TALLINN UNIVERSITY OF TECHNOLOGY
SCHOOL OF ENGINEERING
Department of Materials and Environmental Technology

SPATIO-TEMPORAL ANALYSIS OF DISTRIBUTED ENERGY POTENTIAL IN TALLINN

JAGATUD ENERGIAPOTENSIAALI AJALINE ANALÜÜS TALLINNAS

MASTER THESIS

Student: ARASH SALEHI

Student code: 194226KAYM

Supervisor: Dr. Sambeet Mishra, Postdoctoral Researcher

Co-supervisor: Professor Ivo Palu

Tallinn 2021

(*On the reverse side of title page*)

AUTHOR’S DECLARATION

Hereby I declare, that I have written this thesis independently.
No academic degree has been applied for based on this material. All works, major viewpoints and data of the other authors used in this thesis have been referenced.

27 MAY 2021

Author: Arash Salehi
/signature /

Thesis is in accordance with terms and requirements

27 MAY 2021

Supervisor: Dr.Sambeet Mishra
/signature/

27 MAY 2021

Co-Supervisor: Professor Ivo Palu
/signature/

Accepted for defence

27 MAY 2021

Chairman of theses defence commission: Professor Malle Krunks
/name and signature/

Non-exclusive licence for reproduction and publication of a graduation thesis[1]

I Arash Salehi

1. grant Tallinn University of Technology free licence (non-exclusive licence) for my thesis

SPATIO-TEMPORAL ANALYSIS OF DISTRIBUTED ENERGY POTENTIAL IN TALLINN

supervised by Dr. Sambeet Mishra and Professor Ivo Palu as Co-supervisor of TALTECH.

1.1 to be reproduced for the purposes of preservation and electronic publication of the graduation thesis, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright;

1.2 to be published via the web of Tallinn University of Technology, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright.

2. I am aware that the author also retains the rights specified in clause 1 of the non-exclusive licence.

3. I confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights, the rights arising from the Personal Data Protection Act or rights arising from other legislation.

---

27 MAY 2021

[1] *The non-exclusive licence is not valid during the validity of access restriction indicated in the student's application for restriction on access to the graduation thesis that has been signed by the school's dean, except in case of the university's right to reproduce the thesis for preservation purposes only. If a graduation thesis is based on the joint creative activity of two or more persons and the co-author(s) has/have not granted, by the set deadline, the student defending his/her graduation thesis consent to reproduce and publish the graduation thesis in compliance with clauses 1.1 and 1.2 of the non-exclusive licence, the non-exclusive license shall not be valid for the period.*

TalTech Department`s title

THESIS TASK

Student: Arash Salehi, 194226KAYM

Study programme, KAYM 09/18, Materials and processes for sustainable energetics

main speciality: Processes for sustainable energetics

Supervisor(s): Dr. Sambeet Mishra, Postdoctoral Researcher of TALTECH

Co-supervisor: Professor. Ivo Palu, Programme Director (Electrical Power Engineering and Mechatronics)

Thesis topic:

SPATIO-TEMPORAL ANALYSIS OF DISTRIBUTED ENERGYPOTENTIAL IN TALLINN

JAGATUD ENERGIAPOTENSIAALI AJALINE ANALÜÜSTALLINNAS

Thesis main objectives:

1. Making the base map of Tallinn
2. Geo-spatial analysis
3. Temporal energy consumption analysis
4. Local PV production

Thesis tasks and time schedule:

No	Task description	Deadline
1.	Learning the QGIS software and R programming	Dec2020
2.	Literature review and data collection	Feb2021
3.	Making anlysis	Apr2021

Language: english **Deadline for submission of thesis:** 27 May 2021

Student: Arash Salehi .... 27 MAY 2021

/signature/

Supervisor: Dr. Sambeet Mishra .... 27 MAY 2021

/signature/

Co-supervisor: Professor. Ivo Palu.............. 27 MAY 2021

/signature/

Head of study programme: .... 27 MAY 2021

/signature/

Terms of thesis closed defence and/or restricted access conditions to be formulated on the reverse side

CONTENTS

PREFACE

The topic of this master's thesis was initiated by Arash Salehi and supervised by Dr. Sambeet Mishra Postdoctoral Researcher of TALTECH

This work was for geo-spatial analysis of PV potential in distributed energy system of Tallinn. The geo-spatial analysis part was done in the quantum geographical information system software (QGIS) to define the relation between different regions of Tallinn according to the population, energy consumption and the geographical situation of them.

The main goal of the study was to use the useless rooftop area of existing buildings which has the area more than 3000 square meter to implementing PV modules to produce the portion of electricity demand of Tallinn.

Keywords: geo-spatial analysis, PV, QGIS, energy

1. INTRODUCTION

By increasing the daily number of people in the world, the need for more energy is one of the critical global issues. Expanding the population in any region brings some needs with it. These additional humans need food, dwelling, energy, etc. When we are talking about more food, we need more agricultural fields, water, and companies needed to prepare the required food. So for all these, we need the energy to transport water to fields, make agricultural infrastructure, build new companies that need construction, and new supplies for energy that they need. When we are talking about dwellings, by constructing new buildings, the need for energy supplies will come, and governments need to build sufficient energy infrastructures to supply all the regions everywhere. These are not all the factors when we have a new amount of population. Many other issues will come and they need energy supplies, such as transportation, workplaces, schools, entertainment, etc.

When we are talking about energy, there are different kinds. Some are conventional like oil, gas, coal, etc., which have a limited amount and others are new like solar, hydro, wind, geothermal, and biofuels which are sustainable energy resources.

Oil is liquid and, in comparison with other conventional energy resources, has a higher energy density. Since 1960, the production amount of oil has been increasing due to the Industrial Revolution and the need for more energy annually [1]. Figure 0.1 represents the growth of the oil production over the years in which the growth slope is high till around 2006, which is touching the highest point of the graph which shows the peak of the oil production during these years. Also, it can be seen that the production's prediction is going down yearly because of the efforts of whole world countries to produce cleaner energies to be able to alternate them with oil.

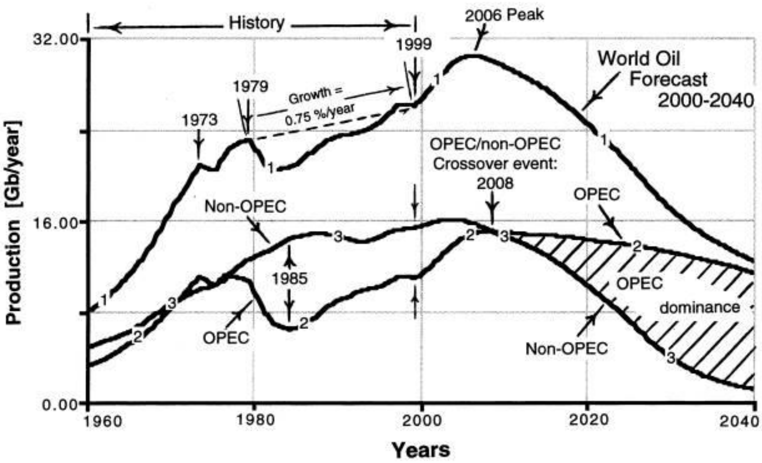

Figure 1.1. Oil production

The other large concern which we all are facing is the issue of keeping our globe cleaner. There is an important competition all over the world these days in the field of energy. The significant aspect of this competition is that efforts indicate countries are cleaner in energy. Countries try to change their policies in order to be greener and have less CO2 emissions annually. Figure 1.2 shows the portion of different countries in CO2 emission per capita.

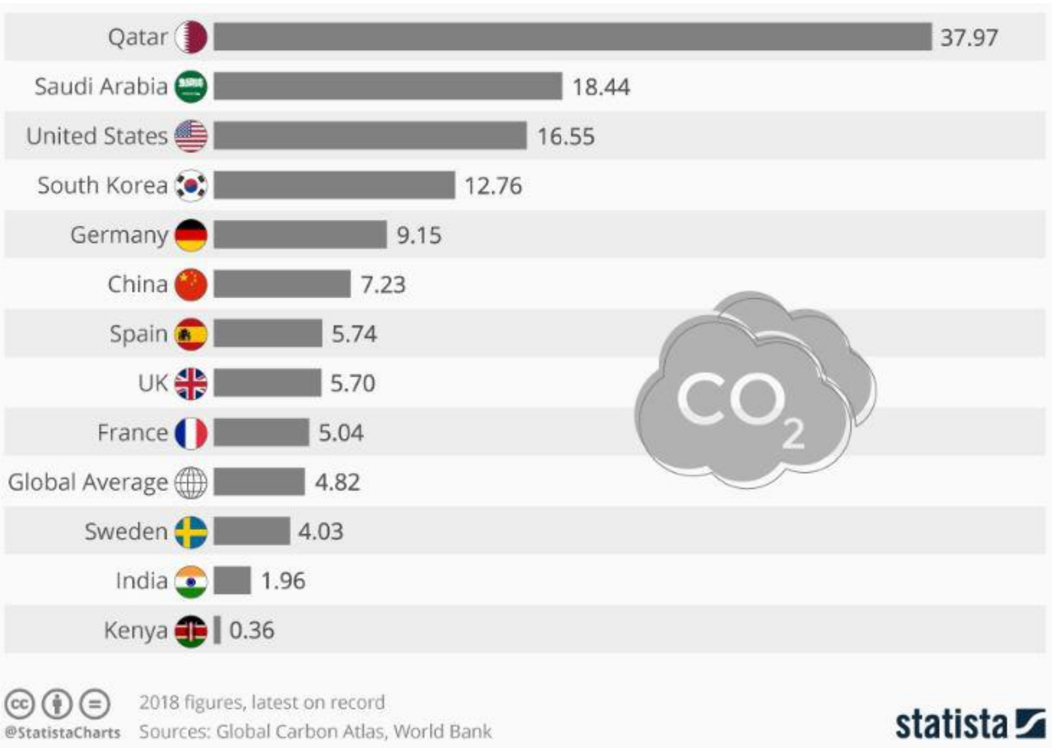

Figure 1.2. The amount of CO2 emission per capita in different countries by 2020

According to recent research [2] we can see the tendency of using sustainable energy resources is very high in the Middle East and Australia. After that, there is a large demand in Russia, Europe, and North America (Figure 1.3).

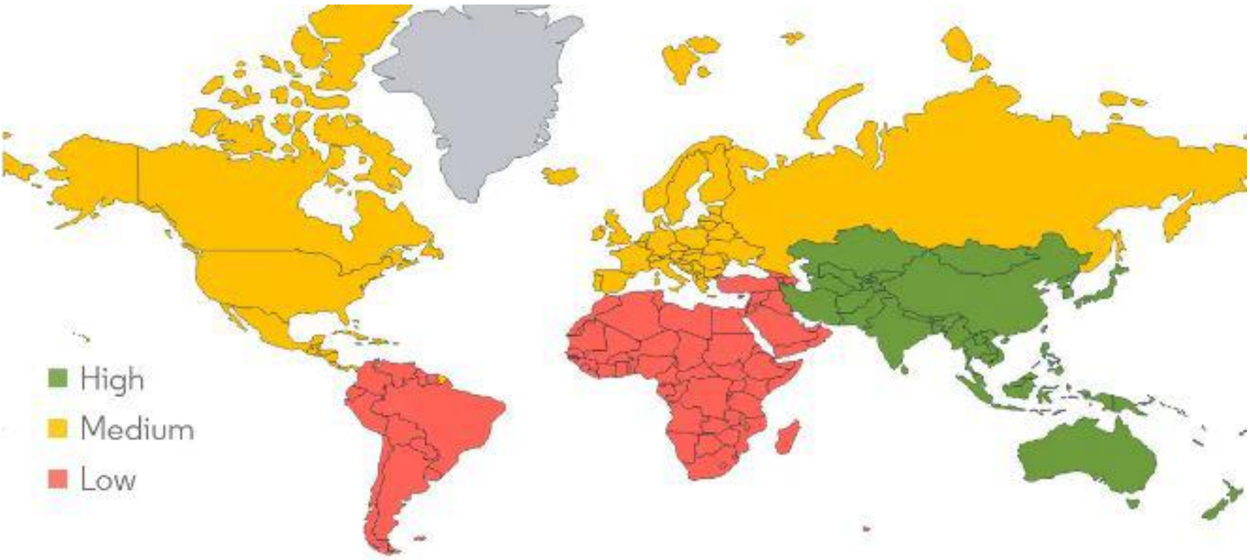

Figure 1.3. World tendency for sustainable energy

This tendency and recent works in a way to be more sustainable have pushed the world to less CO2 emission, which is represented in Figure 1.4, the result of the collected data around the globe.

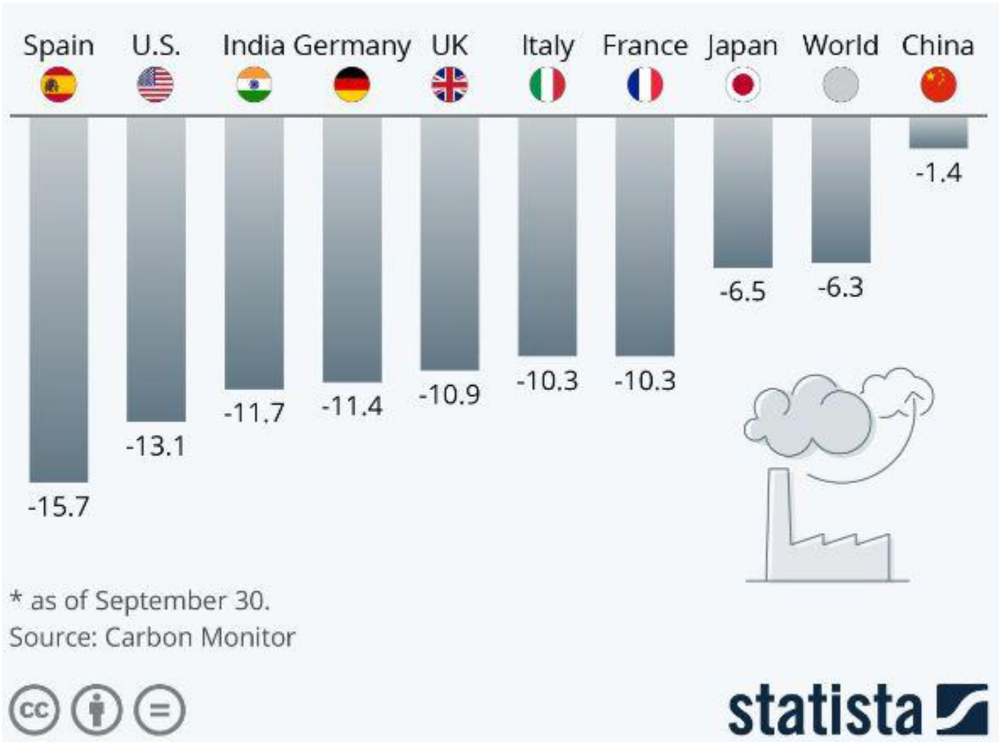

Figure 1.4. Decrease of the CO2 emission by 2020

The increase in energy consumption brings in the issue of efficiency. By being more efficient according to energy production, we could have less energy waste and less pollution. There are many factors which affect the issue of energy. One is the kind of energy production that could be from conventional resources or renewable energy resources. The other factor that could impact the issue of energy production is the distribution method of the countries for power grids. Last but not least is the importance of maintenance. Every system needs regular maintenance to be efficient. This fact is also applicable for power grids to be more reliable to supply their customers.

There are some issues here according to the increasing usage of renewable energy resources. One is the appropriate location for implementing the production units that need accurate research to find these locations and considerable investments in this field. The other issue will be the instructions for the use of these produced energies. It needs some new infrastructure to connect these production units to the primary grid. These new connections also need precise analysis to define a trend and pattern by which we could modify the old energy distribution systems.

The geo-spatial analysis is the way to approach a proper decision, choose the location for production units, and find an appropriate connection to the primary grid. Using the geo analysis will help us assess our desired area and find the best potentials for a generation. Then by analyzing these data and comparing them to the data on consumption, we can make decisions. The consumption is crucial because we are going toward making smaller grids and supply locally to avoid losses and be more reliable in modern energy systems.

2. LITERATURE REVIEW

The future energy demand of the world is likely to increase because of climate change and population growth. By raising the global temperature, the hot seasons are longer, and the demand for cooling is rising accordingly. That’s not all, in the new world the technological improvements, increasing the number of industries, need for more agricultural activities and change in individuals and organizations habits has a critical effect on energy demand [3]. The other factor which is critical according to the growth of energy demand is population. Population changes affect the energy demand. By increasing the rate of humans in any region, the need for more energy is growing [4]. Figure 2.1 has shown the trend line of the population growth all over the world from 1950 to 2048.

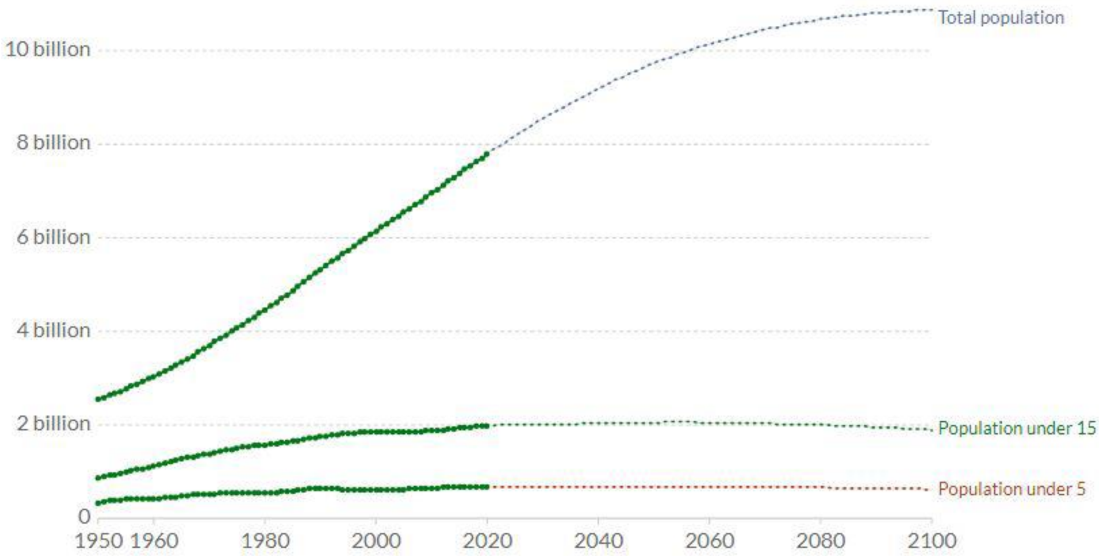

Figure 2.1. The world population growth

On the other hand, by the expansion of countries the industries are also growing. Since the early 1990s, the liberalization of the electricity sector has been an issue for European countries. They try to create competition in production and let customers decide to choose the supplier company. Although one of the significant objectives of the EU's liberalization in the electricity sector is to increase welfare by decreasing the price of electricity and guaranteeing secure energy, and in most European countries, it generally increased [5]. In Figure 2.2 the changing price of electricity for households between 2018 and 2019 is illustrated.

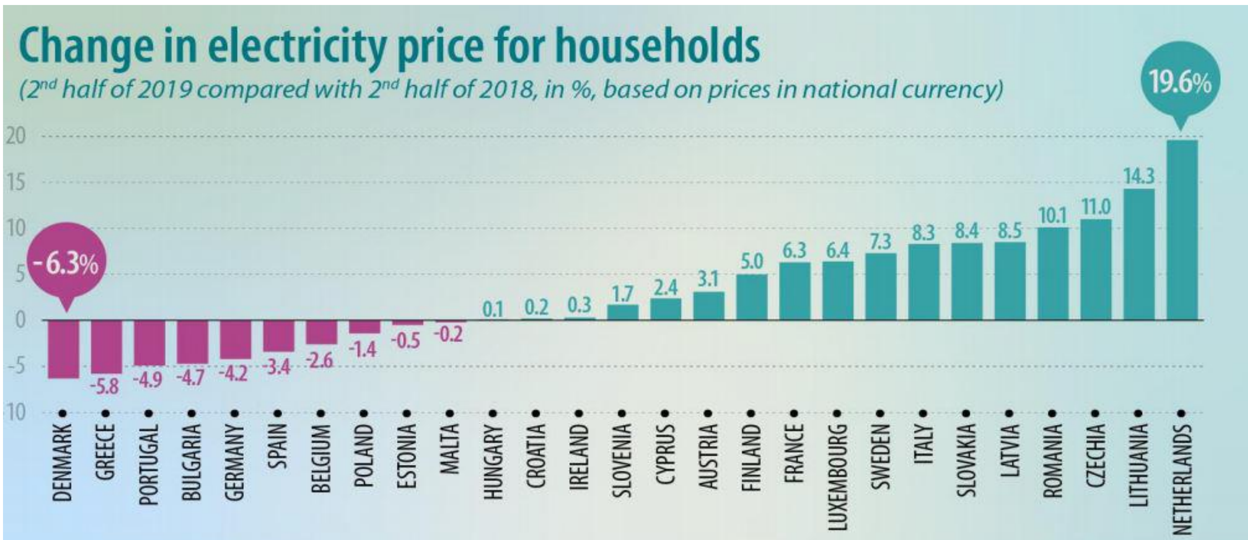

Figure 2.2. The electricity price change in European countries

In addition, to make a competitive product, the EU has set a goal to reach more renewable energy production as sustainable energy resources. Although these new energies are relatively more expensive than the conventional ones, the price of them will decrease by growing the technology. The critical point of renewable energies is the fluctuation of them [6]. To monitor all the actions regarding renewable energies, the Commission of Energy was organized. The national renewable energy action plan commission collects data from members according to renewable energy goals. In this way, the energy trends of the world could be monitored regarding renewable energies annually and present them [7]. In Figure 2.3 you can see the sum energy consumption of European countries in different sectors.

Key factors which play the driving role in the way to be more sustainable according to energy are:

2. Lobby of conventional energy resources
3. Energy efficiency
4. CO2 emission
5. Income

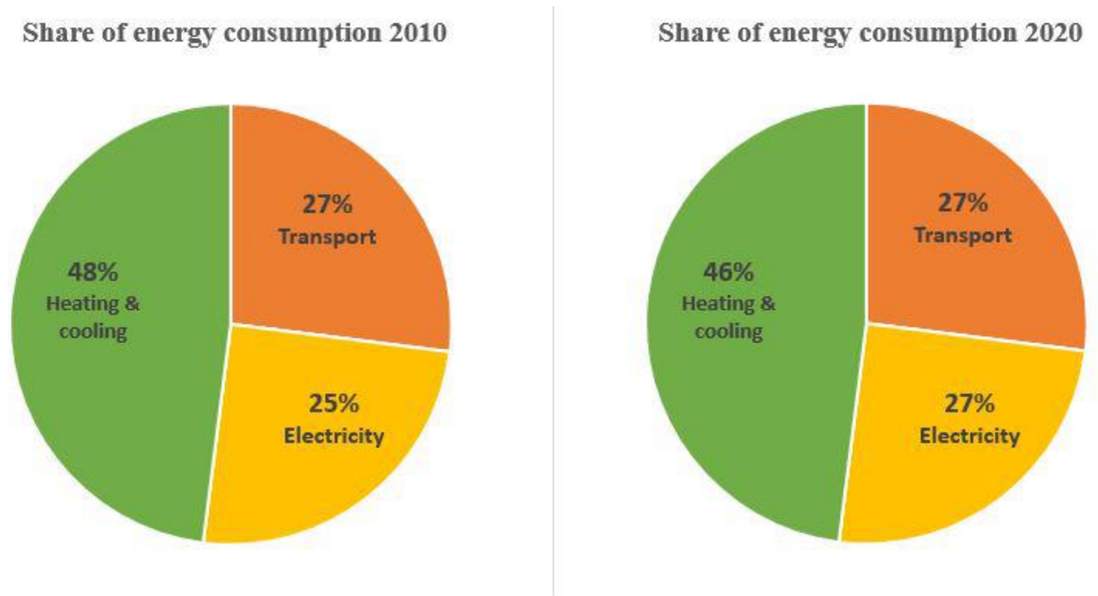

Figure 2.3. Energy usage in different sectors in Europe

The research [8] represented these critical factors as drivers toward renewable energy generation. The pressure of traditional energy sources lobby, from one side and the low efficiency of sustainable energy generation and the tendency to keep the world clean and

less CO_2 emission from the other side and also the income of European people has a direct relation to the amount of their investment in this field.
Among all research about the Europe energy path, there was a study in 2016 to illustrate ways to be sustainable [9]. The primary energy supply, environmental (CO_2 emission), and economy are the factors that have a significant impact in this way. The goal is to reach 100% renewable energy production for Europe by 2050. The solution of the study is to focus on important energy sectors like:

- Electricity
- Transportation
- Heating and cooling

Research shows that 100% renewable energy for Europe is possible by localizing energy production, changing old distribution systems, using district heating and cooling systems, and alternating from conventional transportation fuel vehicles (public and personal) to electrical ones. The result of the study indicates that Europe can reach 80% less CO_2 emission in 2050 in comparison with 1990, and the cost of this system is 3% higher than a conventional system and 12% higher to reach 100% less emissions. As such, a significant portion of this cost is for changing the system, not the energy production itself.
The GIS tool’s significant application in the energy sector is supporting the decisions to be more effective and accurate. In [10] case study, the effect of different factors and the possible constraints was assessed. Using geographical models in this study has represented the potential locations for renewable energy productions. In addition, it helped to provide flexible datasets of the efficiency of the existing electricity grid and implementation of renewable energy units. It helped the local administration to make appropriate decisions for energy production.
In 2019 there was a case study in Italy to assess the Lazio region’s electricity consumption and the possibility of coverage with renewable energy resources [11]. By using some geographical analysis, they reached the conclusion that although they cannot cover more than 30% of the Lazio's Consumption by 2030 because of the population and the amount of energy demand, they do have the analysis for each region which shows that the potential of renewable energy sources can supply some regions need 100% (Figure 2.4).

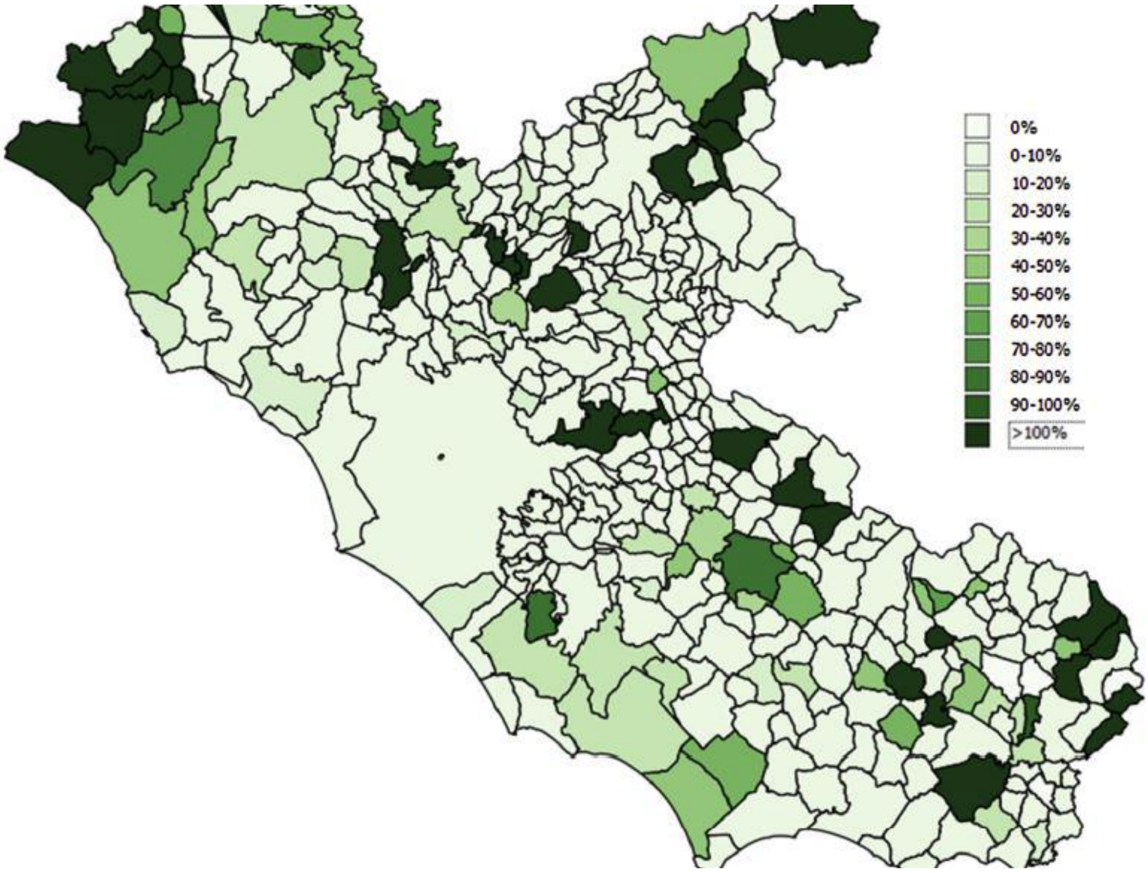

Figure 2.4. Renewable energy potential for Lazio's regions

The SIGAR project is energy-related research for the Spanish electrical generation system for renewable energy resources [12]. The concept is to find the ceiling of renewable generation potential there until the year 2050. GIS tools are key points in this study first to define the distribution of renewable resources, find the restrictions in each location, and calculate the maximum capacity of generation of renewable energies. The results could also be improved according to technological developments as we go forward, but the study shows that by 2050 with this plan, Spanish renewables can generate 100% of the electricity consumption there.

One of the practical tools to execute geo analysis is QGIS. In [13] research the author used this tool to find the appropriate sites for solar applications. In that study, three major factors have been considered to find suitable places for execution. These are:

- Protected areas (Wildlife sanctuaries, World Heritage Sites and National Parks)
- Areas of interests (Tourist spots, Railway stations, Airports)
- Major settlement areas (with a population density above 2000)

Furthermore, by having the target locations available to implement the solar generation units, to analyze it technically by deploying the Analytical Hierarchy Process (AHP) as a

Multi-Criteria Decision-Making (MCDM) tool, the study could define the appropriate sites of the photovoltaic solar units on the map.

The significant potential of GIS tools has been shown in many kinds of research. In a specific investigation [14] for the possibility of renewable energy resources use and energy management in Brazil by using GISA SOL 1.0 (Geographical Information System Applied to Solar Energy) the author made a geo analysis on the implementation of the renewable generation system. For installation of new energy units, a vast number of criteria come to mind like solar resource, rural electrification index, income per capita, proximity to a transmission line, existence of portable water, aptitude for diverse agricultural cultivation, wind resources, restriction of soil use, etc. In this case, the map for analyzing the specific location has 80 different layers to show the impact of various factors, which help to have an accurate analysis and make appropriate decisions.

Due to the proper location of Turkey in the world, the country is rich in terms of renewable energy resources. One study [15] has determined where the solar plant (PV) should be implemented to supply the energy demand at that location. By using the Analytic Hierarchy Process (AHP), one of the Multi-Criteria Decision-Making (MCDM) methods (which was initially developed by Prof. Thomas L. Saaty in 1977) the process of finding the best site locations was executed. In the determination process, many factors have been assessed to find suitable locations. Energy transmission lines, transformer centers, highways, and railways have been assessed using the GIS applications on the base map of Turkey.

In other renewable sectors like biomass the application and innovation of GIS tools have also been represented. In a case study for Milan, Italy [16] by using the invention of the GIS-based method and its implementation for the target location in Milan, the estimation process of the number of wood pruning by-products and help the administrative energy cost efficiency. By this analysis and using the waste of the products for the local energy production with the help of the GIS tools, the government could be cost-efficient for energy production, continue the way to be green, and make good samples of circular economy systems in the energy sector too.

3. OBJECTIVES AND KEY CONTRIBUTIONS

How is the aggregation of the population in Tallinn? What is the amount of energy consumption there? How is the power grid of Tallinn shaped? What is the PV energy production potential of Tallinn and the allocation of them? How can this PV potential affect the energy system of Tallinn? This study has investigated these objectives to make a clear vision of them and provide a proper solution and model. To achieve these, we need to have some essential material. We have to find some statistical resources about Tallinn which should be reliable and precise to make an accurate model. To make these data more understandable Its better to make them visual. By using geo analysis tools and collected data, we can make those visualizations. These are the features of Tallinn which we need to visualize:

- Base map
- Energy
- Power grid model

These data visualizations will help us understand the tendency of different attributes according to their location. It also allows us to be able to have a more accurate geo-spatial analysis. A vast number of factors are involved in these datasets, and each factor could affect the results of the study. Some are related to making a base map, such as:

- Coordination
- Coordinate reference system(CRS) of the maps
- Scale

And others have significant effect on the analysis of the study like:

- Location of energy demand points
- Terrain which demand point is located in
- Road map
- Energy grid map
- Power distribution substation’s location
- PV potential locations
- Time aspect of demands

All these huge amounts of data are essential to make a precise geo-spatial analysis and find a reliable model for the energy system.

4. TOOLS AND DATA

To make the analysis, we need to have tools with essential applications that we are looking to make our analysis and have different data resources for our study. Datasets are the input of our analytical flow. Tools will change and carry out the analysis process on these datasets. Some of the outputs are the result, and some are input for other tools to make more analysis on them (Figure 4.1).

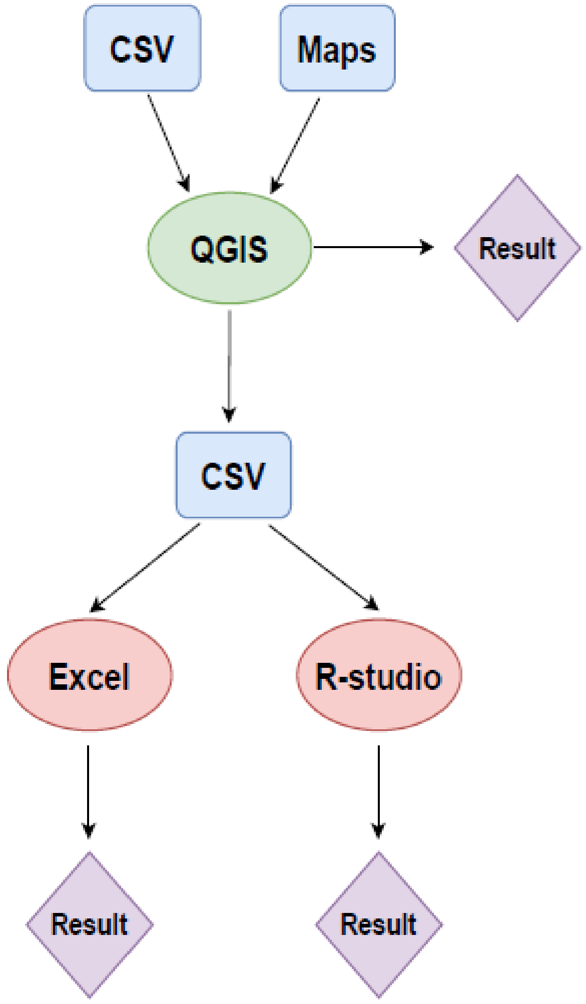

Figure 4.1. Process flowchart of analysis

4.1 Tools and software

As you can see in the flowchart above, we have different tools for processing and finding results in this study, such as:

- The Quantum geographic information system (QGIS) software
- The R-statistical programming environment
- Excel spreadsheet

4.1.1 The quantum geografic information system software

What is QGIS? QGIS is a quantum GIS which is a powerful geographical information system. This software is used to gather, manage and analyze data. In geographical sciences, this platform is used to visualize data and monitor the changes of attributes to make a deep analysis. A different analysis of the QGIS will help us find the relations between different attributes and make decisions [17]. Features that we have with QGIS:

- View data

- Explore data and compose maps
- create, edit, manage and export data
- Analyse data
- Publish maps
- Extend QGIS functionality through plugins
- Programming console

To view data in QGIS, we need to make maps with them. By using datasets, we can make different layers in QGIS with a variety of applications. After making these maps which consist of different layers, we can make a vast number of geo-spatial analyses with them and then publish our work through the QGIS on the internet to be available for others to use. Although we can make our maps and layers according to data we have collected and analyze them, the software gives us the opportunity of using plugins to extend the abilities and functionalities.

Last but not least, there is also a Python programming console inside the software which allows us to make our analysis with our codes [18]. The input materials of QGIS also have a wide variety. When we use a previous job or collected data, it is more usual to use maps in the format of shapefile (.SHP) or use a dataset in excel with CSV format. By considering the Coordinate reference system (CRS) of the maps, we can import them into QGIS, use the made map in our process, import the csv dataset, and make a new map.

As a global GIS software, it also has a connection with different global GIS data centers. We can use the global maps from Google Maps, Google Satellite, and open street maps to allocate our new layers on the actual map of the location.

After all, we could have the result of the analysis as a picture of a map or export it as a map in the format of shapefile to be available for others who use the QGIS or make an excel dataset in the format of a CSV file.

4.1.2 R-statistical programming environment

R-studio is an integrated development environment for R and Python programming language with console, editor for direct code execution, debugging, and tools for plotting. R provides us a workspace to use geo-spatial data as input and make further analysis on them. In R, the CSV files could also be the input material for making different processes. So in R as a programming platform, we can use CSV datasets and all different maps we have generated in QGIS as a shapefile.

4.2 Dataset

One of the essential things for an accurate analysis is a reliable and precise dataset. In this study, we first used the Tallinn data website [19] to collect the base map of Tallinn. This website has a considerable number of different maps of Tallinn. We can find the maps we need in the desired format according to our tools.

4.2.1 Spatial

In this case, we have collected the base map of Tallinn in a shapefile. This map illustrates different regions of this city on the map with the exact CRS(EPSG:4326 - WGS 84 – Geographic). After that, collecting data regarding the layers we need is important. For these data and maps, we used both the Tallinn website and Geoportaal Maaamet [20] . By investigating these websites, we can find different maps and layers in a wide variety of formats that we need according to the tool that we use.

4.2.2 Temporal

After composing maps and modifying the different layers that we need on them, statistics is a significant part. To make the geo-spatial analysis using the geo tools and analytical tools, we need datasets of the study's factors. According to Estonia, the stat website [21] is one of the reliable and accurate data sources. On this website, there are different categories for the necessary data and in each of them, it gives you the option to modify the different columns of datasets that you need. There are various formats for datasets that use the CSV format because we can work on it in the QGIS, R studio, and Excel.

5. METHODOLOGY

In this study, the main goal is to make decisions for the energy production of Tallinn in order to be more efficient and be more sustainable. This work is done to make regional energy demand and production classifications and define the potential regions of Tallinn to implement PV panels for more sustainable energy production.

According to the IRENA organization [22], (International Renewable Energy Agency) statistics, the price of PV electricity production decreases annually. It shows, in comparing the PV and wind, PV production costs almost similar to wind production, and economically using PV generation units for producing the demand of electricity is beneficial (Figure 5.1), not only because of the production cost which is almost the same as the wind but also because of the allocation of these units. In this study, the aim is to use the useless rooftops of existing buildings. On the other hand, for wind parks, there is a need to have vast lands, which makes more cost and needs to have changes in nature.

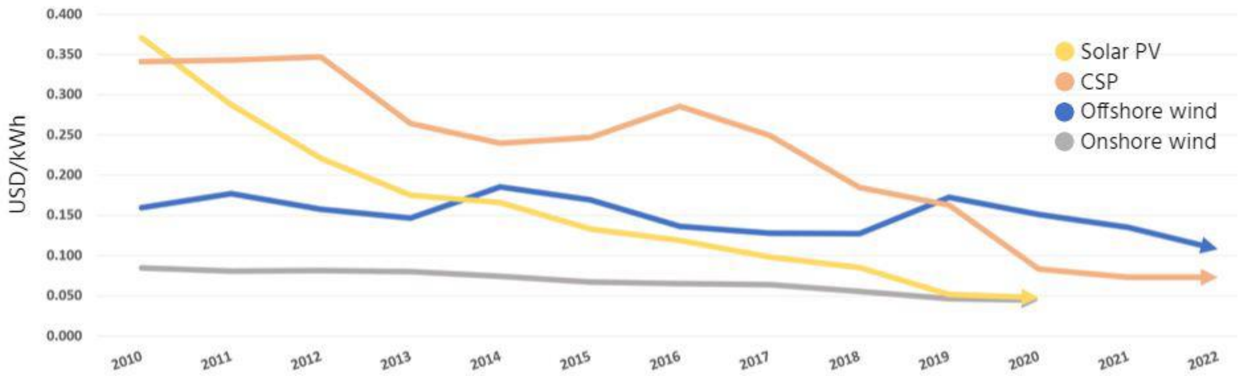

Figure 5.1. Cost of different kinds of electricity production

To reach this goal, there is a need to have an accurate analysis. This analysis has different aspects.

One is the Geo-Spatial analysis part. The relation of the influential factors of demand points or regions connected to space and the geographical condition will be represented. Also, it will give us the ability to monitor these factors in a specific period. There are helpful geographical tools for this kind of analysis. In this study, the QGIS is the tool that has been used to make the base map of Tallinn and other different map layers (Figure 5.2, 5.3, 5.4, 5.5).

Figure 5.2. Administrative map

Figure 5.3. Demand points and grid

Figure 5.4. Building's map

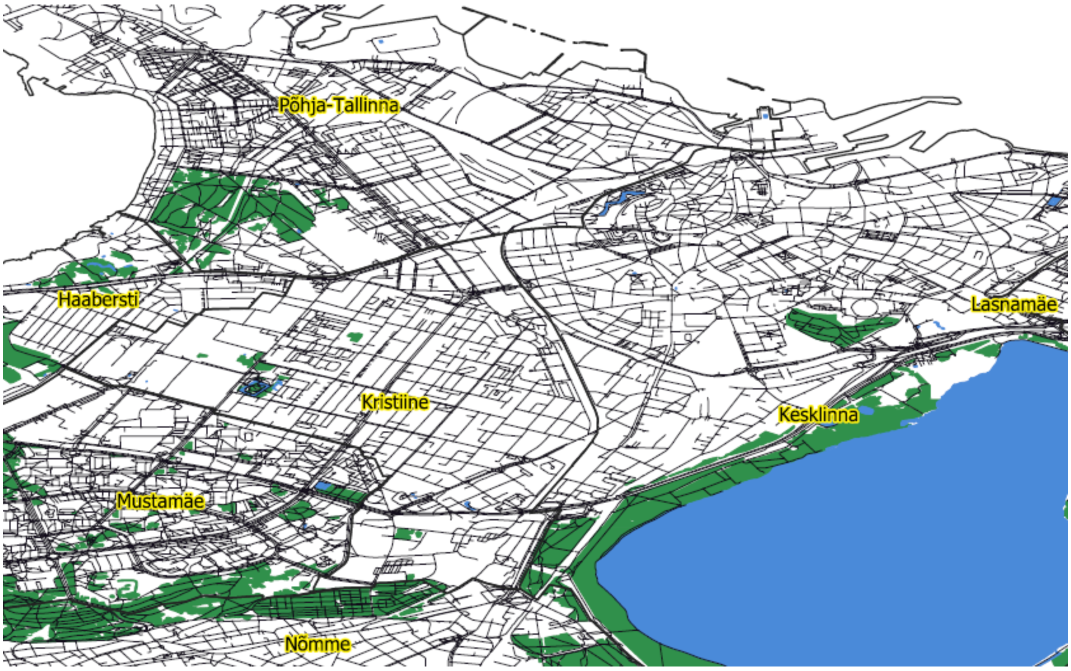

Figure 5.5. Road map

To determine the solar potential of Tallinn, the solar atlas website [23] is a valuable source. In the solar atlas, by choosing a location on the globe, you could have the heat map of the solar irradiation of that location.

Also, it will provide you with some detailed data about the amount of solar radiation and the estimation of the energy production there (Figure 5.6).

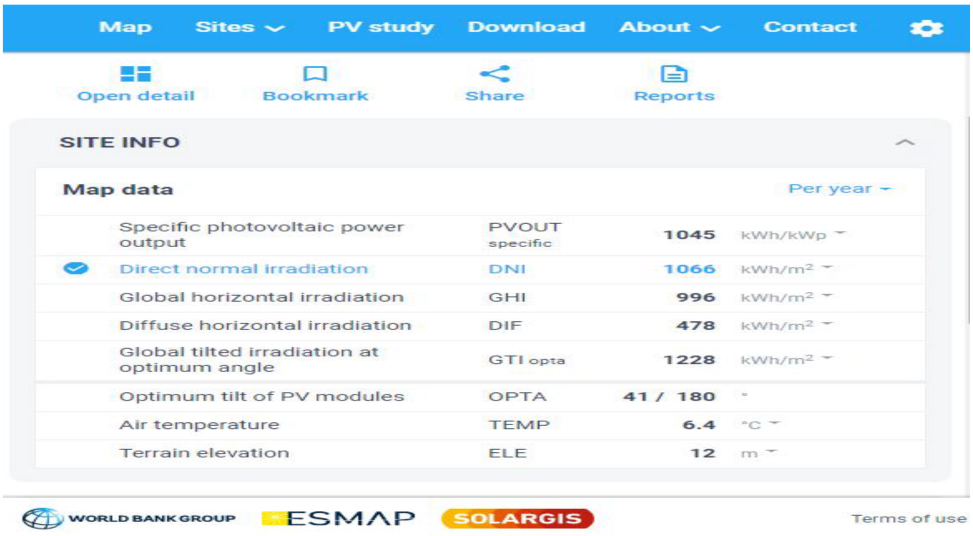

Figure 5.6 Solar atlas data for Tallinn solar potential

The other essential aspect of this analysis is data. The actual data of the factors that are being used need to be collected from reliable data resources of Estonia. As much as the data are accurate, the analysis and classification of the regions according to their attribute's

allocation could be more precise. The resource used in this study is the Stat organization [21], which is one of Estonia's main statistical data resources.
After all, the analysis can be executed. By using data that have been collected from stat and also the QGIS. There is the possibility of making different map layers in the QGIS for various aspects needed for the analysis. The QGIS, as a powerful software, gives us the ability to create different analyses using tools and functionality.

5.1 Space-time relationship

When we are talking about space-time relationships, it is classically referred to as geographically physical interactions. As we know in geographical analysis, each factor has two crucial aspects which will introduce them, one is geographical detail, and the other is time. The point is that these two aspects are not separated from each other and have a close connection to represent each factor in a specific time with a particular location physically.

5.1.1 Space-time path

The space-time path represented how the factor has physical movement concerning time (Figure 5.7). It represents the relation between time and the kind of energy demand, which means that we have which type of energy demand in a specific period, and then illustrates the need to transition the energy to the other demand. In sum, it describes the time and path relation of energy demands for distribution solutions according to the geographical aspect of the demand points.

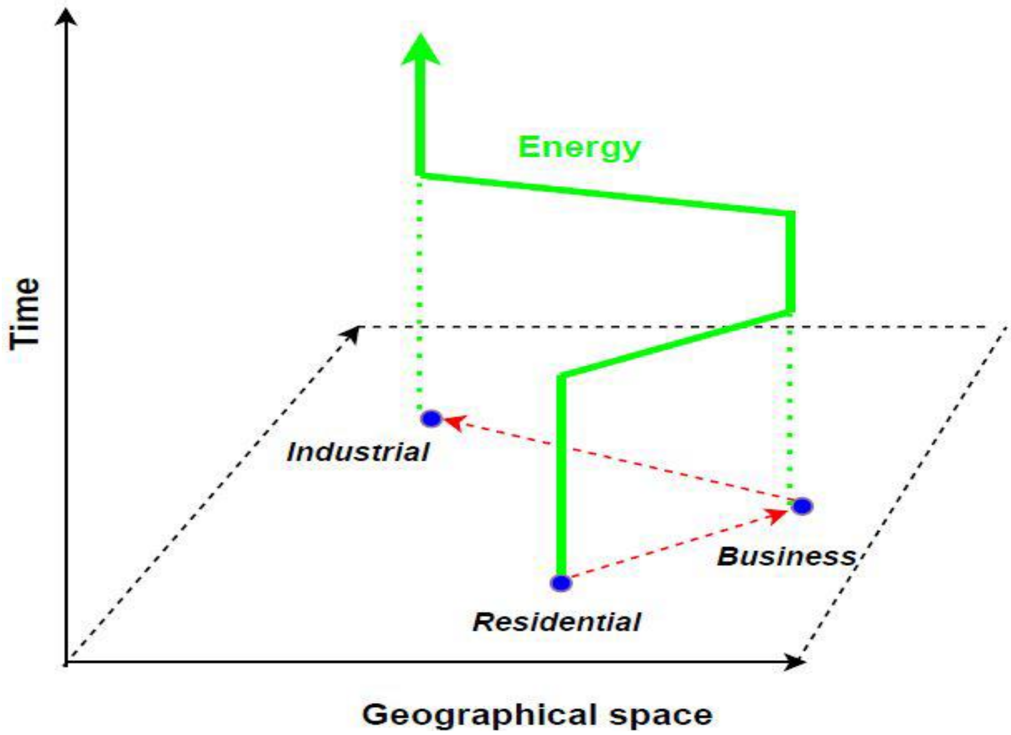

Figure 5.7. Space-time path illustration

5.2 Exploratory spatial data analysis

The rapid development of the GIS softwares helps us make accurate analyses with a huge amount of data. The exploratory spatial data analysis is a way to illustrate our data in a practical way according to the geographical locations. The visualization of data is a way to represent the different factors of data we use in our studies (Figure 5.8). Exploratory spatial data analysis gives us meaningful visualizations of:

- Spatial distributions
- Spatial auto-correlation
- Spatial attributes

With this illustration, we can make different maps and plots. Using these maps and plots, we cannot only show the distribution of the factors but also find the relation between different regions according to a specific factor of the dataset [24].

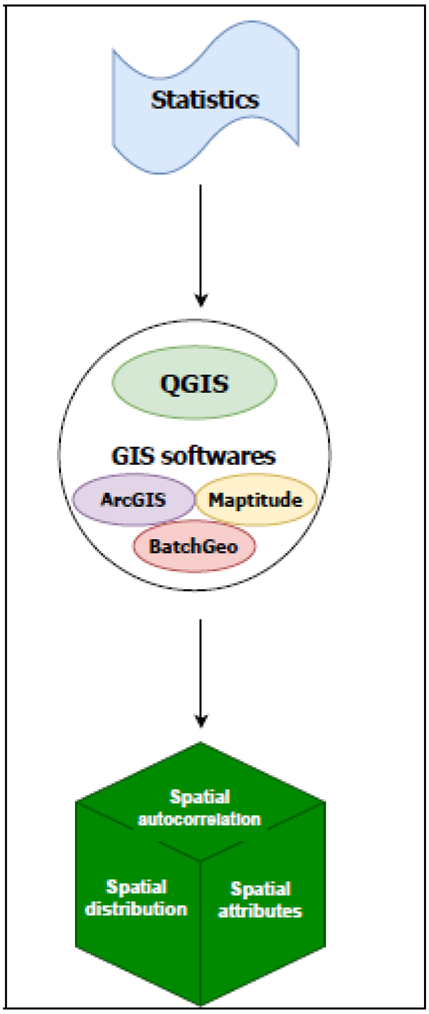

Figure 5.8. Exploratory Spatial Data Analysis diagram

5.2.1 Energy consumption

When talking about private or public facilities, one of the most critical factors is their location. It helps not only to reduce fixed costs but also increase the accessibility and reliability of the facility [25]. To be able to analyze the spatial data of Tallinn energy demand points, we need the allocation of cumulative demand points on the map of Tallinn in the QGIS, which is one of the influential factors for determining energy production unit allocation analysis . With different layer tools from QGIS, we can make this layer on the top of the base map and although these are demand points, they can contain a vast amount of data inside that affect the allocation of production units (Figure 5.9).

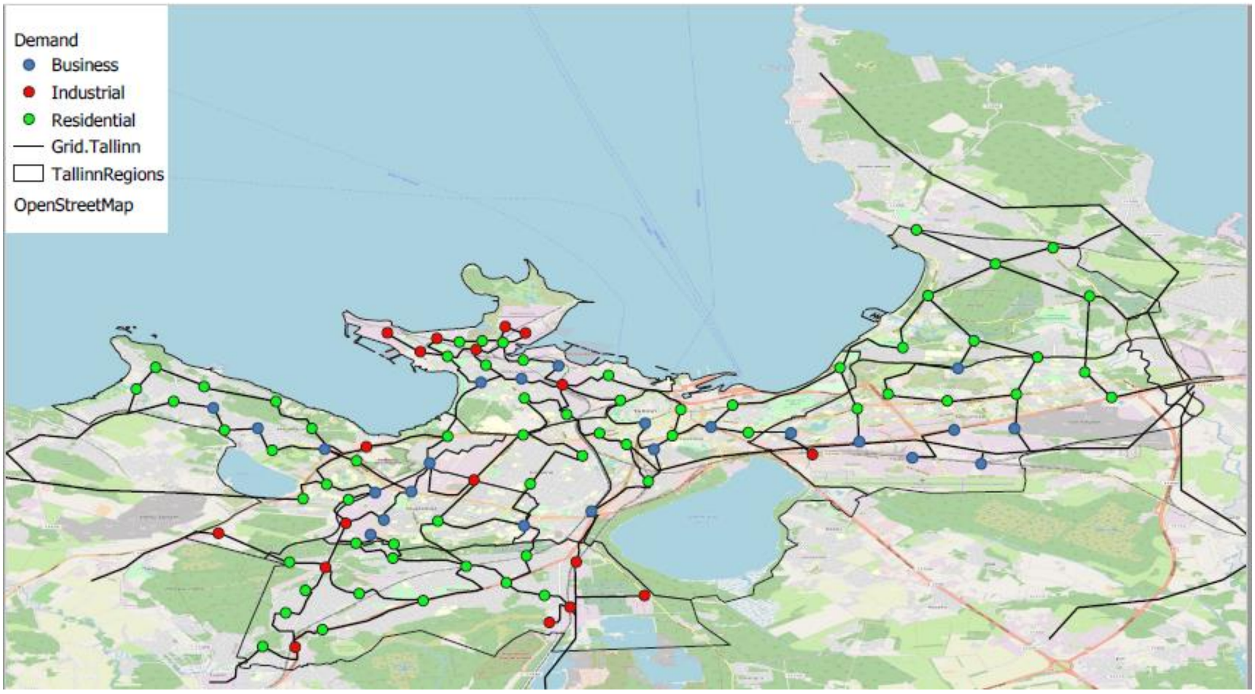

Figure 5.9. Tallinn energy demand points

The accuracy of the analysis is dependent on many items and one of those is to have different kinds of factors which are effective in the way of our analysis. These are the factors of Tallinn energy demand points out (Figure 5.10):

- *ID* which is simply the id number of the demand point on the map to give an identity.
- *longitude* and "latitude" show the location of the point in the specific CRS of the map.
- *population* is the number of people who are living in the region that the demand point located in it and it's important to help for a more accurate estimation for each demand point.
- *region* the region in which the demand is located and it is helpful to make a cumulative data for each region.
- *Kind* is the type of the demand point with three different types of residential, business and industrial, which also can help to determine the energy consumption in each point.
- terrain attribute shows the kind of the terrain in which the point is located.
- airport shows that demand is near the airports or not.
- water shows the availability of water in the location.
- *heat(GWh)* the total amount of heat needed for the whole region where the demand point is located.

- *elec(GWh)* the total amount of electricity needed for the whole region where the demand point is located.
- *solar(GWh)* is the PV generation potential at the regions.

For the last three attributes, the amount is for the whole region where the demand point is located and shows the total energy for heat and electricity demand there.

Demand :: Features Total: 104, Filtered: 104, Selected: 0

	ID	longitude	latitude	population	region	Kind	terrain	airport	water	heat(GWh)	elec(GWh)	solar(GWh)
1	93	24.622600	59.424200	11.600	haabersti	Residential	flat	no	yes	944.10	641.00	10.300
2	90	24.606900	59.430800	11.600	haabersti	Residential	flat	no	yes	944.10	641.00	10.300
3	91	24.617900	59.431500	11.600	haabersti	Business	flat	no	yes	944.10	641.00	10.300
4	92	24.623900	59.440300	11.600	haabersti	Residential	jungle	no	yes	944.10	641.00	10.300
5	85	24.577800	59.444400	11.600	haabersti	Residential	jungle	no	yes	944.10	641.00	10.300
6	86	24.584200	59.451500	11.600	haabersti	Residential	jungle	no	yes	944.10	641.00	10.300
7	82	24.632800	59.408200	11.600	haabersti	Residential	flat	no	yes	944.10	641.00	10.300
8	83	24.640600	59.412900	11.600	haabersti	Residential	flat	no	yes	944.10	641.00	10.300
9	104	24.890300	59.449900	6.360	lasnamae	Residential	jungle	no	yes	517.60	351.40	62.800
10	97	24.899100	59.441600	6.360	lasnamae	Residential	jungle	no	yes	517.60	351.40	62.800
11	98	24.853800	59.460300	15.300	pirita	Residential	jungle	no	yes	1245.00	845.40	15.800
12	9	24.706100	59.462900	10.100	pohja	Industrial	flat	no	yes	822.00	558.10	39.200
13	6	24.691800	59.460200	10.100	pohja	Residential	flat	no	yes	822.00	558.10	39.200
14	7	24.676900	59.461000	10.100	pohja	Industrial	flat	no	yes	822.00	558.10	39.200

Figure 5.10. Attribute table in QGIS

5.2.2 Energy consumption choropleth maps

Choropleth map is one of the materials for spatial data analysis. Historically, this kind of map was used to minimize the number of different colors that they need to make a map, but now it's one of the classification methods in geographical analysis. It has three aspects; Classification with same intervals, classification with same frequencies of objects in the classes and, statistically optimal classification [26].

By using choropleth maps, we can visualize the portion of essential factors for us in our analysis and pick some benchmarks to make an accurate classification of the regions according to a specific characteristic. Also, it gives us the ability to have an excellent comparative method between the critical factors of our research and indicates the direction of our classification [27].

5.3 Sptial weight allocation

Spatial statistics define the space and spatial connection and correlation according to their statistics like area, location, length, etc. Usually, these relations are represented with spatial weights. They are generated in the shape of spatial matrixes.

A spatial weight matrix is generated with a mathematical model and modifies the relationship between our dataset features. While the relationship's real value may vary from feature to feature, it is made according to a unified model and will generate the spatial matrix with rows and columns. Each cell's value has represented the correlation among those two features that have that cross-section cell [28].

There are many different models possible for weighting. Weights could be binary or variable. Binary weighting is generated from a fixed distance of features, space-time window, K nearest neighbors' data. For a particular target, the weight 1 is assigned to all neighbors and 0 to others, but inverse model variables have the range between 0 to 1, so in this case, being closer to the target means the value is closer to 1.

5.3.1 Moran's i matrix

The Moran's I model is one of the spatial correlation analysis models. As mentioned before, for this kind of analysis, we need the spatial correlation matrix at first to execute our analysis model. The Queen weight matrix is one of the various spatial weight matrices which represents the relation between a feature, and it's neighbors by which the Moran's I analysis is possible and gives us the value to understand the indication on the map and make a decision for reasonable allocation for the regions on the map [29].

In the QGIS, there are many tools for spatial allocation analysis. The *Visualist* plugin is one of them which has many functions, and one of those is Moran's I analysis. The Moran's I function of this tool will help us generate maps that can visualize our data analysis. Depending on the data and features that we want to analyze and the Queen weight matrix which will be generated in the QGIS software, this tool will make maps that show the significant aggregation of that feature on different regions of Tallinn.

This analysis will generate two different result maps. One will represent the aggregation of the target feature with proportional values in different regions by taking into account the spatial weight matrix value and the datasets available for that feature. The other one is a map with special legends. The map represents the probability of clusters and outliers [30]. Four different amounts of legend are:

- *HH*: the bunch of statistics with high values according to input data.
- *LL*: the bunch of statistics with low values according to input data.
- *LH*: brings out low values, which are surrounded by high values.
- *HL*: brings out high values, which are surrounded by low values.

The goal of Moran's I analysis is to make a classification of the regions with similar attitudes according to demand points which are located in it and the relation between regions.

5.3.2 K-means Clustering

The most typical and frequent problem in pattern recognition of features and working with statistics is the clustering problem. The clustering method is a way that helps us to find a group of data from our dataset, which are homogeneous points in that given dataset. These groups are called a cluster. In clustering, one critical issue is errors, and we should try to optimize our method to reduce mistakes. The popular clustering method which can minimize errors is the K-means algorithm [31].

K-means algorithm is based on dividing according to initial points which are given to it. This method is proposed by J. B. McQueen and is one of the easiest methods for clustering and is an iterative algorithm that is not only trying to make clusters of similar points but also tries to keep clusters as different as possible too [32].

In the QGIS as a powerful geo analysis software, there is also a plugin for clustering models in which the K-means model exists. With this tool, we can execute the K-means clustering model for different attributes of our data like electricity and heat consumption and PV potential. We could do this for the demand point that we have, to cluster them according to data we have collected from databases of Tallinn for energy consumptions and also PV potential estimation. In the QGIS, we can make different layers and put them on top of each other, and in this case, we can execute the K-means clustering not only for demand points but also for regions (which means polygons on the map) that have all the data of those demand point which are in that region. To compare and decide more accurately, we can merge these two different clustering maps and put them together on a single map.

5.4 Spatial regression

Regression provides us with an ideal case for examining how spatial architecture can help us understand and analyze our data. Typically, spatial structure helps models in one of two ways. The first way that space can affect our data is when the data generation process is clearly spatial. And the other one is when we are analyzing data that we get wrongly, which is common in econometric.

There are three different models for regression:

- *OLS*
- *SAR*
- *SEM*

The first one is ordinary least squares (OLS) which has the bellow function:

Y=XB+e

Y is a vector of dependent variable observation. X is a value matrix of connected variables, B is a vector of coefficients for the regression model, and e is a vector of independent error according to the normal distribution.

The next one is the spatial autoregressive (SAR) model, which is presented the presence of spatial autocorrelation for the data which are connected. However, it will take into account the spatial independence of the error terms [33].

The equation of SAR is:

Y= PWy+XB+e

Where Y is a vector of dependent variable observation, P is the autoregressive coefficient, W is a weight matrix which is spatially lagged dependent variable, X is a matrix of values of connected variables, B is a vector of coefficients for a regression model, and e is a vector of independent error according to the normal distribution. The spatial error model (SEM) is also the other model of spatial regression. The equation of SEM is:

Y=XB+e and e=LWe+u

Y is a vector of dependent variable observation. X is a value matrix of connected variables, B is a vector of coefficients for the regression model, and e is a vector of independent error according to the normal distribution. The equation of e also consists of W, which is a weight matrix spatially lagged error term, and u is a vector of independent errors.

The OLS model is usually used, but there is a need to analyze data with spatial autocorrelation in geo analysis, so the SAR and the SEM due to their functionality to have weight matrices and autocorrelation coefficient in their equations are better choices.

After the analysis, we can generate some plots and graphs to represent the process's result in a visual way. Those are:

- Residuals vs. fitted value graph
- Quantile-Quantile graph
- Scale-location graph
- Standardized residuals vs. leverage graph

When we talk about the regression analysis, the one most frequently used is the residual versus fit value plot. This is a scatter plot representing the residual values and the estimated response to detect non-linearity, errors, and outliers.

The Quantile-Quantile plot is a visual representation model whether two data sets come from a similar distribution group or not. With the quantile, we mean the fraction of points of data below the limit value. For example, the 0.2 quantile implies that 20 percent of the values fall below and 80 percent above the limit indication [34].

Scale-location plot is used to present that if residuals are spread proportionally, and the ranges of inputs are not. The plot consists of fitted values in the horizontal axis and the

root of standardized residuals in the vertical axis. If the points are spread randomly in a horizontal way, it is one feature that can show that the model we used is suitable [35].
The plot of standardized residuals versus leverage finds the influential outliers. Even if they have extreme values, we cannot determine the regression line with them. It shows that they could not significantly impact the result either with or without them in the analysis [36].

6. CASE STUDY AND RESULTS

In this study, the goal is to make the energy distribution system of Tallinn perform more efficiently and try to push it to be more sustainable and greener according to the energy production path. The study has some analysis to make an alternative way for traditional energy production and distribution for this goal: to use potential solar energy by implementing PV panels to generate some portion of energy production and make the system more efficient by making a new distribution model containing small regions, using classification models.

6.1 Geo-spatial information of Tallinn

For making an informed decision, there are many factors which could be effective. Here for monitoring the condition of Tallinn, population and construction have been investigated. These two factors are extremely important to make a sound analysis in accordance with the energy sector in the amount of energy needed to be produced and the method of distribution of the energy.

6.1.1 Population

The data about the population is essential statistics. All the governments have a vast investment in making an accurate dataset related to the population situation of different cities, which can help them decide for the best way to put their investments to make welfare amenities for their people.
In Estonia, this job is for Stat organization. According to the statistics, the population of Estonia is not big but, it is also growing yearly Table 6.1.

Table 6.1 Population of Estonia

Year	Population
2016	1,315,944
2017	1,315,635
2018	1,319,133
2019	1,324,820
2020	1,328,976
2021	1,329,460

Although the statistics are limited, it is clearly visible in figure 6.1 that the number of people in the Harju as the most crowded county of Estonia has a positive index in the beginning and also end of each year which shows the total growth of it.

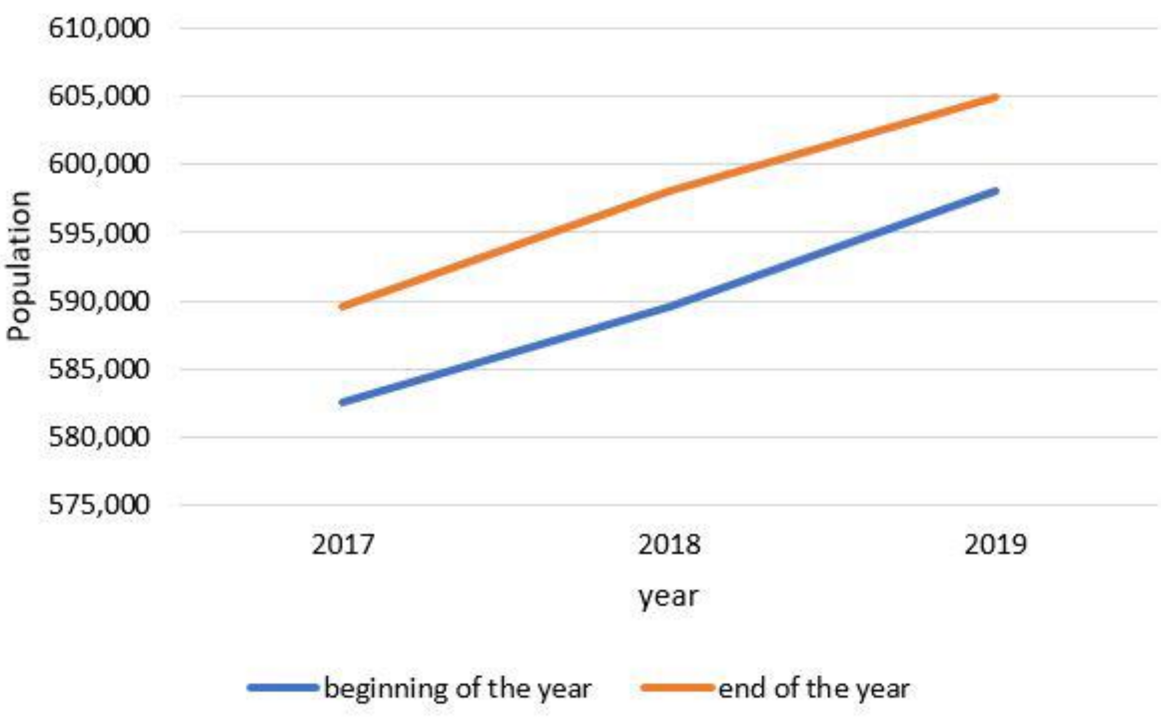

Figure 6.1. Total population of Harju

6.1.2 Building construction

One of the issues will come after the growth of the population is dwelling for them. This issue brings not only a need for constructing new buildings but also the need for new energy infrastructures to supply them. According to statistics this growth can be understood by the number for new building construction permits Figure 6.2.

Figure 6.2. Dwelling construction permit and completed

6.1.3 Residential

In figure 6.3 we can see the rapidly growth in building new residential buildings in Estonia since 2016 to 2020 which is more than thousands of new buildings each year.

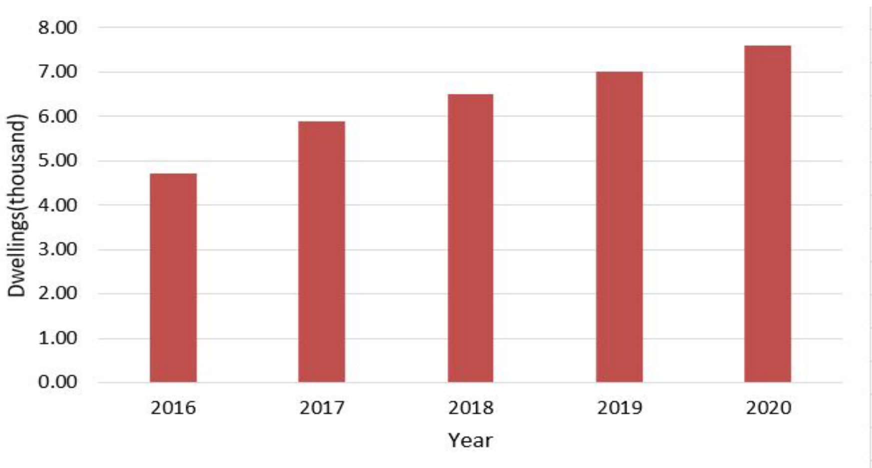

Figure 6.3. Residential constructions of Estonia

6.1.4 Non-residential

Residential is not the only construction which is affecting the need of energy; these new populations also need many other services like, schools, hospitals, stores, work offices, etc. which we can call non-residential buildings and all need to have new construction and also new infrastructure for energy distribution systems. Figure 6.4 shows that the number of non-residential buildings is not increasing frequently, but that the growth is significant in some years.

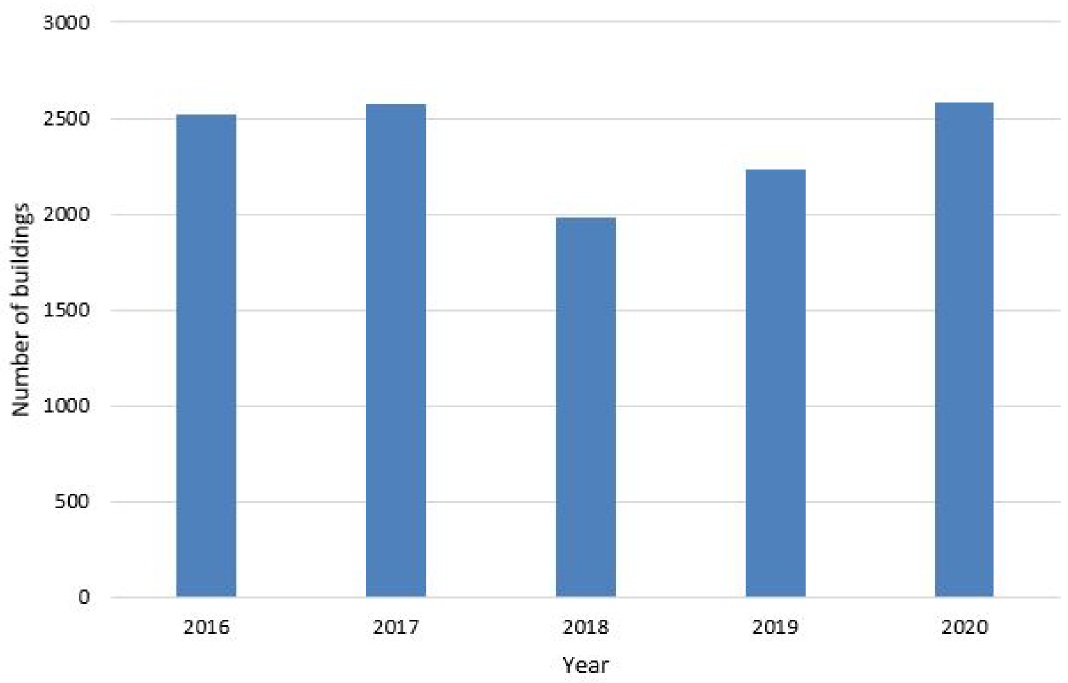

Figure 6.4. Non-residential constructions of Estonia

Also, in Figure 6.5 by using statistics the building layer of the map has been composed on the base map of Tallinn. This layer is a polygon layer which has each building as a component in it and represents them with polygons.

Figure 6.5. Tallinn building layer map in QGIS

Then with functionality of calculation tools in the QGIS figure 6.6 has been made which has the number of buildings in any region of Tallinn.

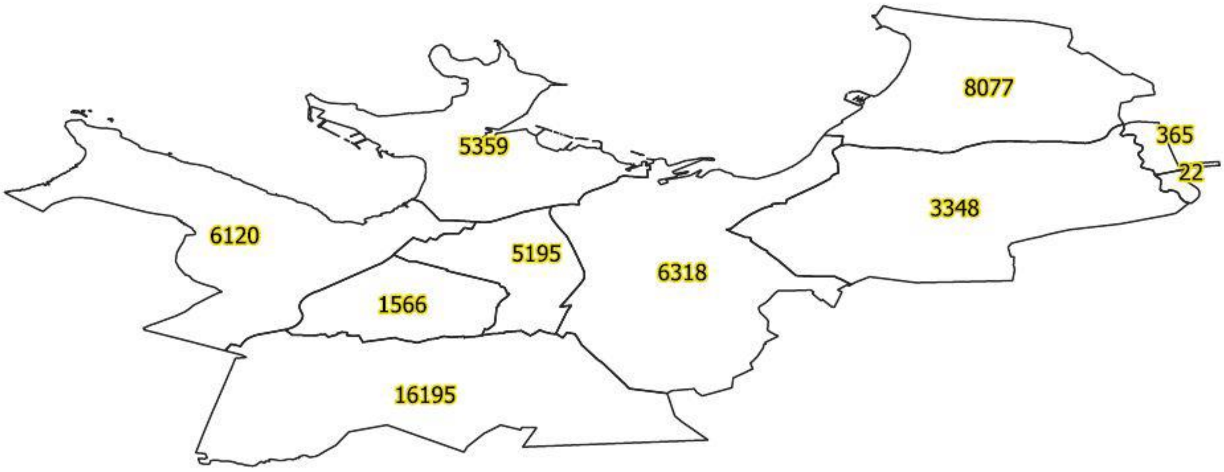

Figure 6.6. Number of buildings in each region of Tallinn

6.2 Patterns and trends analysis

Energy is one of the critical issues of humans in the world. Humans reached this standard level of living during years where it was not possible without using coal, petrol, water power, and other kinds of energy. Since the beginning of the Industrial Revolution, the amount of energy consumption in the world has had a significant rise. Before discovering huge barrel oil reserves and natural gas, coal was the primary energy resource in the world. After that, barrel oil and natural gas are the main energy resources. However, these days the majority of attention is focused on renewable energy resources to keep the world cleaner due to less CO2 emission and being sustainable [37].

6.2.1 Energy production

The factors which indicate the energy issue are the capacity of production and the amount of consumption. The clear part is that by following the growth of population, and the rapid number of new building construction, we need to produce more energy than before, and it is not only the electricity which is needed but also the heat demand will be increased rapidly. By investigating the databases [21] we can reach the good data sets about the energy production of Estonia for years. Figure 6.7 shows the annual amount of energy that was produced in Tallinn from 2015 to 2019.

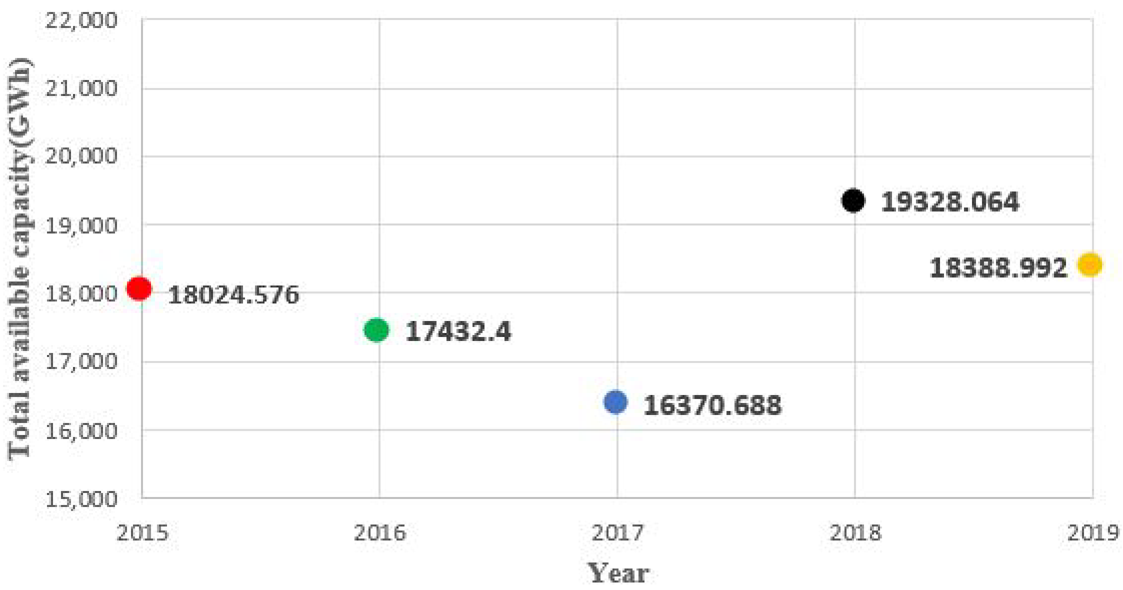

Figure 6.7. Yearly energy production for Tallinn

6.2.2 Energy consumption

Figure 6.8 presents the vast amount of data (more than 9000 data sets) collected from Stat website, about the consumption of energy in Tallinn hourly through the whole year 2019. These amounts are separated into three different kinds of energy consumption. These kinds are:

- Residential
- Industrial
- Business (offices, stores, malls, sport centers, etc.)

According to Figure 6.9, the total consumption of Tallinn is divided into 104 hypothetical demand points, which are located in different regions on the map of this city in the QGIS as a vector layer which gives us the ability to work on the data of these demand points and make our analysis on them. Figure 6.9 shows the way of visualization of these demand points in QGIS maps.

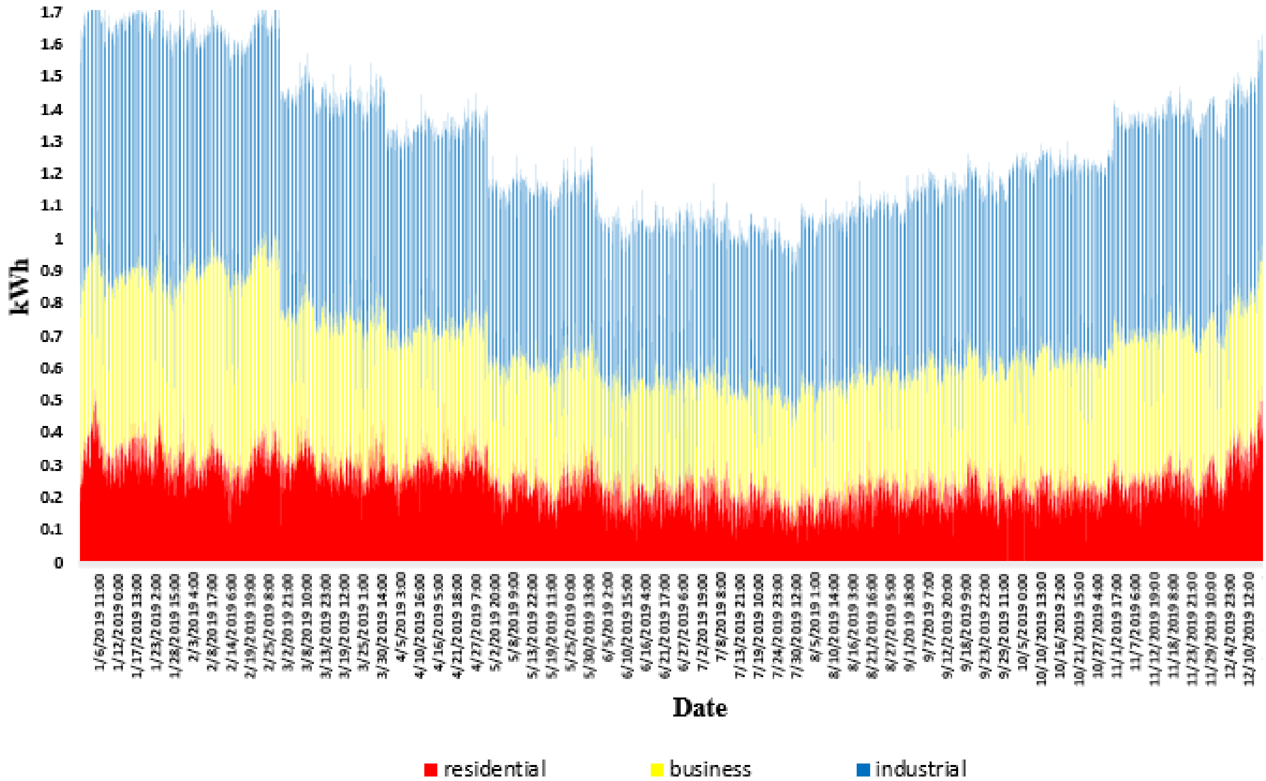

Figure 6.8. Tallinn energy consumption of 2019

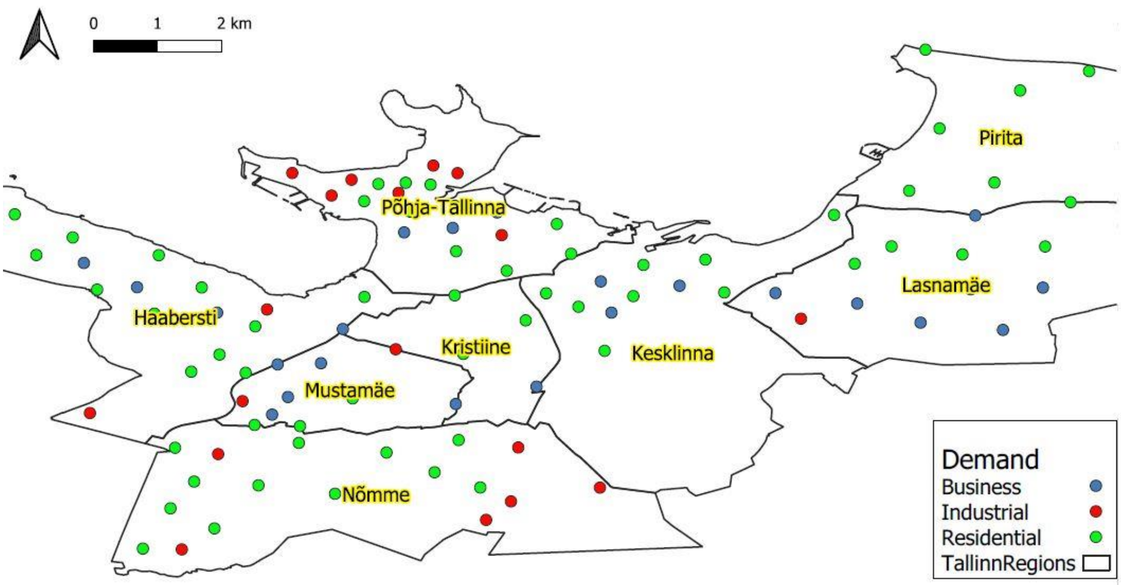

Figure 6.9. Tallinn energy demand points in QGIS

6.2.3 Energy consumption choropleth maps

Below, there are three different choropleth maps for heat consumption, electricity consumption, and also PV generation potential, which show us the dispersion of them on the map. Figure 6.10, 6.11, 6.12 have been generated with QGIS tools to help us have a more accurate analysis for heat and electricity consumption of Tallinn. With these maps, we can compare the portion of the usage in different regions of Tallinn.

Figure 6.10. choropleth map of electricity

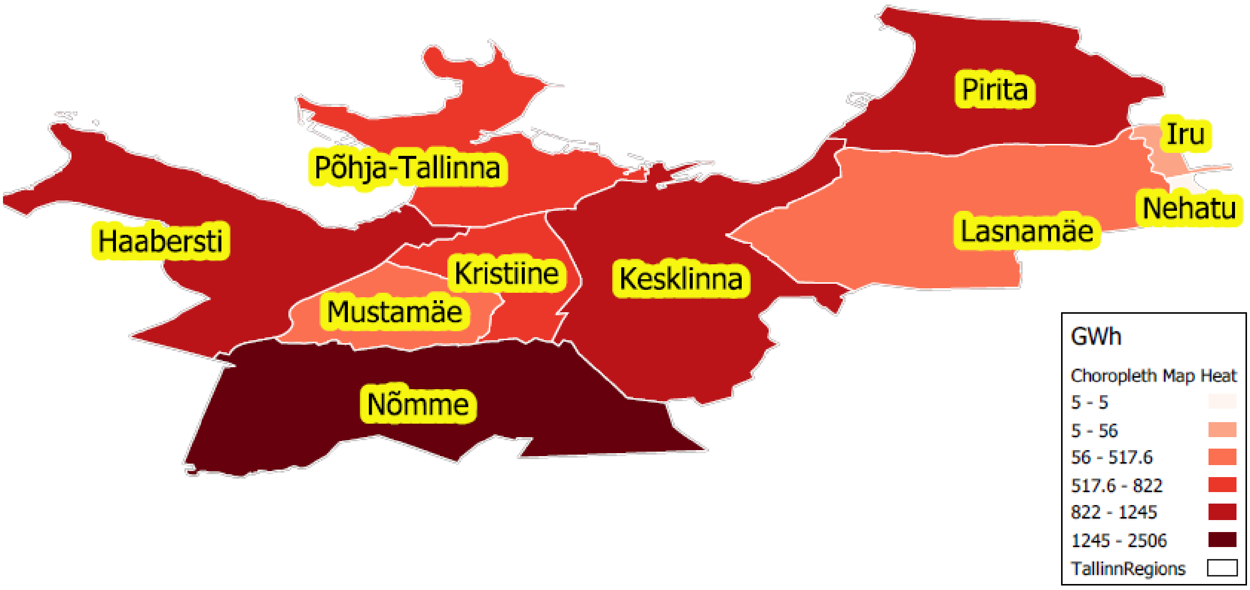

Figure 6.11. choropleth map of heat

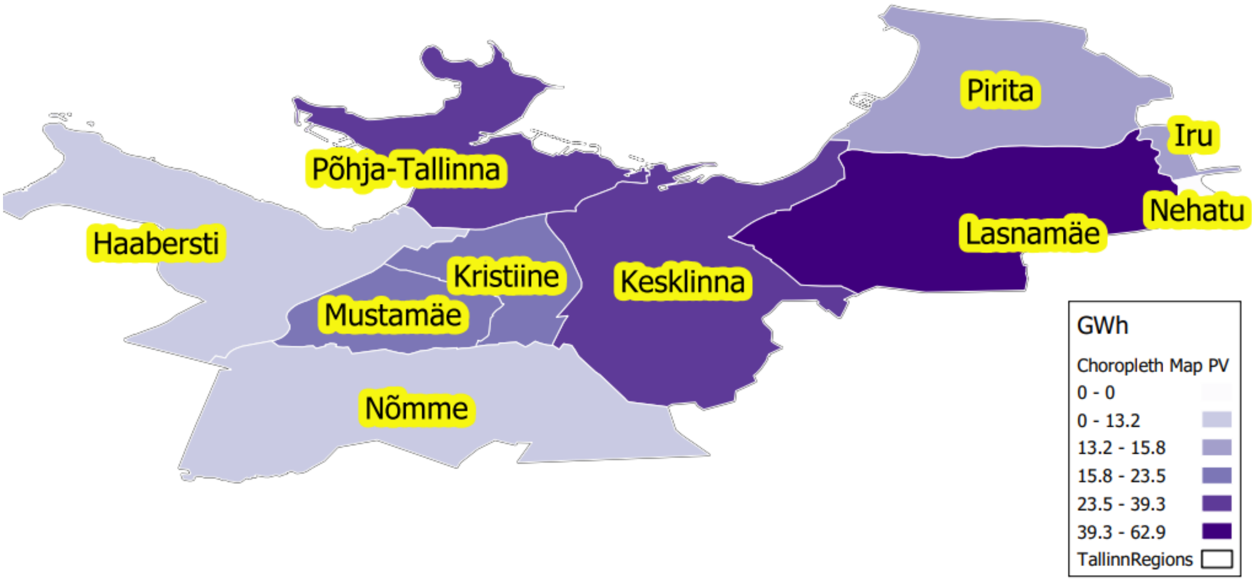

Figure 6.12. choropleth map of PV

6.2.4 Quantile-quantile plots of energy consumption

One of the plots that will help us find the attitude and trends of energy consumption in different regions is the quantile-quantile plot. By using R-studio software and execution of the analysis codes for quantile plots, Figure 6.13 has been generated. This plot illustrated that the average electricity consumption of different regions is around 550 GWh per year. The region with the lowest consumption is Nehatu with 5 GWh electricity, and Nomme has the highest consumption, which is 1702GWh. There is also the heat consumption plot in this plot, which illustrated that the average consumption is 800GWh per year. The region with the lowest consumption is Nehatu with 5GWh, and the highest one is Nomme with 2500GWh.

6.2.5 PV potential

These days because of the concern for the future of the world, most of the governments all over the world are trying to push and encourage energy companies for more renewable energy products and systems, and they also have huge investments in this field too. European countries are the leader in this way to go through green industries and keep the world cleaner than before.

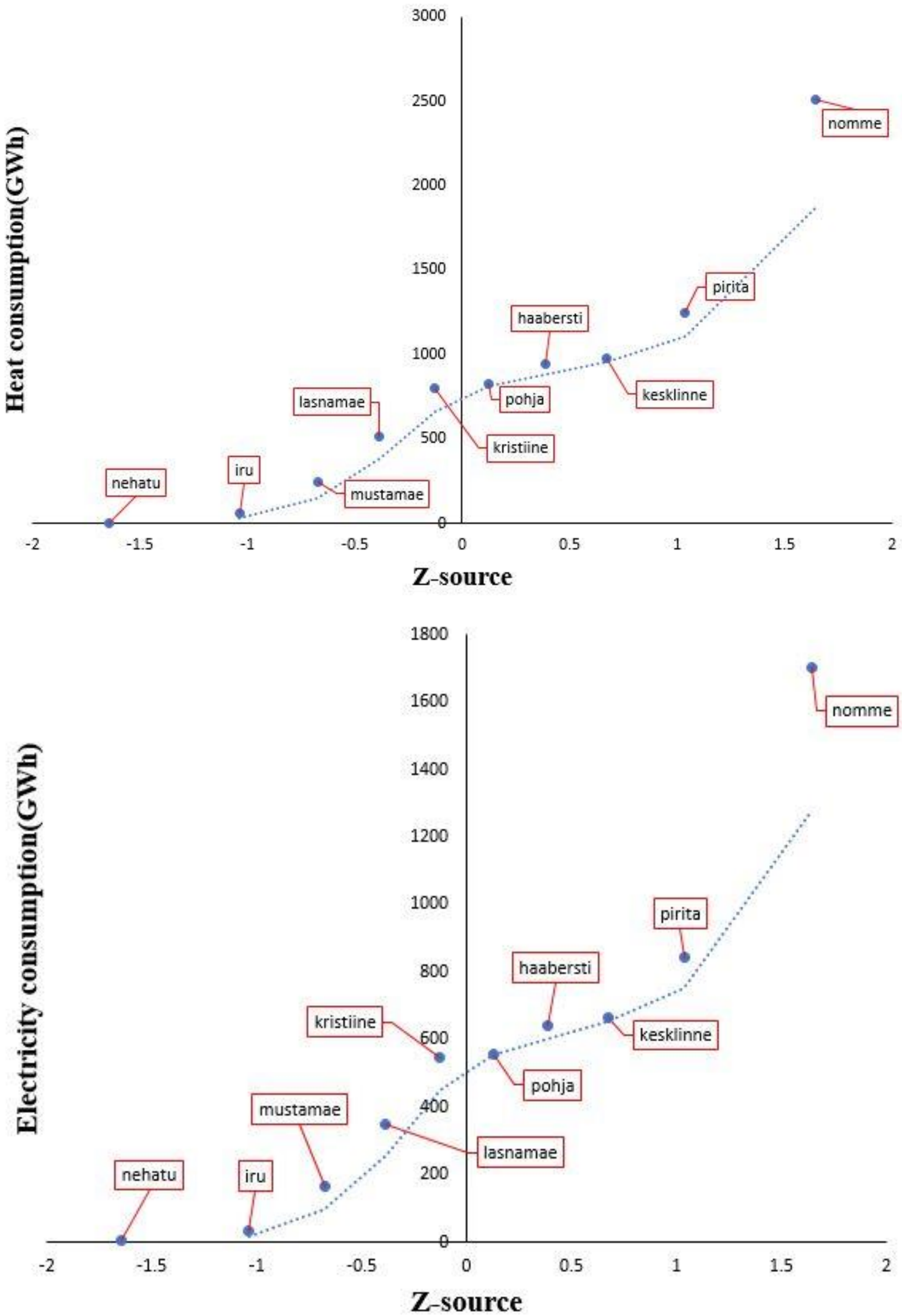

Figure 6.13. Quantile-Quantile plots

In Estonia also there are many activities in this field, and one of the cleanest energies, which is PV, is the point of focus of many companies there. Eesti Energia has a tool on the company's website [38] where we can find out the PV potential of our rooftop area. In Figure 6.14, you can see that in this tool, you need to have some inputs to have an estimation according to them. The data you need to put are, self-consumption, the installation method, the rooftop area and, the direction of the building.

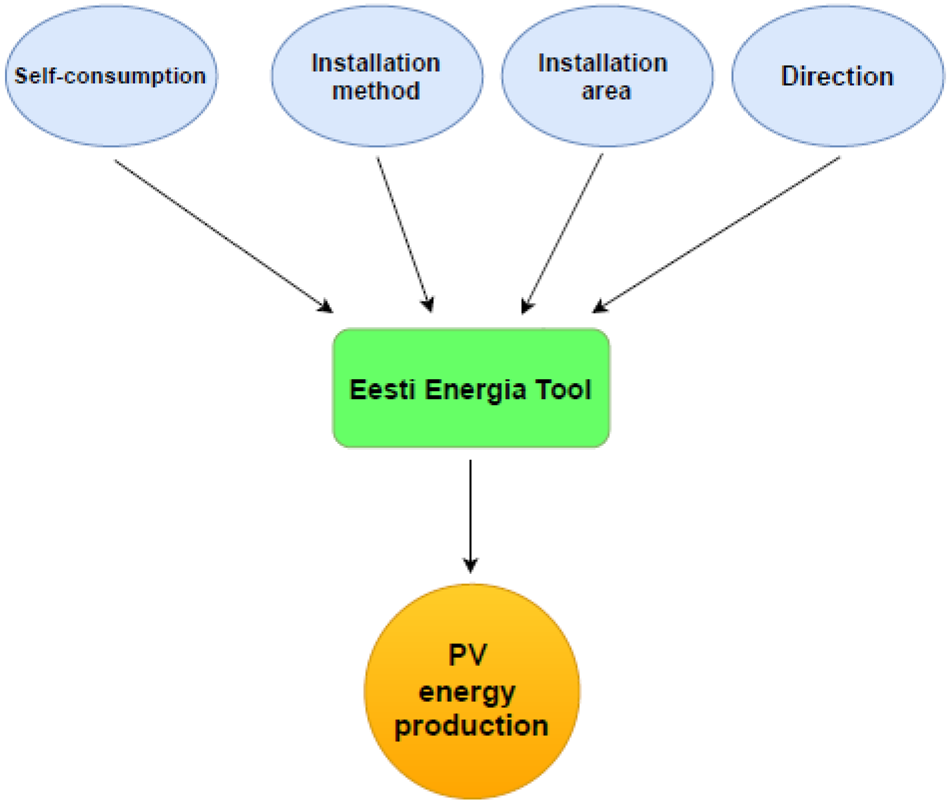

Figure 6.14. Eesti Energia tool mechanism for PV production estimation

Self-consumption is the total amount of energy consumption by the target building during a year in kWh, the installation method where you need to verify the kind of building rooftop that you considered to put PV panels on (flat or gable), the area which is suitable for installation of panels on the roof area in square meter and the direction of the building which indicates the installation direction of the panels according to the available sunlight angle at that region.

For PV systems there are two different connection model (figure 6.15):

- Off-grid
- On-grid

The most significant difference between these two models is batteries bank. In the off-grid model, we need to have a battery bank to store the energy in low-demand times when the demand is higher. But there is no need for a battery bank in on-grid, and the extra electricity production will go directly to the grid [39]. In this study on-grid model has been used without any battery bank.

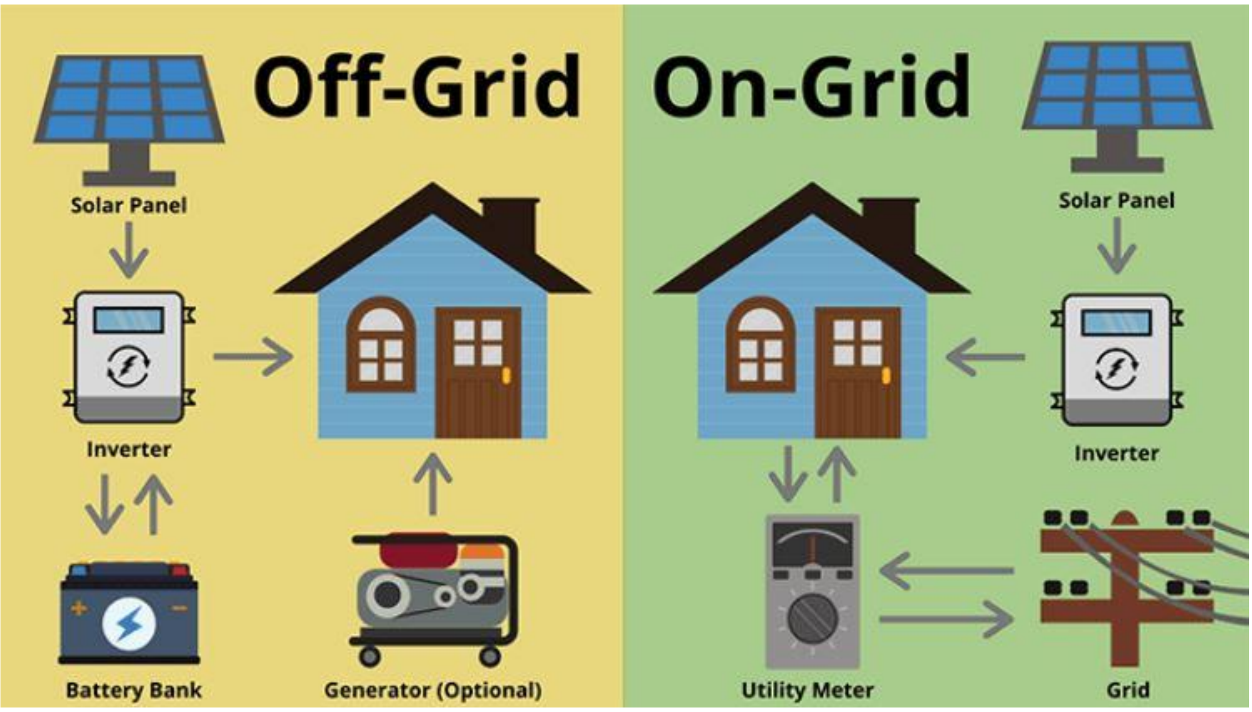

Figure 6.15. On-grid and Off-grid PV models illustration

6.2.6 PV modules

The production estimation of Eesti energia tool is based on:

- Recom panel
- Huawei inverters

Solar panels have at least a 25-year life span. Their capacity decreases about 0.6% per year, and after 25 years, the manufacturer guarantees a power output of around 85%, but inverters should be replaced approximately 15 years after installation [40]. Panels that will be used are for the Recom solar group. The name of the model is Black panther, and these panels are monocrystalline modules. According to Figure 6.16, the efficiency of these modules is more than 97% output for the first year, from year 2 to 24 it will decrease yearly less than 0.65%, and after 25 years the output could be more than 81%.

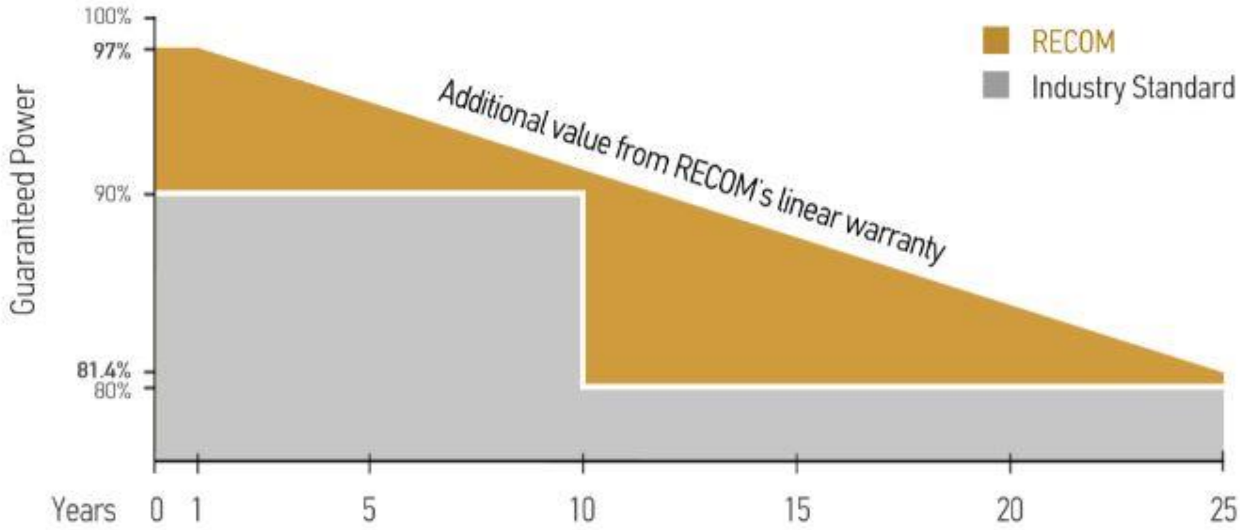

Figure 6.16. Recom modules efficiency diagram

Also, according to data sheets of modules, the power tolerance is 5 watts, and the maximum power voltage 41.7 volts, the maximum current is 9.61 ampere, they operating temperature is -40 to +85 degree centigrade, and the total module efficiency is 20.18%. With the Eesti Energia tool for PV production estimation, we could have some results by putting all the inputs. For a building with a rooftop area of 500 square meters, a flat roof located in the northwest of Tallinn and has the self-consumption of 30MWh in a year, on this rooftop, we can approximately implement around 40kW solar panels. Aside from its own need, it could generate around 20MWh of electricity for the grid too.

By using the QGIS, we have the vast power of selectivity of objectives. In this case, the ability to select polygons is helpful. In the QGIS, buildings are a bunch of polygons in a vector layer on the map, and each polygon represents a specific building. It could have many other data about it inside an attribute table. We need to choose buildings that are useful for us to implement the PV panels on the top of them with the specific area on the rooftop. The QGIS let us do that with geographical tools, which are working with vector layers on the map. We can program the tool in the field calculator, which is one of the useful tools in QGIS, with our method and functions to choose all the polygons of that layer that have the specifications that we are looking for (Figure 6.17).
So, as you can see, there are two choices. One is making a new attribute column on the attribute table with specific functionality or method that we will put for it. The other choice is to update an existing attribute column with our particular data.

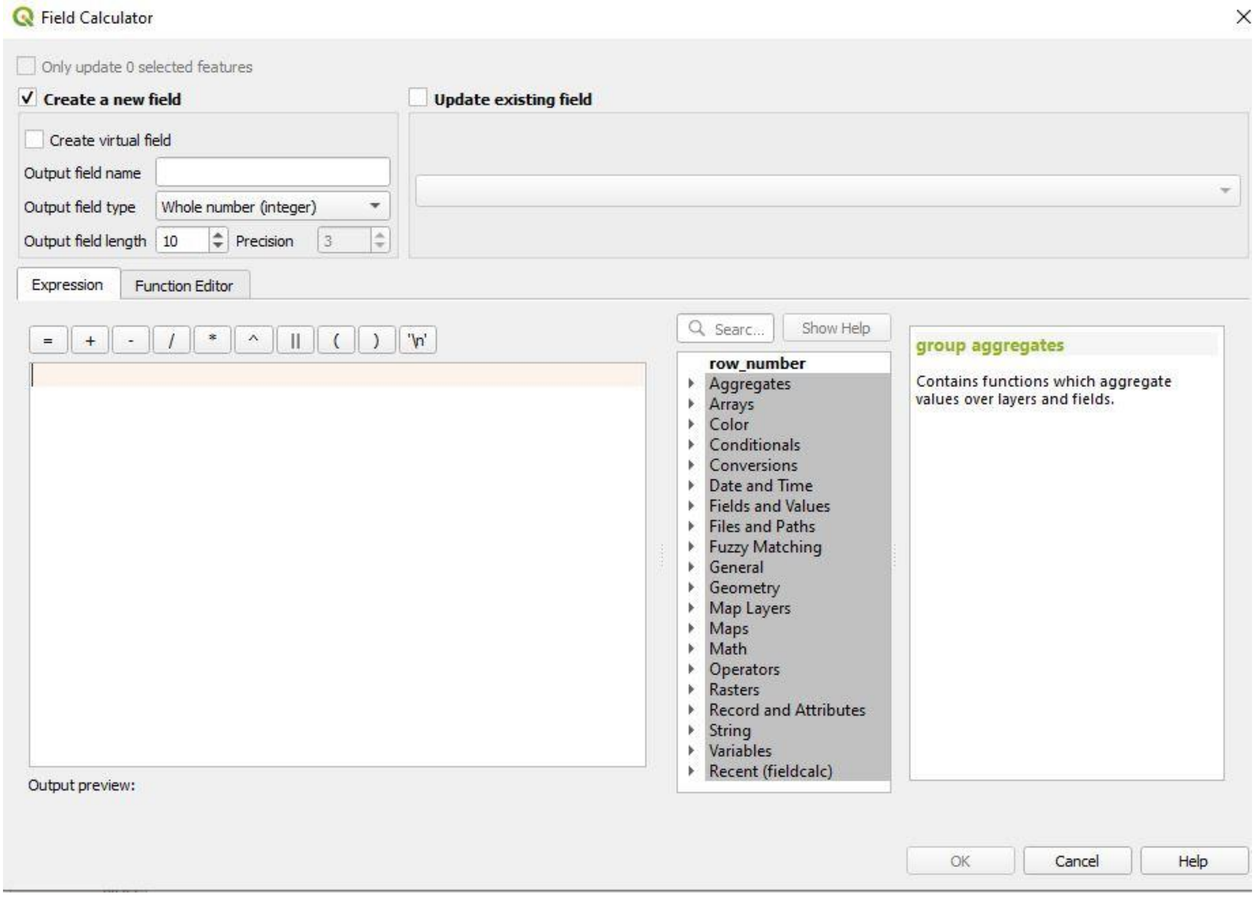

Figure 6.17. Field calculator of the QGIS

To execute each of these structures, we need to place the expression which is selected with our objectives but with the functions of the QGIS or place our selecting program functions in the function editor, which is adapted with the programming language Python. In this case study, the goal is to find buildings with rooftop areas larger than 3000 square meters as potential locations for implementing PV panels. To find these possible locations, we can use the building vector map layer of the QGIS (refer to first chapters where we made it). By making a query in Figure 6.18 for the buildings that we are looking for can distinguish them from others. In Figure 6.19 these target buildings are visible in yellow as a separate layer on the main map of Tallinn's building. In this scenario 521 unique locations in total have been found for PV production.

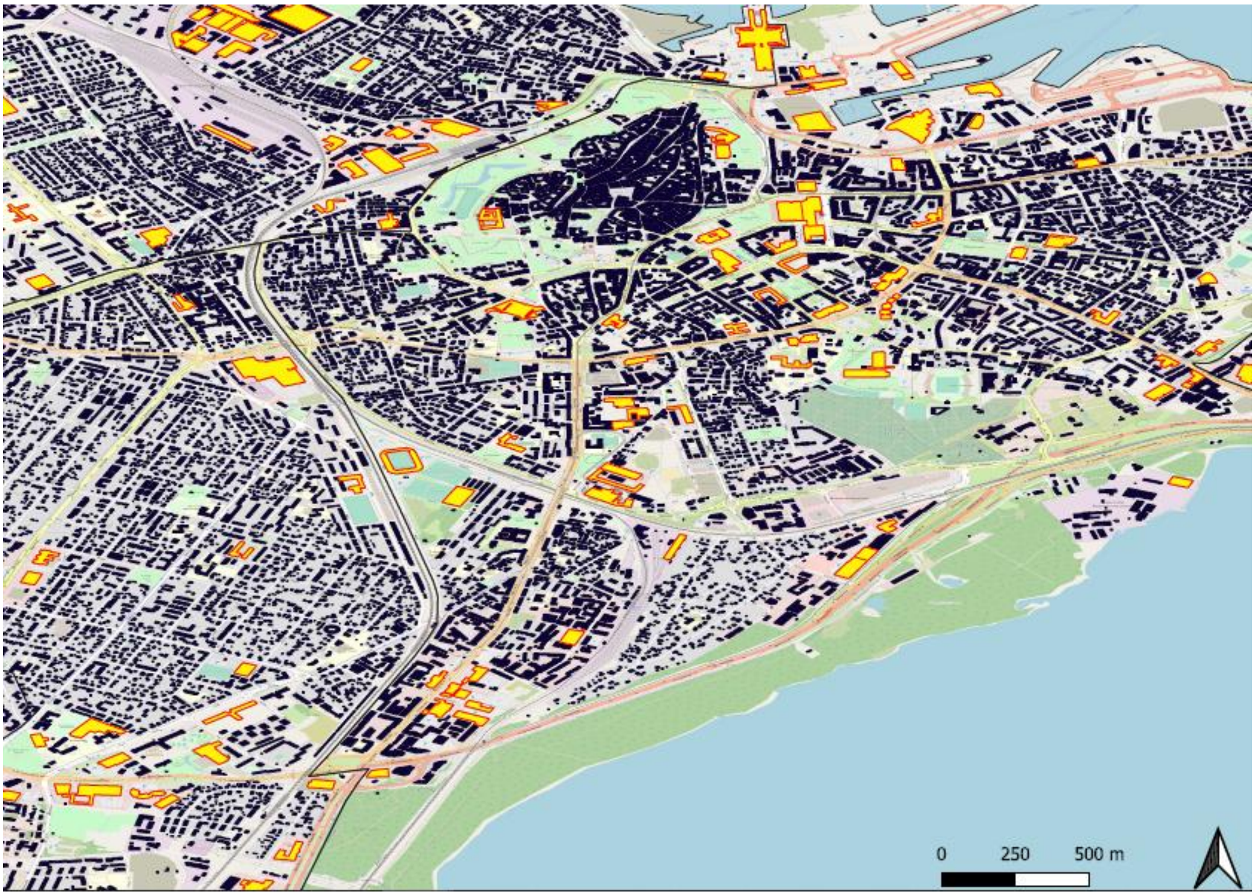

Figure 6.18. Map of Tallinn's buildings with rooftop area wider than 3000 square meter in QGIS

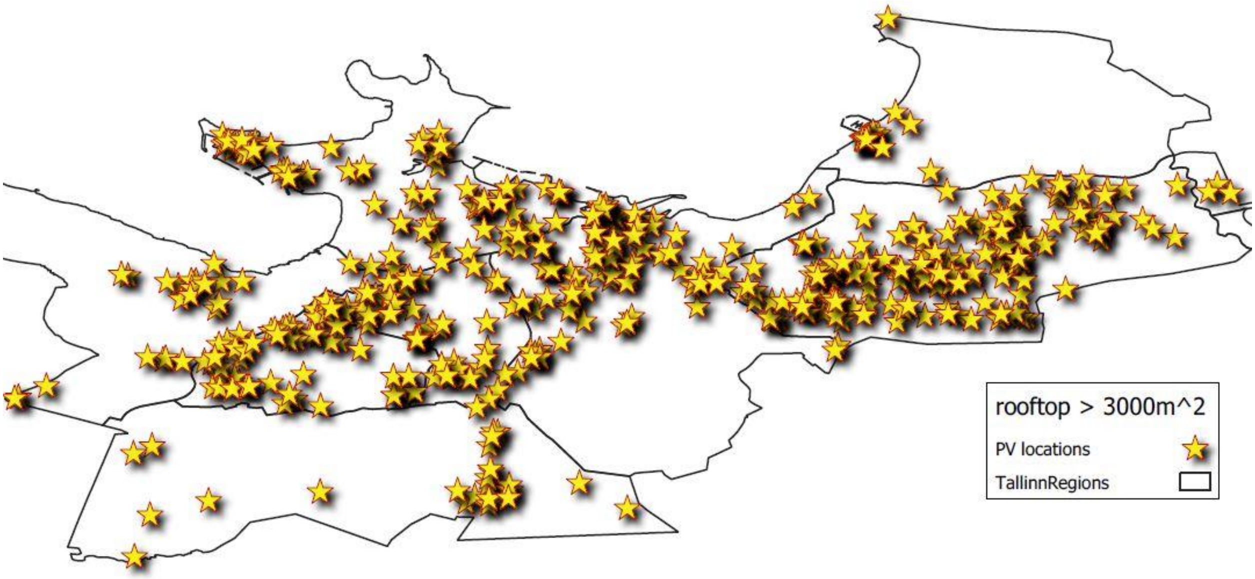

Figure 6.19. All PV locations on the Tallinn's map in QGIS

Figure 6.20 represents the trend of PV production units. Using the R-studio software and execution of the algorithm code for quantile-quantile analysis like what has been done for the electricity and heat consumption, we reached out to this result. It shows that in the Nehatu region, there is no proper PV production unit according to the scenario of this study. Other regions have the production potential between 10-25 GWh per year instead of Pohja and Lasnamae. Pohja has a potential of around 40GWh, and the region with the most potential is Lasnamae with more than 60GWh.

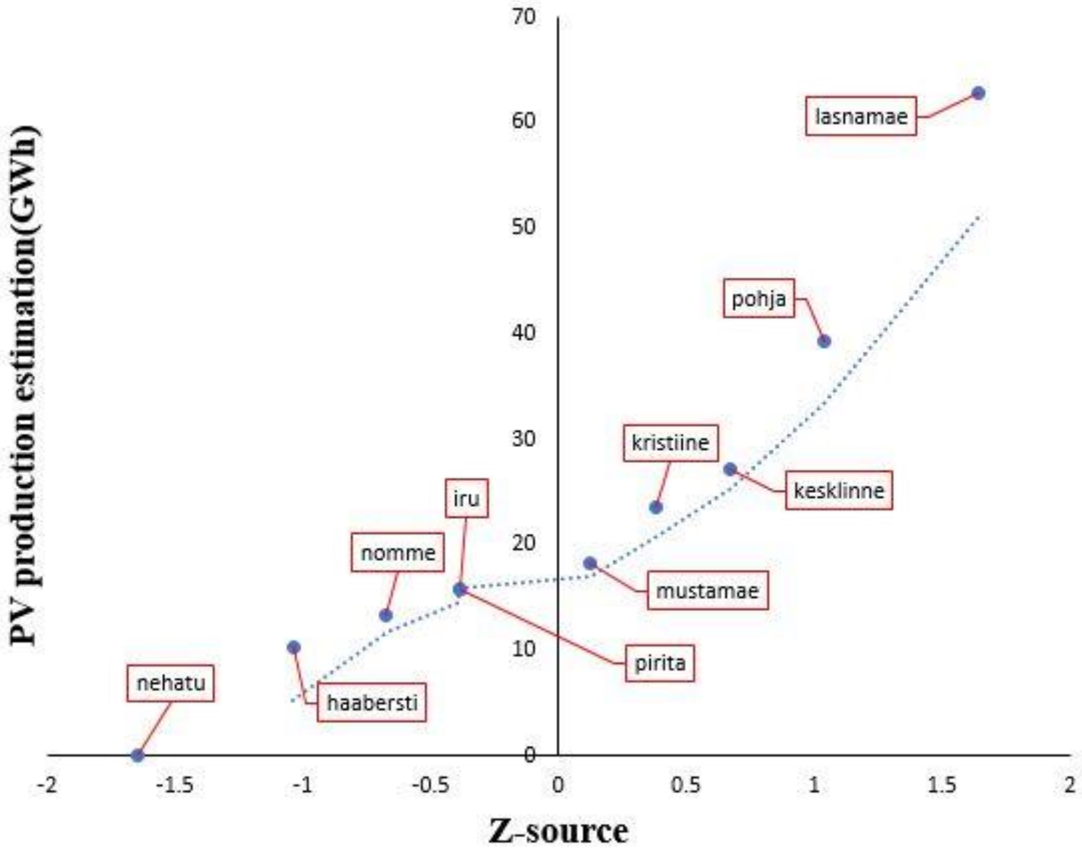

Figure 6.20. PV production Quantile-Quantile plot

6.2.7 TALLINN power grid

Figure 6.21 represents the primary power grid of Tallinn. The map has been illustrated using data from Elering group website [41] and Stat website. Data collected from State has the locations of the grid's distribution substations. The grid line has been drawn by assessing the initial grid lines in the Elering group maps and merging that with the base map of Tallinn, which we have made before, and finding the exact locations on the open street map.

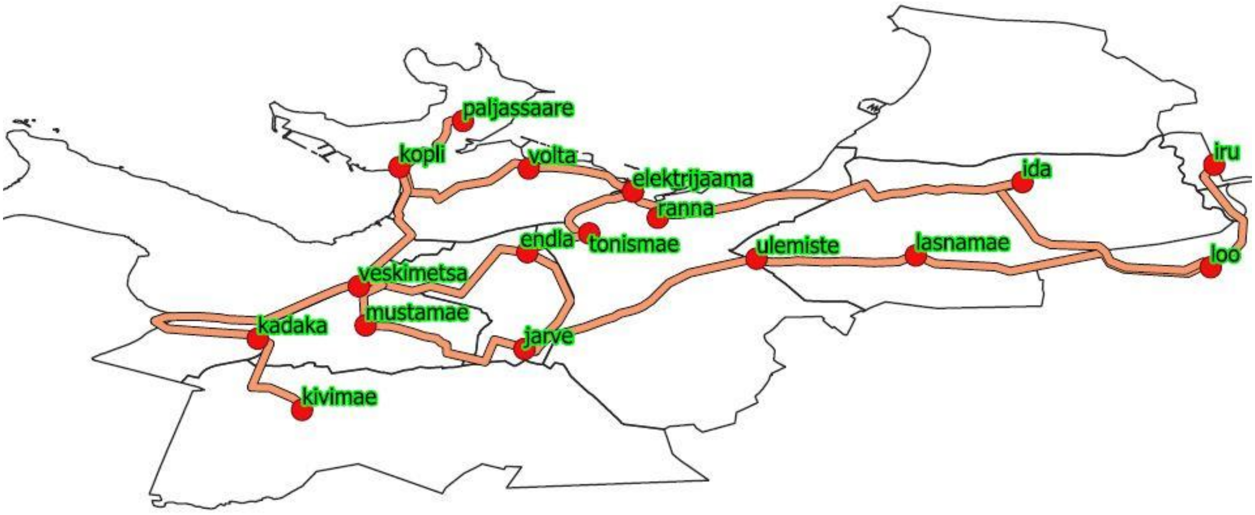

Figure 6.21. Tallinn power grid

6.2.8 PV production units connection to the grid

When the PV production units are allocated in the desired locations, and by having the primary power grid of Tallinn, connecting them can be executed. In QGIS, the PV locations layer and the Tallinn layer's primary grid are generated. By using the functionality of geographical analysis tools, a specific algorithm was executed to make a new layer that illustrated the PV production units path to the nearest primary grid substation of Tallinn. In Figure 6.22, the map represents the path of all the PV units shown by yellow points to the nearest grid substations shown by red points with red lines. In this scenario, the average connection distance of the PV units to the grid was 1.17 kilometers, and the sum of all connection lines was more than 600 kilometers.

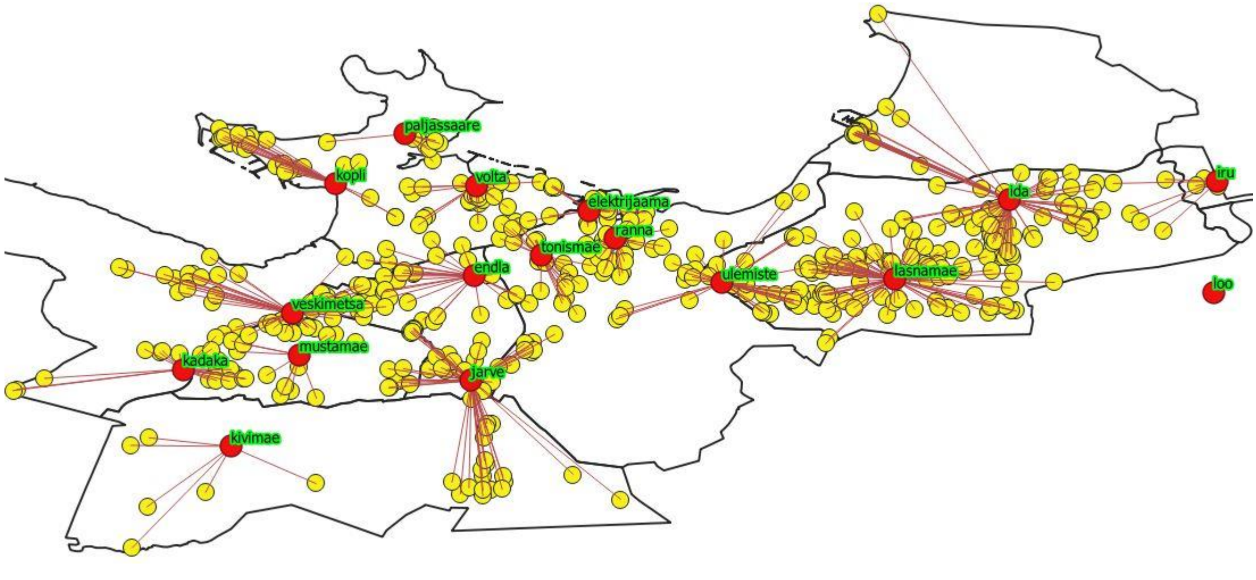

Figure 6.22. PV production units connection to nearest grid substation

6.3 Classification of energy consumption and production in TALLINN

One way to help us to make more accurate decisions about wide issues is through classification. The traditional way is to decide and manage the entire city as a whole. In this case, the classification of regions will help us manage different regions more precisely according to the attitudes of the demand in the region and make informed decisions for production units of energy there. Two different classification algorithms have been used in this study to make the classification considering the energy factor.

6.3.1 Moran's I

In the QGIS, *Visualist* plugin has many functions, and one of them is Moran's I analysis. Figure 6.23, 6.24 have been generated with this tool. The legend of the map (Figure 6.23) shows the significance of the aggregation of energy demand in different regions of Tallinn. To make a Moran's I analysis a proper weight matrix of the data is a significant material that we need. In this case, the analysis was made by the Queen weight matrix of demand point data.

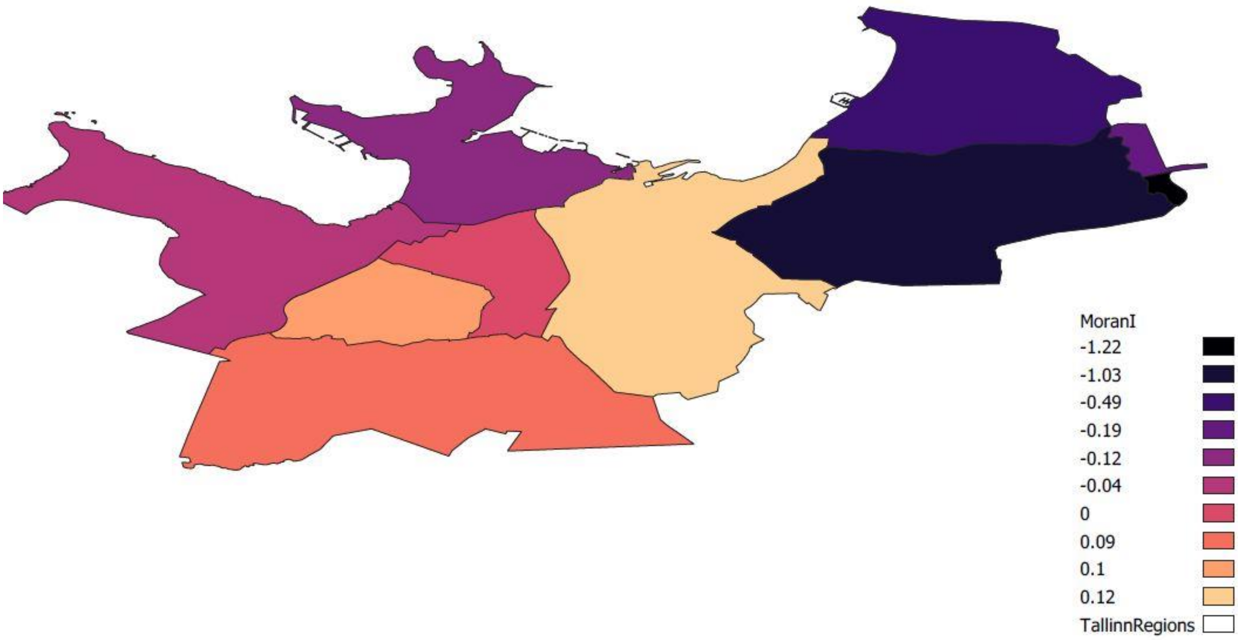

Figure 6.23. Moran's I

The map in Figure 6.24 represents the probability of clusters and outliers, in which HH refers to a cluster of statistics with high values and LL refers to a cluster with low values according to data collected from the demand layer. Also, the LH brings out low values, which are surrounded by high values, and HL is the reverse of it.

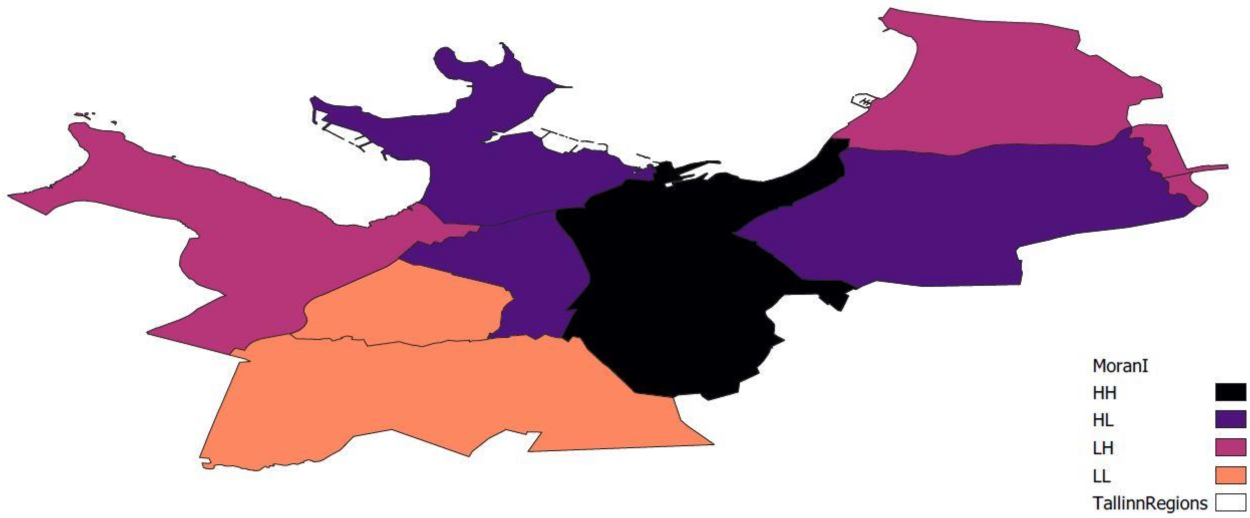

Figure 6.24. Moran’s I for clusters and outliers

The Moran’s I analysis aims to make an appropriate classification of demand points with similar attitudes. This classification can help make this analysis complete by implementing it on the regions and making classified polygons on map. Each contains some of the demand points with similarity and close relations.

6.3.2 K-means

In the QGIS as a powerful Geo analysis software, there is also a plugin for the K-Means algorithm as well, by which the below analyses are presented on the map. Figure 6.25 has been generated with the K-Means analysis for the demand points and has made five different clusters of all demand points. Also, in Figure 6.26, the same analysis was executed on Tallinn regions, in which areas have the data of all the points located in these regions. Both analyses were made to generate 5 clusters. However, the analysis for regions shows four different clusters with K-Means analysis for regions in which there are demand points. One region with no demand point is the other cluster which is malware data.

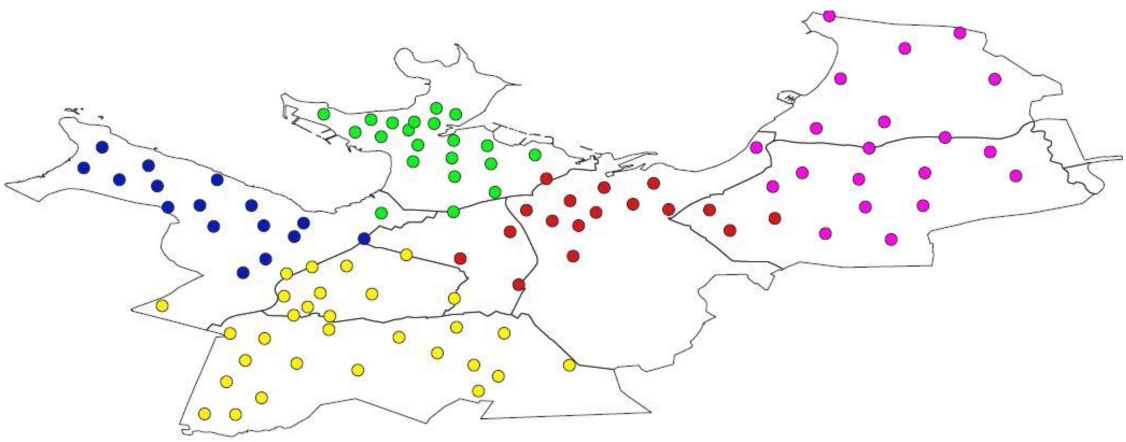

Figure 6.25. K-means for demand points

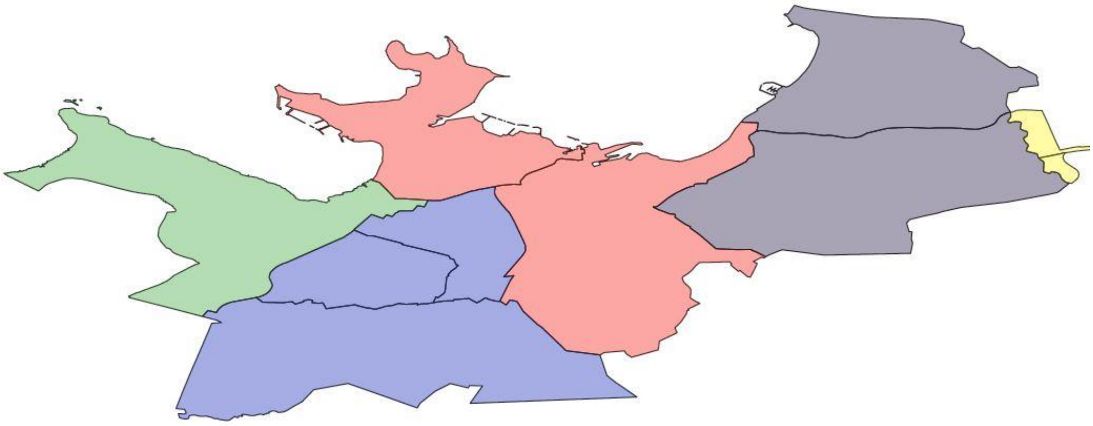

Figure 6.26. K-means for regions(polygons)

By merging these two maps (Figure 6.27) we can realize that all the clusters are almost similar in both demand points and regions, and just the green and the red cluster of demand points could have some relations and maybe have similar attributes that in the regional analysis they are in the same group.

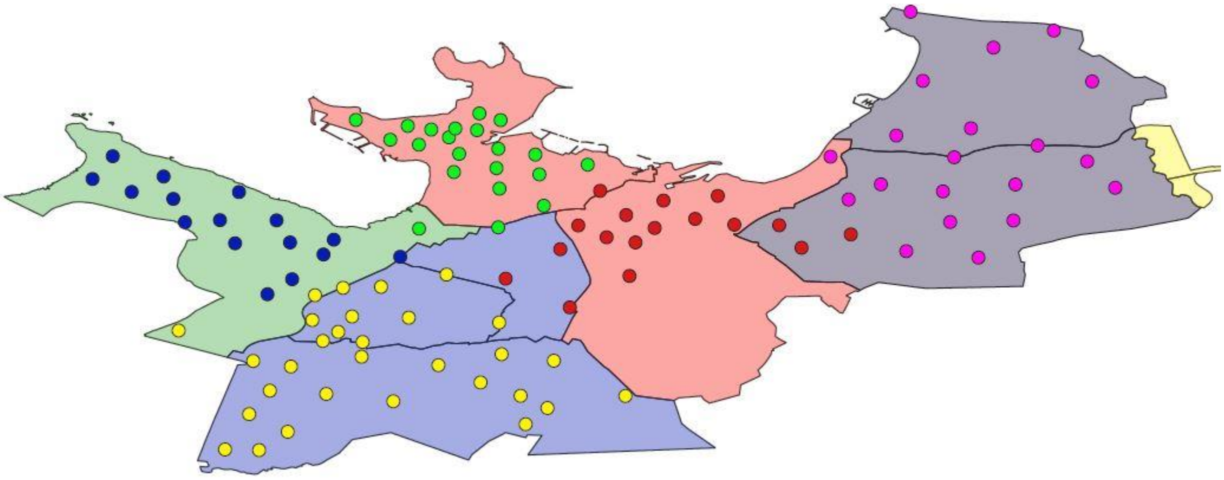

Figure 6.27. Merged K-means of demands and regions

7. DISCUSSIONS AND CONCLUSION

In this study, by collecting spatial data related to Tallinn, the base map has been made with QGIS software. Also, the temporal data of the energy of Tallinn has been collected. Having both spatial data with a base map and temporal data of Tallinn energy has let us execute different geo spatial analyses for Tallinn to make proper decisions for the energy sector. The classification of different regions of Tallinn has been done using Moran's I analysis and K-means analysis. According to Figure 7.27, considering the situation of different regions and the temporal data related to energy consumption of regions, the classification has been done. It illustrated that there are five different groups of regions that have almost similar attributes:

- The Nomme, the Mustamae and the Kristine
- The Haabersti
- The Pohja and the Kesklina
- The Prita and the Lasnamae
- The Iru and the Nehatu

These kinds of classification will help governments manage different regions' energy more easily due to the expansion of cities and the increase in energy demand. Additionally, nowadays, due to the rise in renewable energy production, a good approach is to make the distribution system localized, which helps the system to be more efficient, cost-efficient, and reliable.

To estimate the future energy direction of Tallinn, by using statistics related to the new construction permits. It has been shown that more than 7000 new permits have been granted, which is less than 10% of the current number of buildings in Tallinn. By estimating to finish around 50% of these new constructions and the population growth of Tallinn, which is less than 1%, this study has estimated approximately 2% rise in the amount of energy consumption by considering the population crowed in each region.

In Figure 7.1, the amount of the current energy consumption has been illustrated in green, and the estimated energy consumption in a period of one year has been presented in red. This map also shows that the electricity demand on average is 550 GWh per year for each region. Nomme has the highest demand with around 1700 GWh and the Nehatu with the lowest demand of around 5 GWh.

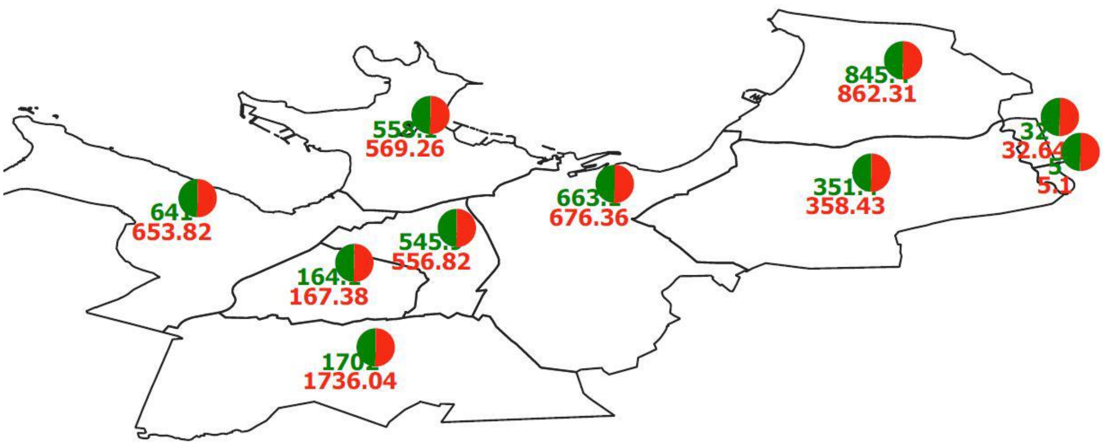

Figure 7.1. Electricity consumption increase in one year

In this study, the scenario is the implementation of PV production units on the rooftop of existing buildings with an area wider than 3000 square meters. To find these potential locations by using a selection algorithm in the QGIS on the base map of Tallinn, which has the buildings layer on it, locations have been allocated. Then with consideration of the power grid of Tallinn and use analysis applications in the QGIS, Figure 6.22 has been composed. It shows that for each PV production unit which distribution substations of the grid are more proper and closer to connect that unit to the grid. This analysis has been done to make the energy system efficient both on the energy and cost sides, decreasing the cost of longer connections and energy losses.
As such, the PV production estimation has been found according to the assessed area for PV units and considering the production amount by using the Eesti Energia tool. In this tool, the modules by which the estimation will be executed are monocrystalline PV modules of Recom company with a total efficiency of around 20%. In Figure 7.2 the diagram is the electricity demand estimation of the coming year. The red parts of the diagram represent the conventional electricity production portion, and the blue part has shown the PV production portion of the total electricity demand of different regions.

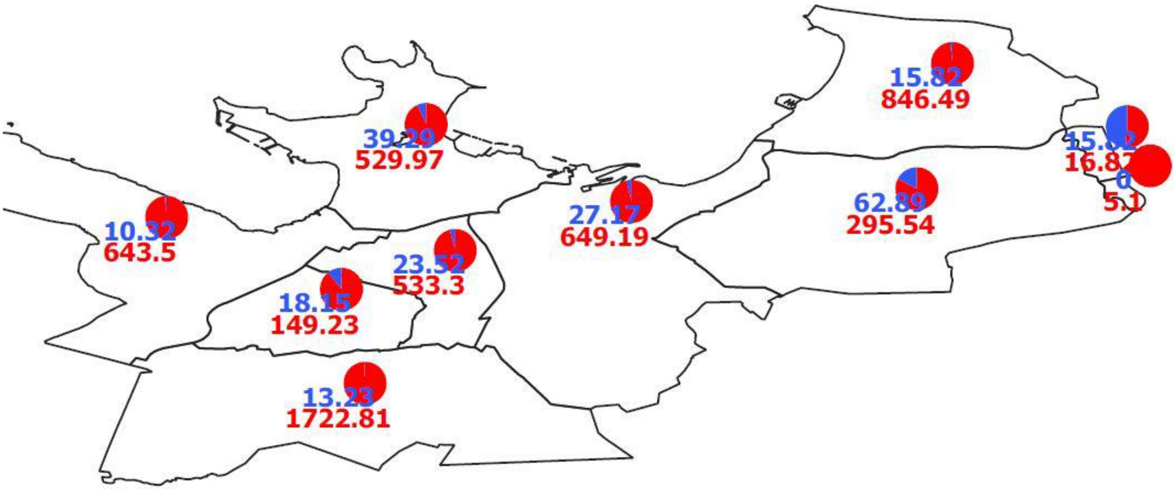

Figure 7.2. Share of the PV production (GWh) in one year according to this study scenario

It has shown that with this scenario, the average PV production capacity is around 13 GWh per year for each region and Lasnamae with more than 60 GWh potential is at the top, while the Nehatu region with 0 has the worst PV potential.
According to Figure 7.3 the region which has the highest potential to produce the electricity demand is Iru with almost 50% of the demand, however the point is that although the PV production is strong there, the demand is also low which results in this ratio. After that Lasnamae and Mustamae have PV production higher than 10% of the whole demand of the regions, then Pohja with around 7%, Kesklinna and Kristine with around 4%, Prita and

Haabersti with almost 2%, Nomme with around 1% and the worst PV potential location according to this scenario is the Nehatu region with 0% production of electricity demand.

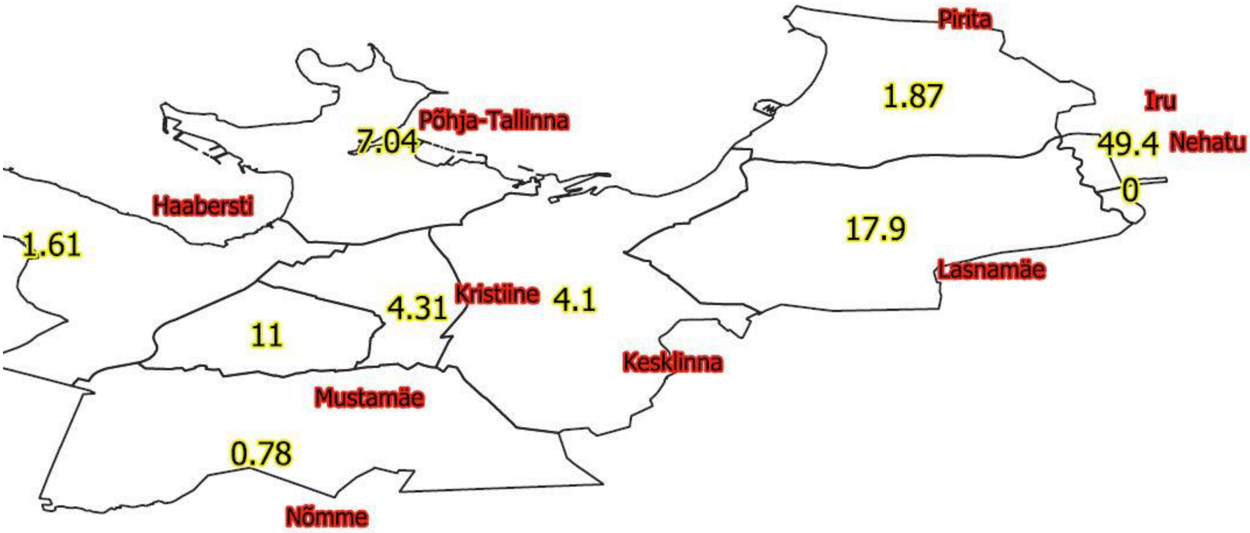

Figure 7.3. PV production percentage of total electricity consumption in each region of Tallinn

Overall, after analysis and making the estimations for future electricity demand of Tallinn and the PV production potential with the scenario of this study, it has been shown that by using these PV potential locations and implementation of PV production units, we can produce 4% of the electricity demand of the whole of Tallinn in a period of one year. This amount could be increased due to the technology development of the solar cells and increasing the efficiency of production units and consumers' efficiency.

7.1 Future works

In this study, the analysis was with the QGIS software and the classification by Moran's I matrix and clustering with the K-means model, but that's not all. For future works, there is the possibility to work with hierarchical clusters or other clustering methods. Also, there is a wide area of research according to the economical part of this study, one of the significant aspects of the energy production process.

On the other hand, other energy production possibilities can be considered as new research topics. For Estonia, this could be wind production due to having good wind potential.

ACKNOWLEDGEMENT

I would like acknowledge the support and advises at every step of the work from my supervisor Dr. Sambeet Mishra and also Professor Ivo Palu as my Co-supervisor. I would also like to thank my parents and friends who were with me in the journey.

SUMMARY

Space and time are both dynamic. While temporal changes are often used as the markers to predict the future, the spatial aspect often is highly related to each other. Energy consumption is gradually increasing, however the rates of change are not the same in various locations. Subsequently energy production potential in Tallinn is also increasing with roof-top PV installations. As such, with the increase in construction, the total energy potential and consumption also increases. This thesis aims to analyze the spatio-temporal energy potential in Tallinn. This starts with a spatial analysis of Tallinn covering the PV power production potential and the energy consumption with reference to the construction of buildings. Then it predicts how a change in the construction area would reflect in both production and consumption. It was found that about 7000 new construction permits have been granted and if 50% of them could finish construction within one year, there is 3-4% new construction in Tallinn. Also there is a population increase of less than 1% in one year for Tallinn. The estimation was done for 2% change in energy consumption potential over a period of one year.

This study aims to use the rooftop area of existing buildings with an area wider than 3000 square meters to implement PV panels to produce electricity locally. The module that is used in this study is the monocrystalline Recom brand with an efficiency of around 20%.

The implementation locations were found with the QGIS software. The result shows that there are around 500 unique locations to make PV production units. Specifically, about more than 4% of the energy consumption of Tallinn could be produced by PV potential according to the scenario of this study.

LIST OF REFERENCES

[1] R. C. Duncan., „World energy production, population growth, and the road to,“ 2001.

[2] „Renewable Energy Market | Growth, Trends, COVID-19 Impact, and Forecasts,“ [Võrgumaterjal]. Available: https://www.mordorintelligence.com/industry-reports/global-renewable-energy-market-industry.

[3] E. D. C. a. I. S. W. Bas J van Ruijven, „Amplification of future energy,“ Nature communications, 2019.

[4] Y. S. e. al, „Effects of climate change on heating and cooling degree days and,“ Climate Research, pp. 135-149, 2016.

[5] „Household and industrial electricity demand in Europe,“ Energy Policy, p. 592–600, 2018.

[6] Y. Scholz, „Renewable energy based electricity supply at low costs: development,“ pp. 11-15, 2012.

[7] A. J.-W. e. al, Renewable and sustainable energy reviews, 2011.

[8] J. A. F. a. J. P. M. António C Marques, Motivations driving renewableenergy in european countries, 2010.

[9] H. L. a. B. V. M. David Connolly, Smart Energy Europe: The technical and economic impact of one potential 100% renewable energy scenario for the European Union, 2016.

[10] D. V. e. al, „Evaluation of renewable energy potential using a GIS decision support system,“ Renewable energy, pp. 333-334, 1998.

[11] F. M. a. D. A. G. Livio de Santoli, A GIS-based model to assess electric energy consumptions and usable renewable energy potential in Lazio region at municipality scale, 2019.

[12] X. G. a. C. a. I. P. P. Javier Domı nguez Bravo, GIS approach to the definition of capacity and generation ceilings of renewable energy technologies, 2007.

[13] D. G. e. al, „Siting high solar potential areas using Q-GIS in West Bengal India,“ Sustainable Energy Technologies and Assessments 42, 2020.

[14] C. T. b. e. al, „A GIS-based decision support tool for renewable energy management and planning in semi-arid rural environments of northeast of Brazil,“ Renewable Energy 35.12, 2010.

[15] T. M. a. Y. G. H Ebru Colak, „Optimal site selection for solar photovoltaic (PV) power plants using GIS and AHP,“ Renewable energy 149, 2020.

[16] G. F. e. al., „Urban greenery management and energy planning,“ Renewable Energy 160, 2020.

[17] „Geographic Information System Mapping Technology,“ [Võrgumaterjal]. Available: https://www.esri.com/en-us/what-is-gis/overview.

[18] „the QGIS,“ [Võrgumaterjal]. Available: https://qgis.org/en/site/.

[19] „tallinn,“ [Võrgumaterjal]. Available: https://www.tallinn.ee/eng/geoportal/Spatial-data.

[20] „Administrative and Settlement Division | Geoportal | Estonian Land Board,“ [Võrgumaterjal]. Available: https://geoportaal.maaamet.ee/eng/Spatial-Data/Administrative-and-Settlement-Division-p312.html.

[21] „Stat,“ [Võrgumaterjal]. Available: https://www.stat.ee/.

[22] „IRENA – International Renewable Energy Agency," [Võrgumaterjal]. Available: https://www.irena.org/.

[23] „global solar atlas," [Võrgumaterjal]. Available: https://globalsolaratlas.info/map.

[24] L. Anselin, „Interactive techniques and exploratory spatial data analysis," 1996.

[25] „Location-allocation analysis—ArcMap | Documentation," [Võrgumaterjal]. Available: https://desktop.arcgis.com/en/arcmap/latest/extensions/network-analyst/location-allocation.htm.

[26] N. A. a. A. S. Gennady Andrienko, „Choropleth maps:classification revisited," 2001.

[27] M. P. Peterson, „An evaluation of unclassed crossed-line choropleth mapping," The American Cartographer, 1979.

[28] „ArcGIS pro," [Võrgumaterjal]. Available: https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/spatial-weights.htm.

[29] S. W. e. al, „Spatial disparity and hierarchical cluster analysis of final energy consumption in China," Energy 197, 2020.

[30] „ESRI," [Võrgumaterjal]. Available: https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/h-how-cluster-and-outlier-analysis-anselin-local-m.htm.

[31] N. V. a. J. J. V. Aristidis Likas, „The global k-means clustering algorithm," Pattern recognition 36.2, 2003.

[32] Y. L. a. H. Wu., „clustering method based on K-means algorithm," Physics Procedia 25, 2012.

[33] S. L. H. a. N. Moore, „A spatial regression analysis of the influence of topography on monthly rainfall in East Africa," International journal of climatology 31-10, 2011.

[34] „engineering statistics handbook," [Võrgumaterjal]. Available: https://www.itl.nist.gov/div898/handbook/eda/section3/eda33o.htm.

[36] „research data services," [Võrgumaterjal]. Available: https://data.library.virginia.edu/diagnostic-plots/.

[37] E. W. Miller, „World Patterns and Trends in Energy Consumption," Journal of Geography 58.6, 1959.

[38] „Eesti energia," [Võrgumaterjal]. Available: https://www.energia.ee/en/era/taastuvenergia/paikesepaneelid.

[39] „paradise energy solution," [Võrgumaterjal]. Available: https://www.paradisesolarenergy.com/blog/difference-between-off-grid-and-on-grid-solar-energy.

[40] „Recom," [Võrgumaterjal]. Available: https://recom-solar.com/products/mono/black-panther-mono-60-72-cells/.

[41] „Elering," [Võrgumaterjal]. Available: https://www.elering.ee/en/map-network.

LIST OF FIGURES

LIST OF TABLES