by water scarcity due to non-organized population growth, unstable politics and global warming; and Algeria is not an exception even with its northern location. Contradictory to other African nations, the northern coast (Mediterranean African countries) are politically stable but still vulnerable to thirst. A detailed study on water use does not exist. For that, a study area like Sedrata would reveal a lot on water consumption behavior not only in the city but the whole country by extrapolation.

The fact that the world is experiencing sustained population growth. Today the world population stands at 7.3 billion people and is expected to reach 8.1 billion by 2025, according to the United Nations report "World Population Prospects: 2012 Revision" (UN, 2015). The damage inflicted on the environment under the weight of the demographic explosion that humanity has known since the nineteenth (XIXth) century, as well as the economic expansion that accompanies it, some will not be able to help but wonder about the future of the environment and the planet.

Hardin (1968) gives an alarmist vision of this situation in a communication under the title team of "tragedy of the commons". This concern is all the more legitimate as it raises the question of supplying future populations with the resources necessary for their survival, in the first-place food (Brown L.R, 1963) and of our responsibility towards their fate (Jonas, 1985).

Demographic and economic growth, has as a corollary another major phenomenon which has also gained more and more importance since the middle of the last century, namely the urbanization of the population. In fact, more than half of the world's population, according to World Bank estimates (World Bank Group, 2015), live in urban settlements, and if we believe the United Nations forecasts published in 2008, the global urbanization rate will reach nearly 60% in 2030 and will be around 70% in 2050. A phenomenon considered worrying for the environment, even if it is difficult to distinguish the impacts linked to urban growth from those linked to demographic and economic growth. the fact remains that cities are the engine of economic development and demographic growth.

Populations and activities are constituting the main place of energy consumption, resources, space and the emission of pollutants and waste. The urban sprawl that characterizes many cities leads to a decrease in cultivable land and the loss of biodiversity. According to the FAO (Food and Agriculture Organization), between 1995 and 2030, 100 million hectares of land will be lost for the benefit of the construction of housing and infrastructure necessary for the growth of the urban population.

The environmental effects of population growth in cities, as well as their repercussions on the quality of life and health of populations are more accentuated in poor or developing countries, because in these countries the authorities often do not have sufficient resources to face the problems generated by urban growth. for example, every year 1.8 million children under 5 die for lack of adequate access to drinking water (Planetoscope, 2012).

Access to drinking water is a major concern in several countries around the world where the survival of the human being depends entirely on it. In human kind history, freshwater resources were the major factor that population based on to choose a city's location. After industrialization, these locations are chosen based on working needs that added more pressure on water resources. Worse still, excessive exploitation of water resources endangers their sustainability by hampering the process of natural groundwater regeneration.

According to FAO (2007), by 2025, 1.8 billion people will live in countries or regions with acute water scarcity, and two-thirds of the world's population will be under water stress conditions, with all the consequences that involving for the health, the living environment of the populations, as well as their socio-economic development (Naimi et al., 2016).

According to water scarcity map, the forty-five countries were grouped into three basic categories of water scarcity.

Group I represents countries that face physical water scarcity in 2025. This means that, even with the highest feasible efficiency and productivity of water use, these countries do not have sufficient water resources to meet their agricultural, domestic, industrial and environmental needs in 2025. Indeed, many of these countries cannot even meet their present needs. Their only options are to invest in expensive desalinization plants.

Group II represents countries that face economic water scarcity in 2025. They have sufficient water resources to meet the 2025 needs but they will have to increase water supplies through additional storage, conveyance and regulation systems by 25 percent or more over the 1995 levels to meet their 2025 needs. Many of these countries face severe financial and development capacity problems in meeting their water needs.

Group III consists of countries that have no physical water scarcity and that will need to develop less than 25 percent more water supplies to meet their 2025 needs. In most cases, this will not pose a substantial problem for them. In fact, several countries in this group could actually decrease their 2025 water supplies from the 1995 levels because of increased water productivity (IWMI).

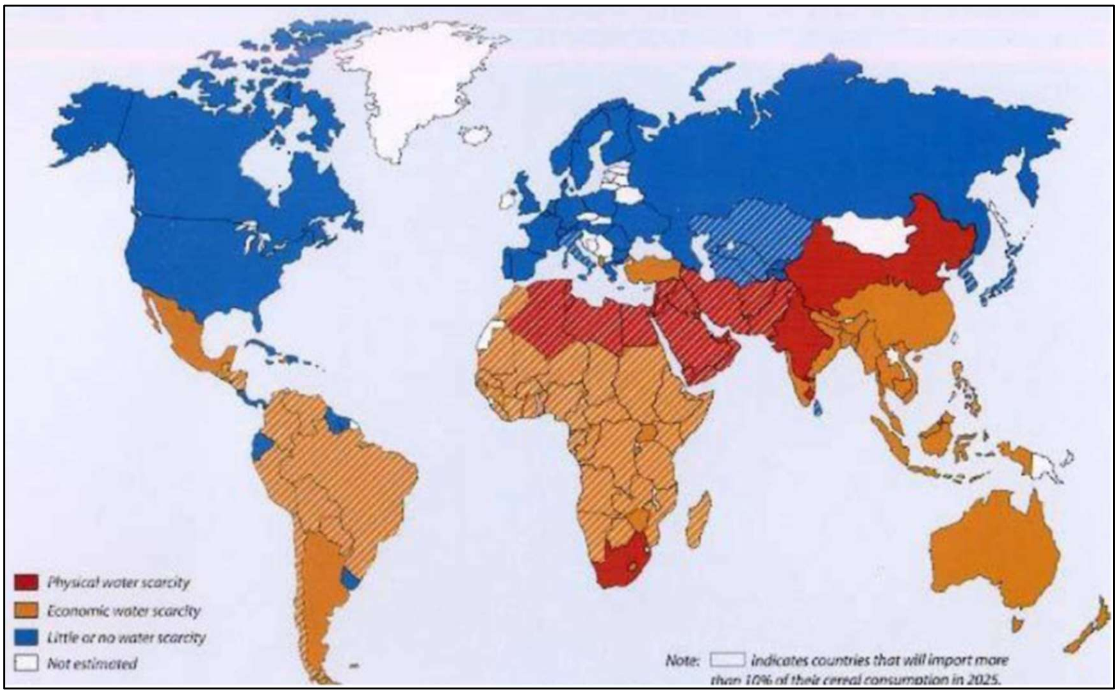


Fig.1 Projected Water Scarcity in 2050 (IWMI: International Water Management Institute)

Moreover, the issue of securing the water supply and the sustainable management of this resource must therefore be included among the priorities of any development strategy. As such, Algeria hosted in Algiers in March 2015 the first ministerial conference of the *5 + 5 Dialogue* on water to define the new water strategy in the Western Mediterranean (Delegation of the European Union in Algeria, 2015). the emphasis was on adopting a common vision to address water issues in order to ensure sustainable development of water resources in this region characterized, especially in the south shore, by a scarcity of water (Naimi *et al.,* 2016).

Until now, little importance has been given to water consumption in Algeria, and no detailed studies exist on household water consumption determinants. A better understanding of factors affecting water use and consumers behaviors will help to satisfy water demand for future planning. The present thesis considers the city of Sedrata, Northeast of Algeria, to assess domestic water use.

2. Problem statement

Sedrata city is repetitive case study of semi-arid Algerian areas, where water supply is not stable. Domestic water is a vital necessity and knowing the determinants will help distribution efficiency and by consequence life quality.

The main goal of statistics is finding parameters governing a phenomena and Water Consumption is no exception. To assess a possible linear association between two variables, correlation analyses is often used thanks to simplicity both to calculate and to interpret ,but in more complex cases such as water consumption more performing methods are required.

The current thesis examines the potential impacts of socio-economic parameters, physical characteristics of housing units and indoor habits of inhabitants on domestic water consumption. The study is performed by three main approaches Classical approach, Artificial Neural Networks and Adaptive Neuro Fuzzy Inference System where their efficiency is also tested.

3. Research objectives

The main aspects under investigation in this research can be started as:

- A descriptive analysis of the present water consumption in Sedrata city.
- To assess the impacts of socio-economic parameters, physical characteristics of housing units and indoor habits of residents on domestic water consumption.
- Explore the principal determinants of water consumption at the household scale.
- To choose the best Artificial neural networks and Adaptive neuro fuzzy inference system models.
- To enhance the water distribution of Sedrata and Souk Ahras by giving the priority to find solutions for the most influential item on domestic water consumption.

4. Thesis structure

The thesis is organized in six chapters with general introduction and general conclusion, as following:

“General Introduction”
The introductory part contains the background of the study, summary on the problem statement, along with research objectives and structure of the research.

Chapter one: “Literature Review”

The second chapter includes a survey of the literature reviews, information related to factors affecting urban water consumption investigated by researchers around the world, the statistical and numerical techniques used by authors to determine the principal factors influencing domestic water use and predict household water consumption.

Chapter Two: “Study Area”

It describes the study area geographically with briefing about its water resources and crisis, piped water supply system and maintenance in Sedrata, historical metrological data analysis.

Chapter Three: “Data Collection”

Chapter three presents brief introduction about the type and the source of collected data, detailed information about domestic water consumption in the study area and the different parameters collected by the questionnaire paper.

Chapter Four: “Analysis Methodology and Tools”

It deals with the methodology used to achieve the objectives of the study, starting from investigating the determinants of household water consumption by using statistical techniques and different tools, ArcGIS, STATISTICA 8, SPSS 19 and MATLAB Software to achieve the goal of this research work.

Chapter Five: “Results and Discussion”

The following chapter contains two parts. Part one explains the findings, results and discussion on domestic water consumption and the impact of each scenario on water use by using different statistical tests and techniques. Part two completes the finding of part one. In the second part, the final results of using ANNs and ANFIS models to predict domestic water consumption. All these results will be discussed and compared to other international studies.

“General Conclusion”

The summary of the findings of the study and major conclusions follows by future recommendations.

5. Thesis Methodology

The methodology adopted for this thesis is shown in the figure 2.

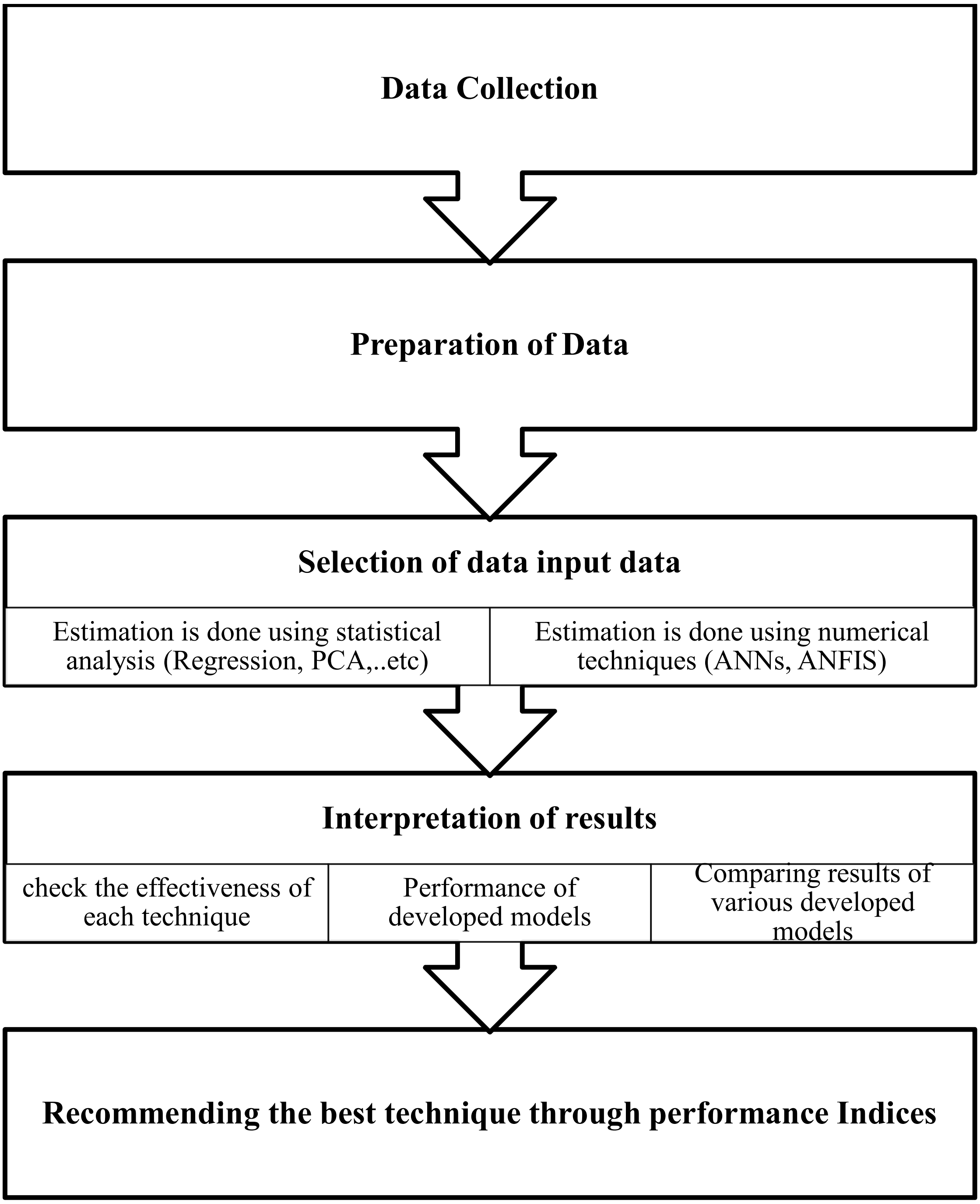


Fig.2 Research adopted methodology

Chapter 1:
Literature Review

1.1.Introduction

In human kind history, freshwater resources were the major factor that population based on to choose a city location. After industrialization, these locations are chosen based on working needs that affects directly water resources. In this century, one of the greatest global challenges is to secure enough fresh water for human beings. With the impressive growth in the world population, especially in the last 70 years, water consumption increased in the same pace. On the other hand, water resources renovation did not follow this pattern. Nowadays, these resources are gradually reducing. The reduction is mainly due to anthropogenic activities and environment degradation .

Our planet has about 70% of its surface covered by water. From this total, 97.5% is salty water, which is inappropriate for the agricultural and industrial activities and human consumption (http://www.ana.gov.br/). This proportion would not be of concern if the balance between the use and the renewal of this resource could remain constant. Meanwhile, this is not what happens and what we can observe is the simultaneous increase of the population and consumption, while the quantity of the disposable potable water worldwide declines as time goes by estimation made by states that in the transition from 1995 to 2025 there will be a substantial decrease of potable water around the world. Currently, in some regions where the lack of water is already a reality, the population lives with serious socioeconomic and healthy problems. In addition, the lack of water causes conflicts throughout the world. Countries like Brazil, rich in water resources, are already affected. Countries in the African continent and Middle East try to suggest recommendations through programs to prevent and mitigate conflict related to the lack of water.

The poverty of water resources around the world encourages us to investigate the parameters affecting domestic water use in order to optimize it. Also, knowing that freshwater is non-renewable resource, the necessity of optimization of water use arise. Several authors extensively investigate water consumption patterns in the developed countries; whilst in the developing countries very little is known (Nauges and Whittington, 2009). Also, global demand for water is continuously increasing due to population growth, industrial development and improvements of economic conditions, while accessible sources keep decreasing in number and capacity. Thus, to satisfy the needs of water consumption in urban areas a study of water use dependence must be conducted. For this reason, researchers have been interested in domestic water consumption by households that are the principal rate of the total volume supplied by the water distribution system in urban areas. According to UNESCO report, water demand arises by 20%yearly with population increase.

Accordingly, the determination of the parameters governing residential water use is an important ingredient for design water demand pattern, for optimize water distribution system and for help future studies in cities design.

By understanding and predicting water use, positive environmental effects can be generated. Also, doing so would contribute to urban water cycle management in terms of strategies for the optimal development of drinking water infrastructure, water consumption control, the environment, and sanitation. Several researchers have focused on forecasting single family residential water use where the urban lifestyle is heavy oriented toward single family homes (Balling et al., 2008; Wentz et al., 2014).

Until now, little importance has been given to water consumption in Algeria, and no studies have been done on household water consumption determinants in Sedrata city, Souk Ahras, North-East of Algeria. A better understanding of factors affecting water use and consumers behaviors will help to satisfy water demand for future planning. The present thesis considers the city of Sedrata, Northeast of Algeria, to assess domestic water use.

The present chapter gives an overview of the water demand forecasting literature and the determinants of household water use appearing in different years in an attempt to identify which forecasting approaches, methods and models are appropriate for urban water utility management decision problems that depend on future levels of demand.

1.2.Framework for Water Demand Forecasting

A set of water demand forecasting literature differentiates forecast practice by the level of planning associated with the forecast (Gardiner and Herrington, 1990), or in accordance with the forecast horizon (Billings and Jones, 2008). In terms of planning level, all water demand forecasting exercises can be used either for strategic, tactical or operational decision making. These respectively concern decisions for capacity expansion, investment planning and system operation, management and optimization. In terms of forecast horizon, water demand forecasting can be categorized as either **long-term, medium-term or short-term** with these horizons being reflective of the general purpose of the forecast. Long-term, medium-term and short-term forecasts are prepared for strategic, tactical and operational decisions respectively (Ghiassi, Zimbra, and Saidane, 2008).

No generally accepted time frame exists for these horizons. For instance, Billings and Jones, 2008 contain different time horizons regarding what constitutes long-term, medium-term and short-term forecasts. One such definition classifies forecasts spanning more than 2 years as *long-term*, those from 3 months to less than 2 years as *medium-term*, and forecasts for 1 to 3 months as *short-term.* (Bougadis et al., 2005; Li and Huicheng, 2010; Nasseri et al., 2011; Donker et al., 2014 and Rinaudo, 2015).

This contrasts with Gardner and Herrington, 1990 who classify these categories as annual forecasts for 10 years or more, annual forecasts for 1 to less than 10 years, and hourly to monthly forecasts up to a year. In terms of application, Ghiassi et al., 2008 prepared monthly demand forecasts for 2 years, weekly demand forecasts for 6 months and daily demand forecasts for 2 weeks, and characterize these as long-term, medium-term and short-term respectively. Researchers focused on forecasting single-family residential water use where the urban lifestyle is heavy oriented toward single-family homes (Wentz et al., 2014; Ouyang et al., 2014; Chang et al., 2017).

1.3.Household Water Consumption Variables and Determinants

Apart from the various planning levels and horizons that tend to complicate urban water demand forecasting, the forecast variable of interest and the determinants of water use are two features that add to the complexity. In the drinking water community, many variables are considered influential in determining water consumption. Although "*A good understanding of the factors influencing demand and reliable estimates of the parameters describing demand behavior and consumption patterns are prerequisites* [to a good forecast]" (Burney et al., 2001). The enormity of these variables can create frustration for water utility mangers. As an illustration of the size and variability of the variables that can be considered, we refer to a report prepared for the Water Research Foundation by Coomes et al., 2010, in which the authors tested the effect of 26 variables on average daily water use for 293 residential customers of the Louisville Water Company.

Besides, in the present research we tested 24 variables on trimester water consumption for 201 individual houses in Sedrata City, Souk Ahras. According to researchers, there are several factors known as drivers or explanatory variables affecting the water consumption. These include socio-economic parameters (population, income, water tariff, etc.), physical characteristics of housing units (area of the house, garden size, etc.), weather data (temperature, precipitation, etc.), study-site infrastructure and system (productivity, technology, etc.), political and administrative factors (application programs, education, and cost) (Gato et al., 2007; Billings and Jones, 2011 and Tiwari and Adamowski, 2013), as well as cultural factors such as consumer preferences and habits. These factors determine water consumption, which means they should be considered input data for forecasting.

1.3.1. Socio-economic parameters

It is imperative to understand what socio-economic and demographic factors. Most empirical studies Renwick and Archibald, 1998; Mayer and DeOreo, 1999; Turner et al., 2009; Beal and Stewart, 2011; Fielding et al., 2012; Halper et al., 2012; Beal et al., 2013; Ouyang et al. 2013; Willis et al., 2013; Matos et al., 2014; Hussein et al., 2016 and Fan et al., 2017 have found residential demand or use for water is influenced by heterogeneity associated with differences in the size of the household and socioeconomic characteristics. The most significant can be categorized into:

1.3.1.1.Household income and water price

Household income has been reported to have a variety of relationships with water consumption (Nieswiadomy and Molina, 1989; Arbues and Villanuà, 2006; Guhathakurta and Gober, 2007; March and Saurf, 2010; Qi and Chang, 2011; Fielding et al., 2012 and Anas et al., 2014). Authors like Kenney et al., 2008; Schleich and Hillenbrand, 2009; Moffat et al. 2011 and Grafton et al., 2011 found that households with **greater incomes** have **greater household water consumption** than households with lower incomes, due mainly to much higher discretionary irrigation end-use demand. But, Beal et al., 2011a outlined a trend of **larger, high-income households** using **less water per capita** than smaller, low-income households. The conflicting results highlight the importance of reporting water demand in water end-use categories (e.g., shower use) and on a per capita basis.

1.3.1.2.Household size

Family size has a variety relationship with water consumption (Arbuès et al., 2010; Beal et al., 2011a; Gato, 2011; Makki et al., 2011; Lee et al., 2012). It has been reported in previous studies that the **increase in household water consumption** is associated with an **increase in the number of people** in the household as Beal et al., 2011b; Beal and Stewart, 2011 and Gato-Trinidad et al., 2011), because in general more people use more water.

In contrast, it has been found that **household per capita consumption decreases** as household **size increases**, due to economies of scale (Beal et al., 2011b; Beal and Stewart, 2011; Russell and Fielding, 2010 and Turner et al., 2009).

1.3.1.3.Retired person

Total indoor water consumption in households **with working residents** is significantly **higher** than that in households with retired residents, and this is mainly due to shower, clothes washer and dishwasher end-use consumption categories (Beal and Stewart, 2011; Beal et al., 2012b, Makki et al., 2011 and Makki et al., 2013). Lyman, 1992 and Willis et al., 2009 hypothesised that retired individuals spend a relatively greater proportion of their time at home and thus have a greater opportunity to use water-dependent appliances.

1.3.1.4.Sex (Gender)

(1 = female and 0 otherwise) is supposed to affect WCP. A positive relationship between WCP and GEN might exist when the respondent is female because they are the ones who take care of domestic household chores such as travelling to other places to fetch water in times of need, hence they will be willing to pay (Moffat et al., 2011).

1.3.1.5.Number of women

Number of women was found to have an effect on indoor water use because women spend more time at home more than men (Mu et al., 1990).

1.3.1.6.Age distribution of household members

WCP for improved water quality and reliability of supply is expected to be positively related to education. The longer time in formal schooling (years), the more people understand better the consequences of using unsafe water and the need to have reliable water supply. Therefore, the educated will be more willing to pay than the illiterate (Moffat et al., 2011). A study showed that houses with **children** may have **higher** water demand as children are more likely to use lawns for play and recreation (Hurd, 2006 and Balling et al., 2008). Although, **children** may use **less water** than **adults** for washing and hygiene (Schleich and Hillenbrand, 2009). Other studies outlined that **teenager** was the key variable of per capita water demand (Schleich and Hillenbrand 2009 and Aquacraft 2015). People of different ages tend to demonstrate different water- related behaviours at home (Makki et al., 2011). There are conflicting results about the effects of age distribution in previous studies. Some, such as Nauges and Thomas, 2000; Martınez-Espineira, 2003; Martins and Fortunato, 2007 and Musolesi and Nosvelli, 2007, find a negative relationship between per capita **water use** and the share of **elderly** people living in households, while other studies, such as Fox et al., 2009; Schleich and Hillenbrand, 2009 and Beal et al., 2011a find that **older** people and **retired individuals** (Lyman, 1992 and Willis et al., 2009) use more water in their homes.

Families with children could be expected to use more water. Outdoors use by children and teenagers might be higher too. Youngsters might use water less carefully, have more showers, and demand more frequent laundering, while retired people might be thriftier. These expectations are confirmed by studies like Nauges and Thomas, 2000 and Kenney et al., 2008 observed that as the mean age of a household increases, so does household water consumption.

Kenney et al., 2008 also outlined the correlation between age, household income and wealth, noting that the increase of water consumption per household is a result of the combination of these variables. Contradictory to previous studies, Wentz et al., 2014 found that age of people was not a significant factor affecting water use. It is well documented that the increase of water use with the increasing in number of occupants is by no means a linear relationship (Heinrich, 2009; Beal et al., 2011a; Gato, 2011 and Lee et al., 2012).

1.3.1.7.Education level

Education level of residents may provide a better estimation of water use (Madanat and Humplick, 1993; House-peters et al., 2010 and Baerenklau et al., 2014 observed that domestic **water consumption** is **positively correlated** with **education level** of occupants. Following from this, higher educational level is a driver of higher per capita water use (Makki et al., 2015).

1.3.1.8.Spatial

Researchers like Wilson and Boehland, 2005; House-Peters et al., 2010 and Breyer et al., 2012 find that **building density** had **negative** impact on domestic **water use**.

Also, **neighbourhood** (neighbours having similar use habits) had significant impact on household water consumption (Wentz et al., 2016 and Gage and Cooper, 2015). In addition, other studies have reported associations between the use of water-efficient technologies in residential dwellings, and reduced water consumption (Heinrich, 2007; Beal and Stewart, 2011; Lee et al., 2011; Water Corporation, 2011; Beal et al., 2013 and Willis et al., 2013).

1.3.2. Physical characteristics of building units

Housing characteristics are the physical features of properties that influence water use efficiency and water needs for the daily life of households, net of influences of household characteristics. Many different housing features have been used, such as:

1.3.2.1.Type of dwelling

Mylopoulos et al., 2004; Domene and Sauri, 2006; Hoffmann et al., 2006 and Fox et al., 2009 demonstrated that the type of dwelling has impact on domestic water usage.

1.3.2.2.Age of dwelling

Household water demand depends also on the age of the house as presented in many studies (Howe and Linaweaver, 1967; Nieswiadomy and Molina, 1988; Nauges and Thomas, 2003; Harlan et al., 2009; Chang et al., 2010; Reynaud, 2013; Ouyang et al., 2014 and Halper et al., 2015). Many other factors were found by different authors of water usage such as:

- **Indoor and outdoor water facilities** (Endter-Wada et al., 2008 and Harlan et al., 2009).
- **Number of bedrooms** (Kenney et al., 2008 and Chang etal., 2010).
- **Lot size** (Pint, 1999; Renwick and Green, 2000; Wentz and Gober, 2007; Chang et al., 2010a; Polebitski et al., 2011 and Halper et al., 2015).
- **Dwelling size** (Nieswiadomy and Molina, 1989; Campbell et al. 2004; Tinker et al. 2005; Mazzanti and Montini 2006; Domene and Sauri, 2006 and House-Peters et al., 2010).
- **Yard size** (House-Peters et al., 2011).
- **Garden size (Turf)** (Nieswiadomy and Molina, 1989; Lyman, 1992 and House-Peters et al., 2010). Researchers as (Giner et al., 2013; Mini et al., 2014 and Gage and Cooper, 2015) found that garden size has a positive impact of domestic water use).
- **Presence of pool and pool size** (Agthe and Billings, 1987; Guhathakurta and Gober, 2007 and Harlan et al., 2009). Also, Halper et al., 2012 observed that household water consumption is negatively correlated with pool size.
- **Landscaping type** (Balling et al., 2008 and Harlan et al., 2009).
- **Property value** (Arbuès et al., 2004).

Variables that measure physical size are generally positively related to residential water use.

1.3.3. Climatic factors

Generally, climatic variables could affect water use ((Balling and Gober, 2007; Franczyk and Chang, 2009; House-Peters et al., 2011; Cole & Stewart, 2013; Misra, 2014; Babel et al., 2014 and Haque et al., 2015a, b).

Researchers considered that **precipitation** has an effective influence on water consumption (Howe and Linaweaver, 1967; Kenney et al., 2008 and Harlan et al., 2009), **temperature** (Danielson, 1979; Guhathakurta and Gober, 2007 and Arbuès et al., 2010) and **evapotranspiration** (Howe and Linaweaver, 1967; Billings and Agthe, 1980 and Farag et al., 2011). Weather patterns influence the consumption of water, with more water being consumed in hot weather, and less during rainy periods. Howe and Linaweaver, 1967 estimated a sprinkling demand model where they specifically took into account summer precipitation and maximum day evapotranspiration. Renwick and Green, 2000 also included seasons and climate in their study.

1.3.4. Environmental

- **Urban heat island** (Gulhathakurta and Gober, 2007; Balling and Cubaque, 2009 and Gober et al, 2012) found that it has a positive impact of water use.
- **Summer** (Chang et al., 2014 and Prandvash and Chang, 2016) demonstrated that summer was found to have a positive effect on domestic water consumption.
- **Drought** (Polebiski and Palmer, 2013 and Breyer and Chang, 2014).

1.3.5. Consumer preferences and habits

In a well-known study, Residential End Uses of Water (REUWS), published in 1999 by the Water Research Foundation and the American Water Works Association, researchers showed that the average DWU of 262.3 litters per capita per day (lpcd) in single-family homes goes into eight end-use components: toilets, faucets, leaks, clothes washers, dishwashers, showers, baths, and other (Mayer et al., 1999).

More detailed information about how and where residential water is consumed (e.g. shower, washing machine, dishwasher, tap, bathtub), is an essential requirement for the development and for a better evaluation of water savings associated with their implementation (Gato, 2006; Kenney, 2008, Heinrich, 2009; Beal and Stewart, 2011; Cole and Stewart, 2012; Willis et al., 2011b, 2013; Makki et al., 2013; Jorgensen et al 2013; Maki et al., 2015; liu et al., 2016; Xue et al., 2017 and Peng et al., 2017). Such as **short shower times** (Jorgensen et al., 2013 and Liu et al., 2016) and **turning off faucet when teeth brushing** (Suero et al., 2012).

The typical selection of economy cycle programmes when using a **dish-washer reduces** the dishwasher end-use water consumption (Beal and Stewart, 2011). The use of **dual flush toilets reduces** toilet end-use water consumption (Walton and Holmes, 2009 and Beal and Stewart, 2011).

For example, use of **efficient showerhead fixtures** results in significant **reductions** in shower end-use consumption (Mayer and DeOreo, 1999; Loh and Coghlan, 2003; Jacobs and Haarhoff, 2004a; Mayer et al., 2004; Roberts, 2005; Turner et al., 2007; Beal and Stewart, 2011; Gato-Trinidad et al., 2011; Beal et al., 2012b; Makki et al., 2013 and Willis et al., 2013).

Moreover, the use of **efficient tap fixtures** and **low-flow tap** add-ons such as flow controllers or restrictors **reduces** tap water end-use consumption (Mayer and DeOreo, 1999; Roberts, 2005; Turner et al., 2005; Cooley et al., 2010; Beal and Stewart, 2011 and Fielding et al., 2012). It has been also reported that the use of **efficient and front- loading washing machines** can result in substantial water **savings** in clothes washer end-use consumption (Gato, 2006; Davis, 2008; Beal and Stewart, 2011; Gato-Trinidad et al., 2011; Lee et al., 2011; Water Corporation, 2011; Beal et al., 2012b and Willis et al., 2013).

Additionally, Gender has also been found to have an influence on shower end-use consumption (Makki et al., 2013). Previous studies have reported that **shower end-use** consumption **increases** in **larger families**, particularly those with **younger** children and **teenagers** (Mayer & DeOreo, 1999; Gato, 2006; Beal and Stewart, 2011; Makki et al., 2013 and Willis et al., 2013). Similarly, **clothes washer** end-use consumption is **positively** related to **household size** and number of **teenagers** and **younger children** in the household (Mayer and DeOreo, 1999; Gato, 2006; Willis et al., 2009d and Beal and Stewart, 2011).

Also, **Tap and toilet end-use** consumption is also **positively** related to **household size**, but in **contrast** to the case of **shower** and **clothes washer** consumption, it **increases** at a higher rate with the addition of higher age occupants such as **adults**, than with the addition of younger children (Mayer and DeOreo, 1999; Gato, 2006 and Beal and Stewart, 2011). Similarly, Beal and Stewart, 2011 noted that **bathing** is commonly associated with families with **younger children**.

Furthermore, **Household size** has also been found to **positively** influences **dishwasher end-use** consumption, although the number of teenagers or younger children has only a weak influence (Mayer and DeOreo, 1999 and Gato, 2006). Also, **household size** is **positively** related to **bath end-use** consumption. However, in a study conducted in Australia, Willis et al., 2009d found that only younger couples and families use bathtubs.

Mayer and DeOreo, 1999 reported **positive** correlations between the **number of employed people** in a household and **shower**, **bath** and **clothes washer end-use** consumption, but **negative** associations of this factor with **tap**, **toilet** and **dishwasher** consumption. Similarly, Makki et al., 2013 suggested that **shower end-use** consumption often represent a large proportion of residential indoor water consumption and it is **positively** correlated with **occupation status**, **education level** and **income level**. Studies at the parcel level report **higher water use** is aligned with **larger irrigation areas, higher incomes, warmer climates, larger house sizes, and a larger household size** (Wilson and Boehland, 2005; Harlan et al., 2009; Gato-Trinidad et al., 2011).

1.4. Forecasting approaches and techniques

Briefly, in the following literature in the field of water consumption modelling has been categorized into studies related to parametric modelling and intelligent modelling. Parametric forecasting methods are appropriate for a straightforward estimation of the relationship between the estimated variable and the explanatory variables. However, they are bounded by some strict assumptions about the given data. Hence, parametric forecasting methods in complex and ambiguous environments lead to a search for intelligent solutions such as artificial neural networks ANNs (Zadeh *et al.,* 2012). Several commonly used techniques have been adopted by water demand studies to identify the influential factors in urban water demand modelling such as;

1.4.1. Regression analysis

Bougadis et al., 2005 investigated the relative performance of regression analysis for short-term peak water demand forecasting. Thus, the use of regression models, as in the Coomes et al., 2010 paper, tend to exert the greatest demand on data collection and management, due to the many factors that are postulated to influence water demand. Also, author researchers used regression analysis in their studies like Fernandes Neto et al., 2005; Dziegielewski and Baumann 2011; Almutaz et al., 2012 and Sebri, 2016.

1.4.2. Factor analysis

Panagopoulos, 2013, used factor analysis to asses factors affecting water consumption in Greece. His results showed that the per connection consumption of large water consumers is decreased when marginal water prices increase.

1.4.3. Sensitivity analysis

Babel *et al., 2014*, forecasted water demand for the Metropolian Waterworks Authority (MWA) in Thailand. Results revealed that climatic variables have very little effect on the annual water demand. However, the monthly demand is significantly affected by climatic variables.

1.4.4. Support Vector Machines

Carlos Peña-Guzmán et al., 2016 forecasted water demand in residential, commercial, and industrial zones in Bogotá, Colombia by using Least-Squares Support Vector Machines.

1.4.5. Principal Component Analysis (PCA)

Alternatively, another multivariate statistical technique, **principal components analysis (PCA)** (Gabriel, 1971; Mahbub *et al.*, 2010 and Haque *et al*., 2015) can be used to investigate the relative importance of various influential factors in water demand modelling. A biplot, the output from PCA, provides a graphical representation of multivariate data, which facilitates easy identification of the dominant factors. There are different approaches to water consumption forecasting including statistical or mathematical techniques. Many new methodologies are developed for modelling the data but current trend seems to be model the data rather than physical process.

For modelling the data, **Artificial Intelligence techniques (AI)** such as **Fuzzy logic (FL), Artificial Neural Networks (ANNs) and Adaptive Neuro Fuzzy Inference System (ANFIS)** are probably the most attractive techniques among the researchers. Thanks to the ability of handling imprecise, fuzzy, noise, probabilistic information to solve complex problem in an efficient manner and less time consuming in modelling complex systems compared to other mathematical models such as regression (Wang, Chau, Cheng, & Qiu, 2009; Pahlavan et al., 2012, Tiwari and Adamowski, 2013 and Sonmez et al., 2018). Artificial neural networks and fuzzy logic techniques of forecasting water demand are advanced methods classified as nonparametric in Billings and Jones, 2008. Also, Koffi, Y.B et al., 2012 forecasted water consumption in Yamoussoukro, Ivory coast using PCA and ANNs.

1.4.6. Artificial Neural Networks (ANNs)

In recent years, Artificial Neural Networks (ANNs) have become extremely popular for prediction and forecasting in a number of areas, including finance, power generation, medicine, water resources and environmental science (Babel and Shinde, 2011; Millie et al., 2014 and Ardjmand et al., 2016, Sakaa et al., 2020). Although the concept of artificial neurons was first introduced in 1943 (McCulloch & Pitts, 1943), research into applications of ANNs has blossomed since the introduction of the back propagation training algorithm for feed forward ANNs in 1986 (Rumelhart et al., 1986a).

According to Tiwari and Adamowski, 2013, the performance advantages of ANNs techniques are attributed to the ability to identify nonlinear relationships among different variables in water demand time series of different characteristics, applied to a very complex data set, used where the structure of the model is unknown and dealing with noisy data. Also, one of the main drawbacks of ANNs methods is, however, their performance limitations in dealing with noisy and non stationary data.

Many reports deal the use of ANNs to forecast water consumption such as Firat et al., 2010; Babel and Shinde., 2011; Bennett et al., 2012, Al-Zahrani and Abo-monasar, 2015; Ghiassi et al., 2017 and Rangel et al., 2017). For more details, Jain et al., 2001 forecasted weekly peak demand, and Jain and Ormsbee, 2002 examined the suitability of ANN models for use in forecasting daily demands; the ANN performance was compared with results produced using conventional time series and regression models.

Also, Pulido-Calvo et al., 2003 found that the ANNs model outperforms regression and time series for total daily water demand of Fuente Palmera, Spain. Furthermore, Bougadis et al., 2005 forecasted peak water demand using three different Artificial Neural Networks (ANNs), regression models and seven-time series models, where results show that ANNs models outperformed regression and time series models. A year later, Babel et al., 2007 developed a model based on the multivariate economic approach using socio-economic characteristics, climatic factors and public water policies and strategies to forecast domestic water demand. The results indicated that the number of connections, water pricing, public education level and average annual rainfall are significant variables of domestic water demand.

Ghiassi et al., 2008 provided urban water demand forecasts using a time series and an autoregressive integrated moving average model. Comparison results showed that a dynamic neural network presents a more accurate and precise forecast rather than the times series forecasts in this study.

Similarly, Adamowski, 2008 developed and compared relative performance of: (i) 39 multiple linear regression models, (ii) 9 autoregressive integrated moving average models and (iii) 39 ANN models; his study concluded that the latter perform the best.

In addition, Oliveira et al., 2009 identified the influencing factors on potable water consumption in the State of Parana-Brazil using a neural representation structure. Firat et al. (2009, 2010) and Yurdusev et al. (2009, 2010) used different ANNs techniques such as General Regression Neural Network (GRNN), Cascade Correlation Neural Network (CCNN), Multilayer Perceptron (MLP), Feed Forward Neural Network (FFNN) and Radial Basis Neural Network (RBNN) to forecast monthly water consumption.

Besides, Mohamed and Al-Mualla, 2010 used a constant rate model to forecast average yearly and monthly water consumption for the city of Umm Al-Quwain, UAE. They concluded that new desalination plants will be needed to cover the expected future water demands.

Moreover, Adamowski et al., 2012 proposed a method based on coupling discrete wavelet transforms and ANNs for daily urban water demand forecasting during summer months. The model was applied in the city of Montreal, Canada, and found more accurate forecasting than the multiple linear and nonlinear regression, autoregressive integrated moving average and ANN models.

Ajbar and Ali, 2013 developed a neural network model for annual and monthly water demand forecasting for the city of Mecca (Saudi Arabia). In addition, the study shed light on the effect of number of visitors to predict water demand.

Bennett et al., 2013 utilized the ANN modelling technique to forecast a residential water end-use demand in Australia for the years from 2005 to 2008. Water end-use data established in the test for over 250 households, which includes demographic, socio-economic and water appliance stock efficiency information. Al-Zahrani and Abo-monasar, (2015) predicted daily water demand in Al-Khobar, Saudi Arabia. Comparison showed that the ANNs models presented results more accurate prediction than other statistical methods like Jain et al., 2001, 2002; Bougadis et al., 2005; Adamowski, 2008 and Adamowski et al., 2012.

1.4.7. Fuzzy Logic (FL)

The fuzzy logic (FL) approach is based on the linguistic uncertainly expression rather than numerical uncertainty. Since Zadeh, 1965 proposed the FL approach to describe complicated systems, it has become popular and has been successfully used in various engineering problems. Altunkaynak et al., 2005 used fuzzy logic approach for forecasting monthly water consumption prediction of the Istanbul city, using Takagi Sugeno method for time series data by considering only one lag as input for the analysis.

Also, Pulido-Calvo and Gutierrez, 2009 combined the ANNs and fuzzy logic to forecast daily water demand at irrigations districts located in Andalucía, Spain. Univariate and multivariate models were compared, in which the former performed better considering the demands and the maximum temperatures of the two previous days. It was found that the multivariate hybrid models performed significantly better than ANNs alone. A.Zadeh et al., 2011 compared the results of fuzzy linear regression and ANNs in forecasting the water consumption in Tehran, Iran from April 2004 to March 2009 and concluded that depending on the temperature one of these methods outperforms the other. They found that in warm days, ANN outperformed the fuzzy linear regression model. However, for cold days, both performed at the same level.

1.4.8. Adaptive Neuro Fuzzy Inference System (ANFIS)

Recently, there appears to be growing interest in hybrid approaches that exploit the strengths of individual methods and aim to reduce model uncertainty (Srinivasulu and Jain, 2009 and Caiado, 2010).

During the past few years, researches attempt to integrate ANNs with other soft computing modelling methods, such as fuzzy sets (FS) (Khajeh and Modarress, 2010; Zhou et al., 2010; Tiwari and Adamowski 2013 and Kant et al. 2013), genetic algorithm (GA) and support vector machine (SVM) to improve the performance of modelling. One of the focused areas in modelling systems is neuro-fuzzy modelling system which is fuzzy system combined with artificial neural networks as a learning algorithm, where the most notable neuro-fuzzy systems is **ANFIS** which stands for adaptive neuro-fuzzy interface system (Jang and Sun, 1995). Fuzzy inference system (FIS) is a rule-based system consisting of three components. These are: (i) a rule-base, containing fuzzy if-then rules, (ii) a data-base, defining the membership functions (MF) and (iii) an inference system that combines the fuzzy rules and produces the system results.

Several authors have used adaptive neuro inference system (ANFIS) models, which integrates ANNs and FL methods, to model water consumption such as (Nayak et al., 2004; Sen et al., 2009; Ahmadi et al., 2014 and Papageorgiou et al., 2015). ANFIS has the potential benefits of both these methods. It eliminates the basic problem in FL design, defining the membership function parameters and design of fuzzy if-then rules, by effectively using the learning capability of ANNs for automatic fuzzy rule generation and parameter optimization (Nayak et al., 2004).

Previous literature shows that ANFIS has been adopted to model and forecast water consumption. Both of Altunkaynak et al., 2005 predicted water consumption in Istanbul at the same year. Also, Kermani and Teshnehlab., 2008 used normalized data for water consumption prediction using ANFIS method and also further, auto regressive model is employed for the analysis and they found that ANFIS model is better than autoregressive model.

In 2009, Tabesh et al., 2009 compared several fuzzy and neuro-fuzzy models to forecast the short-term water demand in Tehran where their results showed that the accuracy of neuro-fuzzy models outperformed the classical fuzzy models. Yurdusev and Firat., 2009 used ANFIS method to forecast monthly water consumption modelling and they have adopted cross correlation method for selection of the input variables. Also, Sen and Altunkaynak., 2009 used Mamdani inference system for modelling of drinking water prediction using different fuzzy sets and rules in the analysis.

Furthermore, a fuzzy neural network was applied by Liu and Huicheng, 2010 to forecast the urban water demand for the city of Dalian, China. The authors concluded that the proposed model was more accurate than the traditional methods. Liu et al., 2012 used a fuzzy clustering technique to forecast the daily water demand of Hangzhou, China.

In addition, Ahmadi et al., 2014 used the fuzzy cognitive map learning method to forecast the water demand. The fuzzy cognitive map models complex systems as a collection of concepts and underlying relationships between concepts. Vijayalaksmi et al., 2015 used a neuro-fuzzy inference system to forecast the water demand for Tamil Nadu, India. In another recent study, Papageorgiou et al., 2015 predicted the water demand by using fuzzy cognitive maps. They optimized the fuzzy cognitive maps' learning process using genetic algorithm.

Many authors concluded that the accuracy of ANFIS models outperformed the other methods like Teshnehlab, 2008 and Liu et al., 2009.

Unfortunately, there is no such study for the case of Sedrata and in particular for the semi-arid area.

Chapter 2:
Study Area

2.1.Introduction

On global scale and for water resources sustainable management, the concept IWRM (Integrated Water Resources Management), was promoted in 2000 by the Global Water Partnership (GWP) in collaboration with the World Water Council as part of its second international forum organized in The Hague in the Netherlands (World Water Council, 2000). Content formulated by the Global Water Partnership, IWRM: "*It is a process that encourages the development and management of the coordinates of water, lands and associated resources, maximizing economic well-being and social which results in a fair manner, without compromising the sustainability of vital projects*". The objective is to satisfy the needs of different uses while preserving the sustainability of resources and the needs that depend on them.

The current chapter details all water consumption relative parameters including extraction, distribution and management in Sedrata city. Before that, a geographic overview of Algeria and the study area is presented.

2.2.Overview of Water in Algeria

In Algeria, natural water potential is about 19 billion cubic meters (BCM) per year, while water demand keeps increasing day after day and approximately 450 cubic meters per capita per year (RAPPORT D'INFORMATION, 2011). Water resources is Algeria vary from north to south. Table 2.1 below demonstrates the available water resources in this country, where renewable surface water has the highest volume (11 BCM) and distributed between north and south.

Surface water inflows are low in the Saharan basin, with a total of 0.5 BCM per year. In contrast, the north relies mainly on surface water, since almost 7 BCM is captured by a number of medium and large dams.

Table 2.1: Water resource availability

Water resource	Volume (BMC)	Region
Renewable surface water	11	North and south
Renewable groundwater	2.5	North
Non-renewable groundwater	6.1	South

Groundwater resources are estimated to total 7.6 BCM, but demand is much higher in the north of the country due to population distribution. Important aquifers in the Sahara meet 96% of water demand in the south (Bouchekima et al., 2008).

In the north, the aquifers are naturally recharged with 1.9 BCM per year, with total withdrawals equal to 2.4 BCM per year, because they are shallow. They are exploited using wells and springs. The deficit is mainly due to a lack of effective groundwater management, linked to poor knowledge of the resource, an increase in the number of illegal wells and a lack of coordination between the water authorities (Food and Agriculture Organization of the United Nations, 2009; British geological survey, 2018).

The south is characterized by two biggest aquifers that are the Complex Terminal and the Continental Interlayer. They are fossil groundwater. Which form the transboundary North-Western Sahara Aquifer System (NWSAS). The Complex Terminal (100-400 meters in deep) and the Continental Interlayer (1,000-1,500 meters in deep) contains significant reserves of 30,000-40,000 BCM (Figure 2.1). These aquifers are exploited mainly using deep boreholes, whereas the shallow ones are exploited using the traditional foggara system (Ibid).

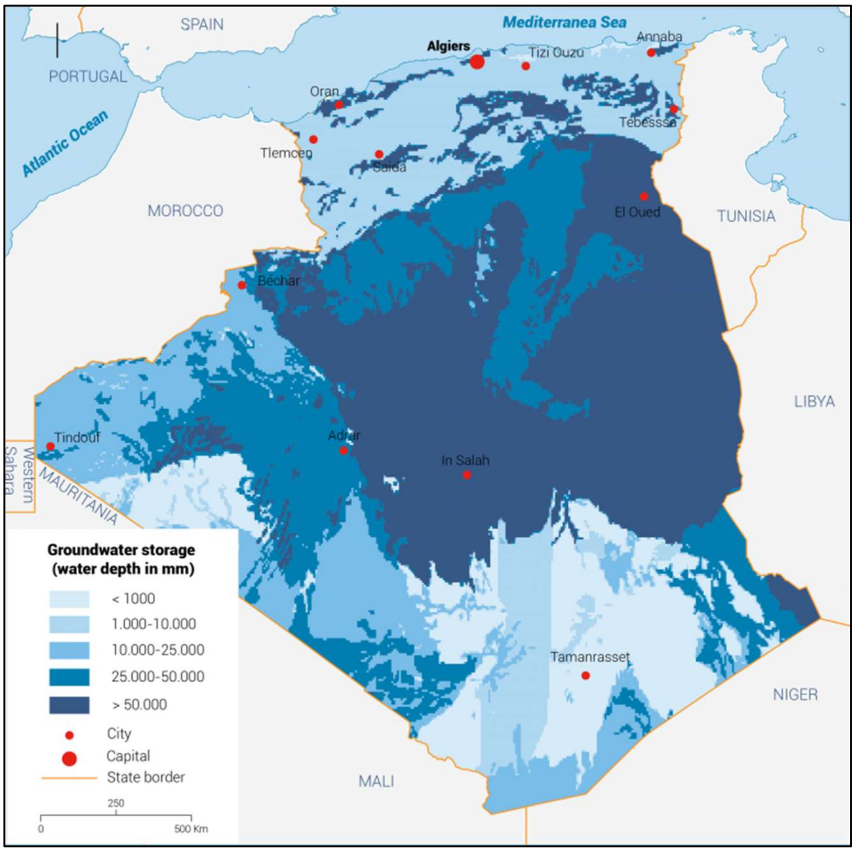


Fig 2.1 Groundwater storage

Currently, the deficit between supply and demand water is around 1.3 BCM. Figure 2.2 illustrates water distribution in Algeria, which are irregular distribution (Fanak, 2020). The overall demand has quadrupled over the last 40 years and currently exceeds more than half the volume of potentially mobilizable resources. At this rate, the maximum limit of hydraulic potential is projected to be reached before 2050 (FAO, GIZ/BGR/OSS, 2016).

The share of drinking water has grown considerably from 16% of overall consumption in 1975 to 36% in 2019. By contrast, during the same period, agriculture's share fell from 80% to 60%, even though it remains the leading consumer (FAO, 2016).

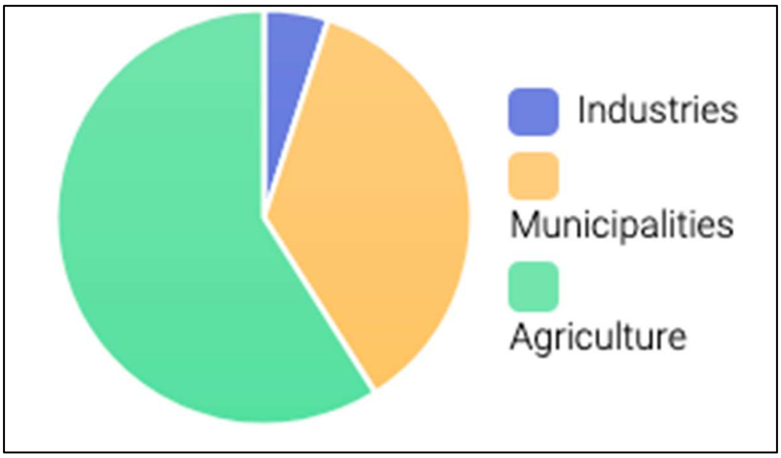


Fig 2.2 Percentage of water consumption by sector

Based on population growth, rural-to-urban migration and per capita water demand projections, the overall water demand for all sectors is projected to be approximately 18.9 BCM by 2030 (Table 2.2) (Hamiche et al., 2016). The planning studies conducted by the Ministry of Water Resources allow the construction of around 70 additional dams across all river basins, reaching a total of 139 reservoirs by 2030 (Ibid).

The Algerian National Strategy for Development of Water Resources gives a significant place to the exploitation of non-conventional water resources. Non-conventional water production is forecasted to be nearly 3 BCM in 2030. This water is intended to be used for watering green spaces and sports fields and developing irrigation around urban areas (Fanak, 2020).

Table 2.2: Changes to future water demand, in BCM (Ibid)

Usage	2011 (BCM)	2030 (BCM)
Drinking (urban, rural and industrial)	2.9	3.5
Irrigation	8.6	15.4
Total	11.5	18.9

Recently, the minister of water resources reported on the finalization of studies for the construction of 23 new dams, 3 major transfers and a dam connection project.

In addition to ongoing studies for the construction of 36 dams and three studies of water transfer from the South to the Highlands. According to him, the forecasted volume of annual needs for 2030 is about 4 BCM for domestic consumption (against 3.2 BCM currently), 8.3 BCM for agriculture (against 7 billion m3 currently) and 0.4 billion m^3 for industry (against 0.3 billion m3 currently). Also, the number of desalination station will increase to 15 large stations (against 11 currently). It will represent 25% of national production by 2030 (Bource, 2020).

2.3.Sedrata location

Sedrata is a municipality and large city in Souk Ahras Province, Algeria, with total population of **54 205** in 2017(ADE). It is semi-arid region characterized by warm and temperate climate with **14.2°C** of temperature and **523mm** of precipitation in average. It is located between 36°07’42” North latitude and 7°31’53” East longitude in the North East of the country, close to the border with Tunisia. It is situated in the West of Souk Ahras. The location map of the study area in sedrata city is given in figure 2.3.

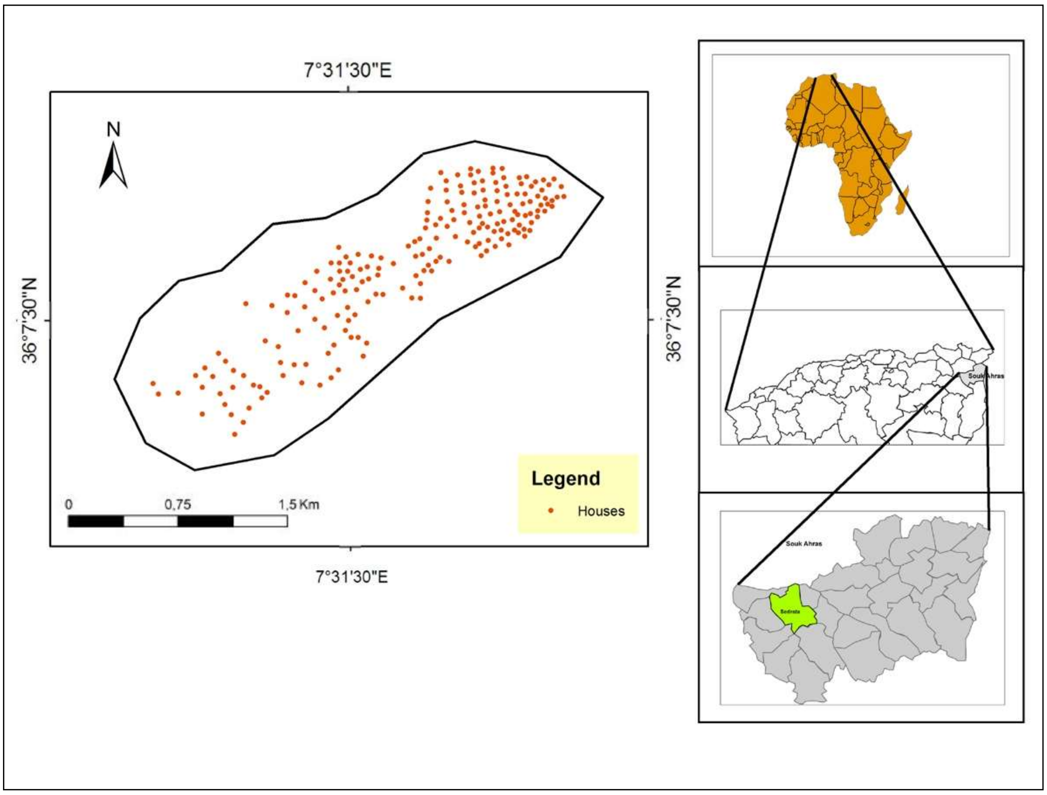


Fig 2.3 Location map of study area in Sedrata city

Souk Ahras province is located in the north-east of this country, 640 km from Algiers, it is limited by the wilaya of El Taref and Guelma from the north, from the west by the wilaya of Oum El Bouaghi, from the south by the wilaya of Tébessa and by Tunisia from east (An 88 km border strip).

The Wilaya of Souk-Ahras covers a total area of 4,360 km^2 with total population equals to 505,337 Inhabitants.

Three types of consumers encountered in Sedrata, domestic, agriculture and industry, however only domestic water usage is considered. It represents **89.17%** (ADE) of the total water usage.

In fact, the type of house is an important background variable determining the water use practice of households. It has a direct link with demand of water. In this study the total population consists of metered domestic consumers of water consumption living in single houses (not collective apartment).

The study area is an explorative one, because no other study has come up in the area.

2.4. Water in Sedrata City

In general, Souk Ahras province continues to grow resulting an increase of water demand. In the region it exists 29 wells under exploitation having capacity of 280 l/s (18 130 m^3/d), 2 boreholes with capacity equals to 41 l/s (260 m^3/d), 9 water sources (25 l/s or 2 000 m^3/D) and 20 pumping station with 584 l/s (50 400 m^3/d). It has also 100 distribution tanks (80 280 m^3), 397 km^2 of adduction linear network and 1669 km of drinking water distribution network.

According to ADE of Souk Ahras, in Sedrata 2 types of local resources. Surface water, treated from Ain Dalia dam which represents 4000 m^3 per day and underground water (1760 m^3 per day). this water is piped to 58 164 inhabitants. From the total produced volume (5760 m^3 per day), only 4614 m^3 per day is distributed to residents of the region. The volume represents 13,77 % of the total distributed volume (33 507 m^3/day) of Souk Ahras province. Table 2.3 shows the distribution program of Sedrata and some regions of Souk Ahras piped by Ain Dalia dam.

One of the main water problems faced in the region is that wells of Guedrane underground water is fully exploited. Also, the piped distribution system is too old. Table 2.4 illustrates current exploitation wells in the region.

As mentioned, the local resources constituted by surface (Ain Dalia Dam) and underground water were no longer sufficient to satisfy the basic needs of the city. Thus, the government then turned to another surface resources, and invested heavily in the implementation of drinking water treatment station at the Oued Charef dam.

The structure of the hydraulic sector is intended to treat a volume of 8,000 m^3 (against 4,000 m^3) of drinking water, in favor of the population of Sedrata, the second urban agglomeration in the border of Souk Ahras.

Table 2.3: Distribution program of piped water from Ain Dalia dam (2018)

Region		Served population (hab)	Produced volume (m^3/d)		Total produced volume (m^3/d)	Distributed volume (m^3/d)	H24(%)	daily		1day/2		1 day/3	
			Dam	Borehole				(%)	hours	(%)	hours	(%)	hours
Souk-Ahras	Souk-Ahras	179 077	22 900	3 900	26 800	24600	-	-	-	70	6	30	6
	Hennancha	9820	200	742	1274	1004	-	80	12	20	2	-	-
	Sedrata	**58 164**	**4000**	**1 760**	**5 760**	**4614**	**-**	**-**	**-**	**30**	**6**	**70**	**5**
	M'Daourouch	37502	1500	600	2 400	1804	-	-	-	-	-	100	4
	Oued Kebrit	5135	400	148	548	402	-	50	4	50	4	-	-
	Zouabi	1965	300	0	300	246	-	0	0	100	8	-	-
	Bir Bouhouch	6944	700	230	930	837	-	-	-	100	5	-	-
Total		**298 607**	**30 000**	**7 432**	**38 012**	**33 507**	**-**	**4**	**-**	**52**	**-**	**44**	**-**

Table 2.4: Current exploitation wells

Wells name	Region	Served region	Theoretical flow (l/s)	Exploitation flow (l/s)	Pumping system (hour)	Exploitation flow (m3/d)
S 15	Khemissa	Khemissa, Sedrata	12	06	6	130
MD 10	M'Daourouch	M'Daourouch	14	10	18	648
BB2	Bir Bouhouch	Bir Bouhouch	05	02	11	79
BB12	Bir Bouhouch	Bir Bouhouch	08	06	12	259
MD9	Guedrane	Sedrata	30	09	18	583
MD11	Guedrane	Sedrata	30	10	18	648

Chapter 3:
Data Collection

3.1. Introduction

Water plays a major role in meeting the day-to-day needs worldwide. Based on gender, age, season and other factors the demand and consumption pattern varies.

Due to the non-controlled cities growth and scarcity of water resources, the water systems could not supply the required daily amount of water needs. Moreover, the region is under construction and that amplifies water demand.

This chapter presents the collected variables and their sources, besides to the techniques employed for collecting the data set. Before the survey conducted, individual discussions and field visits were held to understand the situation of water supply and use in the region.

Estimation of water use determinants requires a reliable measure of water consumption and information on the consumers and their houses. The study was conducted using a questionnaire that contains the purpose of the survey, followed by questions on water usage practices and behaviors. The survey questionnaire is built basing on previous studies and covers all relevant parameters. For better understanding of question by users, it is redacted in two languages and aims only single houses.

The present chapter is divided into statistical description of variables and data preparation. Any statistical analysis starts with standard data preparation techniques. Basic descriptive statistics are produced to note any missing/ abnormal values.

3.2. Description of the variables (Raw data)

Both quantitative and qualitative data are collected and in order to account factors that could explain variations in indoor water use, the selected sample considers many variables. Water consumption is "**the dependent"** variable, while socio-economic parameters, indoor habits of residents, physical characteristics of buildings units and climatic factors are "**the independent"** variables in this study (Figure 3.1). The difference between *dependent* and *independent* is that:

- **Dependent variable:** is variable that may depend on other factors. For example, household water use may change depending on weather factors.
- **Independent variable:** is a variable that does not depend on other factors. For example, weather factors like temperature does not change depending on household water use.

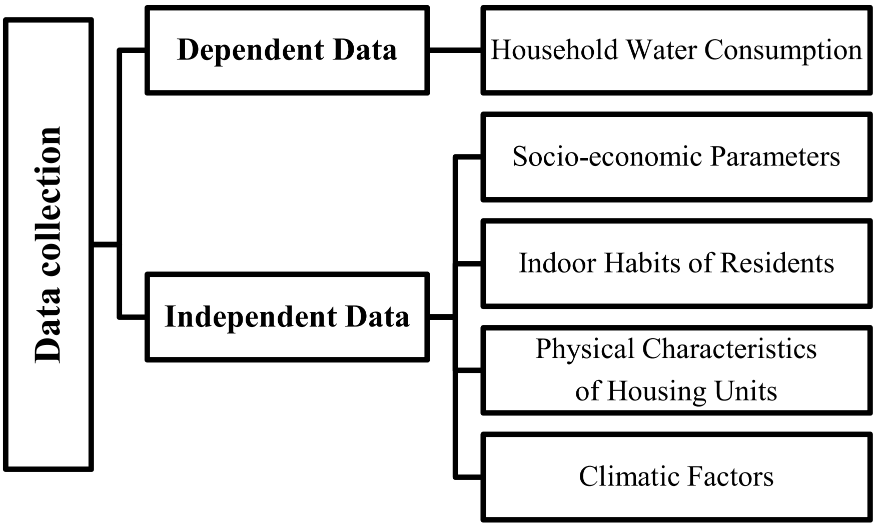


Fig 3.1 Type of collected data

The collected data were chosen based on previous studies as shown in the literature review chapter and they could be categorized into:

3.2.1. Water Consumption (WCP)

Household water consumption is obtained from **"Algérienne des eaux (ADE)"**, authority of water distribution and management, of Souk-Ahras. According to Sadoulet and De Janvry, 1995 the term "Household" varies widely across cultures. Their study defines household as "*the group of people living together and sharing the same kitchen or, in the case of piped water households, using the same piped water*".

WCP data represents trimester values from 2012 to 2017 for 363 single houses selected from more than 500 houses in Sedrata city. The sample size of data is minimized because of the lack of socio-economic variables and physical characteristics of buildings information. WCP is measured by cubic meters.

3.2.2. Socio-economic Parameters (SEP), Indoor Habits of Residents (INH) and Physical Characteristics of Buildings Units (PHC)

To gather socio-economic parameters, indoor habits of people and physical characteristics of houses, the public inquiry represents the best tool.

People are selected randomly to get a representative sample on Sedrata from different types, where a questionnaire is distributed and filled during the period between March and April 2018 (to people living in the study area) based on a list created from preliminary literature review. The questionnaire paper contains questions on all relative parameters (Annex 01).

3.2.2.1. Socio-economic Parameters (SEP)

SEP should be considered in detail in any WCP investigation. In fact, this link between SEP and WCP is logical and confirmed by many authors (Chapter 01).

a. **Family Composition and Gender:** Family composition refers to total number of individuals (HOUS) beside the number of females (FEM) and males (MAL) living in the house. It is expected that a larger household size will increase the tendency to have a higher WCP because household need more water if there is more people. Thus, this variable is expected to have a positive sign to indicate that females are likely to have a higher WCP. This hypothesis is because women usually deal with domestic activities.
b. **Age of Family Members:** This means age of residents in each house. It is divided into four classes of age: under 8 years old (AG1), between 9 to 15 years old (AG2), between 15 to 35 years old (AG3) and older than 35 years old (AG4).
c. **Education Level of Residents:** Educational status of the residents is considered by classifying the level of education into four categories: primary school (PRS), medium school (MDS), high school (HGS) and university level (UNIV). It is expected that higher level of education would decrease WCP thanks to understanding water importance.
d. **Household Income (INC):** This variable is a combination of all incomes from all household members. It is measured in monetary unit, Algerian Dinar, DA/Month.
e. **Car Possession and Usage:** Car possession refers to the existence of cars, their numbers (CARN) and frequency of washing cars per month (WCAR).

3.2.2.2.Indoor Habits (INH)

This INH represents the individual’s personnel and hygienic practices inside the housing units and includes frequency of using toilets per day (UTLT), frequency of washing clothes per week (WCL), frequency of washing dishes per day (WDISH) and frequency of showering for females and males per week (FSHW) and (MSHW), respectively.

3.2.2.3.Physical Characteristics of Building units (PHC)

Besides of residents related characteristics, the PHC has also a possible impact on WCP and describes the *building* and *the garden*.

a. **Building:** It refers to type of building, total area (TAR), building area (size) of the houses (BAR) and number of rooms (ROMN) in every habitation unit. Surface area is measured by squared meters.
b. **Garden Possession and Usage:** Garden possession signifies the existence of garden and their size (GAR), number of times watering the garden (GWAT). It is measured by squared meters.

3.2.3. Climatic Factors (CLF)

Climate is one of the most influential factors on water demand because it dictates consumption. Accordingly, the effects of season and weather must be accounted for any evaluation of water savings. Meteorological data may be directly factored WCP. Most commonly, mean temperature and mean precipitation data are used. Climatic data of Sedrata are obtained from both meteorological station of Souk Ahras and online databases.

a. **Mean Precipitation (PRE)** Precipitation could affect outdoor water use by reducing garden watering for example, and it is measured by millimeter (mm).
b. **Mean Temperature (TEM)** Temperature like precipitation could affect both indoor and outdoor household water use. This variable is measured by (°C).

Other information was collected from official documents, published reports, journals, books and related literature. The previous variables will explain the variation in domestic WCP and table 3.1 summarizes the collected data.

Table 3.1 Variables Description

Type		Variable	Acronym
Dependent		Household Water Consumption (m^3)	WCP
Independent	**Socio-economic indicators (SEP)**	**Family composition and gender**	
		Household size	HOUS
		Number of Female	FEM
		Number of Male	MAL
		Age of family member	
		Under 8 years old	AG1
		Between 9 to 15 years old	AG2
		Between 15 to 35 years old	AG3
		Older than 35 years old	AG4
		Education level of residents	
		Primary School	PRS
		Medium School	MDS
		High School	HGS
		University	UNIV
		Household Income (DA)	INC
		Car possession	
		Number of Cars	CARN
		Number of washing Car (month)	WCAR
	Indoor Habits (INH) Frequency	Washing Clothes (week)	WCL
		Washing Dish (day)	WDISH
		Using Toilets (day)	UTLT
		Female takes shower (week)	FSHW
		Male takes shower (week)	MSHW
	Physical Characteristics of Building (PHC)	**Building**	
		Total Area (m2)	TAR
		Building Area (m2)	BAR
		Number of Rooms	ROMN
		Garden possession	
		Garden Area	GAR
		Number of Garden Watering	GWAT
	Climatic Factors (CLF)	Mean Precipitation (mm)	PRE
		Mean Temperature (°C)	TEM

3.3. Preliminary Data Analysis

In quantitative research, when primary data is collected from surveys, a preliminary data analysis is a critical step required before the actual data analysis such as regression and parametric or non-parametric statistics.

After collecting data, a pre-processing of data (preliminary analysis) was prepared to eliminate outliers, incomplete dataset and to make sure that the subsequent analyses are all valid (Xue et al., 2017). Dataset was pre-processed in the following steps (procedures):

a. Dataset selection

The first step is the selection of the required data. In this work, both of water consumption values that obtained from the authority service and data collected from the questionnaire paper were first amassed together in one Excel sheet. The total number of selected houses was only **363 single houses,** because some residents of the study area were not available during the study visits.

b. Dataset examination

The main purpose of the second step is to check missing values and outliers, where it can happen by chance in any data distribution. Missing values are common in a large dataset, while outliers may indicate measurement errors dataset. In this research, values equal to zero in water consumption are considered as missing values, because it means no residents and no water consumed. Furthermore, housing units with low water consumption (WCP< 2 m^3) are also excluded from further analysis because such values are not realistic and may affect consequently the results. In addition, after evaluating the questionnaire papers the incomplete and incoherent answers are removed from the study.

c. Dataset summarizing

Accordingly, after the two previous steps, **162** houses were chosen and excluded from the first selected 363 houses. The final number of houses used for the study is **201** houses contained a dataset of **4824** valid water consumption values. Also, water consumption values are combined with household information. More details about household parameters are presented in next parts of this chapter.

3.4. Dataset Descriptive Statistics

All descriptive statistics were calculated by "**the Statistical Package of Social Sciences (SPSS 19)"** software and **"STATISTICA 8"** software. More details about SPSS and STATISTICA will be presented in the next chapters.

3.4.1. Water Consumption (WCP)

Description of statistical characteristics of annual water consumption are presented in table 3.2. Results from table 3.2 shows that mean value of water consumption varies between 115,55 m^3 and 139,03 m^3. For min and max values, water consumption varies from 11 m^3 to 36 m^3 and from 301 m^3 to 784 m^3 respectively.

Table 3.2 Descriptive statistics of annual water consumption (2012_2017)

Statistical Parameters	WCP 2012	WCP 2013	WCP 2014	WCP 2015	WCP 2016	WCP 2017
Mean	115,55	130,73	122,57	139,03	138,70	117,10
Median	112,00	125,00	117,00	127,00	126,00	107,00
Mode	144	144	144	120	80[a]	75
Std. Deviation	55,387	63,238	54,745	80,998	72,878	55,981
Variance	3067,715	3999,037	2996,995	6560,606	5311,146	3133,839
Skewness	1,208	2,333	0,648	3,234	1,390	0,983
Std. Error of Skewness	0,172	0,172	0,172	0,172	0,172	0,172
Kurtosis	2,679	12,673	0,245	20,233	2,719	0,838
Std. Error of Kurtosis	0,341	0,341	0,341	0,341	0,341	0,341
Minimum	30	36	11	36	11	30
Maximum	384	578	313	784	429	301
Sum	23226	26277	24637	27945	27880	23536

The analysis of the average per capita water consumption shows that the highest consumption is during the year 2015.

Also, it is evident that there is perceivable variation in WCP availability among the six years.

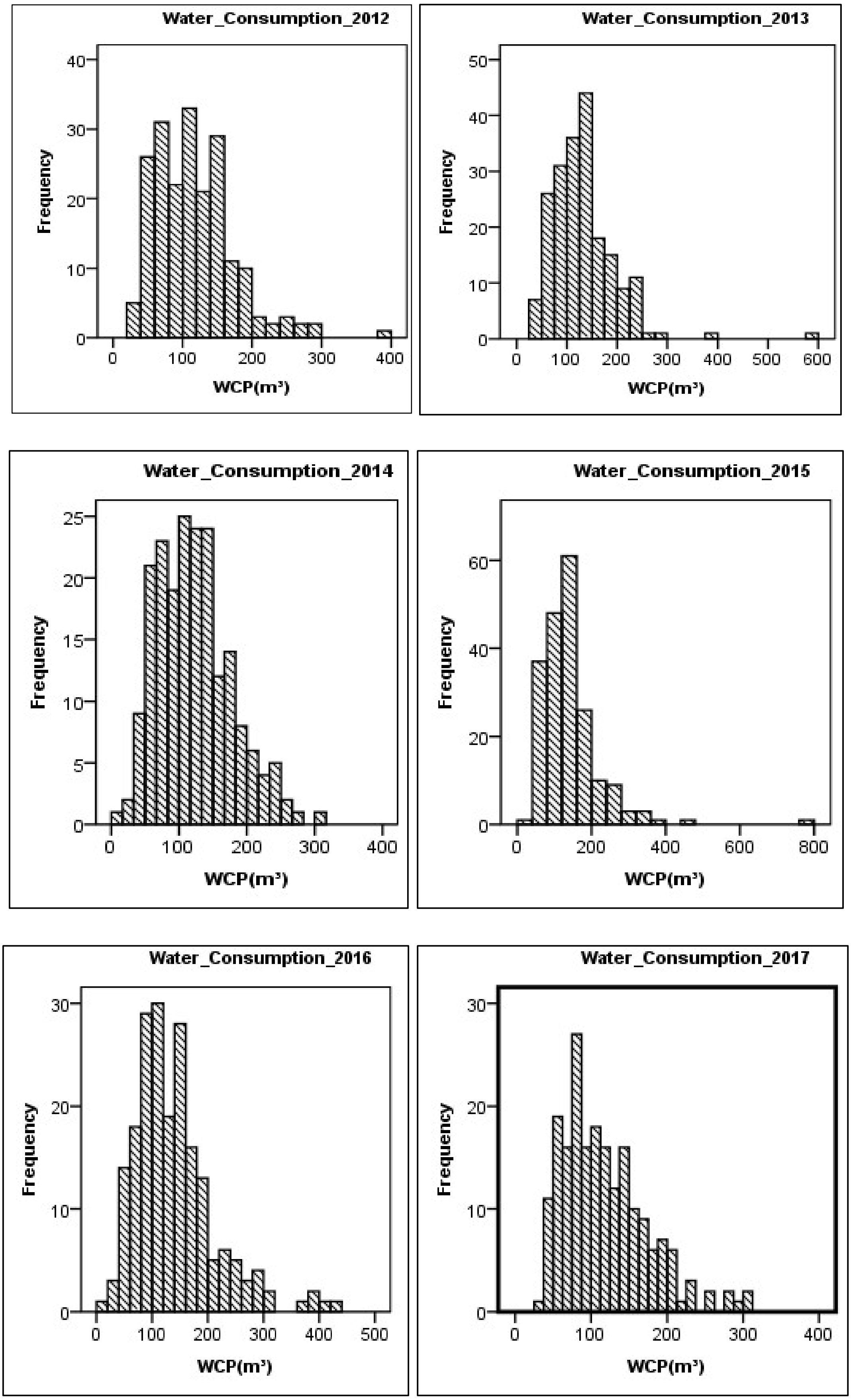


Fig 3.2 Water consumption distribution during the period 2012-2017

In addition to annual consumption, the mean trimester WCP and total WCP from 2012 to 2017 is presented in Table 3.3. Mean WCP values were obtained by calculating the mean value from all trimester WCP of every year, i.e. first value (Mean WCP 1^{st} trimester) was obtained by: WC of 1^{st} trimester in 2012 + WC of 1^{st} trimester in 2013+ WC of 1^{st} trimester in 2014+ WC of 1^{st} trimester in 2015+ WC of 1^{st} trimester in 2016+ WC of 1^{st} trimester in 2017.

Table 3.3 Mean trimester and total water consumption (2012_2017)

Statistical Parameters	**Mean WCP 1^{st}_Trimester**	**Mean WCP 2^{nd}_Trimester**	**Mean WCP 3^{rd}_Trimester**	**Mean WCP 4^{th}_Trimester**
Trimester	**31 Jan to 31 Mar**	**31 Mar to 31 Jun**	**31 Jun to 31 Sep**	**31 Sep to 31 Dec**
Mean	171,78	167,51	204,50	219,89
Median	163,20	160,00	201,00	199,00
Mode	216	216	216	210
Std. Deviation	74,527	74,987	88,120	111,418
Variance	5554,222	5622,985	7765,150	12414,062
Skewness	0,897	1,016	1,331	1,953
Std. Error of Skewness	0,172	0,172	0,172	0,172
Kurtosis	1,012	2,008	3,369	7,910
Std. Error of Kurtosis	0,341	0,341	0,341	0,341
Minimum	38	42	67	41
Maximum	432	523	605	913
Sum	34528	33671	41105	44198

Examination of table 3.3 and figure 3.3 reveals that WCP vary through seasons, where WCP tend to increase with hot seasons.

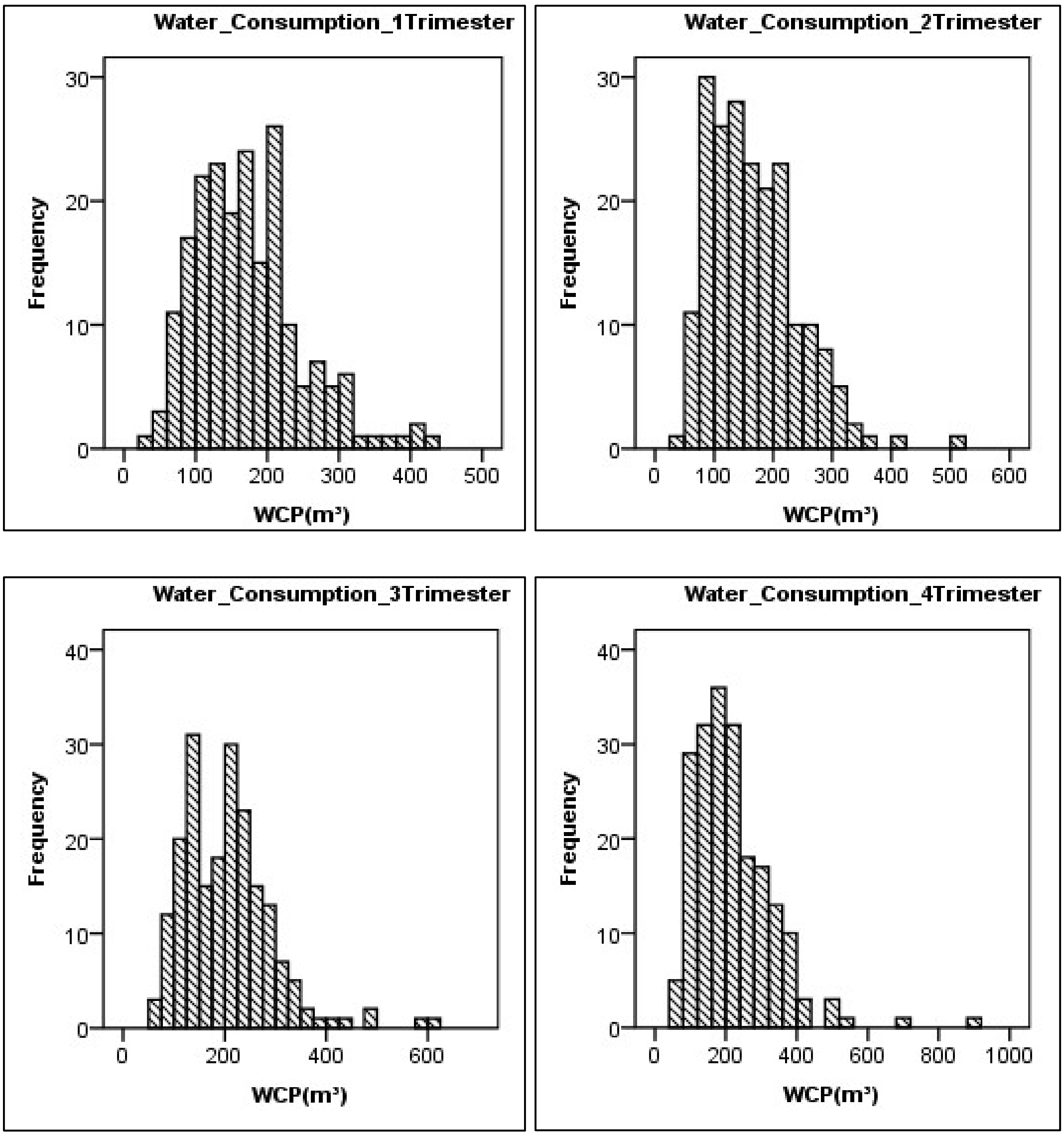


Fig 3.3 Mean trimester water consumption (2012_2017)

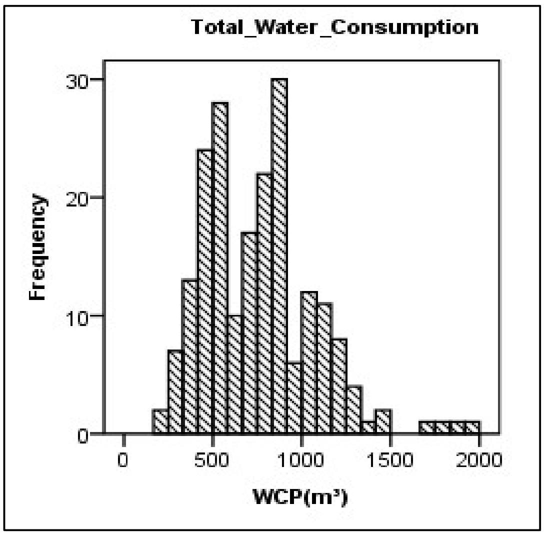


Fig 3.4 Mean total of water consumption (2012_2017)

3.4.2. Socio-economic Parameters (SEP)

Table 3.4 summary the descriptive statistics of the socio-economic parameters

Table 3.4 Descriptive statistics of socio-economic parameters

Statistical Parameters	FEM	MAL	AG1	AG2	AG3	AG4	PRS	MDS	HGS	UNIV	INC	CARN	WCAR
Mean	3,00	2,00	2,00	1,00	2,00	1,00	1,00	1,00	1,00	1,00	53905	2,00	2,00
Median	3,00	2,00	1,00	1,00	2,00	1,00	1,00	1,00	1,00	1,00	40000	1,00	1,00
Mode	3,00	2,00	1	0	1	1	0	1	1	1	40000	1	1
Std. Deviation	1,26	0,89	0,78	0,67	1,12	0,71	0,58	0,67	0,77	0,73	19976	0,69	1,19
Variance	1,58	0,79	0,60	0,45	1,25	0,50	0,34	0,45	0,60	0,53	3,99E8	0,47	1,41
Skewness	0,24	0,38	0,05	0,51	0,46	-0,32	0,73	0,22	0,81	0,09	1,09	-0,33	0,85
Std. Error of Skewness	0,17	0,17	0,17	0,17	0,17	0,17	0,17	0,17	0,17	0,17	0,17	0,17	0,17
Kurtosis	-0,51	-0,16	-0,47	-0,73	-1,21	-0,98	-0,4	-0,78	-0,86	-1,11	-0,13	-0,72	-0,05
Std. Error of Kurtosis	0,34	0,34	0,34	0,34	0,34	0,34	0,34	0,34	0,34	0,34	0,34	0,34	0,34
Minimum	1	0	0	0	1	0	0	0	1	0	$35^{E}3$	0	0
Maximum	6	5	**3**	**2**	4	2	2	2	3	2	$11^{E}4$	3	4
Sum	619	438	257	133	425	242	98	165	323	190	$10835^{E}3$	257	252

3.4.2.1.Household Size (HOUS) and Gender

In term of household size and composition, there is evidence to suggest that Household size is an important indicator for water consumption. People with larger families are expected to use more water.

The data shows that consumption patterns vary also according to gender distribution. Results from table 3.5 and figure 3.5 indicate that the sample consists of **619** (**58.56%**) females and **438** (**41.44%**) males.

Table 3.5 Household size and gender

Gender	Frequency	Percentage
Female	619	58,56 %
Male	438	41,44 %
Total	**1057**	**100 %**

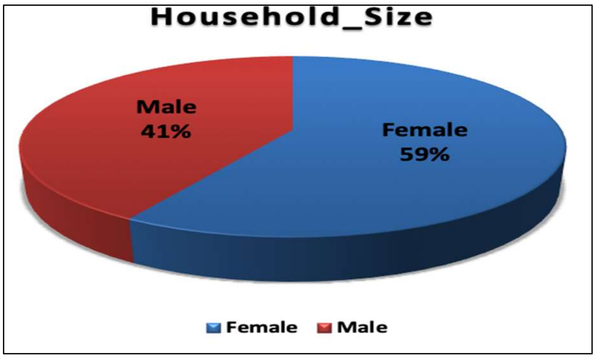


Fig 3.5 Gender in the study area

The majority of houses **(33.83%)** has six members, as shown in table 3.6 and figure 3.6. The minimum and the maximum household members ranges between 2 to 8, respectively. Generally, houses with family members more than seven persons having more than one floor and maybe are not single families (collective houses).

Table 3.6 Distribution of residents (family size)

Family Size	Frequency	Percentage
2 Members	12	5.97 %
3 Members	25	12.44 %
4 Members	27	13.43 %
5 Members	27	13.43 %
6 Members	68	33.83 %
7 Members	29	14.43 %
8 Members	13	6.47 %
Total	**201**	**100**

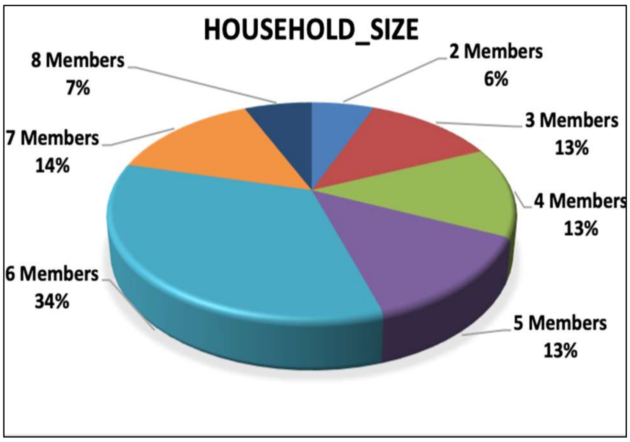


Fig 3.6 Distribution of households

Table 3.4 shows that minimum number of females (FEM) per house is 1, while the maximum number of females is 6. The distribution buildings according to females can be categorized into six main groups (figure 3.7):

- 20 building units with one female (9,95%)
- 52 building units with 2 females (25,87%)
- 53 building units with 3 females (26,37%)
- 51 building units with 4 females (25,37%)
- 18 building units with 5 females (8,96%)
- 7 building units with 6 females (3.48%)

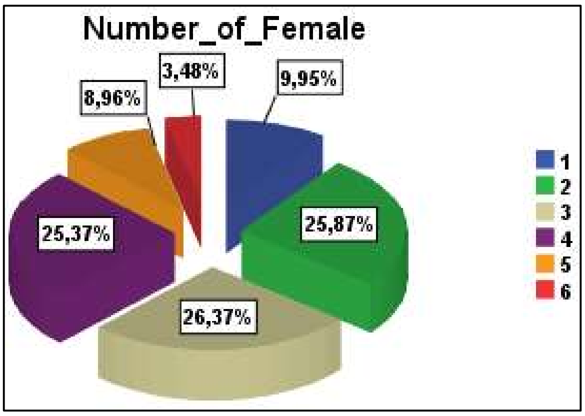


Fig 3.7 Number of female

Similarly to females, number of males (MAL) by house could affect water use in houses too. Examination of table 3.4 shows that minimum and maximum number of males in a house are 0 and 5, respectively.

The mean is equal to 2 males. Observation of Figure 3.8 gives the following classification of houses according to number of males:

- 1 house has no male (0,50 %)
- 44 building units have 1 male (21,89 %)
- 90 building units have 2 males (44,78 %)
- 51 building units have 3 males (25,37 %)
- 14 building units have 4 males (6,97 %)
- 1 building unit has 5 males (0,50 %)

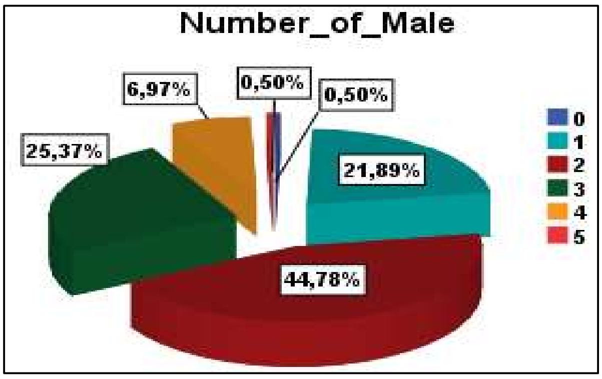


Fig 3.8 Number of males

3.4.2.2.Age of Family Members

Four main categories of age are obtained and their percentages are varied. Children category represents **24%** of the total sample (under 8 years old), with total number of **257 child**, for age category 9-15 years represents **13%** with total number equal to **133** and by this it represents the smallest age group. In addition, **40%** belong to residents having age between 15-35 years old (**425** persons) and it is the largest of the four groups. Finally, **23%** of the samples (with total number equal to **242**) belong to persons with 35 years and more. (Table 3.4 and Figure 3.9).

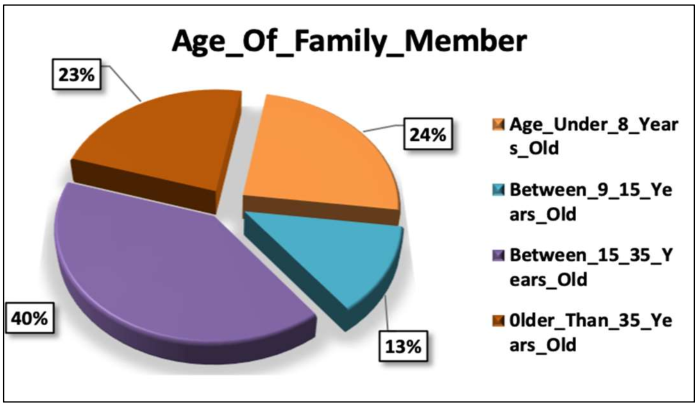


Fig 3.9 Repartition of residents age categories

a. Residents with Age under 8 years old (AG1)

Table 3.4 shows that the minimum, maximum and mean of AG1 that are 0; 3 and 2 respectively. Histogram in figure 3.10 demonstrates the distribution of children. The majority (92) of houses have a single kid in this age range.

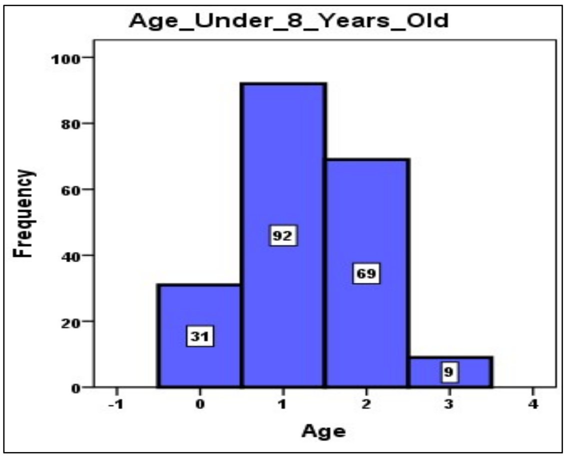


Fig 3.10 Distribution of residents with age under 8 years old

b. Residents with Age between 9-15 years old (AG2)

Table 3.4 shows the minimum, maximum and mean of AG2 that are 0; 2 and 1, respectively. Histogram in figure 3.11 illustrates the distribution of second category of residents. Results shows that most of people have one or none persons in this range of age.

Fig 3.11 Distribution of residents with age between 9-15 years old

c. Residents with Age between 15-35 years old (AG3)

The minimum, maximum and mean of category 3 of age are 1; 4 and 2, respectively (table 3.4). Histogram in figure 3.12 below shows the distribution of third category of residents. Most of houses (83) have one person in this category.

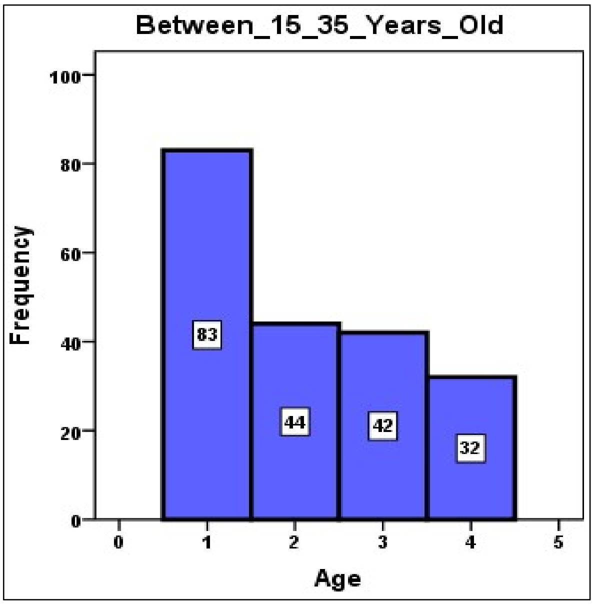


Fig 3.12 Distribution of residents with age between 15-35 years old

d. Residents with Age older than 35 years old (AG4)

Table 3.4 shows the minimum, maximum and mean of AG4 that are 0; 2 and 1, respectively. Histogram in figure 3.13 below shows the distribution of forth category of residents. The main reason behind this result is that the parents fall in this range.

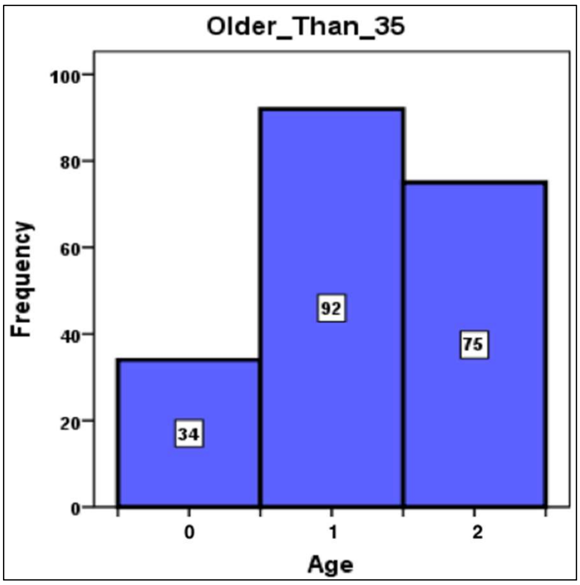


Fig 3.13 Distribution of residents with age more than 35 years old

3.4.2.3.Education Level of Residents

It should be notice that in Algeria , the study program passes by four main grades : primary school in 5 years, medium in four years , secondary (or high school) in three years and finally university program where the lowest title is obtained in three years (bachelor degree) and vary according to field of study and the degree attended. This study program is free and a typical person would pass all the cycle before acceptance in university in 12 years.

Education level affects all daily decisions and tasks, because educated or the literate persons are more aware about the consequences of shortage of water. The distribution of the sample based on educational status of the household is summarized in table 3.4 above and pie chart in figure 3.14 below.

Table 3.4 shows that people in the study zone has at least primary level and quarter of population have higher education (university). 87 % of residents have completed or in medium degree. And more than half have at least a high school level (66%). For university level, some residents are already graduated, others are still studying in the university and the rest of them are employed.

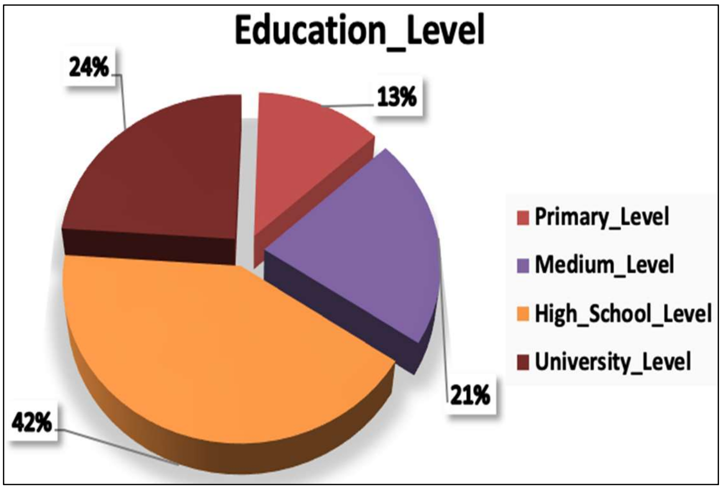


Fig 3.14 distribution of residents based on their educational level

a. **Primary School (PRS)**

Pie chart below demonstrates the repartition of habitation according to their primary level attendants (Figure 3.15).

More precisely 55, 7 % of houses do not have any pupils in primary school, 39,80% have one and only 4,5% have two in primary school.

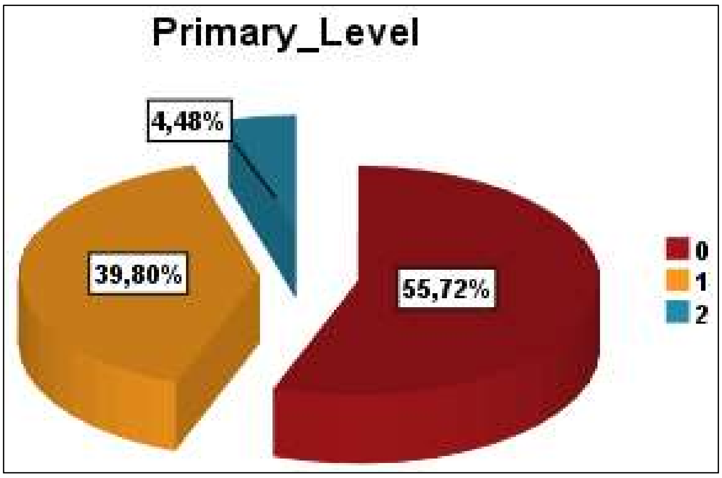


Fig 3.15 Repartition of number of habitations according to primary level

b. Medium School (MDS)

Half (52%) of houses have one student in medium level, and third of them with none and the rest has two.

Pie chart below demonstrates the repartition of number of habitations according to secondary level (Figure 3.16).

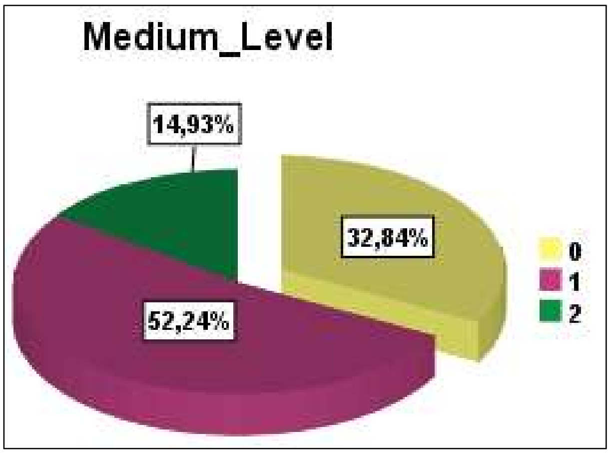


Fig 3.16 Repartition of number of habitations according to secondary level

c. High School (HGS)

For high school level, the population has at least one student in this level (57,21% with one student, 24,88 % with two students and 17,91% with three in high school), the main reason behind this result is that most parents have a high school level.

Pie chart below demonstrates the repartition of number of habitations according to high school level (Figure 3.17).

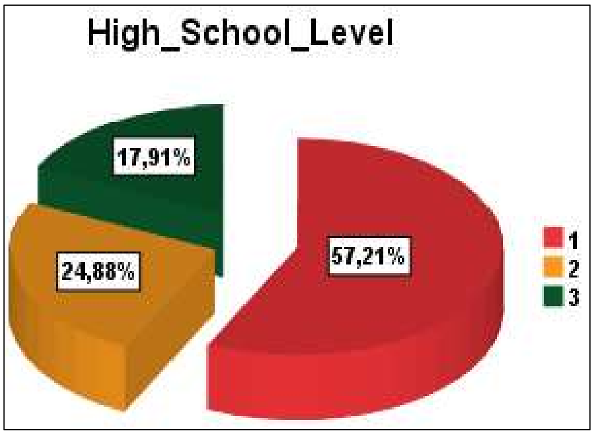


Fig 3.17 Repartition of number of habitations according to high school level

d. University Level (UNIV)

Even though 24% of population are in the university but 29% of houses does not have anyone in university level. This means that there is a tendency only in someone houses to go to university.

Pie chart below demonstrates the repartition of number of habitations according to university level (Figure 3.18).

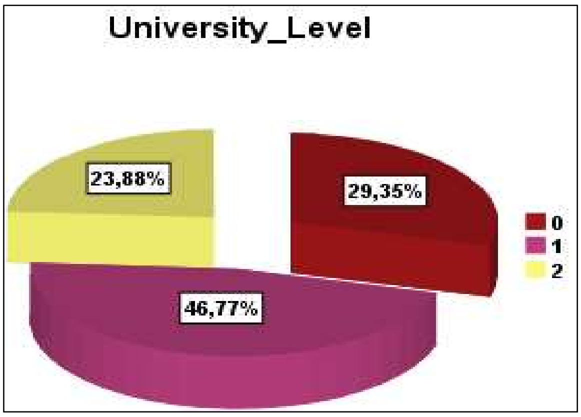


Fig 3.18 Repartition of number of habitations according to university level

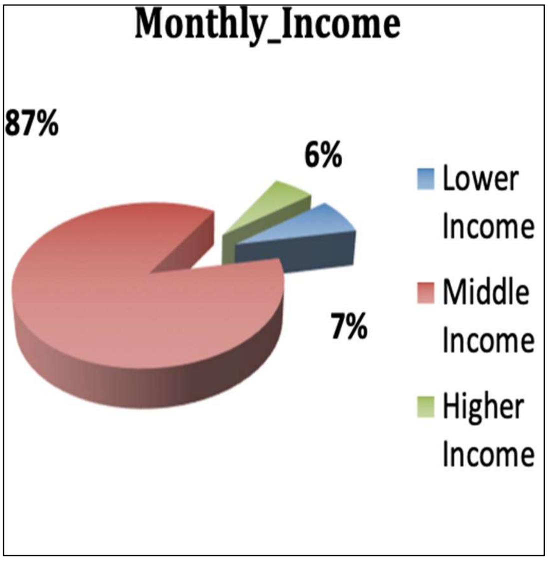
3.4.2.4.Household Income (INC)

Household income is a strong background variable determining water use and demand. Average monthly income in Algeria is around 39.900 DA in 2016 (ONS: National Statistics Office; http://www.ons.dz/). Higher income class constitutes people having average monthly income above 100.000 DA. Households having average monthly income less than 39.900 DA are categorized as lower income class. Generally, people with higher income are more concern about water but their lifestyle is differing from lower income group people.

Table 3.4 shows that the mean of monthly income is about 53 905 DA with minimum equal to 35 000 DA and maximum equal to 110 000 DA. Table 3.7 and figures 3.19 demonstrated that **6 %** of the respondents fall in the high-income class, while **7%** of respondents belong to class of lower income and **87%** falls in middle-income category.

Table 3.7 Distribution of sample based on monthly income

Income Classes	Interval	Frequency	Percentage
Lower Income	< 39 900 DA	15	7 %
Middle Income	40 000_990 000 DA	174	87 %
Higher Income	> 100 000 DA	12	6 %
Total		**201**	**100 %**



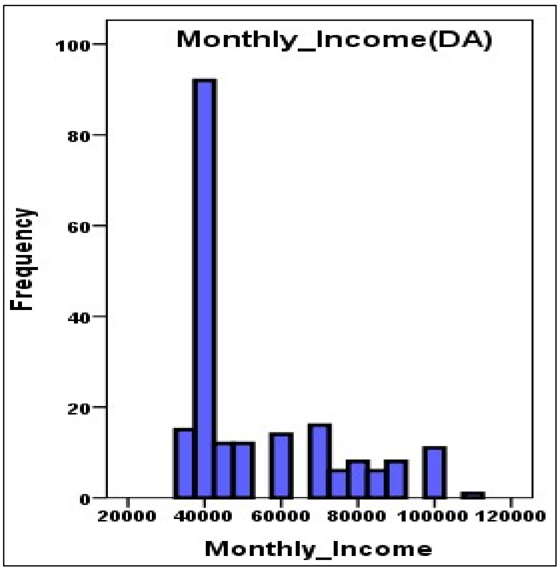
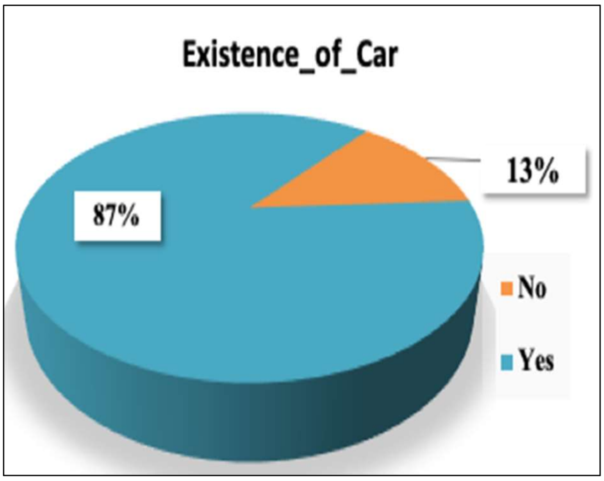


Fig 3.19 Distribution of sample based on income categorization

3.4.2.5.Cars related parameters

Three parameters of cars related are: existence, number (CARN) and washing frequency. For car possession (EXC) **87%** of households have at least one car (figure 3.20).



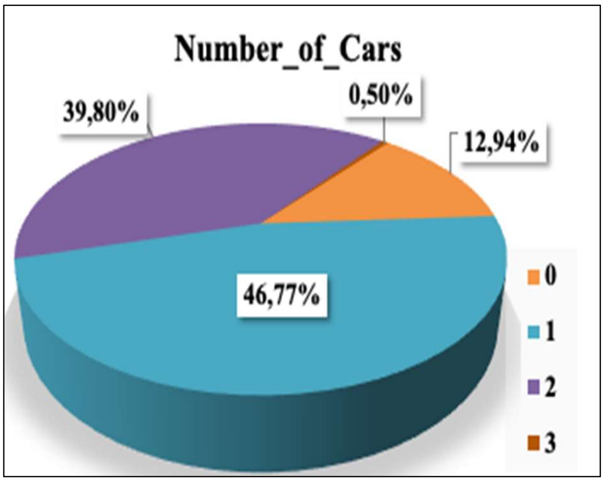


Fig 3.20 Cars possession

Fig 3.21 Repartition of buildings according to cars number

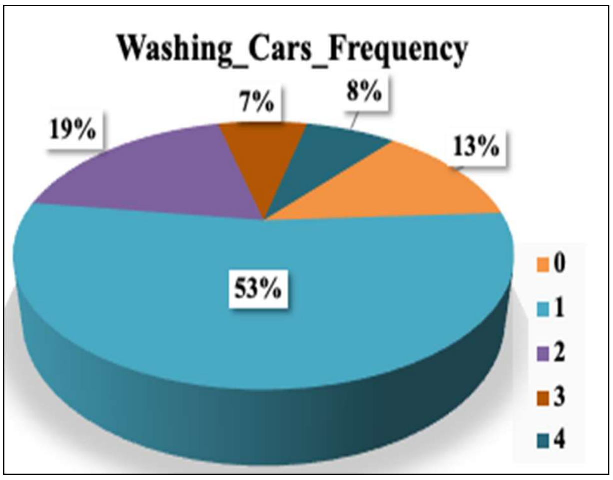


Fig 3.22 Frequency of washing cars per month

From table 3.4 above, mean value of number of cars is equal to **two** cars, minimum and maximum number of cars are equal to zero and three cars per house, respectively. Figure 3.21 demonstrates the repartition of cars in the study area.

For cars washing behaviors (WCAR), figure 3.22 demonstrates the frequency of washing cars in houses. **53%** washing their cars one time per month, **19%** washing their cars **2** times per month, **7%** washing their cars **3** times per month and **8%** of residents washing their cars **4** times per month. Table 3.4 shows that the mean of WCAR is equal to **two**.

3.4.3. Indoor Habits of Residents (INH)

Indoor Habits of Residents (INH) summarize all water related practices frequencies: Washing Dishes (WDISH), Washing Clothes (WCL), Using Toilets (UTLT), Shower for Female (FSHW) and Shower for Male (MSHW). The statistical characteristics of domestic water component is presented in table 3.8. The daily and weekly water usage was determined based on residents answers, which it collected according to the questionnaire paper.

Table 3.8 Household water usage

Statistical Parameters	WDISH (Day)	WCL (Week)	UTLT (Day)	FSHW (Week)	MSHW (Week)
Mean	3,00	2,00	4,00	2,00	2,00
Median	3,00	2,00	4,00	2,00	2,00
Mode	3	2	4	2	2
Std. Deviation	0,61	0,69	0,95	0,99	0,95
Variance	0,38	0,48	0,89	0,99	0,90
Skewness	-1,13	0,84	0,62	0,59	0,53
Std. Error of Skewness	0,17	0,17	0,17	0,17	0,17
Kurtosis	0,24	0,87	0,38	1,16	-0,52
Std. Error of Kurtosis	0,34	0,34	0,34	0,34	0,34
Minimum	1	1	3	1	1
Maximum	3	4	7	7	5
Sum	517	342	871	480	413

3.4.3.1.Washing Clothes (WCL)

Table 3.8 above shows that the mean of WCL per week is equal to **two**, the minimum and maximum frequency of washing clothes are **one** and **four** times per week. Bar chart in figure 3.23 depicted the frequency (number of time) of washing clothes per households as following:

- **41.29%** (83) wash their clothes once a week
- **49.26%** (99) wash their clothes twice a week
- **7.46%** (15) wash their clothes 3 times per week
- **1.99%** (4) wash their clothes 4 times per week.

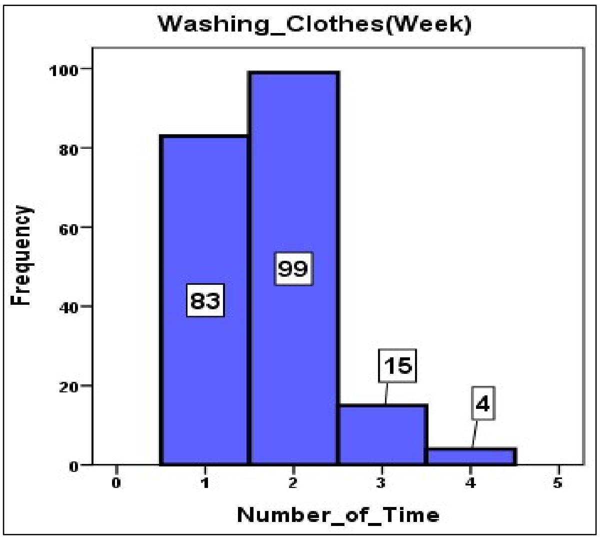


Fig 3.23 Distribution of households according to the frequency washing clothes

The following distribution explained by Algerian behaviors in washing clothes that generally correspond to weekend only and due also to water distribution system (frequency of water supplying).

3.4.3.2.Washing Dishes (WDISH)

The frequency of dishwashing ranges between once and three times a day and shows that the houses could be categorized into three main groups (figure 3.24). The reason why majority of houses have a frequency of three times a day correspond to three main dishes a day (breakfast, lunch and dinner).

- **6.47% (13)** wash dishes one time per day.
- **29.85% (60)** wash dishes **2** times per day.
- **63.68% (128)** wash dishes **3** times per day.

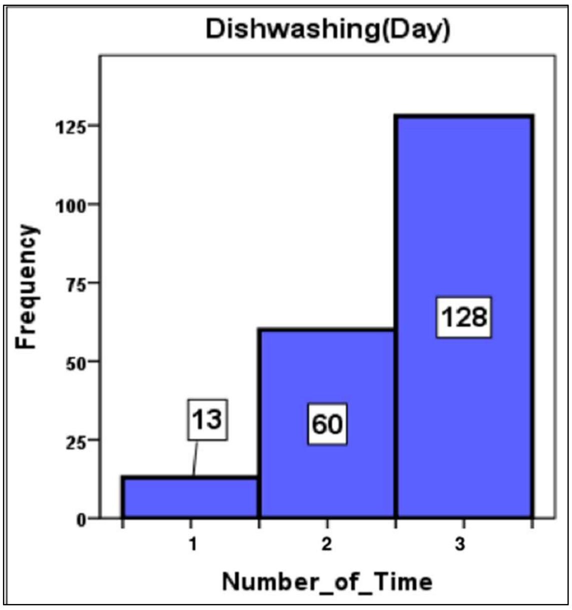


Fig 3.24 Distribution of households according to the frequency of washing dishes

3.4.3.3.Using Toilets (UTLT)

Data in the table 3.8 above shows that, the mean usage of toilet is **four**. The minimum and maximum frequency of using toilets are **three** and **seven** times per day. Bar chart in Figure 3.25 depicted the frequency of using toilets per day per person as following:

- **17.91% (36)** use toilets **3** times per day.
- **43.28% (87)** use toilets **4** times per day.
- **29.35% (59)** use toilets **5** times per day.
- **6.47% (13)** use toilets **6** times per day.
- **2.99% (6)** use toilets **7** times per day.

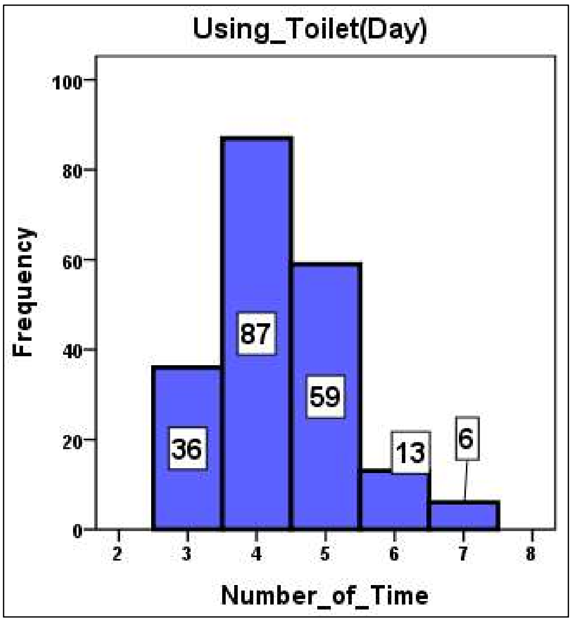


Fig 3.25 Distribution of households according to the frequency of using toilets

3.4.3.4.Showering

Showering frequencies vary considerably from house to house and person to person. Showering is also related to gender. For Females (FSHW) the bath use mean is 2 (table 3.8) above).

The minimum and maximum frequency of taking shower are **1** and **7** times per week. Bar chart in Figure 3.26 showed the frequency of taking shower per week per households as following:

- **19. 90% (40)** of females take shower for once weekly.
- **35. 82% (72)** of females take shower for 2 times per week.
- **31.84% (64)** of females take shower for 3 times per week.
- **11.44% (23)** of females take shower for 4 times per week.
- **0.50% (1)** of females take shower for 5 times per week.
- **0.50% (1)** of females take shower daily.

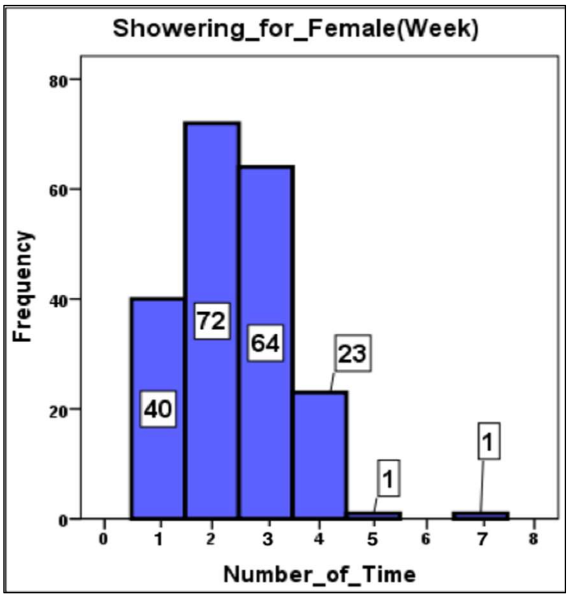


Fig 3.26 Distribution of households according to the frequency of showering for female

Data in the table 3.8 above shows that, the mean of MSHW is equal to **2**. The minimum and maximum frequency of taking shower for males are **1** and **5** times per week. Bar chart in Figure 3.27 showed the frequency (number of time) of showering for males per week per households as following:

- **33.83% (68)** of males take shower for one time per week.
- **34.82% (70)** of males take shower for 2 times per week.
- **23.88% (48)** of males take shower for 3 times per week.
- **6.97% (14)** of males take shower for 4 times per week.
- **0.50% (1)** of males take shower for 5 times per week.

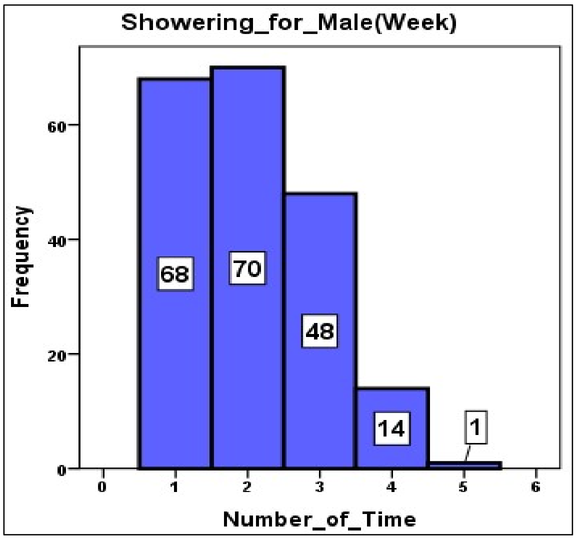


Fig 3.27 Distribution of households according to the frequency of showering for males

3.4.3.5.Who use More Water Inside Every House, Male or Female?

Generally, women use more water inside every house. Highest number of liters used by women is for the household activities. Data reflect that around **47.26% (95)** of people answered that women use more water than men for domestic purpose, whereas around **44.78% (90)** of answers goes to male and the rest of percentage **7.96% (16)** is for those answering that both male and female use the same quantity of water (Figure 3.28). Women are requiring more water because they are involved in performing maximum number of domestic works.

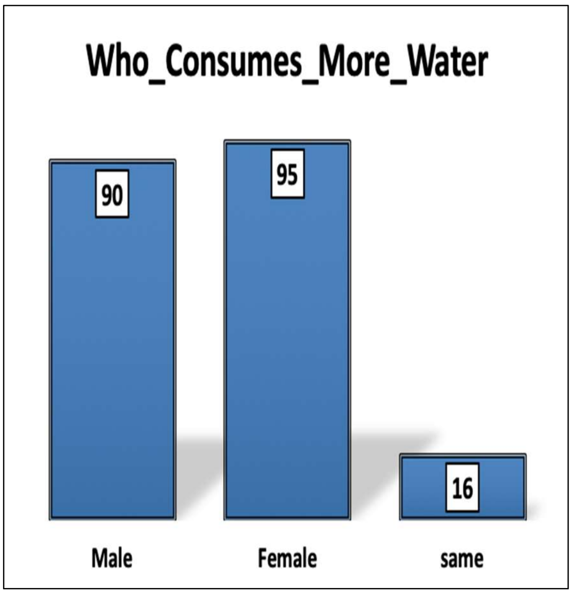


Fig 3.28 Opinion poll about the use of water between male and female

These values of domestic water use were expected considering the different factors influencing consumption, besides the period of records was associated to nearly summer conditions, in which high temperatures occurred.

3.4.4. Physical Characteristics of Housing Units (PHC)

Beside of human activities, the physical characteristics of building units could affect WCP. Five main characteristics are covered: Total Area (TAR), Building Area (BAR), Number of Rooms (ROMN), Garden Area (GAR) and Garden Watering program (GWAT).

Table 3.9 Summary of physical characteristics of houses

Statistical Parameters	TAR (m²)	BAR (m²)	ROMN	GAR (m²)	GWAT
Mean	186,67	164,97	6,00	21,70	2,00
Median	200,00	164,00	5,00	16,00	2,00
Mode	80	90	4	20	2
Std. Deviation	77,29	75,81	3,09	18,14	0,79
Variance	5973,33	5746,91	9,54	328,98	0,62
Skewness	0,15	0,19	0,92	2,16	0,56
Std. Error of Skewness	0,17	0,17	0,17	0,17	0,17
Kurtosis	-1,27	-1,19	-0,20	4,12	-0,22
Std. Error of Kurtosis	0,34	0,34	0,34	0,34	0,34
Minimum	80	40	2	2	1
Maximum	320	302	13	80	4
Sum	37520	33158	1193	4362	381

3.4.4.1. Buildings

Building unit is considering as significant indicator of water use. The type of houses influences the water use. Moreover, to consider all variables such as garden parameters, “single houses” are selected for the study. Figure 3.29 shows the repartition of houses according to their total area (TAR). Table 3.9 demonstrates that the minimum size of houses is **80 m^2** and the maximum size is **320 m^2**. While, the mean is equal to **186.67 m^2**. The samples show big variety with sixteen categories.

Table 3.9 demonstrates also the minimum size of building is **40 m^2** and the maximum size is **302 m^2**. Figure 3.30 shows the repartition of houses according to their building area (BAR). The biggest three categories according to BAT are: 100 m², 160 m² and 280 m² with 17,65%, 14,71% and 17,65% respectively.

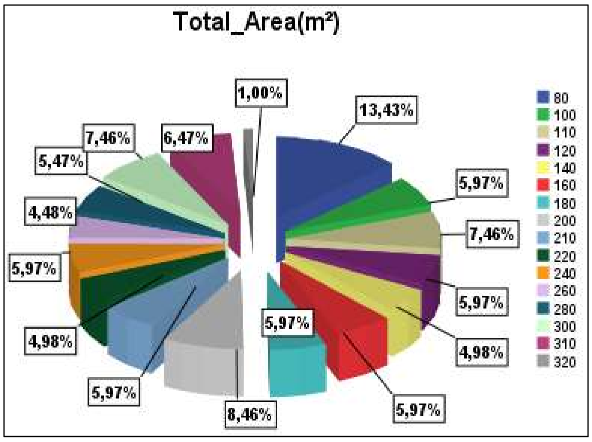


Fig 3.29 Distribution of houses according to their total area

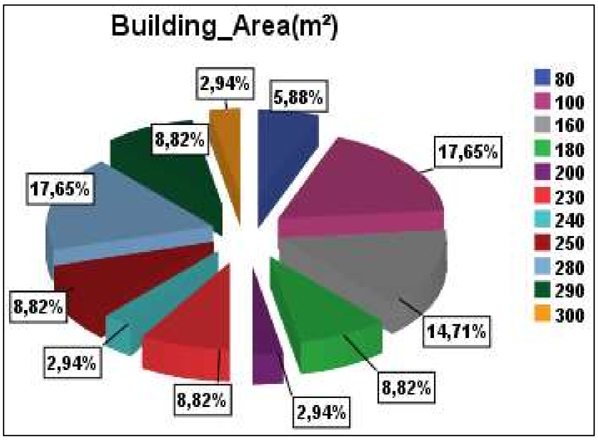


Fig 3.30 Distribution of houses according to their building size

Number of rooms per household (ROMN) is also key factor in physical characteristics, it is related directly life quality. Table 3.9 demonstrates that ROMN ranges between **2** and **13** rooms. The mean is **6** per house. Pie chart in Figure 3.31 shows the repartition of houses according to their number of rooms. The biggest category is 4 rooms and represents 22,39%.

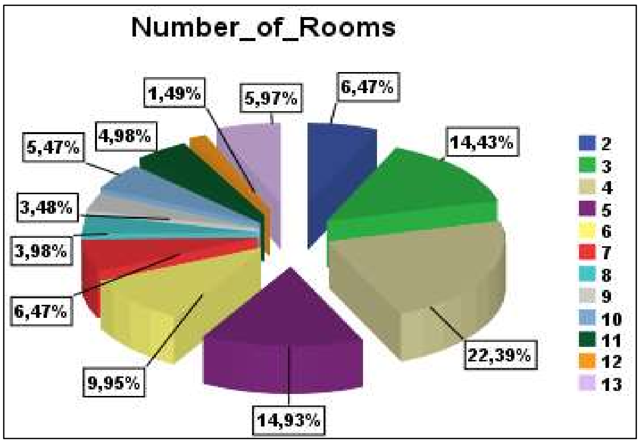


Fig 3.31 Distribution of houses according to their number of rooms

3.4.4.2. Garden Factors

Like building area, garden is an important independent indicator influencing water use. A larger quantity of water is used for gardening and lawn watering. Generally, houses built with garden are higher consumers of water compared to houses without any type of garden. In the present work, all households have a garden.

To measure the influence of garden on WCP, the area of garden (GAR) and watering frequency (GWAT) are used.

The mean of GAR is **21.70 m^2** with a maximum of **80m^2** and 13 groups. Figure 3.32 shows the repartition of houses according to their garden size. Three main size groups exist are 12, 16 and 20 square meters.

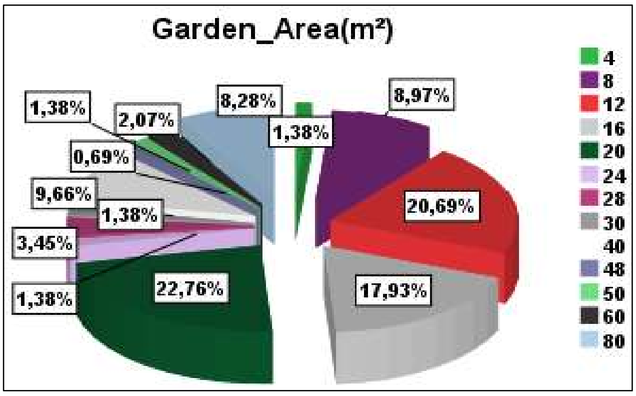


Fig 3.32 Distribution of houses according to their surface area

Residents of study area are categorized into four groups. 33,83% tend to water their garden once a month, almost a half with 45,77% twice a month, 16,41 % with three times and the smallest portion with 3% four times a month (figure 3,33). Garden watering is also related to weather and other climatic factors.

The general mean of watering frequency is 2 (twice a month).

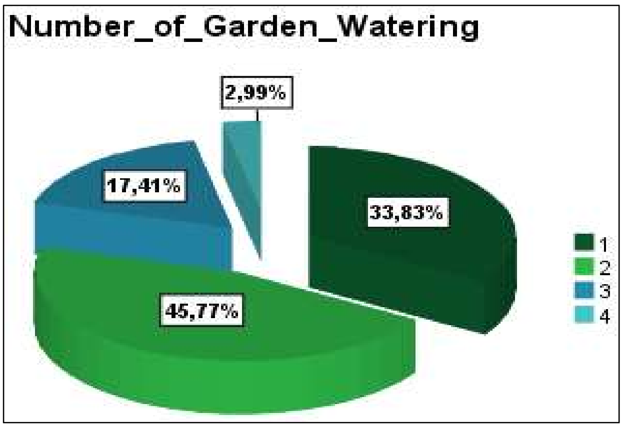


Fig 3.33 Distribution of houses according to their number of time watering garden

3.4.5. Climatic Factors (CLF)

Climatic factors dictate the WCP by influencing the human behaviours in form of changing hygienic and outdoor practices (people tend to bath more in warm season, water gardening, etc.).

The climate in Sedrata is warm and temperate in general. The rain falls mostly in the winter, with relatively little rain in worm seasons. Sedrata climate is classified as Csa by the Köppen-Geiger system. The average annual temperature is 14.2 °C | 57.7 °F. The annual rainfall is 523 mm | 20.6 inches (Website of climate data).

The impact of climate on domestic WCP was assessed through the monthly mean precipitation 'PRE' and mean temperature 'TEM'. The values of mean CLF are illustrated in figure 3.34. The link between temperature and precipitation is clear where hot seasons are characterized with low precipitation and vice versa.

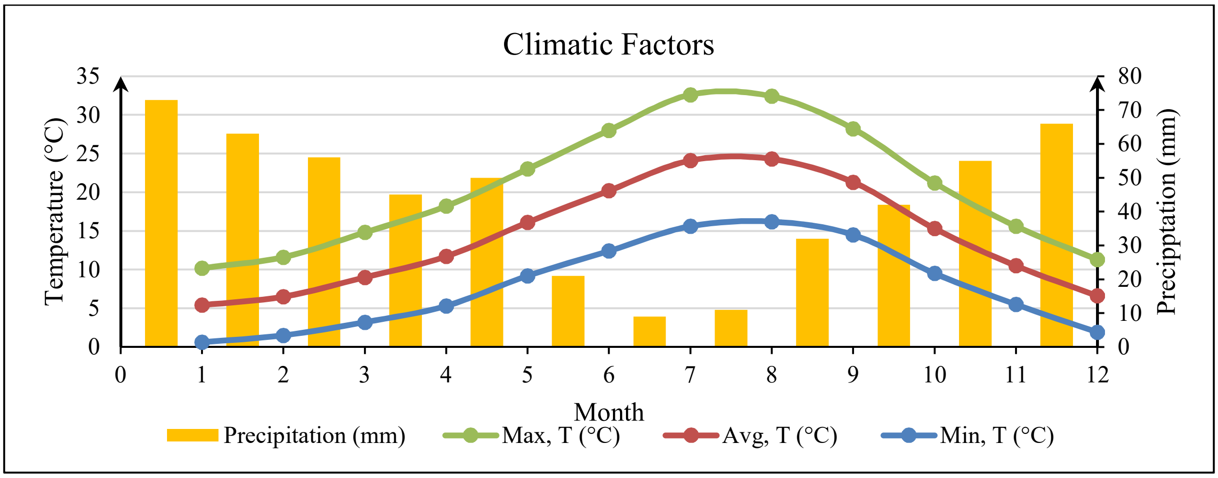


Fig 3.34 Monthly precipitation (mm) and temperature (°C)

The highest rainfall **PRE** (January **(73mm)** and December **(66mm))** correspond to coldest months, and same remark for lowest PRE and TEM (figure 3.34). This variation is typical in north Algeria and covers all the region of south Mediterranean.

About 88% of total annual rainfall occurs during the period from September to May and the rest 12% in the remaining period. The rainfall in the region shows spatial disparity and due to very rough and uneven terrain, droughts in some parts of the region and floods in some others.

On yearly scale, precipitation is highly non predictable for example the difference between 2015 and 2016 is 480 mm that represents 44,7% decrease.

3.5. Data Preparation

In any statistical analysis, the collected data must be prepared. The figure 3.35 below shows the adopted methodology of data preparation:

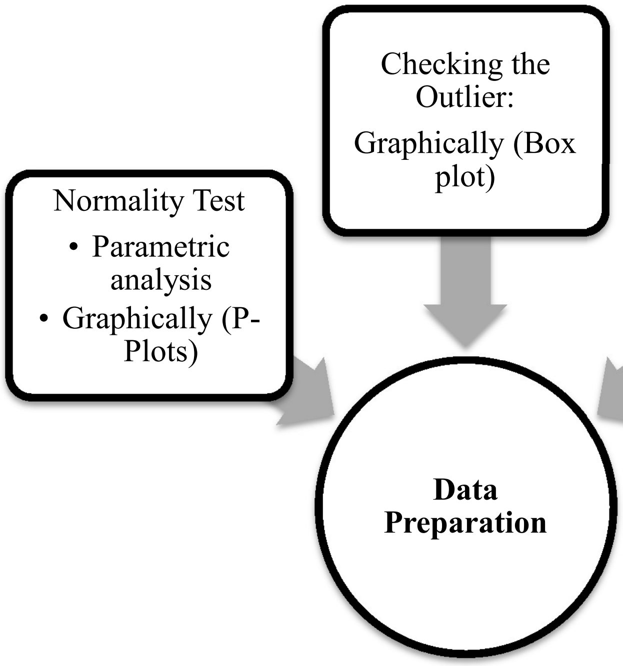


Fig 3.35 Data preparation methodology

3.5.1. Checking the Normality of Parameters

In most traditional statistical models, the data must follow a normal distribution before the model coefficients can be estimated efficiently. If this is not the case, the distribution of data should be applied. The normal distribution is one of the most important and the most widely used example of a continuous random variable. Normal distribution has a bell-shaped curve with the center of the bell located at the arithmetic mean **(μ)**.

The standard deviation **(σ)** controls the depth of this bell. It expressed by **N (μ, σ²)** and defined by two arithmetic parameters; **mean (μ)** and **variance (σ²)**.

To confirm the normality distribution, **"Skewness"** and **"Kurtosis"** are calculated using Statistical Package for the Social Sciences (SPSS) Software.

Skewness characterizes the degree of asymmetry of a distribution around its mean. Positive skewness indicates a distribution with an asymmetric tail extending towards more positive values. Negative skewness indicates a distribution with an asymmetric tail extending towards more negative values (Microsoft, 1996). Skewness value indicates the distribution side and closeness of it to zero means symmetric distribution of parameter. In normal distribution, the **mean** and **median** values should be **very close** to each other and produce a skewness statistic of about Zero.

Kurtosis characterizes the relative peakedness or flatness of a distribution compared to the normal distribution. Positive kurtosis indicates a relatively peaked (leptokurtic, too tall) distribution and negative kurtosis indicates a relatively flat (platykurtic) distribution or even concave if the value is large enough (Microsoft, 1996). Normal distributions produce a kurtosis statistic of about **zero**.

For a random variable *x*, when *log (x)* probability distribution is normal, probability distribution if *x* is defined as logarithmic normal distribution (log-normal distribution). In this case, the distribution of parameter deviates from the symmetry. Also, checking the normality can be done using *Z* score of *Kurtosis* and *Skewness* of each variable. It calculated by the following formula:

$$Z_S = \frac{Skewness}{\sqrt{\frac{6}{N}}} \qquad \textbf{(3.1)}$$

$$Z_K = \frac{Kurtosis}{\sqrt{\frac{24}{N}}} \qquad \textbf{(3.2)}$$

Z_k and Z_S values should be within ±2.58 for p=0.01 or ±1.96 for p=0.05 for considering data as normal distributed. In equation below, Z normal distribution is calculated by using a table. Z_{95} value means the Z value corresponding to 0.95 probability value (=1.65).

$$Z = \frac{X - \mu}{\sigma} \qquad \textbf{(3.3)}$$

With: x is x_{95}; μ is mean value; σ is standard deviation and Z is Z_{95} value.

According to results in tables 3.2; 3.3; 3.4; 3.8 and 3.9 the distribution of each parameter is given in table 3.10 below. Table 3.10 shows that WCP 2014, WCP 2017 and WCP of first trimester have small and positive skewness and kurtosis, which indicate a normal distribution of data. The rest of water consumption variables have big and positive skewness and kurtosis, which indicate reasonably, have non-normal distribution.

For socio-economic parameters, all the parameters have small and negative kurtosis. Only household size, the fourth age category and number of cars have small and negative skewness. Whilst, the rest of socio-economic parameters have positive skewness. As a result, all these parameters have **normal** data distribution.

Total area, building area, number of rooms and frequency of garden watering **in physical characteristics** of building units have small positive skewness and small negative kurtosis. These parameters indicate **normal** data distribution. Only garden area has big positive kurtosis and skewness which reasonably has **non-normal** distribution. All **indoor habits** variables have normal distribution with small positive skewness and kurtosis for washing clothes, using toilets and shower for female frequencies. While, the frequency of dishwashing has small negative skewness with small positive kurtosis. Also, shower for male has small skewness and small negative kurtosis.

Table 3.10: Statistical distribution of parameters

Statistical Values	**Skewness**	**Kurtosis**	**Distribution**
Water Consumption			
WCP_2012	1,21	2,68	**Non-Normal**
WCP_2013	2,33	12,67	**Non -Normal**
WCP_2014	0,65	0,25	Normal
WCP_2015	3,23	20,23	**Non -Normal**
WCP_2016	1,39	2,72	**Non -Normal**
WCP_2017	0,98	0,84	Normal
Mean_WCP_1_Trimester From 31 Jan to 31 Mar	0,89	1,01	Normal
Mean_WCP_2_Trimester From 31 Mar to 31 Jun	1,02	2,01	**Non -Normal**
Mean_WCP_3_Trimester From 31 Jun to 31 Sep	1,33	3,37	**Non -Normal**
Mean_WCP_4_Trimester From 31 Sep to 31 Dec	1,95	7,91	**Non -Normal**
Socio-economic parameters, Physical characteristics of buildings & Indoor habits			
FEM	0,24	-0,51	Normal
MAL	0,38	-0,16	Normal
HOUS	-0,37	-0,1	Normal
AG1	0,05	-0,47	Normal
AG2	0,51	-0,73	Normal
AG3	0,46	-1,21	Normal
AG4	-0,32	-0,98	Normal
PRS	0,73	-0,44	Normal
MDS	0,22	-0,78	Normal
HGS	0,81	-0,6	Normal
UNIV	0,08	-1,10	Normal
INC	1,09	-1,13	Normal
CARN	-0,33	-0,72	Normal
WCAR	0,85	-0,05	Normal
WDISH	-1,13	0,24	Normal
WCL	0,84	0,87	Normal
UTLT	0,62	0,38	Normal
FSHW	0,59	1,16	Normal
MSHW	0,53	-0,52	Normal
TAR	0,15	-1,27	Normal
BAR	0,19	-1,19	Normal
ROMN	0,92	-0,20	Normal
GAR	2,16	4,12	**Non -Normal**
GWAT	0,56	-0,22	Normal

To confirm the results from table 3.10 above, graphs like Histograms and probability plots are obtained for this purpose.

Probability plots are for variables cumulative proportions against the cumulative proportions of any of a number of test distributions. They are generally used to determine whether the distribution of a variable matches a given distribution. The points cluster around a straight line if variable matches the test distribution. P-P graphs plot the cumulative probabilities (values range from 0 to 1), with observed probabilities (cumulative proportion of cases). In normal P-P plots, normal distribution of data set is on y-axis (expected cumulative probabilities), while in log-normal P-P plots, log-normal distribution of data set is on y-axis (expected cumulative probabilities).

Basing on these normal and log-normal P-P plots, distribution of data set is confirmed graphically by related their distribution type in figures 3.36 below.

Fig 3.36 P-P plots for variables (ANNEX 03)

3.5.2. Checking the outliers

The best technique to identify the outliers is the **"box plot"**, as illustrated in Figures 3.37 below.

Fig 3.37 Box plot for variables (ANNEX 04)

The box plots of variables are examined as part of the diagnostic phase of data preparation, to conduct the statistical and numerical techniques. The figure 3.37 shows the existence of outliers in some variables. This last should be removed from data sets.

3.6. Conclusion

Water Scarcity in semi-arid areas is a major problem encountered by city planning. Moreover, providing enough and drinkable water all day still challenging and to better rationalize water usage, all parameters governing water consumption should be considered in detail.

Water Consumption is clearly a multi-variable function where each variable has a different weight and pattern. For measuring reliably their impact on WCP, all related parameters are collected directly from authorities or from questionnaire for indoor/outdoor consumers habits.

The pre-processing task ensured that the obtained data is representative and the subsequent analyses are all valid. In fact, after removing outliers 201 household remained with dataset of 4824 valid water consumption values.

Parameters affecting WCP are categorized into three: indoor habits, socio-economic parameters and physical characteristics of buildings. This distinction between inputs helps to asses separately the significance of every parameter.

One of the most important statistical features is the normality where tests like Kurtosis and Skewness are employed. Finally, when data does not follow a normal distribution, it must be normalized because the AI modelling requires a normality distribution.

Chapter 4:
Analysis Methodology and Tools

4.1.Introduction

In almost studies performed in industrialized countries, the residential water demand function is specified as a single equation linking tap water use (the dependent variable) to water price and a vector of demand shifters (like household socio-economic characteristics, housing features, climatologic variables, etc.) to control for heterogeneity of preferences and other variables affecting water demand (Agthe and Biillings, 1987).

Most of the models that are employed in residential water demand study both in the developed and developing countries are regression models. They typically use the form:

Q=f (P, Z) where P is the price variable and Z are the factors or a range of shifters of demand such as income, household demographics and other characteristics such as weather variables, etc. (Arbuès et al., 2004).

The current chapter presents the proposed methodology for assessing the relationships between WCP and the other parameters by using statistical techniques. The subject of statistics is based around the idea that when it exists a big set of data and the purpose is to analyze that set in terms of relationships between individual points.

The techniques used in this thesis have the following objectives:

- Determine the possible association between the variables.
- Establishing the cause-effect relation between the data set.
- Measuring the strength of association between the variables and the direction of the relationship.
- Create and chose the best model for predicting residential water in the study area.

4.2.General Methodology

The chapter will attempt to give some elementary background mathematical skills that will be required to understand the process of many statistical analyses. These analyses will explore the effect of some parameters on water use. Firstly, the methodology covered mathematical techniques used for the investigation, followed by an explanation about the reason why such technique may be used and what the result of the operation tells about the data. Besides, in the second part an overview of the software, tools, used during the analysis. For more details, multivariate analysis will be conducted to see if there is any predictive relationship between water use and the other factors. The examination of results will be presented in the next chapter of this thesis.

There is a huge variety of machine learning algorithms existing nowadays.

The methodology is composed of many algorithms used for this research and presented in the following steps:

Firstly, to figure out which variables are connected together and to study the strength of a relationship between these continuous variables, **Correlation Analysis** will be used for these purposes followed by **Correlation Matrix**. This last will used to compute the correlation coefficients between variables.

Second step is **ANOVAs Test.** An overview about the use of ANOVA in statistics will be presented. The purpose of this test is to assess any differences in WCP in houses according to every independent variable.

In the third step, **Factor Analysis (FA)** is used, followed by **Principal Components Analysis (PCA)** in the fourth step. This analysis uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. Then the fifth step is **Cluster Analysis (CA)**, which divides data into homogeneous and distinct groups (clusters). FA, PCA & CA were uses because they summarize data so that relationships and patterns can be easily interpreted and understood.

The two final steps represent **the Artificial Neural Networks (ANNs) and the Adaptive Neuro Fuzzy Inference System (ANFIS)**. The two techniques were used to configure or to reject the previous tests results, *i.e.,* to assess the main determinants of household water use.

Figure 4.1 presents the proposed methodology used in the present thesis. It attempts to better understand the relationship between household water use and indoor or outdoor factors.

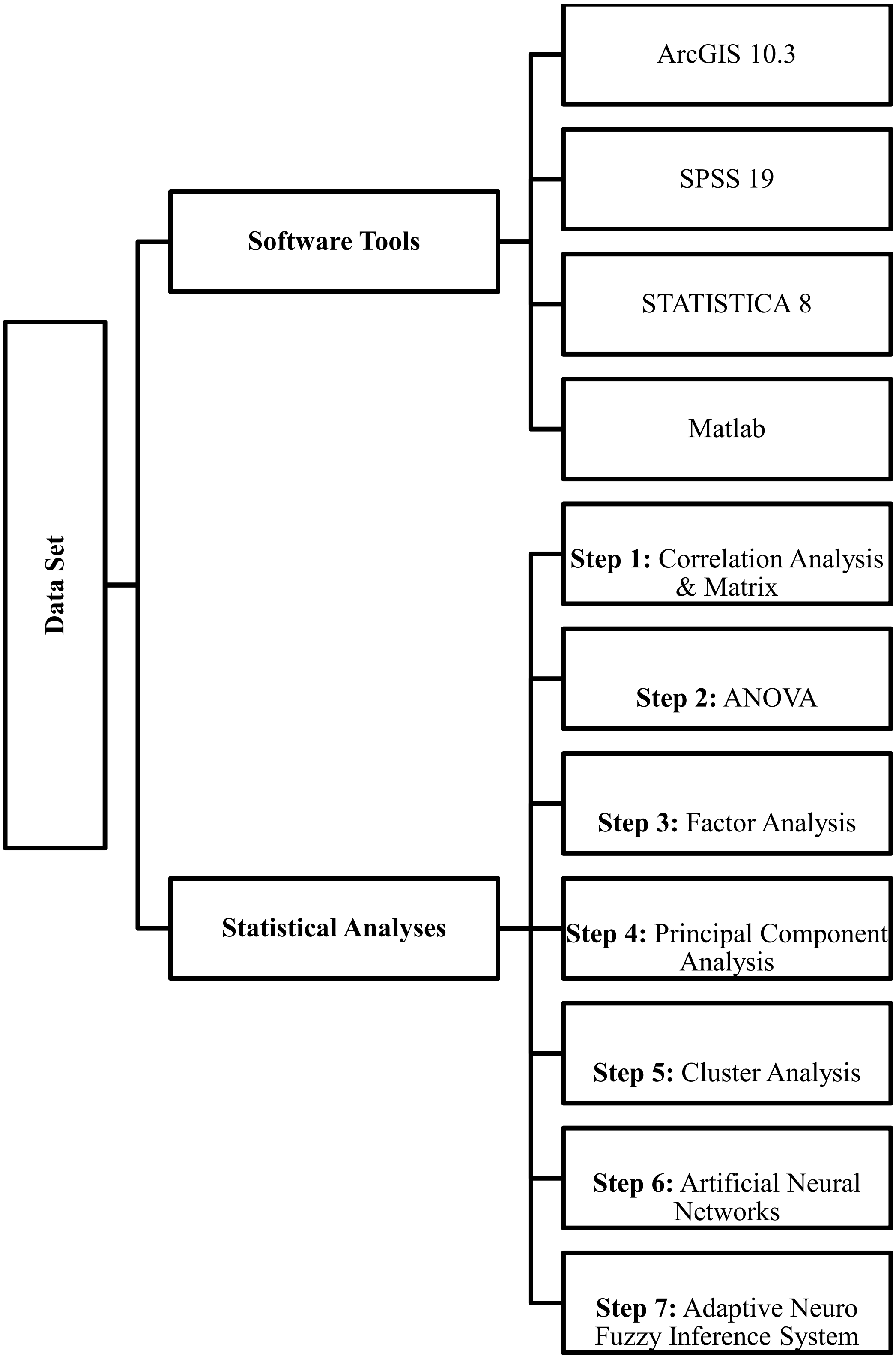


Fig 4.1 Proposed methodology

4.3. Statistical Analysis

4.3.1. Correlation Analysis and Correlation Matrix

The correlation analysis is the statistical technique used to study the closeness of the relationship between two or more variables. The variables are correlated when the movement of another variable accompanies the movement of one variable. The purpose of such analysis is to find out if any change in the independent variable results in the change in the dependent variable or not. Hence, with the correlation analysis the degree of relationship between these variables can be measured in one figure. Usually in statistics, four types of correlations could be measured for this purpose: Pearson correlation, Kendall rank correlation, Spearman correlation and the point-Biserial correlation. In this research the focus will be only on Pearson correlation because is the most widely used.

a) Karl Pearson's coefficient of correlation

The Pearson correlation is widely used mathematical method wherein the numerical expression is used to calculate the degree and direction of the relationship between linear related variables. The coefficient is popularly known as a **Pearsonian Coefficient of Correlation** denoted by **"r"**. If the relationship between two variables X and Y is to be ascertained, then the following formula is used:

$$\boldsymbol{r} = \frac{\sum(\boldsymbol{X}-\bar{\boldsymbol{X}})(\boldsymbol{Y}-\bar{\boldsymbol{Y}})}{\sqrt{\sum(\boldsymbol{X}-\bar{\boldsymbol{X}})^2}\sqrt{(\boldsymbol{Y}-\bar{\boldsymbol{Y}})^2}} \qquad \textbf{(4.1)}$$

Where: $\bar{X}$ is the mean of X variable and $\bar{Y}$ is the mean of Y variable. The value of "r" is always lies between$\pm\mathbf{1}$, such as:

r = +1 indicates perfect and positive correlation.

r= -1 indicates perfect and negative correlation.

r= 0 means no correlation.

Regression analysis is widely used for prediction and forecasting, where use has substantial overlap with the field of machine learning. It is also used to understand which among the independent variables are related to the dependent variable, and to explore the forms of these relationships. The degree of which the variables are correlated to each other depends on the Regression Line, *i.e.* all the points plotted are connected via a line in the manner that the distance from the line to the points is the smallest. The differences between correlation and regression are:

- The correlation coefficient measures the "degree of relationship" between variables, X and Y whereas the regression analysis studies the "nature of relationship" between variables.

- Correlation coefficient does not clearly indicate the cause/effect relationship between the variables, i.e. it cannot be said with certainty that one variable is the cause, and the other is the effect. Otherwise, the regression analysis clearly indicates the cause-effect relationship between the variables.

b) Regression line

The regression line is the line that the best fits the data, such that the overall distance from the line to the points (variable values) plotted on a graph is the smallest. In other words, a line used to minimize the squared deviations of predictions is called as the regression line.

c) Regression equation

The regression equation is the algebraic expression of the regression lines. It is used to predict the values of the dependent variable from the given values of independent variables. It can be expressed as follows:

$$\mathbf{Y = a + b\,X} \qquad \mathbf{(4.2)}$$

Where Y is the dependent variable (or sometimes, the outcome, target or criterion variable), X is the independent variable (or sometimes, the predictor, explanatory or regressor variables), a and b are the two unknown constants that determine the position of the line.

d) Assumptions

- **Assumption 1:** the dependent variable should be measured on a continuous scale (i.e. measured in hours, in Kg…etc).
- **Assumption 2:** the existence of two or more independent variables.
- **Assumption 3:** the existence of independence of variations (i.e. independence of residuals), which can be easily checked by using the Durbin-Watson statistics.
- **Assumption 4**: linear relationship between (a) the dependent variable and each of your independent variables, and (b) the dependent variable and the independent variables collectively. Using scatterplot and partial regression plots can check these relationships.
- **Assumption 5:** the data needs to show homoscedasticity, which is where the variance along the line of best fit remains similar as the moving along the line.
- **Assumption 6:** the data must not be multicollinearity, which means the existence of two or more correlated independent variables with each other.
- **Assumption 7:** no significant outliers in the data.
- **Assumption 8:** the dataset must be approximately normally distributed (i.e. fit the shape of a bell curve).

4.3.2. ANOVA Test

The acronym ANOVA refers to **analysis of variance** and is a statistical procedure used to test the degree to which two or more groups vary or differ in an experiment. In other words, they help to figure out if there is a need to reject the null hypothesis or accept the alternate hypothesis. Basically, a null hypothesis is the assumption that there will be no differences between groups that are tested. There are two main types of ANOVA, one-way and two-way test. One-way and two-way refers to the number of independent variables in the dataset.

a. One-way ANOVA

Used between two groups to see if there's a difference between them. It used to compare two means from two independent groups using the f-distribution. The null hypothesis for the test is that the two means are equal. Therefore, a significant result means that the two means are unequal.

b. Two-way ANOVA

Used when you have two groups, *i.e.* has two independent variables (can have multiple levels). Levels are different groups in the same independent variable.

The f value in ANOVA is a tool to help answering the question "is the variance between the means of two populations significantly different?" The f ratio is a test statistic, calculated as:

F value = variance of the group means (Means square between)/ mean of the within group variances (Mean squared error).

c. Assumptions

- The population must be closer to normal distribution
- Samples must be independent
- Population variances must be equal
- Groups must have equal sample sizes

4.3.3. Factor Analysis (FA)

In this part, a method will be used to restructure the dataset by reducing the number of variables; often called "data reduction" or "dimension reduction" technique. The way that the information contained is measured is by considering the variability within and co-variation across variables that is the variance and co-variance (correlation).

Either the reduction might be by discovering that a particular linear compensation of the variables accounts for a large percentage of the total variability in the data or by discovering that several of the variables reflect another "latent variable".

Broadly the process used three ways:

- To discover the linear combinations that reflects the most variation in the data.
- To discover if the original variables are organized in a particular way reflecting another "latent variable" (called exploratory factor analysis_ EFA).
- To confirm a belief about how the original variables are organized in a particular way.

The factor analysis is a technique of clumping subgroups of variables together based on their correlations and often just by looking at the correlation matrix and spotting clusters of high correlations between groups of variables can show what the factors are going to be. (Note that: Latent variable= construct= factor).

a. Perform factor analysis

In some type of statistic associated probability density function to produce a *p value*. Two such statistics are **the Bartlett test of Sphericity** and **the Kaiser-Meyer-Olkin** measure of sampling adequacy (usually called the **MSA**).

The Bartlett test of Sphericity compares the correlation matrix with a matrix of zero correlations (called the identity matrix). It is used to reject or not the hypothesis according to which the variables are not correlated. While, The MSA does not produce a p value but some researchers like Norman & Streiner in p198 recommend that it's better to remove variables with an MSA below 0.7.

b. Extracting the factor analysis

Principal components and Principal axis are two factoring extraction methods. PCA is not a type of factor analysis but it gives very similar results.

c. Rotation for factor analysis

Rotations can be applied on the factors. Several methods are available including Varimax, Quartimax, Promax, etc.

4.3.4. Principal Component Analysis (PCA)

PCA is a powerful and popular multivariate analysis method that can identify patterns in datasets with quantitative variables, and expressing the data in such a way as to highlight their similarities and differences. The aim of this technique is the extraction of the important information from this data set and representation of it with a set of new orthogonal variables that are called as principal components (PCs) (Abdi, et al., 2010).

a) The process of conducting a PCA

- The inter-correlations amongst the items are calculated yielding a correlation matrix.
- The inter-correlated items or "factors" are extracted from the correlation matrix to yield "principal components".
- These "factors" are rotated for purposes of analysis and interpretation.

b) Eigenvalues and inertia

Eigenvalues are the amount of information (inertia) summarized in every dimension. The first dimension contains the highest amount of inertia. The number of Eigenvalues is equal to the number of non-null Eigenvalues.

c) The correlation circle or variables chart

The correlation circle shows the correlations between the components and the initial variables. The variables can be displayed in the shape of vectors.

d) The Biplots

The Biplots represent the observations and variables simultaneously in the new space. There are different types of Biplots; correlation biplot, distance biplot and symmetric biplot.

e) Assumptions for conducting FA, PCA and CA

- A large enough sample size to allow the correlations to converge into mutually exclusive "factors".
- Normality and linearity of dataset
- The sample must be relatively homogeneous
- No significant outliers

4.3.5. Cluster Analysis (CA)

One of the useful analytical tools is unsupervised clusterization, where the data is classified by the algorithm into specified number of classes based on internal patterns. It can be used to search for the subtypes and subclasses for researched process, value or compound (Likas et al., 2003). There are a number of different methods that can be used:

a. Hierarchical methods

- *Agglomerative methods*: in which subjects start in their own separate cluster. The two 'closest' (most similar) clusters are then combined and this is done repeatedly until all subjects are in one cluster. At the end the optimum number of clusters is then chosen out of all cluster solutions.
- *Divisive methods:* in which all subjects start in the same cluster and the above strategy is applied in reverse until every subject is in a separate cluster.

b. Non-hierarchical methods (Known as K-means clustering methods)

In this work the concentrate will be on the former rather than the latter.

When carrying out a hierarchical cluster analysis, the process can be represented on a diagram known as a Dendrogram. This diagram illustrates which clusters have been joined at each stage of the analysis and the distance between clusters at the time of joining. If there is a large jump in the distance between clusters from one stage to another then this suggests that at one stage clusters that are relatively close together were jointed whereas, at the following stage, the cluster that were jointed were relatively far apart. This implies that the optimum number of clusters may be the number present just before that large jump in distance. The data is classified firstly by setting k centroids, which will be the core to the searched classes. Then, the grouping is done by minimizing the sum of squares of distances between data and the corresponding cluster centroid. At each of iteration cluster center is recalculated until the best position is reached (Likas et al., 2003).

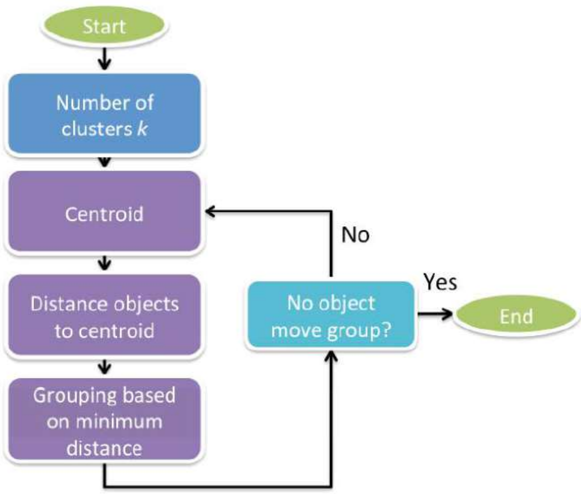


Fig 4.2 K-means clustering algorithm (Likas et al., 2003)

4.3.6. Artificial Neural Networks (ANNs)

This part of the chapter introduces brief information about artificial neural networks. Over the past, there has been an explosion of interest in neural networks. It started with successful application of this powerful technique across a wide range of problem domains, in areas as diverse as finance, medicine, engineering, geology and even physics. The artificial neural networks (ANNs) are one of the most popular machine learning models nowadays. It was first introduced in 1943 (McCulloch and Pitts, 1943) research into application of ANNs has blossomed since the introduction of back propagation training algorithm for feed forward ANNs in 1986 (Rumelhart *et al.*, 1985).

The basic of this model is in several layers that are made up of a number of interconnected nodes, containing the activation function. Data is fed into the input layer (independent) and is transformed by weights and neurons through the hidden layers (one or more hidden layers), then sent to the output layer (dependent). Produced output is sent to outer world as result. ANNs have different types such as Multilayer Perceptron Neural Networks (MLP) and General Regression Neural Network (GRNN). In this paper, MPL model was adopted since its training pattern propagates several times until a lower error is achieved.

a. General Characteristic of ANNs

More recently, artificial intelligence techniques (AI) are adopted increasingly in view of their ability to do non-linear curve fitting and applicability to a very complex data set, it used where the structure of the model is unknown and dealing with noisy data (Tiwari and Adamowski, 2013). Additionally, “Soft Computing” such as Artificial Neural Networks (ANNs) and Fuzzy Logic (FL) are more efficient and less time consuming in modeling complex systems (Pahlavan *et al.*, 2012; Sonmez *et al.*, 2018).

ANNs are adaptive where they take data and learn from it. They can reduce development time by learning underlying relationships even if they are difficult to find and describe. They can also solve problems that lack existing solutions. ANNs can generalize, where they can correctly process data that only broadly resembles the data they were trained on originally. Similarly, they can handle imperfect or incomplete data, providing a measure of fault tolerance. Generalization is particularly useful in practical applications because real world data is noisy. ANNs can capture complex interactions among the input variables in a system. They are highly parallel, i.e., their numerous identical, independent operations can be executed simultaneously (Aggrawal and Song, 1997).

b. Disadvantages of ANNs

Like any approach, ANNs has also disadvantages. They can be difficult to account for their results. ANNs are like human experts and express opinions that they cannot easily explain. In addition, training methods are imperfectly understood, where few definite rules exist for choosing the optimum architecture and there is no definite way of finding the best solutions which also depend in practice on the accuracy of the training data used. They can consume large amounts of computer time, especially during training (Aggrawal and Song, 1997).

c. Basic Components of ANNs

The basic of ANNs model is in several layers that are made up of a number of interconnected nodes, containing the activation function. The training set is presented to a model through input layer; one or more hidden layers perform processing by the system of weighted connections, taking each of the inputs for calculation and finally output gives the fitted function (Cheng and Titterington 1994).

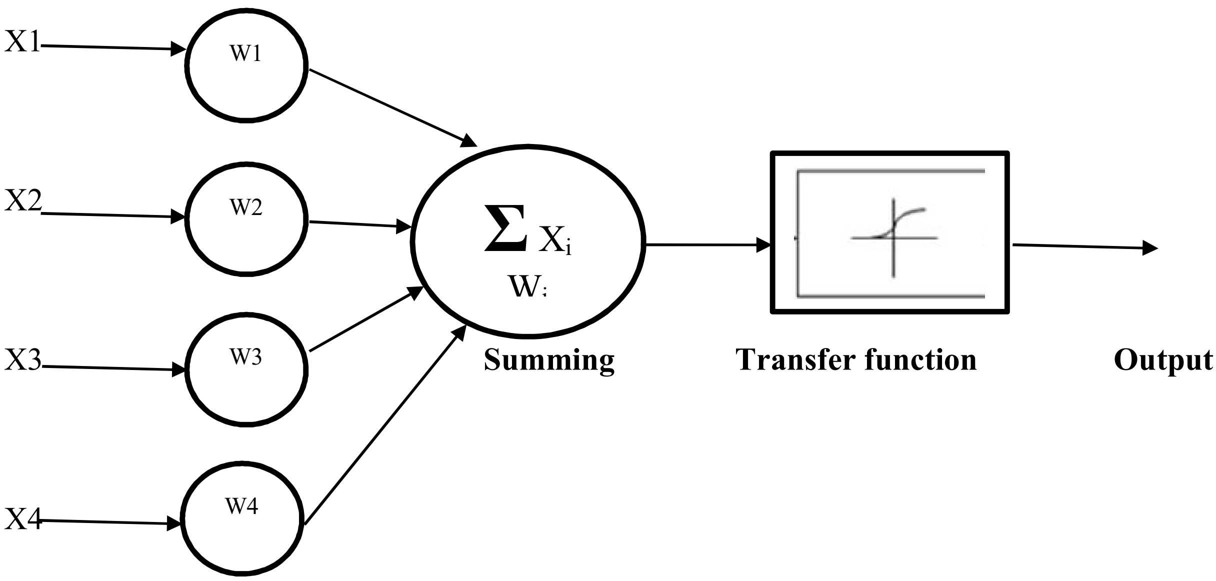


Fig 4.3 Schematic representation of an artificial neuron

The components of ANNs are:

- **Inputs**

The inputs (X_1, X_2, X_3… X_i) are elements of ANNs, which take data from outer world. They do not have any other functions than transform data to the next step. A neuron can have unlimited number of inputs but, there must be only one output of every neuron.

- **Weights**

The weights (W_1, W_2, W_3…W_i) are most important elements of mathematical neuron by itself and ANNs generally. Because learned data by the networks is stored on the weights. Accordingly, network's synaptic weight vector expressing as W= $[W_i]_{nx}$ consists convertible values. Typically, initial values of weights are selected as a random value in (-1, 1) range. How to set up of a weight for learning relationship between the given variables is decided based on selected learning rule. Weights of ANNs can be thought as synapse in biological nervous system.

- **Summation Function**

The summation function is responsible for summation of all data coming from outer world and related weights. It is shown in equation (5.3):

$$\mathbf{vi = \sum_{i=1}^{n} xi.wi - \theta} \qquad \textbf{(4.3)}$$

Summation function transfers the created weighted input (v_i) to activation function. (θ) Value expresses the threshold value. Use of threshold value in summation function is not obligatory. Some other functions that can be used of summation function are given in the following table.

Table 4.1 Some functions can be used instead of summation function

Functions	Weights input (v_i)
Multiplication	$v_i = \prod_{i=1}^{n} xiwi$
Maximum	v_i= max (x_i w_i)
Minimum	v_i= min (x_i w_i)
Signum	v_i= sgn (x_i W_i)

- **Activation (Transfer) Function**

The activation function is responsible for activation of coming weighted input (v_i) and determination of the final output value. It varies based on type of problem and there is no a universal formula for which type of activation function should be used. The choice of activation function depends largely on available data and what the designed learning of network. Sigmoid and hyperbolic tangent functions are most commonly used. The equation below is the mathematical expression of activation function:

$$\mathbf{f\,(v_i) = y} \qquad \textbf{(4.4)}$$

- **Output Function**

The output function is responsible for transfer of output value of activation function to outer world as network's final output value or to other connected neuron as their input values.

The artificial neuron components come together in three layers and parallel to each other for forming ANNs, not randomly, as follow:

- **Input layer:** Neurons transfer the information coming from outer world to hidden layers.
- **Hidden layers:** Information coming from input layer is processed and sent to output layer. There can be more than one hidden layer.
- **Output layer:** In this final layer, process elements. Transferred information from hidden layers are handled and produce output for given data in input layer. Produced output is sent to outer world as result (Lippmann, 1987). Layers are connected together with weights, as shown in the next figure.

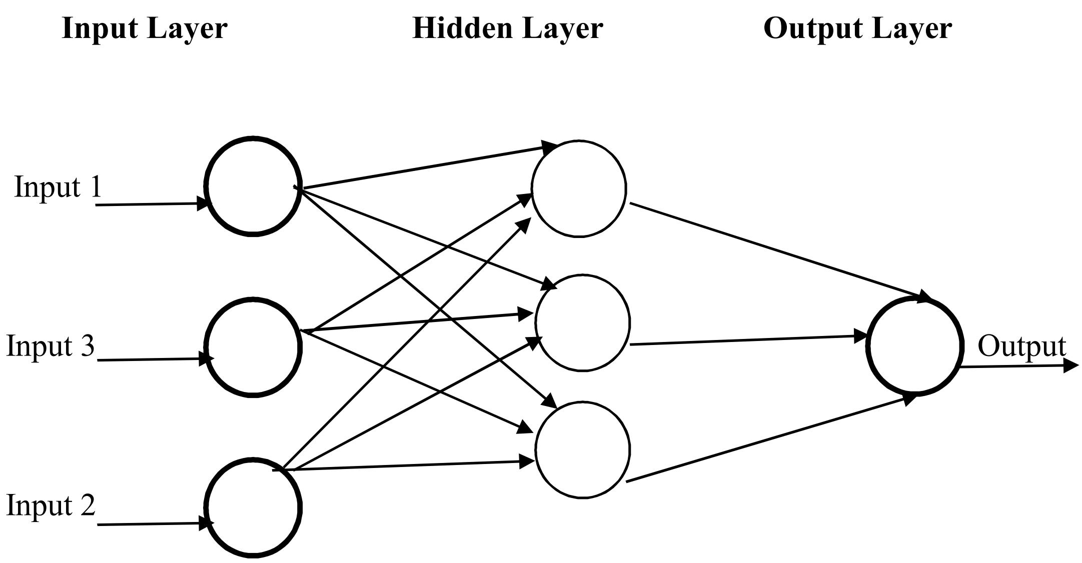


Fig 4.4 Artificial neural network representation with 3 layers

d. Types of ANN Based on Structure

ANNs are divided into two types: Feed Forward ANNs and Back Forward ANNs.

- **Feed Forward Networks**

In this group, neurons are located into layers generally. Inputs are sent from one layer to next one by one-way weights. Due to the one-way connection, a return back of the next layer's output to previous layer as input is impossible. Multi-layer Perceptron (MLP) and Learning Vector Quantization (LVQ) networks are examples for feed forward networks.

- **Feed-Back Networks**

In this group, outputs of hidden and output layer are fed to input layer or previous hidden layers. In this way, inputs can be transferred in both forward and reverse directions. This type networks have dynamic memory and an output in a moment reflects inputs at that moment and previous inputs.

They are suitable for prediction applications and quite successful in the estimation of various types of time series. Examples of these networks are SOM (Self Organization Map), Elman and Jordan networks.

e. Selection of Learning and ANNs Structure

The selection of appropriate ANNs structure depends on considered learning algorithm. In the next table, the differences between network categories and the common network structures.

Table 4.2 Network type and their intended use (Anderson, et al., 1992)

Indented Use of Network	Networks	Use of Network
Prediction	▪ Back-propagation ▪ Delta bar delta ▪ Extended delta bar delta ▪ Directed random search ▪ Higher order neural networks ▪ Self-organizing map (SOM) into back-propagation	Using of inputs values for prediction of some output
Classification	▪ Learning vector quantization (LVQ) ▪ Counter-propagation ▪ Probabilistic neural networks	Using of inputs for classification
Data Association	▪ Hopfield ▪ Bolzmann machine ▪ Hamming network ▪ Bidirectional associative memory ▪ Spatio-temporal pattern recognition	Determining of incorrect values and completing of missing values in input data
Data Conceptualization	▪ Adaptive resonance network (ART) ▪ Self-organizing map (SOM)	Analyze of for derivation of grouping relationships
Data Filtering	▪ Recirculation	Smooth of an input signal

f. Identification of Performance Function

Performance functions calculate the cumulative values between the target outputs values and created outputs by the network. According to these calculated values, how the network close to the pattern of training set is observed and using these values changes connection weights. Mean square error (MSE) is commonly used performance function in feed forward networks.

$$\mathbf{MSE}=\frac{1}{n}\sum_{i=1}^{n}[\mathbf{e}\ (\mathbf{t})\ ^{2}] \qquad \mathbf{(4.5)}$$

Mean Error (ME), Root Mean Square (RMSE) and Mean Absolute Error (MAE) are some of the other performance functions can be used, their equations are:

$$\mathbf{ME}=\frac{1}{n}\sum_{i=1}^{n}\mathbf{e}\ (\mathbf{t}) \qquad \mathbf{(4.6)}$$

$$\mathbf{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}[\mathbf{e}(\mathbf{t})^{2}]} \qquad \mathbf{(4.7)}$$

$$\mathbf{MAE}=\frac{1}{n}\sum_{i=1}^{n}|\mathbf{e}(\mathbf{t})| \qquad \mathbf{(4.8)}$$

$$\mathbf{e_t}=\frac{y_t-t_t}{t_t}\times \mathbf{100} \qquad \textbf{Then MAE}=\frac{1}{n}\sum_{i=1}^{n}|e_t| \qquad \mathbf{(4.9)}$$

Where; e (t) is forecast error at period t, n is number of periods, y_t is forecast and t_t is actual result at period t (Agami *et al*., 2009).

4.3.7. Fuzzy Logic (FL)

Fuzzy logic is primarily concerned with the manipulation of sets with non-deterministic boundaries (Zadeh, 1994). In this regard, fuzzy logic can be used for approximate reasoning. Since in forecasting one usually deals with uncertain outcomes, fuzzy models are one class of widely used tools in this area. Fuzzy logic is especially useful in modeling qualitative and imprecise data and systems, such as human reasoning processes, making decisions based upon vague or imprecise data and considering uncertainty at various levels (Zadeh, 1994 and Iyatomi & Hagiwara, 2004).

4.3.8. Adaptive Neuro Fuzzy Inference System (ANFIS)

Along with fuzzy logic and models, the hybridization of fuzzy logic and ANNs, resulting in neuro-fuzzy models, has also been widely used in water demand forecasting. Neuro-fuzzy models possess the pattern recognition ability of ANNs and the reasoning ability of fuzzy logic (Nauck & Kruse 1997).

Fuzzy logic is employed to describe human thinking and reasoning in a mathematical framework. The main problem with fuzzy logic is that there is no systematic procedure to define the membership function parameters. The construction of the fuzzy rule necessitates the definition of premises and consequences as fuzzy sets. On the other hand, an ANNs has the ability to learn from input and output pairs and adapt to it in an interactive manner (Yurdusev et al., 2009). In recent years, the ANFIS method, which integrates ANNs and FL methods, has been developed. ANFIS has the potential benefits of both these methods in a single framework. ANFIS eliminates the basic problem in fuzzy system design, defining the membership function parameters and design of fuzzy if-then rules, by effectively using the learning capability of ANNs for automatic fuzzy rule generation and parameter optimization (Nayak et al., 2004).

In the present work, the ANFIS methodology is proposed to self-organize model structure and to adapt parameters of the fuzzy system water consumption prediction and rules governing water consumption in the region. It has the advantage of allowing the extraction of fuzzy rules from numerical data. The main drawback of the ANFIS prediction model is the time requested for training structure and determining parameters (Chang & Chang, 2006 and Sen, 2001).

a. ANFIS Components

Fuzzy inference system (FIS) is a rule-based system consisting of three components. These are: A rule-base, containing fuzzy if-then rules, a data-base, defining the membership functions (MF) and an inference system that combines the fuzzy rules and produces the system results (Sen, 2001).

b. ANFIS Structure

- The first phase of FL modeling is the determination of membership function (MF) for input and output variables.
- The second phase is the construction of fuzzy rules.
- The last phase is the determination of output characteristics, output MF and model results (Yurdusev et al., 2009). A general structure for FIS is shown in figure below.

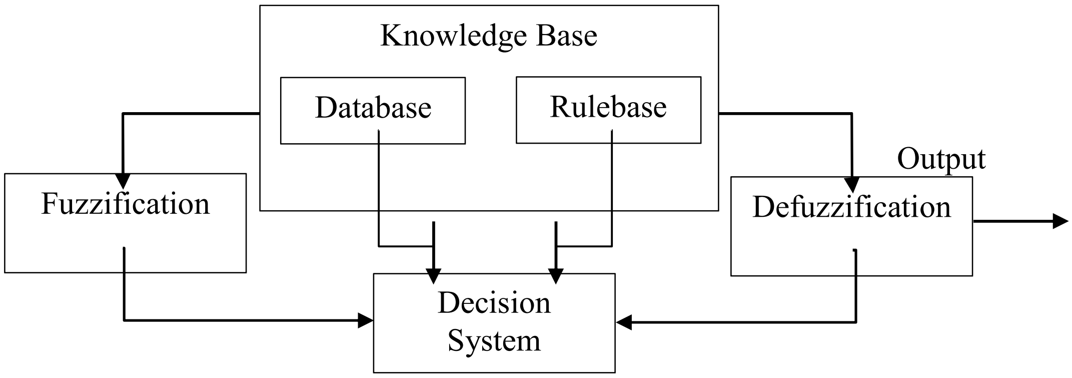


Fig 4.5 The general structure of the fuzzy inference system (Yurdusev et al., 2009)

- **Fuzzification** is a way of mapping numeric input variables into linguistic fuzzy sets. Fuzzification is a mathematical procedure for converting an element in the universe of discourse into the membership value of the fuzzy set. Universe discourse includes a system consisting of a number of input and output variables (Yurdusev et al., 2009).
- **Defuzzification** converts the conclusions of the inference mechanism into numerical values. The aim of defuzzification is always to transform a fuzzy set into a crisp number (Yurdusev et al., 2009).

c. ANFIS types

There are two types of FISs, described in the literature, Sugeno-Takagi FIS and Mamdani FIS. In this study, Sugeno Takagi FIS model is used for predicting water consumption.

The most important difference between these systems is definition of the consequence parameter. The consequence parameter in Sugeno FIS is either a linear equation, called ''first-order Sugeno FIS", or constant coefficient, ''zero-or- der Sugeno FIS (Jang et al., 1997).

d. ANFIS Process

ANFIS uses the learning ability of ANNs to define the input–output relationship and the fuzzy rules are constructed by determining the input structure. The system results are obtained through the reasoning capability of FL (Yurdusev et al., 2009).

The steps of conducting the process are:

- Generate the initial FIS model by choosing one of the following partitioning techniques: Grid partition (Generates a single-output Sugeno-type FIS by using grid partitioning on the data) or Sub. clustering (Generates an initial model for ANFIS training by first applying subtractive clustering on the data).
- There are two methods that ANFIS learning employs for updating membership function parameters: Backpropagation for all parameters (a steepest descent method) A hybrid method consisting of backpropagation for the parameters associated with the input membership functions, and least squares estimation for the parameters associated with the output membership functions.
- Training Error: The training error is the difference between the training data output value, and the output of the fuzzy inference system corresponding to the same training data input value, (the one associated with that training data output value). The training error records the root mean squared error (RMSE) of the training data set at each epoch.
- Checking Error: The checking error is the difference between the checking data output value, and the output of the fuzzy inference system corresponding to the same checking data input value, which is the one associated with that checking data output value. The checking error records the RMSE for the checking data at each epoch.

4.4. Software and Tools

The second part of this chapter gives more details about the tools used in this research. The used statistical software are : ArcGIS, SPSS, STATISTICA and MATLAB Software.

4.4.1. Arc GIS 10.3

ArcGIS is a geographic information system (GIS) for working with maps and geographic information. It is designed for creating and using maps, compiling geographic data, analyzing mapped information, sharing and discovering geographic information, using maps and geographic information in a range of applications and mapping geographic information in a database. The system provides an infrastructure for making maps and geographic information available throughout an organization, across a community. Esri released ArcGIS version 10.3 in December 2014 that used in this research.

In the present paper, ArcGIS is used to realize the thematic analysis, to illustrate the variation of WCP through the years, to demonstrate the variation of the socio-economic variables, the physical characteristics of housing units and the indoor habits of residents in the study area.

4.4.2. Statistical Package for the Social Sciences (SPSS 19)

The software name originally stood for Statistical Package for the Social Sciences (SPSS) reflecting the original market, although the software is now popular in other fields as well. SPSS statistics is software package widely used for statistical analysis. Market researchers, health researchers, survey companies, government, education researchers, and others use it.

The SPSS Statistic version 19 is used in this research. Statistical input has a two-dimensional table structure, where the rows represent cases (such as individuals or households) and the columns represent measurements (such as age, gender, household income). Only two data types are defined: *numeric and text*.

The first procedure or step in SPSS is the creation of the dependent and the independent variables. Output results are varying, according to how many variables exist, from tables, graphs like scatterplots, histograms, pie charts, etc. The SPSS software can conduct the statistical description of variables, regression analysis, ANOVA test, etc.

4.4.3. STATISTICA 8

STATISTICA is statistical software used around the world in more than 20 countries. Stat Soft STATISTICA line of software has gained unprecedented recognition by users and reviews. STATISTICA is a software package that offers much to both students of and professionals. It has a friendly graphical user interface, make it easy to be learnt, in addition to its extensive help facility.

It produces statistical results in high quality output format, seamlessly integrating with several Word-processor and Spreadsheet programs, under the Microsoft Windows environment. It has limitations, through, but these are more likely to be felt by advanced users and statisticians who need fine control over the computational process (STATISTICA, 2011).

With the available water consumption, socio-economic, physical and meteorological data during the period of study, the relationship between water consumption and the independent variables can be defined by many statistical technique

s. STATISTICA 8 can provide lots of statistical analysis including: basic, multiple regressions, ANOVA, principal component analysis, cluster analysis, nonparametric and distribution fittings among other statistics. Also, it enables statisticians and researchers to conduct data mining, in order to check their model's accuracy. One of the perfect abilities of this statistical software is the high-quality graphs, where it is able to produce histograms, scatter plots, mean plot, box plot, and availability and line graphs.

4.4.4. MATLAB

All network training and testing processes in this thesis was performed in MATLAB, which provides a solid ANNs toolbox. This toolbox has built-in functions to construct, train, and save the network.

Matlab is a multi-paradigm numerical computing environment and proprietary programming language developed by MathWorks. Matlab allows matrix manipulations, plotting of functions and data, implementation of algorithms, creation of user interfaces, and interfacing with programs written in other languages.

The application of the whole analysis on dataset, besides to the presentation of results and their interpretation is presented with details in following chapters.

Chapter 5:
Results and Discussion

Part I:
Regression Analysis, ANOVAs Test, Principal Component Analysis, Cluster Analysis Factor Analysis

5.1. Introduction

The present section of thesis presents the results of correlation analysis, ANOVA, cluster analysis, factor analysis and principal component analysis between water consumption and different parameters. In addition, graphically presentation of results obtained in section 1. Results of Artificial Neural Networks with Adaptive Neuro Fuzzy Inference System are presented in part 2 of this chapter. The chosen variables are used for assessing the main determinants of domestic water use. As mentioned in the previous chapter, it exists many independent variables where they are grouped into four categories.

5.2. Creation of Input Scenarios

To cover all possible scenarios, four scenarios are created based on the categories of factors mentioned in details in previous section of the thesis that are (table 5.1):

Table 5.1: Scenarios adopted for WCP analysis

Scenarios	**Parameters**	**Data Sets**	**Variation**
Scenario One	**Socio-Economic Parameters**	Household size, number of female and male, monthly income, the four categories of household age, the four categories of education level and car habits.	vary according to each house
Scenario Two	**Physical Characteristics of Housing Units**	Total area of the house, building area of the house, garden area, number of rooms and frequency of garden watering.	
Scenario Three	**Indoor Habits**	Frequency of washing dish per day, washing clothes (laundry) per week, using toilet per day and shower for male and female per week.	
Scenario Four	**Climatic Factors**	mean precipitation and mean temperature	constant for all houses

The three first scenarios are human related wile the climatic factors are site dependants and does not vary with houses, and by consequence only the first 3 categories are used for the statistical analysis and numerical techniques in this thesis.

5.3.Correlation Analysis and Matrix

Correlation analysis is conducted to see if there is any predictive relationship between water consumption and the three scenarios and the inter-relationship between all parameters.

5.3.1. Scenario1

To evaluate the correlation, the G.DE Landsheere, 1979 (Belhassen et al., 2016) scale is used (table 5.2).

The relationship between socio-economic parameters and per capita total water usage is demonstrated in table 5.3, table 5.4 and figures 5.1.

Water consumption is *very strongly correlated* with monthly income. Also, it's *strongly correlated* with number of females, household size, two categories of education level HGS and UNIV. it has *medium correlation* with the two age categories AG1 and AG3, primary school and car numbers. Furthermore, WCP has *very week correlation* between number of males, the two age categories AG2 and AG4, medium school in education level and the frequency of washing cars.

In addition, results of table demonstrate an intercorrelation between variables. For example, a strong relationship between number of females and monthly income, between university education level and household size, etc.

Table 5.2: Classification of correlation strength (according to G.DE LANDSHEERE ,1979)

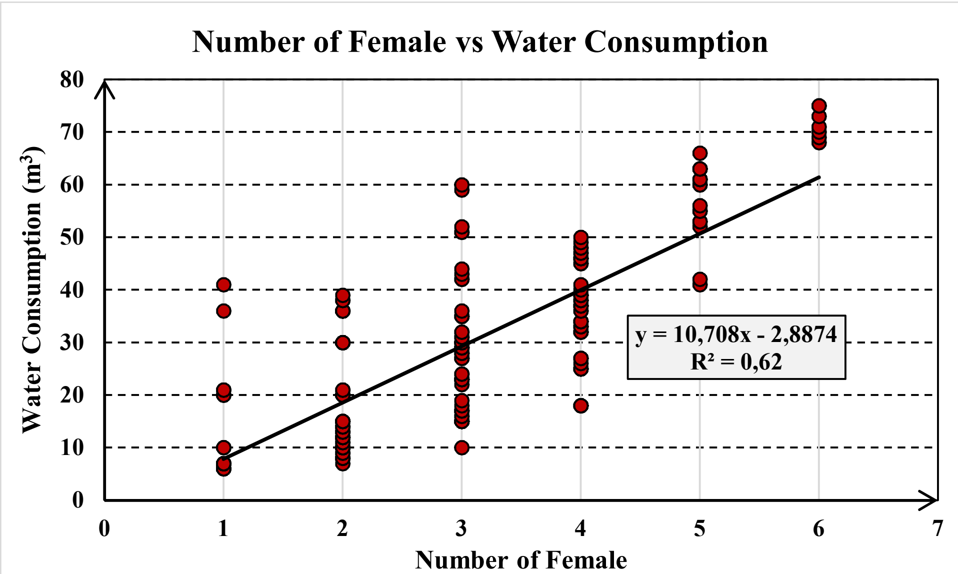
R > 0,8	Very strong correlation
If 0,6 <R< 0,8	Strong correlation
If 0,4 < R < 0,6	Medium correlation
If 0,25 < R < 0,4	Week correlation
If R < 0,25	Very week correlation

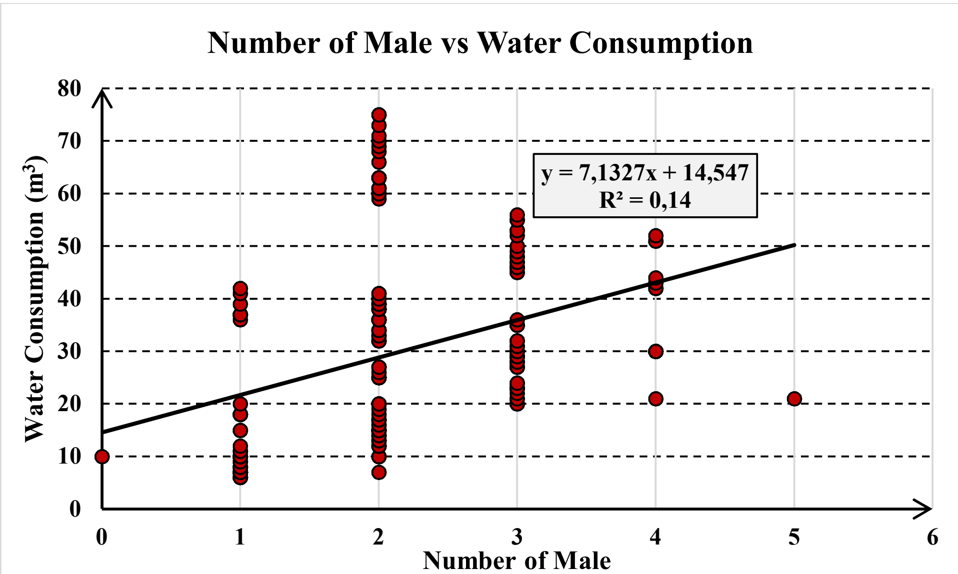
Table 5.3: Correlation matrix between socio-economic parameters and WCP

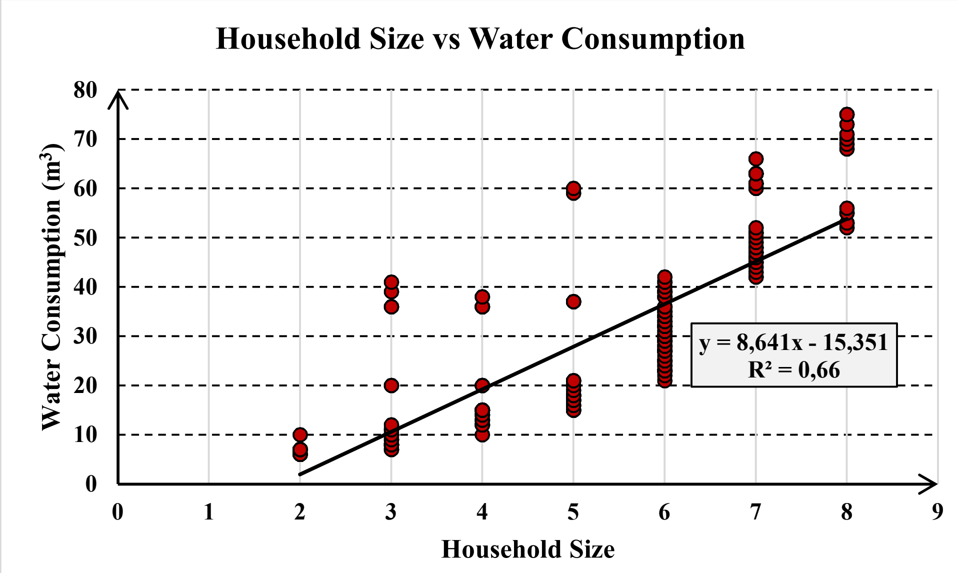
Marked correlations are significant at p < 0,05															
Variable	**WCP**	**FEM**	**MAL**	**HOUS**	**AG1**	**AG2**	**AG3**	**AG4**	**PRS**	**MDS**	**HGS**	**UNIV**	**INC**	**CARN**	**WCAR**
WCP	1,00														
FEM	**0,62**	1,00													
MAL	0,14	0,11	1,00												
HOUS	**0,66**	**0,84**	**0,63**	1,00											
AG1	**0,50**	**0,60**	0,30	**0,63**	1,00										
AG2	0,03	-0,02	0,32	0,16	-0,30	1,00									
AG3	**0,57**	**0,64**	0,32	**0,67**	0,33	-0,20	1,00								
AG4	0,01	0,26	0,31	0,38	0,11	0,07	-0,21	1,00							
PRS	**0,53**	0,48	0,35	**0,57**	**0,60**	-0,15	**0,54**	-0,07	1,00						
MDS	0,01	0,12	0,27	0,25	-0,22	**0,84**	0,03	-0,03	-0,13	1,00					
HGS	**0,61**	**0,61**	0,16	**0,56**	0,31	-0,20	**0,83**	-0,17	**0,58**	-0,18	1,00				
UNIV	**0,63**	**0,76**	**0,52**	**0,87**	0,44	0,02	**0,75**	0,30	**0,50**	0,15	**0,63**	1,00			
INC	**0,80**	**0,68**	0,24	**0,66**	0,46	-0,18	**0,77**	-0,05	**0,61**	-0,14	**0,88**	**0,72**	1,00		
CARN	**0,50**	**0,50**	0,32	**0,57**	**0,55**	-0,16	**0,57**	-0,06	**0,53**	-0,12	**0,53**	**0,53**	**0,64**	1,00	
WCAR	0,01	0,07	0,08	0,10	0,02	-0,06	0,12	0,07	-0,03	-0,07	0,10	0,13	0,12	0,12	1,00

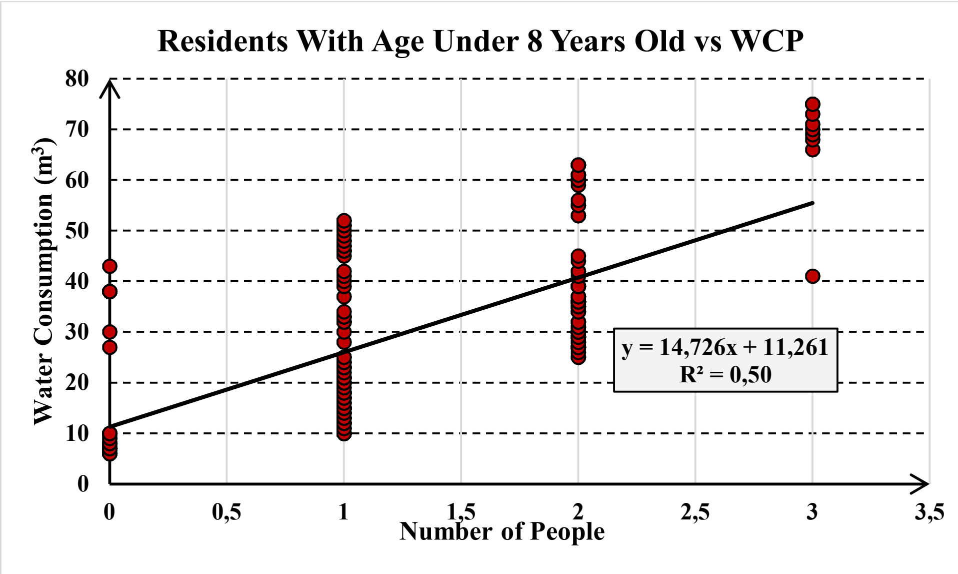
Table 5.4: Classification of socioeconomic parameters according to correlation strength

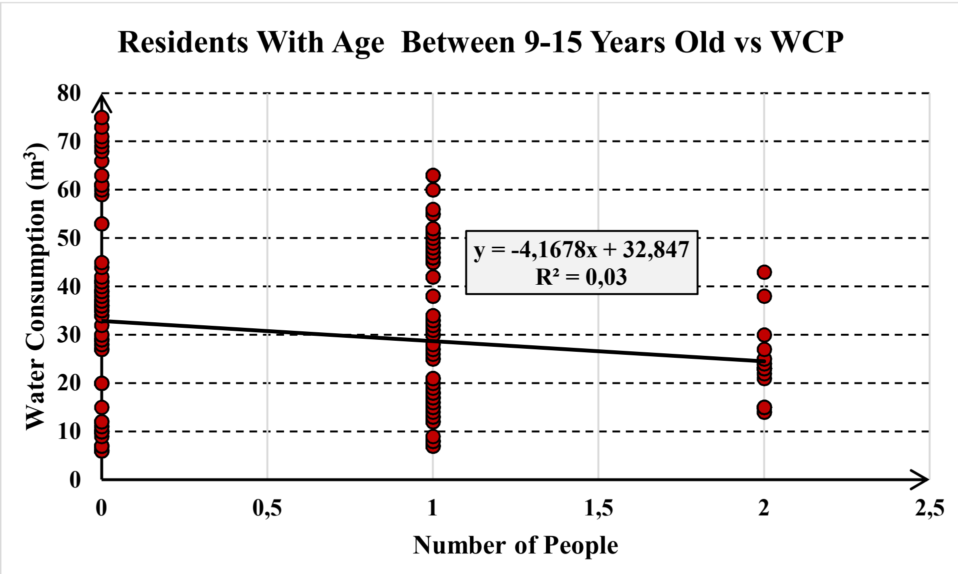
Variables	INC	HOUS	UNIV	FEM	HGS	AG3	PRS	CARN	AG1	MAL	AG2	WCAR	AG4	MDS
Correlation with WCP	0,80	0,66	0,63	0,62	0,61	0,57	0,53	0,50	0,50	0,14	0,03	0,01	0,01	0,01
Correlation strength	Very strong	Strong				Medium				Very week				

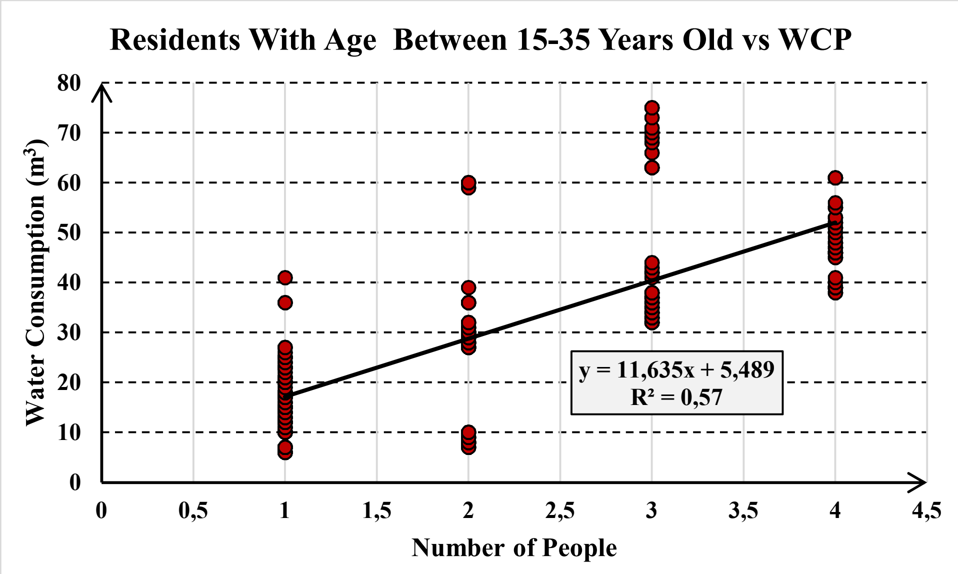


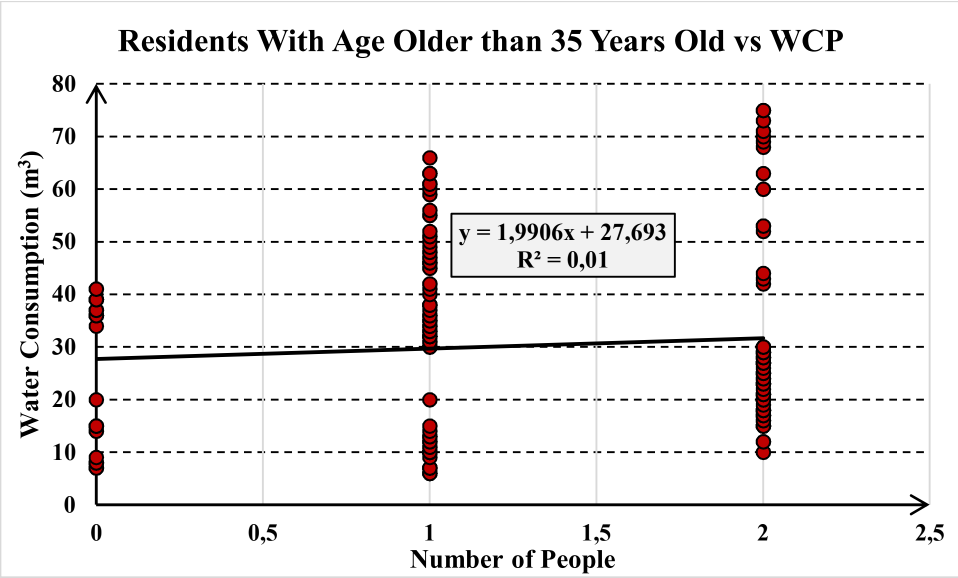


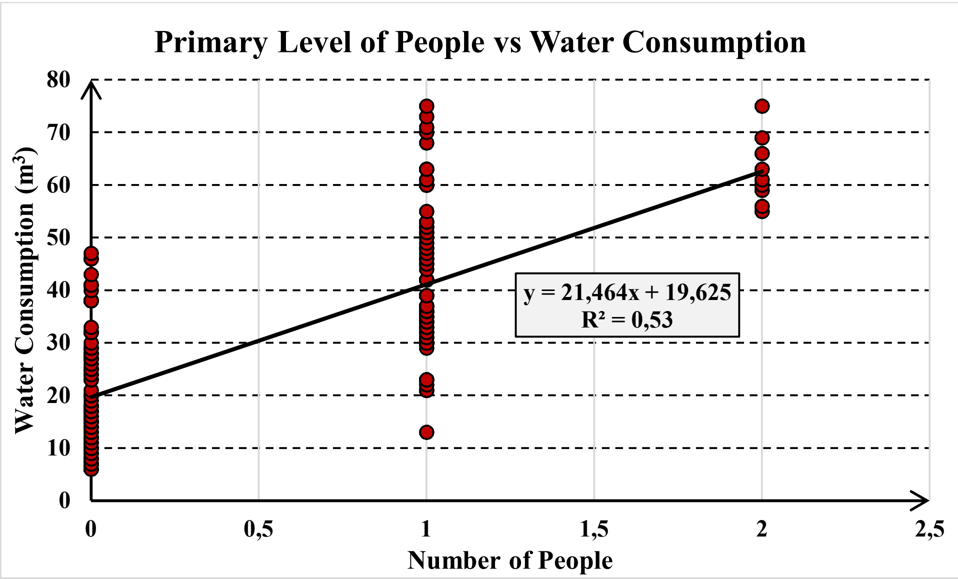

Household Size vs Water Consumption
Water Consumption (m3)
80
70
60
50
40
30
20
10
0
0
1
2
3
4
5
6
7
8
9
Household Size
y = 8,641x - 15,351
R² = 0,66

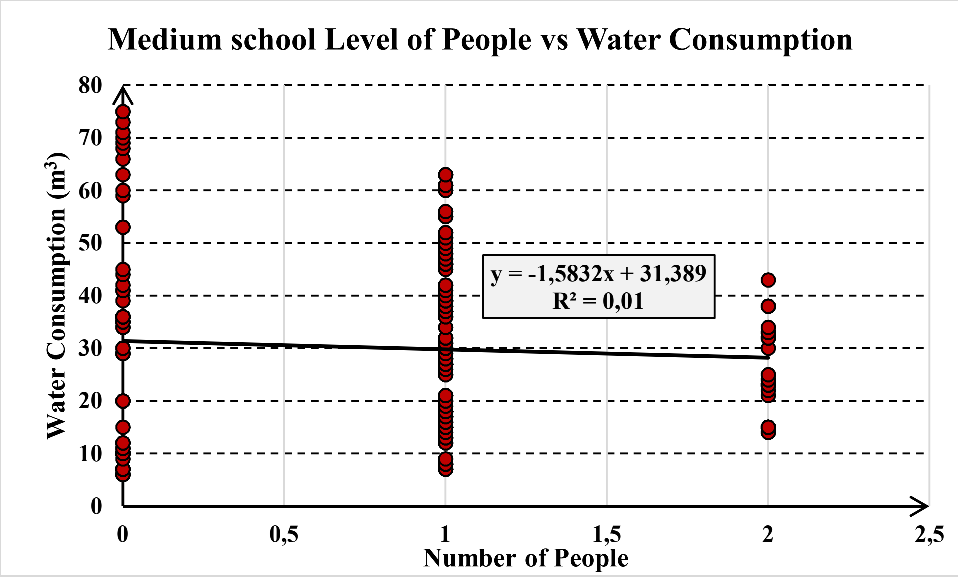

Residents With Age Under 8 Years Old vs WCP
Water Consumption (m3)
80
70
60
50
40
30
20
10
0
0
0,5
1
1,5
2
2,5
3
3,5
Number of People
y = 14,726x + 11,261
R² = 0,50

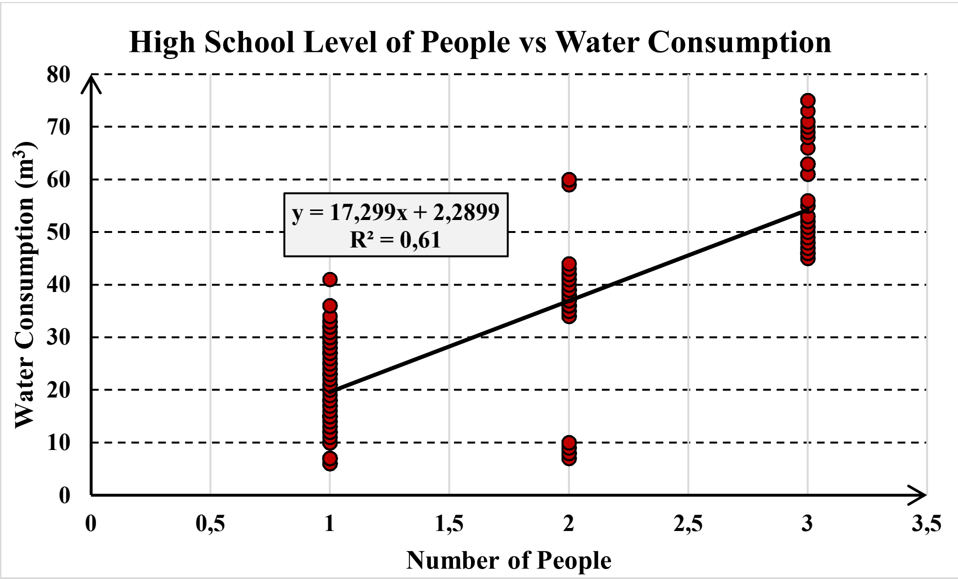

Residents With Age Between 9-15 Years Old vs WCP
Water Consumption (m3)
80
70
60
50
40
30
20
10
0
0
0,5
1
1,5
2
2,5
Number of People
y = -4,1678x + 32,847
R² = 0,03

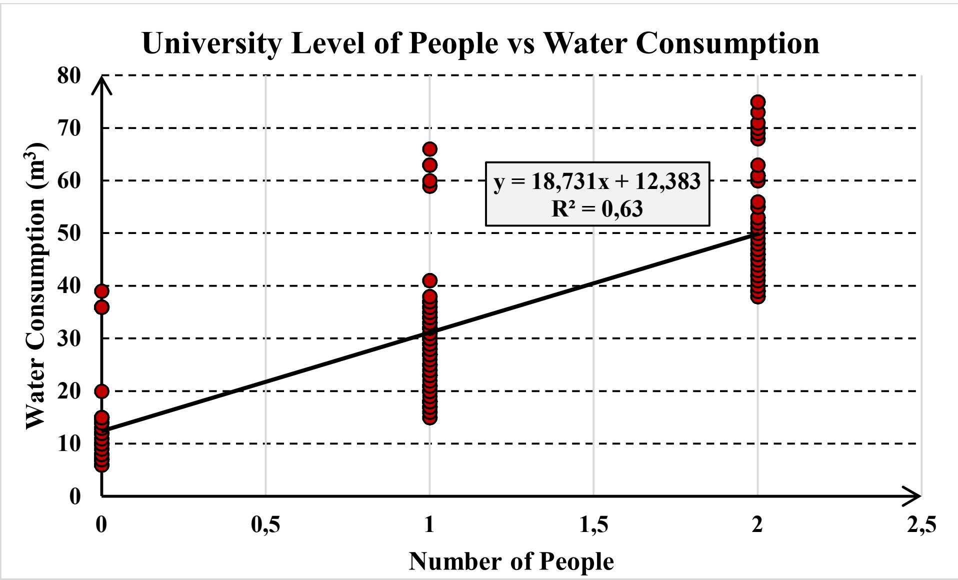

Residents With Age Between 15-35 Years Old vs WCP
Water Consumption (m³)
80
70
60
50
40
30
20
10
0
0
0,5
1
1,5
2
2,5
3
3,5
4
4,5
Number of People
y = 11,635x + 5,489
R² = 0,57

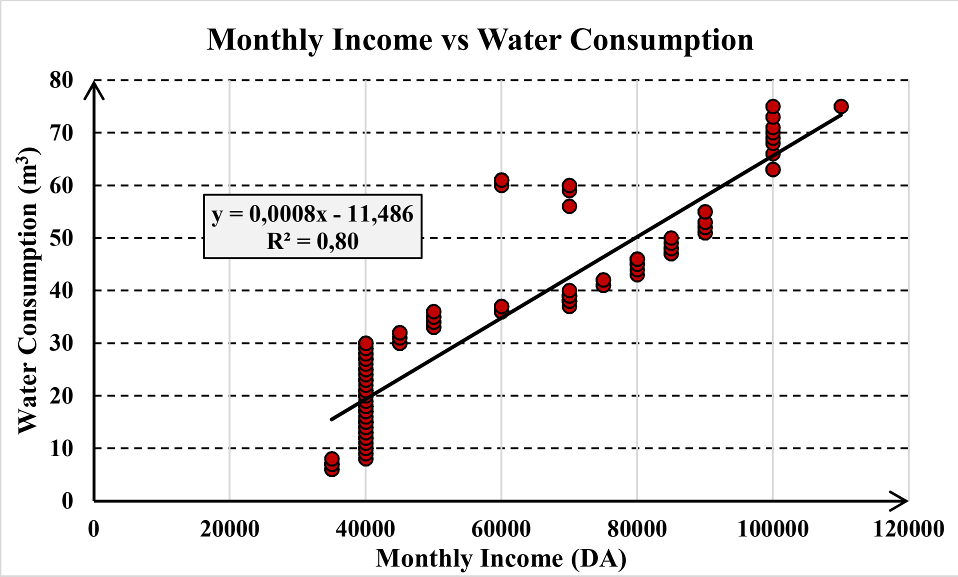

Residents With Age Older than 35 Years Old vs WCP
Water Consumption (m³)
80
70
60
50
40
30
20
10
0
0
0,5
1
1,5
2
2,5
Number of People
y = 1,9906x + 27,693
R² = 0,01

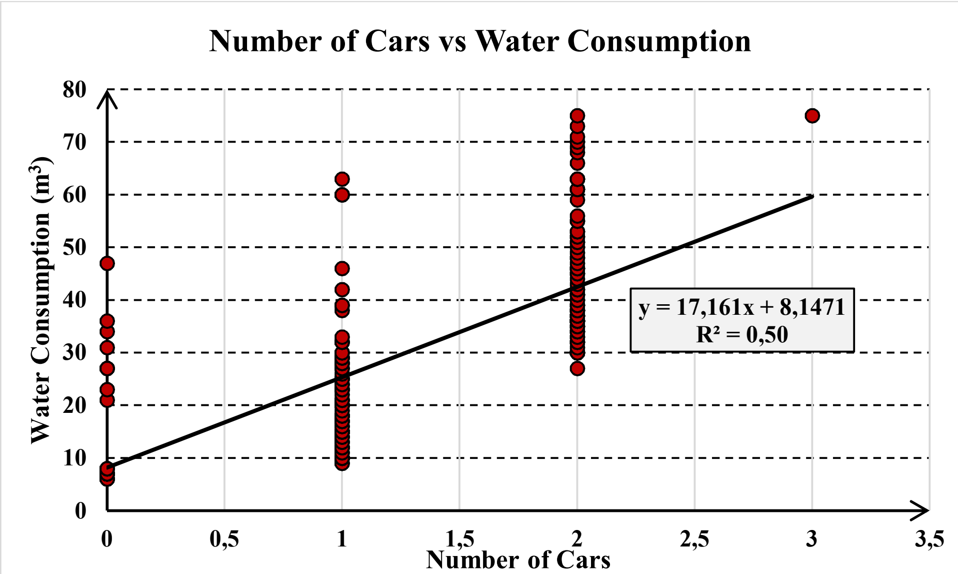

Primary Level of People vs Water Consumption
Water Consumption (m³)
80
70
60
50
40
30
20
10
0
0
0,5
1
1,5
2
2,5
Number of People
y = 21,464x + 19,625
R² = 0,53


Medium school Level of People vs Water Consumption
y = -1,5832x + 31,389
R² = 0,01
Water Consumption (m³)
Number of People
0
10
20
30
40
50
60
70
80
0
0,5
1
1,5
2
2,5


High School Level of People vs Water Consumption
y = 17,299x + 2,2899
R² = 0,61
Water Consumption (m³)
Number of People
0
10
20
30
40
50
60
70
80
0
0,5
1
1,5
2
2,5
3
3,5


University Level of People vs Water Consumption
y = 18,731x + 12,383
R² = 0,63
Water Consumption (m³)
Number of People
0
10
20
30
40
50
60
70
80
0
0,5
1
1,5
2
2,5





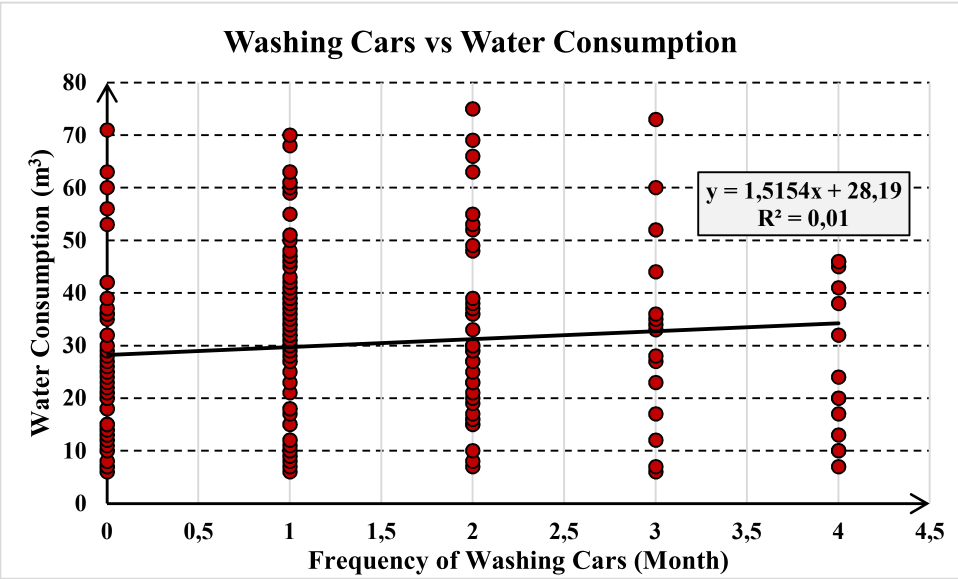


Fig 5.1 Scatterplots between water consumption and socio-economic parameters

5.3.2. Scenario2

The relationship between physical characteristics of housing units and indoor water consumption is illustrated in figures 5.2, correlation matrix in table 5.5 and table 5.6. Results reveals that:

Water consumption is *very strongly correlated* with total area of the house, building area and number of rooms. Also, it's very weekly correlated with garden possession.

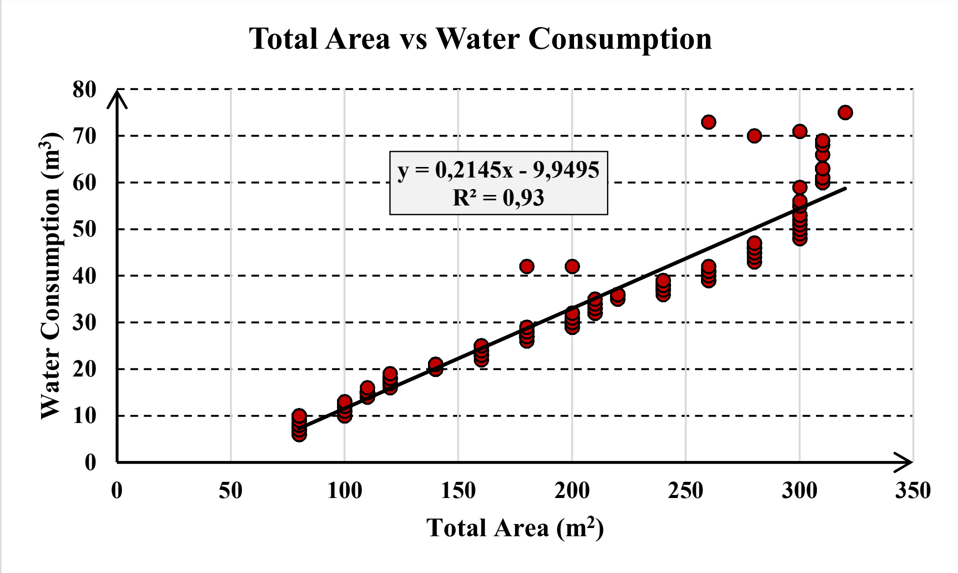
Also, an inter-relationship between variables exists where total area and building area are very strongly related to number of rooms.

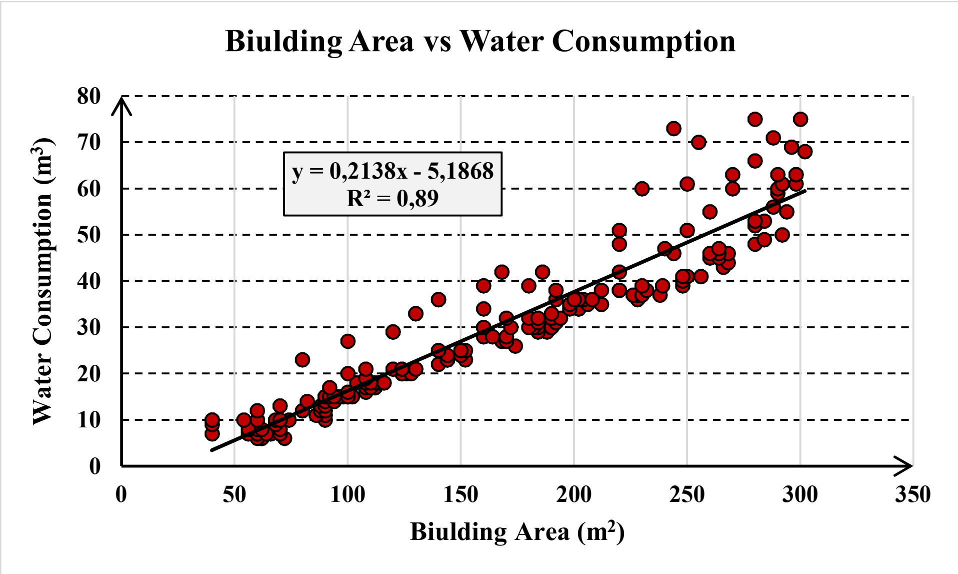
Table 5.5: Correlation matrix between physical characteristics of building units and WCP

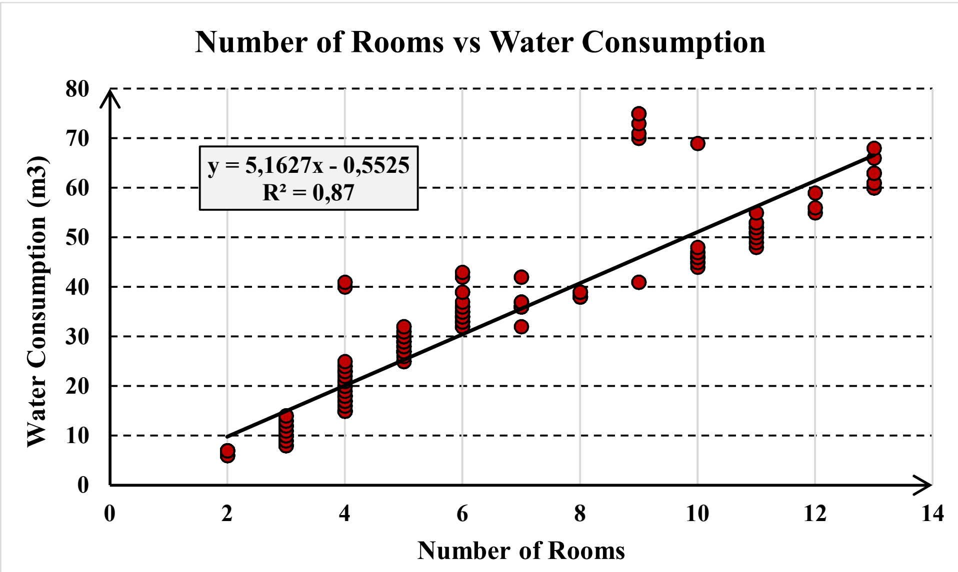
Marked correlations are significant at $p < 0,05$						
Variable	**WCP**	**TAR**	**BAR**	**ROMN**	**GAR**	**GWAT**
WCP	1,00					
TAR	**0,93**	1,00				
BAR	**0,89**	**0,97**	1,00			
ROMN	**0,87**	**0,93**	**0,90**	1,00		
GAR	0,03	0,20	-0,04	0,18	1,00	
GWAT	2E-5	0,02	-0,06	0,02	0,35	1,00

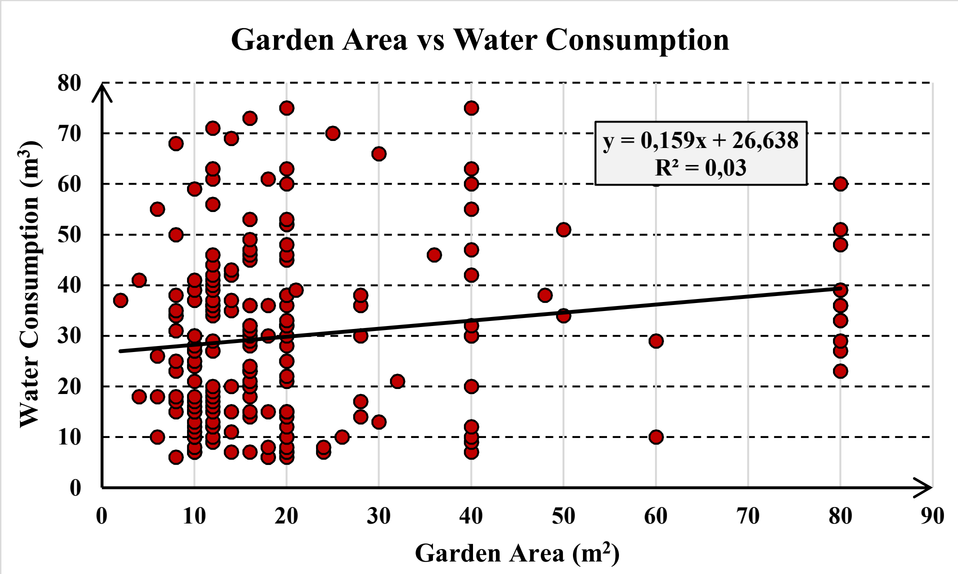
Table 5.6: Classification of physical characteristics of housing units according to correlation strength

Variables	TAR	BAR	ROMN	GAR	GWAT
Correlation with WCP	0,93	0,89	0,87	0,03	2E-5
Correlation strength	Very strong			Very week	


Total Area vs Water Consumption
y = 0,2145x - 9,9495
R² = 0,93
Water Consumption (m³)
Total Area (m²)
0
10
20
30
40
50
60
70
80
0
50
100
150
200
250
300
350


Biulding Area vs Water Consumption
y = 0,2138x - 5,1868
R² = 0,89
Water Consumption (m³)
Biulding Area (m²)
0
10
20
30
40
50
60
70
80
0
50
100
150
200
250
300
350


Number of Rooms vs Water Consumption
y = 5,1627x - 0,5525
R² = 0,87
Water Consumption (m3)
Number of Rooms
0
10
20
30
40
50
60
70
80
0
2
4
6
8
10
12
14



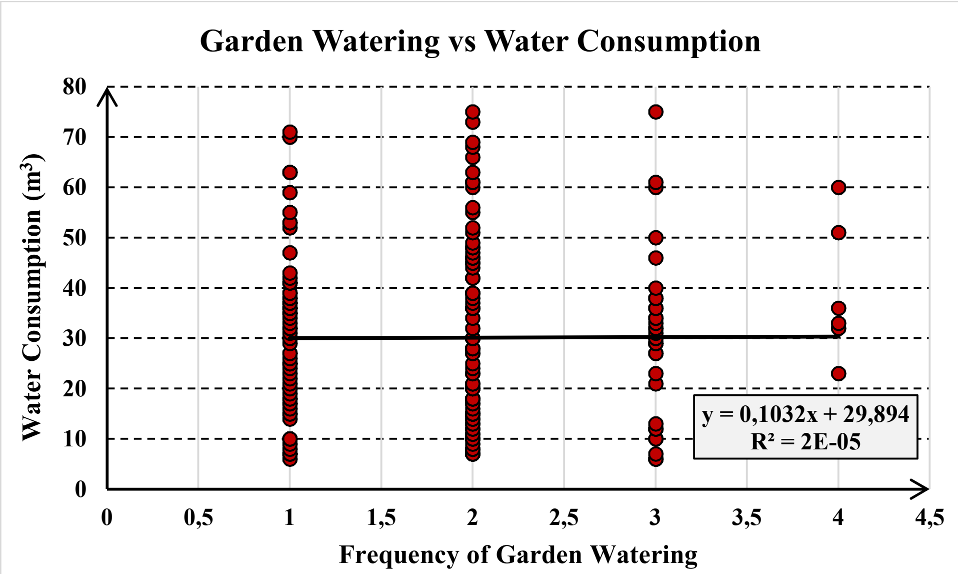


Fig 5.2 Scatterplots between water consumption and physical characteristics of building units

5.3.3. Scenario3

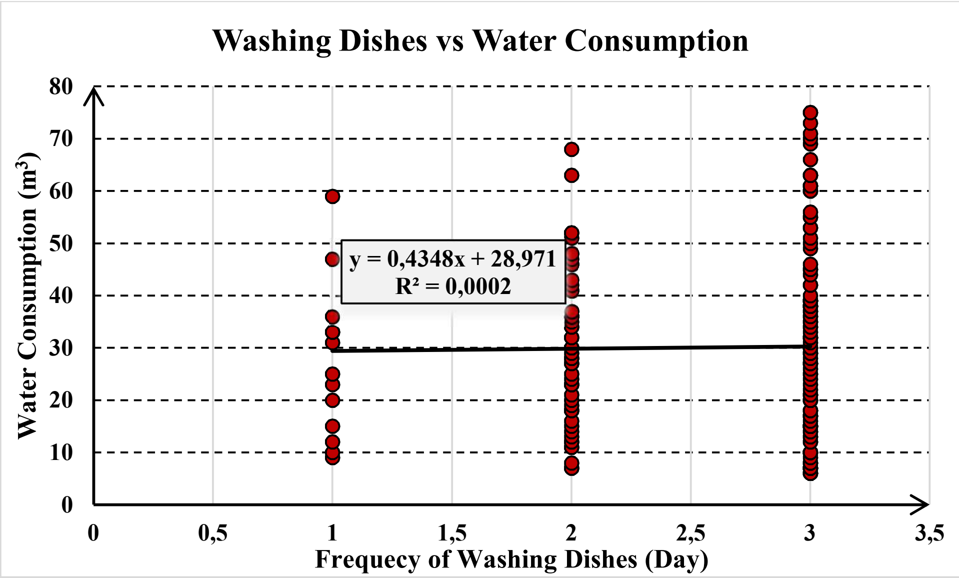
Data in table 5.7, table 5.8 and figures 5.3 shows that the indoor habits of people have very week correlation with WCP. Also, the variables are very weekly correlated with each other because the indoor habits haven't the same scale duration. For example, the duration of washing dish is per day and washing clothes is per week.

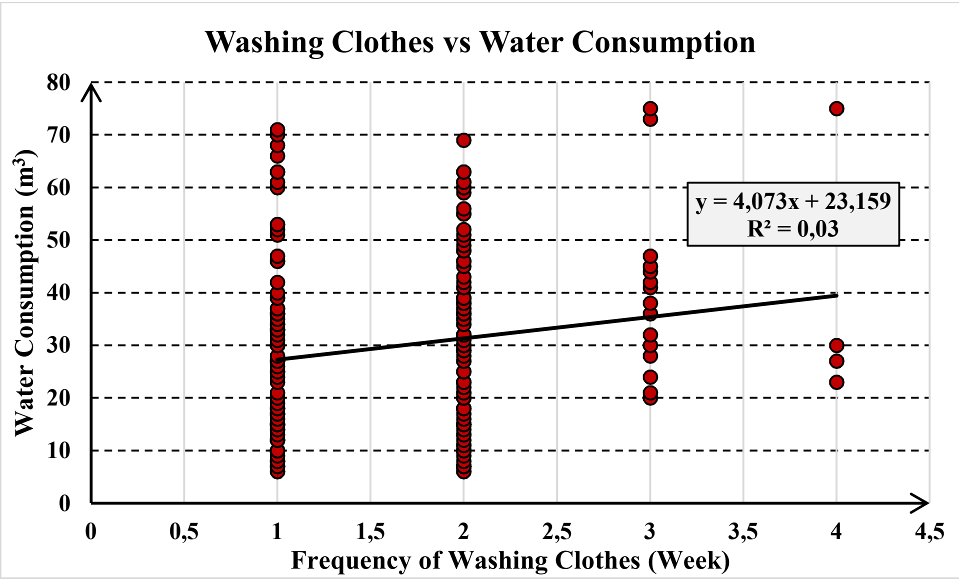
Table 5.7: Correlation matrix between indoor habits of residents and WCP

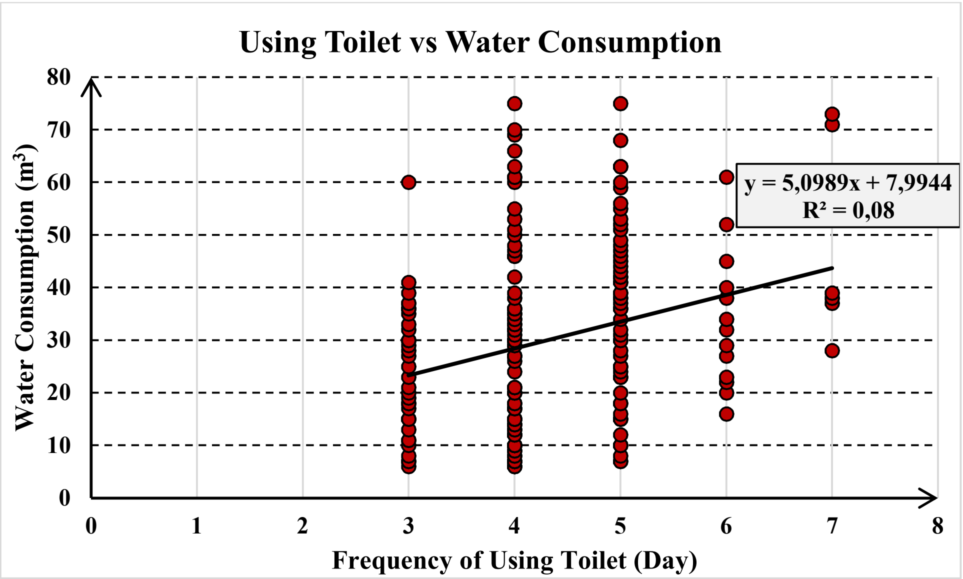
Marked correlations are significant at $p < 0{,}05$						
Variable	**WCP**	**WDISH**	**WCL**	**UTLT**	**FSHW**	**MSHW**
WCP	1,00					
WDISH	0,0002	1,00				
WCL	0,03	0,10	1,00			
UTLT	0,08	0,06	0,22	1,00		
FSHW	0,002	-0,18	0,08	-0,16	1,00	
MSHW	0,03	-0,30	0,15	0,08	0,42	1,00

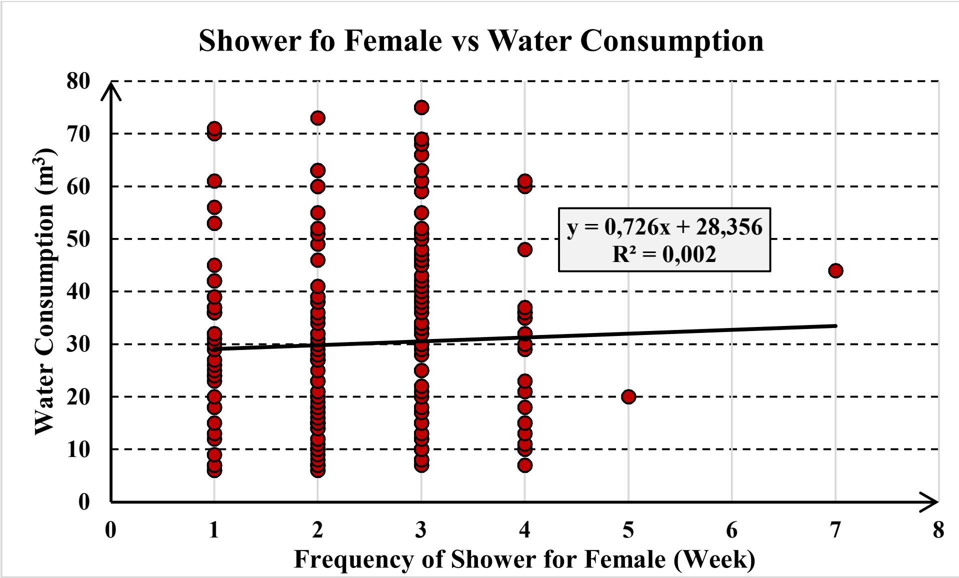
Table 5.8: Classification of indoor habits of residents according to correlation strength

Variables	WDISH	WCL	UTLT	FSHW	MSHW
Correlation with WCP	0,0002	0,03	0,08	0,002	0,03
Correlation strength	Very week				


Washing Dishes vs Water Consumption
Water Consumption (m³)
0
10
20
30
40
50
60
70
80
y = 0,4348x + 28,971
R² = 0,0002
0
0,5
1
1,5
2
2,5
3
3,5
Frequecy of Washing Dishes (Day)


Washing Clothes vs Water Consumption
Water Consumption (m³)
0
10
20
30
40
50
60
70
80
y = 4,073x + 23,159
R² = 0,03
0
0,5
1
1,5
2
2,5
3
3,5
4
4,5
Frequency of Washing Clothes (Week)


Using Toilet vs Water Consumption
Water Consumption (m³)
0
10
20
30
40
50
60
70
80
y = 5,0989x + 7,9944
R² = 0,08
0
1
2
3
4
5
6
7
8
Frequency of Using Toilet (Day)



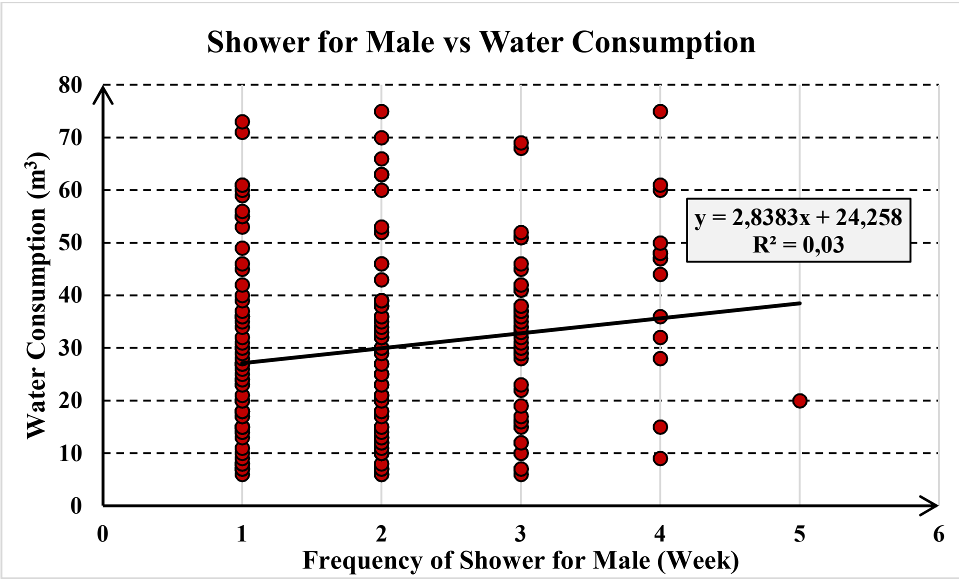


Fig 5.3 Scatterplots between water consumption and indoor habits of residents

5.4.Analysis of Variance (ANOVA)

The purpose of using ANOVA in this research work, is to answer on the following research questions:

- Is there a difference in household water use for houses having 1, 2 or more family members?
- Is there any difference in WCP for houses that having 1, 2 and 3 females and males members?
- Is there any difference in WCP for houses that haven't cars or having 1, 2 or more cars in their building?
- Is there any difference in WCP for houses that have different building or garden size?
- For indoor habits, is there any difference in WCP between houses that have different frequency for washing clothes or dishwashing, etc.?

ANOVA will indicate whether there are significant differences in the mean of WCP across the categories in each factor. In this section to perform the test, SPSS Statistics was used. Below is the output of the ANOVA to compare the means of the dependent variable, WCP, and the other parameters.

5.4.1. Homogeneity test

Results of homogeneity test of variance are presented in Table 5.9 for the three scenarios.

Variance in scores is the same for all groups in every variable. For example, the variance score is equal to 0,036 for each of 6 groups of number of females (FEM). Description of the groups of each variable is presented in Table 6.6 in the ANNEX. Generally, if the values of Sig are less than 0, 05 consultations of Robust Tests of equality of means are obligatory.

Table 5.9: Test of homogeneity of variance

Test of Homogeneity of Variances WCP (m³)				
Variables	**Levene Statistic**	**df1**	**df2**	**Sig.**
HOUS	7,734	6	194	**0,000**
FEM	2,438	5	195	**0,036**
MAL	7,249[a]	3	195	**0,000**
AG1	1,417	3	197	0,239
AG2	9,367	2	198	**0,000**
AG3	14,196	3	197	**0,000**
AG4	0,635	2	198	0,531
PRS	4,108	2	198	**0,018**
MDS	14,553	2	198	**0,000**
HGS	5,462	2	198	**0,005**
UNIV	2,211	2	198	0,112
INC	18,212[a]	10	189	**0,000**
CARN	1,568[a]	2	197	0,211
WCAR	0,647	4	196	0,629
TAR	3,723	15	185	**0,000**
BAR				
ROMN	10,971	11	189	**0,000**
GAR	1,855[a]	15	178	**0,031**
GWAT	1,257	3	197	0,290
WDISH	3,932	2	198	**0,021**
WCL	0,543	3	197	0,653
UTLT	2,073	4	196	**0,086**
FSHW	0,324[a]	3	195	0,808
MSHW	0,867[a]	3	196	0,459
a. Groups with only one case are ignored in computing the test of homogeneity of variance for WCP (m³).				

Table 5.10: Description for groups of variables

(ANNEX 04)

5.4.2. ANOVA

The purpose of using ANOVA in this research work is to test and determine the possible influence that independent variables may have on the dependent output. In other terms, does the considered parameter Xi has an impact on WCP and whether it is a systematic factor ? ANOVA measures the differences in the mean of WCP across the categories in each factor. In this section to perform the test, SPSS Statistics is used. Below is the output of the ANOVA to compare the means of the dependent variable, WCP, and the other parameters.

Table 5.11: ANOVA results

ANOVA WCP (m³)						
Variables		**Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
HOUS	**Between Groups**	44995,179	6	7499,197	103,216	0,000
	Within Groups	14095,209	194	72,656		
	Total	59090,388	200			
FEM	**Between Groups**	39523,081	5	7904,616	78,774	0,000
	Within Groups	19567,307	195	100,345		
	Total	59090,388	200			
MAL	**Between Groups**	13067,689	5	2613,538	11,074	0,000
	Within Groups	46022,699	195	236,014		
	Total	59090,388	200			
AG1	**Between Groups**	28199,077	3	9399,692	59,944	0,000
	Within Groups	30891,311	197	156,809		
	Total	59090,388	200			
AG2	**Between Groups**	1727,150	2	863,575	2,981	**0,053**
	Within Groups	57363,238	198	289,713		
	Total	59090,388	200			
AG3	**Between Groups**	35564,770	3	11854,923	99,271	0,000
	Within Groups	23525,618	197	119,419		
	Total	59090,388	200			
AG4	**Between Groups**	1471,853	2	735,927	2,529	**0,082**
	Within Groups	57618,535	198	291,003		
	Total	59090,388	200			
PRS	**Between Groups**	31427,709	2	15713,855	112,474	0,000
	Within Groups	27662,679	198	139,710		
	Total	59090,388	200			
MDS	**Between Groups**	669,368	2	334,684	1,134	**0,324**
	Within Groups	58421,020	198	295,056		
	Total	59090,388	200			
HGS	**Between Groups**	36888,247	2	18444,123	164,486	0,000
	Within Groups	22202,141	198	112,132		
	Total	59090,388	200			
UNIV	**Between Groups**	37837,748	2	18918,874	176,257	0,000

	Within Groups	21252,640	198	107,337		
	Total	59090,388	200			
INC	**Between Groups**	51879,403	11	4716,309	123,615	0,000
	Within Groups	7210,985	189	38,153		
	Total	59090,388	200			
CARN	**Between Groups**	30830,952	3	10276,984	71,642	0,000
	Within Groups	28259,436	197	143,449		
	Total	59090,388	200			
WCAR	**Between Groups**	5322,191	4	1330,548	4,850	**0,001**
	Within Groups	53768,197	196	274,328		
	Total	59090,388	200			
TAR	**Between Groups**	56535,533	15	3769,036	272,920	0,000
	Within Groups	2554,855	185	13,810		
	Total	59090,388	200			
BAR	**Between Groups**	56298,875	85	662,340	27,286	0,000
	Within Groups	2791,513	115	24,274		
	Total	59090,388	200			
NROM	**Between Groups**	55211,961	11	5019,269	244,594	0,000
	Within Groups	3878,427	189	20,521		
	Total	59090,388	200			
GAR	**Between Groups**	8186,769	22	372,126	1,301	**0,176**
	Within Groups	50903,619	178	285,975		
	Total	59090,388	200			
GWAT	**Between Groups**	645,138	3	215,046	0,725	**0,538**
	Within Groups	58445,250	197	296,676		
	Total	59090,388	200			
WDISH	**Between Groups**	143,994	2	71,997	0,242	**0,785**
	Within Groups	58946,394	198	297,709		
	Total	59090,388	200			
WCL	**Between Groups**	2074,392	3	691,464	2,389	**0,070**
	Within Groups	57015,996	197	289,421		
	Total	59090,388	200			
UTLT	**Between Groups**	5305,642	4	1326,411	4,834	**0,001**
	Within Groups	53784,746	196	274,412		
	Total	59090,388	200			
FSHW	**Between Groups**	2913,088	5	582,618	2,022	**0,077**
	Within Groups	56177,300	195	288,089		
	Total	59090,388	200			
MSHW	**Between Groups**	2827,377	4	706,844	2,462	**0,047**
	Within Groups	56263,011	196	287,056		
	Total	59090,388	200			

5.4.3. Post hoc tests (Multiple comparisons)

The results from the one-way ANOVA do not indicate which of the groups differ from one another. As a result, on many cases it is of interest to follow the analysis with a post hoc test or a planned comparison among particular means. If several comparisons between pairs of means are made, it is a good idea to use a test, such as the Tukey, that controls for alpha inflation.

Results from post hoc tests, table 5.12, reveal a significant difference between HOUS groups, AG1 groups, AG2 groups, AG3 groups, AG4 groups, FEM groups, PRS groups, MDS groups, HGS groups, UNIV groups, WCAR groups, TAR groups, ROMN groups, GWAT groups, UTLT groups WCL groups and WDISH groups.

Table 5.12: Post hoc tests_ multiple comparisons

(ANNEX 05)

5.4.4. Effect size

The effect size is calculated from the information provided in the ANOVA table (Table 5.13). It calculated by the following formula:

Eta squared= **Sum of squares between** groups divided by **Total sum of squares**.

According to Cohen (1988):

- If Eta squared= 0.01 it is **small effect**
- If Eta squared= 0.06 it is **medium effect**
- If Eta squared= 0.14 it is **large effect**

Results from table 5.13 of effect size shows that all of AG2, AG4, MDS, GWAT, WDISH, WCL, FSHW and MSHW have small effect. The following variables WCAR and UTLT have medium effect. The rest of parameters HOUS, FEM, MAL, AG1, AG3, PRS, HGS, UNIV, INC, CARN, TAR, BAR, NROM and GAR have large effect.

A one-way between groups analysis of variance was conducted to explore the impact of socio-economic parameters like household size, family income, education level besides to physical characteristics of building units like building area of the house and some indoor habits of residents such as washing clothes and showers, etc., on domestic water consumption. Every variable was divided into groups (3groups, 4groups, etc.) where each variable has its own groups (for example: for male there are 5 groups according to number of males in every house).

There was a statistically significant difference at the *p-value* for some groups of variables (for example: it was equal to 0.053, 0.324 and 0.785 for AG2, MDS and WDISH, respectively).

The effect size, calculated using Eta squared, indicated three type of effect according to Cohen where some has small effect like AG4 and GWAT, medium effect like WCAR and large effect like INC and HOUS. Post-hoc comparisons using the Tukey HSD test indicated that the mean score for each group was significantly different from the other groups in the same variable.

Table 5.13: Effect size

Variables	Sum of Squares		Eta squared
	Between Groups	Total	
HOUS	44995,179	59090,388	**0,76**
FEM	39523,081	59090,388	**0,67**
MAL	13067,689	59090,388	**0,22**
AG1	28199,077	59090,388	**0,48**
AG2	1727,150	59090,388	**0,03**
AG3	35564,770	59090,388	**0,60**
AG4	1471,853	59090,388	**0,02**
PRS	31427,709	59090,388	**0,53**
MDS	669,368	59090,388	**0,01**
HGS	36888,247	59090,388	**0,62**
UNIV	37837,748	59090,388	**0,64**
INC	51879,403	59090,388	**0,88**
CARN	30830,952	59090,388	**0,52**
WCAR	5322,191	59090,388	**0,09**
TAR	56535,533	59090,388	**0,96**
BAR	56298,875	59090,388	**0,95**
NROM	55211,961	59090,388	**0,93**
GAR	8186,769	59090,388	**0,14**
GWAT	645,138	59090,388	**0,01**
WDISH	143,994	59090,388	**0,002**
WCL	2074,392	59090,388	**0,04**
UTLT	5305,642	59090,388	**0,09**
FSHW	2913,088	59090,388	**0,05**
MSHW	2827,377	59090,388	**0,05**

5.5.Cluster Analysis

The cluster analysis is mainly used to organize observations and variables in the same category of data set, to more meaningful groups so each group is more-or-less homogeneous and distinct from other clusters. It was carried out for the three scenarios. The linkage of tree clustering was selected so that Euclidean distance between two clusters is determined by the distance of the furthest cases of these two clusters.

5.5.1. Scenario 1

These conclusions concerning the number of clusters and their membership were reached through a visual inspection in figure 5.4. It suggests that the socio-economic parameters form two clusters. The first consists of monthly income. The second contains water consumption, household size, number of females and males, the four categories of age, the four categories of education level and number of cars.

It can be noted that the two clusters are quite distinct from each other. From this inspection, the first cluster can be identified as **"Income"** cluster whilst the second cluster is labeled as "**Household member categories**" cluster.

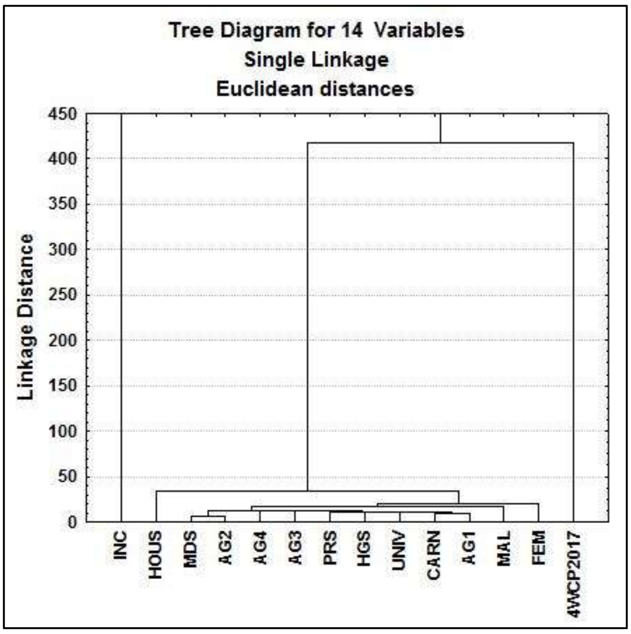


Fig 5.4 Dendrogram of the single link method applied for Scenario1

5.5.2. Scenario 2

Examination of figure 5.5 demonstrates that physical characteristics of housing units formulate two dissimilar groups of variables.

Total area and building area formulated the first cluster. The second cluster consists of WCP, number of rooms, garden area and frequency of garden watering. Based on this examination, the first cluster can be identified as "**Surface area**" cluster whilst the second cluster is labeled as "**Water distribution**" cluster.

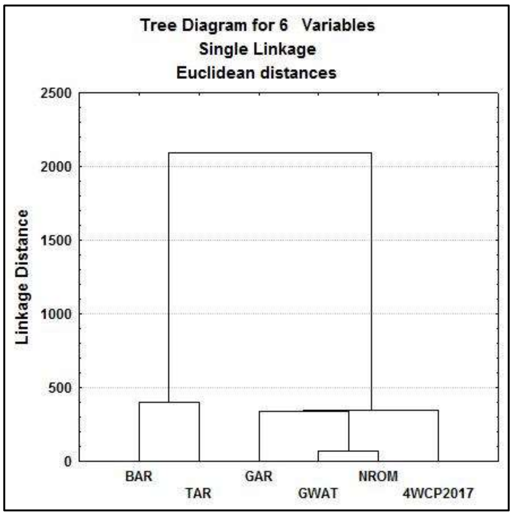


Fig 5.5 Dendrogram of the single link method applied for Scenario2

5.5.3. Scenario 3

Examination of figure 5.6 indicates two different clusters of indoor habits. The first cluster has WCP. The second cluster consists of MSHW, FSHW, WCL, UTLT, WCAR and WDISH. According to results of this scenario, the first cluster can be identified as "**Water consumption**" cluster and the second can be labeled as "**Personal usage**" cluster.

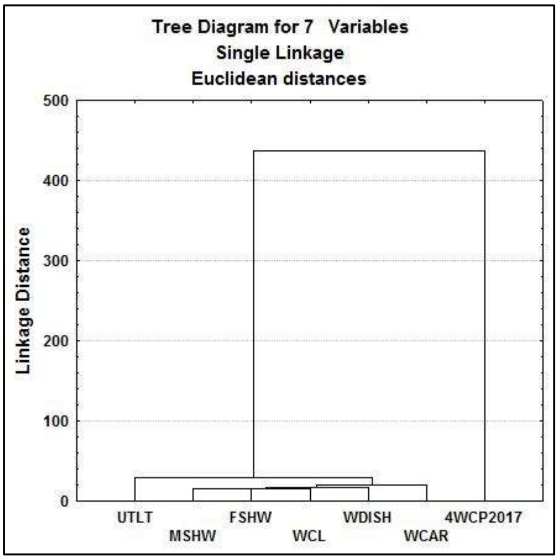


Fig 5.6 Dendrogram of the single link method applied for Scenario3

5.6.Principal Component Analysis (PCA)

The main purpose of PCA is to reduce the number of variables into a smaller number of dimensions (factors) and to classify variables and clusters of observations with similar characteristics with respect to these factors. Interpretation of the PC is based on finding which variables are most strongly correlated with each component, (*i.e.,* which of these numbers are large). Interpretation of the PCA results is based on the three scenarios.

5.6.1. Scenario 1

The **first scenario** is composed of 14 variables as shown in table 5.14, and thus the sum of all Eigenvalues is equal to 12. The number of factors was chosen in accordance to Kaiser's criterion and Cattell's scree test.

The scree plot in figure 5.7 indicates that the point where the continuous drop in Eigenvalues levels off is at factor 3. Therefore, three factors were chosen for analysis with cumulative variance of **77.47%.** The remaining Eigenvalues each account for less than **25%** of the total variance.

Table 5.14: Eigenvalues of correlation matrix, and related statistics-Scenario1

Eigenvalues of correlation matrix, and related statistics -Scenario1				
Value number	**Eigenvalue**	**% Total variance**	**Cumulative eigenvalue**	**Cumulative %**
1	7,044216	50,31583	7,04422	50,3158
2	2,363299	16,88071	9,40751	67,1965
3	1,438426	10,27447	10,84594	**77,4710**
4	0,928810	6,63436	11,77475	84,1054
5	0,746320	5,33085	12,52107	89,4362
6	0,481419	3,43871	13,00249	92,8749
7	0,399712	2,85509	13,40220	95,7300
8	0,250292	1,78780	13,65249	97,5178
9	0,155430	1,11022	13,80792	98,6280
10	0,100693	0,71923	13,90862	99,3473
11	0,050008	0,35720	13,95862	99,7045
12	0,041375	0,29554	14,00000	100,0000

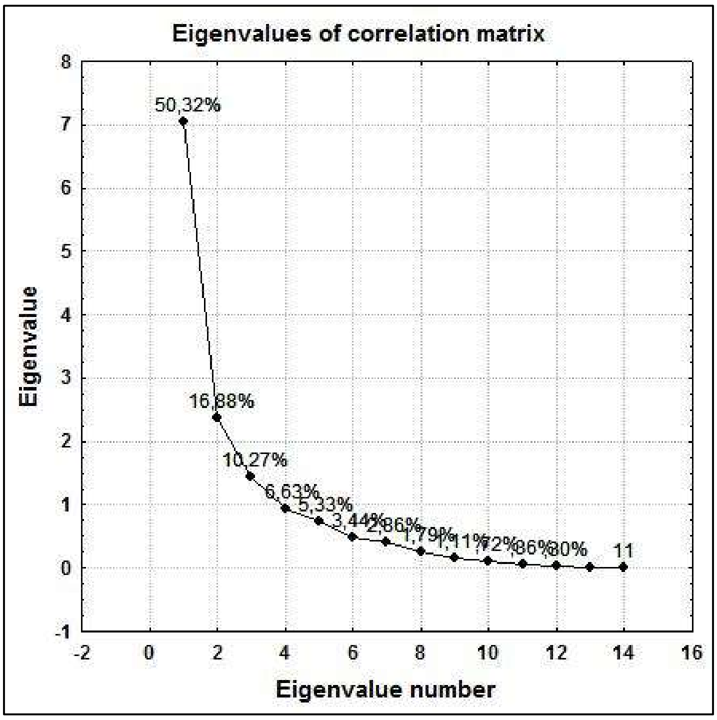


Fig 5.7 Eigenvalues of correlation matrix-Scenario1

The correlation between the principal components and the original variables are copied into the following table 5.15 for the Scenario1. Table 5.15 presents variances of factors and their loadings from variables.

Table 5.15: Factor-variable correlations (factor loadings)-Scenario1

Factor-variable correlations (factor loadings)-Scenario1			
Variable	**Factor 1**	**Factor 2**	**Factor 3**
WCP	**-0,959535**	-0,062106	-0,000679
FEM	**-0,825547**	0,110976	0,077618
MAL	-0,453011	**0,553028**	0,1988
HOUS	**-0,891466**	0,390065	0,169589
AG1	**-0,670351**	-0,154968	0,418533
AG2	0,132487	**0,875136**	-0,304989
AG3	**-0,83407**	-0,101097	-0,368774
AG4	-0,107173	0,394834	**0,797124**
PRS	**-0,739463**	-0,145702	-0,001512
MDS	0,027681	**0,853501**	-0,403908
HGS	**-0,806809**	-0,253965	-0,346726
UNIV	**-0,866912**	0,261089	0,056277
INC	**-0,884758**	-0,178942	-0,194394
CARN	**-0,731519**	-0,136975	-0,014491

The **first principal** component (first factor) corresponds to the largest eigenvalue **(7.04)** and accounts for approximately **50.32%** of the total variance. It is most correlated with water consumption, number of females, household size, first and the third age categories (AG1 and AG3), three education levels (primary, high school and university level), monthly income and car numbers (negative correlation). This first principal component increases with decreasing in the ten mentioned variables. This suggests that all the criteria vary together. If one increases, then the remaining ones tend to increase as well.

The **second factor** corresponding to the second eigenvalue **(2.37)** accounts for **16.88%** of the total variance. It is correlated with number of males, the second age category (AG2) and medium school for education level (positive correlation). This second PC increases with increasing in the three variables. The **third factor** corresponding to the eigenvalue **1.44** accounts for **10.27%.** It is correlated with the fourth age category (positive correlation). The third PC increases with increasing in this variable. Results show that water consumption is very strongly correlated with the first factor.

To complete the analysis, correlation circle (or variables chart) shows the correlations between the components and the initial variables. Figures 5.8 displays coordinates for the three factors. The current analysis is based on correlations, the largest factor coordinate (variable-factor correlation) that can occur is equal to 1, and also, the sum of all squared factor coordinates for a variable (squared correlations between the variables and all factors) cannot exceed 1.

Based on the magnitude of the factors coordinates (variable-factor correlations) for the variables in the analysis, the **first factor** can be labeled as "**household water consumption determinants**". **Second factor** can be labeled as "**household male teenagers**" and the **third factor** can be labeled as **"old residents"**. Figure 5.9 shows the factor coordinates for all houses.

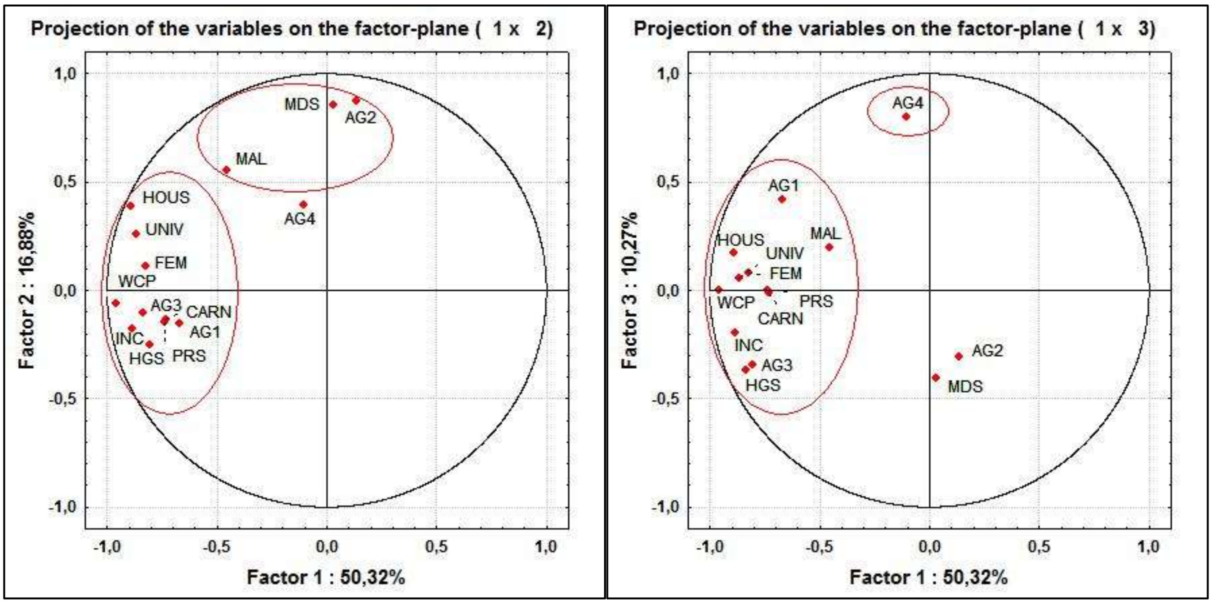


Fig 5.8 Projection of the variables on the factor-plane (1*2) and (1*3)-Scenario1

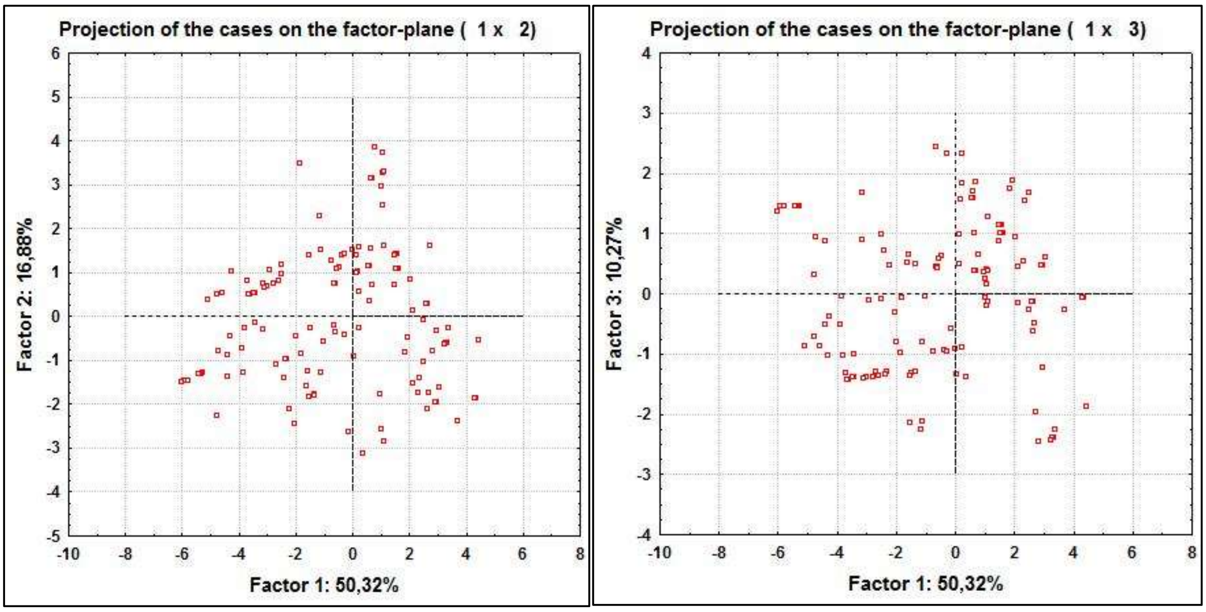


Fig 5.9 Projection of the cases on the factor-plane (1*2) and (1*3) -Scenario1

5.6.2. Scenario 2

The scree plot in figure 5.10 indicates that the point where the continuous drop in Eigenvalues levels off is at factor 2. Therefore, two factors were selected for analysis with cumulative variance of **97.59%** (table 5.16)**.**

Table 5.16: Eigenvalues of correlation matrix, and related statistics-Scenario2

Eigenvalues of correlation matrix, and related statistics-Scenario2				
Value number	**Eigenvalue**	**% Total variance**	**Cumulative eigenvalue**	**Cumulative %**
1	3,841056	64,01760	3,841056	64,0176
2	1,366578	22,77630	5,207634	86,7939
3	0,648320	10,80533	5,855954	97,5992
4	0,098915	1,64858	5,954869	99,2478
5	0,045131	0,75218	6,000000	100,0000

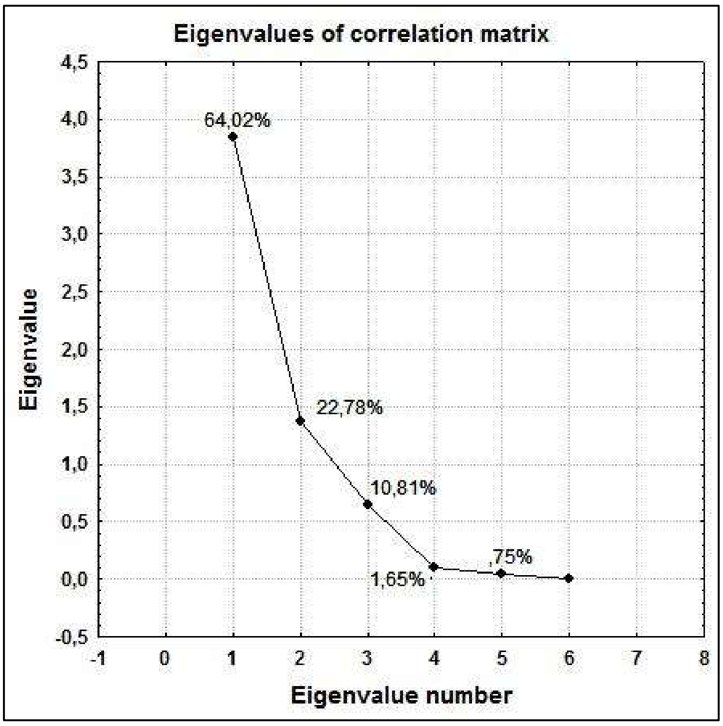


Fig 5.10 Eigenvalues of correlation matrix-Scenario2

Table 5.17 and figure 5.11 present variances of factors and their loadings from variables. The first factor corresponds to the largest eigenvalue **(3.84)** and accounts for **64.08%** of the total variance. It is most correlated with water consumption, the total area of the house, building area and number of rooms (positive correlation). The **first factor** labeled as **"household water consumption determinants**". The **second factor** corresponding to the eigenvalue **(1.37)** and accounts for **22.78%** for the total variance. It is correlated with garden area and frequency of garden watering (positive correlation) and can be labeled as "**garden area**". Figure 5.12 presents the factor coordinates for all houses.

Table 5.17: Factor-variable correlations (factor loadings)-Scenario2

Factor-variable correlations (factor loadings)-Scenario2		
Variable	**Factor 1**	**Factor 2**
WCP	**0,982343**	-0,011908
TAR	**0,990859**	0,013854
BAR	**0,967676**	-0,180934
ROMN	**0,96225**	0,013168
GAR	0,177688	**0,815258**
GWAT	0,018983	**0,817734**

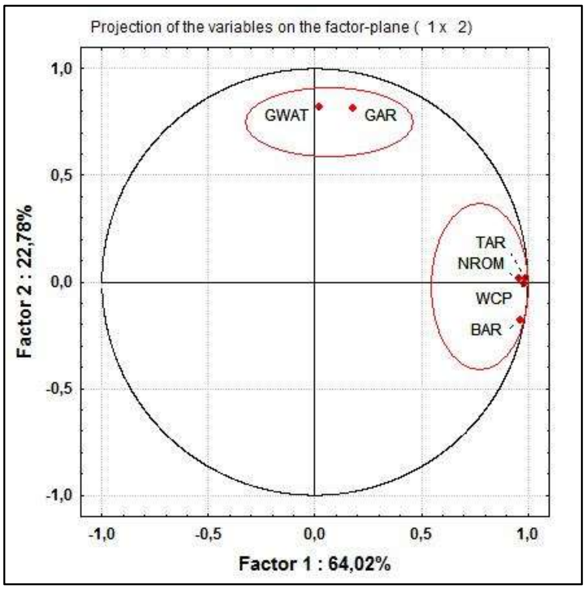


Figure 5.11 Projection of the variables on the factor-plane (1*2)-Scenario2

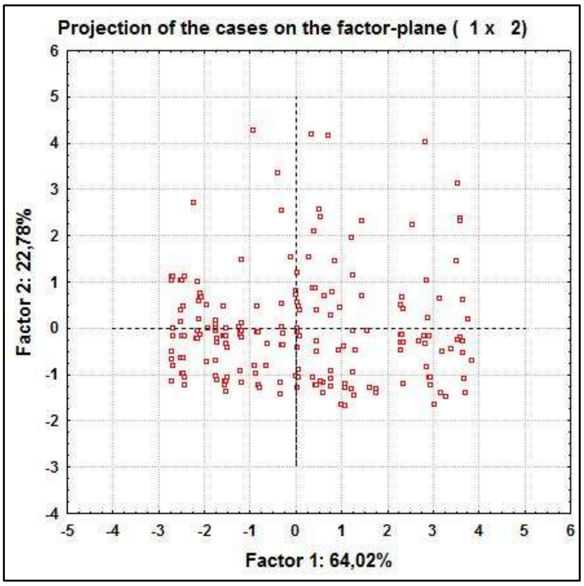


Fig 5.12 Projection of the cases on the factor-plane (1*2)-Scenario2

5.6.3. Scenario 3

From the eigenvalues of correlation matrix of the scenario3 (Table 5.18) and the scree plot (Figure 5.13), three factors were chosen for analysis with a variance of **60.90%.**

Table 5.18: Eigenvalues of correlation matrix, and related statistics-Scenario3

Eigenvalues of correlation matrix, and related statistics-Scenario3				
Value number	**Eigenvalue**	**% Total variance**	**Cumulative eigenvalue**	**Cumulative %**
1	1,829360	26,13371	1,829360	26,1337
2	1,476067	21,08667	3,305426	47,2204
3	0,957885	13,68407	4,263311	60,9044
4	0,871560	12,45086	5,134872	73,3553
5	0,758398	10,83425	5,893269	84,1896
6	0,624276	8,91823	6,517545	93,1078
7	0,482455	6,89221	7,000000	100,0000

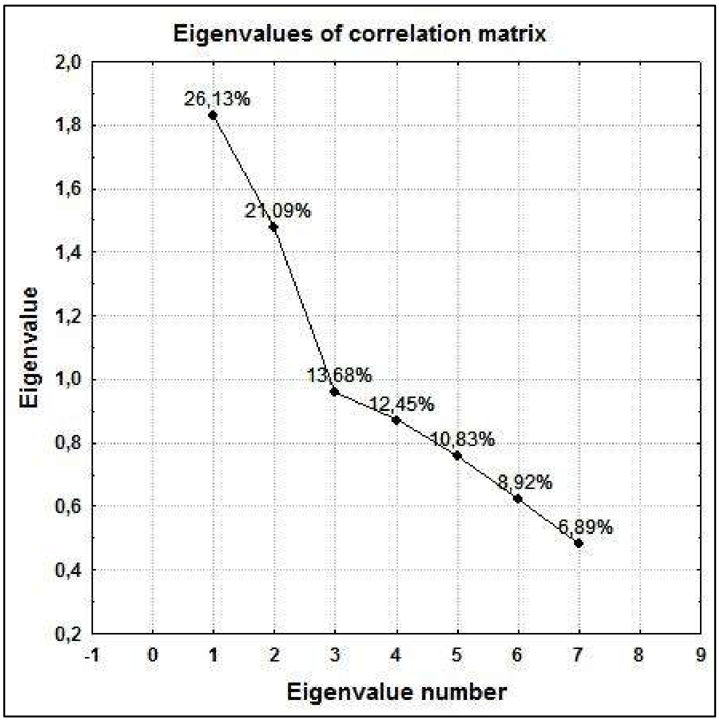


Fig 5.13 Eigenvalues of correlation matrix-Scenario3

Table 5.19 and figure 5.14 present variances of factors and their loadings from variables.

The first factor corresponds to eigenvalues **(1.83)** and accounts for **26.13%** of the total variance.

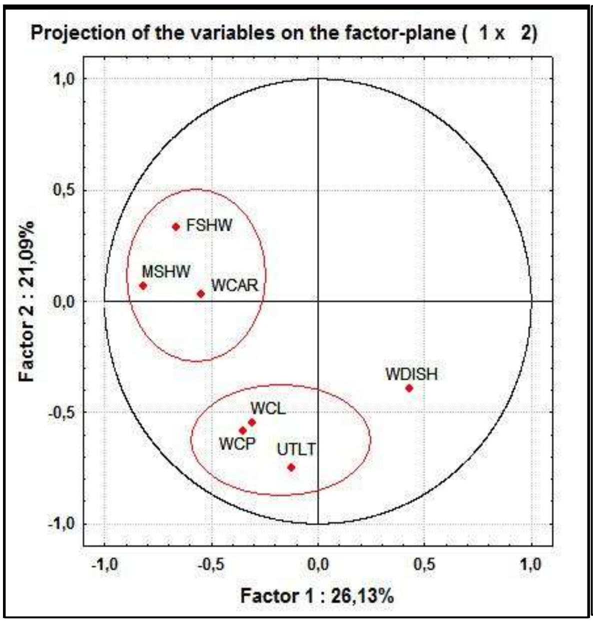
It is correlated with frequency of washing cars frequency of showering for males and females (negative correlation). The **first factor** labeled as **"household water consumption determinants**".

The second factor has eigenvalue equal to **1.48** and accounts for **21.09%** for the total variance. It is correlated with water consumption, frequency of using toilet and washing clothes (negative correlation). The **second factor** can be labeled as "**indoor habits**".

The **third factor** corresponding to the eigenvalue **0.96** accounts **13.68%** of the total variance. It is correlated with frequency of dishwashing, (negative correlation) and it can be labeled as "**dishwashing habit**". Figure 6.15 presents the factor coordinates for all houses.

Table 5.19: Factor-variable correlations (factor loadings)-Scenario3

Factor-variable correlations (factor loadings)-Scenario3			
Variable	**Factor 1**	**Factor 2**	**Factor 3**
WCP	-0,350886	**-0,585081**	0,20578
WCAR	**-0,546793**	0,033772	-0,327041
WDISH	0,43145	-0,39262	**-0,6795**
WCL	-0,305113	**-0,549009**	-0,379842
UTLT	-0,120419	**-0,750049**	0,343907
FSHW	**-0,66476**	0,331665	-0,276783
MSHW	**-0,819517**	0,066853	0,087764



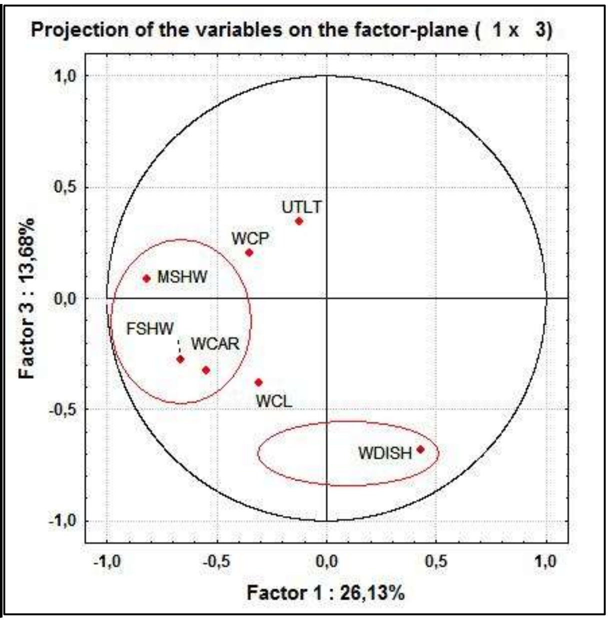


Fig 5.14 Projection of the variables on the factor-plane (1*2) and (1*3)-Scenario3

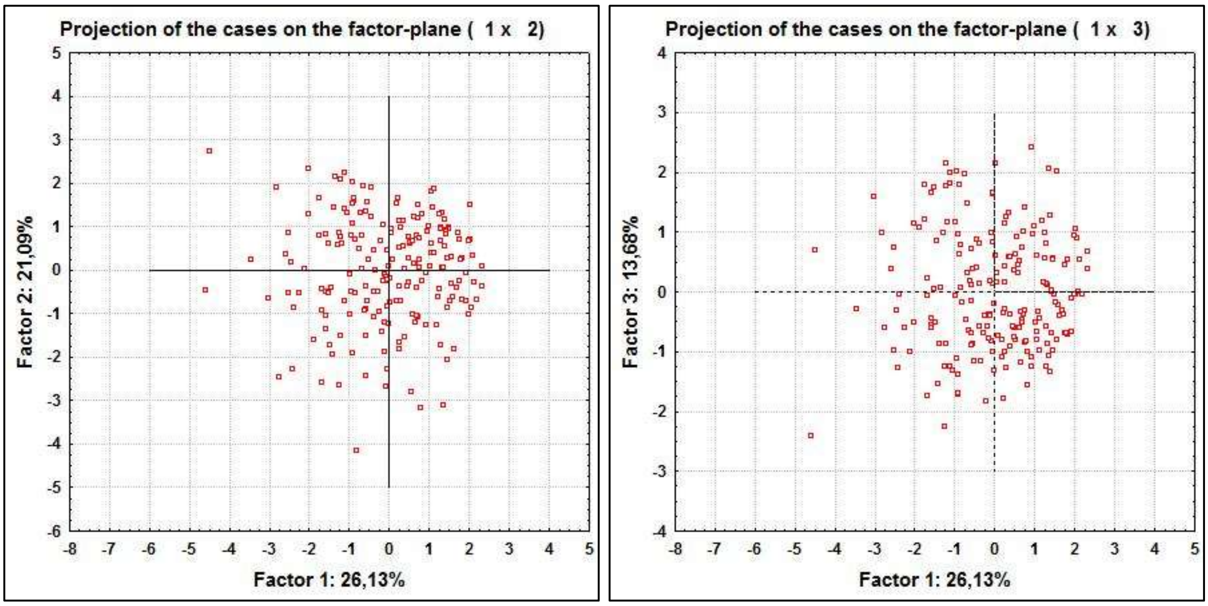


Fig 5.15 Projection of the cases on the factor-plane (1*2) and (1*3)-Scenario3

5.7.Comparison between Cluster analysis and PCA results

Results between the cluster analysis and PCA have similarities and differences for the three scenarios. PCA provides more details about groups of variables (factors) and association of cases (houses) with the corresponding variables. Also, it gives the weight of each group of variables reflected by the variance value and presents the variables loadings on factors reflecting their significance and priority.

The results of PCA analysis can be applied for formulating priority strategy programs to handle the water stresses is specified geographic areas. However, the cluster analysis can be used as early exploratory tool to investigate the hierarchy and shapes of possible groups of cases and corresponding variables (Jalala, 2005).

5.8. Factor Analysis (FA)

Factor analysis is like cluster analysis and PCA. The purpose of using FA is to compare their results with PCA results. This technique reduces the number of observed variables per scenario to a smaller number of unobserved latent factors, which are correlated with each other, and classifies variables within these factors (Jalala, 2005).

The interpretability of factors can be improved through rotation. Rotation maximizes the loading of each variable on one of the extracted factors whilst minimizing the loading on all other factors. It works through changing the absolute values of the variables while keeping their differential values constant. The Varimax normalized rotation was adopted.

5.8.1. Scenario 1

The table 5.20 below and the figure 5.16 illustrate the three factor rotated solution with the cross-loadings of their classified variables.

Factor 1 includes inter-correlated observed variables that are water consumption, household size, number of female, the first and the third age categories (AG1 & AG3), three education levels (primary, hight school and university level), monthly income and number of cars. The factor 1 represents "**householdwater consumption determinants"**.

Factor 2 has the second age category (AG2) and medium school (education level). The factor 2 represents **"household teenagers"**.

Factor 3 has two variables that are number of males and the fourth age category (AG4). The factor 3 represents **"old males residents"**.

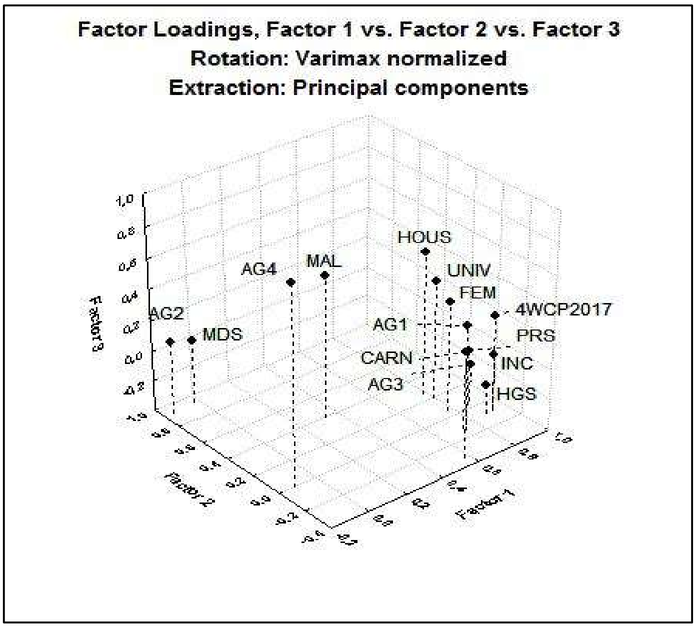


Fig 5.16 Factor loadings, factor1 vs. factor2 vs. factor3_Scenario1

Table 5.20: Factor loadings_Scenario1

Factor Loadings_Scenario1 (Varimax normalized) Extraction: Principal components (Marked loadings are >0,700000)			
Variable	**Factor - 1**	**Factor - 2**	**Factor - 3**
WCP	**0,935863**	-0,065523	0,210789
FEM	**0,770921**	0,056661	0,319904
MAL	0,338176	0,403796	**0,522653**
HOUS	**0,785882**	0,265836	0,536045
AG1	**0,569825**	-0,330757	0,463092
AG2	-0,146857	**0,921259**	0,078422
AG3	**0,901857**	0,063202	-0,156673
AG4	-0,116874	0,001511	**0,888327**
PRS	**0,731261**	-0,137856	0,119541
MDS	-0,020692	**0,944383**	0,009247
HGS	**0,885779**	-0,083416	-0,20999
UNIV	**0,800777**	0,200347	0,376149
INC	**0,919155**	-0,084118	-0,026155
CARN	**0,725638**	-0,124216	0,110051
Expl.Var	6,707824	2,183804	1,954311
Prp.Totl	0,47913	0,155986	0,139594

5.8.2. Scenario 2

Table 5.21 and figure 5.17 shows that the **first factor** contains water consumption, the total area of the house, building area and number of rooms. The factor 1 represents **"household water consumption determinants"**. **Factor2** has garden area and frequency of garden watering. This factor represents **"garden area"**.

Table 5.21: Factor loadings_Scenario2

Factor Loadings_Scenario2 (Varimax normalized) Extraction: Principal components (Marked loadings are >0,700000)		
Variable	**Factor - 1**	**Factor - 2**
WCP	**0,98018**	0,066226
TAR	**0,986622**	0,092583
BAR	**0,978998**	-0,103431
ROMN	**0,958158**	0,089625
GAR	0,112313	**0,826804**
GWAT	-0,046087	**0,816655**
Expl.Var	3,825417	1,382217
Prp.Totl	0,637569	0,23037

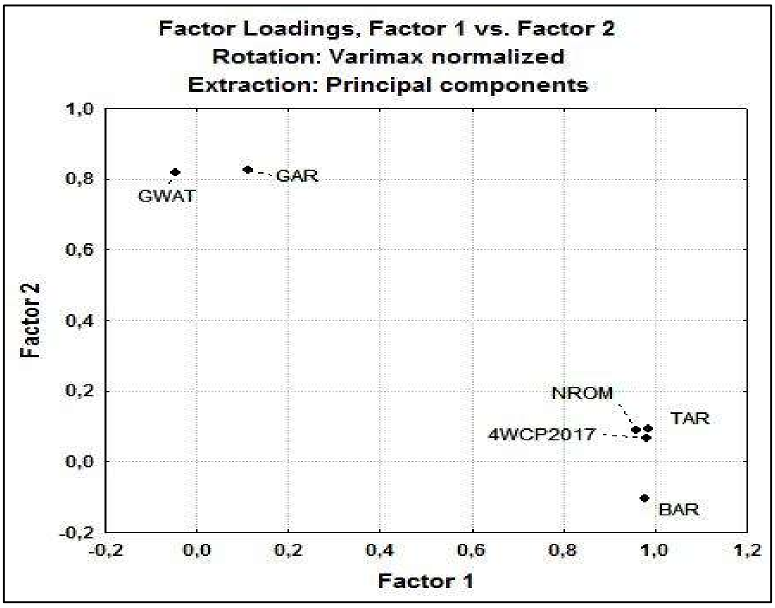


Fig 5.17 Factor loadings, factor1 vs. factor2_Scenario2

5.8.3. Scenario 3

In the **last scenario**, table 5.22 and figure 5.18 show that **factor1** includes frequency of washing cars, dishwashing and showering for females and males. The first factor represents **"indoor habits"**. **Factor 2** has water consumption, frequency of washing clothes and using toilet. The second factor represents **"household water consumption determinants"**.

Table 5.22: Factor loadings_Scenario3

Factor Loadings_Scenario3 (Varimax normalized) Extraction: Principal components (Marked loadings are >0,700000)		
Variable	**Factor - 1**	**Factor - 2**
WCP	0,116463	**0,672218**
WCAR	**0,522224**	0,165546
WDISH	**-0,543949**	0,210757
WCL	0,086765	**0,622075**
UTLT	-0,157968	**0,743048**
FSHW	**0,739616**	-0,06982
MSHW	**0,788544**	0,232969
Expl.Var	1,783479	1,521948
Prp.Totl	0,254783	0,217421

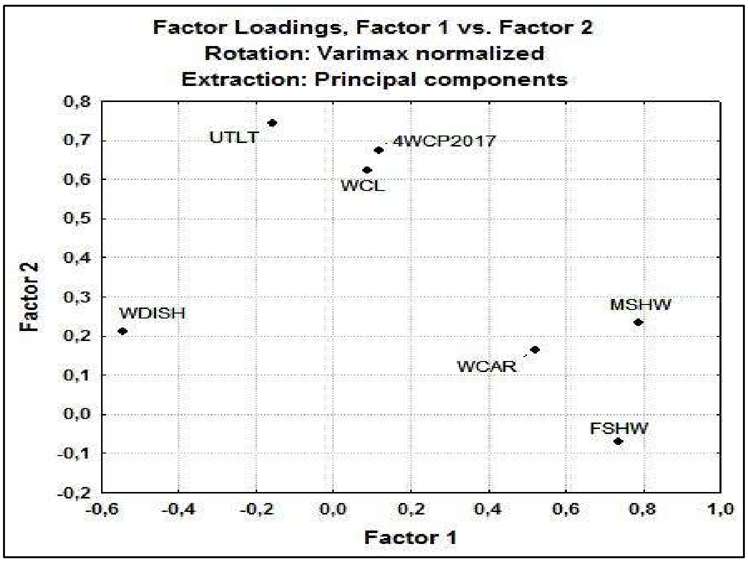


Fig 5.18 Factor loadings, factor1 vs. factor2_Scenario3

5.9.Comparison between FA and PCA results

Comparison with the PCA results for the three scenarios revealed that the results of factor analysis in scenario1 and scenario2 are similar to the PCA results, but they have different factor loadings. In scenario3 results are different.

5.10. Geographical distribution of results

Figures below illustrate the distribution of 12 explanatory variables in the study area with indoor habits of residents. As demonstrated in the maps, water consumption is highly compatible with the determinants. Moreover, the increasing in these factors affects directly the water usage.

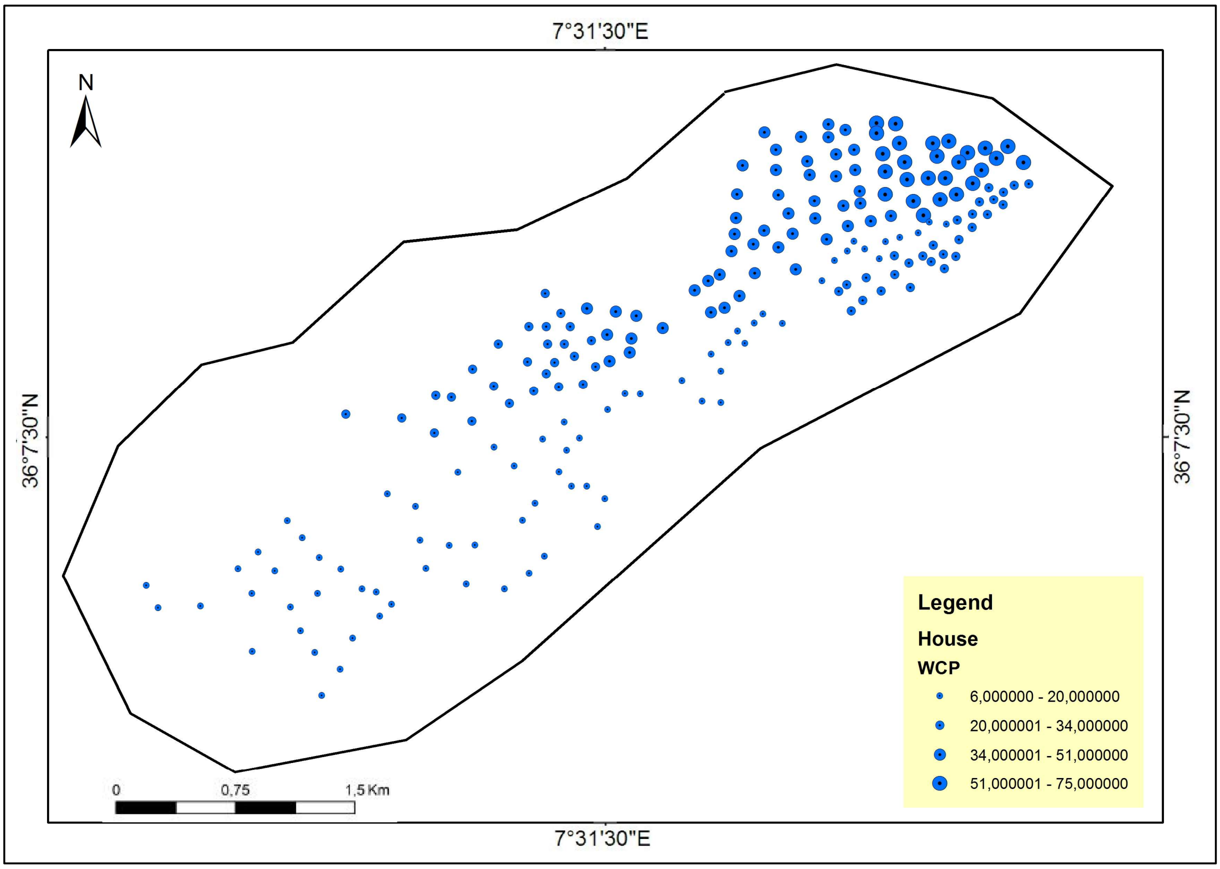


Fig 5.19 Water consumption distribution

The figure 5.19 represents the geographical distribution of WCP across the study area. It shows that the higher users are located on the north-east of the city, while the lowest consumers are located in the south west.

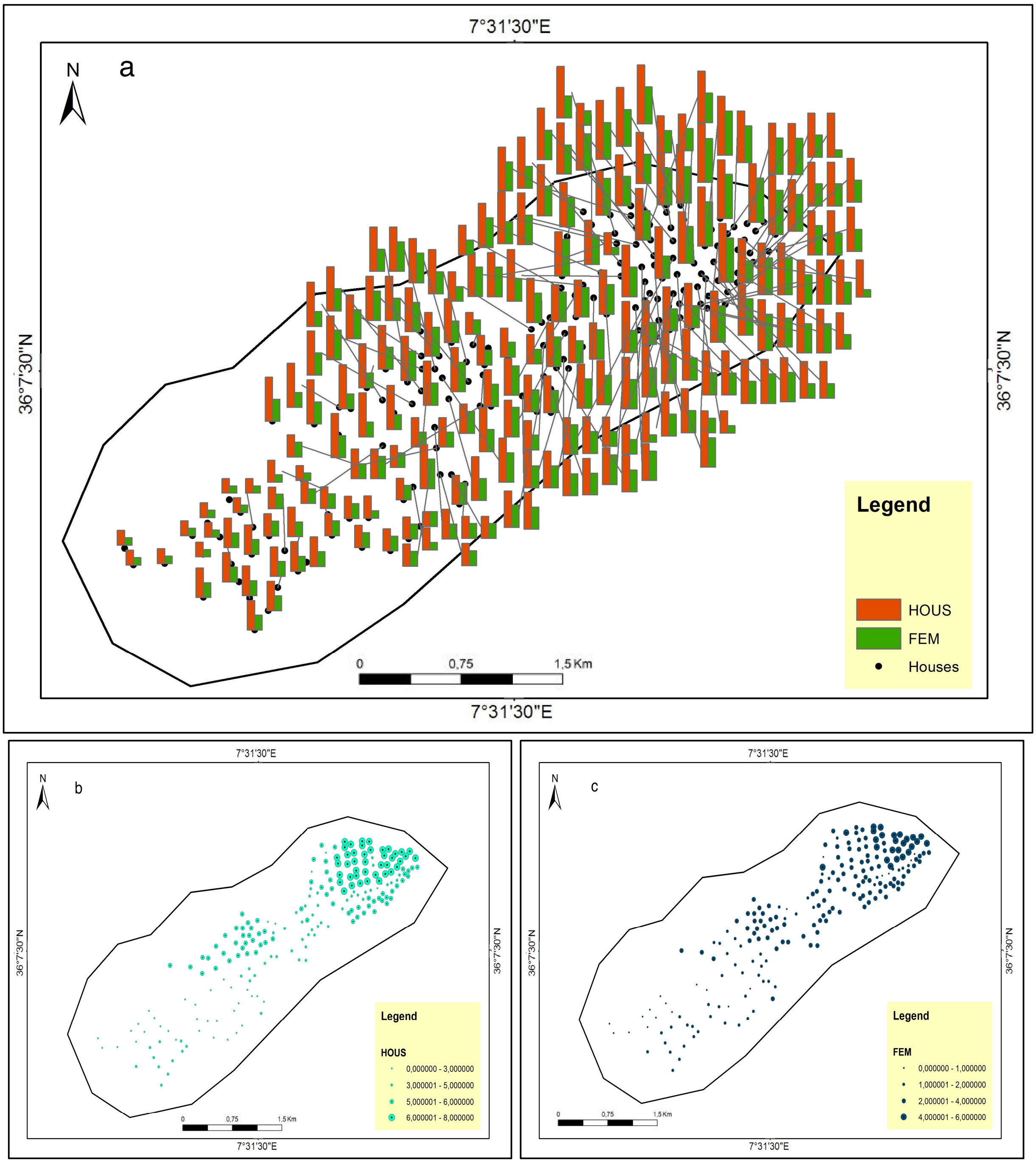


Fig 5.20 Distribution of household size and number of females

The gender distribution (number of females by household size) is almost the same across the study area (figures 5.20 a, b and. c), but the total resident per house (In orange color) is highly variable and agrees mainly the WCP geographic distribution.

Another factor related to resident that found to be correlated to water usage is education level. In fact, the houses with high number of residents have at least two categories of the education level (primary/high school, high school/university) and sometimes the three (primary/high school/university) figures 5.21 a, b, c and d.

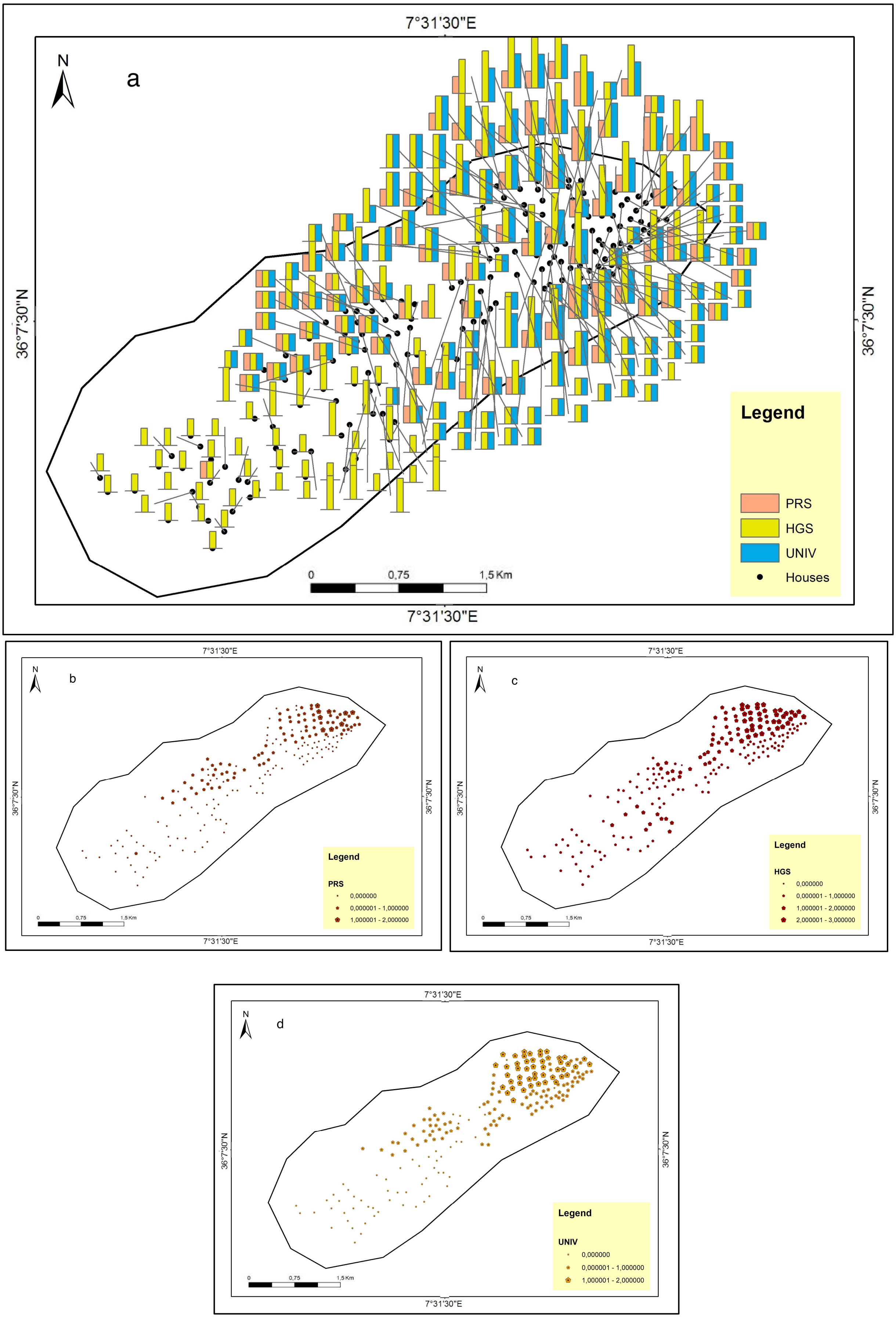


Fig 5.21 Distribution of education level

Also, age has a big effect on WCP as confirmed in figures 5.22 a, b and c below. Another remark could be made basing on these figures (figures 5.23 and figures 5.24) is that houses located on the south west are relatively younger families.

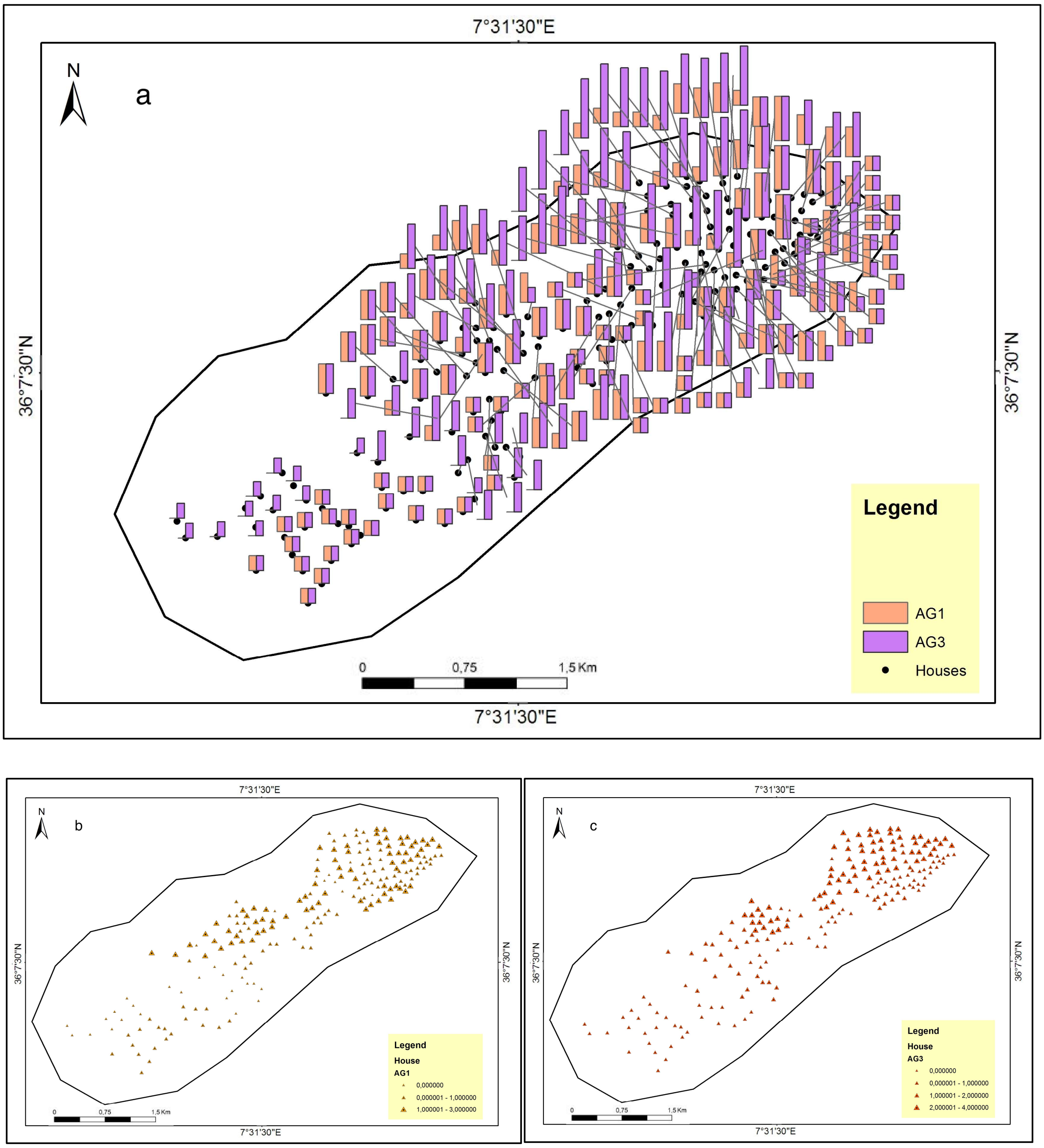


Fig 5.22 Distribution of household age

Figure 5.23 below demonstrates the distribution of monthly income in the region. The graphically distribution of income is similar to WCP distribution (figure 5.19). in fact, income is one of the most important determinants of water usage. In addition, figure 6.24 represent car possession. This last is similar to income distribution.

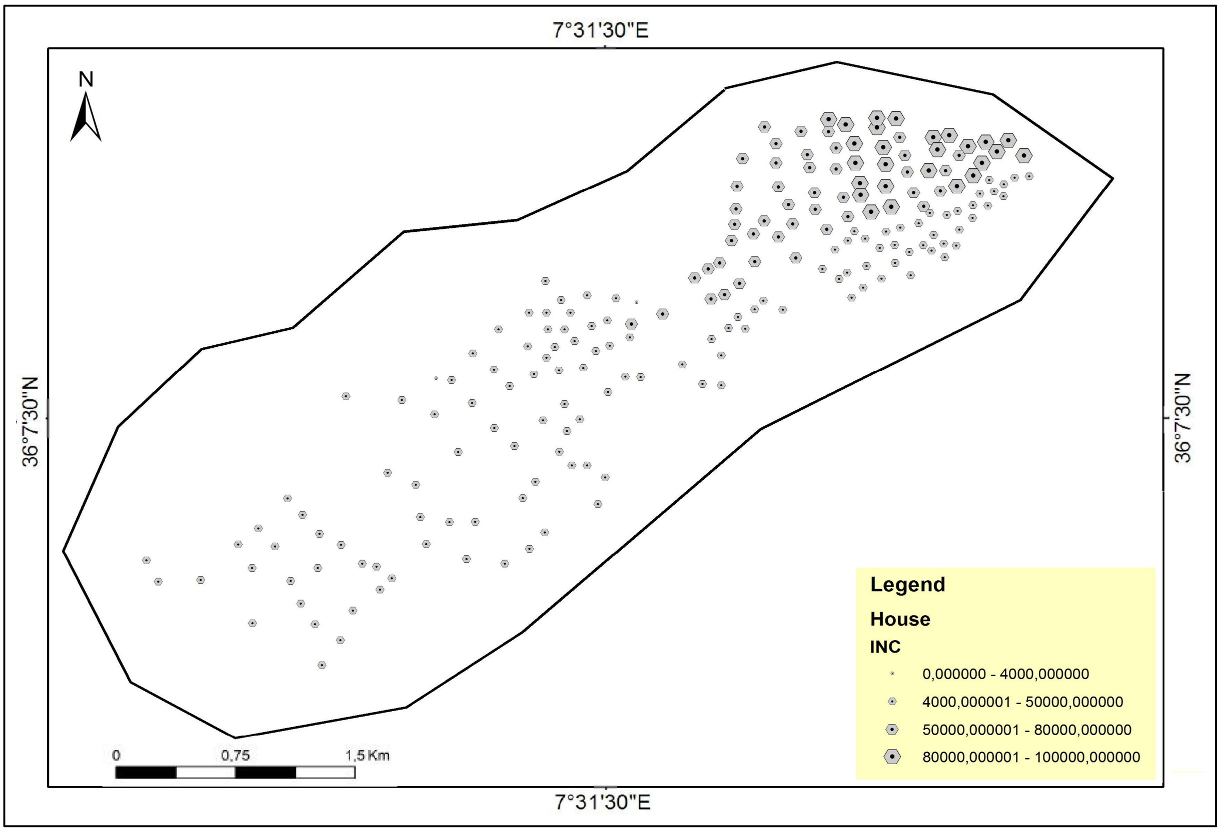


Fig 5.23 Monthly income distribution

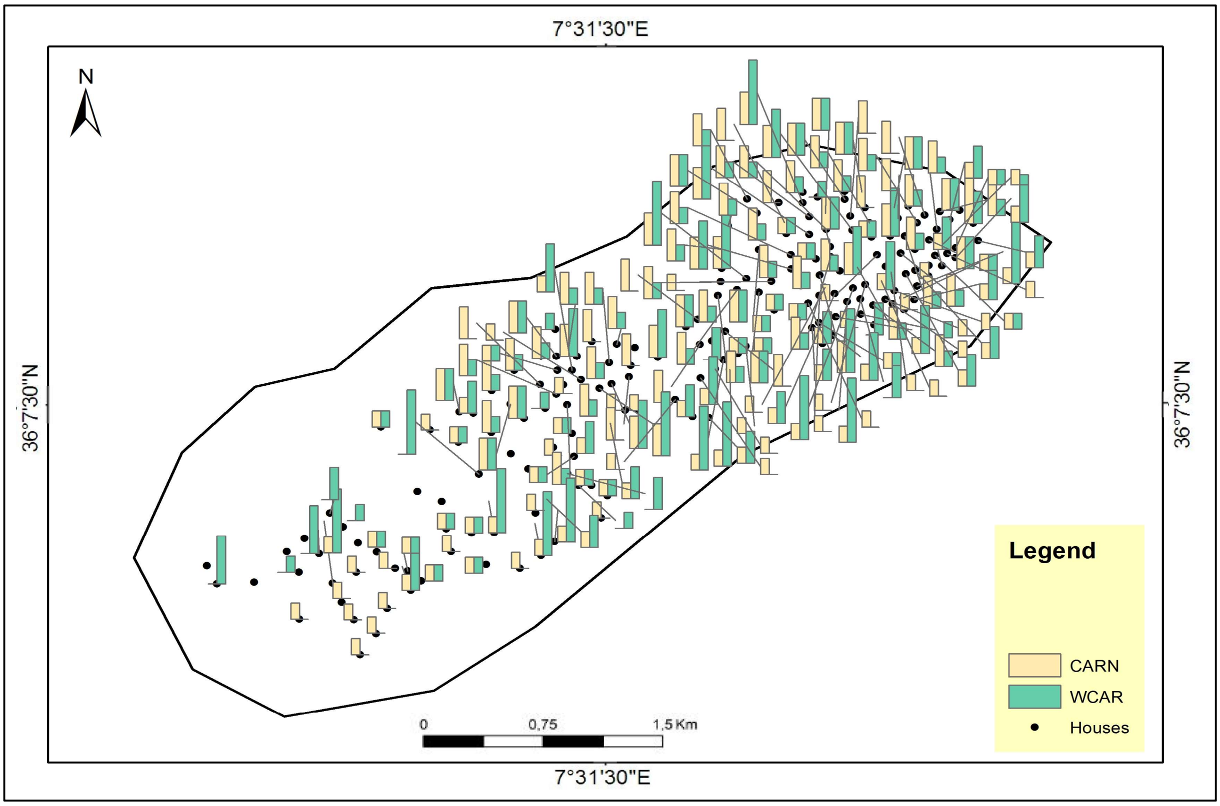


Fig 5.24 Car possession distribution

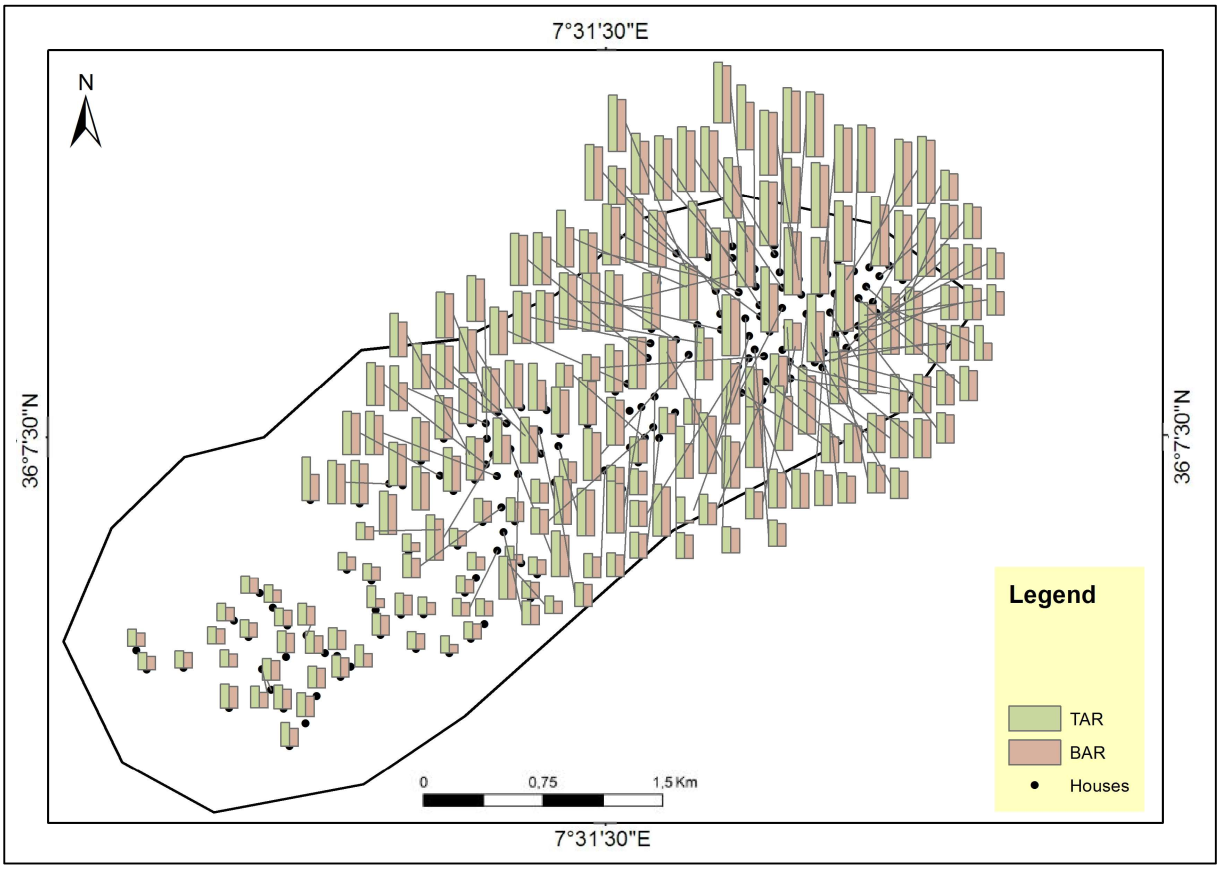


Fig 5.25 Distribution of total area and building area

By comparison, building, total area and number of rooms are related to each other (figure 5.25 and figure 5.26), this link between these parameters is logical (more the total surface, the bigger the house and number of rooms). The graphical distribution shows three main categories big, medium and small: the north and middle, north east to south east and finally south west area respectively. Figure 5.25 also demonstrates the link between WCP and TAR.

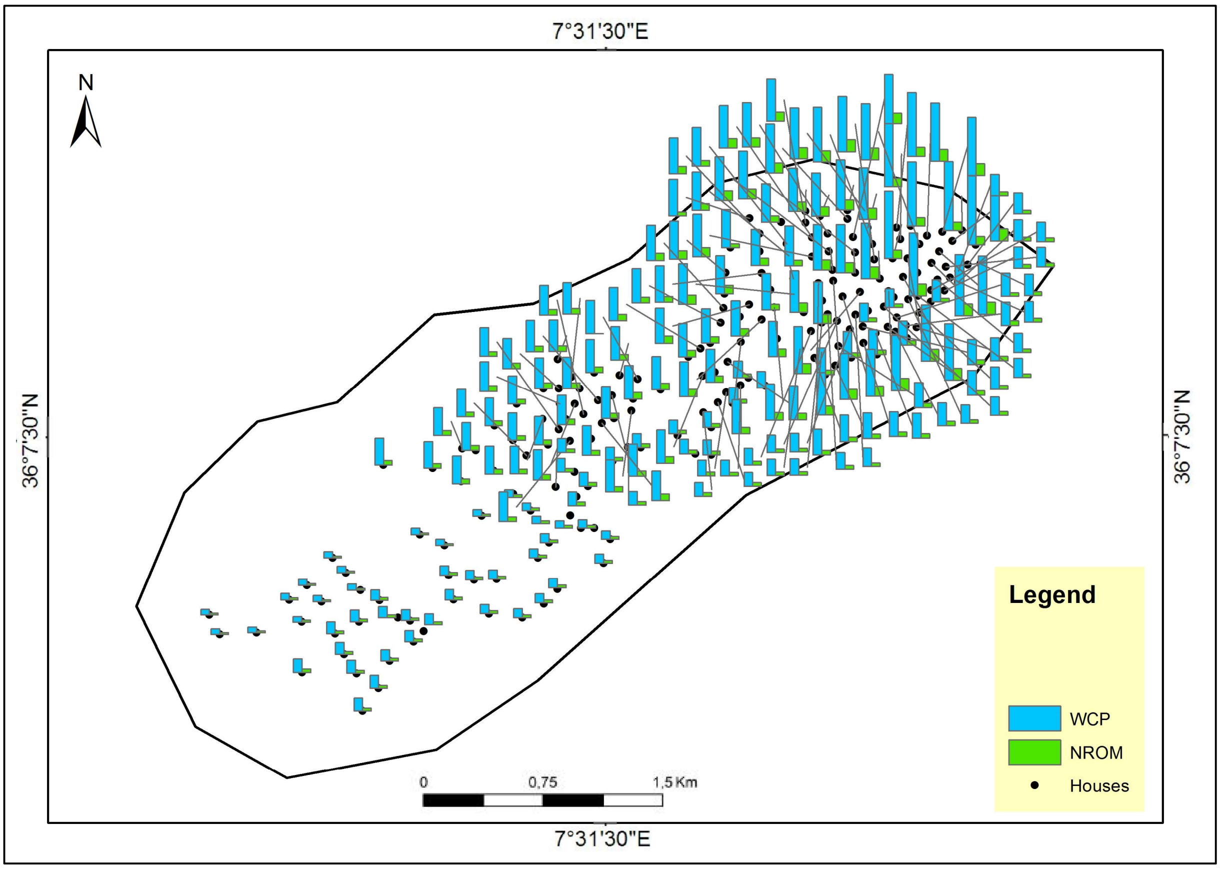
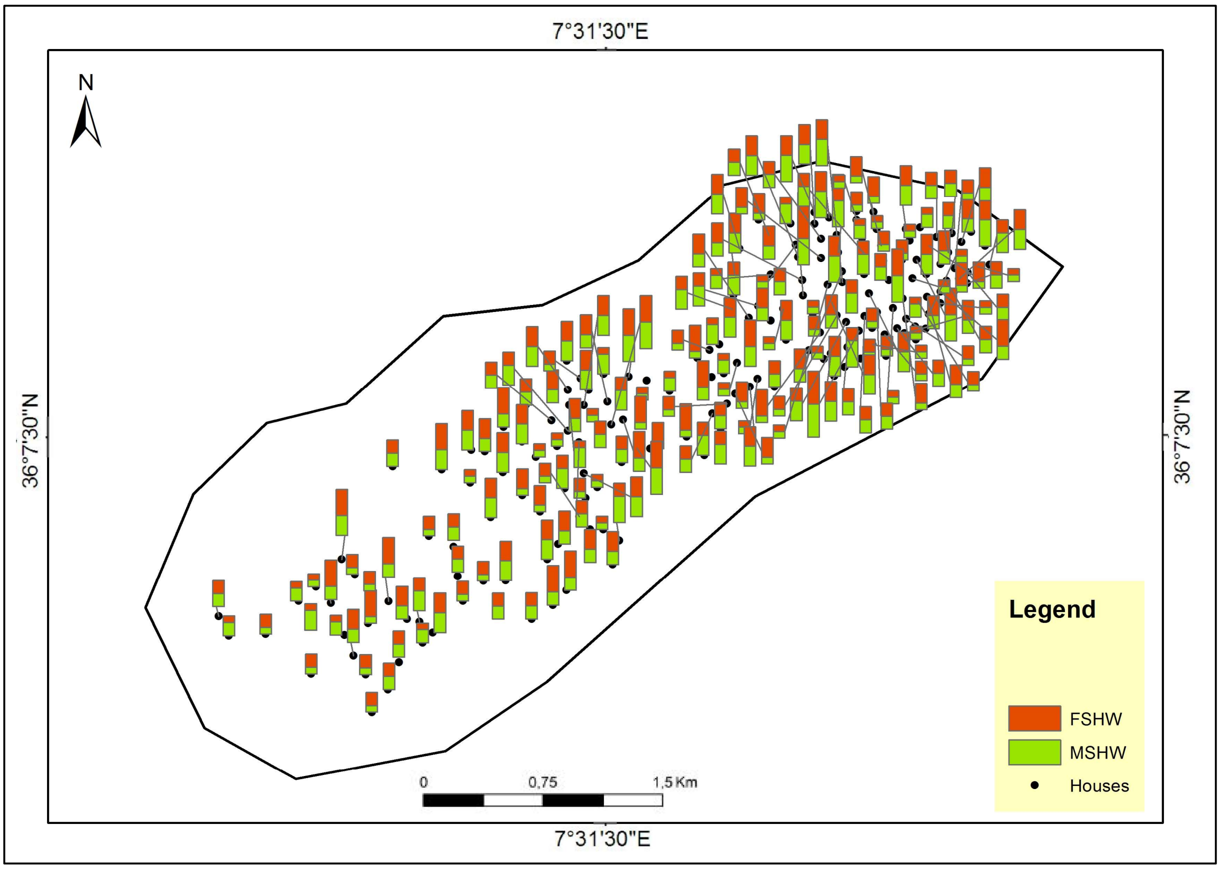


Fig 5.26 Relationship between water consumption and number of rooms

The following maps shows indoor habits distribution.



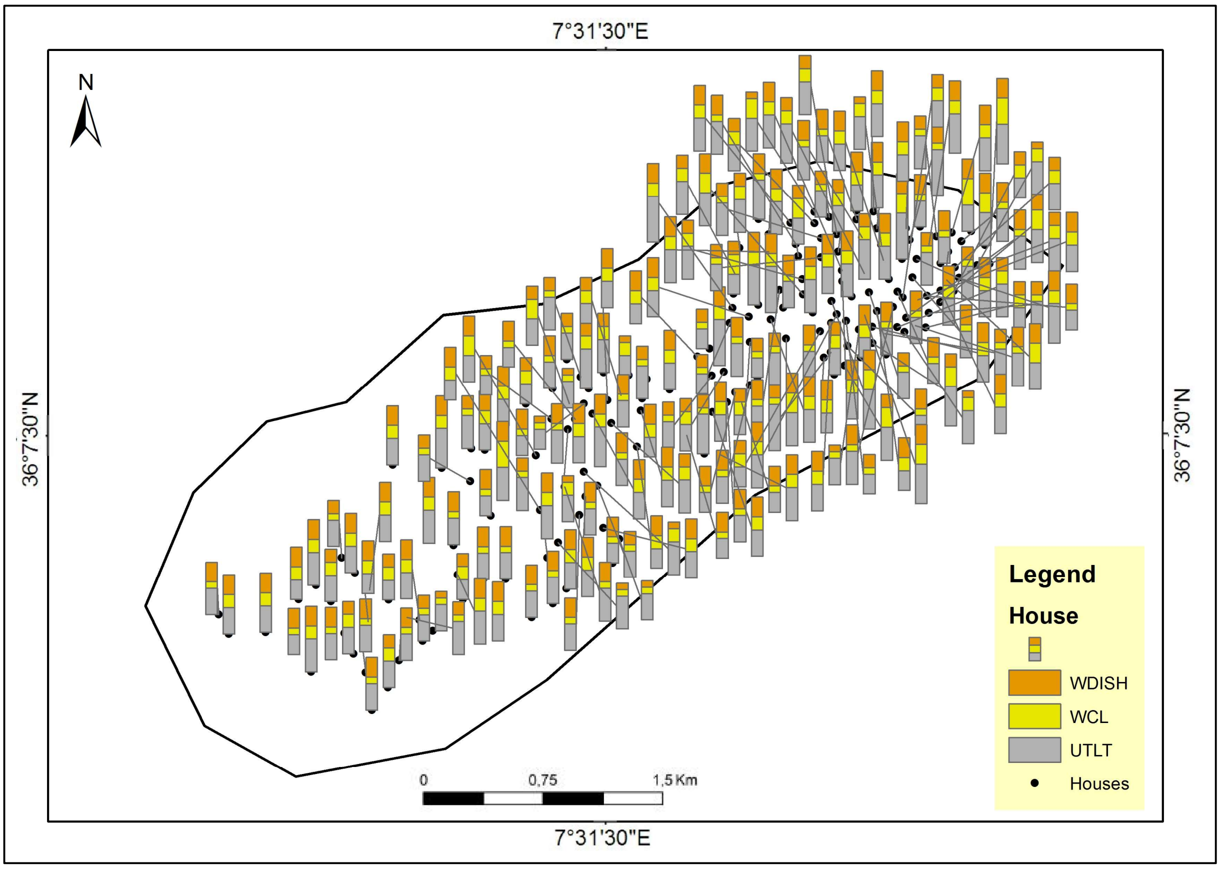


Fig 5.27 Indoor habits distribution

5.11. Conclusion 1

The present section of thesis presents the results of correlation analysis, ANOVA, cluster analysis, factor analysis and principal component analysis between water consumption and different parameters. The correlation analysis gives the following conclusions:

- Household size and monthly income are very strongly correlated with water consumption.
- Females use more water.
- In age groups; children (AG1) and adult people (AG3) are responsible for water usage.
- Education level of residents has an impact on water use, where residents having primary, high school and university education level consume more water than those having medium school.
- Houses with cars tend to use more water.
- Water consumption is very strongly correlated with total area of the house, building area and number of rooms. As consequence, all the previous parameters are considered as determinants of domestic water consumption.

For ANOVA, results support the previous results and demonstrate the influence of tested inputs on WCP.

Comparison between cluster, factor and principal component analyses demonstrate a variety of results with different factor loadings. Also, principal component analysis and factor analysis have similar results indicating :

- Strong correlation between water consumption and number of females, household size, first and the third age categories (AG1 and AG3), three education levels (primary, high school and university level), monthly income and car numbers.
- Water use very strongly correlated with total area of the house, building area and number of rooms.
- For indoor habits water consumption is associated with frequency of washing clothes and using toilets.

In conclusion, the five tests demonstrate the impact of some parameters of the three scenarios on household water consumption. As a result, 12 variables were chosen as determinants of household water use that are: number of females, household size, first and the third age categories (AG1 and AG3), three education levels (primary, high school and university level), monthly income, car numbers, total area of the house, building area and number of rooms. In the part II of this chapter, the focus will be only on the 12 explanatory variables.

Part II:
Application of Artificial Neural Network (ANNs), Adaptive Neuro Fuzzy Inference System (ANFIS) to predict water consumption

As stated in the Chapter 4, the modeling is conducted by two phases: Artificial Neural works (ANNs) Phase and Adaptive Neuro Fuzzy Inference System (ANFIS) Phase.

5.12. Artificial Neural Networks (ANNs)

In this part of thesis, the ANNs models are used for characterizing and prioritizing the effective parameters affecting water consumption (WCP) in the three scenarios. The applied methodology is illustrated in figure 5.28.

The neural network for identifier is designed as a three-layer neural network (input/ hidden/output layers). The neuron numbers in the hidden layers can be chosen depending on the practical training result. The training model adaptive neural network approach based on Back propagation algorithm is applied to assess household water consumption determinants.

Dataset	
Input parameters: The three scenarios of parameters	**Output parameter:** Water consumption

Input parameters: Scenarios	
• **Scenario1 (S1):** Socio-economic parameters • **Scenario2 (S2):** Physical characteristics of housing units	• **Scenario3 (S1+S2):** Socio-economic parameters + Physical characteristics of housing units

Data selection in Matlab Software		
Training data size: **(50%; 60%; 70% and 80%)**	Testing data size: **(10%; 15%; 20% 25%)**	Validation Data size: **(10%; 15%; 20% 25%)**

Preparation for training the neural network		
Network type: Multilayer perceptron (**MLP)**	Rising the number of network training: **50**	Select the best network result based on **5** performance criteria for each scenario

Results obtained	
MES (mean square error)	**R** (Correlation coefficient)

Fig 5.28 The applied methodology for ANNs approach

5.12.1. Input parameters

It is expected that the sensitivity of WCP differs accordingly to the considered input parameters, and this mainly because the variation of correlation strength between WCP and each parameter (Part 01 of Chapter 05).

- **Data initialization**

Before creating the input vectors and output vector, the data must be in normalized form. This step in modelling aims to convert the input and output values into representative range of values for the computation using ANNs. It was applied already in the previous Chapter 03. As mentioned in part I of this chapter, the total selected variables (explanatory variables) are 12 that are categorized into three scenarios. in the current work, the back-propagation feed-forward MLP with sigmoidal-type activation function (equation 5.1) thanks to its popularity and mostly high performance compared to other networks (Lippmann, 1987).

$$f(z) = \frac{1}{1+\exp^{-z}} \qquad \textbf{(5.1)}$$

Performance function are one of the important factors affects the learning performance. For feed forward network, **"Mean Square Error (MSE)"** is commonly used performance function. This last calculate the cumulative values between the target outputs values and created outputs by the network.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left[e(t)^2\right] \qquad \textbf{(5.2)}$$

Where; *e (t)* is forecast error at period *t*, *n* is number of periods (Agami et al., 2009)

a. Scenario1 (S1)

The target output Water Consumption (WCP) can be considered as a function of: household size (HOUS), number of female (FEM), monthly income (INC), two age categories (AG1 and AG3), three categories of education level (PRS, HGS & UNIV) and car numbers (CARN). The MLP network can be expressed by:

WCP= f (HOUS, FEM, INC, AG1, AG3, PRS, HGS, UNIV, CARN) **(5.3)**

AG2 AG4 MAL MDS are removed from further analysis because of their week correlation to WCP. Removing such inputs would improve the process and reduces the time.

b. Scenario2 (S2)

Scenario 2 represents the explanatory variables of physical characteristics of housing units. In this scenario, WCP is considered a function of the following parameters: total area of the house (TAR), building area (BAR) and number of rooms (NROM). The MLP network can be expressed by:

WCP= f (TAR, BAR, NROM) **(5.4)**

Also, because of their week impact on WCP, Garden area (GAR) and watering frequency (GWAT) are no more taken into consideration.

c. Scenario3 (S1+S2)

To confirm the sensibility of WCP to number of inputs, all the factors are considered in this scenario.

This scenario is a combination between the first and the second scenarios where the total explanatory variables are 12 parameters. The MLP network can be expressed by the following function:

WCP= f (HOUS, FEM, INC, AG1, AG3, PRS, HGS, UNIV, CARN, TAR, BAR, NROM) **(5.5)**

5.12.2. Training, testing and validating the neural networks

The neural network models are trained to learn the forward dynamics of water consumption. 12 inputs and one output (WCP) are selected as the identifier model.

5.12.2.1. Scenario1: Models architectures and their performances

Results in table 5.23 shows the four selected models for the first scenario. Also, the training, testing and validation **"MSE"** and the training, testing and validation **"R"**.

Table 5.23: Summary of the architectures and the performance of the Scenario1 models

Models	**M1**	**M2**	**M3**	**M4**
Input layer	S1	S1	S1	S1
Training size	50 %	60 %	70 %	80 %
Testing size	25 %	20 %	15 %	10 %
Validation size	25 %	20 %	15 %	10 %
Structure	(9 : 7 :1)	(9 : 6 : 1)	(9 : 6 : 1)	(9 : 5 : 1)
Hidden layer	7	6	6	5
Training MSE	0.576	0.240	0.165	0.199
Validation MSE	1.186	0.401	0.199	0.539
Testing MSE	1.384	0.732	0.441	0.175
All MSE	**0.928**	**0.370**	**0.211**	**0.230**
Training R	0.99	0.99	0.99	0.99
Validation R	0.98	0.99	0.99	0.99
Testing R	0.98	0.99	0.99	0.99
All R	**0.98**	**0.99**	**0.99**	**0.99**

Results from the above table 5.23 shows that the four models have a good correlation coefficient (> 0.97) between WCP and the selected socio-economic parameters.

Figure 5.29 represents the architecture of model 1. The model has nine inputs, seven hidden layers and one output (WCP).

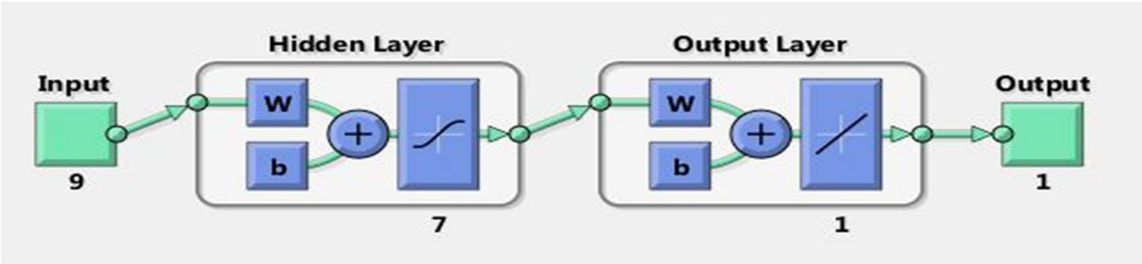


Fig 5.29 Neural network structure of model 1-scenario 1

Figure 5.30 demonstrates the Target-Output graphs of training, testing and validation phases with their correlation coefficients of the model 1-scenario 1.

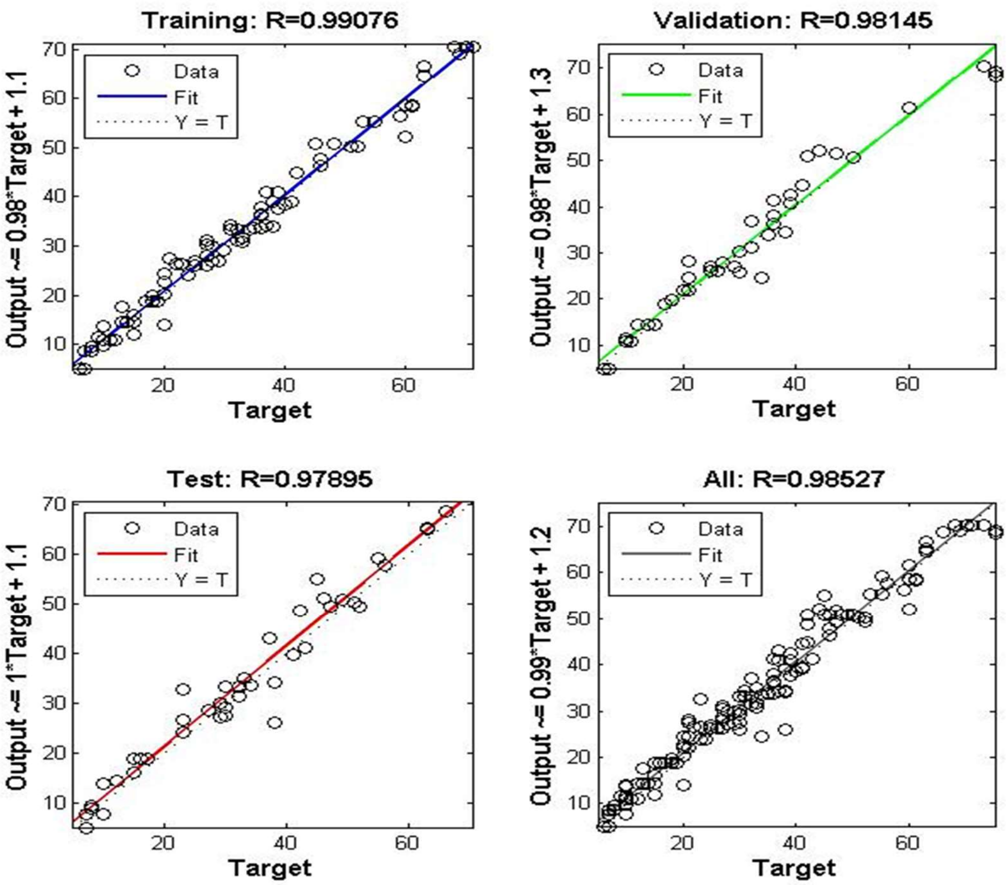


Fig 5.30 Neural network Target-Output graphs of training, testing and validation phases of model 1-scenario 1

Figure 5.31 shows the architecture of model 2. This model has nine inputs, 6 hidden layers and one output (WCP).

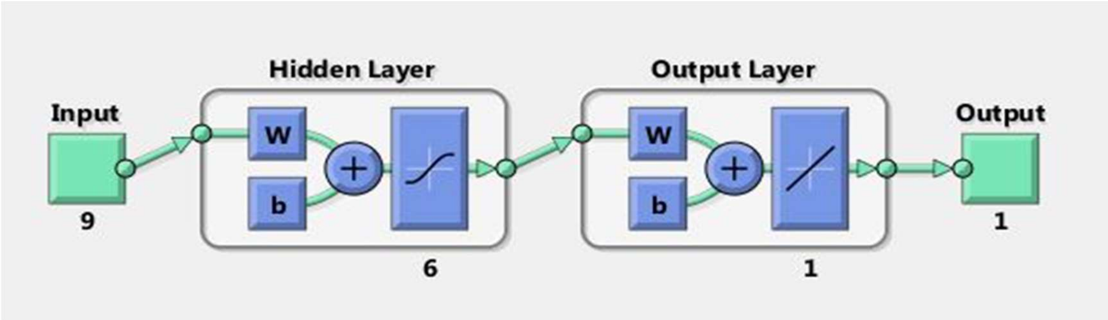


Fig 5.31 Neural network structure of model 2-scenario 1

Figure 5.32 represents the Target-Output graphs of training, testing and validation phases with their correlation coefficients of the model 2-scenario 1.

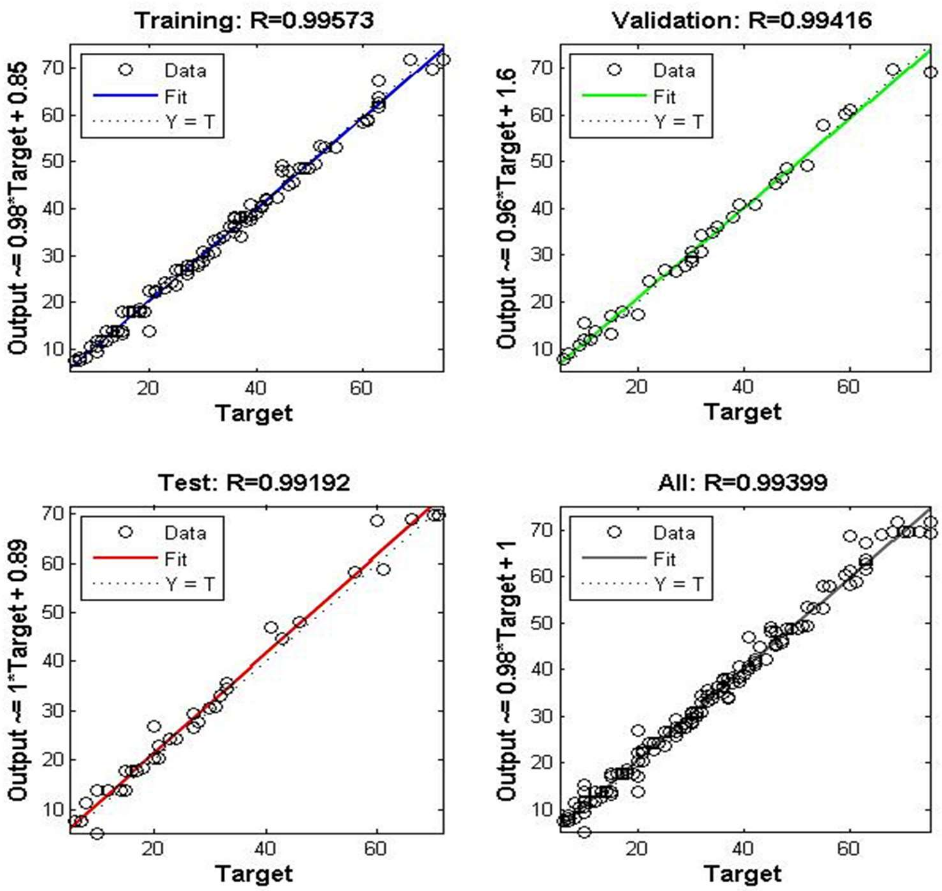


Fig 5.32 Neural network Target-Output graphs of training, testing and validation phases of model 2-scenario 1

Figure 5.33 shows the architecture of model 3. Model 3 has nine inputs, 6 hidden layers and one output (WCP).

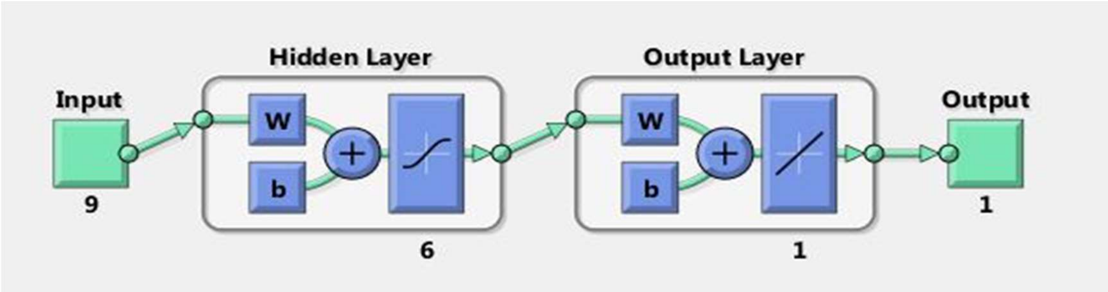


Fig 5.33 Neural network structure of model 3-scenario 1

Figure 5.34 shows the Target-Output graphs of training, testing and validation phases with their correlation coefficients of the model 3-scenario 1.

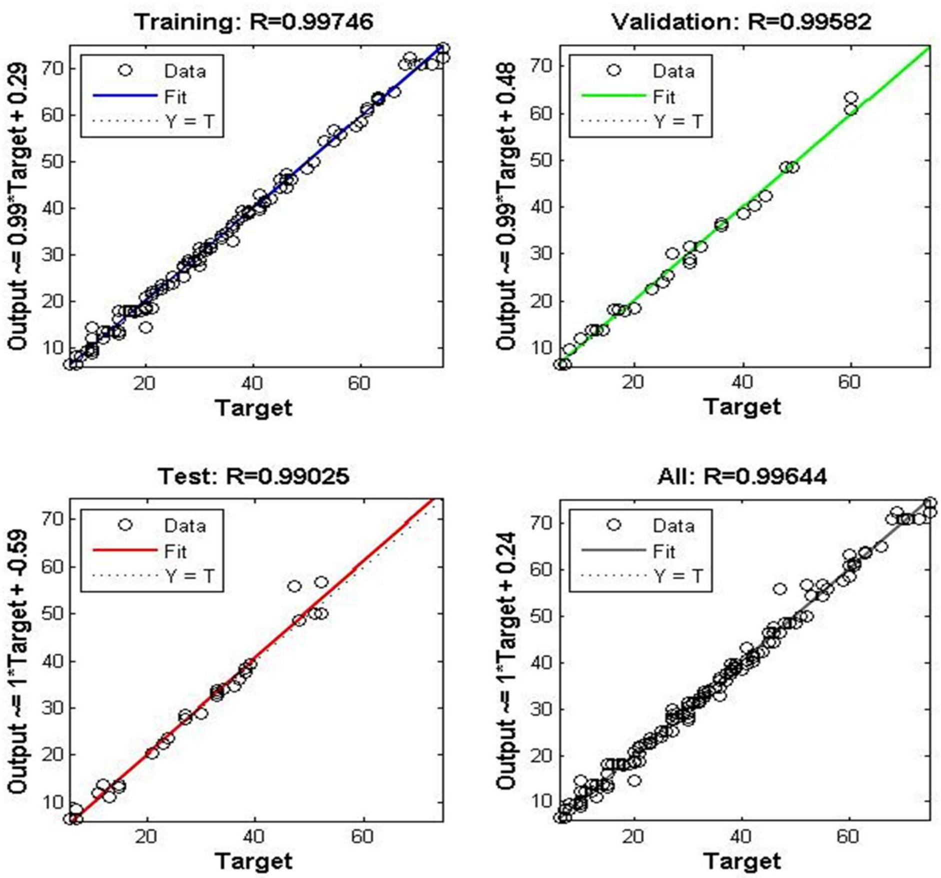


Fig 5.34 Neural network Target-Output graphs of training, testing and validation phases of model 3-scenario 1

Figure 5.35 represents the architecture of model 4. The last model in scenario 1 has nine inputs, 5 hidden layers and one output (WCP).

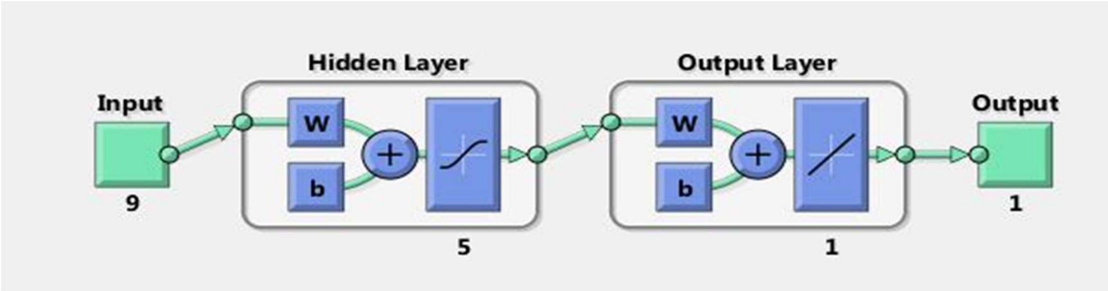


Fig 5.35 Neural network structure of model 4-scenario 1

Figure 5.36 illustrates the Target-Output graphs of training, testing and validation phases with their correlation coefficients of the model 4-scenario 1.

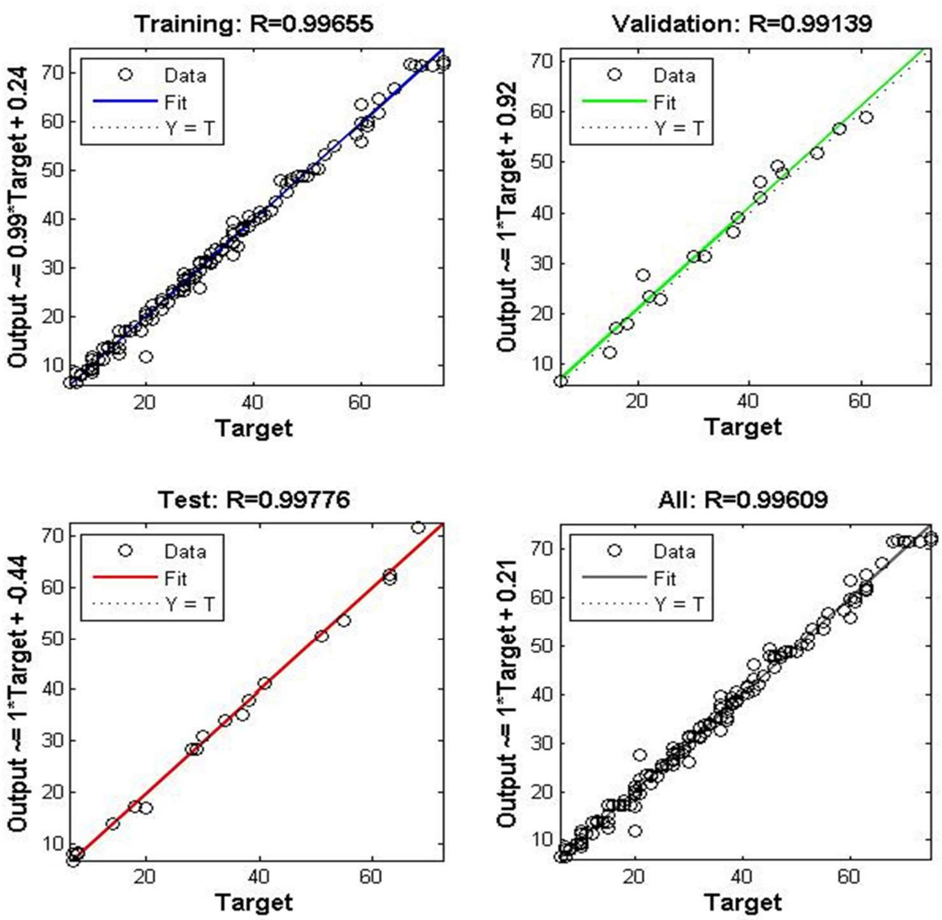


Fig 5.36 Neural network Target-Output graphs of training, testing and validation phases of model 4-scenario 1

5.12.2.2. Scenario2: Models architectures and their performances

Table 5.24 demonstrates the five selected models for the second scenario, besides to the training, testing and validation **"MSE"**, besides to the training, testing and validation **"R"**.

Table 5.24: Summary of the architectures and the performance of the scenario2 models

Models	**M5**	**M6**	**M7**	**M8**	**M9**
Input layer	S2	S2	S2	S2	S2
Training size	50 %	60 %	70 %	80 %	60 %
Testing size	25 %	20 %	15 %	10 %	30 %
Validation size	25 %	20 %	15 %	10 %	10 %
Structure	(3 4 1)	(3 2 1)	(3 2 1)	(3 2 1)	(3 4 1)
Hidden layer	4	2	2	2	4
Training MSE	1.526	1.755	1.541	1.389	1.155
Validation MSE	0.741	0.529	0.449	0.432	0.572
Testing MSE	0.562	0.249	0.576	0.443	0.164
All MSE	**0.912**	**1.211**	**1.234**	**1.200**	**0.882**
Training R	0.97	0.97	0.97	0.98	0.98
Validation R	0.98	0.99	0.99	0.99	0.99
Testing R	0.98	0.99	0.99	0.99	0.99
All R	**0.98**	**0.98**	**0.98**	**0.98**	**0.98**

Table 5.24 indicates that the five models (M5, M6, M7, M8 and M9) have a good correlation coefficient (> 0.97) between WCP and the selected physical characteristics of building units variables.

Figure 5.37 represents the architecture of model 5. The first model in scenario 2 has three inputs, 4 hidden layers and one output (WCP).

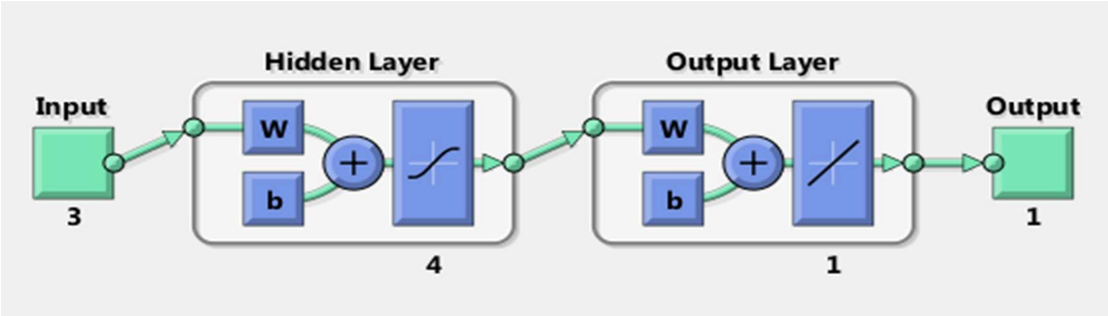


Fig 5.37 Neural network structure of model 5-scenario 2

Figure 5.38 illustrates the Target-Output graphs of training, testing and validation phases with their correlation coefficients of the model 5-scenario 2.

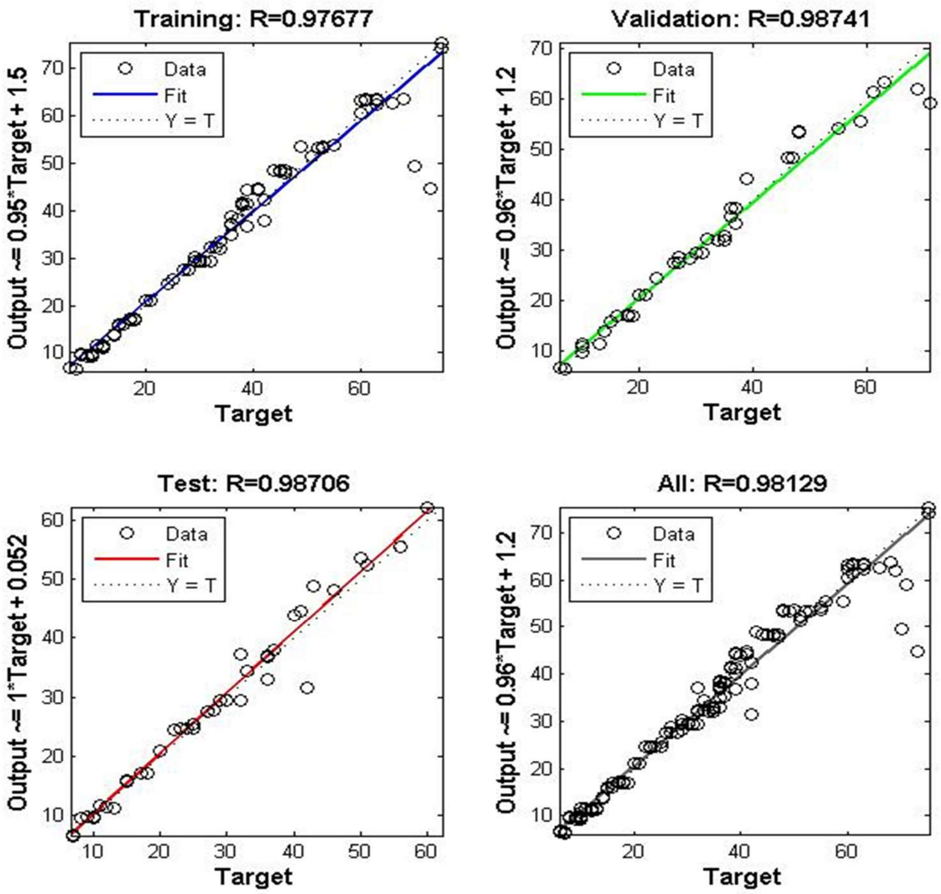


Fig 5.38 Neural network Target-Output graphs of training, testing and validation phases of model 5-scenario 2

Figure 5.39 illustrates the architecture of model 6 from scenario 2. It has three inputs, 2 hidden layers and one output (WCP).

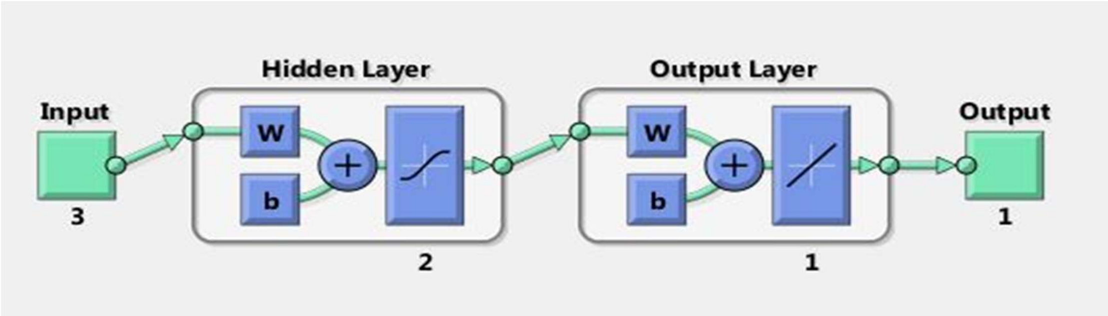


Fig 5.39 Neural network structure of model 6-scenario 2

The Target-Output graphs of training, testing and validation phases with their correlation coefficients of the model 6-scenario 2 are illustrated in figure 5.40 below.

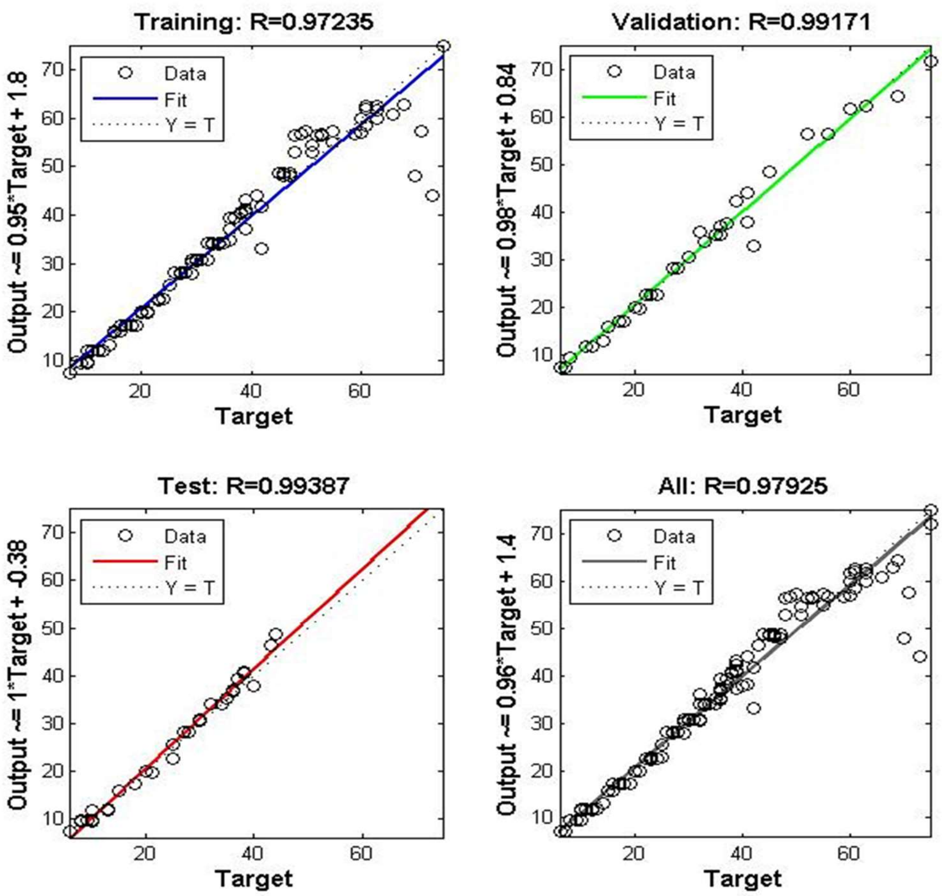


Fig 5.40 Neural network Target-Output graphs of training, testing and validation phases of model 6-scenario 2

Figure 5.41 demonstrates the architecture of model 7 from scenario 2, where this model has three inputs, 2 hidden layers and one output (WCP).

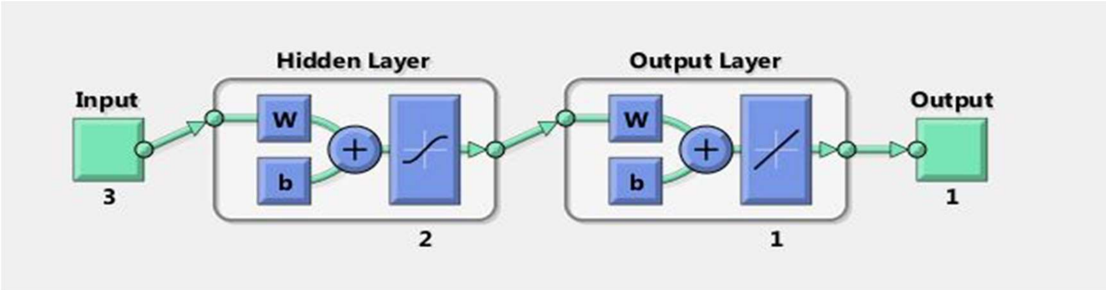


Fig 5.41 Neural network structure of model 7-scenario 2

Figure 5.42 shows the Target-Output graphs of training, testing and validation phases with their correlation coefficients of the model 7-scenario 2.

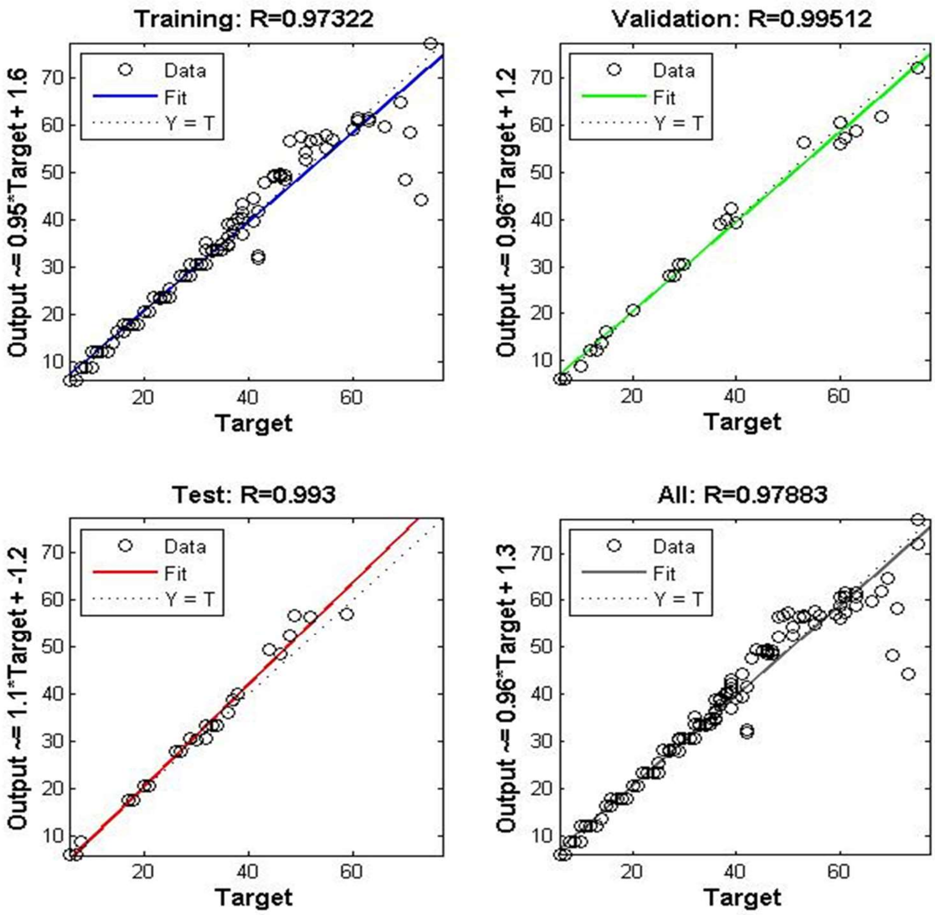


Fig 5.42 Neural network Target-Output graphs of training, testing and validation phases of model 7-scenario 2

The architecture of model 8-scenario 2 is given in figure 5.43 below. Model 8 has three inputs, 2 hidden layers and one output (WCP).

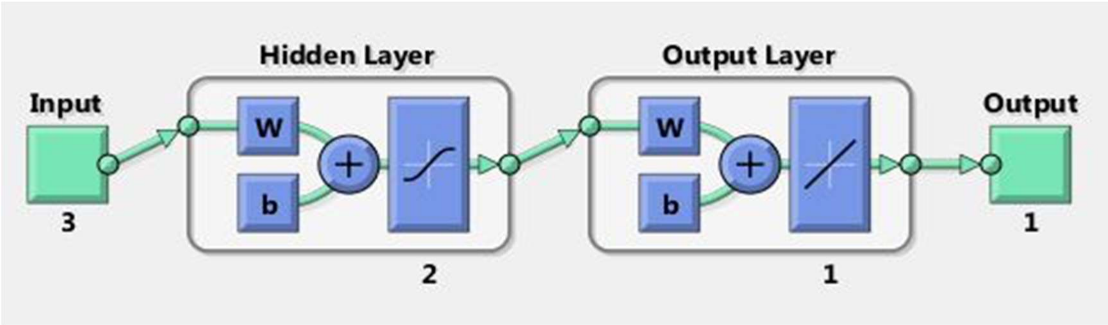


Fig 5.43 Neural network structure of model 8-scenario 2

Figure 5.44 illustrates the Target-Output graphs of training, testing and validation phases with their correlation coefficients of the model 8-scenario 2.

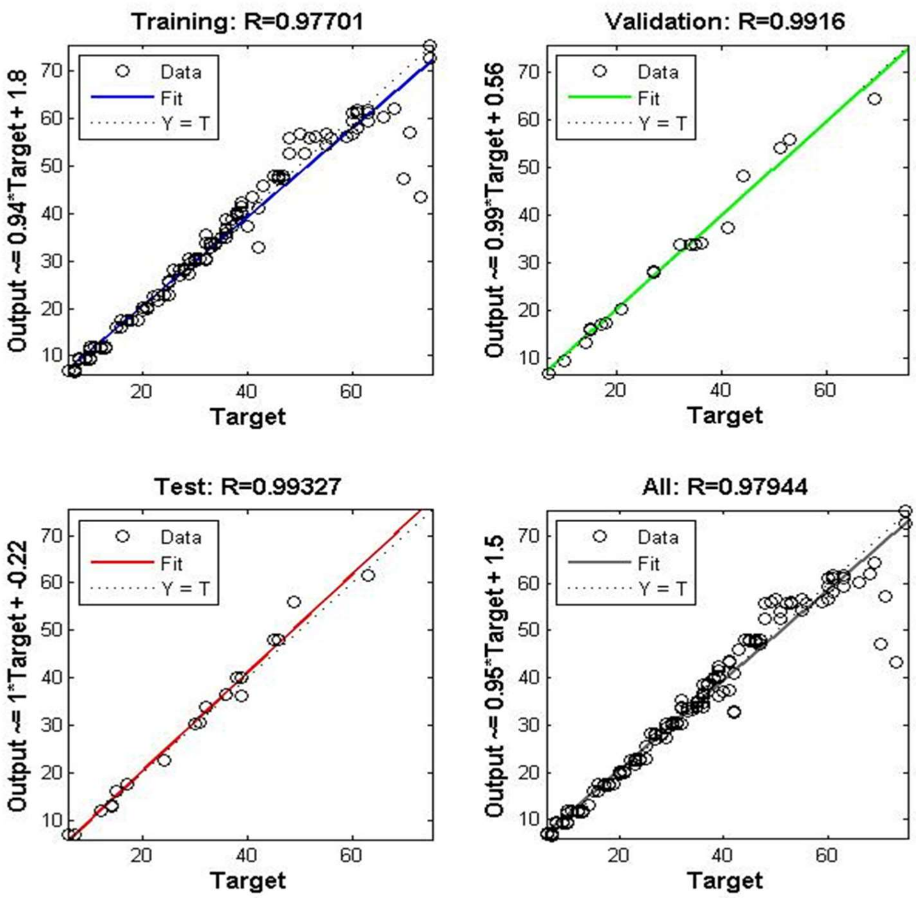


Fig 5.44 Neural network Target-Output graphs of training, testing and validation phases of model 8-scenario 2

Figure 5.45 shows the architecture of model 9 from scenario 2, where this last has three inputs, 4 hidden layers and one output (WCP).

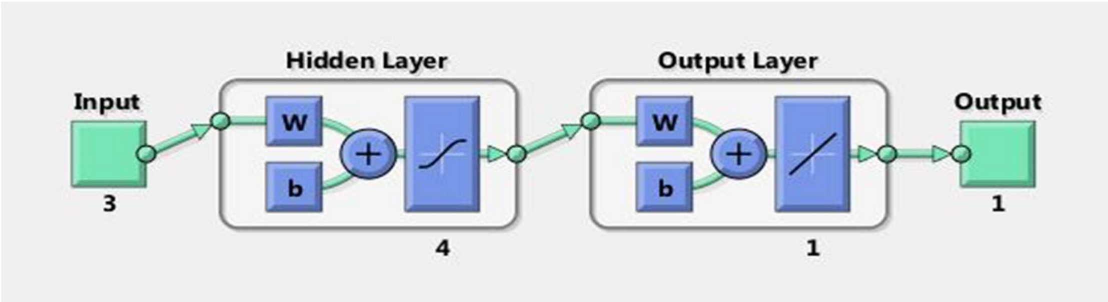


Fig 5.45 Neural network structure of model 4-scenario 1

The Target-Output graphs of training, testing and validation phases with their correlation coefficients of the model 9-scenario 2 are illustrated in figure 5.46.

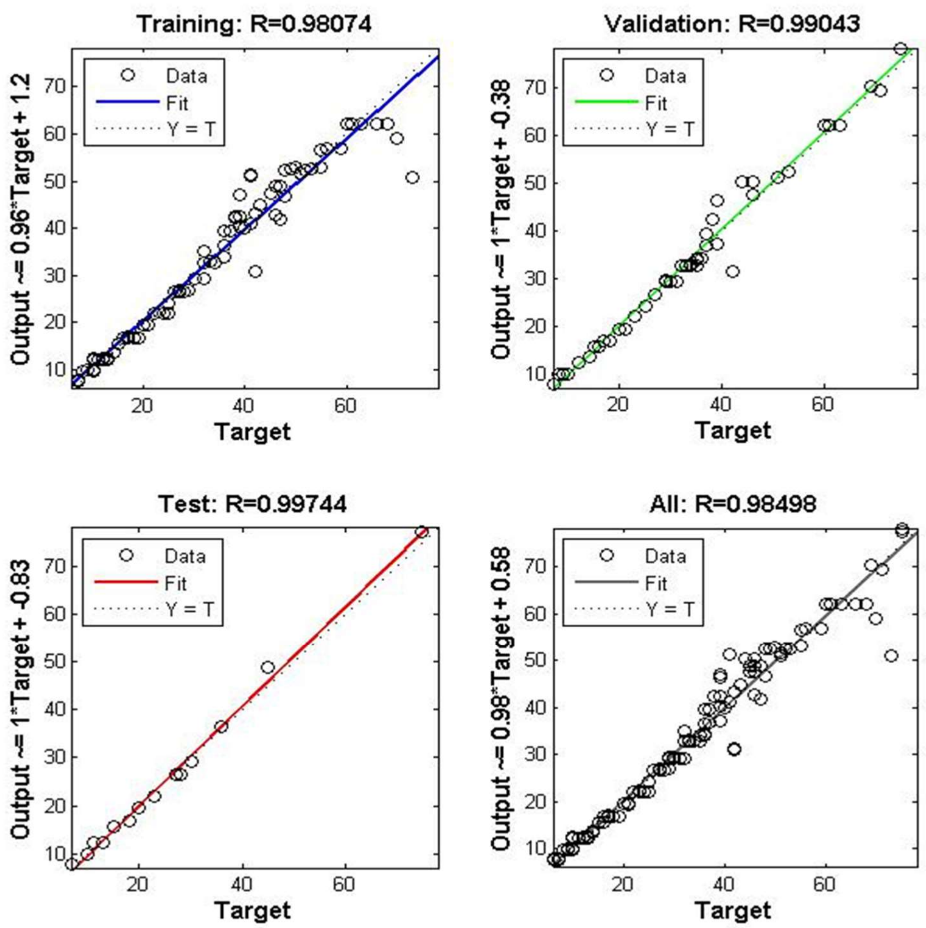


Fig 5.46 Neural network Target-Output graphs of training, testing and validation phases of model 9-scenario 2

5.12.2.3. Scenario3: Models architectures and their performances

The six selected models for the third scenario, the training, testing and validation **"MSE"** and the training, testing and validation **"R"** are demonstrated in table 5.25 below.

Table 5.25: Summary of the architectures and the performance of the scenario3 models

Models	**M10**	**M11**	**M12**	**M13**	**M14**	**M15**
Input layer	S1+S2	S1+S2	S1+S2	S1+S2	S1+S2	S1+S2
Training size	50 %	60 %	70 %	80 %	80 %	60 %
Testing size	25 %	20 %	15 %	10 %	10 %	25 %
Validation size	25 %	20 %	15 %	10 %	10 %	15 %
Structure	(12 10 1)	(12 9 1)	(12 6 1)	(12 6 1)	(12 4 1)	(12 7 1)
Hidden layer	10	9	6	6	4	7
Training MSE	0.235	0.713	0.117	0.322	0.137	0.027
Validation MSE	0.365	0.376	0.303	0.417	0.184	0.301
Testing MSE	0.752	0.191	0.189	0.621	0.134	0.189
All MSE	**0.396**	**0.156**	**0.156**	**0.091**	**0.141**	**0.119**
Training R	0.99	0.99	0.99	0.99	0.99	0.99
Validation R	0.99	0.99	0.99	0.99	0.99	0.99
Testing R	0.99	0.99	0.99	0.99	0.99	0.99
All R	**0.99**	**0.99**	**0.99**	**0.99**	**0.99**	**0.99**

Table 5.25 indicates that the six models (M10, M11, M12, M13, M14 and M15), where the inputs are the combination between the socio-economic parameters and the physical characteristics of housing units have a good correlation coefficient (equal to 0.9) between them and WCP.

Figure 5.47 shows the architecture of model 10 from scenario 3, where this last has twelve inputs, 10 hidden layers and one output (WCP).

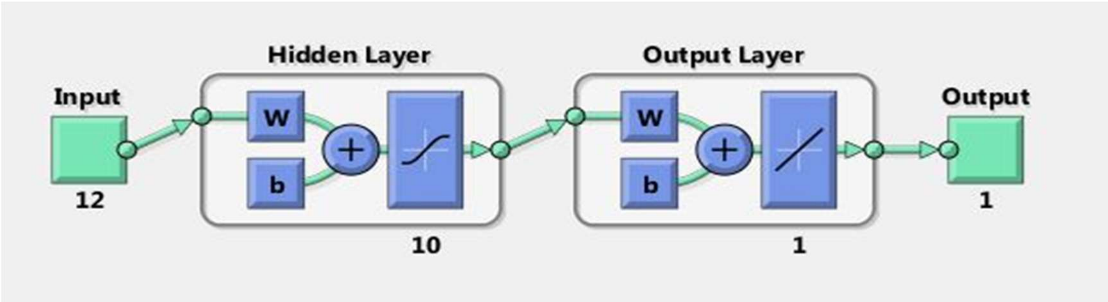


Fig 5.47 Neural network structure of model 10-scenario 3

The Target-Output graphs of training, testing and validation phases with their correlation coefficients of the model 10-scenario3 are illustrated in figure 5.48.

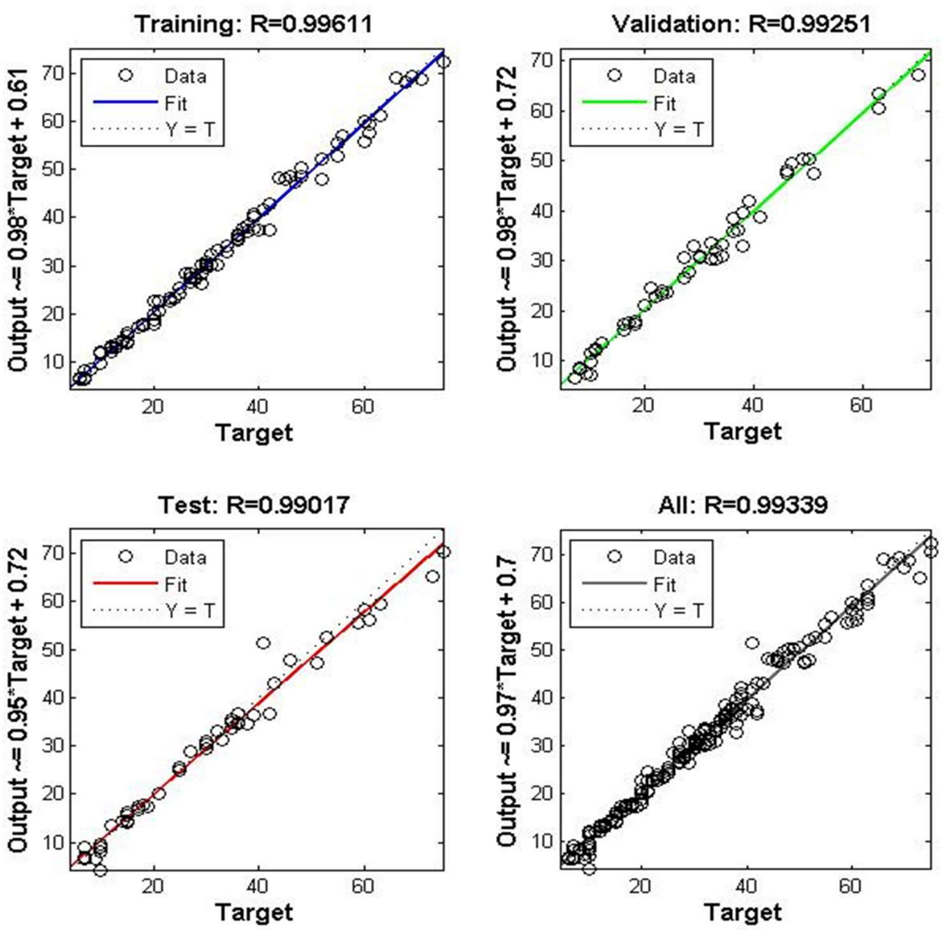


Fig 5.48 Neural network Target-Output graphs of training, testing and validation phases of model 10-scenario 3

Figure 5.49 demonstrates the architecture of model 11 from scenario 3. Model 11has twelve inputs, 9 hidden layers and one output (WCP).

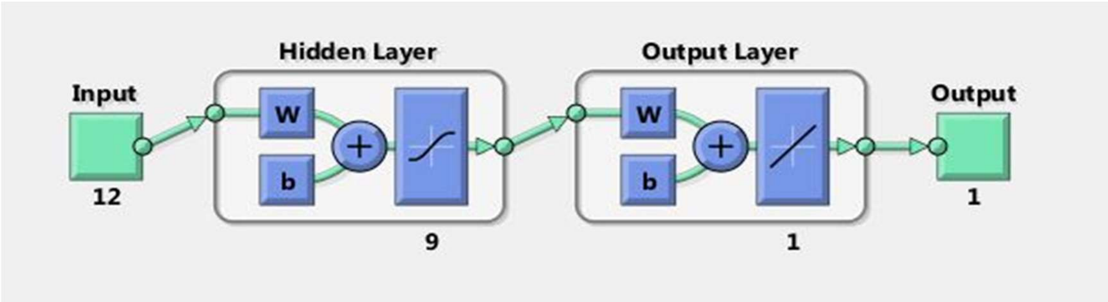


Fig 5.49 Neural network structure of model 4-scenario 3

Figure 5.50 shows the Target-Output graphs of training, testing and validation phases with their correlation coefficients of the model 11-scenario3.

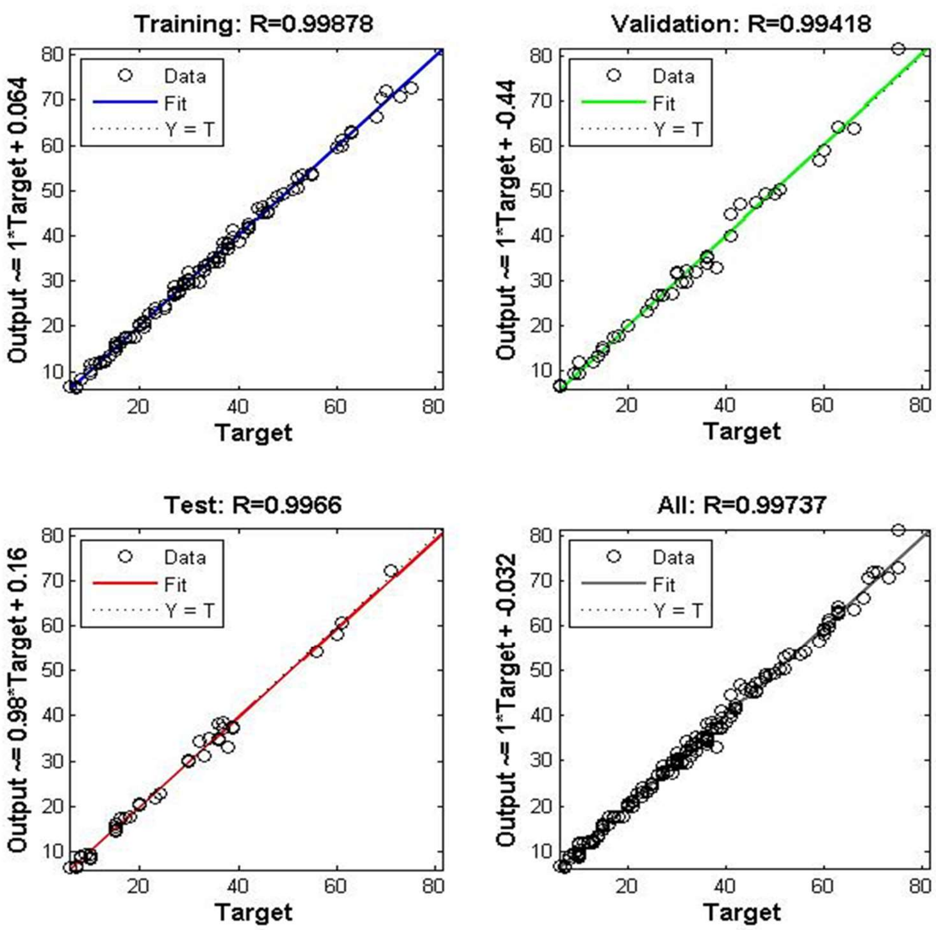


Fig 5.50 Neural network Target-Output graphs of training, testing and validation phases of model 11-scenario 3

The architecture of model 12 from scenario 3 is illustrated in figure 5.51 below, where it has twelve inputs, 6 hidden layers and one output (WCP).

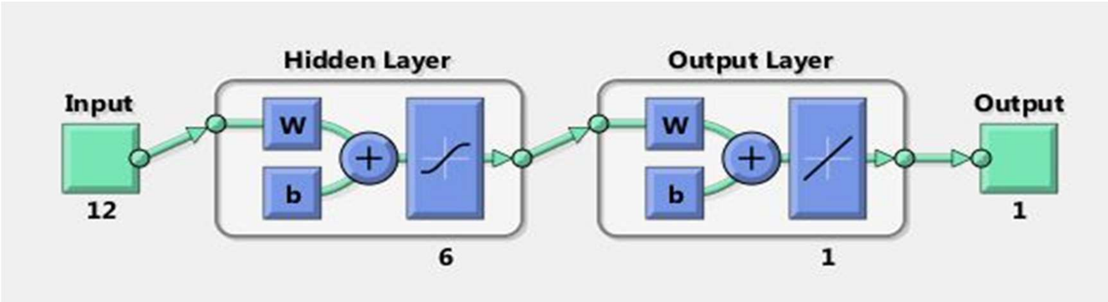


Fig 5.51 Neural network structure of model 12-scenario 3

The following graphs in figure 5.52 represents Target-Output graphs of training, testing and validation phases with their correlation coefficients of the model 12-scenario3.

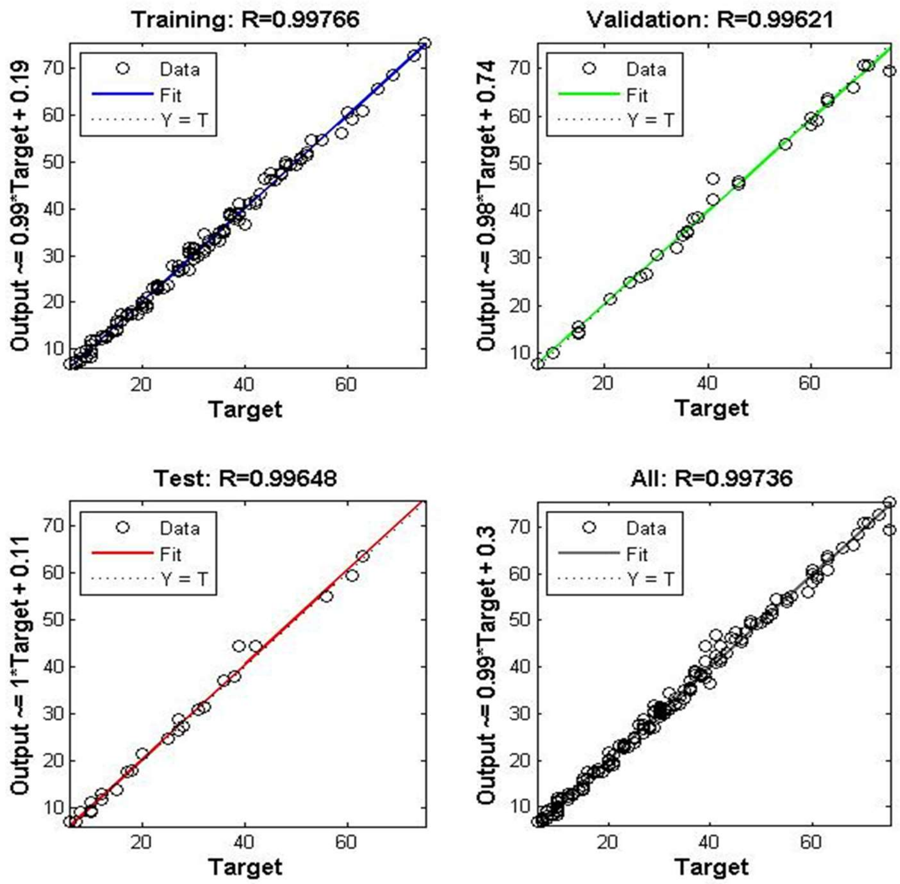


Fig 5.52 Neural network Target-Output graphs of training, testing and validation phases of model 12-scenario 3

Figure 5.53 illustrates the architecture of model 13 from scenario 3, where this last has twelve inputs, 6 hidden layers and one output (WCP).

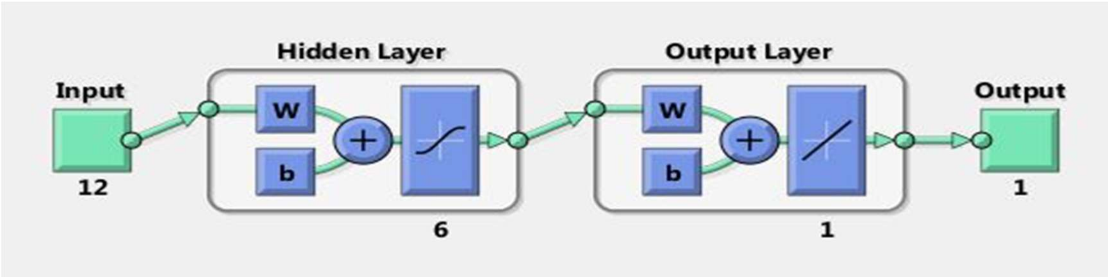


Fig 5.53 Neural network structure of model 4-scenario 3

The Target-Output graphs of training, testing and validation phases with their correlation coefficients of the model 13-scenario3 are illustrated in figure 5.54.

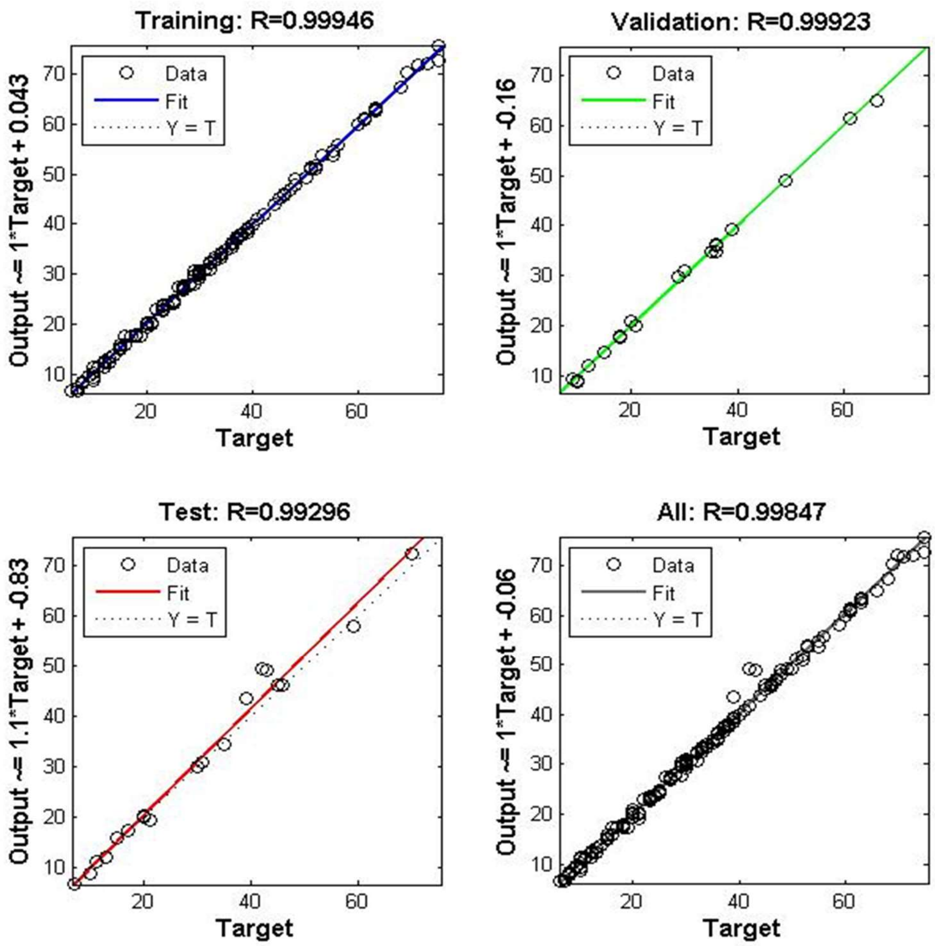


Fig 5.54 Neural network Target-Output graphs of training, testing and validation phases of model 13-scenario 3

Figure 5.55 shows the architecture of model 14 from scenario 3. It has twelve inputs, 4 hidden layers and one output (WCP).

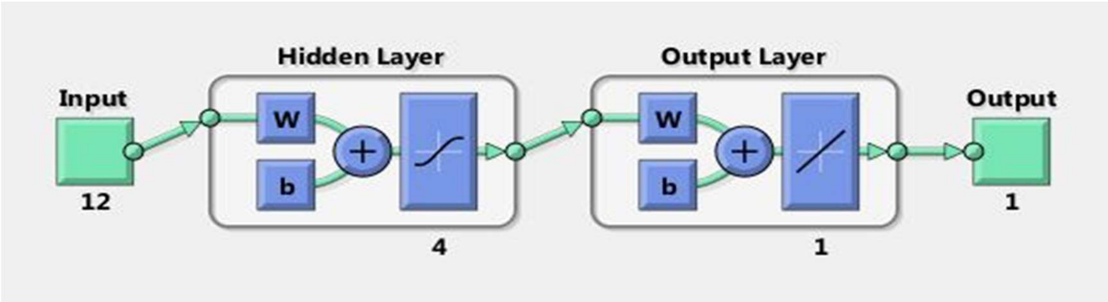


Fig 5.55 Neural network structure of model 14-scenario 3

The Target-Output graphs of training, testing and validation phases with their correlation coefficients of the model 14-scenario3 are illustrated in figure 5.56.

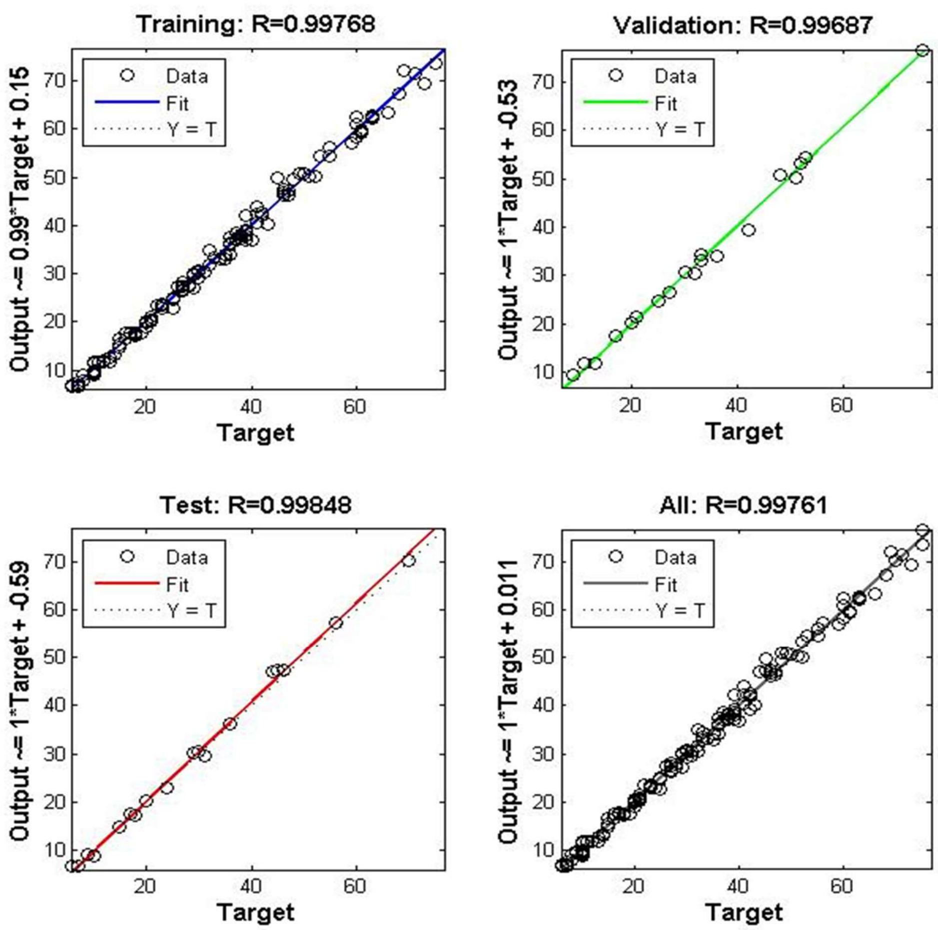


Fig 5.56 Neural network Target-Output graphs of training, testing and validation phases of model 14-scenario 3

Figure 5.57 shows the architecture of model 15 from scenario 3, where this last has twelve inputs, 7 hidden layers and one output (WCP).

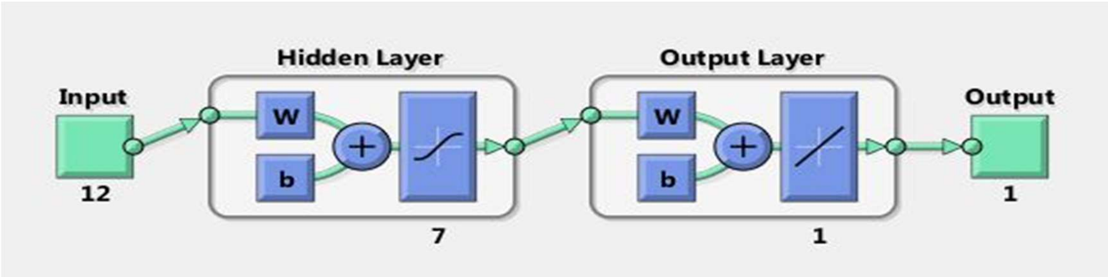


Fig 5.57 Neural network structure of model 15-scenario 3

The Target-Output graphs of training, testing and validation phases with their correlation coefficients of the model 15-scenario3 are illustrated in figure 5.58.

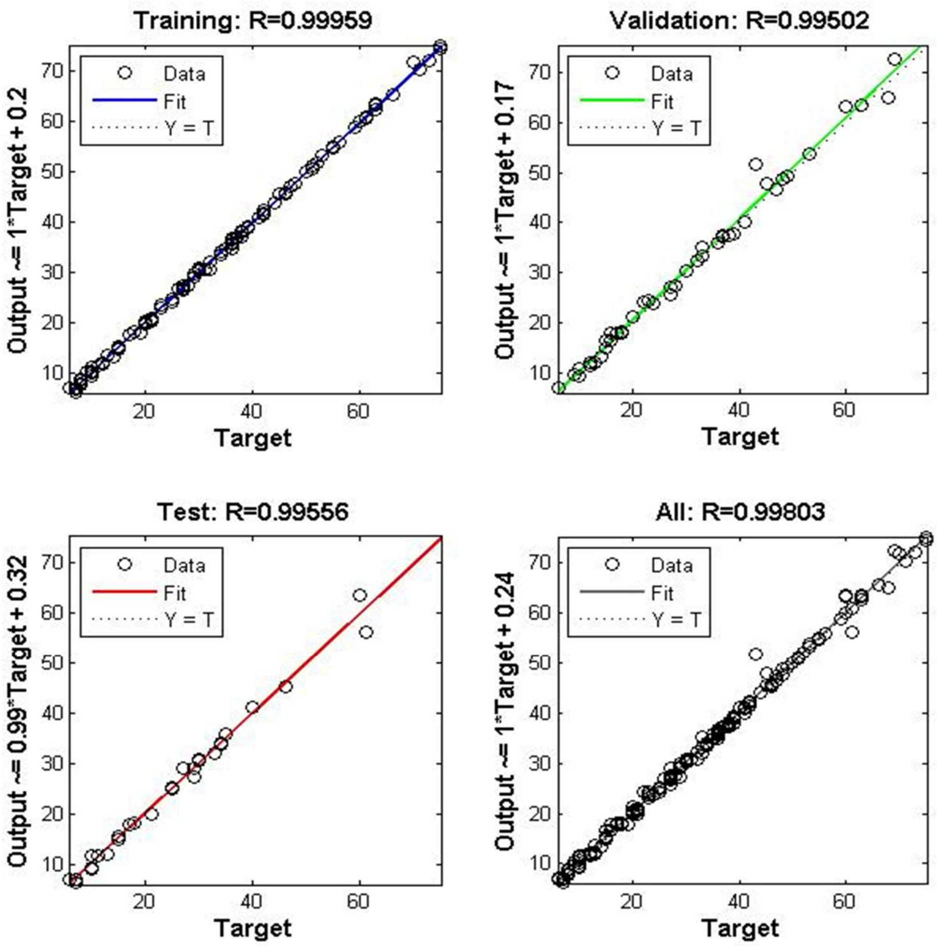


Fig 5.58 Neural network Target-Output graphs of training, testing and validation phases of model 15-scenario 3

5.12.3. ANNs predicted models and their performances

In order to test the accuracy and predictive ability of the 15 ANNs models for each household, actual data set (observed) and its predicted values are plotted and displayed in figures 5.59, 5.60 and 5.61 below. It can be seen that the predicted values are well fitted with the original data.

Moreover, and by combining the two groups of variables (socio-economic parameters and physical characteristics of building units), the neural models would gain more information than it would be separately. A possible explanation is that the inter correlation between variables.
In addition, the correlation coefficients are equal to 0.99 in the training, testing and validation phases for the 6 last models.

In general, the combination of variables improves significantly the model performance, that is why the choice of input variables is utmost important. Other studies aiming to model water consumption but with different variables agrees with the obtained results Bougadis et al., 2005, Al-Zahrani and Abo-monasar,, 2015 and El Masri et al., 2016).

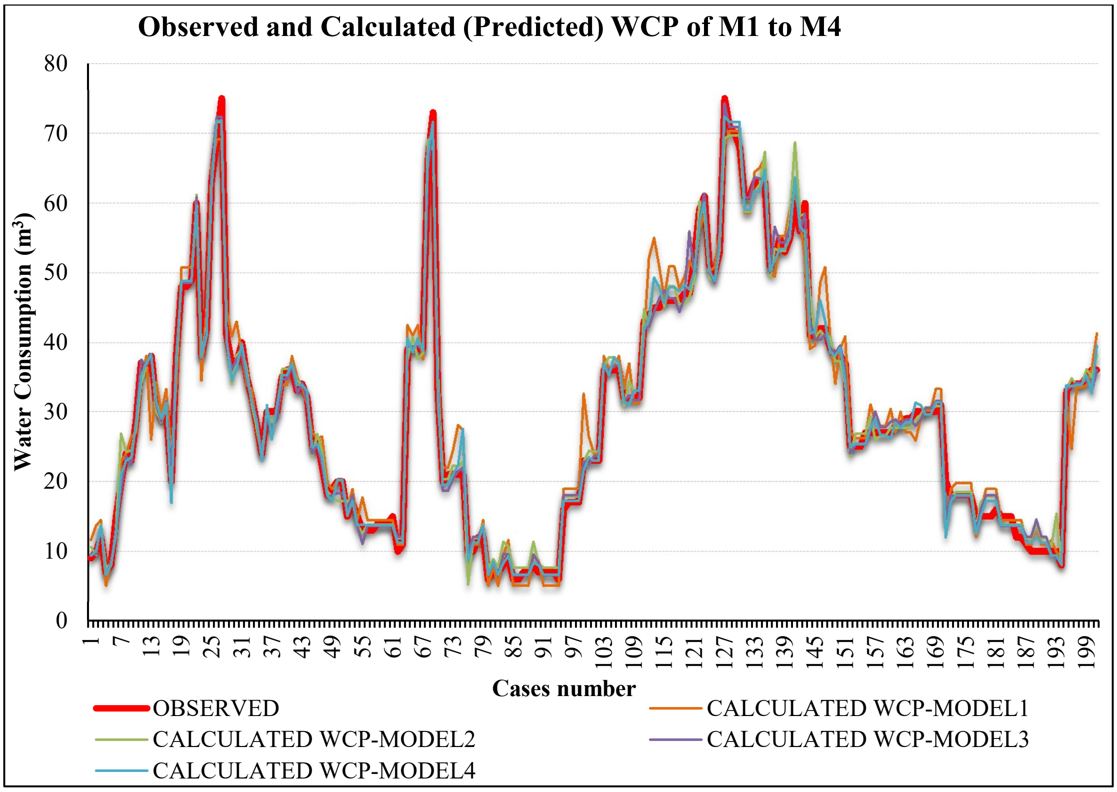


Fig 5.59 Comparison between Observed and Calculated (Predicted) WCP of the models M1, M2, M3 and M4

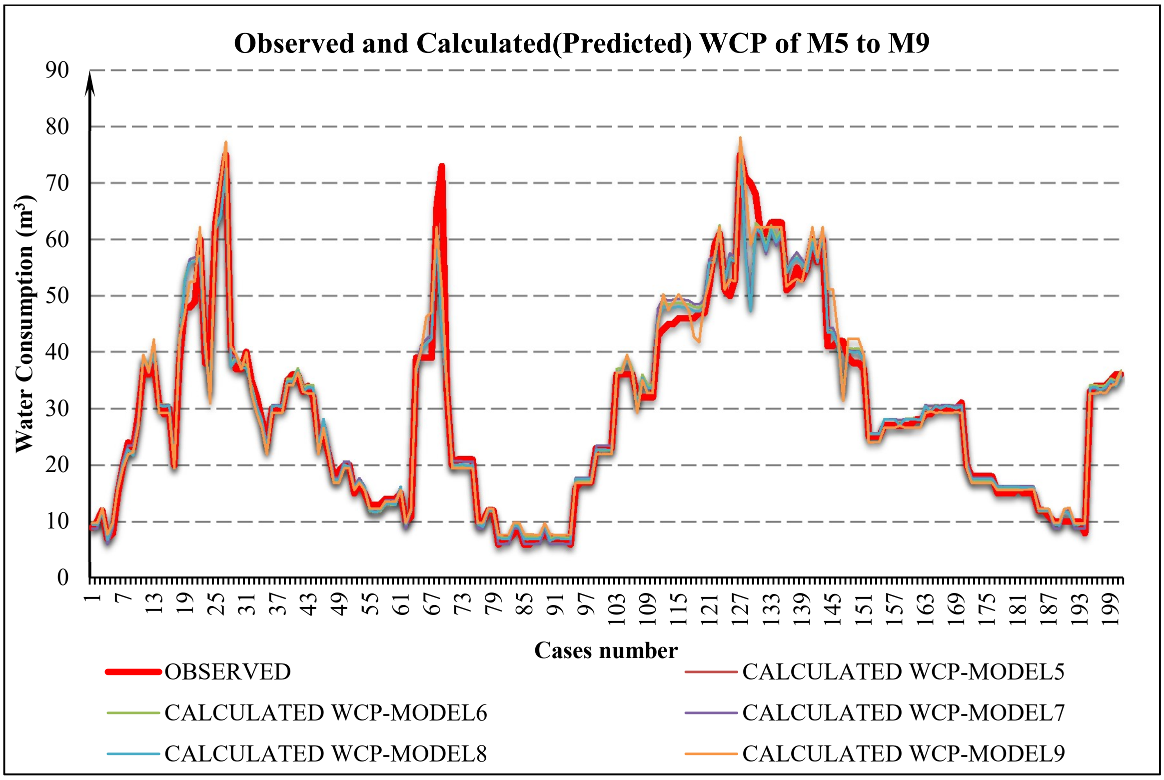


Fig 5.60 Comparison between Observed and Calculated (Predicted) WCP of the models M5, M6, M7, M8 and M9

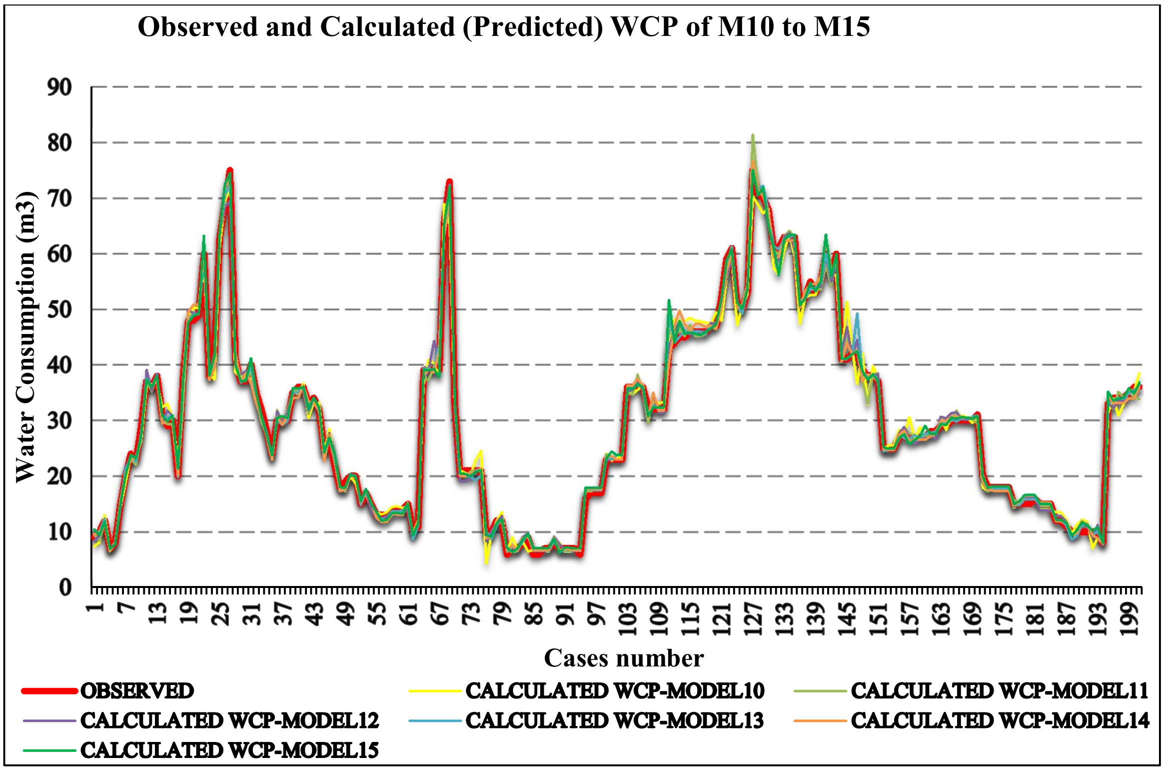


Fig 5.61 Comparison between Observed and Calculated (Predicted) WCP of the models M10, M11, M12, M13, M14 and M15

Basing on the correlation coefficient, 15 models on total are selected with values greater than 0.95 for all the phases and more efficient in forecasting water consumption of Sedrata city.

The Mean Square Error (MSE) is an excellent indicator of models performance and lower values are preferable especially during the training phase.
Values of MSE for the three scenarios are illustrated in (figures 5.62, 5.63 and 5.64). In training phase, the MSE coefficient of the last models (M10, M11, M12, M13, M14 and M15) are the smallest with 0.235, 0.713, 0.117, 0.322, 0.137 and 0.027 respectively. Furthermore, the MSE values for the models M1, M2, M3 and M4 are low too (0.576, 0.240, 0.165 and 0.199, respectively) compared with the MSE values of M5, M6, M7, M8 and M9 (1.526, 1.755, 1.541, 1.389 and 1.155, respectively). Taking these results into account, the last models (M10 to M15) appear to be the most efficient for forecasting domestic water consumption in study zone. In testing phase, the correlation coefficient values of the best models (M10 to M15) are similar and equal to 0.99. The MSE values obtained during the test phase also allow the rest models to be compared (Tables 5.23, 5.24 and 5.25).

Summing up the results, the models M10-M15 represent the best performing models. The main reason behind is the inputs and neural structure, where the best models have higher number of inputs.

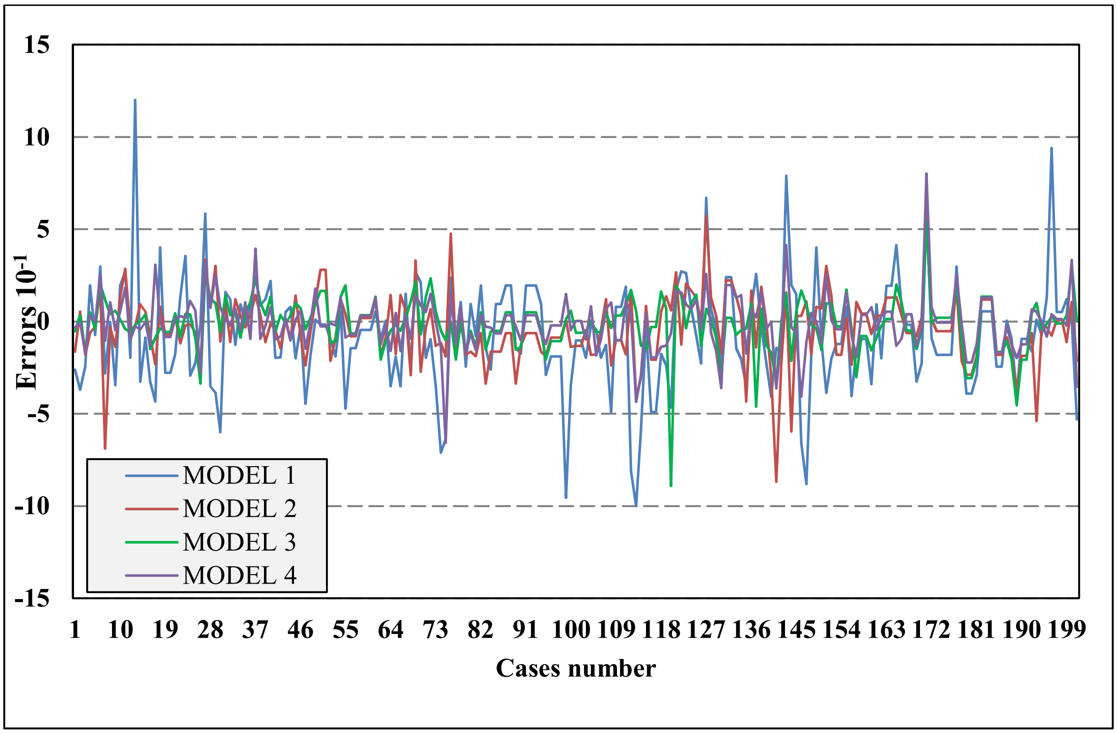


Fig 5.62 errors values of the models M1, M2, M3 and M4

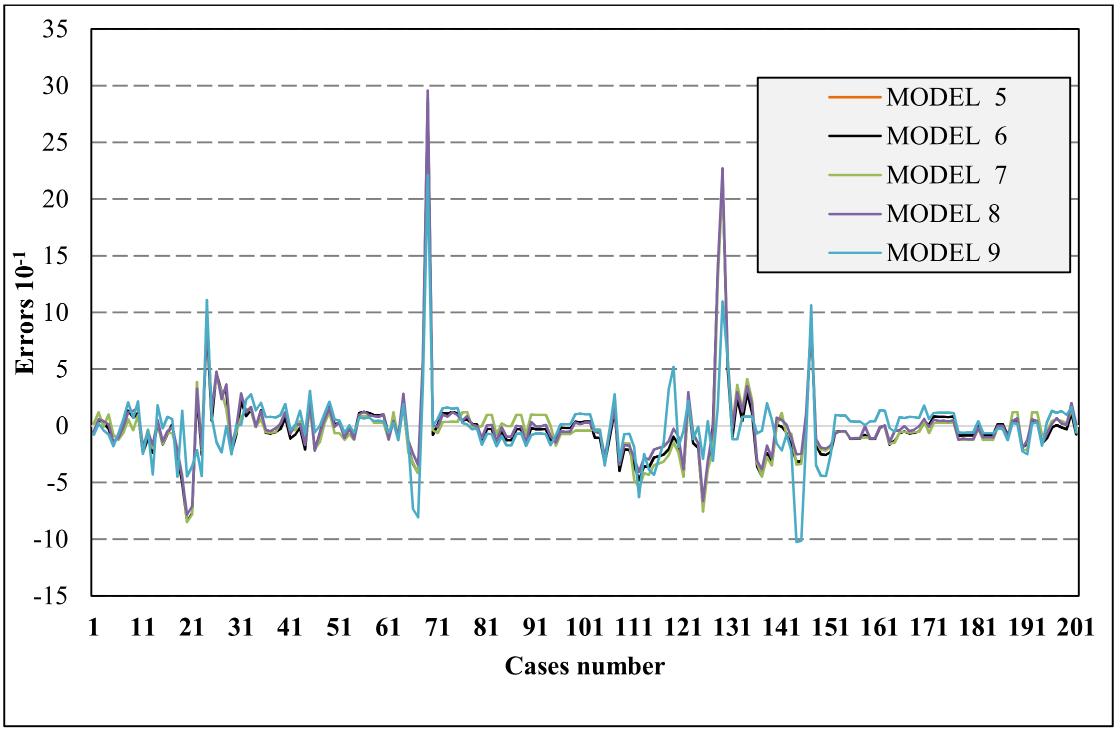


Fig 5.63 Errors values of the models M5, M6, M7, M8 and M9

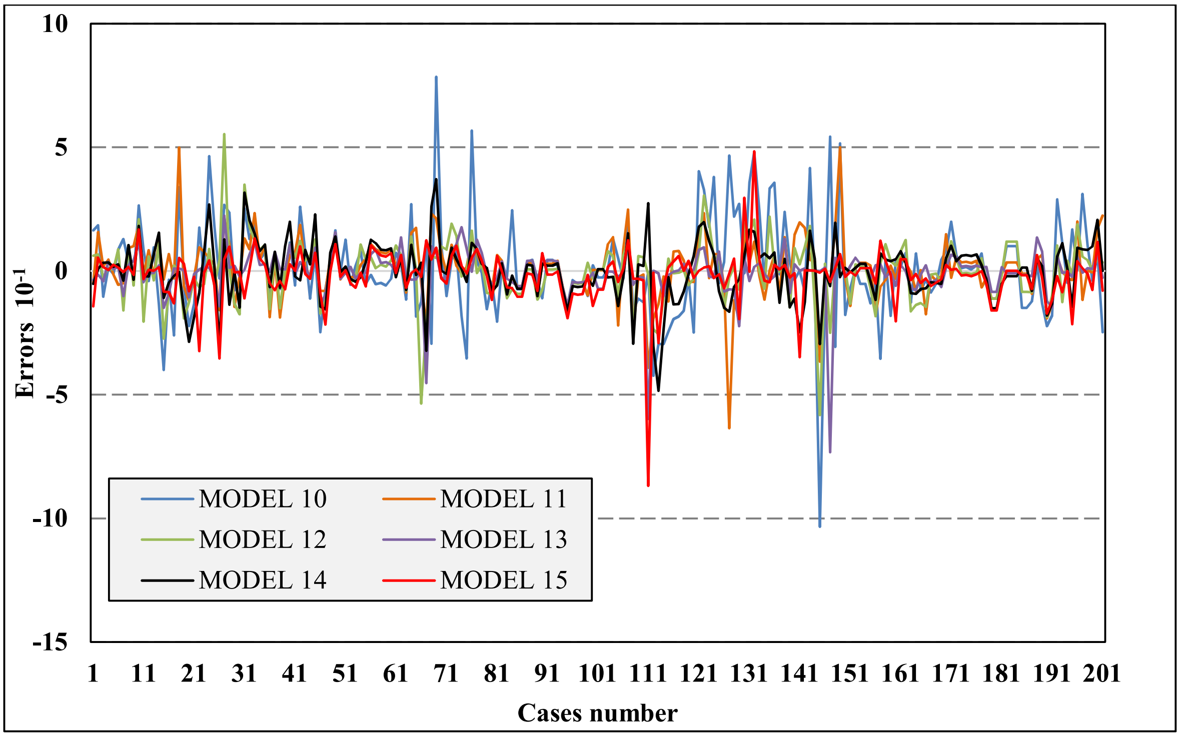


Fig 5.64 errors values the errors of the models M10, M11, M12, M13, M14 and M15

5.13. Adaptive Neuro Fuzzy Inference System (ANFIS)

Majority of research papers dealing with WCP forecasting use ANNs, while the use of Adaptive Neuro Fuzzy Inference System (ANFIS) to treat such problems is relatively new. In this part of the thesis, ANFIS is adopted to build general rules on water consumption behavior and patterns in a simplified manner. Following from that, the obtained results helped to build a forecasting model.

5.13.1. Input selection and models construction

For modeling with ANFIS, input combination of 8 scenarios are selected to cover all possible combinations. The reason behind this selection is the limitation of the model (number of inputs). S1, S2, S3 and S4 were conducted in order to evaluate the effect of socio-economic parameters on WCP. On the other hand, the selection of S5 aims to determine the effect of physical characteristics of building units on WCP. As it can be inferred S6, S7 and S8 are the input combination that mixed socio-economic parameters with physical characteristics of building units in order to assess the effect of them on WCP. The selected models and their inputs are represented in table 5.26.

Table 5.26: Classification of neuro fuzzy models for inputs data

Scenarios	HOUS	FEM	INC	AG1	AG3	CARN	PRS	HGS	UNIV	TAR	BAR	ROMN
S1	**X**	**X**	**X**									
S2				**X**	**X**	**X**						
S3							**X**	**X**	**X**			
S4	**X**					**X**			**X**			
S5										**X**	**X**	**X**
S6		**X**	**X**								**X**	
S7	**X**		**X**									**X**
S8					**X**			**X**			**X**	

5.13.2. Training, Testing and Checking of ANFIS models

The input data of ANFIS were divided into three sections : *training, testing* and *checking sections* in the same manner similar to ANNs (for training data size: 50%, 60%, 70% and 80%, for testing data size: 10%, 15%, 20% and 25%, and for validating data size 10%, 15%, 20% and 25%). The selected sections are shown in table 5.27.

ANFIS of MATLAB Fuzzy Logic is not designed to create models in an automated way; rather the trial-and-error approach was used to find the best model. In each time the ANFIS parameters were changed manually, besides the choice of ANFIS structure type, membership function types, training algorithm alternatives of ANFIS were limited compared to ANNs.

Table 5.27: classification of neuro fuzzy models for inputs data

Scenarios	**Training inputs**	**Testing inputs**	**Checking inputs**
S1	50%	25%	25%
S2	60%	20%	20%
S3	70%	15%	15%
S4	80%	10%	10%
S5	50%	25%	25%
S6	70%	15%	15%
S7	60%	20%	20%
S8	50%	25%	25%

To construct fuzzy models the fuzzy toolbox of MATLAB (FIS) is used. ANFIS is able to construct models with both subtractive clustering and grid partition categories. The toolbox grid partition approach is used for model construction and producing fuzzy rules. The topology of the ANFIS model with grid partition including 3 inputs for the eight scenarios S1-S8, where every input has 3 input-membership function, as shown in figure 5.65.

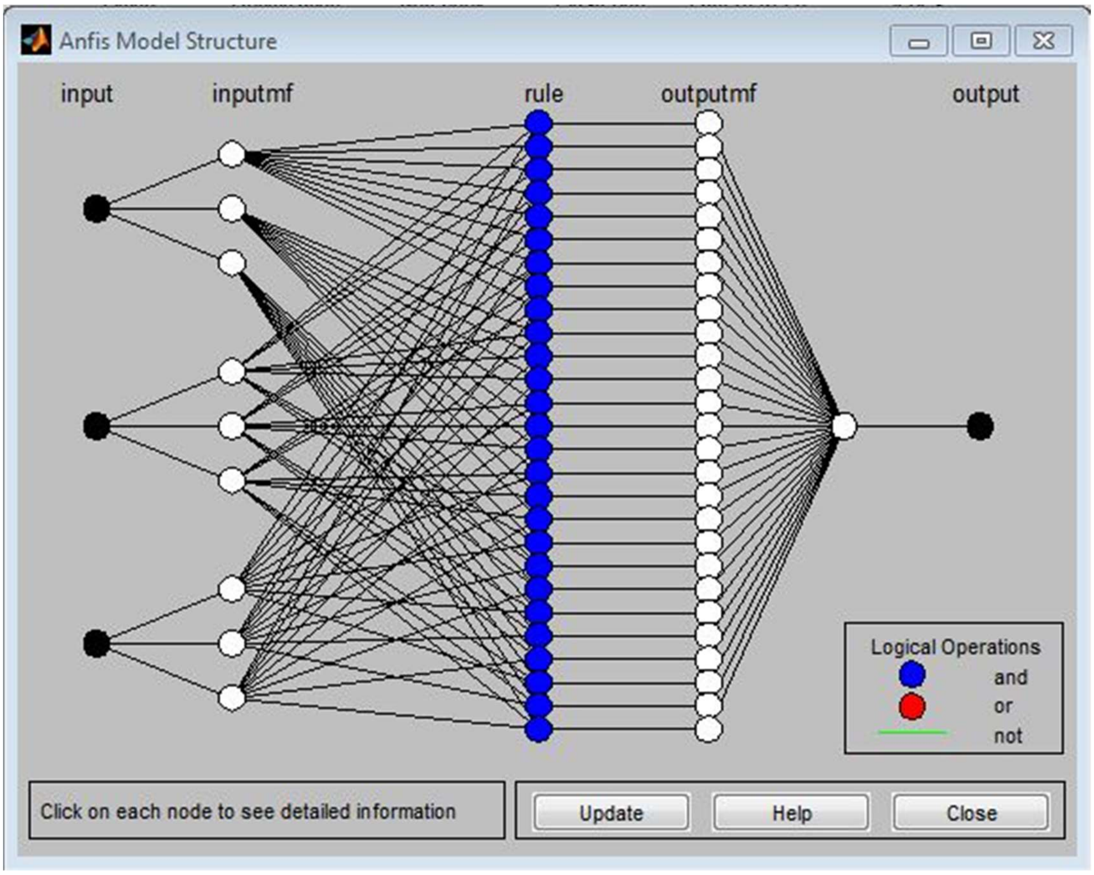


Fig 5.65 Topology of fuzzy and neuro-fuzzy models with grid partition (3 inputs)

The model can be described as follows:

- Layer 1 (input layer): No computation is considered in this layer. Each node, which corresponds to one input variable, only transmits input values to the next layer directly. For example, there are 3 input variables in scenario 4 that are household size (HOUS), university level of residents (UNIV) and car numbers (CARN).
- Layer 2 (fuzzification layer): each node in this layer corresponds one linguistic label to one of the input variables in layer 2. In other words, the output link represents the membership value that specifies the degree to which an input value belongs to a fuzzy set. In the subtractive grid category, a bell-shaped Gaussian function is used for the membership function in the following for (Jang, 1993):

$$\mu_{i,j}\left(X_j\right) = \exp\left(-\frac{\left(X_j - c_{ij}\right)^2}{2\sigma_{ij}^{\ 2}}\right) j = 1,2 \qquad (5.6)$$

Where, $\mu_{i,j}$ is the membership function or the degree of the membership of variable j in the *i-th* fuzzy implication rule; $\mu_{\ i,j}$ (Xj) indicates membership degree of the i-th input variable (X_j); c_{ij} and σ_{ij} represent the center and the half width of the membership function for the *j-th* variable and the *i-th* fuzzy implication rule, respectively.

- Layer 3 (rule antecedent): a node represents the antecedent part of a rule. Normally a T-norm operator is used in this node. Output of the layer 3 represents the firing strength of the corresponding fuzzy rule.
- Layer 4 (combination and defuzzification layer): the single node computes the overall output as a summation of all the incoming signals. The output of N fuzzy implication rules is obtained as follows for grid partition and subtractive clustering:

$$S = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} s_{ij} w_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}} \qquad (5.7)$$

In which w_i is the firing strength of the *i-th* fuzzy implication rule, n is the number of the clusters, S_i is the daily water demand estimation value from the *i-th* fuzzy implication rule and S is the estimated daily water demand (Jang, 1993).

Train and validation data are used for constructing neuro fuzzy models, while checking data controls the training error.

Figure 5.66 shows the training, testing and checking data with FIS output and training error for neuro fuzzy in scenario1 with 300 epochs.

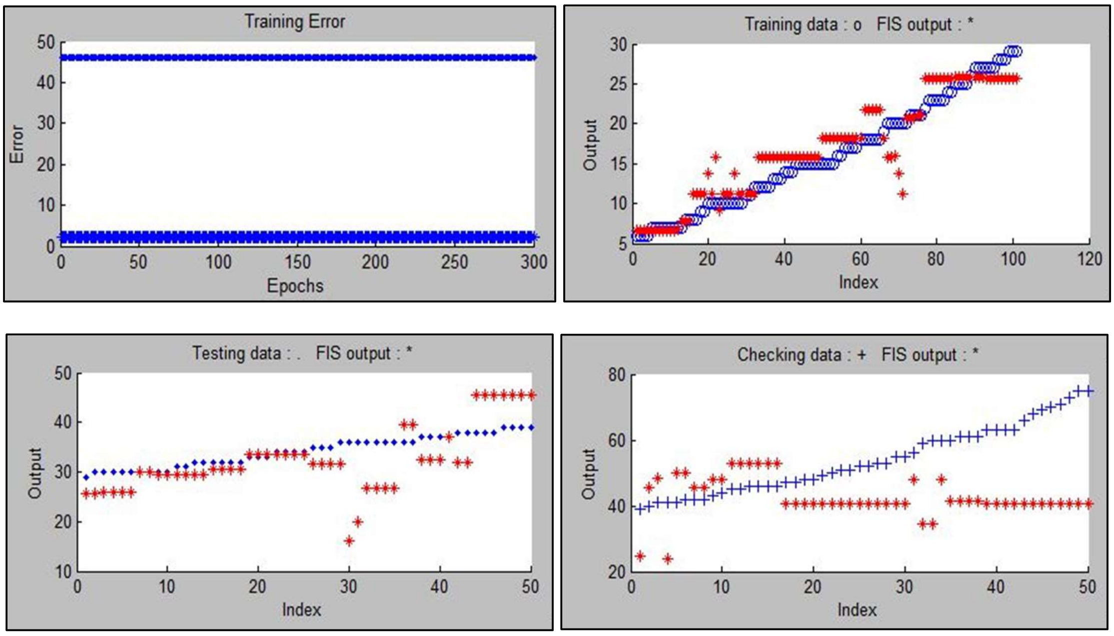


Fig 5.66 Training error and training, testing and checking data with FIS output, scenario 1

The training, testing and checking data with FIS output and training error for neuro fuzzy in scenario 2 are presented in figure 5.67.

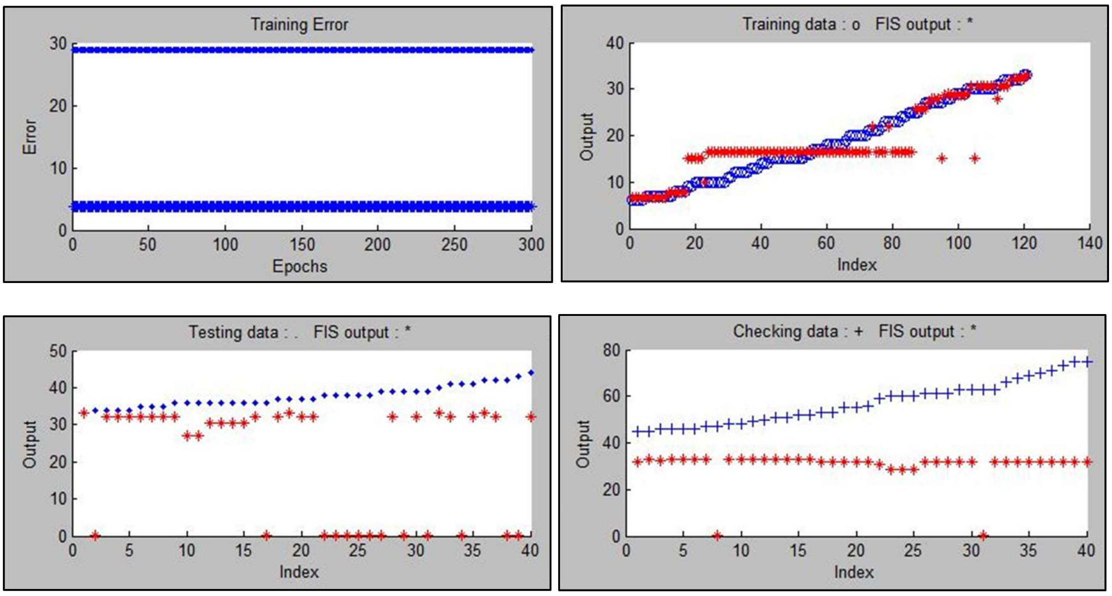


Fig 5.67 Training error and training, testing and checking data with FIS output, scenario 2

The training, testing and checking data with FIS output and training error for neuro fuzzy in scenario 3 are presented in figure 5.68.

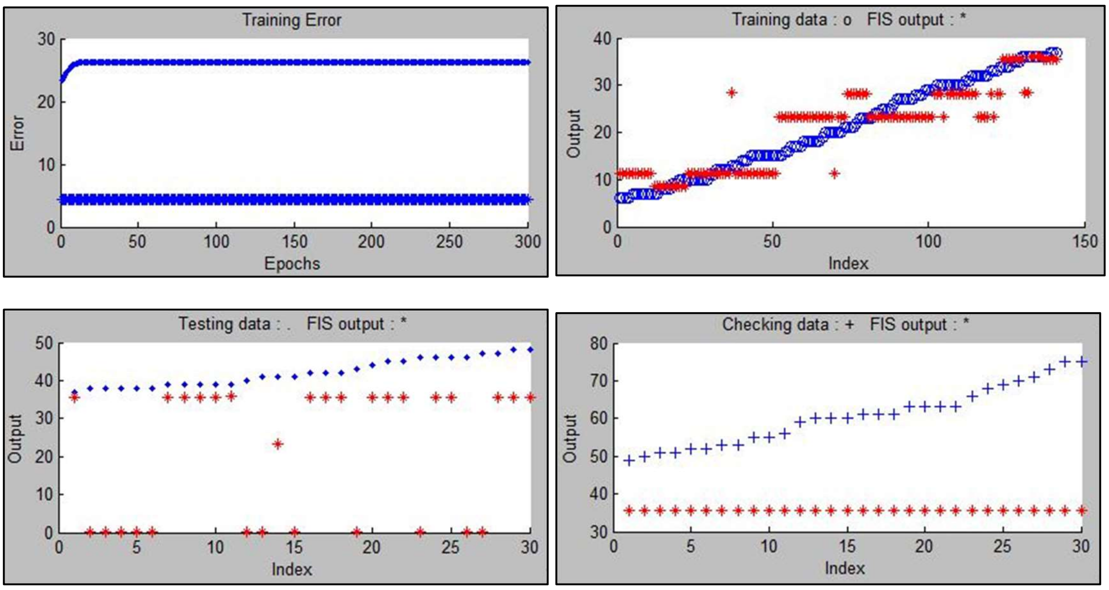


Fig 5.68 Training error and training, testing and checking data with FIS output, scenario 3

The training, testing and checking data with FIS output and training error for neuro fuzzy in scenario 4 are presented in figure 5.69.

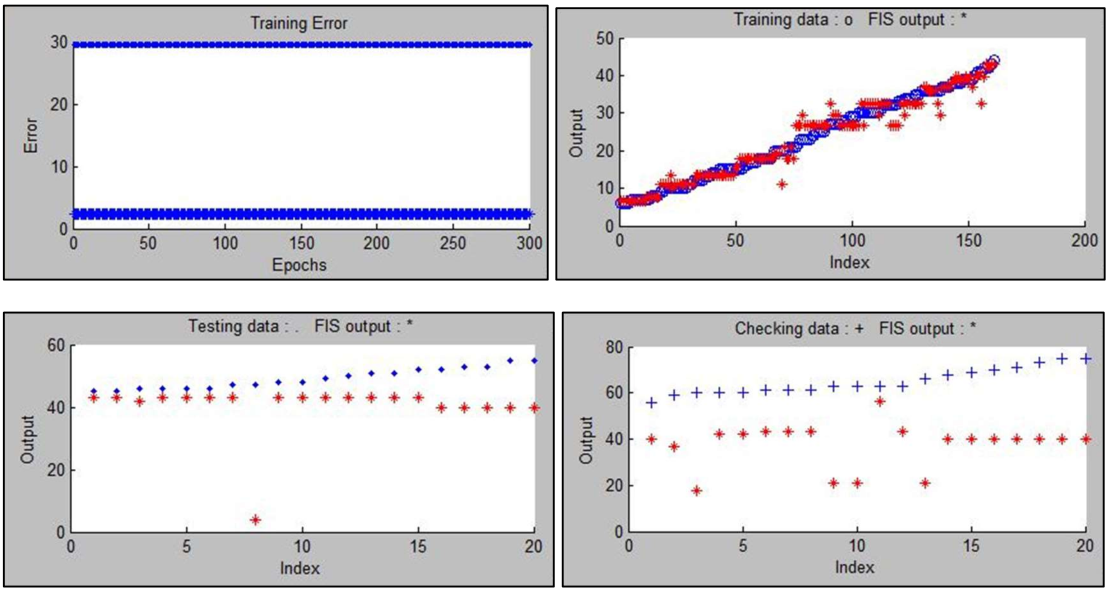


Fig 5.69 Training error and training, testing and checking data with FIS output, scenario 4

Figure 5.70 demonstrates the training, testing and checking data with FIS output. Also, it shows the training error for neuro fuzzy in scenario 5.

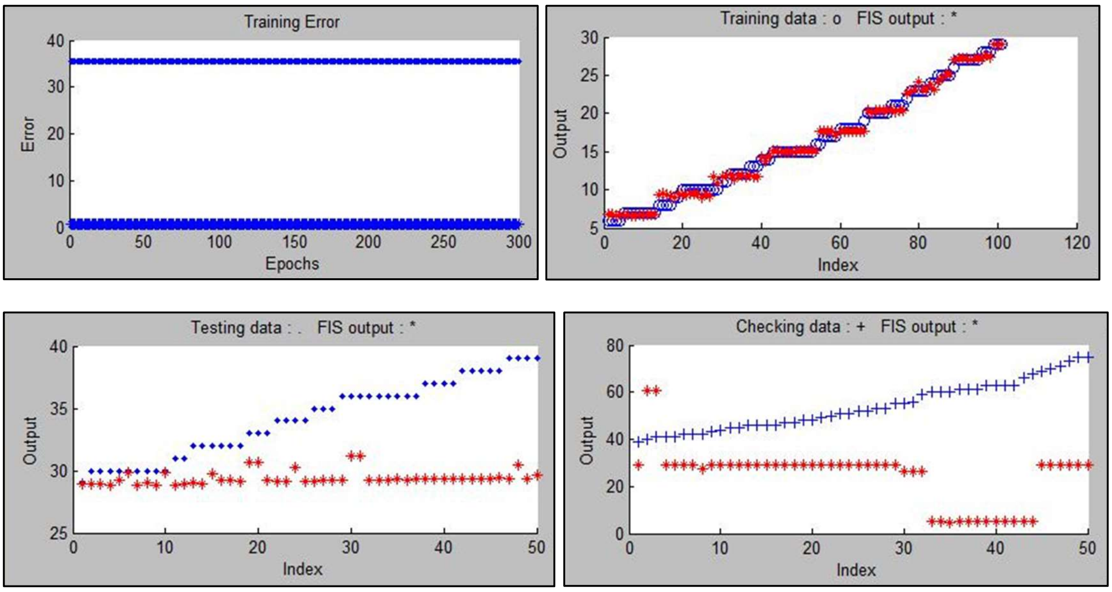


Fig 5.70 Training error and training, testing and checking data with FIS output, scenario 5

Figure 5.71 illustrates the training, testing and checking data with FIS output. In addition, it shows the training error for neuro fuzzy in scenario 6.

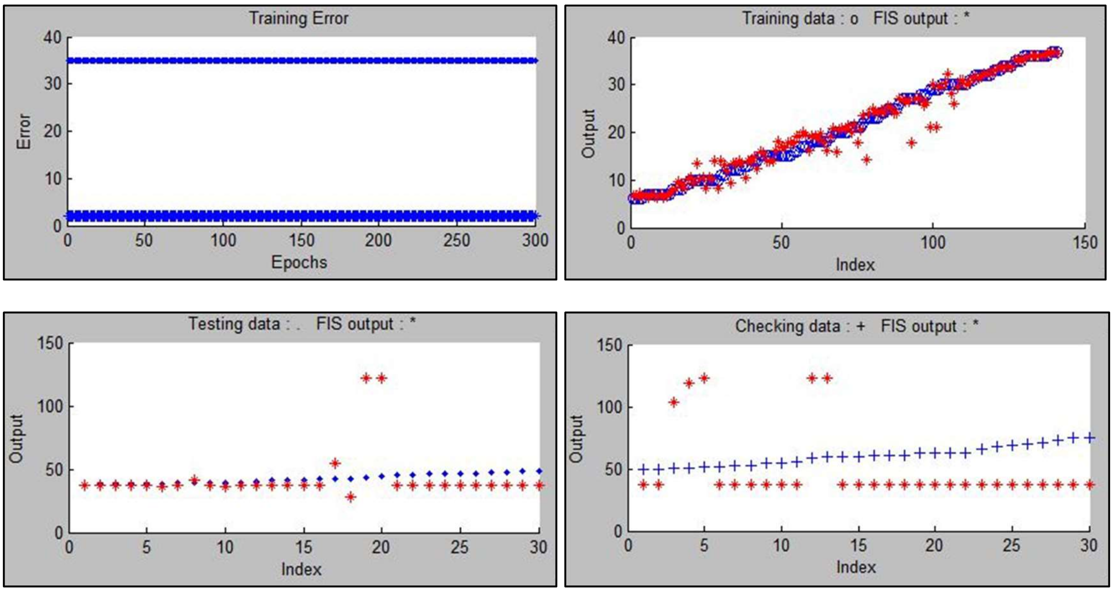


Fig 5.71 Training error and training, testing and checking data with FIS output, scenario 6

Figure 5.72 demonstrates the training, testing and checking data with FIS output. Also, it shows the training error for neuro fuzzy in scenario 7.

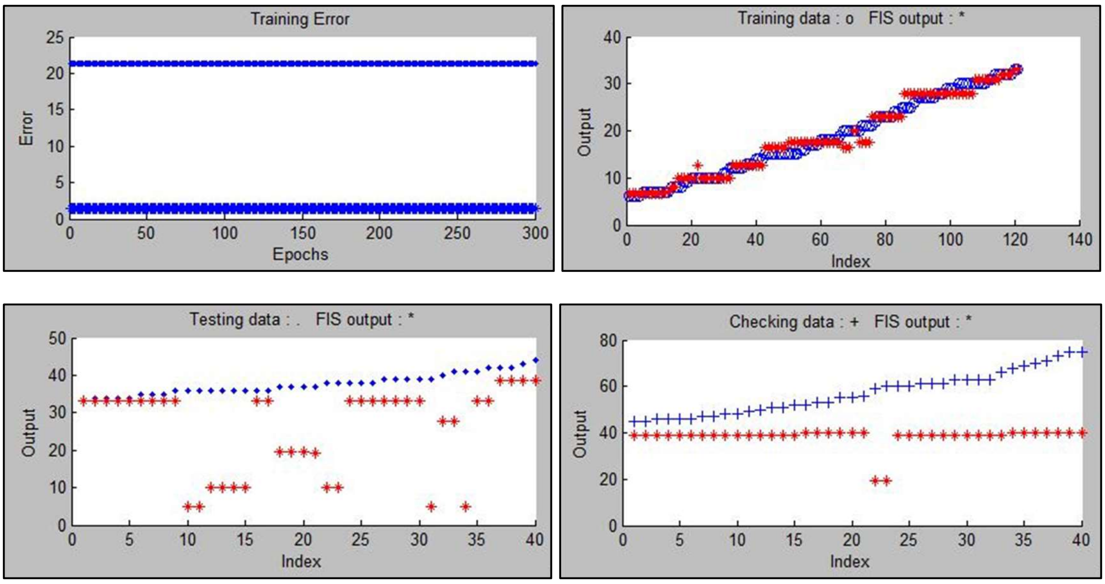


Fig 5.72 Training error and training, testing and checking data with FIS output, scenario 7

Figure 5.73 demonstrates the training, testing and checking data with FIS output. Also, it shows the training error for neuro fuzzy in scenario 8.

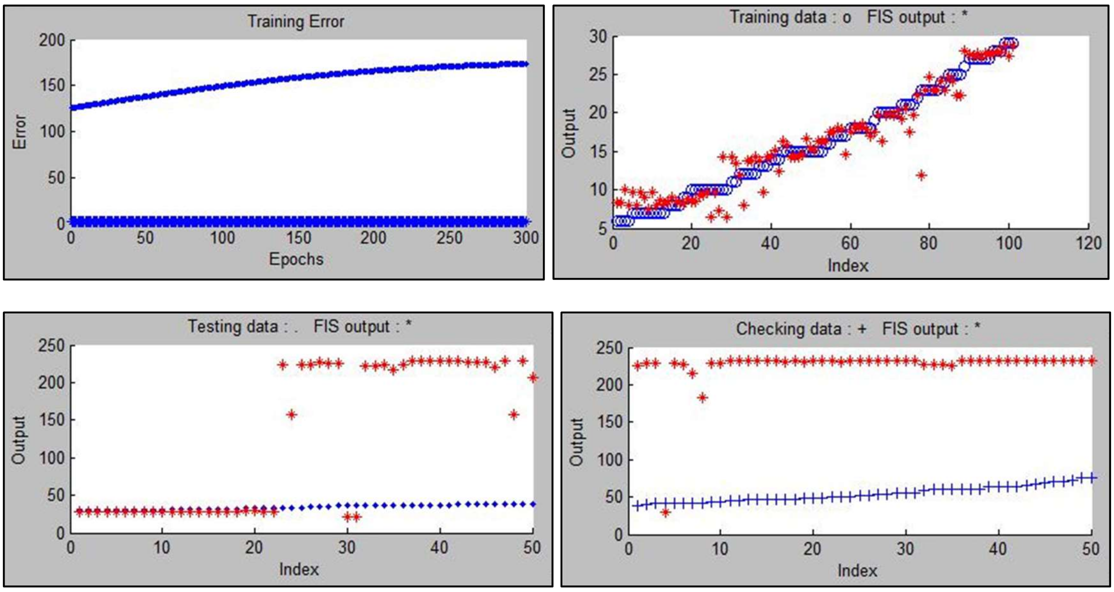


Fig 5.73 Training error and training, testing and checking data with FIS output, scenario 8

In this figures, minimum validation error indicates the optimum number of epochs (300) for construction and testing of the models.

5.13.3. Fuzzy Rule based and Membership functions

The dataset is applied on the Fuzzy logic for the classification purpose. Figures 5.74 shows the modelling of Fuzzy Logic which helps in the classification of the WCP classes.

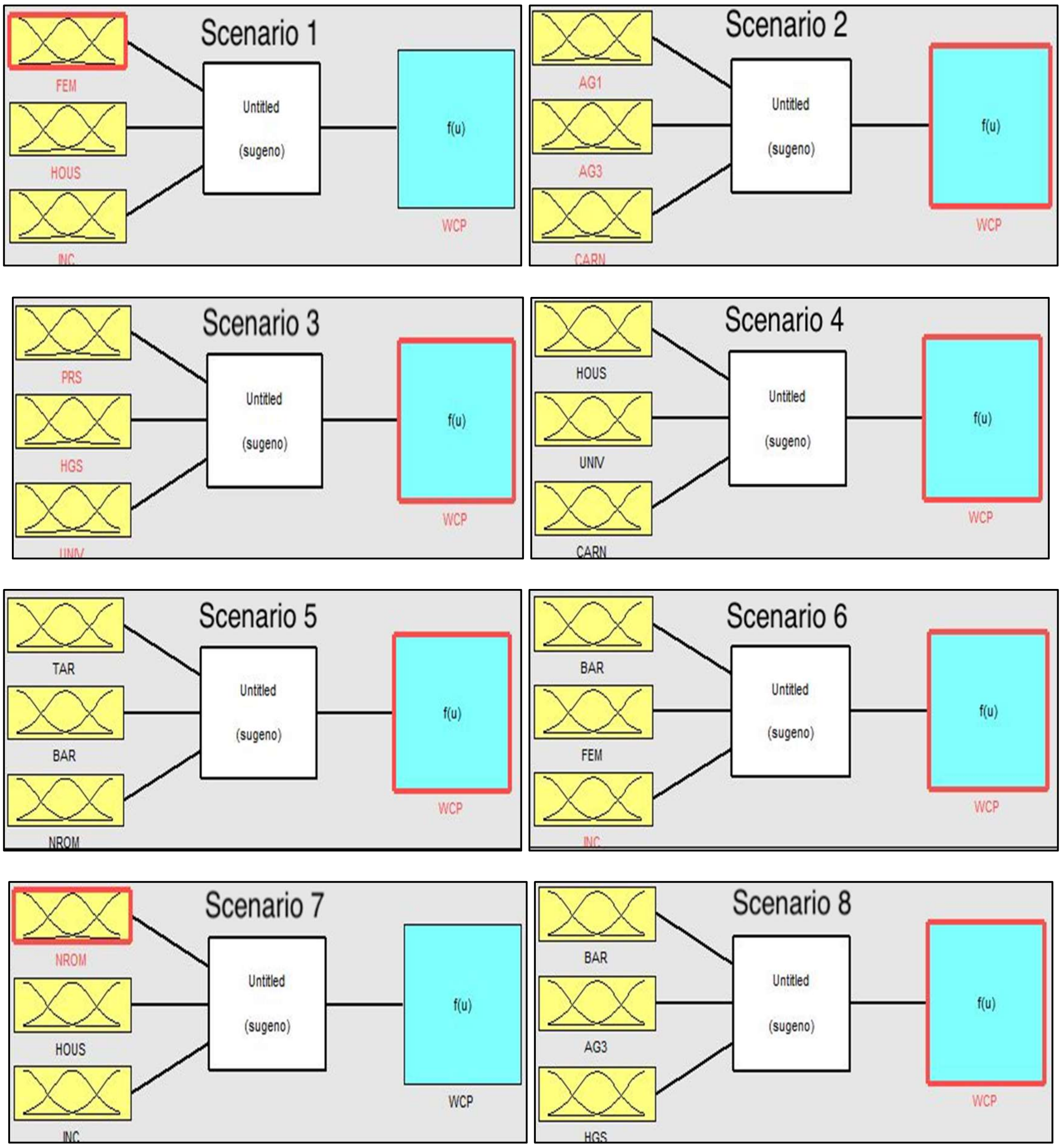


Fig 5.74 Fuzzy rule-based models

The Sugeno model is used in the system. The Sugeno model is computationally efficient, and works well with optimization and adaptive techniques, so it is popular for water problems.

Figures 5.75-5.82 depicted the membership functions for the inputs and the output. Gaussian functions are used for the inputs because of its smoothness.

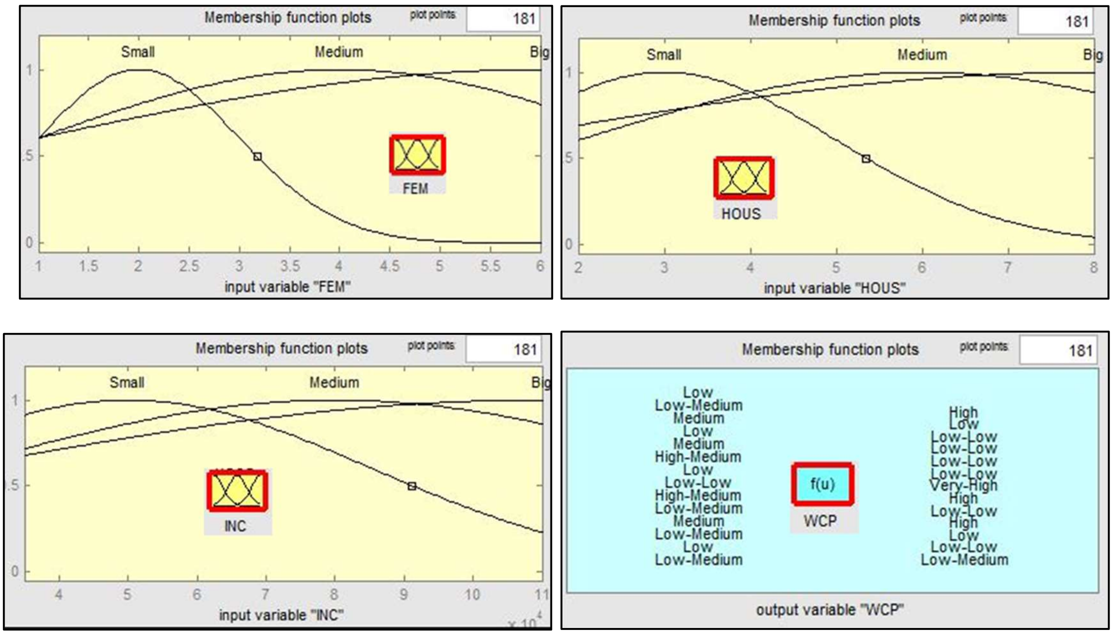


Fig 5.75 Membership functions for inputs and output, scenario 1

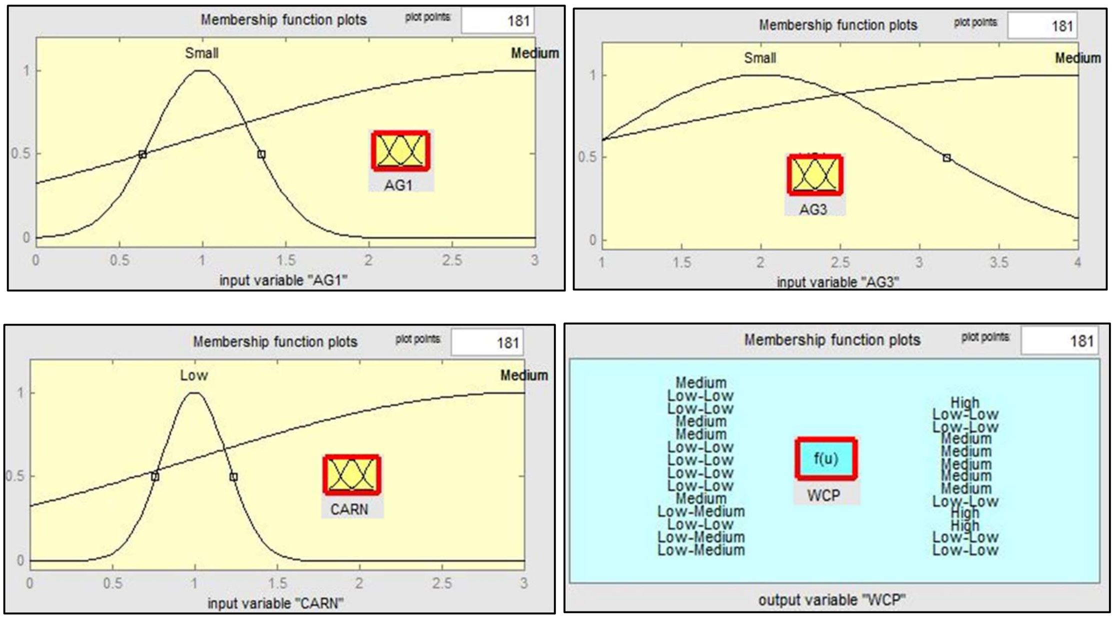


Fig 5.76 Membership functions for inputs and output, scenario 2

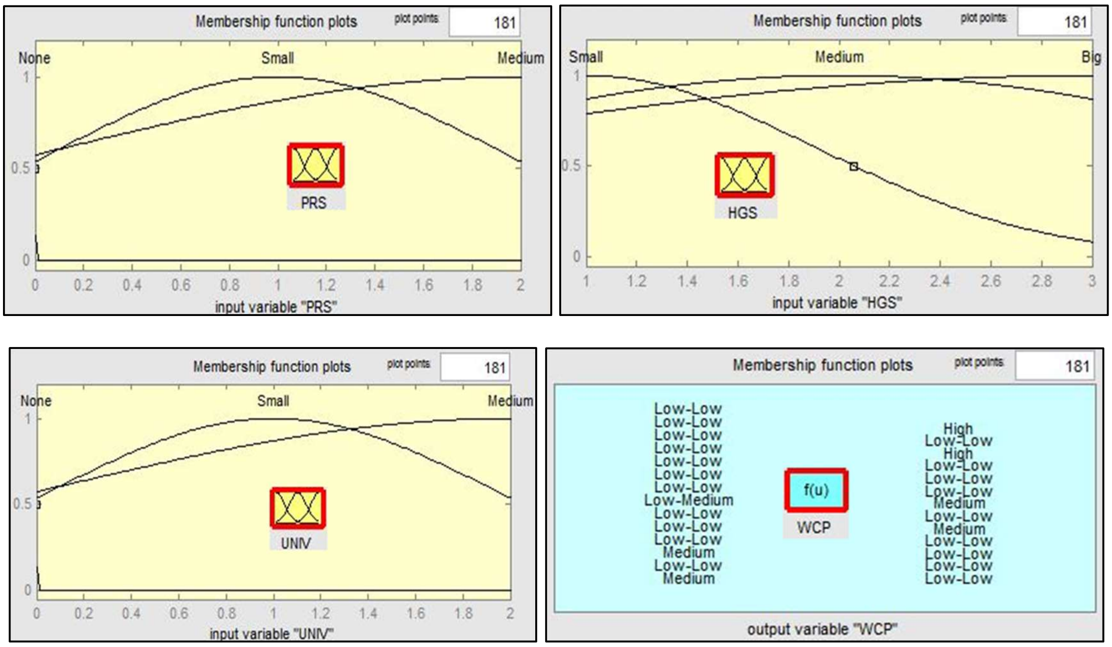


Fig 5.77 Membership functions for inputs and output, scenario 3

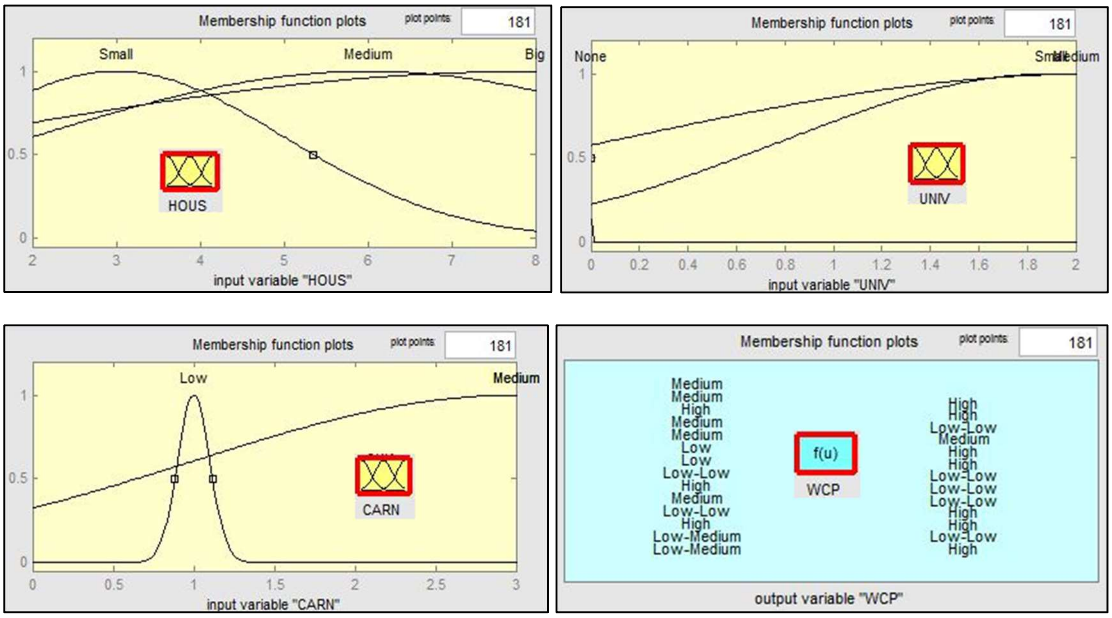


Fig 5.78 Membership functions for inputs and output, scenario 4

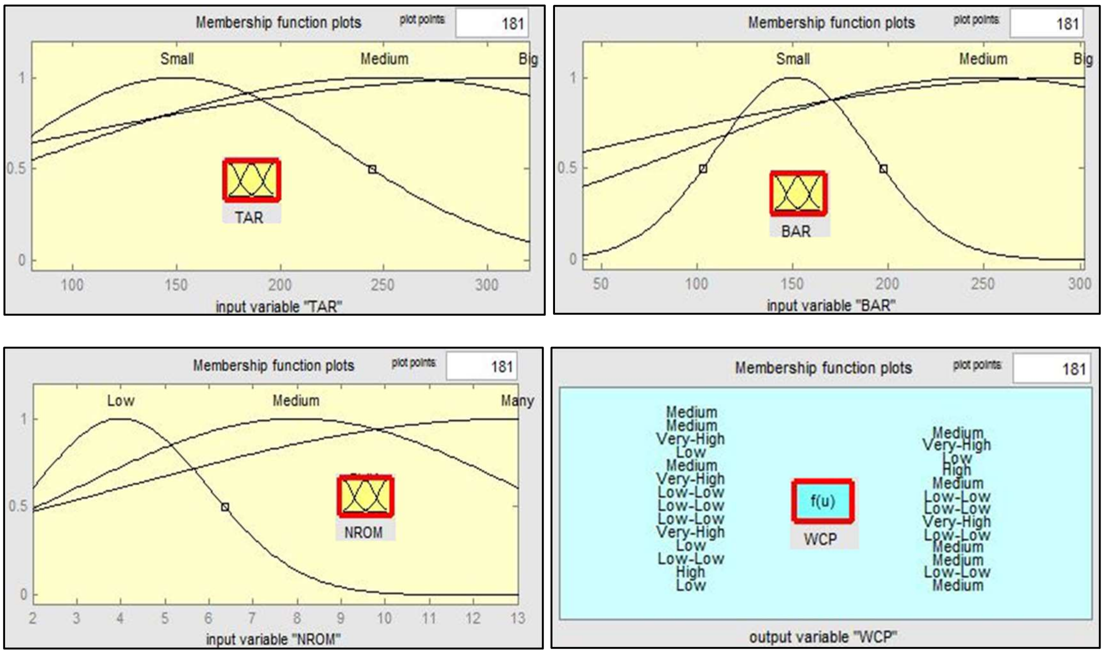


Fig 5.79 Membership functions for inputs and output, scenario 5

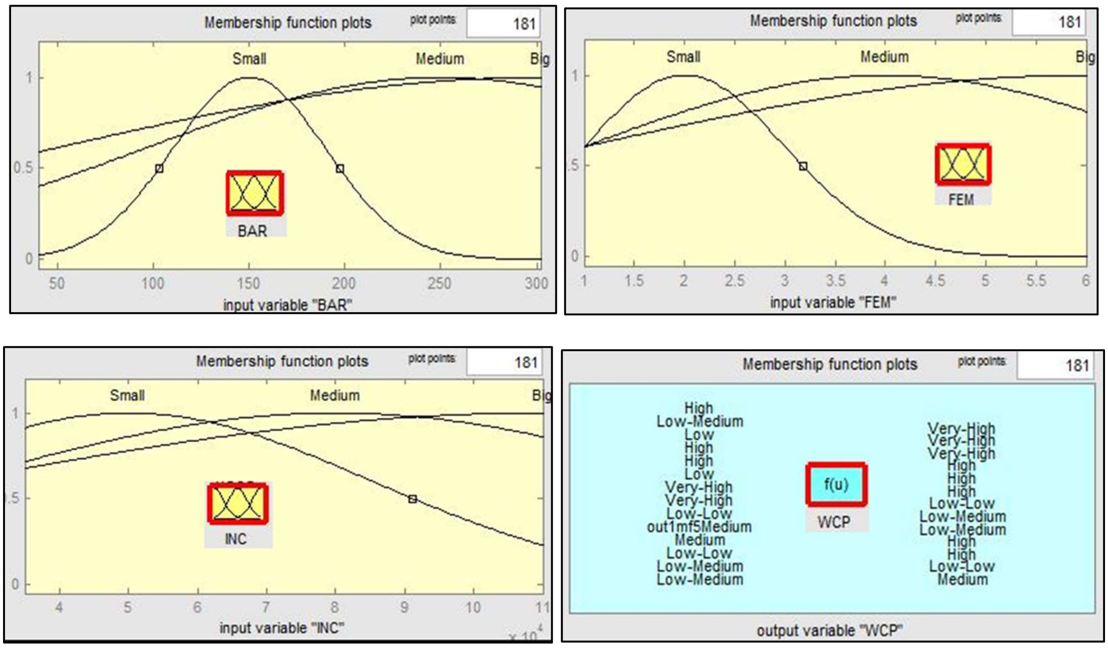


Fig 5.80 Membership functions for inputs and output, scenario 6

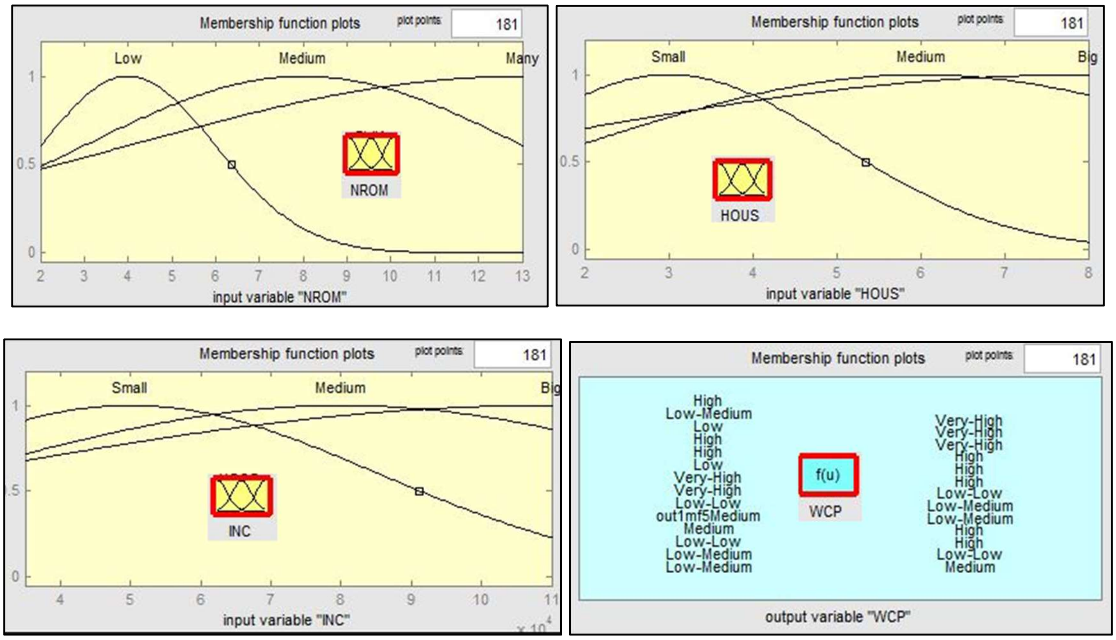


Fig 5.81 Membership functions for inputs and output, scenario 7

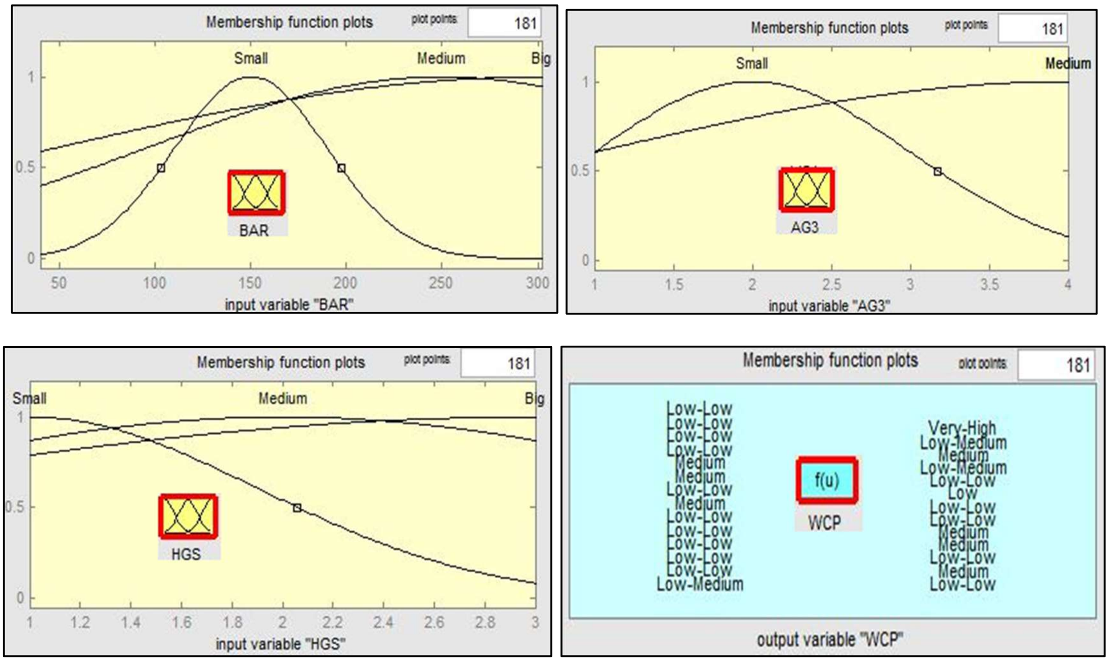


Fig 5.82 Membership functions for inputs and output, scenario 8

5.13.4. Fuzzy Rules

The eight models had 27 rules. Tables 5.28-5.36 demonstrates the fuzzy rules of water consumption. For the inputs, the fuzzy set is divided into small, medium, and big for females, Household size, Income, Age Category 1, Age Category 3, Car Number, Primary School, High School, University, Total Area of the house, Building Area. For Number of Rooms, the fuzzy set is divided into Few, Medium and Many.

As the output factor, the fuzzy sets are divided into following values: low– low, low, low–medium, medium, high–medium, high, and very high.

Table 5.28 Fuzzy rules for water consumption of scenario 1

Rules	Females	Household size	Income	Water Consumption
1	Small	Small	Small	Low-Medium
2	Small	Small	Medium	Low
3	Small	Small	Big	Low-Medium
4	Small	Medium	Small	Medium
5	Small	Medium	Medium	Low-Medium
6	Small	Medium	Big	High-Medium
7	Small	Big	Small	Low-Low
8	Small	Big	Medium	Low
9	Medium	Big	Big	High-Medium
10	Medium	Small	Small	Medium
11	Medium	Small	Medium	Low
12	Medium	Small	Big	Medium
13	Medium	Medium	Small	Low-Medium
14	Medium	Medium	Medium	Low
15	Medium	Medium	Big	Low-Medium
16	Medium	Big	Small	Low-Low
17	Medium	Big	Medium	Low
18	Medium	Big	Big	High
19	Big	Small	Small	Low-Low
20	Big	Small	Medium	High
21	Big	Small	Big	Very-High
22	Big	Medium	Small	Low-Low
23	Big	Medium	Medium	Low-Low
24	Big	Medium	Big	Low-Low
25	Big	Big	Small	Low-Low
26	Big	Big	Medium	Low
27	Big	Big	Big	High

Table 5.29 Fuzzy rules for water consumption of scenario 2

Rules	Age Category 1	Age Category 3	Car Number	Water Consumption
1	Small	Small	Low	Low-Medium
2	Small	Small	Medium	Low-Medium
3	Small	Small	Medium	Low-Medium
4	Small	Medium	Low	Low-Medium
5	Small	Medium	Medium	Medium
6	Small	Medium	Medium	Low-Low
7	Small	Medium	Low	Low-Low
8	Small	Medium	Medium	Low-Low
9	Small	Medium	Medium	Low-Low
10	Medium	Small	Low	Medium
11	Medium	Small	Medium	Medium
12	Medium	Small	Medium	Low-Low
13	Medium	Medium	Low	Low-Low
14	Medium	Medium	Medium	Medium
15	Medium	Medium	Medium	Low-Low
16	Medium	Medium	Low	Low-Low
17	Medium	Medium	Medium	High
18	Medium	Medium	Medium	High
19	Medium	Small	Low	Low-Low
20	Medium	Small	Medium	Medium
21	Medium	Small	Medium	Medium
22	Medium	Medium	Low	Medium
23	Medium	Medium	Medium	Medium
24	Medium	Medium	Medium	Medium
25	Medium	Medium	Low	Low-Low
26	Medium	Medium	Medium	Low-Low
27	Medium	Medium	Medium	High

Table 5.30 Fuzzy rules for water consumption of scenario 3

Rules	Primary School	High School	University	Water Consumption
1	None	Small	None	Medium
2	None	Small	Small	Low-Low
3	None	Small	Medium	Medium
4	None	Medium	None	Low-Low
5	None	Medium	Small	Low-Low
6	None	Medium	Medium	Low-Low
7	None	Big	None	Low-Medium
8	None	Big	Small	Low-Low
9	None	Big	Medium	Low-Low
10	Small	Small	None	Low-Low
11	Small	Small	Small	Low-Low
12	Small	Small	Medium	Low-Low
13	Small	Medium	None	Low-Low
14	Small	Medium	Small	Low-Low
15	Small	Medium	Medium	Low-Low
16	Small	Big	None	Low-Low
17	Small	Big	Small	Low-Low
18	Small	Big	Medium	Low-Low
19	Medium	Small	None	Medium
20	Medium	Small	Small	Low-Low
21	Medium	Small	Medium	Medium
22	Medium	Medium	None	Low-Low
23	Medium	Medium	Small	Low-Low
24	Medium	Medium	Medium	Low-Low
25	Medium	Big	None	High
26	Medium	Big	Small	Low-Low
27	Medium	Big	Medium	High

Table 5.31 Fuzzy rules for water consumption of scenario 4

Rules	Household size	University	Car Number	Water Consumption
1	Small	None	Low	Low-Medium
2	Small	None	Medium	Low-Medium
3	Small	None	High	High
4	Small	Small	Low	Low-Low
5	Small	Small	Medium	Medium
6	Small	Small	Medium	High
7	Small	Medium	Low	Low-Low
8	Small	Medium	Medium	Low
9	Small	Medium	Medium	Low
10	Medium	None	Low	Medium
11	Medium	None	Medium	Medium
12	Medium	None	Medium	High
13	Medium	Small	Low	Medium
14	Medium	Small	Medium	Medium
15	Medium	Small	Medium	High
16	Medium	Medium	Low	Low-Low
17	Medium	Medium	Medium	High
18	Medium	Medium	Medium	High
19	Big	None	Low	Low-Low
20	Big	None	Medium	Low-Low
21	Big	None	Medium	Low-Low
22	Big	Small	Low	High
23	Big	Small	Medium	High
24	Big	Small	Medium	Medium
25	Big	Medium	Low	Low-Low
26	Big	Medium	Medium	High
27	Big	Medium	Medium	High

Table 5.32 Fuzzy rules for water consumption of scenario 5

Rules	Total Area of the house	Building Area	Number of Rooms	Water Consumption
1	Small	Small	Few	Low
2	Small	Small	Medium	High
3	Small	Small	Many	Low-Low
4	Small	Medium	Few	Low
5	Small	Medium	Medium	Very-High
6	Small	Medium	Many	Low-Low
7	Small	Big	Few	Low-Low
8	Small	Big	Medium	Low-Low
9	Small	Big	Many	Very-High
10	Medium	Small	Few	Medium
11	Medium	Small	Medium	Low
12	Medium	Small	Many	Very-High
13	Medium	Medium	Few	Medium
14	Medium	Medium	Medium	Medium
15	Medium	Medium	Many	Medium
16	Medium	Big	Few	Low-Low
17	Medium	Big	Medium	Medium
18	Medium	Big	Many	Medium
19	Big	Small	Few	Low-Low
20	Big	Small	Medium	Very-High
21	Big	Small	Many	Low-Low
22	Big	Medium	Few	Low-Low
23	Big	Medium	Medium	Medium
24	Big	Medium	Many	High
25	Big	Big	Few	Low
26	Big	Big	Medium	Very-High
27	Big	Big	Many	Medium

Table 5.33 Fuzzy rules for water consumption of scenario 6

Rules	Building Area	Females	Income	Water Consumption
1	Small	Small	Small	Low
2	Small	Small	Medium	Low-Medium
3	Small	Small	Big	High
4	Small	Medium	Small	Very-High
5	Small	Medium	Medium	Low-Low
6	Small	Medium	Big	Low-Low
7	Small	Big	Small	Low-Medium
8	Small	Big	Medium	Low-Low
9	Small	Big	Big	Very-High
10	Medium	Small	Small	Low-Medium
11	Medium	Small	Medium	High
12	Medium	Small	Big	High
13	Medium	Medium	Small	Low
14	Medium	Medium	Medium	Very-High
15	Medium	Medium	Big	Low-Low
16	Medium	Big	Small	Medium
17	Medium	Big	Medium	High
18	Medium	Big	Big	Medium
19	Big	Small	Small	Very-High
20	Big	Small	Medium	Low-Medium
21	Big	Small	Big	High
22	Big	Medium	Small	High
23	Big	Medium	Medium	Low-Medium
24	Big	Medium	Big	Very-High
25	Big	Big	Small	Medium
26	Big	Big	Medium	High
27	Big	Big	Big	High

Table 5.34 Fuzzy rules for water consumption of scenario 7

Rules	Number of Rooms	Household size	Income	Water Consumption
1	Few	Small	Small	Low-Medium
2	Few	Small	Medium	Low-Medium
3	Few	Small	Big	Low-Low
4	Few	Medium	Small	Medium
5	Few	Medium	Medium	Medium
6	Few	Medium	Big	Low-Low
7	Few	Big	Small	Very-High
8	Few	Big	Medium	Very-High
9	Few	Big	Big	Low
10	Medium	Small	Small	High
11	Medium	Small	Medium	High
12	Medium	Small	Big	Low
13	Medium	Medium	Small	Low-Medium
14	Medium	Medium	Medium	High
15	Medium	Medium	Big	Medium
16	Medium	Big	Small	Low-Low
17	Medium	Big	Medium	High
18	Medium	Big	Big	High
19	Many	Small	Small	Low-Medium
20	Many	Small	Medium	Low-Medium
21	Many	Small	Big	Low-Low
22	Many	Medium	Small	High
23	Many	Medium	Medium	High
24	Many	Medium	Big	High
25	Many	Big	Small	Very-High
26	Many	Big	Medium	Very-High
27	Many	Big	Big	Very-High

Table 5.35 Fuzzy rules for water consumption of scenario 8

Rules	Building Area	Age Category 3	High School	Water Consumption
1	Small	Small	Small	Low-Medium
2	Small	Small	Medium	Low-Low
3	Small	Small	Big	Low-Low
4	Small	Medium	Small	Low-Low
5	Small	Medium	Medium	Low-Low
6	Small	Medium	Big	Low-Low
7	Small	Medium	Small	Medium
8	Small	Medium	Medium	Low-Low
9	Small	Medium	Big	Medium
10	Medium	Small	Small	Medium
11	Medium	Small	Medium	Low-Low
12	Medium	Small	Big	Low-Low
13	Medium	Medium	Small	Low-Low
14	Medium	Medium	Medium	Low-Low
15	Medium	Medium	Big	Low-Low
16	Medium	Medium	Small	Medium
17	Medium	Medium	Medium	Low-Low
18	Medium	Medium	Big	Medium
19	Big	Small	Small	Medium
20	Big	Small	Medium	Low-Low
21	Big	Small	Big	Low-Low
22	Big	Medium	Small	Low
23	Big	Medium	Medium	Low-Low
24	Big	Medium	Big	Low-Medium
25	Big	Medium	Small	Medium
26	Big	Medium	Medium	Low-Medium
27	Big	Medium	Big	Very-High

Results of the eight fuzzy models are depicted in Table 5.36. The advantages of the models are their simple structure and easy representation. The disadvantages of them, had large error values, different inputs in each scenario and disability to assess the variation of water consumption with all the parameters.

Table 5.36: Results of fuzzy models

Modes	Average Errors		
	Training	Testing	Checking
S1	2.47	5.79	17.38
S2	3.69	22.63	28.97
S3	4.33	27.01	26.21
S4	2.29	12.69	29.39
S5	0.63	5.66	35.53
S6	1.94	21.31	35.07
S7	1.41	16.39	21.38
S8	2.01	132.62	173.55

Model S5 has small training error compared to other models. It can evaluate the variation of water consumption, however, error values for the other scenarios in the three phases are still high. Figures 5.83 below demonstrates the forecasted water consumption from ANFIS for model 6.

As shown in this figure 5.83, the quantity of water consumed is around 38.5 m^3, where the building area equal to 280 m^2, number of females is 4 and monthly income is about 100 000 DA per month. Figures 5.84 illustrates some examples of water consumption predictions.

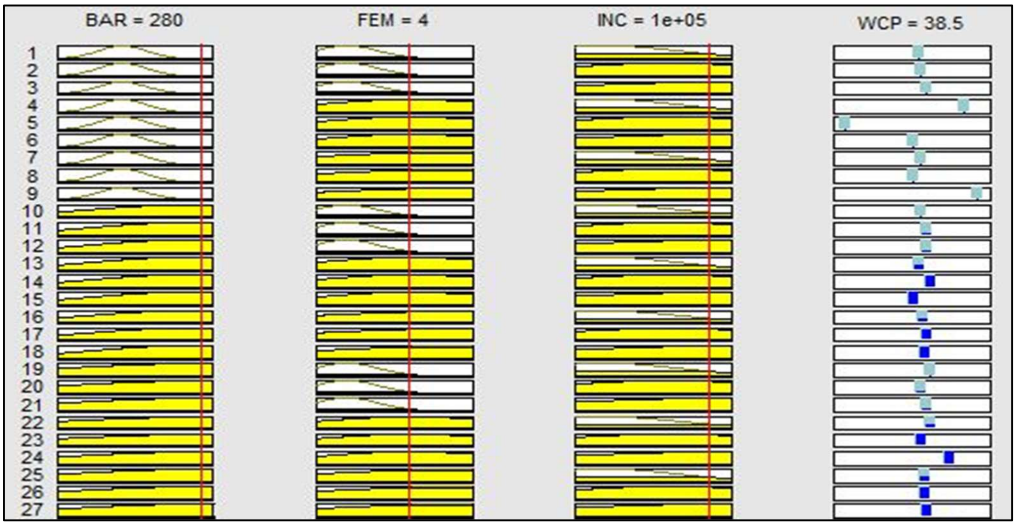


Fig 5.83 Water consumption prediction in scenario 6

Fig 5.84 Water consumption predictions

(ANNEX 06)

The three-dimensional (3D) relationship of females, household size, income and water consumption for scenario 1 is shown in figure 5.85.

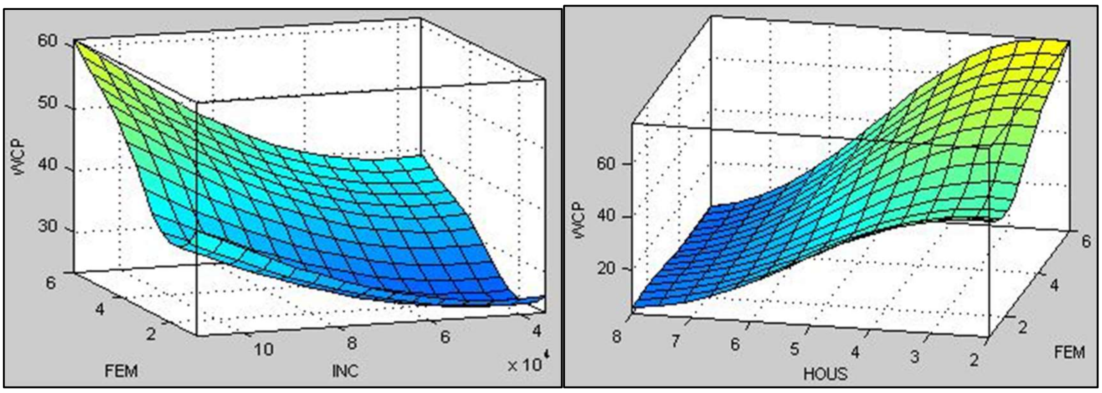


Fig 5.85 Relationship between females, household size, income and water consumption

The three-dimensional (3D) relationship of age category 1, age category 3, car number and water consumption for scenario 2 is shown in figure 5.86.

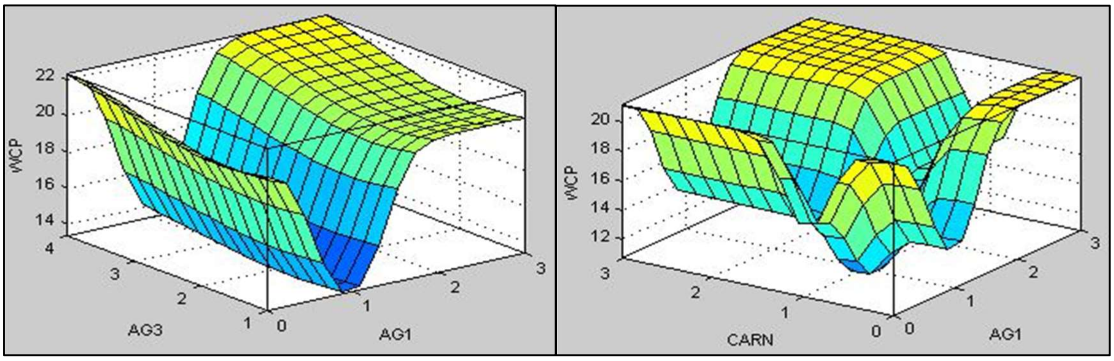


Fig 5.86 Relationship between age category 1, age category 3, car number and water consumption

The three-dimensional (3D) relationship of primary school, high school, university and water consumption for scenario 3 is shown in figure 5.87.

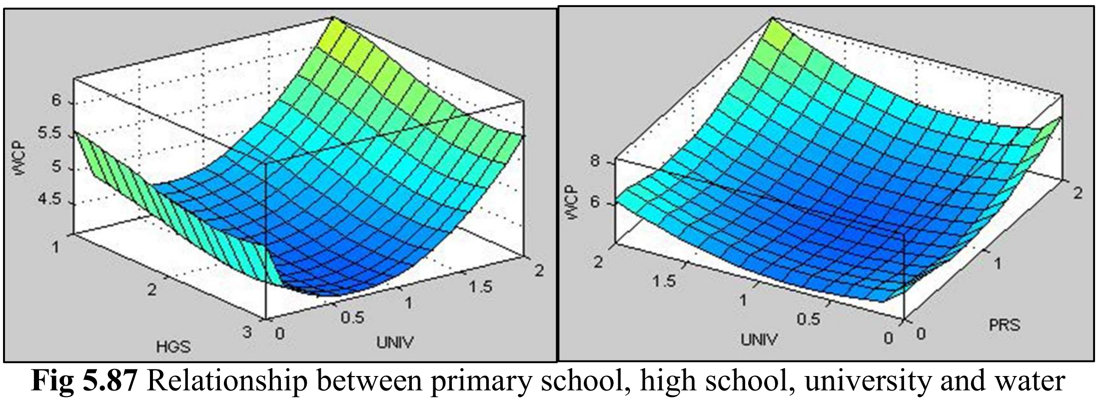


Fig 5.87 Relationship between primary school, high school, university and water consumption

The three-dimensional (3D) relationship of university level, household size, car numbers and water consumption for scenario 4 is shown in figure 5.88.

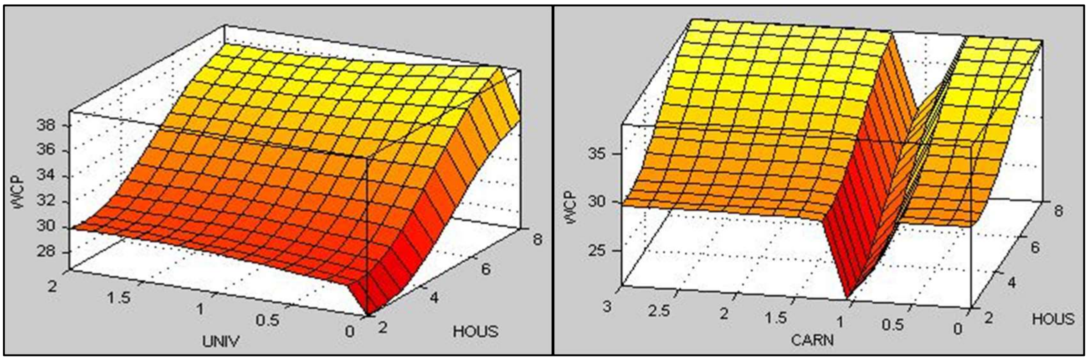


Fig 5.88 Relationship between university level, household size, car numbers and water consumption

The three-dimensional (3d) relationship of total area of the house, building area, number of rooms and water consumption for scenario 5 is shown in figure 5.89.

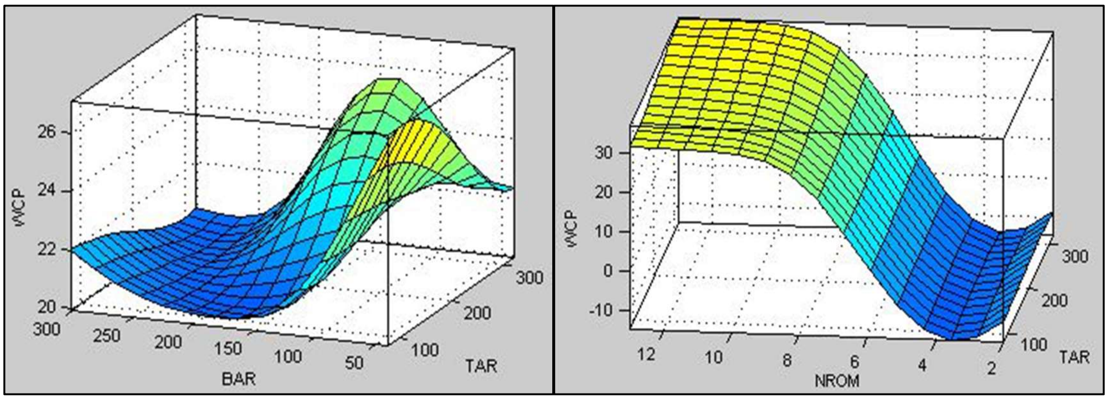


Fig 5.89 Relationship between total area of the house, building area, number of rooms and water consumption

The three-dimensional (3d) relationship of building area, income, females and water consumption for scenario 6 is depicted in figure 5.90.

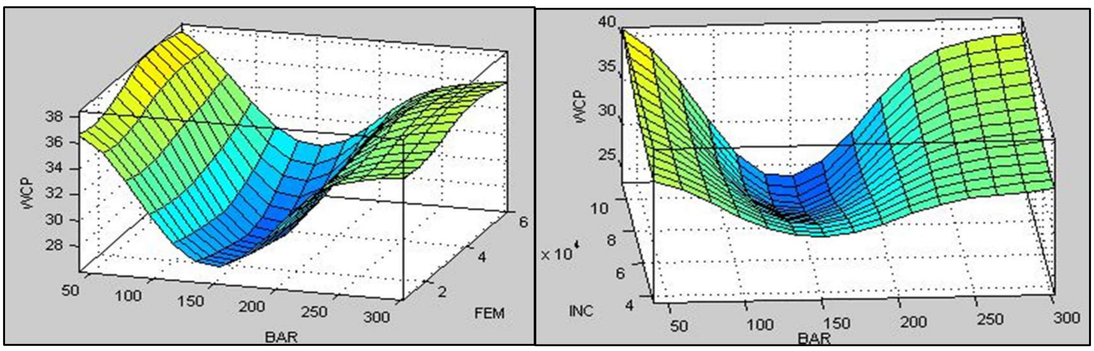


Fig 5.90 Relationship between building area, income, females and water consumption

The three-dimensional (3d) relationship of number of rooms, income, household size and water consumption for scenario 7 is shown in figure 5.91.

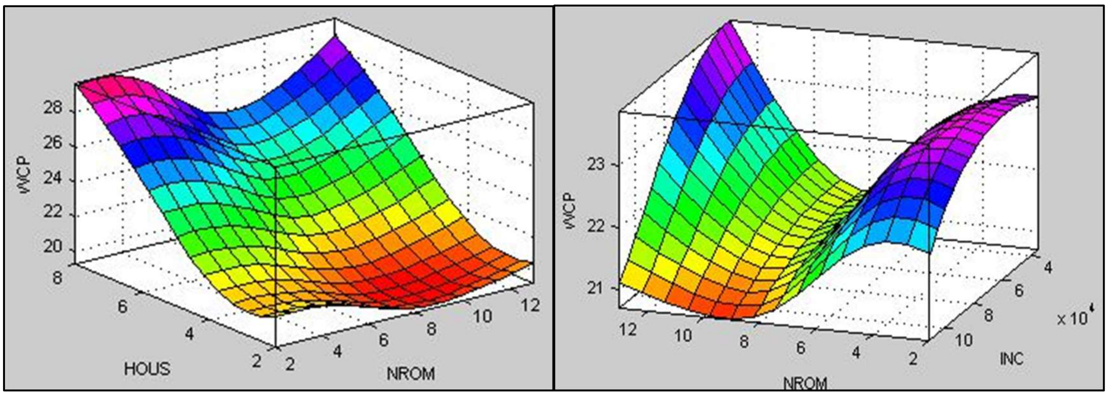


Fig 5.91 Relationship between number of rooms, income, household size and water consumption

The three-dimensional (3d) relationship of building area, high school, age category 3 and water consumption for scenario 8 is shown in figure 5.92.

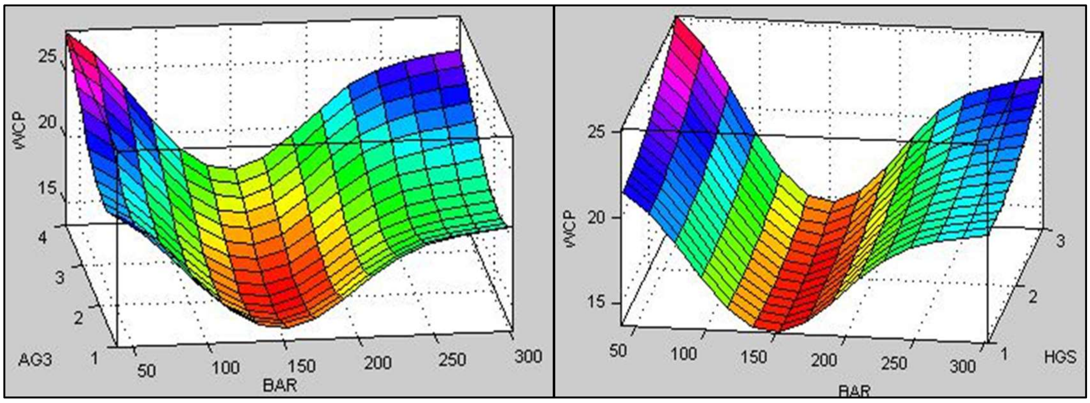


Fig 5.92 Relationship between building area, high school, age category 3 and water consumption

The fuzzy system can deal with uncertain and vague data. The system uses three input factors and develops 27 rules for the eight models to forecast trimester water consumption. To obtain the optimum structure of the model, input data are divided into different sections. This needs several models to be constructed. The results show that the best structure is related to the grid partition of 3-3-3. It is also observed that with increasing the partitions, the model error indicators are higher. The reason may be the fact that with increasing the partitions, the number of required rules is increased exponentially, leading to complexity of the model. For instance, the number of required rules for a model with partition of 3-3-3 are equal to 27. The eight selected models can be used for forecasting Sedrata water consumption without needing to estimate any indoor habits or climatic input parameters. MSE indices is better in model 5 than the other models.

As it can be inferred from the performance parameters that depicted in tables above of ANFIS results in each scenario, the performance of ANFIS was so low in terms of prediction performance of WCP compared to ANNs. The reasons behind pure WCP predictions with ANFIS may be due to software limitations of ANFIS, lack of script that may contain the efforts to develop a successful model and user errors due to low degree of user-friendliness of ANFIS.

5.14. Conclusion 2

The research is conducted through over 50 ANNs models in hope to define the optimum models and architecture. To judge performance of each scenario, indicators like **Mean Square Error** and **Correlation Coefficient** are employed. The inputs in every scenario are selected in way that covers all possible combinations and exclude the non-influential variables for efficiency reasons. In socio-economic scenario, models with architecture (training-testing-validation) of (50 - 25 – 25), (60 -20 -20), (70 -15 -15), (80 -10 -10) and corresponding to hidden layers 7 / 6 / 6 / 5 respectively are the best. For physical characteristics scenario, the best architectures are (50-25-25), (60- 20 -20), (70- 15 -15), (80 -10- 10), (60-30-10) with hidden layers 4, 2, 2, 2, 4 respectively. Another major finding is that when combining all inputs at once, the performance of modelling improves significantly.

As discussed already in literature review, the fuzzy approach can deal better with uncertain and vague data. The obtained results helped to obtain 27 rules for the eight models to forecast trimester water consumption basing on three input factors. To obtain the optimum structure of the model, input data are divided into different sections. This needs several models to be constructed. The results show that the best structure is related to the grid partition of 3-3-3.

Moreover, the usage of adaptive Sugeno fuzzy and neuro-fuzzy inference system (ANFIS) models is preferable thanks to their simple structure and easy representation.

The general conclusion from this line of work is that models developed from neural networks had good perform. This superiority in performance is attributed to the ability of neural networks to efficiently capture non-linearities.

General Conclusion

General conclusion

Defining the key parameters governing the water consumption will help the cities planners to improve life conditions. Such task is not easy because of the huge number of inputs and consequently large error margin. For that, preliminary statistical tests are employed to remove all outliers and data noise .for the present case study, the data set is reduced by approximately thirty percent.

Estimation of water use determinants requires a reliable measure of water consumption and information on the consumers and their houses. The study was conducted using a questionnaire that contains the purpose of the survey, followed by questions on water usage practices and behaviors. The survey questionnaire is built basing on previous studies and covers all relevant parameters.

A dependable assessment of water consumption requires a profound understanding of the links between the considered inputs. This is achieved in the current thesis by applying a variety of statistical techniques and tools. In fact, five methods are applied and the results demonstrated that basing on correlation analysis, ANOVA, cluster analysis, factor analysis and principal component analysis:

- Household size and monthly income are very correlated with water consumption.
- Females use more water.
- In age groups; children and adults are responsible for water usage.
- Education level of residents has an impact on water use.
- Houses with cars tend to use more water.
- Water consumption is correlated with physical characteristics of household.
- For indoor habits water consumption is associated with frequency of washing clothes and using toilets.

As consequence, twelve (12) variables out of sixteen (16) initial are proven to be determinants of domestic water usage: nine (09) are socioeconomic while the rest three (03) are physical characteristics of households.

The application of Artificial Neural Networks for modelling domestic water demand requires in the first place the definition of optimum models and architectures and for those goal indicators like **Mean Square Error** and **Correlation Coefficient** used.

The inputs in every scenario are selected in way that covers all possible combinations and exclude the non-influential variables for efficiency reasons. Another major finding is that when combining all inputs at once, the performance of modelling improves significantly.

For the the fuzzy approach, the obtained results helped to obtain 27 rules for the eight models to forecast trimester water consumption basing on three input factors. To obtain the optimum structure of the model, input data are divided into different sections. The results show that the best structure is related to the grid partition of 3-3-3. Moreover, the usage of adaptive Sugeno fuzzy and neuro-fuzzy inference system (ANFIS) models is preferable thanks to their simple structure and easy representation.

References

References

1. A. zadeh, A; Neshat, N; Hamidipour, H. Hybrid Fuzzy Regression–Artificial Neural Network for Improvement of Short-Term Water Consumption Estimation and Forecasting in Uncertain and Complex Environments: Case of a Large Metropolitan City. J. Water Resour. Plan. Manag. **2012**, 138, 71–75.
2. Abdi H and Williams, L J, 2010. Principal component analysis, interdisciplinary reviews: Computational Statistics, Vol.2, No.4, 387-515. **https://doi.org/10.1002/wics.101.**
3. Adamowski, J.F., 2008. Peak daily water demand forecast modeling using artificial neural networks. Journal of Water Resources Planning and Management 134 (2), 119–128.
4. Adamowski, Jan, Hiu Fung Chan, Shiv O Prasher, Bogdan Ozga-zielinski, and Anna Sliusarieva. "Comparison of Multiple Linear and Nonlinear Regression, Autoregressive Integrated Moving Average, Artificial Neural Network, and Wavelet Artificial Neural Network Methods for Urban Water Demand Forecasting in Montreal , Canada." WATER RESOURCES RESEARCH 48, no. W01528 (2012). **https://doi.org/10.1029/2010WR009945.**
5. ADE; Algérienne des eaux, ADE", water authority distribution and management, of Souk Ahras from the year 2012 to 2017.
6. Agami, Nedaa, Amir Atiya, Mohamed Saleh, and Hisham El-Shishiny. "A Neural Network Based Dynamic Forecasting Model for Trend Impact Analysis." Technological Forecasting and Social Change 76 (2009): 952–62. **https://doi.org/10.1016/j.techfore.2008.12.004.**
7. Aggrawl, Raj and Yonghua, Song. "Artificial neural networks in power systems". POWER ENGINEERING JOURNAL, June 1997.
8. Agthe, D.E., Billings, R.B., 1987. Equity, price elasticity, and household income under increasing block rates for water. American Journal of Economics and Sociology 46 (3), 273–286.
9. Agthe, Donald E. and R.Bruce Billings, "Equity, Price Elasticity and Household Income under Increasing Block Rates for Water", American Journal of Economics and Sociology, Vol.46, No.3, July 1987.
10. Ahmadi, S., Alizadeh, S., Forouzideh, N., Yeh, C.-H., Martin, R., & Papageorgiou, E. (2014). ICLA imperialist competitive learning algorithm for fuzzy cognitive map: application to water demand forecasting. In: 2014 I.E. International Conference on Fuzzy Systems (FUZZ-IEEE), (pp. 1041–1048): IEEE.
11. Ajbar AH, Ali EM (2013) Prediction of municipal water production in touristic Mecca City in Saudi Arabia using neural networks. J King Saud Univ Eng Sci 27(1):83–91
12. Al-Zahrani, M.A., Abo-Monasar, A. Urban Residential Water Demand Prediction Based on Artificial Neural Networks and Time Series Models. Water Resour Manage **29,** 3651–3662 (2015). **https://doi.org/10.1007/s11269-015-1021-z.**
13. Almutaz I, Ajbar A, Ali E (2012b) Determinants of residential water demand in an arid and oil rich country: a case study of Riyadh city in Saudi Arabia. Int J Phys Serv 7(43):5787–5796
14. Almutaz I, Ajbar A, Khalid Y, Ali E (2012a) A probabilistic forecast of water demand for a tourist and desalination dependent city: case of Mecca, Saudi Arabia. Desalination 294:53–59. **https://doi.org/10.1016/j.desal.2012.03.010.**
15. Altunkaynak, A., Özger, M. & Çakmakci, M. Water Consumption Prediction of Istanbul City by Using Fuzzy Logic Approach. Water Resour Manage **19,** 641–654 (2005). **https://doi.org/10.1007/s11269-005-7371-1.**
16. Anderson, J. S. ; Lall, S. P. ; Anderson, D. M. ; Chandrasoma, J., 1992. Apparent and true availability of amino acids from common feed ingredients for Atlantic salmon (*Salmo salar*) reared in sea water. Aquaculture, 108 (1-2): 111-124
17. Aquacraft Inc. Application of end use study data for development of residential demand models. 2015. Retrieved from**:** **http://www.aquacraft.com/wp-content/uploads/2015/09/Residential-Models.pdf.**
18. Arbuès, F., R. Barbera_n, and I. Villanu_a, 2004. Price Impact on Urban Residential Water Demand: A Dynamic Panel Data Approach. Water Resources Research 40: W11402.

19. Arbuès, F., Villanuà, I., 2006. Potential for pricing policies in water resource management: estimation of urban residential water demand in Zaragoza, Spain. Urban Stud. 43 (13), 2421–2442.
20. Arbuès, F., Villanuà, I., Barberàn, R., 2010. Household size and residential water demand: an empirical approach. Aust. J. Agric. Resour. Econ. 54 (1), 61–80.
21. Ardjmand, E., Millie, D. F., Ghalehkhondabi, I., Young II, W. A., & Weckman, G. R. (2016). A state-based sensitivity analysis for distinguishing the global importance of predictor variables in artificial neural networks. Advances in artificial neural systems. **https://doi.org/10.1155/2016/2303181.**
22. Aristidis Likas, Nikos Vlassis, Jakob Verbeek. The global k-means clustering algorithm. Pattern Recognition, Elsevier, 2003, 36 (2), pp.451 - 461.
23. Babel, M.S., Gupta, A.D., Pradhan, P., 2007. A multivariate econometric approach for domestic water demand modeling: an application to Kathmandu, Nepal. Water Resour. Manage. 21, 573–589.
24. Babel, M.S., Maporn, N., Shinde, V.R., 2014. Incorporating future climatic and socioeconomic factors in water demand forecasting: a case study in Bangkok. Water Resour. Manage. 28, 2049–2062.
25. Babel, M.S., Shinde, V.R., 2011. Identifying prominent explanatory variables for water demand prediction using artificial neural networks: a case study of Bangkok. Water Resour. Manage. 25 (6), 1653–1676.
26. Bachir Sakaa, Hicham Chaffai & Azzedine Hani (2020): ANNs approach to identify water demand drivers for Saf-Saf river basin, Journal of Applied Water Engineering and Research, **DOI: 10.1080/23249676.2020.1719220**
27. Baerenklau, K.A., Schwabe, K.A., Dinar, A., 2014. The residential water demand effect of increasing block rate water budgets. Landsc. Ecol. 90, 683–699.
28. Balling Jr, R C, P Gober, and N Jones. "Sensitivity of Residential Water Consumption to Variations in Climate: An Intraurban Analysis of Phoenix, Arizona." WATER RESOURCES RESEARCH, 44, no. W10401 (2008). **https://doi.org/10.1029/2007WR006722.**
29. Balling, R.C., Cubaque, H.C., 2009. Estimating future residential water consumption in Phoenix, Arizona based on simulated changes in climate. Phys. Geogr. 30, 308–323.
30. Beal C, Makki AA, Stewart RA. Identifying the drivers of water consumption: a summary of results from the South-East Queensland residential end use study, Science Forum and Stakeholder Engagement: Building Linkages Collaboration and Science Quality. Urban Water Secur Res Alliance 2012b:126–32.
31. Beal C, Stewart RA. South East Queensland residential end use study: finalreport. Urban Water Security Research Alliance. Technical Report No. 47; 2011 **http://www.urbanwateralliance.org.au/publications/UWSRA-tr.**
32. Beal CD, Stewart RA, Gardner J, Fielding K, Spinks A, McCrae R. Mind or machine? Examining the drivers of residential water end-use efficiency. J Aust Water Assoc 2013;40(3):66–70.
33. Beal CD, Stewart RA. Identifying residential water end-uses underpinning peak day and peak hour demand. J Water Resour Plan Manage 2013, 1943–5452.0000357, **http://dx.doi.org/10.1061/(ASCE)WR.**
34. Beal, C., Stewart, R., Huang, T., & Rey E., (2011a), Applying smart metering technology to disaggregate residential water end uses in South-East Queensland, Water: Journal of the Australian Water Association, 38(1), 80-84.
35. Belhassen, K, Nouiri, I, Jeridi, A, Ridene, S (2016). « Prévision de la demande en eau potable journalière par les réseaux de neurones artificiels: Cas du système de répartition du grand TunisT unisie ». Conference paper: Premier symposium international de l'association des géographes Tunisien "TANGEO" At: Hammamet Tunisie. **https://www.researchgate.net/publication/311375897.**

36. Bennett, C. Stewart, R.A. and Beal, C.D. (2012) ANN-Based residential water end-use demand forecasting model. Expert Systems with Application, Volume 40, Issue 4, 1014-1023.
37. Billings R.B. and Jones, C.V. Forecasting Urban Water Demand, American Water Works Association, 2011.
38. Billings, B., Jones, C., 2008. Forecasting urban water demand. 2nd edition. American Water works Association, USA.
39. Billings, R.B., Agthe, D.E., 1980. Price elasticities for water: a case of increasing block rates. Land Economics 56 (1), 73–84.
40. Bouchekima B, Bechki D, Bouguettaia H, Boughali S and Tayeb Meftah M, 2008. 'The underground brackish waters in South Algeria: Potential and viable resources'. Laboratoire de Développement des Energies Nouvelles et Renouvelables dans les Zones Arides Sahariennes, Université de Ouargla.
41. Bougadis, John, Kaz Adamowski, and Roman Diduch. "Short-Term Municipal Water Demand Forecasting." Hydrological Processes 19, no. 1 (2005): 137–48. https://doi.org/10.1002/hyp.5763.
42. Breyer, B., Chang, H., 2014. Urban water consumption and weather variation in the Portland, Oregon metropolitan area. Urban Clim. 9, 1–18.
43. Breyer, B., Chang, H., Parandvash, H., 2012. Land-use, temperature and single-family residential water use patterns in Portland, Oregon and Phoenix, Arizona. Appl. Geogr. 35, 142–151.
44. British Geological Survey, 2018. Africa Groundwater Atlas: Hydrogeology of Algeria.
45. Brown, L. R. (1963). Man, land and food. Washington, DC: United States Department of Agriculture.
46. Burney N, Mukhopadhyay A, Al-Mussallam N, Akber A, & Al-Awadi E (2001). Forecasting of Freshwater Demand in Kuwait. The Arabian Journal for Science and Engineering, Volume 26, Number 2B.
47. Caiado, J. (2010). "Performance of combined double seasonal univariate time series models for forecasting water demand." J. Hydrol. Eng. 215–222, **10.1061/(ASCE)HE.1943-5584.0000182.**
48. Campbell HE, Johnson RM and Larson EH. Prices devices people, or rules: The relative effectiveness of policy instruments in water conservation. Review of Policy Research. 2004; 21: 637–662.
49. Chang, F.J., Chang, Y.T., 2006. Adaptive neuro-fuzzy inference system for prediction of water level in reservoir. Advances in Water Resources 29, 1–10.
50. Chang, H., Parandvash, H., Shandas, V., 2010a. Spatial variations of single family residen- tial water consumption in Portland, Oregon. Urban Geogr. 31, 953–972.
51. Chang, H., Praskievicz, S., Parandvash, H., 2014. Sensitivity of urban water consumption to weather climate variability at multiple temporal scales: The case of Portland. Intl. J. Geospatial Environ. Res. 1.

52. Chang, Heejun, Matthew Ryan Bonnette, Philip Stoker, Britt Crow-Miller, and Elizabeth Wentz. "Determinants of Single-Family Residential Water Use across Scales in Four Western Cities." Science of the Total Environment 596–597 (2017): 451–64. https://doi.org/http://dx.doi.org/10.1016/j.scitotenv.2017.03.164.
53. Cheng. B and Titterington. D.M, "Neural Networks: A Review from Statistical Perspective", *Statistical Science*, vol. 9, no. 1, pp. 2-54, 1994.
54. Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
55. Cole, G., Stewart, R.A., 2013. Smart meter enabled disaggregation of urban peak water demand: precursor to effective urban water planning. Urban Water J. 10 (3), 174–194.
56. Cooley H, Christian-Smith J, Gleick PH, Cohen MJ, Heberger M, Ross N, et al. Cal- ifornia's next million acre-feet: saving water, energy, and money. Oakland, September: Pacific Institute; 2010.

57. Coomes, P., Rockway, T., Rivard, J & Kornstein, B. (2010). North American residential water usage trends since 1992. Dever, Co: Water Research Foundation.
58. Danielson, L.E., 1979. An Analysis of Residential Demand for Water Using Micro Time-Series Data. Water Resources Research 15:763-767.
59. Davis LW. Durable goods and residential demand for energy and water: evidence from a field trial. RAND J Econ 2008; 39: 530–46.
60. Délégation de l'union européenne en Algérie (2015). Délégation de l'union européenne en Algérie. Récupéré sur l'union européenne participe à la 1 ère conférence ministérielle du dialogue 5+5 sur l'eau : **http://eeas.europa.ea/delegations/algeria/press_corner//all_news/news/2015/conference_eau_fr.htm.**
61. Domene, E., SaurÍ, D., 2006. Urbanization and water consumption: influencing factors in the Metropolitan Region of Barcelona. Urban Stud. 43, 1605–1623.
62. Donkor E. A., Mazzuchi T. A., Soyer R., and Roberson A., "Urban water demand forecasting: review of methods and models," Journal of Water Resources Planning and Management, vol. 140, no. 2, pp. 146–159, 2014.
63. Dziegielewski, B., Chowdhury, F.J., 2011. Scenario-based forecast of regional water demands in Northeastern Illinois. J. Water Resour. Plann. Manage. 138 (2), 80–89.
64. Endter-Wada, J., J. Kurtzman, S.P. Keenan, R.K. Kjelgren, and C.M.U. Neale, 2008. Situational Waste in Landscape Watering: Residential and Business Water Use in an Urban Utah Community. Journal of the American Water Resources Association 44:902-920.
65. Fan, Liangxin, Lingtong Gai, Yan Tong, and Ruihua Li. "Urban Water Consumption and Its Influencing Factors in China : Evidence from 286 Cities." Journal of Cleaner Production 166 (2017): 124–33. **https://doi.org/10.1016/j.jclepro.2017.08.044.**
66. FAO Aquastat, 2015. Algeria.
67. Farag, F.A., C.M.U. Neale, R.K. Kjelgren, and J. Endter-Wada, 2011. Quantifying Urban Landscape Water Conservation Poten- tial Using High Resolution Remote Sensing and GIS. Photo- grammetric Engineering and Remote Sensing 77:1113-1122.
68. Fayez EL MASRI, (2016), A DATA-BASED MODEL FOR THE DOMESTIC WATER DEMAND IN PALESTINIAN TERRITORY. Lille University of Science and Technology, France. **http://www.theses.fr/2016LIL10101.**
69. Fernandes Neto, M., Naghettini, M., Von Sperling, M., Libaˆnio, M., 2005. Assessing the relevance of intervening parameters on the per capita water consumption rates in Brazilian urban communities. Water Sci. Technol. Water Supply 5, 9–15.
70. Fielding, Kelly S, Sally Russell, Anneliese Spinks, and Aditi Mankad. "Determinants of Household Water Conservation: The Role of Demographic, Infrastructure , Behavior , and Psychosocial Variables." WATER RESOURCES RESEARCH 48 (2012). **https://doi.org/10.1029/2012WR012398.**
71. Firat, M., Güngör, M. Monthly total sediment forecasting using adaptive neuro fuzzy inference system. Stoch Environ Res Risk Assess **24,** 259–270 (2010**). https://doi.org/10.1007/s00477-009-0315-1.**
72. Firat, M., Yurdusev, M.A., Turan, M.E., 2009. Evaluation of artificial neural network techniques for municipal water consumption modeling. Water Resour. Manage. 23 (4), 617–632.
73. Food and Agriculture Organization of the United Nations, 2009. Groundwater Management in Algeria: Draft Synthesis Report.
74. Fox, C., McIntosh, B., Jeffrey, P., 2009. Classifying households for water demand forecasting using physical property characteristics. Land Use Policy 26 (3), 558–568.
75. Franczyk, J., Chang, H., 2009. Spatial analysis of water use in Oregon, USA, 1985–2005. Water Resour. Manage. 23 (4), 755–774.
76. Gabriel, K.R., 1971. The biplot graphic display of matrices with application to principal component analysis. Biometrika 58, 453–467.
77. Gage, E., Cooper, D.J., 2015. The influence of land cover, vertical structure, and socioeco- nomic factors on outdoor water use in a Western US city. Water Resour. Manag. 29, 3877–3890.

78. Gardiner, V., Herrington, P., 1990. Water demand forecasting. 1st edition. Taylor and Francis Ltd, London, UK.
79. Gato S. Forecasting urban residential water demand [PhD Thesis]. School of Civil, Environmental and Chemical Engineering. RMIT University; 2006.
80. Gato S., Jayasuriya N., and Roberts P, “Temperature and rainfall thresholds for base use urban water demand modelling,” Journal of Hydrology, vol. 337, no. 3-4, pp. 364–376, 2007.
81. Gato-Trinidad, Shirley, Niranjali Jayasuriya, and Peter Roberts. “Understanding Urban Residential End Uses of Water.” Water Science & Technology 64.1 (2011): 36–43. **https://doi.org/10.2166/wst.2011.436.**
82. Ghiassi, M, David K Zimbra, and H Saidane. “Urban Water Demand Forecasting with a Dynamic Artificial Neural Network Model.” WATER RESOURCES PLANNING AND MANAGEMENT 134, no. 2 (2008): 138–46. **https://doi.org/10.1061/(ASCE)0733-9496(2008)134:2(138).**
83. Ghiassi, M, F Fa’al, and A Abrishamchi. “Large Metropolitan Water Demand Forecasting Using DAN2, FTDNN, and KNN Models : A Case Study of the City of Tehran , Iran.” Urban Water Journal 14, no. 6 (2017): 655–59. **https://doi.org/10.1080/1573062X.2016.1223858.**
84. Giner, N.M., Polsky, C., Pontius, R.G., Runfola, D.M., 2013. Understanding the social determinants of lawn landscapes: a fine-resolution spatial statistical analysis in suburban Boston, Massachusetts, USA. Landsc. Urban Plan. 111, 25–33.
85. GIZ/BGR/OSS, 2016. *Projet CREM: Etude d’évaluation du secteur de l’eau en Algérie, Etat des Lieux.*
86. Gober, P., Middel, A., Brazel, A., Myint, S., Chang, H., Duh, J.D., House-Peters, L., 2012. Tradeoffs between water conservation and temperature amelioration in Phoenix and Portland: implications for urban sustainability. Urban Geogr. 33, 1030–1054.
87. Grafton RQ, Ward MB, To H and Kompas T. Determinants of residential water consumption: Evidence and analysis from a 10-country household survey. Water Resources Research. 2011; 47(8).
88. Guhathakurta, S. and P. Gober, 2007. The Impact of the Phoenix Urban Heat Island on Residential Water Use. Journal of the American Planning Association 73:317-329.
89. Halper, E.B., Dall'erba, S., Bark, R.H., Scott, C.A., Yool, S.R., 2015. Effects of irrigated parks on outdoor residential water use in a semi-arid city. Landsc. Urban Plan. 134, 210–220.
90. Halper, E.B., Scott, C.A., Yool, S.R., 2012. Correlating vegetation, water use and surface temperature in a semi-arid city: a multiscale analysis of the impacts of irrigation by single-family residences. Geogr. Anal. 44, 235–257.
91. Hamiche A, Stambouli A and Flazi S, 2016. ‘A review on the water and energy sectors in Algeria: Current forecasts, scenario and sustainability issues’. *Renewable and Sustainable Energy Reviews*, 41:261-276.
92. Haque, M.M. et al. Assessing the significance of climate and community factors on urban water demand. Inter- national Journal of Sustainable Built Environment (2015), **http://dx.doi.org/10.1016/j.ijsbe.2015.11.001.**
93. Hardin, G. (1968). The tragedy of the commons. Science, 1968, Vol. 162, no 3859, p. 1234-1248. Sciences, 162 (3859), 1243-1248.
94. Harlan, S.L., S.T. Yabiku, L. Larsen, and A.J. Brazel, 2009. House- hold Water Consumption in an Arid City: Affluence, Affordance, and Attitudes. Society & Natural Resources 22:691-709.
95. Heinrich M. Auckland water use study – monitoring of water end uses. SB10 New Zealand; 2009.
96. Heinrich M. Water end use and efficiency project (WEEP) – final report. BRANZ Study Report 159. New Zealand: Branz, Judge ford; 2007.
97. Hoffmann, M., A. Worthington, and H. Higgs, 2006. Urban Water Demand with Fixed Volumetric Charging in a Large Municipal- ity: The Case of Brisbane, Australia. Australian Journal of Agri- cultural and Resource Economics 50:347-359.

98. House-Peters, Lily A, and Heejun Chang. “Urban Water Demand Modeling: Review of Concepts, Methods, and Organizing Principles.” WATER RESOURCES RESEARCH 47, no. W05401 (2011). **https://doi.org/10.1029/2010WR009624.**
99. Howe, Charles W., and F. P. Linaweaver Jr. 1967. "The Impact of Price on Residential Water Demand and Its Relation to System Design and Price Structure." Water Resources Research 3 (1):13-31.
100. Hurd, B.H., 2006. Water Conservation and Residential Landscapes: Household Preferences, Household Choices. Journal of Agricul- tural and Resource Economics 31:173-192.
101. Hussien, W.A., Memon, F.A. & Savic, D.A. Assessing and Modelling the Influence of Household Characteristics on Per Capita Water Consumption. Water Resour Manage **30,** 2931–2955 (2016). **https://doi.org/10.1007/s11269-016-1314-x.**
102. Ibid.
103. Iyatomi, H., & Hagiwara, M. (2004). Adaptive fuzzy inference neural network. Pattern Recognition, 37(10), 2049–2057.
104. Jacobs H, Haarhoff J. Application of a residential end-use model for estimating cold and hot water demand, wastewater flow and salinity. Water S A 2004a; 30:305–16.
105. Jain, A., and Ormsbee, L. E. (2002). “Short-term water demand forecast modeling techniques: Conventional methods versus AI.” Water Res. Manage., 15(5), 299–321.
106. Jain, A., Varshney, A., Joshi, U., 2001. Short-term water demand forecast modeling at IIT Kanpur using artificial neural networks. Water Resources Management 15 (5), 299–321.
107. Jalala. S, (2005). « Characterizing the Multi-criteria Parameters of Integrated Water Management Model in the Semi-arid Mediterranean Region: Application to Gaza Strip as a case study ». University of Lille for Science and Technology 132-164
108. Jang, J. R., and C. T. Sun. 1995. Neuro-fuzzy modeling and control. In Proceedings of the IEEE 83:378–406. Raleigh, NC, USA
109. Jang, J. S. R. (1993). ANFIS: Adaptive-network-based fuzzy inference systems. *IEEE Transaction Systems, Man and Cybernetics*, Vol. 23, pp. 665-685.
110. Jang, J.S.R., Sun, C.T., Mizutani, E., 1997. Neuro-Fuzzy and Soft Computing. PrenticeHall, ISBN 0-13-261066-3 (607p.).
111. Jonas, H. (1985). The imperative of responsibility: In search of an ethics for the technological age. Chicago: University of Chicago Press.
112. Jorgensen BS, Martin JF, Pearce M, Willis E. Some difficulties and inconsistencies when using habit strength and reasoned action variables in models of metered household water conservation. J Environ Manage 2013a; 115: 124–35.
113. Jorgensen BS, Martin JF, Pearce MW, Willis EM. Predicting household water consumption with individual-level variables. Environ Behav 2013b:1–26, **http://dx.doi.org/10.1177/0013916513482462.**
114. Kant, A., et al. (2013). “Comparison of multi-objective evolutionary neural network, adaptive neuro-fuzzy inference system and bootstrap- based neural network for flood forecasting.” Neural Comput. Appl., 23(S1), 231–246.
115. Kenney, Douglas S., Christopher Goemans, Roberta Klein, Jessica Lowrey, and Kevin Reidy. “Residential Water Demand Management: Lessons from Aurora, Colorado.” Journal of the American Water Resources Association 44, no. 1 (2008): 192–207. **https://doi.org/10.1111/j.1752-1688.2007.00147.x.**
116. Kermani Z. Teshnehlab M (2008) Using adaptive Neuro Fuzzy inference system for hydrological time series. App Soft Comput 8(2008): 928-936.
117. Koffi, Y.B., K.E. Ahoussi, A.M. Kouassi, L.C. Kpangui and J. Biémi, 2012. Modélisation de la consommation en eau potable dans les capitales africaines au sud du sahara: Application des réseaux de neurones formels a la ville de yamoussoukro, capitale politique de la côte d’ivoire. Journal of Asian Scientific Research, 2(10): 562-573.
118. Lee M, Tansel B, Balbin M. Influence of residential water use efficiency measures on household water demand: a four-year longitudinal study. Resour Conserv Recycl 2011; 56:1–6.

119. Lee M, Tansel B. Life cycle-based analysis of demands and emissions for residential water-using appliances. J Environ Manage 2012; 101:75–81.
120. Li W. and Huicheng Z., “Urban water demand forecasting based on HP filter and fuzzy neural network,” Journal of Hydroinformatics, vol. 12, no. 2, pp. 172–184, 2010.
121. Liu, H., Deng, T., & Zhang, H. (2009). Research on forecasting method of urban water demand based on fuzzy theory. In fuzzy systems and knowledge discovery, 2009. FSKD'09. Sixth International Conference on, (Vol. 6, pp. 389–395): IEEE.
122. Liu, J.-Q., Cheng, W.-P., & Zhang, T.-Q. (2012). Principal factor analysis for forecasting diurnal water-demand pattern using combined rough-set and fuzzy-clustering technique. Journal of Water Resources Planning and Management, 139(1), 23–33.
123. Liu, A., Giurco, D., Mukheibir, P., 2016. Urban water conservation through customised water and end-use information. J. Clean. Prod. 112, 3164–3175.
124. Lippmann, R.P. “An introduction to Computing with Neural Nets.” IEEE ASSP Magazine (1987): pp 4-22.
125. Loh M, Coghlan P. Domestic water use study: in Perth, Western Australia, 1998–2001. Western Australia: Water Corporation; 2003.
126. Lyman, R.A., 1992. Peak and off-peak residential water demand. Water Resources Research 28 (9), 2159–2167.
127. Madanat S, Humplick F (1993) A model of household choice of water supply systems in developing countries. Water Resour Res 29(5):1353–1358.
128. Mahbub, P., Ayoko, G.A., Goonetilleke, A., Egodawatta, P., Kokot, S., 2010. Impacts of traffic and rainfall characteristics on heavy metals build-up and wash-off from urban roads. Environ. Sci. Technol. 44 (23), 8904–8910.
129. Makki AA, Stewart RA, Panuwatwanich K, Beal C. Revealing the determinants of shower water end use consumption: enabling better targeted urban water con- servation strategies. J Clean Prod 2013; 60:129–46. https://doi.org/10.1016/j. jclepro.2011.08.007.
130. Makki, Anas A, Rodney A Stewart, Cara D Beal, and Kriengsak Panuwatwanich. “Novel Bottom-up Urban Water Demand Forecasting Model : Revealing the Determinants, Drivers and Predictors of Residential Indoor End-Use Consumption.” “Resources, Conservation & Recycling” 95 (2015): 15–37. https://doi.org/10.1016/j.resconrec.2014.11.009.
131. Makki, Anas A, Rodney A Stewart, Kriengsak Panuwatwanich, and Cara Beal. “Revealing the Determinants of Shower Water End Use Consumption: Enabling Better Targeted Urban Water Conservation Strategies.” Journal of Cleaner Production 60 (2011): 129–46. https://doi.org/10.1016/j.jclepro.2011.08.007.
132. Martins, R. and A. Fortunato, 2007. Residential Water Demand under Block Rates A Portuguese Case Study. Water Policy 9:217-230.
133. Martınez-Espineira, R., 2003. Estimating Water Demand under Increasing-Block Tariffs Using Aggregate Data and Proportions of Users per Block. Environmental & Resource Economics 26:5- 23.
134. Mary E. Renwick and Sandra O. Archibald, 1998. Demand Side Management Policies for Residential Water Use: Who Bears the Conservation Burden? Land Economics, Vol. 74, No. 3, pp. 343-359. http://www.jstor.org/stable/3147117.
135. Matos, Cristina, Carlos A Teixeira, Ricardo Bento, João Varajão, and Isabel Bentes. “An Exploratory Study on the Influence of Socio-Demographic Characteristics on Water End Uses inside Buildings.” Science of the Total Environment 466–467 (2014): 467–74. https://doi.org/10.1016/j.scitotenv.2013.07.036.
136. Mayer PW, DeOreo WB, Towler E, Martien L, Lewis D. Tampa water department residential water conservation study: the impacts of high efficiency plumbing fixture retrofits in single-family homes. A report prepared for Tampa Water Department and the United States Environmental Protection Agency; 2004.
137. Mayer PW, DeOreo WB. Residential end uses of water. American Water Works Association; 1999.

138. Mazzanti, M. and A. Montini, 2006. The Determinants of Residen- tial Water Demand: Empirical Evidence for a Panel of Italian Municipalities. Applied Economics Letters 13:107-111.
139. McCulloch, Warren S, and Walter H Pitts. "A Logical Calculus of The Ideas Immanent In Nervous Activity." Bulltin of Mathematical Biophysics 5 (1943): 115–33.
140. Microsoft [Computer software]. (1996). *Excel*. Redmond, WA: Microsoft Corporation.
141. Millie, D. F., Weckman, G. R., Fahnenstiel, G. L., Carrick, H. J., Ardjmand, E., Young, W. A., et al. (2014). Using artificial intelligence for CyanoHAB niche modeling: discovery and visualization of microcystis–environmental associations within western Lake Erie. Canadian Journal of Fisheries and Aquatic Sciences, 71(11), 1642–1654.
142. Mini, C., Hogue, T.S., Pincetl, S., 2014. Patterns and controlling factors of residential water use in Los Angeles, California. Water Policy 16, 1054–1069.
143. Misra, A.K., 2014. Climate change and challenges of water and food security. Int. J. Sustainable Built Environ. 3 (1), 153–165.
144. Moffat, B., Motlaleng, G. & Thukuza, A. (2011). Households willingness to pay for improved water quality and reliability of supply in Chobe ward, Maun. Botswana Journal of Economics, 8 (12): 45-61.
145. Mohamed M, Al-Mualla A (2010) Water demand forecasting in Umm Al-Quwain using the constant rate model. Desalination 259:161–168.
146. Mu, Xinming, Dale WHITTINGTON, and John Briscoe. "Modeling Village Water Demand Behavior: A Discrete Choice Approach." WATER RESOURCES RESEARCH, 26, no. 4 (1990): 521–29. **https://doi.org/00O43.1397/90/89WR-03 110505.00.**
147. Musolesi, A. and M. Nosvelli, 2007. Dynamics of Residential Water Consumption in a Panel of Italian Municipalities. Applied Economics Letters 14:441-444.
148. Mylopoulos, Y.A., A.K. Mentes, and L. Theodossiou, 2004. Modeling Residential Water Demand Using Household Data: A Cubic Approach. Water International 29:105-113.
149. Naimi Ait Aoudia Meriem, Modélisation, « Modélisation d'un système d'indicateurs pour l'évaluation de la capacité de charge urbaine relative aux ressources hydriques de la wilaya d'Alger », 2016. École polytechnique d'architecture et d'urbanisme, Alger.
150. Nasseri M., Moeini A., and Tabesh M., "Forecasting monthly urban water demand using Extended Kalman Filter and Genetic Programming," Expert Systems with Applications, vol. 38, no. 6, pp. 7387–7395, 2011.
151. Nauck, D., & Kruse, R. (1997). A neuro-fuzzy method to learn fuzzy classification rules from data. Fuzzy Sets and Systems, 89(3), 277–288.
152. Nauges, C. and A. Thomas, 2003. Long-Run Study of Residential Water Consumption. Environmental & Resource Economics 26:25-43.
153. Nauges, C., & Whittington, D. (2009). Estimation of water demand in developing countries: An overview. World Bank Research Observer, 25(2), 263-294. [lkp016]. **https://doi.org/10.1093/wbro/lkp016.**
154. Nauges, C., Thomas, A., 2000. Privately-operated water utilities, municipal price negotiation, and estimation of residential water demand: the case of France. Land Economics 76 (1), 68–85.
155. Nayak, P., Sudheer, K., Rangan, D., Ramasastri, K., 2004. A neuro-fuzzy computing technique for modeling hydrological time series. J. Hydrol. 291 (1–2), 52–66.
156. Nieswiadomy, Michael L., and David J. Molina. 1989. "Comparing Residential Water Demand Estimates Under Decreasing and Increasing Block Rates Using Household Demand Data." Land Economics 65 (Aug.):280-89.
157. Oliveira, D. M., Oliveira, A. L., Neri Nobre, C., and Zarate, L. E. (2009). "The usage of artificial neural networks in the classification and forecast of potable water consumption." Int. Jt. Conf. on Neural Netw., Atlanta, GA, 2331–2338.

158. Ouyang, Yun, Elizabeth A Wentz, Benjamin L Ruddell, and Sharon L Harlan. "A MULTI-SCALE ANALYSIS OF SINGLE-FAMILY RESIDENTIAL WATER USE IN THE PHOENIX METROPOLITAN AREA." JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION AMERICAN 50, no. 2 (2014): 448–67. **https://doi.org/10.1111/jawr.12133.**

159. Pahlavan, R, M Omid, and A Akram. "Application of Data Envelopment Analysis for Performance Assessment and Energy Efficiency Improvement Opportunities in Greenhouses Cucumber Production." J. Agr. Sci. Tech 14 (2012): 1465–75.

160. Panagopoulos, George P. "Assessing the Impacts of Socio-Economic and Hydrological Factors on Urban Water Demand: A Multivariate Statistical Approach." JOURNAL OF HYDROLOGY, 2013. **https://doi.org/10.1016/j.jhydrol.2013.10.036.**

161. Papageorgiou, E. I., Poczęta, K., & Laspidou, C. (2015). Application of fuzzy cognitive maps to water demand pre- diction. In: Fuzzy Systems (FUZZ-IEEE), 2015 I.E. International Conference on, (pp. 1–8): IEEE.

162. Peña-Guzmán, C., Melgarejo, J. and Prat, D. Forecasting Water Demand in Residential, Commercial, and Industrial Zones in Bogotá, Colombia, Using Least-Squares Support Vector Machines. Mathematical problems in engineering, **2016**, 1–10. **http://dx.doi.org/10.1155/2016/5712347.**

163. Pint, E., 1999. Household responses to increased water rates during the California drought. Land Economics 75 (2), 246–266.

164. Planetoscope. (2012). Récupéré sur Planetoscope. Statistiques mondiales en temps rééel : **http://www.planetoscope.com/mortalite/213-nombre-d-enfants-qui-meurent-faute-d-eau-potable.html.**

165. Polebitski, A., Palmer, R.N., Waddell, P., 2011. Evaluating water demands under climate change and transitions in the urban environment. J. Water Resour. Plan. Manag. 137, 249–257.

166. Polebitski, A., Palmer, R.N., 2013. Analysis and predictive models of single-family custom- er response to water curtailments during drought. J. Am. Water Resour. Assoc. 49, 40–51.

167. Prandvash, H., Chang, H., 2016. Analysis of long-term climate change on per capita water demand in urban versus suburban areas in the Portland metropolitan area, USA. J. Hydrol. 538, 574–586.

168. Pulido-calvo, J Roldan, R Lopez-Luque, J.C Gutierrez-Estrada, and I. "Demand Forecasting for Irrigation Water Distribution Systems." JOURNAL OF IRRIGATION AND DRAINAGE ENGINEERING 129, no. 6 (2003): 422–31. **https://doi.org/10.1061/(ASCE)0733-9437(2003)129.**

169. Pulido-Calvo, I., & Gutierrez-Estrada, J. C. (2009). Improved irrigation water demand forecasting using a soft-computing hybrid model. Biosystems Engineering, 102(2), 202–218.

170. Qi C, Chang NB. System dynamics modeling for municipal water demand estima- tion in an urban region under uncertain economic impacts. J Environ Manage 2011;92(6):1628–41.

171. Radhakrishnan. N, URBAN WATER SUPPLY MANAGEMENT: A STUDY OF CENTRALISED AND COMMUNITY BASED SYSTEMS IN CALICUT CITY, 2003, Cochin University Of Science And Technology Cochin-682 022, Kerala p55-65, p112-126

172. Rangel, Hector Rodriguez, Vicenç Puig, Rodrigo Lopez Farias, and Juan J Flores. "Short-Term Demand Forecast Using a Bank of Neural Network Models Trained Using Genetic Algorithms for the Optimal Management of Drinking Water Networks." J. Hydro inform 19 (2017): 1–16. **https://doi.org/10.2166/hydro.2016.199.**

173. Rapport d'Information déposé en application de l'article 145 du Règlement par la Commission des Affaires Étrangères de France en conclusion des travaux d'une mission d'information constituée le 5 octobre 2010 sur La géopolitique de l'eau, 2011.

174. Renwick, M.E., Green, R., 2000. Do residential water demand side management policies measure up? An analysis of eight California water agencies. Journal of Environmental Economics and Management 40 (1), 37–55.

175. Reynaud, A., 2013. Assessing the impact of price and non-price policies on residential water demand: a case study in Wisconsin. Int. J. Water Resour. Dev. 29, 415–433.

176. Rinaudo, Jean-daniel. "Long-Term Water Demand Forecasting." Understanding and Managing Urban Water in Transition 15 (2015): 239–68. https://doi.org/10.1007/978-94-017-9801-3_11.
177. Roberts P. Yarra Valley Water 2004 residential end use measurement study. Final report, June 2005; http://www.yvw.com.au/yvw/groups/public/ documents/document/yvw1001680.pdf.
178. Rumelhart, David E, Hinton Geoffrey E, and Williams Ronald J. "Learning International Representations by Error Propagation," 1985. In: Rumelhart, D. E., McClelland, J. L (Eds.), Parallell Distributed Precessing: MIT Press, Cambridge.
179. Russell S, Fielding K. Water demand management research: a psychological perspective. Water Resour Res 2010;46(5):1–12, W05302, doi:05310.01029/ 02009WR008408.
180. Sadoulet, E. & De Janvry, A. (1995). Quantitative development policy analysis: Johns Hopkins University Press Baltimore.
181. Schleich, J., Hillenbrand, T., 2009. Determinants of residential water demand in Germany. Ecol. Econ. 68 (6), 1756–1769.
182. Sebri, M. Forecasting urban water demand: A meta-regression analysis. J. Environ. Manag. 2016, 183, 777–785.
183. Sen Z, Altunkaynak A (2009) Fuzzy system modeling of drinking water consumption prediction. Espert Syst Appl 36(2009): 11745-11752.
184. Sen, Z., 2001. Fuzzy Logic and Foundation. BKS Publisher, ISBN 9758509233 (in Turkish, 172 p.).
185. Shiken: JALT Testing & Evaluation SIG Newsletter, 1 (1) April 1997 (p. 20-23).
186. Sonmez, Adem Yavuz, Semih Kale, Rahmi Can Ozdemir, and Ali Eslem Kadak. "An Adaptive Neuro-Fuzzy Inference System (ANFIS) to Predict of Cadmium (Cd) Concentrations in the Filyos River, Turkey." Turkish Journal of Fisheries and Aquatic Sciences 18 (2018): 1333–43. https://doi.org/10.4194/1303-2712-v18.
187. Srinivasulu, S., and Jain, A. (2009). "River flow prediction using an inte- grated approach." J. Hydrol. Eng., 10.1061/(ASCE)1084-0699(2009) 14:1(75), 75–83.
188. Suero, F.J., Mayer, P.W., Rosenberg, D.E., 2012. Estimating and verifying United States households' potential to conserve water. J. Water Resour. Plan. Manag. 138, 299–306.
189. Tabesh, M., & Dini, M.M. (2009). FUZZY AND NEURO-FUZZY MODELS FOR SHORT-TERM WATER DEMAND FORECASTING IN TEHRAN. Iranian Journal of Science & Technology, transaction B: Eng 33 (61-77).
190. Teshnehlab M, Shoorehdeli MA, Sedigh AK (2008) Novel hybrid learning algorithms for tuning ANFIS parameters as an identifier using fuzzy PSO. In: Proceedings of 2008 IEEE international conference on networking, sensing and control, ICNSC, pp 111–116. https://doi.org/10.1109/ICNSC.2008.4525193.
191. Tinker A, Bame S, Burt Rand Speed M. Impact of non-behavioral fixed effects on water use: Weather and economic construction differences on residential water use in Austin, Texas. Electronic Green Journal. 2005; 1(22).
192. Tiwari, Mukesh K, and Jan Adamowski. "Urban Water Demand Forecasting and Uncertainty Assessment Using Ensemble Wavelet-Bootstrap-Neural Network Models." WATER RESOURCES RESEARCH, 49, no. September (2013): 6486–6507. https://doi.org/10.1002/wrcr.20517.
193. Turner A, White S, Beatty K, Gregory A, Cubillo F. Results of the largest residential demand management program in Australia. Water Sci Technol: Water Supply 2005; 5:249–56.
194. Turner A, Hausler G, Carrard N, Kazaglis A, White S, Hughes A, et al. Review of water supply-demand options for South East Queensland. Sydney and Cardno, Brisbane: Institute for Sustainable Futures; 2007.
195. Turner A, Fyfe J, Retamal M, White S, Coates A. The one to one water savings program unpacking residential high-water usage. In: IWA efficient 09 conference; 2009.
196. UNESCO. "Global Water Resources under Increasing Pressure from Rapidly Growing Demands and Climate Change, According to New UN World Water Development Report." Villa la Colombella, Perugia, Italy, n.d. www.unesco.org/water/wwap.

197. United Nations (2013). World Population Prospects: The 2012 Revision (Vol. Volume I: Comprehensive Tables). New York: United Nations.
198. Vijayalaksmi, D., & Babu, K. J. (2015). Water supply system demand forecasting using adaptive neuro-fuzzy inference system. Aquatic Procedia, 4, 950–956.
199. Walton C, Holmes K. How much water efficiency does $321 million buy? In: International Water Association (IWA), Australian Water Association, Sydney, Australia, editors. Proceedings of the 5th IWA specialist conference, efficient 2009.
200. Wang, W.-C., Chau, K.-W., Cheng, C.-T., & Qiu, L. (2009). A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. Journal of Hydrology, 374(3-4), 294-306.
201. Water Corporation. Perth residential water use study 2008/2009. West- ern Australia: Water Forever, Water Corporation;2011**http://www.water.wa.gov.au/PublicationStore/first/98576.pdf.**
202. Wentz, E.A., Gober, P., 2007. Determinants of small-area water consumption for the City of Phoenix, Arizona. Water Resour. Manag. 21, 1849–1863.
203. Wentz, Elizabeth A, Angela J Wills, Won Kyung Kim, Soe W Myint, Gober Patricia, and Robert C Balling Jr. "Factors Influencing Water Consumption in Multifamily Housing in Tempe, Arizona." The Professional Geographer 66, no. 3 (2014): 501–10. **https://doi.org/10.1080/00330124.2013.805627.**
204. Wentz, E.A., Rode, S., Li, X., Tellman, E.M., Turner, B.L., 2016. Impact of Homeowner Association (HOA) landscaping guidelines on residential water use. Water Resour. Res. 52, 3373–3386.
205. Willis, R, R A Stewart, K Panuwatwanich, B Capati, and D Giurco. "GOLD COAST DOMESTIC WATER END USE STUDY." Water 36, no. 6 (2009): 79–85.
206. Willis RM. Domestic water end use study: an investigation of the water savings attributed to demand management strategies and dual reticulated recy- cled water systems [PhD Thesis]. School of Engineering. Griffith University; 2011.
207. Willis RM, Stewart RA, Williams P, Hacker C, Emmonds S, Capati G. Residential potable and recycled water end uses in a dual reticulated supply system. Desali- nation 2011b; 272:201–11.
208. Willis R. M., Stewart R. A., Giurco D. P., Talebpour M. R. & Mousavinejad A (2013), End use water consumption in households: impact of socio-demographic factors and efficient devices: Journal of cleaner production, 60 107-115, **https://doi.org/10.1016/j.jclepro.2011.08.006.**
209. Wilson, A., Boehland, J., 2005. Small is beautiful US house size, resource use, and the en- vironment. J. Ind. Ecol. 9, 277–287.
210. Worldwide Governance Indicators (2014). Acces Governance Indicators. Récupéré sur Worldwide Governance Indicators: **http://info.worldbank.org/governance/wgi/index.aspx#home.**
211. Xue, Peng, Tianzhen Hong, Bing Dong, and Cheuk Ming Mak. "A Preliminary Investigation of Water Usage Behavior in Single-Family Homes." Energy Technologies Area, 2017.
212. Yurdusev MA, Firat M (2009) Adaptive neuro fuzzy inference system approach for municipal water consumption modeling: an application to Izmir, Turkey. J Hydrol 365: 225–234.
213. Yurdusev, M. A., Firat, M., Mermer, M., & Turan, M. E. (2009). Water use prediction by radial and feed-forward neural nets. In: Proceedings of the Institution of Civil Engineers-Water Management (vol. 162, pp. 179–188, Vol. 3): Thomas Telford Ltd.
214. Yurdusev MA, Firat M, Turan ME (2010) Generalized regression neural networks for municipal water consumption prediction. J Stat Comput Simul 80:477–478.
215. Zadeh L A. Fuzzy sets. Inform. Contr. 8:338-53, 1965. [Dept. Electrical Engineering and Electronics Res. Lab., Univ. California, Berkeley, CA]
216. Zadeh, L. A. (1994). Fuzzy logic, neural networks, and soft computing. Communications of the ACM, 37(3), 77– 85.
217. Zhou, J., & Yang, K. (2010). General regression neural network forecasting model based on PSO algorithm in water demand. In: Knowledge Acquisition and Modeling (KAM), 2010 3rd International Symposium on, (pp. 51–54): IEEE.

Web site

1. http://www.ana.gov.br/.
2. https://en.climate-data.org/africa/algeria/souk-ahras/sedrata-45684/#climate-table.
3. http://www.salaryexplorer.com/.
4. https://en.climate-data.org/.
5. http://www.ons.dz/.
6. https://en.wikipedia.org/wiki/2018.
7. https://fr.wikipedia.org/wiki/Statistica.
8. https://fr.wikipedia.org/wiki/SPSS.
9. https://en.wikipedia.org/wiki/MATLAB.
10. https://water.fanack.com/algeria/water-use/.
11. IWMI: www.iwmi.org.
12. http://www.iwmi.cgiar.org/About_IWMI/Strategic_Documents/Annual_Reports/1999-2000/Scientific-WS2025.pdf.
13. https://water.fanack.com/algeria/water-resources/#_ftn1.
14. http://bourse-dz.com/ressources-en-eau-le-plan-necib-pour-2030/.

Annexes

Annexe 01 : *Fiche d'enquête*

Évaluer ma consommation d'eau : Nom **(الاسم)** :

Vous souhaitez avoir une estimation de votre consommation d'eau ?Savez-vous combien vous consommez d'eau au quotidien?Le numéro de maison **(رقم المنزل)**: Le code de maison :

RUE(الحي) :Cite 228 LOGTS **Cite 176 LOGTS** **Cite Saada**

Cite El AmeriaCite 12 LOGTS

Caractéristiques physiques des logements :

- Type de logements**(نوع المنزل)** : M. individuelle **(منزل ارضي)** A. bâtiment **(في عمارة)**
- Surface de logement **(مساحة المنزل)**:
- Jardin **(حديقة)** Non **(لا)**: Oui **(نعم)** : Surface **(المساحة)**
- Combien de fois par mois arrosez-vous votre jardin? **(كم عدد المرات تسقي الحديقة)**
- Piscine **(مسبح)**: Non **(لا)** Oui **(نعم)** Volume **(حجم المسبح)**
- Nombre des chambres **(عدد الغرف)**

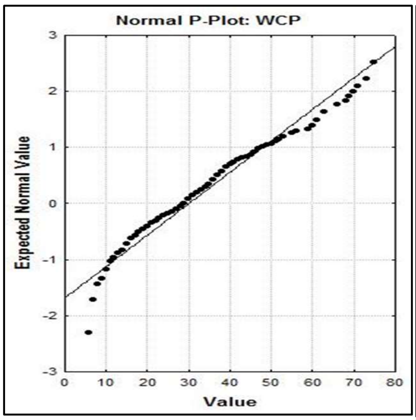
Indicateurs socioéconomiques :

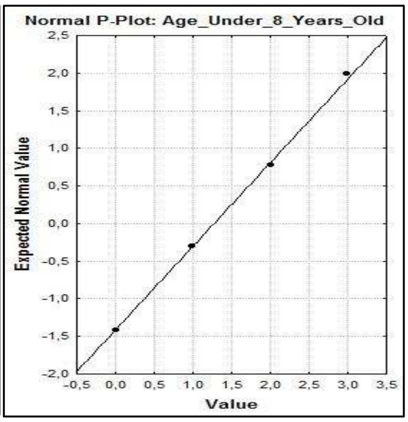
- Combien de personnes compte votre foyer?**(عدد أفراد الأسرة)** Male**(ذكور)** Femelle**(إناث)**
- Age **(العمر)** : ≤8 9-15 15-35 ≥35
- Niveau d'éducation **(المستوى التعليمي)** : Primaire **(ابتدائي)** moyen **(متوسط)** Lycée**(ثانوي)** Universitaire **(جامعي)**
- Revenu familial mensuel**(الدخل العائلي الشهري DA)** :
- Combien de voiture avez-vous dans votre foyer **(عدد السيارات)**:
- Combien de fois par an lavez-vous votre voiture? **(كم مرة في العام تغسل السيارة)**
- Combien de fois par jour lavez-vous votre vaisselle à main **(كم مرة في اليوم تغسل الأواني)**?
- Combien de lessives par semaine faites-vous? **(كم مرة في الأسبوع تغسل الثياب)**
- Combien de fois par jour en moyenne une personne de votre foyer tire la chasse ?**(كم مرة في اليوم يستخدم أفراد أسرتك المرحاض)**
- Dans votre foyer vous prenez combien de bains (prenez-vous en moyenne par semaine pour l'ensemble de votre foyer)?**(كم مرة في الأسبوع يستحم أفراد أسرتك في الحوض)**
- Dans votre foyer, les hommes utilisent-ils plus d'eau que les femmes ? **(هل الرجال يستهلكوا الماء اكثر من النساء)**Oui **(نعم)** Non**(لا)** même quantité**(نفس الكمية)**
- Combien de douches par semaine faites-vous? **(كم مرة في الأسبوع يستحم أفراد أسرتك)**
  Male **(ذكور)** Femelle **(إناث)**

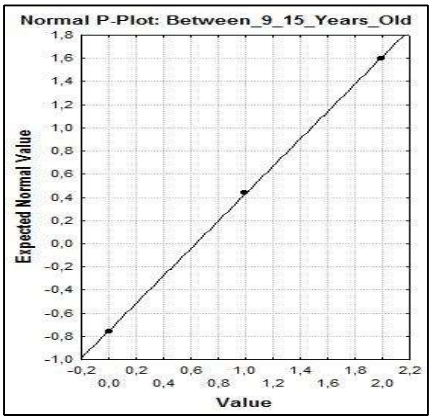
ANNEX 02: Statistic description of variables

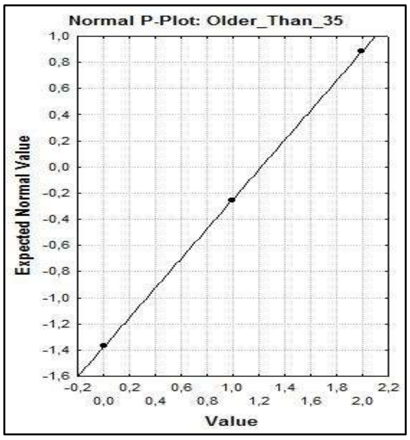
Variable	Acronym	Minimum	Maximum	Mean	Std. Deviation	Skewness	Kurtosis
Household Water Consumption (m^3)	WCP	6	75	30.09	17.19	0.62	-0.30
Socio-economic indicator (SEP)							
Family Composition and Gender							
Household size	HOUS	2	8	5	1.62	-0.37	-0.71
Number of Female	FEM	1	6	3	1.26	0.24	-0.51
Number of Male	MAL	0	5	2	0.89	0.38	-0.16
Age Categories							
Under 8 years old	AG1	0	3	1	0.78	0.05	-0.47
Between 9 to 15 years old	AG2	0	2	1	0.67	0.51	-0.73
Between 15 to 35 years old	AG3	1	4	2	1.12	0.46	-1.21
Older than 35 years old	AG4	0	2	1	0.71	-0.32	-0.98
Education level							
Primary School	PRS	0	2	1	0.58	0.73	-0.44
Medium School	MDS	0	2	1	0.67	0.22	-0.78
High School	HGS	1	3	1	0.77	0.81	-0.86
University	UNIV	0	2	1	0.73	0.09	-0.11
Household Income (DA)	INC	350 00	11 000 0	53905.47	19976	1.09	-0.13
Car Possession							
Number of Cars	CARN	0	3	2	0.69	-0.33	-0.72
Washing Cars frequency (month)	WCAR	0	4	2	1.19	0.85	-0.05
Indoor Habits (INH)							
Clothes wash frequency (week)	WCL	1	4	2	0.69	0.84	0.87
Dishwashing frequency (day)	WDISH	1	3	3	0.61	-1.13	0.24
Toilets use frequency (day)	UTLT	3	7	4	0.95	0.62	0.38
Shower frequency for Female (week)	FSHW	1	7	2	1	0.59	1.16
Shower frequency for Male (week)	MSHW	1	5	2	0.95	0.53	-0.52
Physical Characteristics of Buildings (PHC)							
Total Area (m2)	TAR	80	320	186.6	77.29	0.15	-1.27
Building Area (m2)	BAR	40	302	164.9	75.81	0.19	-1.19
Number of Rooms	ROMN	2	13	6	3.09	0.92	-0.20
Garden Possession							
Garden Area	GAR	2	80	21.7	18.14	**2.16**	**4.11**
Garden Watering frequency	GWAT	1	4	2	0.79	0.56	-0.22

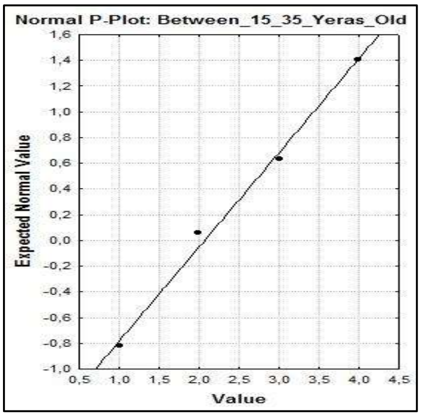
ANNEX 03: **P-Plots of variables**

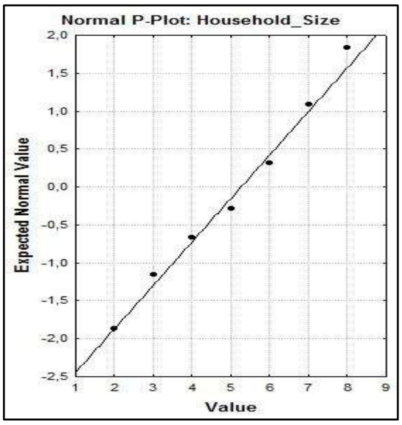


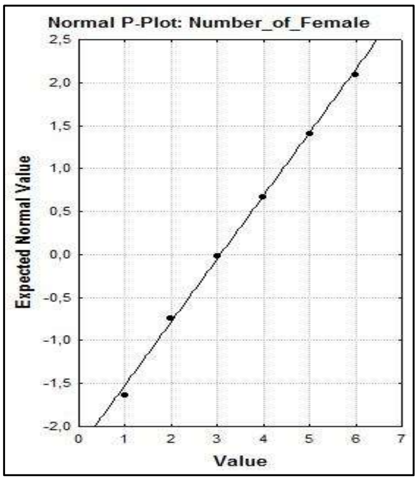


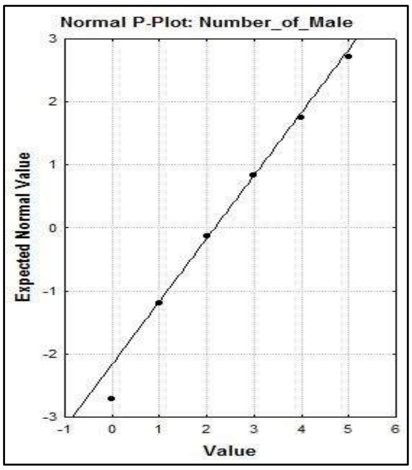


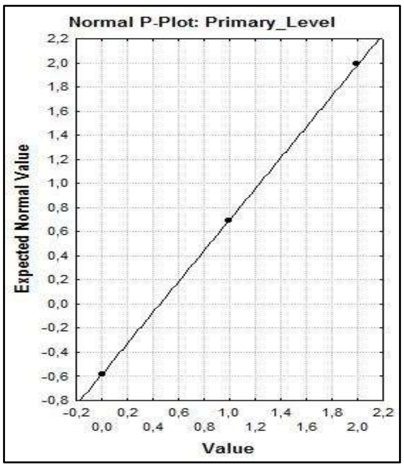


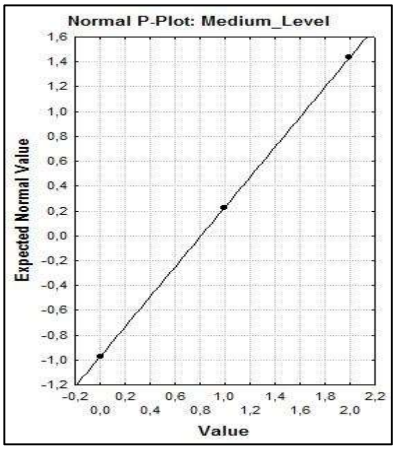


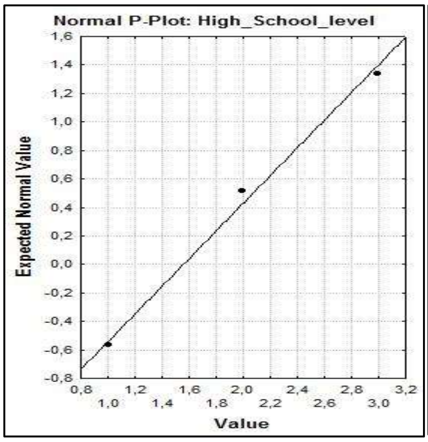


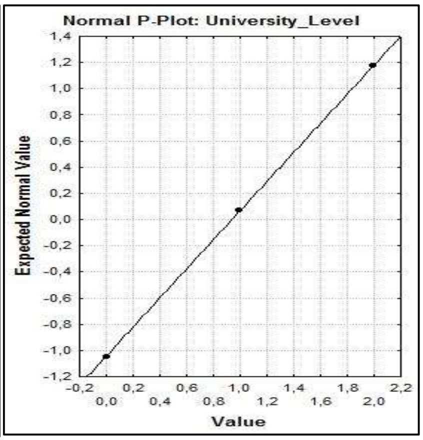


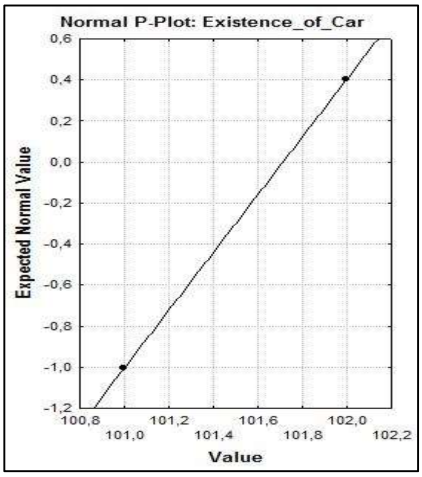


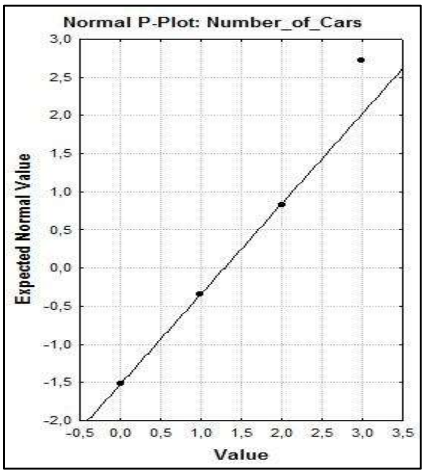


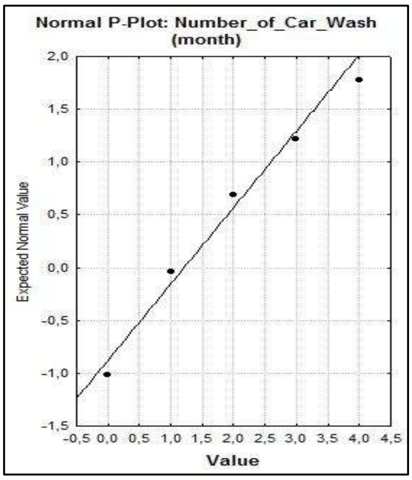

Normal P-Plot: Medium_Level
Expected Normal Value
Value

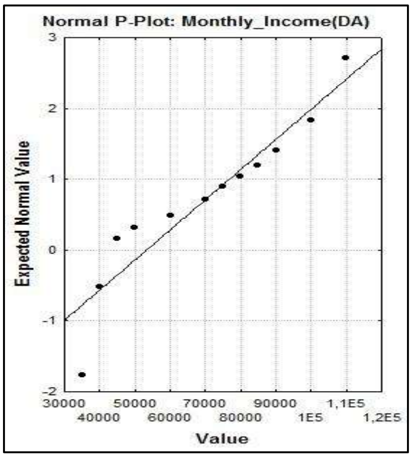

Normal P-Plot: High_School_level
Expected Normal Value
Value

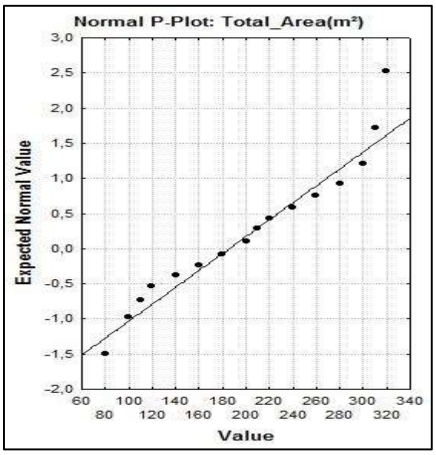

Normal P-Plot: University_Level
Expected Normal Value
Value

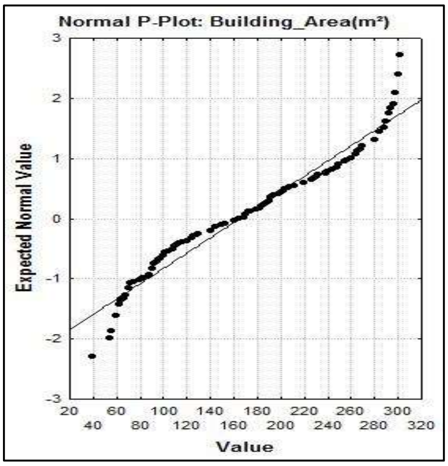

Normal P-Plot: Existence_of_Car
Expected Normal Value
Value

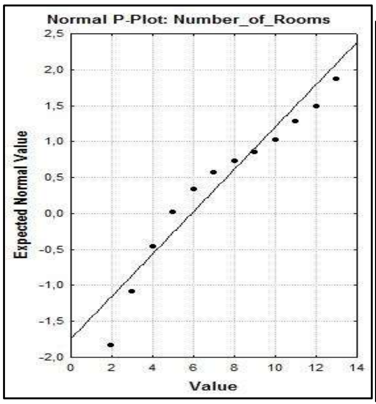

Normal P-Plot: Number_of_Cars
Expected Normal Value
Value

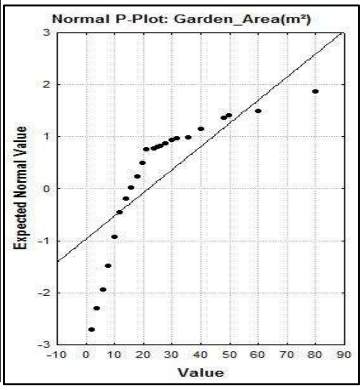

Normal P-Plot: Number_of_Car_Wash (month)
Expected Normal Value
Value

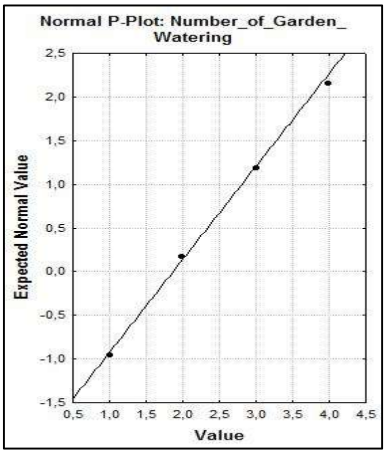

Normal P-Plot: Monthly_Income(DA)
Expected Normal Value
Value


Normal P-Plot: Total_Area(m²)
Expected Normal Value
Value


Normal P-Plot: Building_Area(m²)
Expected Normal Value
Value


Normal P-Plot: Number_of_Rooms
Expected Normal Value
Value


Normal P-Plot: Garden_Area(m²)
Expected Normal Value
Value


Normal P-Plot: Number_of_Garden_ Watering
Expected Normal Value
Value

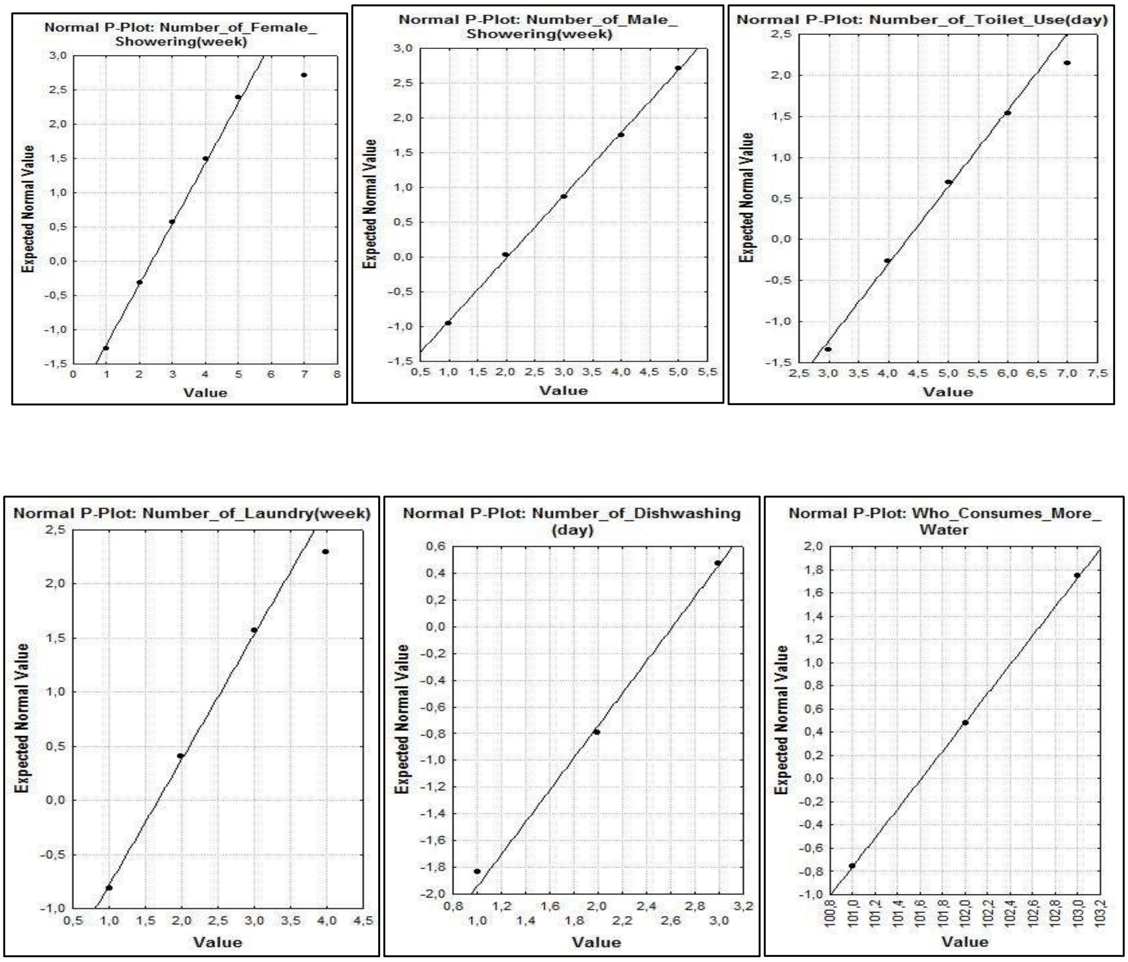
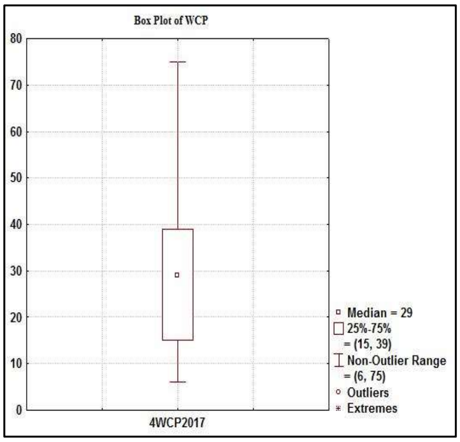
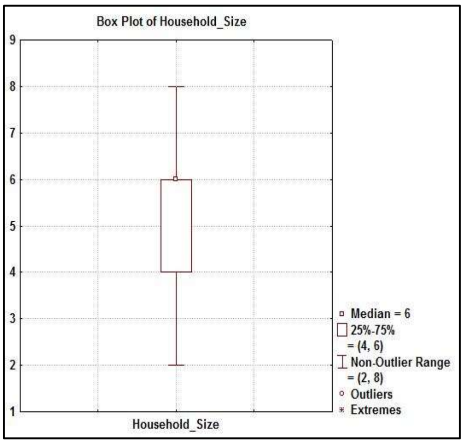
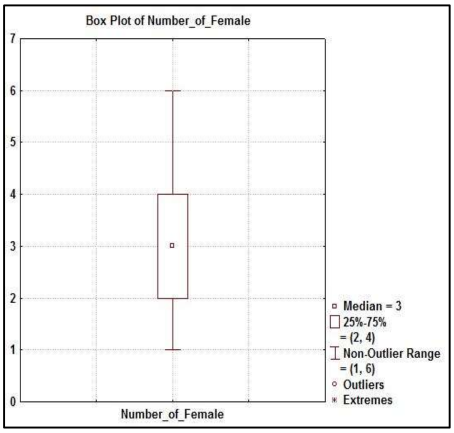


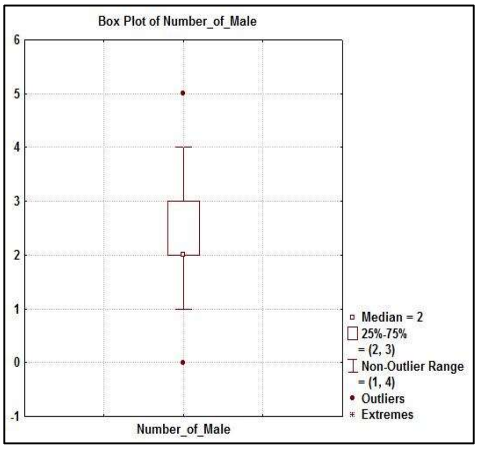
Fig 3.36 P-P plots for variables

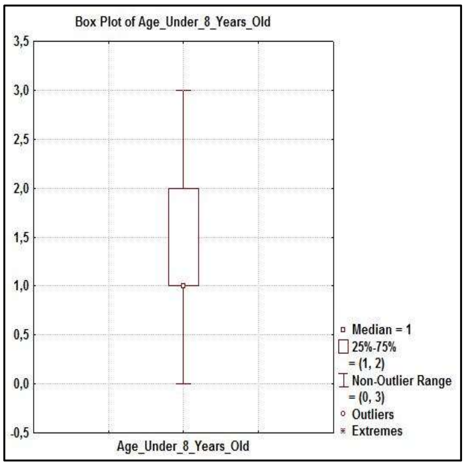
ANNEX 04: **Box Plots of variables**

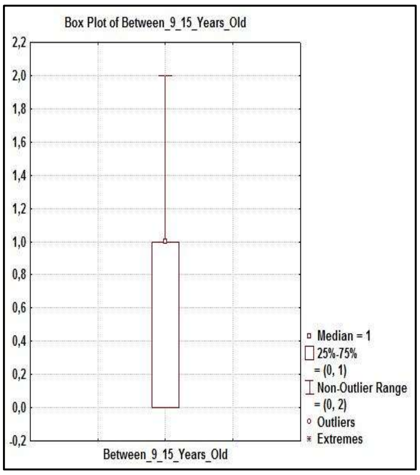


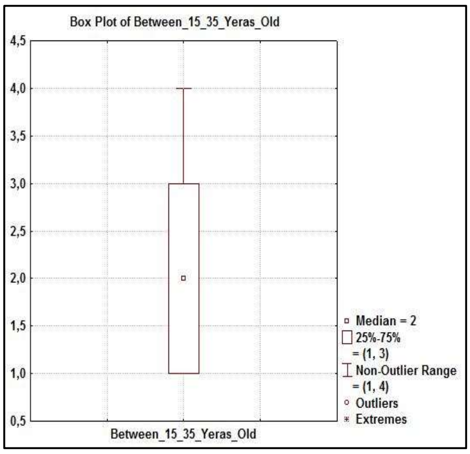


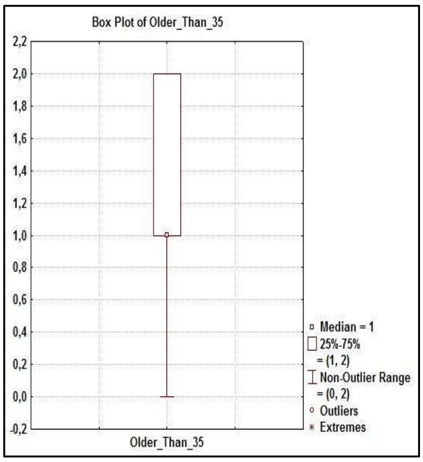


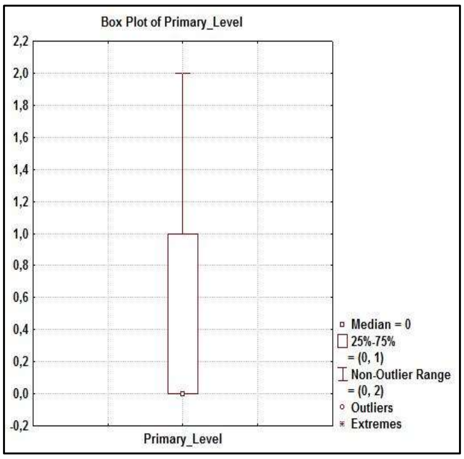


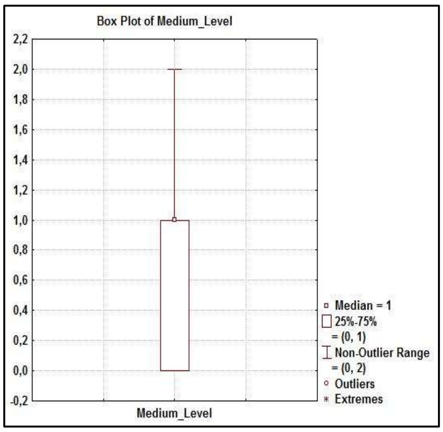


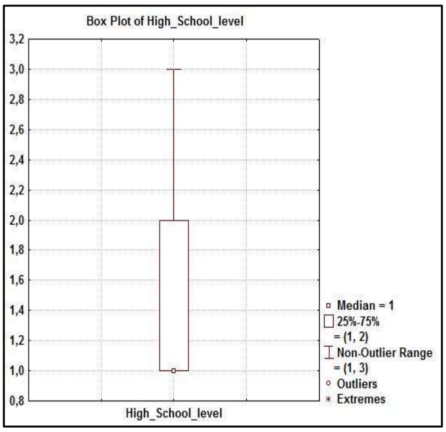


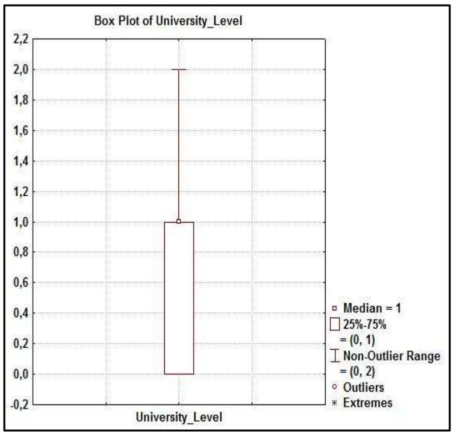


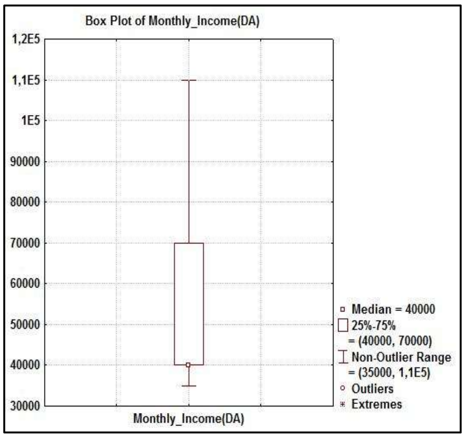


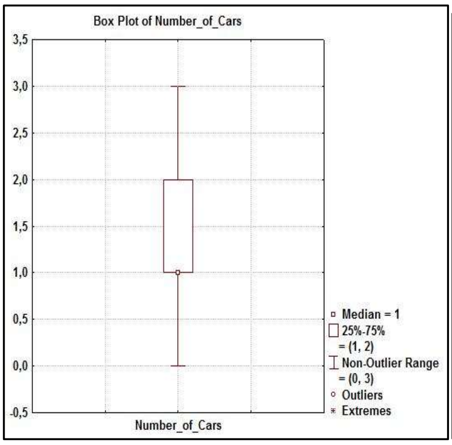


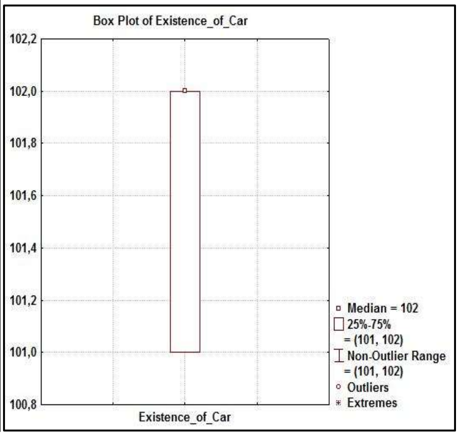

Box Plot of Medium_Level
Median = 1
25%-75% = (0, 1)
Non-Outlier Range = (0, 2)
Outliers
Extremes
Medium_Level

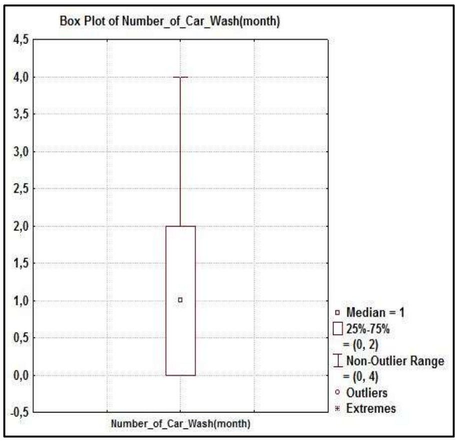

Box Plot of High_School_level
Median = 1
25%-75% = (1, 2)
Non-Outlier Range = (1, 3)
Outliers
Extremes
High_School_level

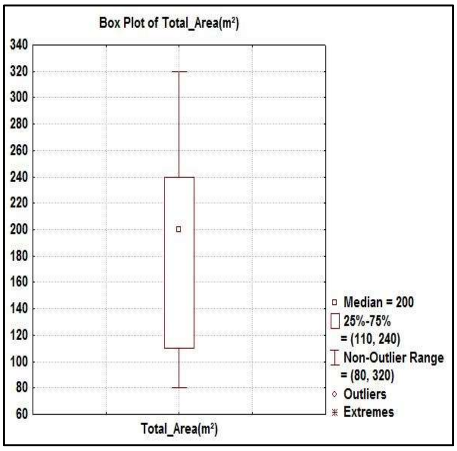

Box Plot of University_Level
Median = 1
25%-75% = (0, 1)
Non-Outlier Range = (0, 2)
Outliers
Extremes
University_Level

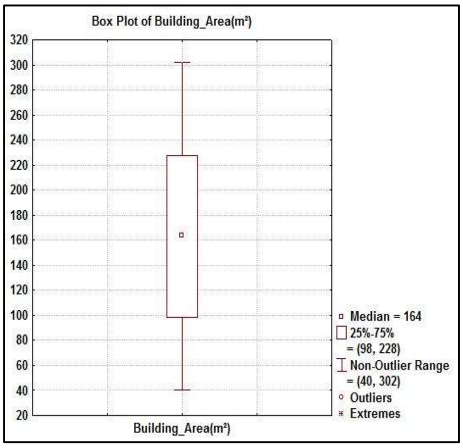

Box Plot of Monthly_Income(DA)
Median = 40000
25%-75% = (40000, 70000)
Non-Outlier Range = (35000, 1,1E5)
Outliers
Extremes
Monthly_Income(DA)

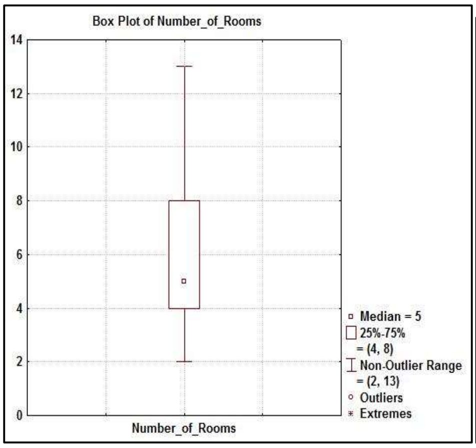

Box Plot of Number_of_Cars
Median = 1
25%-75% = (1, 2)
Non-Outlier Range = (0, 3)
Outliers
Extremes
Number_of_Cars

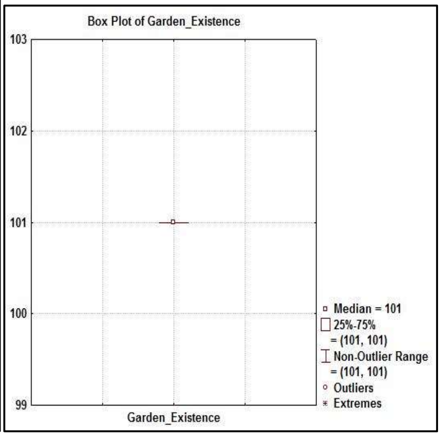

Box Plot of Existence_of_Car
Median = 102
25%-75% = (101, 102)
Non-Outlier Range = (101, 102)
Outliers
Extremes
Existence_of_Car

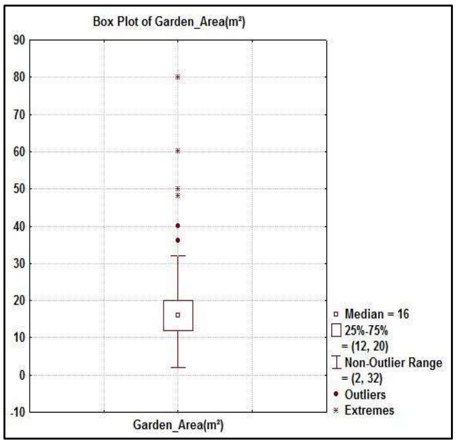

Box Plot of Number_of_Car_Wash(month)
Median = 1
25%-75% = (0, 2)
Non-Outlier Range = (0, 4)
Outliers
Extremes
Number_of_Car_Wash(month)


Box Plot of Total_Area(m²)
Median = 200
25%-75% = (110, 240)
Non-Outlier Range = (80, 320)
Outliers
Extremes
Total_Area(m²)


Box Plot of Building_Area(m²)
Median = 164
25%-75% = (98, 228)
Non-Outlier Range = (40, 302)
Outliers
Extremes
Building_Area(m²)


Box Plot of Number_of_Rooms
Median = 5
25%-75% = (4, 8)
Non-Outlier Range = (2, 13)
Outliers
Extremes
Number_of_Rooms


Box Plot of Garden_Existence
Median = 101
25%-75% = (101, 101)
Non-Outlier Range = (101, 101)
Outliers
Extremes
Garden_Existence


Box Plot of Garden_Area(m²)
Median = 16
25%-75% = (12, 20)
Non-Outlier Range = (2, 32)
Outliers
Extremes
Garden_Area(m²)

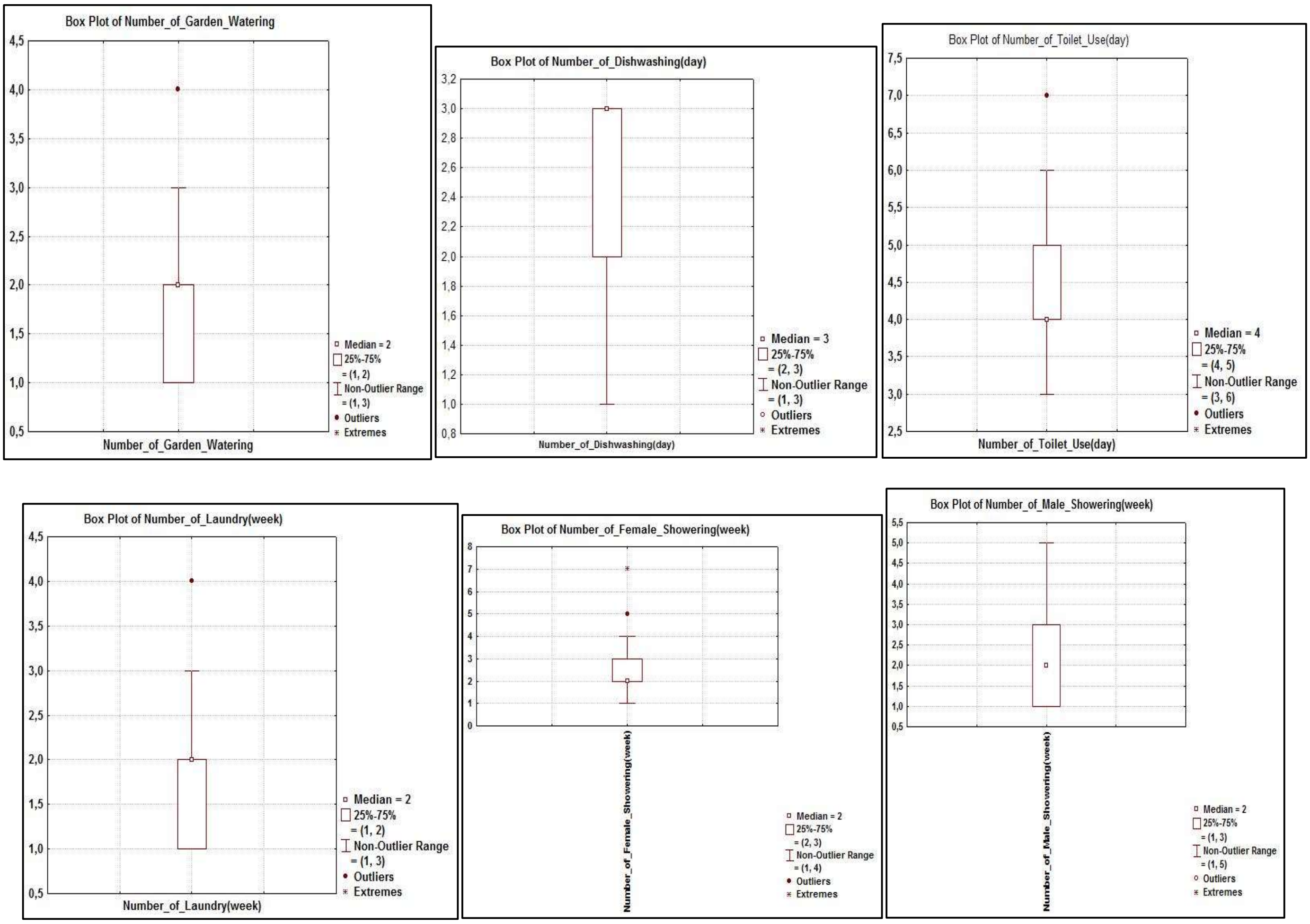


Fig 3.37 Box plot for variables

ANNEX 04: Description for groups of variables

Table 5.10: Description for groups of variables

WCP (m³)									
Variables	Groups	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
HOUS Descriptives	2	12	6,92	1,084	,313	6,23	7,61	6	10
	3	25	14,80	10,840	2,168	10,33	19,27	7	41
	4	27	19,41	9,704	1,868	15,57	23,25	10	38
	5	27	23,59	12,595	2,424	18,61	28,58	15	60
	6	68	30,88	5,321	,645	29,59	32,17	21	42
	7	29	52,03	7,872	1,462	49,04	55,03	42	66
	8	13	63,46	9,386	2,603	57,79	69,13	52	75
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75
FEM Descriptives	1	20	12,95	10,364	2,318	8,10	17,80	6	41
	2	52	18,50	10,216	1,417	15,66	21,34	7	39
	3	53	28,38	11,806	1,622	25,12	31,63	10	60
	4	51	35,18	9,022	1,263	32,64	37,71	18	50
	5	18	57,11	7,020	1,655	53,62	60,60	41	66
	6	7	71,57	2,820	1,066	68,96	74,18	68	75
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75
MAL Descriptives	0	1	10,00	.	.	.	.	10	10
	1	44	15,45	11,553	1,742	11,94	18,97	6	42
	2	90	33,54	19,108	2,014	29,54	37,55	7	75
	3	51	35,14	11,393	1,595	31,93	38,34	20	56
	4	14	37,57	9,990	2,670	31,80	43,34	21	52
	5	1	21,00	.	.	.	.	21	21
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75
AG1 Descriptives	0	31	15,06	12,995	2,334	10,30	19,83	6	43
	1	92	24,61	12,935	1,349	21,93	27,29	10	52
	2	69	39,26	11,963	1,440	36,39	42,13	25	63
	3	9	67,56	10,418	3,473	59,55	75,56	41	75
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75
AG2 Descriptives	0	90	32,38	19,443	2,049	28,31	36,45	6	75
	1	89	29,63	16,003	1,696	26,26	33,00	7	63
	2	22	22,59	7,507	1,600	19,26	25,92	14	43
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75
AG3 Descriptives	1	83	16,63	7,071	,776	15,08	18,17	6	41
	2	44	27,73	14,205	2,141	23,41	32,05	7	60
	3	42	45,50	15,063	2,324	40,81	50,19	32	75
	4	32	48,03	6,832	1,208	45,57	50,49	38	61
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75
AG4 Descriptives	0	34	24,41	13,701	2,350	19,63	29,19	7	41
	1	92	32,11	17,798	1,856	28,42	35,79	6	66
	2	75	30,19	17,476	2,018	26,17	34,21	10	75
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75

PRS Descriptives	0	112	19,63	10,771	1,018	17,62	21,65	6	47
	1	80	41,06	13,527	1,512	38,05	44,07	13	75
	2	9	62,67	6,423	2,141	57,73	67,60	55	75
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75
MDS Descriptives	0	66	30,30	21,584	2,657	25,00	35,61	6	75
	1	105	31,17	15,840	1,546	28,11	34,24	7	63
	2	30	25,83	8,396	1,533	22,70	28,97	14	43
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75
HGS Descriptives	1	115	20,39	8,608	,803	18,80	21,98	6	41
	2	50	33,20	14,666	2,074	29,03	37,37	7	60
	3	36	56,75	9,584	1,597	53,51	59,99	45	75
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75
UNIV Descriptives	0	59	13,73	8,965	1,167	11,39	16,07	6	39
	1	94	29,43	10,892	1,123	27,19	31,66	15	66
	2	48	51,50	10,875	1,570	48,34	54,66	38	75
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75
TAR Descriptives	80	27	8,04	1,506	,290	7,44	8,63	6	10
	100	12	11,75	1,055	,305	11,08	12,42	10	13
	110	15	14,87	,516	,133	14,58	15,15	14	16
	120	12	17,58	,793	,229	17,08	18,09	16	19
	140	10	20,40	,516	,163	20,03	20,77	20	21
	160	12	23,75	1,055	,305	23,08	24,42	22	25
	180	12	28,58	4,295	1,240	25,85	31,31	26	42
	200	17	30,88	2,998	,727	29,34	32,42	29	42
	210	12	33,17	1,030	,297	32,51	33,82	32	35
	220	10	35,80	,422	,133	35,50	36,10	35	36
	240	12	37,67	,888	,256	37,10	38,23	36	39
	260	9	43,89	10,971	3,657	35,46	52,32	39	73
	280	11	47,73	7,485	2,257	42,70	52,76	43	70
	300	15	53,53	5,718	1,476	50,37	56,70	48	71
	310	13	62,92	3,013	,836	61,10	64,74	60	69
	320	2	75,00	,000	,000	75,00	75,00	75	75
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75
BAR Descriptives	40	4	9,00	1,414	,707	6,75	11,25	7	10
	54	1	10,00	.	.	.	.	10	10
	56	2	7,50	,707	,500	1,15	13,85	7	8
	60	7	8,57	2,149	,812	6,58	10,56	6	12
	62	3	6,67	1,155	,667	3,80	9,54	6	8
	64	1	7,00	.	.	.	.	7	7
	66	1	7,00	.	.	.	.	7	7
	68	2	9,50	,707	,500	3,15	15,85	9	10
	70	7	8,86	2,268	,857	6,76	10,95	7	13
	72	1	6,00	.	.	.	.	6	6
	74	1	10,00	.	.	.	.	10	10
	80	2	17,50	7,778	5,500	-52,38	87,38	12	23
	82	1	14,00	.	.	.	.	14	14
	86	1	11,00	.	.	.	.	11	11
	88	3	12,33	,577	,333	10,90	13,77	12	13
	90	8	13,13	1,959	,693	11,49	14,76	10	15
	92	2	16,00	1,414	1,000	3,29	28,71	15	17

	94	2	14,50	,707	,500	8,15	20,85	14	15
	96	1	15,00	.	.	.	.	15	15
	98	1	15,00	.	.	.	.	15	15
	100	6	18,00	4,817	1,966	12,95	23,05	15	27
	102	1	15,00	.	.	.	.	15	15
	104	1	18,00	.	.	.	.	18	18
	108	5	18,20	1,924	,860	15,81	20,59	16	21
	110	2	17,50	,707	,500	11,15	23,85	17	18
	112	2	17,50	,707	,500	11,15	23,85	17	18
	114	1	18,00	.	.	.	.	18	18
	116	1	18,00	.	.	.	.	18	18
	120	3	26,33	4,619	2,667	14,86	37,81	21	29
	124	4	20,25	,500	,250	19,45	21,05	20	21
	126	1	20,00	.	.	.	.	20	20
	128	1	20,00	.	.	.	.	20	20
	130	3	29,00	6,928	4,000	11,79	46,21	21	33
	140	5	28,80	6,686	2,990	20,50	37,10	22	36
	144	4	23,25	,500	,250	22,45	24,05	23	24
	150	2	24,50	,707	,500	18,15	30,85	24	25
	152	2	24,00	1,414	1,000	11,29	36,71	23	25
	160	5	32,20	4,382	1,960	26,76	37,64	28	39
	164	1	28,00	.	.	.	.	28	28
	168	3	32,00	8,660	5,000	10,49	53,51	27	42
	170	5	28,20	2,168	,970	25,51	30,89	27	32
	172	1	30,00	.	.	.	.	30	30
	174	1	26,00	.	.	.	.	26	26
	180	3	33,67	4,726	2,728	21,93	45,41	30	39
	182	1	30,00	.	.	.	.	30	30
	184	4	30,50	1,291	,645	28,45	32,55	29	32
	186	1	42,00	.	.	.	.	42	42
	188	1	29,00	.	.	.	.	29	29
	190	6	31,17	1,329	,543	29,77	32,56	30	33
	192	3	35,00	3,606	2,082	26,04	43,96	31	38
	194	1	32,00	.	.	.	.	32	32
	198	2	34,50	,707	,500	28,15	40,85	34	35
	200	1	36,00	.	.	.	.	36	36
	202	3	34,67	1,155	,667	31,80	37,54	34	36
	204	2	36,00	,000	,000	36,00	36,00	36	36
	206	1	35,00	.	.	.	.	35	35
	208	1	36,00	.	.	.	.	36	36
	212	2	36,50	2,121	1,500	17,44	55,56	35	38
	220	5	43,40	5,899	2,638	36,08	50,72	38	51
	226	1	37,00	.	.	.	.	37	37
	228	2	36,50	,707	,500	30,15	42,85	36	37
	230	3	45,33	12,741	7,356	13,68	76,98	37	60
	232	1	38,00	.	.	.	.	38	38
	238	1	37,00	.	.	.	.	37	37
	239	1	39,00	.	.	.	.	39	39
	240	1	47,00	.	.	.	.	47	47
	244	2	59,50	19,092	13,500	-112,03	231,03	46	73
	248	3	40,00	1,000	,577	37,52	42,48	39	41

	250	3	51,00	10,000	5,774	26,16	75,84	41	61
	255	1	70,00	.	.	.	.	70	70
	256	1	41,00	.	.	.	.	41	41
	260	3	48,67	5,508	3,180	34,99	62,35	45	55
	264	3	46,00	1,000	,577	43,52	48,48	45	47
	266	1	43,00	.	.	.	.	43	43
	268	2	45,00	1,414	1,000	32,29	57,71	44	46
	270	2	61,50	2,121	1,500	42,44	80,56	60	63
	280	6	57,67	10,482	4,279	46,67	68,67	48	75
	284	2	51,00	2,828	2,000	25,59	76,41	49	53
	288	2	63,50	10,607	7,500	-31,80	158,80	56	71
	290	3	60,67	2,082	1,202	55,50	65,84	59	63
	292	2	55,50	7,778	5,500	-14,38	125,38	50	61
	294	1	55,00	.	.	.	.	55	55
	296	1	69,00	.	.	.	.	69	69
	298	3	62,33	1,155	,667	59,46	65,20	61	63
	300	1	75,00	.	.	.	.	75	75
	302	1	68,00	.	.	.	.	68	68
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75
GAR Descriptives	2	1	37,00	.	.	.	.	37	37
	4	2	29,50	16,263	11,500	-116,62	175,62	18	41
	6	4	27,25	19,619	9,810	-3,97	58,47	10	55
	8	13	30,31	16,209	4,496	20,51	40,10	6	68
	10	30	21,20	12,277	2,241	16,62	25,78	7	59
	12	30	32,93	17,665	3,225	26,34	39,53	9	71
	14	9	31,00	19,729	6,576	15,83	46,17	7	69
	16	26	30,58	14,566	2,857	24,69	36,46	7	73
	18	7	23,14	20,530	7,760	4,16	42,13	6	61
	20	33	30,45	18,427	3,208	23,92	36,99	6	75
	21	1	39,00	.	.	.	.	39	39
	24	2	7,50	,707	,500	1,15	13,85	7	8
	25	1	70,00	.	.	.	.	70	70
	26	1	10,00	.	.	.	.	10	10
	28	5	27,00	10,954	4,899	13,40	40,60	14	38
	30	2	39,50	37,477	26,500	-297,21	376,21	13	66
	32	1	21,00	.	.	.	.	21	21
	36	1	46,00	.	.	.	.	46	46
	40	14	35,14	22,360	5,976	22,23	48,05	7	75
	48	1	38,00	.	.	.	.	38	38
	50	2	42,50	12,021	8,500	-65,50	150,50	34	51
	60	3	33,33	25,775	14,881	-30,69	97,36	10	61
	80	12	37,83	10,616	3,065	31,09	44,58	23	60
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75
GWAT Descriptives	1	68	30,94	16,387	1,987	26,97	34,91	6	71
	2	92	29,28	18,215	1,899	25,51	33,05	7	75
	3	35	29,00	16,562	2,799	23,31	34,69	6	75
	4	6	39,17	13,674	5,582	24,82	53,52	23	60
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75
ROMN Descriptives	2	13	6,69	,480	,133	6,40	6,98	6	7
	3	29	10,79	1,859	,345	10,09	11,50	8	14
	4	45	19,71	5,467	,815	18,07	21,35	15	41

	5	30	28,63	1,991	,364	27,89	29,38	25	32
	6	20	35,40	3,169	,709	33,92	36,88	32	43
	7	13	36,85	2,609	,724	35,27	38,42	32	42
	8	8	38,38	,518	,183	37,94	38,81	38	39
	9	7	63,71	15,628	5,907	49,26	78,17	41	75
	10	11	48,09	7,021	2,117	43,37	52,81	44	69
	11	10	51,40	2,066	,653	49,92	52,88	48	55
	12	3	56,67	2,082	1,202	51,50	61,84	55	59
	13	12	62,42	2,503	,723	60,83	64,01	60	68
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75
INC Descriptives	35000	15	6,87	,640	,165	6,51	7,22	6	8
	40000	92	18,65	6,629	,691	17,28	20,03	8	30
	45000	12	31,17	,937	,271	30,57	31,76	30	32
	50000	12	34,33	1,073	,310	33,65	35,02	33	36
	60000	14	43,29	11,472	3,066	36,66	49,91	36	61
	70000	16	43,56	9,121	2,280	38,70	48,42	37	60
	75000	6	41,50	,548	,224	40,93	42,07	41	42
	80000	8	45,13	1,126	,398	44,18	46,07	43	46
	85000	6	48,17	1,169	,477	46,94	49,39	47	50
	90000	8	52,75	1,581	,559	51,43	54,07	51	55
	100000	11	67,64	4,365	1,316	64,70	70,57	63	75
	110000	1	75,00	.	.	.	.	75	75
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75
CARN Descriptives	No	26	15,08	12,182	2,389	10,16	20,00	6	47
	1	94	21,64	11,629	1,199	19,26	24,02	9	63
	2	80	44,34	12,311	1,376	41,60	47,08	27	75
	3	1	75,00	.	.	.	.	75	75
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75
WCAR Descriptives	No	62	23,23	15,034	1,909	19,41	27,04	6	71
	1	71	34,52	16,502	1,958	30,62	38,43	6	70
	2	38	33,97	18,616	3,020	27,85	40,09	7	75
	3	15	32,47	19,056	4,920	21,91	43,02	6	73
	4	15	25,27	14,670	3,788	17,14	33,39	7	46
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75
WDISH Descriptives	1	13	27,15	14,843	4,117	18,18	36,12	9	59
	2	60	30,82	13,967	1,803	27,21	34,42	7	68
	3	128	30,05	18,780	1,660	26,76	33,33	6	75
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75
WCL Descriptives	1	83	27,95	17,520	1,923	24,13	31,78	6	71
	2	99	30,07	16,387	1,647	26,80	33,34	6	69
	3	15	39,73	16,373	4,227	30,67	48,80	20	75
	4	4	38,75	24,336	12,168	,03	77,47	23	75
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75
UTLT Descriptives	3	36	24,31	12,224	2,037	20,17	28,44	6	60
	4	87	27,38	17,866	1,915	23,57	31,19	6	75
	5	59	35,02	17,201	2,239	30,53	39,50	7	75
	6	13	33,77	13,192	3,659	25,80	41,74	16	61
	7	6	47,67	19,263	7,864	27,45	67,88	28	73
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75

		N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
FSHW Descriptives	1	40	31,33	16,703	2,641	25,98	36,67	6	71
	2	72	26,26	16,497	1,944	22,39	30,14	6	73
	3	64	34,61	18,052	2,256	30,10	39,12	7	75
	4	23	27,17	15,733	3,280	20,37	33,98	7	61
	5	1	20,00	.	.	.	.	20	20
	7	1	44,00	.	.	.	.	44	44
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75
MSHW Descriptives	1	68	28,21	17,209	2,087	24,04	32,37	6	73
	2	70	28,70	17,899	2,139	24,43	32,97	6	75
	3	48	31,27	14,738	2,127	26,99	35,55	6	69
	4	14	42,86	17,767	4,748	32,60	53,12	9	75
	5	1	20,00	.	.	.	.	20	20
	Total	201	30,09	17,189	1,212	27,70	32,48	6	75

ANNEX 05: Post hoc tests_ multiple comparisons

Table 5.12 Post hoc tests_ multiple comparisons

Multiple Comparisons
Tukey HSD

(I) AG3	(J) AG3	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
1	2	-11,101*	2,038	,000	-16,38	-5,82
	3	-28,873*	2,069	,000	-34,24	-23,51
	4	-31,405*	2,274	,000	-37,30	-25,51
2	1	11,101*	2,038	,000	5,82	16,38
	3	-17,773*	2,357	,000	-23,88	-11,66
	4	-20,304*	2,539	,000	-26,88	-13,73
3	1	28,873*	2,069	,000	23,51	34,24
	2	17,773*	2,357	,000	11,66	23,88
	4	-2,531	2,564	,757	-9,18	4,11
4	1	31,405*	2,274	,000	25,51	37,30

	2	20,304*	2,539	,000	13,73	26,88
	3	2,531	2,564	,757	-4,11	9,18

Multiple Comparisons
Tukey HSD

(I) AG2	(J) AG2	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
0	1	2,749	2,544	,527	-3,26	8,76
	2	9,787*	4,048	,043	,23	19,35
1	0	-2,749	2,544	,527	-8,76	3,26
	2	7,038	4,053	,194	-2,53	16,61
2	0	-9,787*	4,048	,043	-19,35	-,23
	1	-7,038	4,053	,194	-16,61	2,53

Multiple Comparisons
Tukey HSD

(I) FEM	(J) FEM	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	-5,550	2,636	,289	-13,14	2,04
	3	-15,427*	2,629	,000	-22,99	-7,86
	4	-22,226*	2,643	,000	-29,83	-14,62
	5	-44,161*	3,255	,000	-53,53	-34,79
	6	-58,621*	4,399	,000	-71,28	-45,96
2	1	5,550	2,636	,289	-2,04	13,14
	3	-9,877*	1,955	,000	-15,51	-4,25
	4	-16,676*	1,974	,000	-22,36	-10,99
	5	-38,611*	2,739	,000	-46,50	-30,73
	6	-53,071*	4,033	,000	-64,68	-41,46
3	1	15,427*	2,629	,000	7,86	22,99
	2	9,877*	1,955	,000	4,25	15,51
	4	-6,799*	1,965	,009	-12,45	-1,14
	5	-28,734*	2,733	,000	-36,60	-20,87
	6	-43,194*	4,028	,000	-54,79	-31,60
4	1	22,226*	2,643	,000	14,62	29,83
	2	16,676*	1,974	,000	10,99	22,36
	3	6,799*	1,965	,009	1,14	12,45
	5	-21,935*	2,746	,000	-29,84	-14,03
	6	-36,395*	4,038	,000	-48,02	-24,77
5	1	44,161*	3,255	,000	34,79	53,53
	2	38,611*	2,739	,000	30,73	46,50
	3	28,734*	2,733	,000	20,87	36,60
	4	21,935*	2,746	,000	14,03	29,84
	6	-14,460*	4,462	,017	-27,30	-1,62
6	1	58,621*	4,399	,000	45,96	71,28
	2	53,071*	4,033	,000	41,46	64,68
	3	43,194*	4,028	,000	31,60	54,79
	4	36,395*	4,038	,000	24,77	48,02
	5	14,460*	4,462	,017	1,62	27,30

Multiple Comparisons
Tukey HSD

(I) UTLT(day)	(J) UTLT(day)	Mean Difference (I-J)	Std. Error	Sig.
3	4	-3,074	3,283	,882
	5	-10,711*	3,503	,021

	6	-9,464	5,360	,397
	7	-23,361*	7,305	,014
4	3	3,074	3,283	,882
	5	-7,638	2,794	,053
	6	-6,390	4,926	,693
	7	-20,287*	6,992	,033
5	3	10,711*	3,503	,021
	4	7,638	2,794	,053
	6	1,248	5,075	,999
	7	-12,650	7,098	,387
6	3	9,464	5,360	,397
	4	6,390	4,926	,693
	5	-1,248	5,075	,999
	7	-13,897	8,176	,436
7	3	23,361*	7,305	,014
	4	20,287*	6,992	,033
	5	12,650	7,098	,387
	6	13,897	8,176	,436

Multiple Comparisons
Tukey HSD

(I) AG1	(J) AG1	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
0	1	-9,544*	2,601	,002	-16,28	-2,81
	2	-24,196*	2,708	,000	-31,21	-17,18
	3	-52,491*	4,741	,000	-64,78	-40,21
1	0	9,544*	2,601	,002	2,81	16,28
	2	-14,652*	1,994	,000	-19,82	-9,48
	3	-42,947*	4,374	,000	-54,28	-31,61
2	0	24,196*	2,708	,000	17,18	31,21
	1	14,652*	1,994	,000	9,48	19,82
	3	-28,295*	4,438	,000	-39,79	-16,80
3	0	52,491*	4,741	,000	40,21	64,78
	1	42,947*	4,374	,000	31,61	54,28
	2	28,295*	4,438	,000	16,80	39,79

Multiple Comparisons
Tukey HSD

(I) MDS	(J) MDS	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
0	1	-,868	2,698	,945	-7,24	5,50
	2	4,470	3,782	,465	-4,46	13,40
1	0	,868	2,698	,945	-5,50	7,24
	2	5,338	3,556	,293	-3,06	13,74
2	0	-4,470	3,782	,465	-13,40	4,46
	1	-5,338	3,556	,293	-13,74	3,06

Multiple Comparisons
Tukey HSD

(I) AG4	(J) AG4	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
0	1	-7,697	3,424	,066	-15,78	,39
	2	-5,775	3,527	,232	-14,10	2,55
1	0	7,697	3,424	,066	-,39	15,78
	2	1,922	2,654	,749	-4,35	8,19
2	0	5,775	3,527	,232	-2,55	14,10
	1	-1,922	2,654	,749	-8,19	4,35

Multiple Comparisons
Tukey HSD

(I) GWAT	(J) GWAT	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
1	2	1,659	2,755	,931	-5,48	8,80
	3	1,941	3,583	,949	-7,34	11,23
	4	-8,225	7,335	,677	-27,23	10,78
2	1	-1,659	2,755	,931	-8,80	5,48
	3	,283	3,421	1,000	-8,58	9,15
	4	-9,884	7,257	,525	-28,69	8,92
3	1	-1,941	3,583	,949	-11,23	7,34
	2	-,283	3,421	1,000	-9,15	8,58
	4	-10,167	7,611	,541	-29,89	9,55
4	1	8,225	7,335	,677	-10,78	27,23
	2	9,884	7,257	,525	-8,92	28,69
	3	10,167	7,611	,541	-9,55	29,89

Multiple Comparisons
Tukey HSD

(I) HGS	(J) HGS	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
1	2	-12,809*	1,794	,000	-17,04	-8,57
	3	-36,359*	2,022	,000	-41,13	-31,58
2	1	12,809*	1,794	,000	8,57	17,04
	3	-23,550*	2,315	,000	-29,02	-18,08
3	1	36,359*	2,022	,000	31,58	41,13
	2	23,550*	2,315	,000	18,08	29,02

Multiple Comparisons
Tukey HSD

(I) PRS	(J) PRS	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
0	1	-21,429*	1,730	,000	-25,51	-17,34
	2	-43,033*	4,095	,000	-52,70	-33,36
1	0	21,429*	1,730	,000	17,34	25,51
	2	-21,604*	4,156	,000	-31,42	-11,79
2	0	43,033*	4,095	,000	33,36	52,70
	1	21,604*	4,156	,000	11,79	31,42

Multiple Comparisons
Tukey HSD

(I) HOUS	(J) HOUS	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
2	3	-7,883	2,993	,122	-16,80	1,04
	4	-12,491*	2,957	,001	-21,30	-3,68
	5	-16,676*	2,957	,000	-25,49	-7,86
	6	-23,966*	2,669	,000	-31,92	-16,01
	7	-45,118*	2,926	,000	-53,84	-36,40
	8	-56,545*	3,412	,000	-66,71	-46,38
3	2	7,883	2,993	,122	-1,04	16,80
	4	-4,607	2,366	,452	-11,66	2,44
	5	-8,793*	2,366	,005	-15,84	-1,74
	6	-16,082*	1,994	,000	-22,02	-10,14
	7	-37,234*	2,326	,000	-44,17	-30,30
	8	-48,662*	2,915	,000	-57,35	-39,98
4	2	12,491*	2,957	,001	3,68	21,30

	3	4,607	2,366	,452	-2,44	11,66
	5	-4,185	2,320	,547	-11,10	2,73
	6	-11,475*	1,939	,000	-17,25	-5,70
	7	-32,627*	2,280	,000	-39,42	-25,84
	8	-44,054*	2,877	,000	-52,63	-35,48
5	2	16,676*	2,957	,000	7,86	25,49
	3	8,793*	2,366	,005	1,74	15,84
	4	4,185	2,320	,547	-2,73	11,10
	6	-7,290*	1,939	,004	-13,07	-1,51
	7	-28,442*	2,280	,000	-35,23	-21,65
	8	-39,869*	2,877	,000	-48,44	-31,30
6	2	23,966*	2,669	,000	16,01	31,92
	3	16,082*	1,994	,000	10,14	22,02
	4	11,475*	1,939	,000	5,70	17,25
	5	7,290*	1,939	,004	1,51	13,07
	7	-21,152*	1,890	,000	-26,78	-15,52
	8	-32,579*	2,580	,000	-40,27	-24,89
7	2	45,118*	2,926	,000	36,40	53,84
	3	37,234*	2,326	,000	30,30	44,17
	4	32,627*	2,280	,000	25,84	39,42
	5	28,442*	2,280	,000	21,65	35,23
	6	21,152*	1,890	,000	15,52	26,78
	8	-11,427*	2,845	,002	-19,90	-2,95
8	2	56,545*	3,412	,000	46,38	66,71
	3	48,662*	2,915	,000	39,98	57,35
	4	44,054*	2,877	,000	35,48	52,63
	5	39,869*	2,877	,000	31,30	48,44
	6	32,579*	2,580	,000	24,89	40,27
	7	11,427*	2,845	,002	2,95	19,90

Multiple Comparisons

Tukey HSD

(I) ROMN	(J) ROMN	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
2	3	-4,101	1,512	,229	-9,10	,90
	4	-13,019*	1,426	,000	-17,74	-8,30
	5	-21,941*	1,504	,000	-26,92	-16,96
	6	-28,708*	1,614	,000	-34,05	-23,37
	7	-30,154*	1,777	,000	-36,03	-24,27
	8	-31,683*	2,036	,000	-38,42	-24,95
	9	-57,022*	2,124	,000	-64,05	-49,99
	10	-41,399*	1,856	,000	-47,54	-35,26
	11	-44,708*	1,905	,000	-51,01	-38,40
	12	-49,974*	2,902	,000	-59,58	-40,37
	13	-55,724*	1,813	,000	-61,73	-49,72
3	2	4,101	1,512	,229	-,90	9,10
	4	-8,918*	1,079	,000	-12,49	-5,35
	5	-17,840*	1,180	,000	-21,74	-13,94
	6	-24,607*	1,317	,000	-28,96	-20,25
	7	-26,053*	1,512	,000	-31,06	-21,05
	8	-27,582*	1,809	,000	-33,57	-21,59
	9	-52,921*	1,908	,000	-59,23	-46,61
	10	-37,298*	1,604	,000	-42,61	-31,99
	11	-40,607*	1,661	,000	-46,10	-35,11
	12	-45,874*	2,747	,000	-54,97	-36,78
	13	-51,624*	1,555	,000	-56,77	-46,48
4	2	13,019*	1,426	,000	8,30	17,74
	3	8,918*	1,079	,000	5,35	12,49
	5	-8,922*	1,068	,000	-12,46	-5,39

	6	-15,689*	1,217	,000	-19,72	-11,66
	7	-17,135*	1,426	,000	-21,86	-12,41
	8	-18,664*	1,738	,000	-24,42	-12,91
	9	-44,003*	1,841	,000	-50,09	-37,91
	10	-28,380*	1,524	,000	-33,42	-23,34
	11	-31,689*	1,584	,000	-36,93	-26,45
	12	-36,956*	2,701	,000	-45,90	-28,02
	13	-42,706*	1,472	,000	-47,58	-37,83
5	2	21,941*	1,504	,000	16,96	26,92
	3	17,840*	1,180	,000	13,94	21,74
	4	8,922*	1,068	,000	5,39	12,46
	6	-6,767*	1,308	,000	-11,09	-2,44
	7	-8,213*	1,504	,000	-13,19	-3,23
	8	-9,742*	1,803	,000	-15,71	-3,78
	9	-35,081*	1,901	,000	-41,37	-28,79
	10	-19,458*	1,597	,000	-24,74	-14,17
	11	-22,767*	1,654	,000	-28,24	-17,29
	12	-28,033*	2,743	,000	-37,11	-18,96
	13	-33,783*	1,547	,000	-38,90	-28,66
6	2	28,708*	1,614	,000	23,37	34,05
	3	24,607*	1,317	,000	20,25	28,96
	4	15,689*	1,217	,000	11,66	19,72
	5	6,767*	1,308	,000	2,44	11,09
	7	-1,446	1,614	,999	-6,79	3,89
	8	-2,975	1,895	,918	-9,25	3,30
	9	-28,314*	1,989	,000	-34,90	-21,73
	10	-12,691*	1,700	,000	-18,32	-7,06
	11	-16,000*	1,754	,000	-21,81	-10,19
	12	-21,267*	2,805	,000	-30,55	-11,98
	13	-27,017*	1,654	,000	-32,49	-21,54
7	2	30,154*	1,777	,000	24,27	36,03
	3	26,053*	1,512	,000	21,05	31,06
	4	17,135*	1,426	,000	12,41	21,86
	5	8,213*	1,504	,000	3,23	13,19
	6	1,446	1,614	,999	-3,89	6,79
	8	-1,529	2,036	1,000	-8,27	5,21
	9	-26,868*	2,124	,000	-33,90	-19,84
	10	-11,245*	1,856	,000	-17,39	-5,10
	11	-14,554*	1,905	,000	-20,86	-8,25
	12	-19,821*	2,902	,000	-29,42	-10,22
	13	-25,571*	1,813	,000	-31,57	-19,57
8	2	31,683*	2,036	,000	24,95	38,42
	3	27,582*	1,809	,000	21,59	33,57
	4	18,664*	1,738	,000	12,91	24,42
	5	9,742*	1,803	,000	3,78	15,71
	6	2,975	1,895	,918	-3,30	9,25
	7	1,529	2,036	1,000	-5,21	8,27
	9	-25,339*	2,344	,000	-33,10	-17,58
	10	-9,716*	2,105	,000	-16,68	-2,75
	11	-13,025*	2,149	,000	-20,14	-5,91
	12	-18,292*	3,067	,000	-28,44	-8,14
	13	-24,042*	2,068	,000	-30,88	-17,20
9	2	57,022*	2,124	,000	49,99	64,05
	3	52,921*	1,908	,000	46,61	59,23
	4	44,003*	1,841	,000	37,91	50,09
	5	35,081*	1,901	,000	28,79	41,37
	6	28,314*	1,989	,000	21,73	34,90
	7	26,868*	2,124	,000	19,84	33,90
	8	25,339*	2,344	,000	17,58	33,10
	10	15,623*	2,190	,000	8,37	22,87
	11	12,314*	2,232	,000	4,93	19,70
	12	7,048	3,126	,513	-3,30	17,39
	13	1,298	2,154	1,000	-5,83	8,43

10	2	41,399*	1,856	,000	35,26	47,54
	3	37,298*	1,604	,000	31,99	42,61
	4	28,380*	1,524	,000	23,34	33,42
	5	19,458*	1,597	,000	14,17	24,74
	6	12,691*	1,700	,000	7,06	18,32
	7	11,245*	1,856	,000	5,10	17,39
	8	9,716*	2,105	,000	2,75	16,68
	9	-15,623*	2,190	,000	-22,87	-8,37
	11	-3,309	1,979	,879	-9,86	3,24
	12	-8,576	2,951	,147	-18,34	1,19
	13	-14,326*	1,891	,000	-20,58	-8,07
11	2	44,708*	1,905	,000	38,40	51,01
	3	40,607*	1,661	,000	35,11	46,10
	4	31,689*	1,584	,000	26,45	36,93
	5	22,767*	1,654	,000	17,29	28,24
	6	16,000*	1,754	,000	10,19	21,81
	7	14,554*	1,905	,000	8,25	20,86
	8	13,025*	2,149	,000	5,91	20,14
	9	-12,314*	2,232	,000	-19,70	-4,93
	10	3,309	1,979	,879	-3,24	9,86
	12	-5,267	2,982	,834	-15,14	4,60
	13	-11,017*	1,940	,000	-17,44	-4,60
12	2	49,974*	2,902	,000	40,37	59,58
	3	45,874*	2,747	,000	36,78	54,97
	4	36,956*	2,701	,000	28,02	45,90
	5	28,033*	2,743	,000	18,96	37,11
	6	21,267*	2,805	,000	11,98	30,55
	7	19,821*	2,902	,000	10,22	29,42
	8	18,292*	3,067	,000	8,14	28,44
	9	-7,048	3,126	,513	-17,39	3,30
	10	8,576	2,951	,147	-1,19	18,34
	11	5,267	2,982	,834	-4,60	15,14
	13	-5,750	2,924	,715	-15,43	3,93
13	2	55,724*	1,813	,000	49,72	61,73
	3	51,624*	1,555	,000	46,48	56,77
	4	42,706*	1,472	,000	37,83	47,58
	5	33,783*	1,547	,000	28,66	38,90
	6	27,017*	1,654	,000	21,54	32,49
	7	25,571*	1,813	,000	19,57	31,57
	8	24,042*	2,068	,000	17,20	30,88
	9	-1,298	2,154	1,000	-8,43	5,83
	10	14,326*	1,891	,000	8,07	20,58
	11	11,017*	1,940	,000	4,60	17,44
	12	5,750	2,924	,715	-3,93	15,43

Multiple Comparisons

Tukey HSD

(I) WCL(week)	(J) WCL(week)	Mean Difference (I-J)	Std. Error	Sig.
1	2	-2,119	2,532	,837
	3	-11,782	4,773	,068
	4	-10,798	8,709	,602
2	1	2,119	2,532	,837
	3	-9,663	4,714	,173
	4	-8,679	8,676	,749
3	1	11,782	4,773	,068
	2	9,663	4,714	,173
	4	,983	9,573	1,000
4	1	10,798	8,709	,602
	2	8,679	8,676	,749
	3	-,983	9,573	1,000

Multiple Comparisons
Tukey HSD

(I) UNIV	(J) UNIV	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
0	1	-15,697*	1,721	,000	-19,76	-11,63
	2	-37,771*	2,014	,000	-42,53	-33,02
1	0	15,697*	1,721	,000	11,63	19,76
	2	-22,074*	1,838	,000	-26,41	-17,73
2	0	37,771*	2,014	,000	33,02	42,53
	1	22,074*	1,838	,000	17,73	26,41

Multiple Comparisons
Tukey HSD

(I) TAR(m²)	(J) TAR(m²)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
80	100	-3,713	1,289	,236	-8,19	,77
	110	-6,830*	1,197	,000	-10,99	-2,67
	120	-9,546*	1,289	,000	-14,03	-5,07
	140	-12,363*	1,376	,000	-17,14	-7,58
	160	-15,713*	1,289	,000	-20,19	-11,23
	180	-20,546*	1,289	,000	-25,03	-16,07
	200	-22,845*	1,151	,000	-26,84	-18,85
	210	-25,130*	1,289	,000	-29,61	-20,65
	220	-27,763*	1,376	,000	-32,54	-22,98
	240	-29,630*	1,289	,000	-34,11	-25,15
	260	-35,852*	1,430	,000	-40,82	-30,88
	280	-39,690*	1,329	,000	-44,31	-35,07
	300	-45,496*	1,197	,000	-49,65	-41,34
	310	-54,886*	1,255	,000	-59,24	-50,53
	320	-66,963*	2,723	,000	-76,42	-57,50
100	80	3,713	1,289	,236	-,77	8,19
	110	-3,117	1,439	,720	-8,12	1,88
	120	-5,833*	1,517	,015	-11,10	-,56
	140	-8,650*	1,591	,000	-14,18	-3,12
	160	-12,000*	1,517	,000	-17,27	-6,73
	180	-16,833*	1,517	,000	-22,10	-11,56
	200	-19,132*	1,401	,000	-24,00	-14,26
	210	-21,417*	1,517	,000	-26,69	-16,15
	220	-24,050*	1,591	,000	-29,58	-18,52
	240	-25,917*	1,517	,000	-31,19	-20,65
	260	-32,139*	1,639	,000	-37,83	-26,45
	280	-35,977*	1,551	,000	-41,37	-30,59
	300	-41,783*	1,439	,000	-46,78	-36,78
	310	-51,173*	1,488	,000	-56,34	-46,00
	320	-63,250*	2,838	,000	-73,11	-53,39
110	80	6,830*	1,197	,000	2,67	10,99
	100	3,117	1,439	,720	-1,88	8,12
	120	-2,717	1,439	,878	-7,72	2,28
	140	-5,533*	1,517	,029	-10,80	-,26
	160	-8,883*	1,439	,000	-13,88	-3,88
	180	-13,717*	1,439	,000	-18,72	-8,72
	200	-16,016*	1,316	,000	-20,59	-11,44
	210	-18,300*	1,439	,000	-23,30	-13,30
	220	-20,933*	1,517	,000	-26,20	-15,66
	240	-22,800*	1,439	,000	-27,80	-17,80
	260	-29,022*	1,567	,000	-34,47	-23,58
	280	-32,861*	1,475	,000	-37,99	-27,74
	300	-38,667*	1,357	,000	-43,38	-33,95
	310	-48,056*	1,408	,000	-52,95	-43,16

	320	-60,133*	2,797	,000	-69,85	-50,42
120	80	9,546*	1,289	,000	5,07	14,03
	100	5,833*	1,517	,015	,56	11,10
	110	2,717	1,439	,878	-2,28	7,72
	140	-2,817	1,591	,924	-8,34	2,71
	160	-6,167*	1,517	,007	-11,44	-,90
	180	-11,000*	1,517	,000	-16,27	-5,73
	200	-13,299*	1,401	,000	-18,17	-8,43
	210	-15,583*	1,517	,000	-20,85	-10,31
	220	-18,217*	1,591	,000	-23,74	-12,69
	240	-20,083*	1,517	,000	-25,35	-14,81
	260	-26,306*	1,639	,000	-32,00	-20,61
	280	-30,144*	1,551	,000	-35,53	-24,76
	300	-35,950*	1,439	,000	-40,95	-30,95
	310	-45,340*	1,488	,000	-50,51	-40,17
	320	-57,417*	2,838	,000	-67,28	-47,56
140	80	12,363*	1,376	,000	7,58	17,14
	100	8,650*	1,591	,000	3,12	14,18
	110	5,533*	1,517	,029	,26	10,80
	120	2,817	1,591	,924	-2,71	8,34
	160	-3,350	1,591	,759	-8,88	2,18
	180	-8,183*	1,591	,000	-13,71	-2,66
	200	-10,482*	1,481	,000	-15,63	-5,34
	210	-12,767*	1,591	,000	-18,29	-7,24
	220	-15,400*	1,662	,000	-21,17	-9,63
	240	-17,267*	1,591	,000	-22,79	-11,74
	260	-23,489*	1,707	,000	-29,42	-17,56
	280	-27,327*	1,624	,000	-32,97	-21,69
	300	-33,133*	1,517	,000	-38,40	-27,86
	310	-42,523*	1,563	,000	-47,95	-37,09
	320	-54,600*	2,879	,000	-64,60	-44,60
160	80	15,713*	1,289	,000	11,23	20,19
	100	12,000*	1,517	,000	6,73	17,27
	110	8,883*	1,439	,000	3,88	13,88
	120	6,167*	1,517	,007	,90	11,44
	140	3,350	1,591	,759	-2,18	8,88
	180	-4,833	1,517	,113	-10,10	,44
	200	-7,132*	1,401	,000	-12,00	-2,26
	210	-9,417*	1,517	,000	-14,69	-4,15
	220	-12,050*	1,591	,000	-17,58	-6,52
	240	-13,917*	1,517	,000	-19,19	-8,65
	260	-20,139*	1,639	,000	-25,83	-14,45
	280	-23,977*	1,551	,000	-29,37	-18,59
	300	-29,783*	1,439	,000	-34,78	-24,78
	310	-39,173*	1,488	,000	-44,34	-34,00
	320	-51,250*	2,838	,000	-61,11	-41,39
180	80	20,546*	1,289	,000	16,07	25,03
	100	16,833*	1,517	,000	11,56	22,10
	110	13,717*	1,439	,000	8,72	18,72
	120	11,000*	1,517	,000	5,73	16,27
	140	8,183*	1,591	,000	2,66	13,71
	160	4,833	1,517	,113	-,44	10,10
	200	-2,299	1,401	,959	-7,17	2,57
	210	-4,583	1,517	,171	-9,85	,69
	220	-7,217*	1,591	,001	-12,74	-1,69
	240	-9,083*	1,517	,000	-14,35	-3,81
	260	-15,306*	1,639	,000	-21,00	-9,61
	280	-19,144*	1,551	,000	-24,53	-13,76
	300	-24,950*	1,439	,000	-29,95	-19,95
	310	-34,340*	1,488	,000	-39,51	-29,17
	320	-46,417*	2,838	,000	-56,28	-36,56
200	80	22,845*	1,151	,000	18,85	26,84
	100	19,132*	1,401	,000	14,26	24,00

	110	16,016*	1,316	,000	11,44	20,59
	120	13,299*	1,401	,000	8,43	18,17
	140	10,482*	1,481	,000	5,34	15,63
	160	7,132*	1,401	,000	2,26	12,00
	180	2,299	1,401	,959	-2,57	7,17
	210	-2,284	1,401	,961	-7,15	2,58
	220	-4,918	1,481	,078	-10,06	,23
	240	-6,784*	1,401	,000	-11,65	-1,92
	260	-13,007*	1,532	,000	-18,33	-7,68
	280	-16,845*	1,438	,000	-21,84	-11,85
	300	-22,651*	1,316	,000	-27,22	-18,08
	310	-32,041*	1,369	,000	-36,80	-27,28
	320	-44,118*	2,778	,000	-53,77	-34,47
210	80	25,130*	1,289	,000	20,65	29,61
	100	21,417*	1,517	,000	16,15	26,69
	110	18,300*	1,439	,000	13,30	23,30
	120	15,583*	1,517	,000	10,31	20,85
	140	12,767*	1,591	,000	7,24	18,29
	160	9,417*	1,517	,000	4,15	14,69
	180	4,583	1,517	,171	-,69	9,85
	200	2,284	1,401	,961	-2,58	7,15
	220	-2,633	1,591	,955	-8,16	2,89
	240	-4,500	1,517	,194	-9,77	,77
	260	-10,722*	1,639	,000	-16,41	-5,03
	280	-14,561*	1,551	,000	-19,95	-9,17
	300	-20,367*	1,439	,000	-25,37	-15,37
	310	-29,756*	1,488	,000	-34,92	-24,59
	320	-41,833*	2,838	,000	-51,69	-31,97
220	80	27,763*	1,376	,000	22,98	32,54
	100	24,050*	1,591	,000	18,52	29,58
	110	20,933*	1,517	,000	15,66	26,20
	120	18,217*	1,591	,000	12,69	23,74
	140	15,400*	1,662	,000	9,63	21,17
	160	12,050*	1,591	,000	6,52	17,58
	180	7,217*	1,591	,001	1,69	12,74
	200	4,918	1,481	,078	-,23	10,06
	210	2,633	1,591	,955	-2,89	8,16
	240	-1,867	1,591	,999	-7,39	3,66
	260	-8,089*	1,707	,000	-14,02	-2,16
	280	-11,927*	1,624	,000	-17,57	-6,29
	300	-17,733*	1,517	,000	-23,00	-12,46
	310	-27,123*	1,563	,000	-32,55	-21,69
	320	-39,200*	2,879	,000	-49,20	-29,20
240	80	29,630*	1,289	,000	25,15	34,11
	100	25,917*	1,517	,000	20,65	31,19
	110	22,800*	1,439	,000	17,80	27,80
	120	20,083*	1,517	,000	14,81	25,35
	140	17,267*	1,591	,000	11,74	22,79
	160	13,917*	1,517	,000	8,65	19,19
	180	9,083*	1,517	,000	3,81	14,35
	200	6,784*	1,401	,000	1,92	11,65
	210	4,500	1,517	,194	-,77	9,77
	220	1,867	1,591	,999	-3,66	7,39
	260	-6,222*	1,639	,018	-11,91	-,53
	280	-10,061*	1,551	,000	-15,45	-4,67
	300	-15,867*	1,439	,000	-20,87	-10,87
	310	-25,256*	1,488	,000	-30,42	-20,09
	320	-37,333*	2,838	,000	-47,19	-27,47
260	80	35,852*	1,430	,000	30,88	40,82
	100	32,139*	1,639	,000	26,45	37,83
	110	29,022*	1,567	,000	23,58	34,47
	120	26,306*	1,639	,000	20,61	32,00
	140	23,489*	1,707	,000	17,56	29,42

	160	20,139*	1,639	,000	14,45	25,83
	180	15,306*	1,639	,000	9,61	21,00
	200	13,007*	1,532	,000	7,68	18,33
	210	10,722*	1,639	,000	5,03	16,41
	220	8,089*	1,707	,000	2,16	14,02
	240	6,222*	1,639	,018	,53	11,91
	280	-3,838	1,670	,626	-9,64	1,96
	300	-9,644*	1,567	,000	-15,09	-4,20
	310	-19,034*	1,611	,000	-24,63	-13,44
	320	-31,111*	2,905	,000	-41,20	-21,02
280	80	39,690*	1,329	,000	35,07	44,31
	100	35,977*	1,551	,000	30,59	41,37
	110	32,861*	1,475	,000	27,74	37,99
	120	30,144*	1,551	,000	24,76	35,53
	140	27,327*	1,624	,000	21,69	32,97
	160	23,977*	1,551	,000	18,59	29,37
	180	19,144*	1,551	,000	13,76	24,53
	200	16,845*	1,438	,000	11,85	21,84
	210	14,561*	1,551	,000	9,17	19,95
	220	11,927*	1,624	,000	6,29	17,57
	240	10,061*	1,551	,000	4,67	15,45
	260	3,838	1,670	,626	-1,96	9,64
	300	-5,806*	1,475	,011	-10,93	-,68
	310	-15,196*	1,522	,000	-20,48	-9,91
	320	-27,273*	2,857	,000	-37,20	-17,35
300	80	45,496*	1,197	,000	41,34	49,65
	100	41,783*	1,439	,000	36,78	46,78
	110	38,667*	1,357	,000	33,95	43,38
	120	35,950*	1,439	,000	30,95	40,95
	140	33,133*	1,517	,000	27,86	38,40
	160	29,783*	1,439	,000	24,78	34,78
	180	24,950*	1,439	,000	19,95	29,95
	200	22,651*	1,316	,000	18,08	27,22
	210	20,367*	1,439	,000	15,37	25,37
	220	17,733*	1,517	,000	12,46	23,00
	240	15,867*	1,439	,000	10,87	20,87
	260	9,644*	1,567	,000	4,20	15,09
	280	5,806*	1,475	,011	,68	10,93
	310	-9,390*	1,408	,000	-14,28	-4,50
	320	-21,467*	2,797	,000	-31,18	-11,75
310	80	54,886*	1,255	,000	50,53	59,24
	100	51,173*	1,488	,000	46,00	56,34
	110	48,056*	1,408	,000	43,16	52,95
	120	45,340*	1,488	,000	40,17	50,51
	140	42,523*	1,563	,000	37,09	47,95
	160	39,173*	1,488	,000	34,00	44,34
	180	34,340*	1,488	,000	29,17	39,51
	200	32,041*	1,369	,000	27,28	36,80
	210	29,756*	1,488	,000	24,59	34,92
	220	27,123*	1,563	,000	21,69	32,55
	240	25,256*	1,488	,000	20,09	30,42
	260	19,034*	1,611	,000	13,44	24,63
	280	15,196*	1,522	,000	9,91	20,48
	300	9,390*	1,408	,000	4,50	14,28
	320	-12,077*	2,823	,003	-21,88	-2,27
320	80	66,963*	2,723	,000	57,50	76,42
	100	63,250*	2,838	,000	53,39	73,11
	110	60,133*	2,797	,000	50,42	69,85
	120	57,417*	2,838	,000	47,56	67,28
	140	54,600*	2,879	,000	44,60	64,60
	160	51,250*	2,838	,000	41,39	61,11
	180	46,417*	2,838	,000	36,56	56,28

	200	44,118*	2,778	,000	34,47	53,77
	210	41,833*	2,838	,000	31,97	51,69
	220	39,200*	2,879	,000	29,20	49,20
	240	37,333*	2,838	,000	27,47	47,19
	260	31,111*	2,905	,000	21,02	41,20
	280	27,273*	2,857	,000	17,35	37,20
	300	21,467*	2,797	,000	11,75	31,18
	310	12,077*	2,823	,003	2,27	21,88

Multiple Comparisons
Tukey HSD

(I) WCAR(month)	(J) WCAR(month)	Mean Difference (I-J)	Std. Error	Sig.
No	1	-11,295*	2,879	,001
	2	-10,748*	3,412	,016
	3	-9,241	4,766	,300
	4	-2,041	4,766	,993
1	No	11,295*	2,879	,001
	2	,547	3,329	1,000
	3	2,054	4,707	,992
	4	9,254	4,707	,286
2	No	10,748*	3,412	,016
	1	-,547	3,329	1,000
	3	1,507	5,051	,998
	4	8,707	5,051	,422
3	No	9,241	4,766	,300
	1	-2,054	4,707	,992
	2	-1,507	5,051	,998
	4	7,200	6,048	,757
4	No	2,041	4,766	,993
	1	-9,254	4,707	,286
	2	-8,707	5,051	,422
	3	-7,200	6,048	,757

Multiple Comparisons
Tukey HSD

(I) WDISH(day)	(J) WDISH(day)	Mean Difference (I-J)	Std. Error	Sig.
1	2	-3,663	5,278	,767
	3	-2,893	5,023	,833
2	1	3,663	5,278	,767
	3	,770	2,700	,956
3	1	2,893	5,023	,833
	2	-,770	2,700	,956