
FIRST INTERACTIVE SOLAR & WIND ENERGY ATLAS FOR NORTHEAST MEXICO

MASTER THESIS

Author: Fernando Gustavo Lappe
Mansilla

Hamburg University of Applied
Sciences (HAW – HAMBURG)
M. Eng. Renewable Energy Systems

ACKNOWLEDGEMENTS

First of all I would like to thank God for giving me the strength and patience to complete this challenge that I set myself a couple of years ago. I would also like to thank my parents for encouraging me at all times. This experience would never been possible without the support of the CONACyT, the Mexican government and the government of Nuevo Leon (Technology Institute I²T²) to which I'm also very thankful.

I would also like to thank my tutors Prof. Dr. Anna Rodenhausen and Prof. Dr.-Ing. Thomas Schiemann for the support given.

.....y esta va por México.

**Fernando Lappe
Hamburg, Germany
24.02.2014**

TABLE OF CONTENTS - THEORY

ABSTRACT	1
INTRODUCTION.....	3
Chapter 1: Country of Interest.....	5
1.1 Description of the Country – Mexico.....	5
1.1.1 Tamaulipas	6
1.1.2 Nuevo Leon	6
1.1.3 Coahuila.....	6
1.1.4 Chihuahua	6
1.1.5 Durango	7
1.1.6 Zacatecas	7
1.1.7 San Luis Potosi.....	7
Chapter 2: Renewable Energy in Mexico.....	10
2.1 Renewable Power Installed Capacity.....	10
2.2 Solar Energy in Mexico as of Today.....	11
2.3 Wind Energy in Mexico as of Today.....	11
2.4 Infrastructure for Renewables in Mexico	12
2.4.1 Mexican National Interconnected System – NIS.....	12
2.4.2 Mexican National Highway and Roads System	13
Chapter 3: How Do Solar and Wind Energy Work?	14
3.1 Solar Energy – The Sun	14
3.1.1 The Sun as a Fusion Reactor	14
3.1.2 Solar Resources in the Earth	19
3.1.3 Wind Resources in the Earth	23
3.1.3.1 Influence of Surroundings and Height	25
3.1.3.2 Power Content of Wind	27
Chapter 4: Spatial Interpolation.....	32
4.1 Data Used in the Project	32
4.2 Automatic Meteorological Stations - AMS	34
4.3 Spatial Interpolation.....	36
4.4 Techniques.....	38
4.4.1 Global interpolators.....	38
4.4.1.1 Classification Techniques.....	38
4.4.1.2 Trend Surface Analysis	38

4.4.1.3	Regression Techniques.....	40
4.4.2	Other Global interpolators.....	41
4.4.3	Local Interpolators	41
4.4.3.1	Thiessen Polygons.....	41
4.4.3.2	Weighted Moving Average Methods.....	44
4.4.3.3	Kriging	46
Chapter 5:	Factors Affecting the Performance of Spatial Interpolation Methods.....	53
5.1	Sampling Design and Sample Spatial Distribution	53
5.1.1	High Density	53
5.1.2	Low Density	53
5.1.2.1	Sample Spatial Distribution	54
5.1.3	Sample Size and Spatial Interpolation Methods	55
5.2	Data Quality.....	55
5.2.1	Distribution.....	55
5.2.2	Isotropism and Anisotropism.....	56
5.2.3	Variance and Range	56
5.2.4	Accuracy.....	57
5.2.5	Spatial Correlation and Other Factors.....	58
Chapter 6:	Procedure for Elaborating the Tool.....	59
6.1	Description of the Tool	59
6.2	Building the Data Base.....	60
6.2.1	First Step – Standardizing formats.....	60
6.2.2	Second Step – Calculating monthly averages and arranging information.....	60
6.2.3	Third Step – Yearly data bases for each State	61
6.2.4	Fourth Step – Concentration of variables of interest.....	62
6.3	Construction of Maps	63
6.3.1	First Step – Finding and downloading maps	64
6.3.2	Second Step – Modifying maps to fit program needs.....	64
6.3.3	Third Step – Extraction of State maps	65
6.3.4	Fourth Step – Adding municipalities to State maps	65
6.3.5	Fifth Step – Division of municipalities.....	66
6.3.6	Sixth Step – Attaching municipalities to State maps	67
6.3.7	Seventh Step – Merging maps.....	67
6.4	Interpolations.....	68

6.4.1	First Step – Information input for QGIS.....	68
6.4.2	Second Step - Interpolations.....	68
6.4.3	Third Step - Sampling.....	69
6.5	Distances Calculation.....	70
6.5.1	First Step – Finding and downloading maps.....	70
6.5.2	Second Step – Choosing method for calculation.....	71
6.5.3	Third Step – Loading coordinates to MatLab and distances calculation.....	72
6.6	Maps Projection.....	73
6.6.1	First Step – Merging the information.....	73
6.6.2	Second Step – Assigning merged information to corresponding cell.....	74
Chapter 7:	Results and Discussion.....	75
	Bibliography.....	A
	Annexes.....	D
	Instructions to Use Tool.....	E
	Features and Options.....	E
	Thematic map.....	E
	Cloropleth Map.....	E
	Proportional Symbol Map.....	E
	Select Country or Other Map Area Using the Map.....	F
	Map Legend.....	F
	Map Zoom.....	G
	Indicators Panel.....	G
	Select Category.....	G
	Indicator Bar.....	H
	Graph panel.....	H
	Bar Chart / Column Chart.....	H
	Time Series.....	H
	Vertical Bubble Chart.....	H
	Scatter Plot (Bubble Chart).....	H
	Select X-Axis / Y-Axis Indicator.....	I
	Search.....	I
	Adjust Graph Size.....	J
	Adjust Graph Scale.....	J
	Options Panel.....	J

Data-Table Panel	K
Selection Panel	K
Select Regions	L
Select Button	L
Deselect All Button	L
Refresh Button	L
Remove Button.....	L
Story Panel.....	L
Time Slider	M
Interface Options	M
View Panel	M
Shrink / Enlarge.....	M
Help / Minimize / Close.....	M
Adjust Graph Panel Size	N
Full Screen	N
Changing Category	N

TABLE OF CONTENTS - FIGURES

Figure 1 Maps of Mexico showing the States used in this work. On the left the white colored part shows the States analyzed in this work, on black the rest of the country. On the right on white are all the states from where the weather stations were considered, on black the rest of the country.....	4
Figure 2 Mexico's political division. As guidance to indicate the position of the seven analyzed States in this work.....	9
Figure 3 On the left the renewable power installed capacity in Mexico per technology is shown, on the right the growth rate in the last years of each technology is to be seen.....	10
Figure 4 Map of the main transmission lines in Mexico. Circled is the id number of each transmission line and over the lines its capacity.....	13
Figure 5 Representation of nuclear fusion in the Sun.....	16
Figure 6 Graphical representation of the parameters used to calculate irradiance at the Earth. The center of the small circle is the position of the Sun. The position of the Earth is depicted in the figure.....	19
Figure 7 Spectral irradiance of Sun's light according to its wavelength at AM 0 (extraterrestrial) and AM 1.5 (terrestrial).....	21
Figure 8 Representation of position of Sun and corresponding AM (at a lesser height or angle with respect to the ground, the AM is greater).....	22
Figure 9 Graph of global irradiance throughout the day at different dates. Here the differences of the irradiance in summer and winter are to be seen.....	23
Figure 10 Depiction of Coreolis effect which affects wind direction. This is one of the main aspects that influences the direction of the wind. On the northern hemisphere, this effect deviates the wind to the right. On the southern hemisphere it is the opposite.....	25
Figure 11 Sketch of changes in wind speed when flowing through wind turbine. As the diameter gets bigger the wind speed slows down up to one third of the original velocity.....	29
Figure 12 Graphical representation of the number of weather stations in each of the target States of the work.....	32
Figure 13 Graphical representation of the number of weather stations that were installed in the period from 2002-2012.....	33
Figure 14 Graphical representation of the accumulated number of weather stations in the period from 2000-2012.....	33
Figure 15 Location of weather stations used. A total of ninety weather stations were used for this work. Some were located on the target States and some on the surrounding States.....	34
Figure 16 Depiction of the two types of weather stations used with all their parts, structure and sensors.....	36

Figure 17 Delaunay triangulation with circumcircles around the red sample data. The resulting interpolated TIN surface created from vector points is shown on the right ..	42
Figure 18 Formation of Thiessen polygons out of initial sketch.....	43
Figure 19 Map used to show the pycnophylactic interpolation technique.....	43
Figure 20 Inverse Distance Weighted interpolation based on weighted sample point distance (left). Interpolated IDW surface from vector points (right).....	45
Figure 21 An example of a semivariogram illustrated by an exponential model	48
Figure 22 Examples of four commonly used variogram models: (a) spherical; (b) exponential; (c) linear; and (d) Gaussian.....	50
Figure 24 First format found in the information obtained. The different variables and their order can be seen. This format was chosen to work with, meaning that all the obtained spreadsheets and different variables with other formats were arranged this way.....	60
Figure 25 Second format found in the information obtained. The different variables and their order can be seen. The spreadsheets with this order were arranged to match the format of Figure 24. An approximate of 5,400 files had to be rearranged	60
Figure 26 Depiction of the monthly average summary used for each weather station and each variable. An approximate of 500 files were obtained at the end with this format. This format allowed the concentration and easy manipulation of the information	61
Figure 27 Screenshot of yearly averages done for each weather station (Aguascalientes).....	61
Figure 28 Screenshot of yearly averages done for each weather station (Zacatecas).....	62
Figure 29 Summary of the variables of interest for different years. Information of this file was used to carryout the interpolations, from which the results were later uploaded to Stat Planet Plus.....	62
Figure 30 Screenshot of CSV file uploaded to QGIS to carryout interpolations	63
Figure 31 First maps obtained by INEGI. The one on the left represents the states and the one on the right the municipalities of the country. From the map of the left, the states had to be extracted and then merged with the map of the right to obtain the municipalities in the State. Around 400 maps had to be created.....	64
Figure 32 Screenshot of the modifications that had to be made to the original files. A lot of names had to be rewritten to fit the requirements of the program	65
Figure 33 Some of the States that were extracted from whole map of Mexico. As shown in Figure 31	65
Figure 34 Some of the States with municipalities that were extracted from whole map of Mexico. As shown in Figure 31.....	66
Figure 35 Procedure showing the division of the municipalities	66

Figure 36 Representation of one divided municipality of the State of San Luis Potosi. The map of the municipality had to be attached to the map of the State This had to be done around 450 times, one time of each municipality..... 67

Figure 37 Depiction of the information that had to be uploaded to QGIS to carryout the interpolations. The map on the right contains all the weather stations and weather information used. This was done for each year (5)..... 68

Figure 38 Image of the first obtained results. Blue points are the weather stations used and the other points the interpolated results. From this a sampling was done to each of the centroids of the ten square kilometers squares to obtain a result..... 69

Figure 39 Screenshot of results changing the coefficients of the IDW method, 10 were used in total..... 70

Figure 40 Maps of transmission lines and roads obtained by INEGI. On the left the transmission lines (400 and 230 kV) and on the right the highways and roads are depicted..... 70

Figure 41 Triangle for representation Pythagoras Theorem used to calculate distances 71

Figure 42 Maps of transmission lines and roads converted to points..... 72

Figure 43 Screenshot of Excel VBA file used to arrange information. This information was uploaded to Stat Planet Plus..... 73

Figure 44 Screenshot of arranged information ready for programming into StatPlanet Plus 74

TABLE OF CONTENTS - TABLES

Table 1 Renewable power installed in Mexico by technology Here it is also shown which fraction of each technology comprises each generation scheme 11

Table 2 As reference for the reader in this table are shown different parameters relevant to this topic. This table serves to differentiate and distinguish some of the parameters that will be later used. Units of measurement are presented to indicate the reader the quantities measured by each parameter 15

Table 3 Different relevant facts from the Sun and the Earth..... 15

Table 4 Mass of particles involved in nuclear fusion reactions in the Sun. These are relevant for further calculations in this topic 17

Table 5 Reduction of Sun's intensity related to its height. This shows the loss of intensity as the height of the Sun increases. This due that at lesser height the light has to travel a greater distance 22

Table 6 Roughness length and its description according to ground class 27

Table 7 Calculation results when obtaining wind speed velocities by changing ground class and height 27

ABSTRACT

Due to the fact that Mexico has over 71,000 MW of potential renewable energy installed capacity and that as of today there are only around 12,000 MW installed (Mexican Energy Ministry, 2013) (barely 17% of the total), it was decided to investigate the potential causes of this situation with the goal of contributing to the promotion of development of renewable energy projects and the exploitation of said potential. During this analysis it was found that the National Weather Ministry of Mexico has only 188 weather stations that are managed by the National Board of that Ministry (there are over 406 more but are managed by other private/public bodies), which means that there are very wide regions across the whole Mexican territory where the weather conditions (wind speed and solar irradiance) are not known. This is considered to be one of the different factors that impede and obstruct the development of big scale renewable energy projects in that country, prohibiting also the establishment of a turning point towards a more green and environmental friendly way of producing electricity.

For this project, 7 States in Mexico (803,097 km²) have been analyzed. Each one of those 354 municipalities has been divided into 100 km² squares. For each centroid of those squares the yearly mean wind speed and solar irradiance conditions were calculated for the period 2008-2012. The distances to the next transmission line (400 and/ or 230 kV) and next highway (or accessible road) were also calculated.

Therefore, two methods of spatial interpolation were compared to determine their suitability for estimating the yearly mean wind speeds and solar irradiance for the Northeast States of Mexico (Tamaulipas, Nuevo Leon, Coahuila, Chihuahua, Durango, Zacatecas and San Luis Potosi), from data recorded at nearly 90 locations across the mentioned region (for calculation purposes, surrounding States were also considered Aguascalientes, Guanajuato, Hidalgo, Jalisco, Nayarit, Queretaro, Sinaloa, Sonora and Veracruz). The eventual purpose of producing such surfaces was to help estimate, identify and recognize the areas with the best conditions for developing big scale renewable energy projects, meaning the ones with the most favorable environmental conditions, the ones closer to highways and to transmission lines.

The interpolation techniques included two local, exact and deterministic methods: Inverse Distance Weighting (IDW) and Triangulated Irregular Network (TIN). Quantitative assessment of the continuous surfaces showed that there was a large difference between the accuracy of interpolation methods used, being IDW the most suitable for the analyzed region.

The results of that method were used for the further construction of the First Interactive Solar & Wind Energy Atlas for Northeast Mexico.

INTRODUCTION

Weather data are generally recorded at point locations, so estimating data values at other locations requires some form of spatial interpolation. A variety of deterministic and geostatistical interpolation methods are available to estimate variables at unsampled locations but, depending on the spatial attributes of the data, accuracies vary widely among methods. The final use of any interpolated variable surface must also be taken into account because different methods result in different surfaces (Willmot, 1985).

Spatial interpolation is more worthwhile if a sufficient density of weather stations is available across the study area. The density of the network required depends upon the variable to be estimated. Wind speed, for example, is more variable over shorter distances than temperature or relative humidity, and hence would be expected to require a denser network of monitoring sites to achieve accurate and precise interpolated surfaces.

While there have been comparisons of interpolation methods for temperature and precipitation, (Phillips et al., 1992; Collins and Bolstad, 1996; Goovaerts, 2000; Price et al., 2000; Jarvis and Stuart, 2001; Vicente-Serrano et al., 2003) few research efforts have been directed towards comparing the effectiveness of different spatial interpolators in predicting wind speed and solar radiation. In this study two spatial interpolation techniques – inverse distance weighting (IDW) and triangulated irregular network (TIN) were used.

The purpose was to determine which method created the best representation of reality for wind speed and solar irradiance across the northeast of Mexico (Tamaulipas, Nuevo Leon, Coahuila, Chihuahua, Zacatecas, Durango and San Luis Potosi). For the calculations, measurements done by the weather stations in the northeast States and also from surrounding states (Sonora, Sinaloa, Nayarit, Jalisco, Veracruz, Hidalgo, Queretaro, Guanajuato and Aguascalientes) were used. The objective of this project was to develop a renewable energy atlas for that region and to identify the best places for the development of big scale renewable energy projects, meaning the places with the best weather conditions and the ones closest to highways and transmission lines (W. Luo, M. C. Taylor and S. R. Parker, 2007).



Figure 1 Maps of Mexico showing the States used in this work. On the left the white colored part shows the States analyzed in this work, on black the rest of the country. On the right on white are all the states from where the weather stations were considered, on black the rest of the country

Chapter 1: Country of Interest

1.1 Description of the Country – Mexico

Mexico, officially the United Mexican States is a federal constitutional republic in North America. It is bordered on the north by (Mexican Embassy in the United States, 2012) the United States; on the south and west by the Pacific Ocean; on the southeast by Guatemala, Belize, and the Caribbean Sea; and on the east by the Gulf of Mexico (Webster M. , 2007). Covering almost two million square kilometers (CIA, 2013), Mexico is the fifth largest country in the Americas by total area and the 13th largest independent nation in the world (INEGI, 2010).

With an estimated population of over 113 million, it is the eleventh most populous and the most populous Spanish-speaking country in the world and the second most populous country in Latin America. Mexico is a federation comprising thirty-one states and a Federal District, the capital city.

In pre-Columbian Mexico many cultures matured into advanced civilizations such as the Olmec, the Toltec, the Teotihuacan, the Zapotec, the Maya and the Aztec before first contact with Europeans. In 1521, the Spanish Empire conquered and colonized the territory from its base in Mexico Tenochtitlan, which was administered as the Viceroyalty of New Spain.

Mexico has one of the world's largest economies; it is the tenth largest oil producer in the world, the largest silver producer in the world and is considered both a regional power and middle power. In addition, Mexico was the first Latin American member of the Organization for Economic Co-operation and Development OECD (since 1994), and considered an upper-middle income country by the World Bank. Mexico is considered a newly industrialized country and an emerging power (World Bank, 2011).

It has the fourteenth largest nominal GDP and the tenth largest GDP by purchasing power parity. The economy is strongly linked to those of its North American Free Trade Agreement (NAFTA) partners, especially the United States of America. Mexico ranks sixth in the world and first in the Americas by number of UNESCO World Heritage Sites with 32, and in 2010 was the tenth most visited country in the world with 22.5 million international arrivals per year (World Tourism Organization, 2011)

According to Goldman Sachs, by 2050 Mexico is expected to become the world's fifth largest economy. PricewaterhouseCoopers (PwC) estimated in January 2013 that by 2050 Mexico could be the world's seventh largest economy.

1.1.1 Tamaulipas

As of 1990 pop. 2,249,581, 79,601 sq. km, NE Mexico, on the Gulf of Mexico. Ciudad Victoria is the capital. The central and western parts of the state are in the mountains of the Sierra Madre Oriental. In the north and south are arable plains, particularly in the long panhandle beginning at Nuevo Laredo and following the Rio Grande opposite Texas to Matamoros. The coast is low, sandy, fringed with lagoons, and (except for Matamoros and Tampico) only lightly inhabited by fishermen and a few resorts. The extreme southwestern mining area borders on the vast semiarid basins of central Mexico. Except in the elevated interior, the climate is hot and humid. (The Columbia Encyclopedia, 2013).

1.1.2 Nuevo Leon

As of 1990 pop. 3,098,736, 65,102 sq. km, N Mexico. Monterrey is the capital. The southern and western parts of the state are traversed by the Sierra Madre Oriental, but some of the extreme western portions lie within the vast, semiarid basin lands of N Mexico, which are cultivable under irrigation. Much of the north is arid cactus country, but to the east, where the plains sweep down toward the lowlands of Tamaulipas and are crossed by several large rivers, the land is suitable for agriculture. (The Columbia Encyclopedia, 2013).

1.1.3 Coahuila

1990 pop. 1,972,340, 150,394 sq. km, N Mexico, on the northward bulge of the Rio Grande, S of Texas. Saltillo is the capital. In the eastern part of the state, where peaks of the Sierra Madre Oriental rise, are quantities of silver, copper, lead, iron, and zinc. Coahuila is an important coal-producing state and a leading national producer of iron and steel. A considerable portion of the Laguna District lies in this area. Torreón is the chief metropolis. (The Columbia Encyclopedia, 2013).

1.1.4 Chihuahua

1990 pop. 2,441,873, 245,612 sq. km, N Mexico, on the border of N.Mex. and Texas. The city of Chihuahua is the capital. Largest of the Mexican states, Chihuahua is

divided into two regions—the mountains of the Sierra Madre Occidental to the west, and the vast, cactus-and-greasewood desert basins, broken by scattered barren ranges, to the north and east. In extreme E Chihuahua and W Coahuila is a desolate basin, the Bolsón de Mapimí. At Nuevo Casas Grandes, in NW Chihuahua, is Paquimé, a vast and important archaeological site.

Chihuahua is a leading national mineral producer; the mines of the Sierra Madre yield silver, gold, zinc, lead, and other minerals and constitute the state's most valuable industry. (The Columbia Encyclopedia, 2013).

1.1.5 Durango

1990 pop. 1,349,378, 123,520 sq. km, N central Mexico. The city of Durango is the capital. The western half of the state is dominated by the Sierra Madre Occidental. These mountains contain deposits of many different minerals, and the mines extend north into the state of Chihuahua and south into Zacatecas. Durango is a leading national producer of ferrous metals. On the border of Coahuila is the fertile Laguna District, where vast desert basin lands are irrigated by the Nazas River. Gómez Palacio is the main settlement in this region. (The Columbia Encyclopedia, 2013).

1.1.6 Zacatecas

1990 pop. 1,276,329, 72,844 sq. km, N central Mexico. Zacatecas is the capital. Lying on the central plateau, Zacatecas is a state of semiarid plains and mountains. The Sierra Madre Occidental dominates the western half, and a transverse spur (often over 3,048 m high) of the same range, crossing the state from west to east, divides it. Rainfall is light and vegetation scanty. The absence of large rivers to support irrigation has limited agriculture. Zacatecas is one of Mexico's largest producers of mineral wealth (The Columbia Encyclopedia, 2013).

1.1.7 San Luis Potosi

1990 pop. 2,003,187, 63,240 sq. km central Mexico. San Luis Potosí is the capital. Most of the state lies on the eastern tablelands of Mexico's central plateau. Except in the humid tropical Pánuco River valley in the extreme east, near the Gulf of Mexico, the climate is mild and dry. Generally level, with an average elevation of 1,829 m, the plateau is broken by spurs of the Sierra Madre Oriental; it is largely desert in the north.

Rainfall is generally light, and rivers are few; thus, despite fertile soil, agriculture is practiced mainly for subsistence. (The Columbia Encyclopedia, 2013).

STATES AND CAPITALS

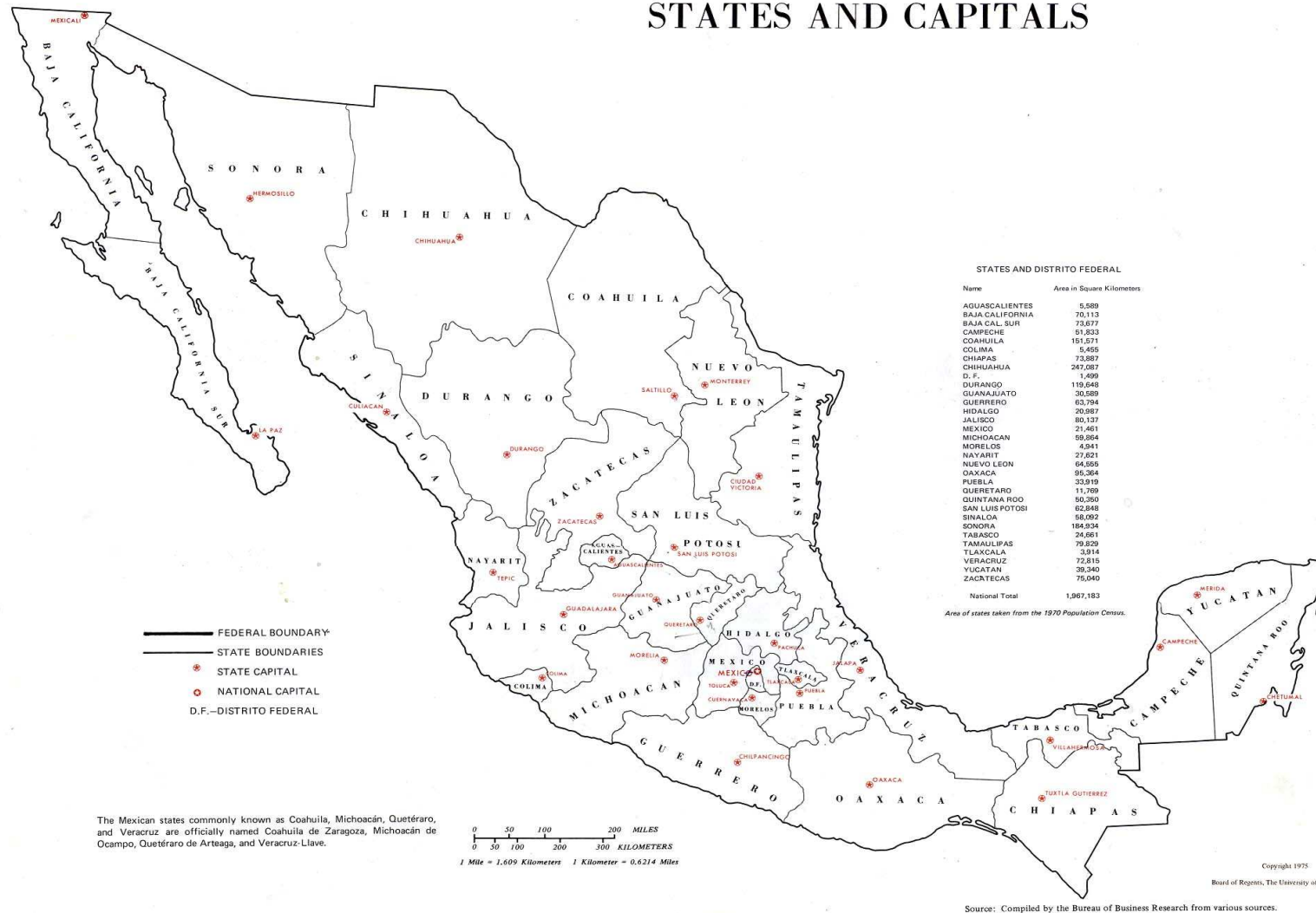


Figure 2 Mexico's political division. As guidance to indicate the position of the seven analyzed States in this work

Chapter 2: Renewable Energy in Mexico

By the beginning of 2012, the power generation installed capacity from renewable energy sources was 14,357 MW. Of this installed capacity, 87.3% belongs to the public sector and 12.7% is owned by privates.

2.1 Renewable Power Installed Capacity

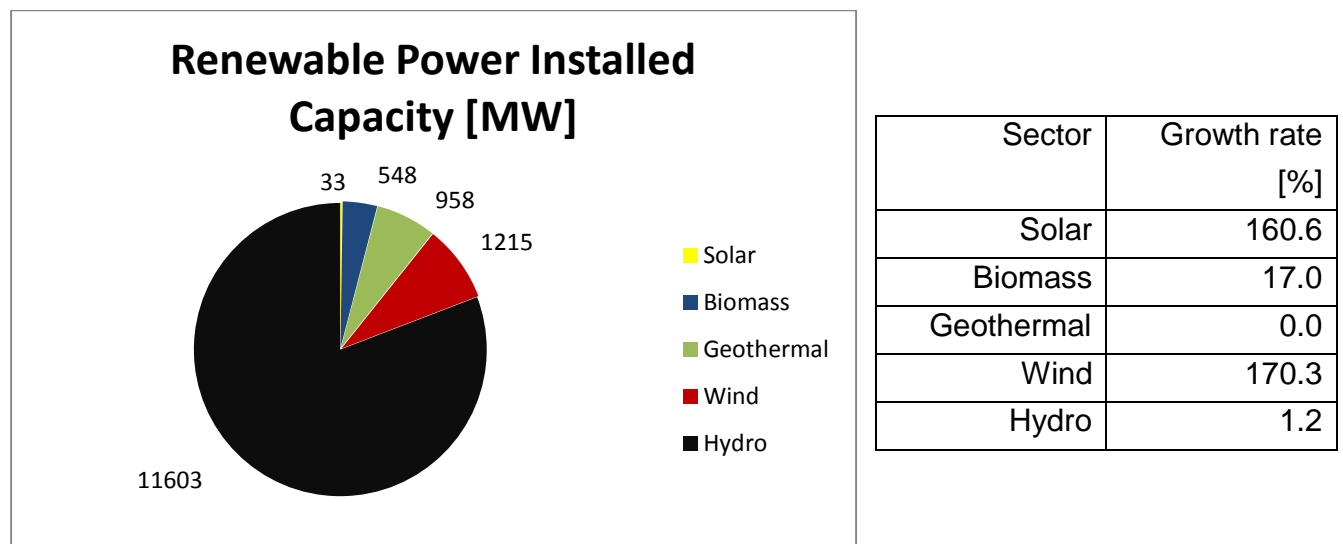


Figure 3 On the left the renewable power installed capacity in Mexico per technology is shown, on the right the growth rate in the last years of each technology is to be seen

The growth rate is calculated based on the authorized capacity under construction in February of 2012. The wind and the solar energy sectors showed considerable increase, of 170% and 160% respectively.

By the year of 2025, there will be an estimated of 30,000 MW of total power installed capacity from renewable energies. This means an increase of 18,716 MW in current installed capacity, led by wind and hydropower sources, with 60 and 24% share, respectively. This installed capacity will make up 15.3% of the total National Electric Network. This forecast is distributed among the different renewable sources as follows:

Energy	Public service [MW]	Self Sufficiency [MW]	Distributed Generation [MW]	Total [MW]	%
Wind	2,023	8,264	991	11,278	60
Hydro	3,531	575	435	4,541	24
PV	5	601	1,567	2,173	12
Concentrated Solar	14	601	55	69	0.4
Biomass	-	324	194	518	3
Geothermal	102	-	35	137	0.7
Total	5,675	9,764	3,277	18,716	100

Table 1 Renewable power installed in Mexico by technology Here it is also shown which fraction of each technology comprises each generation scheme

As of the beginning of the year 2012, there were around 204 renewable energy power stations in operation or under construction, with a total installed capacity of 5,505 MW.

2.2 Solar Energy in Mexico as of Today

Until the end of 2011, the total installed capacity of photovoltaic systems in Mexico was of 32 MW, these mainly used for rural electrification, water supply in the residential sector, water pumping and in the industrial and commercial sectors they are used for exterior lightning, emergency system powering, etc.

After the publication of the regulatory instruments that facilitate the interconnection of photovoltaic systems to the grid, the Federal Electricity Commission registered an additional installed capacity on small and medium scale by 3.48 MW in the period 2010 – 2011. The grow of these systems in the residential and service sector was of 763% (1.34 MW) in 2010 and 128% (1.95 MW) in 2011. The electricity generation by photovoltaic systems off the grid was of 0.2 MW (5.71%). The mean plant load factor rounded the 0.207.

2.3 Wind Energy in Mexico as of Today

The Mexican Federal Electricity Commission maintained in operation until the end of 2012 the wind power plants of La Venta (in the State of Oaxaca, 84.7 MW), Guerrero Negro (in the State of Baja California Sur, 0.6 MW) and one 1.5 MW wind mill installed for the Conference of

Parts celebrated in Cancun, giving a total installed capacity under the scheme of public service of over less than 90 MW. At the moment, 2,280 MW of capacity are being installed in the States of Baja California, Nuevo Leon, Oaxaca, San Luis Potosi, Tamaulipas and Veracruz.

The number of projects developed in the State of Oaxaca allowed the planning and construction of new transmission infrastructure to take advantage of the wind resource at the Tehuantepec Isthmus. (Mexican Energy Ministry, 2013).

2.4 Infrastructure for Renewables in Mexico

2.4.1 Mexican National Interconnected System – NIS

The infrastructure of the NIS makes possible the transformation, transmission, distribution and commercialization of electrical energy across the whole country. This infrastructure is operated by control areas of the National Electricity Commission and they maintain the reliability of the system.

At the same time, these areas supervise that the demand and supply of energy are balanced all the time. Until the end of 2011, the transmission and distribution grid had a total length of 845,201 km. The transmission grid is composed by 230 – 400 kV lines (5.9%), 69 – 161 kV lines (5.8%), 2.4 – 34.5 kV (47.7%), low tension grid (30.4%) and the lines that were operated by the today extinct company “Luz y Fuerza” (10.2%).

Between 2001 and 2011 the national transmission and distribution grid was expanded by 201,271 km. The branches that registered the greatest expansion in the grid were the 13.8 kV lines (raise of 57,236 km), the lines of the extinct “Luz y Fuerza” (raise of 57,166 km) and low tension lines (raise of 41,783 km) (Mexican Energy Ministry, 2013).



Figure 4 Map of the main transmission lines in Mexico. Circled is the id number of each transmission line and over the lines its capacity.

2.4.2 Mexican National Highway and Roads System

In Mexico the road network is the most used transport infrastructure due to its flexibility and great extension. The 370 thousand of kilometers of interurban roads are composed of by highways, freeways, roads and rural roads. They allow the connectivity in practically all the cities in the country independently of their number of inhabitants or economical relevance. The national road network is roughly composed by 50,000 kilometers of federal jurisdiction highways (from which almost 9 thousand are toll highways), around 80,000 kilometers of state jurisdiction highways, 170,000 kilometers of rural roads and 70,000 kilometers of normal roads.

In Mexico there have been identified 14 main highway corridors, which together they add up almost 20,000 kilometers and run from north to south and east to west.

Those corridors connect all the capital cities, the main metropolitan concentrations, middle - sized cities, ports and border accesses to the United States, Guatemala and Belize (Trade & Logistics Innovation Center, 2011).

Chapter 3: How Do Solar and Wind Energy Work?

3.1 Solar Energy – The Sun

The Sun is by far the most significant source of renewable energy. This abundant source of energy can be utilized directly by solar thermal or photovoltaic systems. In principle, wind power and hydroelectricity are transformed solar energy and are sometimes called indirect solar energy. This chapter is dedicated to solar irradiation and solar irradiance they are very important for many renewable energy systems and are one of the main pillars of this work.

Daylight quantities refer only to the visible part of light, whereas solar energy also includes invisible ultraviolet and infrared light. It is important to mention that solar irradiation refers to the amount of energy from the Sun received per square meter in the surface of the Earth¹. In practice, it has units of $kWh/(m^2 \cdot day)$ or $kWh/(m^2 \cdot year)$. In contrast, solar irradiance refers only to the power received per square meter on the Earth surface². It refers to W/m^2 . To convert from irradiation to irradiance the values have to be multiplied times 0.041666666 (365 days / 8760 (hours/year)).

3.1.1 The Sun as a Fusion Reactor

The Sun is the central point of our solar system; it has probably been in existence for 5 billion years and is expected to survive for a further 5 billion years.

Radiant physical quantities			Radiant physical quantities		
Name	Symbol	Unit	Name	Symbol	Unit
Radiant energy	Q_e	$Ws = J$	Quantity of light	Q_v	$Im s$
Radiant flux / radiant power	ϕ_e	W	Luminous flux	ϕ_v	Im
Specific emission	M_e	W/m^2	Luminous exitance	M_v	Im/m^2

¹ Energy is measured in Wh.

² Power is measured in W.

Radiant intensity	I_e	W/sr	Luminous intensity	I_v	$cd = lm/sr$
Radiance	L_e	$W/(m^2 sr)$	Luminance	L_v	cd/m^2
Irradiance	$E_e G$	W/m^2	Illuminance	$E_e v$	$lx = lm/m^2$
Irradiation	H_e	$Wh/(m^2 s)$	Light exposure	H_v	$lx s$

Table 2 As reference for the reader in this table are shown different parameters relevant to this topic. This table serves to differentiate and distinguish some of the parameters that will be later used. Units of measurement are presented to indicate the reader the quantities measured by each parameter

Note: W = watt; m = meter; s = second; sr = steradian; lm = lumen; lx = lux; cd =candela.
Source: DIN, 1982; ISO, 1993.

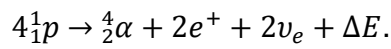
	Sun	Earth	Ratio
Diameter (km)	1,392,520	12,756	1:109
Circumference (km)	4,374,097	40,075	1:109
Surface (km ²)	6.0874×10^{12}	5.101×10^8	1:11,934
Volume (km ³)	1.4123×10^{18}	1.0833×10^{12}	1:303,670
Mass (kg)	1.9891×10^{30}	5.9742×10^{24}	1:332,946
Average density (g/cm ³)	1.409	5.516	1:0.26
Gravity surface (m/s ²)	274.0	9.81	1:28
Surface temperature (K)	5777	288	1:367
Center temperature (K)	15,000,000	6,700	1:2,200

Table 3 Different relevant facts from the Sun and the Earth

The Sun consists of about 80% hydrogen, 20% helium and only 0.1% other elements. Table 3 contains data about the Sun in comparison to the Earth.

Nuclear fusion processes create the radiant power of the Sun. During these processes, four hydrogen nuclei (protons 1p) fuse to form one helium nucleus (alpha particle $^4\alpha$). The alpha particle consists of two neutrons 1n and two positively charged protons 1p . Furthermore, this

reaction produces two positrons e^+ and two neutrinos ν_e and generates energy. The equation of the gross reaction next illustrated:



Equation 1

Comparing the masses of the atomic particles before and after the reaction shows that the total mass is reduced by the reaction. Table 4 shows the necessary particle masses for the calculation of the mass difference. The mass of the neutrinos ν_e can be ignored in this calculation and the mass of a positron e^+ is the same as that of an electron e^- .

The mass difference Δm will be calculated by:

$$\Delta m = 4m(\text{}^1\text{p}) - m(\text{}^4\alpha) - 2m(e^+).$$

Equation 2

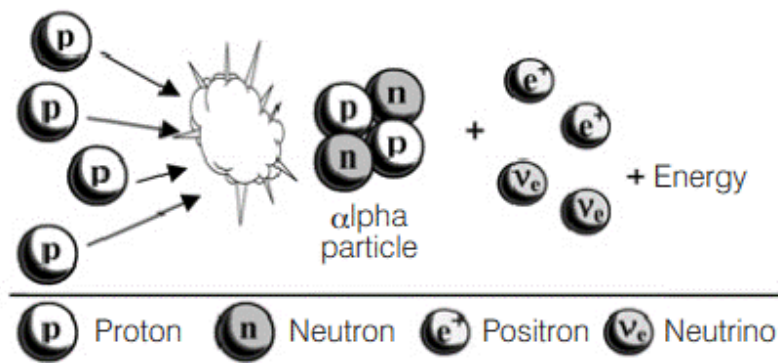


Figure 5 Representation of nuclear fusion in the Sun

Particle or nuclide	Mass	Particle or nuclide	Mass
Electron (e^-)	0.00054858 u	Hydrogen (^1H)	1.007825032 u
Proton (^1p)	1.00727647 u	Helium (^4He)	4.002603250 u

Neutron (1n)	1.00866492 u	Alpha particle (${}^4\alpha$)	4.001506083 u
---------------------	--------------	---------------------------------	---------------

Table 4 Mass of particles involved in nuclear fusion reactions in the Sun. These are relevant for further calculations in this topic

For this reaction the result is:

$$\Delta m = 4 \cdot 1.00727647 \text{ u} - 4.0015060883 \text{ u} - 2 \cdot 0.00054858 \text{ u} = 0.026502663 \text{ u}.$$

Equation 3

Thus, the total mass of all particles after the fusion is less than that before. The mass difference is converted into energy ΔE , with the relationship:

$$\Delta E = \Delta m \cdot c^2.$$

Equation 4

With the speed of light $c = 2.99792478 \cdot 10^8 \text{ m/s}$, this equation determines the energy released by the fusion as $\Delta E = 3.955 \cdot 10^{-12} \text{ J} = 24.687 \text{ MeV}$. The binding energy E_b of a nucleus explains the different masses after the fusion as well as the energy difference. An atomic nucleus consists of N neutrons 1n and Z protons 1p . To maintain equilibrium, this binding energy has to be released during the assembly of a nucleus with protons and neutrons.

The mass difference of the alpha particle and the two neutrons together with the two protons determines the binding energy of a helium nucleus. So far, only the atomic nuclei have been considered; the electrons in the atomic shell have not been taken into account. There is one electron in the atomic shell of a hydrogen atom ${}^1\text{H}$, while there are two electrons in the helium atom ${}^4\text{He}$.

During the nuclear fusion process, two of the four electrons of the hydrogen atoms become the atomic shell electrons of the helium atom. The two other electrons and the positrons convert directly into energy. This radiative energy is four times the equivalent mass of an electron of 2.044 MeV.

The total energy released during the reaction is thus 26.731 MeV. This very small amount of the energy does not appear to be significant at first glance; however, the enormous number of fusing nuclei results in the release of vast quantities of energy.

The Sun loses 4.3 million metric tons of mass per second (*i.e.* $\Delta m = 4.3 \cdot 10^9 \text{ kg/s}$) this results in the solar radiant power $\phi_{e,S}$:

$$\phi_{e,S} = \Delta m \cdot c^2 = 3.845 \cdot 10^{26} \text{ W}.$$

Equation 5

This value divided by the Sun's surface area, A_S , provides the specific emission of the Sun:

$$M_{e,S} = \frac{\phi_{e,S}}{A_S} = 63.11 \frac{\text{MW}}{\text{m}^2}.$$

Equation 6

Every square meter of the Sun's surface emits a radiant power of 63.11 MW. One fifth of a square kilometer of the Sun's surface emits radiant energy of 400 EJ per year. This amount of energy is equal to the total primary energy demand on Earth at present. The Sun's irradiance can be approximated to that of a black body. The Stefan–Boltzmann law states that:

$$M_e(T) = \sigma \cdot T^4.$$

Equation 7

It can be used to estimate the surface temperature of the Sun, T_{Sun} . With the Stefan–Boltzmann constant $\sigma = 5.67051 \cdot 10^8 \text{ W/(m}^2\text{K}^4)$, it becomes:

$$T_{\text{sun}} = \sqrt[4]{\frac{M_{e,S}}{\sigma}} = 5,777 \text{ K}.$$

Equation 8

The surface A_{SE} , of a sphere with the Sun as its center and with a radius equal to the average distance from the Earth to the center of the Sun ($r_{se} = 1.5 \cdot 10^8 \text{ km}$) receives the same total radiant power as the surface of the Sun A_S (Figure 6). However, the specific emission $M_{e,S}$, or the energy density measured over one square meter, is much higher at the Sun's surface than at the sphere surrounding the Sun.

With $M_{e,S} \cdot A_s = E_e \cdot A_{SE}$ and substituting $A_{SE} = 4 \cdot \pi \cdot r_{SE}^2$, the irradiance at the Earth, E_e , finally becomes:

$$E_e = E_{e,S} \cdot \frac{A_s}{A_{SE}} = M_{e,S} \cdot \frac{r_s^2}{r_{SE}^2}.$$

Equation 9

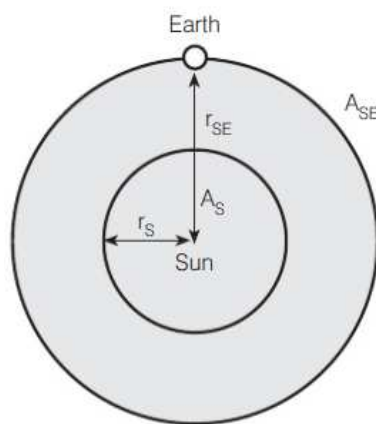


Figure 6 Graphical representation of the parameters used to calculate irradiance at the Earth. The center of the small circle is the position of the Sun. The position of the Earth is depicted in the figure

This determines the extraterrestrial irradiance experienced at Earth's orbital distance from the Sun. However, the distance between the Sun and Earth is not constant throughout the year. It varies between $1.47 \cdot 10^8$ and $1.52 \cdot 10^8$ km. This causes a variation in the irradiance, E_e , of between $1,325 \text{ W/m}^2$ and $1,420 \text{ W/m}^2$. The average value, called the solar constant:

$$E_0 = 1.367 \pm 2 \text{ W/m}^2.$$

Equation 10

This value can be measured outside the Earth's atmosphere on a surface perpendicular to the solar radiation.

3.1.2 Solar Resources in the Earth

Values measured on the surface of Earth are usually lower than the solar constant.

Various influences of the atmosphere reduce the irradiance. They are:

- Reduction due to reflection by the atmosphere,
- Reduction due to absorption in the atmosphere (mainly O₃, H₂O, O₂ and CO₂) and,
- Reduction due to Rayleigh scattering reduction due to Mie scattering.

The absorption of light by different gases in the atmosphere, such as water vapor, Ozone and Carbon Dioxide, is highly selective and influences only some parts of the spectrum. Figure 7 shows the spectrum outside the atmosphere (AM 0) and at the surface of the Earth (AM 1.5). The spectrum describes the composition of the light and the contribution of the different wavelengths to the total irradiance.

Seven per cent of the extraterrestrial spectrum (AM 0) falls in the ultraviolet range, 47 per cent in the visible range and 46 per cent in the infrared range. The terrestrial spectrum AM 1.5 shows significant reductions at certain wavelengths caused by absorption by different atmospheric gases. Molecular air particles with diameters smaller than the wavelength of light cause Rayleigh scattering.

The influence of Rayleigh scattering rises with decreasing light wave length. Dust particles and other air pollution cause Mie scattering. The diameter of these particles is larger than the wavelength of the light. Mie scattering depends significantly on location; in high mountain regions it is relatively low, whereas in industrial regions it is usually high. Table 5 shows the contributions of Mie and Rayleigh scattering and absorption for different Sun heights γ_s . Climatic influences such as clouds, snow, and rain or fog can cause additional reductions.

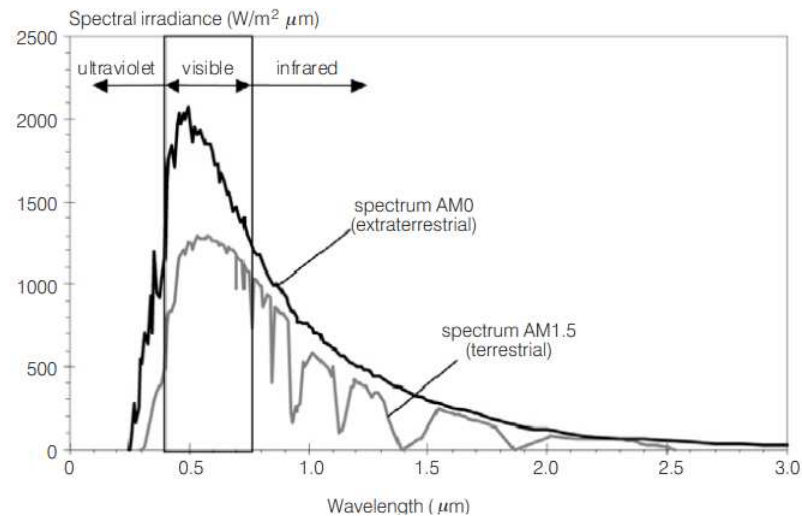


Figure 7 Spectral irradiance of Sun's light according to its wavelength at AM 0 (extraterrestrial) and AM 1.5 (terrestrial)

The relationship between the Sun height γ_s , and the air mass (AM) is:

$$AM = \frac{1}{\sin \gamma_s}$$

Equation 11

The AM value is a unit less measure of the length of the path of light through the atmosphere; it is expressed in multiples of the thickness of atmosphere. If the Sun is at its zenith, AM is equal to 1, i.e. the light is passing vertically through the atmosphere. The AM value outside the atmosphere is zero.

Figure 8 shows the highest position of the Sun at solar noon and the corresponding AM values for various days of a year for Berlin and Cairo. The elevation of the Sun also influences the irradiation received at the surface of the Earth, which is thus dependent on the time of the year.

Sun height (γ_s)	Air Mass (AM)	Absorption (%)	Rayleigh scattering (%)	Mie scattering (%)	Total reduction (%)
90°	1.00	8.7	9.4	0 - 25.6	17.3 – 38.5
60°	1.15	9.2	10.5	0.7 – 29.5	19.4 – 42.8
30°	2.00	11.2	16.3	4.1 – 44.9	28.8 – 59.1

10°	5.76	16.2	31.9	15.4 – 74.3	51.8 – 85.4
5°	11.5	19.5	42.5	26.4 – 86.5	65.1 – 93.8

Table 5 Reduction of Sun's intensity related to its height. This shows the loss of intensity as the height of the Sun increases. This due that at lesser height the light has to travel a greater distance

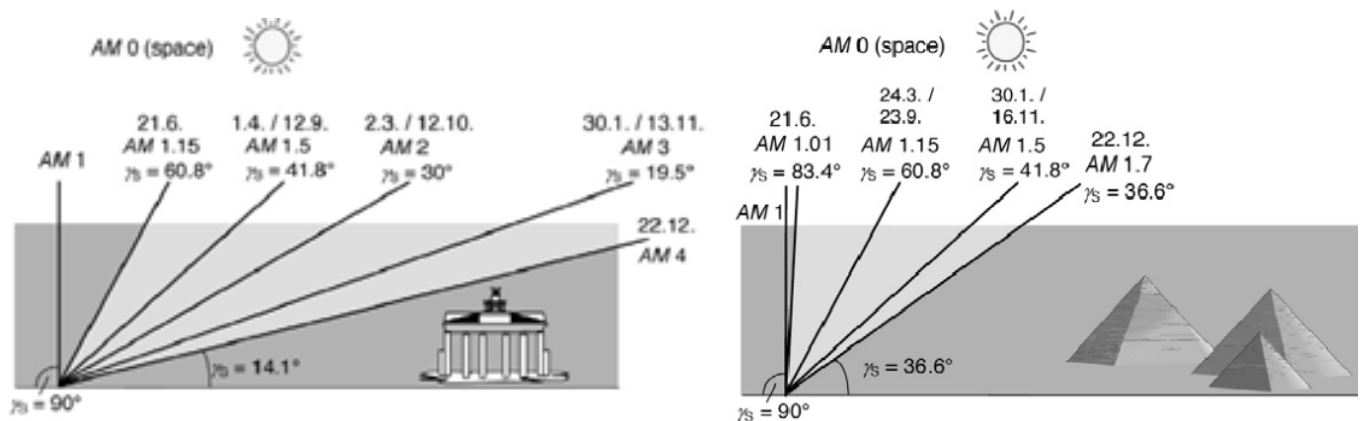


Figure 8 Representation of position of Sun and corresponding AM (at a lesser height or angle with respect to the ground, the AM is greater)

Clouds and weather are important as well. The daily irradiance in central Europe can reach values above 312.5 W/m^2 in summer, whereas single days in winter can have less than 4.16 W/m^2 . Figure 9 shows the variation in the irradiance for a cloudless day in summer (2 July) and in winter (28 December) as well as a very cloudy day in winter (22 December) for Karlsruhe in southern Germany. As an example in Ciudad Juarez, Mexico the irradiance in winter rounds $3.07 \text{ kWh/m}^2/\text{day}$ in December and then increases to $7.36 \text{ kWh/m}^2/\text{day}$ in summer (June) (Solar Electricity Handbook, 2013).

The annual irradiance varies significantly throughout the world. For instance, in America there are large differences between north and south. In the north, differences between summer and winter are much higher than in the south. In Winnipeg, Manitoba (Canada) the ratio of global irradiation (total irradiance on a horizontal surface on Earth) in June to global irradiation in December is 5:1, whereas in Ciudad Juarez (Mexico) this ratio is only 2:1.

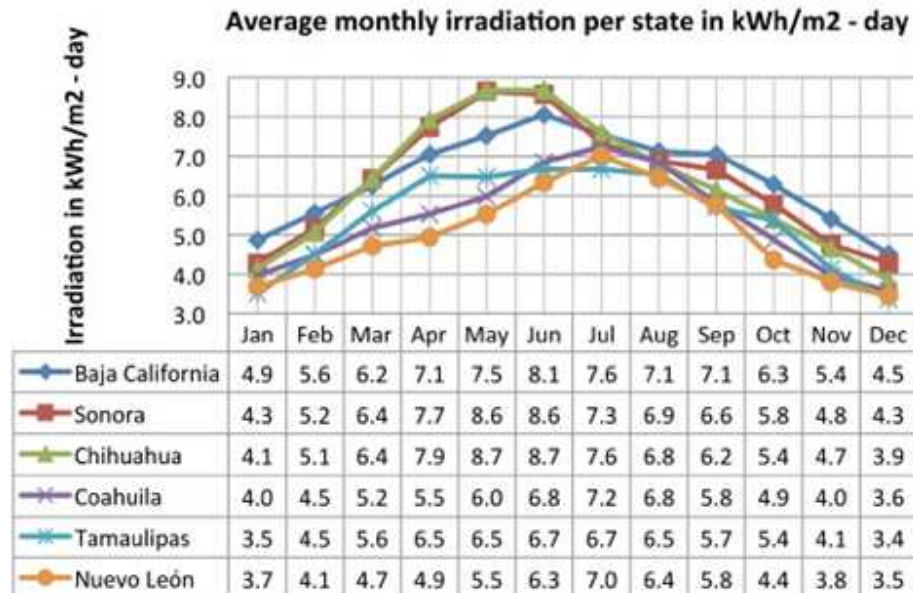


Figure 9 Graph of global irradiance throughout the day at different dates. Here the differences of the irradiance in summer and winter are to be seen

However, the latitude can only give a rough indication of the annual irradiation because local effects have a major impact on the energy reaching Earth’s surface.

For instance, the annual irradiation in Stockholm (Sweden) and Berlin (Germany) are nearly the same, although Stockholm’s latitude is 7° higher than Berlin. On the other hand, the annual irradiation in London, which is south of Berlin, is significantly lower. The annual irradiance in the Sahara is about 272.26 W/m². The total annual irradiation received by the surface of the Sahara (around 8.7 million km²) is nearly 200 times higher than the global annual primary energy demand; in fact the global primary energy demand could be provided by collecting the solar energy received by 48,500 km² of the Sahara, an area slightly larger than Switzerland, or one-ninth that of California. These numbers clearly show that it is possible to provide the whole global energy demand solely by solar energy.

3.1.3 Wind Resources in the Earth

The Sun is also responsible for the formation of wind. At all times, huge amounts of energy in form of heat and light from the Sun reaches the Earth. In order for the Earth not to melt because of the continuous accumulation of energy, a considerable amount of it is reflected back to the outer space. More energy strikes the Earth at the Equator because its distance from the Sun is

smaller than the one to the poles, where the situation is the inverse. At the poles, more energy is reflected back to the outer space than that reaching the Earth. Because of this, a huge energy transport from the Equator to the poles takes place.

This transport is carried out naturally from huge exchanges of air masses. Enormous worldwide air circulations pump the heat from the Equator to the poles; these are called Hadley circulation cells and are depicted on Figure 10. The rotation of the Earth deviates these air mass circulations, so in the tropical sea regions above the Equator a relatively constant wind coming from the north east blows.

A side these so called equilibrium circulations, there are other smaller currents of air that appear because of the influences of high and low atmospheric pressure regions, these are influenced and deviated by the Coreolis effect. Because of this effect, the air masses at the north semicircle of the Earth will be deviated to the right and masses at the south semicircle would be deviated to the left. These currents would then move in a whirling motion towards low atmospheric pressure regions.

The sunlight heats the land more quickly than the water during the day. The results are pressure differentials and compensating winds in the direction of the land. These winds can reach up to 50 km inland. During the night the land cools much faster than the sea; this causes compensating winds in the opposite direction.

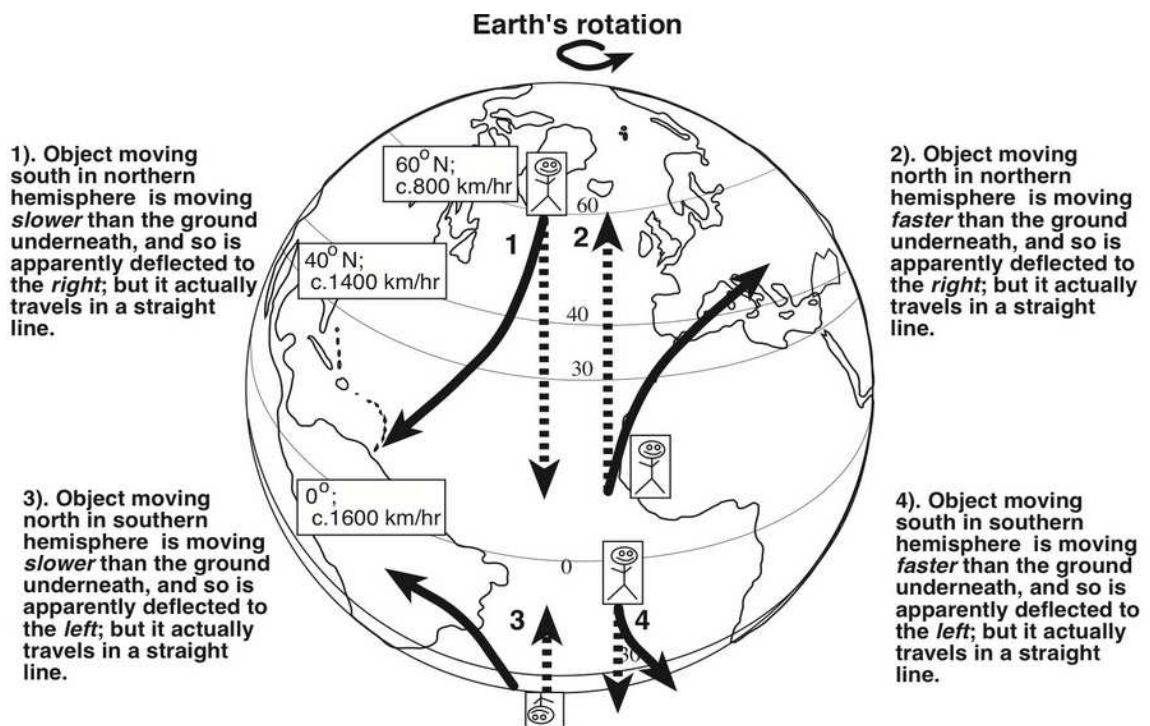


Figure 10 Depiction of Coriolis effect which affects wind direction. This is one of the main aspects that influences the direction of the wind. On the northern hemisphere, this effect deviates the wind to the right. On the southern hemisphere it is the opposite

3.1.3.1 Influence of Surroundings and Height

The wind speed is usually recorded at a height of 10 m. Changes in elevation can change the wind speed in a distance of only a few hundred meters. Hills or mountains influence the wind speed significantly. On top of a mountain or on the luff side, which is perpendicular to the wind, the wind speed can rise to become double the uninfluenced value. In contrast, the wind speed is much lower on the lee side.

Obstacles, plants or hills near a wind generator site can slow the wind significantly. Single obstacles are no problem if the total rotor area is over three times higher than the obstacle or if there is sufficient distance between the wind generator and the obstacle. However, this distance can be up to 35 times the obstacle height. Without proper clearance, wind turbulence can reduce the usable wind energy.

The wind speed increases with the height from ground because the wind is slowed down by the roughness of the ground. Wind generators usually have hub heights of more than 10 meters.

For the estimation of the wind potential, additional wind speed measurements at other heights are necessary. However, if the type of ground cover is known, the wind speed at other heights can be calculated. The wind speed $v(h_2)$ at height h_2 can be calculated directly with the roughness length, z_0 of the ground cover and the wind speed $v(h_1)$ at height h_1 :

$$v(h_2) = v(h_1) \cdot \frac{\ln\left(\frac{h_2-d}{z_0}\right)}{\ln\left(\frac{h_1-d}{z_0}\right)}$$

Equation 12

Obstacles can cause a displacement of the boundary layer from the ground. This displacement can be considered by the parameter d . For widely scattered obstacles, parameter d is zero. In other cases d can be estimated as 70 per cent of the obstacle height. The roughness length z_0 describes the height at which the wind is slowed to zero. In other words, surfaces with a large roughness length have a large effect on the wind. Table 6 shows the classification of various ground classes depending on the roughness length.

The following example shows the influence of the ground cover. The same wind speed $v(h_1) = 10$ m/s at a height $h_1 = 50$ m above different ground classes is assumed. Equation 12 is applied to calculate the wind speed $v(h_2)$ at a height of $h_2 = 10$ m. The displacement d for the boundary layer from the ground must be considered for higher obstacles in ground classes 6 to 8. Table 7 shows the calculated results. The wind speed decreases significantly with rising roughness lengths z_0 ; thus, it does not make any sense to install windpower plants in built-up areas or large forests.

The wind speed also increases significantly with height. For instance, the wind speed at a height of 50 m is 30 per cent higher than at 10 m for ground class 4. This must be considered for the installation of large wind turbines.

The usable wind speed at the top of large wind towers is much higher than at the common measurement height of 10 m. Wind turbines of the megawatt class come with hub heights of between 50 and 70 m for coastal areas (ground class 1 to 3) and even higher for inland areas with higher roughness lengths.

This example should not give the impression that the wind speed is already independent of the ground at a height of 50 m. The wind speed usually becomes independent of the height,

where the wind becomes known as geostrophic wind, at heights significantly exceeding 100 m from the ground. Finally, the power law of Hellmann is another relation for the vertical distribution of wind speeds.

Ground class	Roughness length z_0 in m	Description
1.- Sea	0.0002	Open sea
2.- Smooth	0.005	Mud flats
3.- Open	0.03	Open flat terrain, pasture
4.- Open to rough	0.1	Agricultural land with low population
5.- Rough	0.25	Agricultural land with high population
6.- Very rough	0.5	Park landscape with bushes and trees
7.- Closed	1	Regular obstacles (woods, village, suburbs)
8.- Inner city	2	Centers of cities with low and high buildings

Table 6 Roughness length and its description according to ground class

Ground class	z_0	d	$v(h_2)$ @ 10 m	Ground class	z_0	d	$v(h_2)$ @ 10 m
1	0.0002	0 m	8.71 m/s	5	0.25	0 m	6.96 m/s
2	0.005	0 m	8.25 m/s	6	0.5	3 m	5.81 m/s
3	0.03	0 m	7.83 m/s	7	1	5 m	4.23 m/s
4	0.1	0 m	7.41 m/s	8	2	6 m	2.24 m/s

Table 7 Calculation results when obtaining wind speed velocities by changing ground class and height

With $z = \sqrt{h_1 \cdot h_2}$ and $a = \frac{1}{\ln\left(\frac{z}{z_0}\right)}$ the Equation 12 becomes $\frac{v(h_2)}{v(h_1)} = \left(\frac{h_2}{h_1}\right)^a$. For $z = 10$ m and $z_0 = 0.01$ m, the parameter a is about 1/7; this equation is then called a 1/7 power law. However, this power law is only valid if the displacement d of the boundary layer from the ground is equal to zero.

3.1.3.2 Power Content of Wind

The kinetic energy E carried by a wind with speed v is given by the general equation:

$$E = \frac{1}{2} m \cdot v^2.$$

Equation 13

The power P that the wind contains is calculated by differentiating the energy with respect to time. For a constant wind speed v the power is:

$$P = \dot{E} = \frac{1}{2} \dot{m} \cdot v^2.$$

Equation 14

The density ρ and volume V determine the mass:

$$m = \rho \cdot V.$$

Equation 15

The derivative with respect to time results in the air mass flow:

$$\dot{m} = \rho \cdot \dot{V} = \rho \cdot A \cdot \dot{s} = \rho \cdot A \cdot v.$$

Equation 16

This mass of air with density ρ flows through an area A with speed v . Hence the power of the wind becomes:

$$P = \frac{1}{2} \rho \cdot A \cdot v^3.$$

Equation 17

For the utilization of wind power a technical system such as a wind turbine should take as much power from the wind as possible. This turbine slows the wind from speed v_1 to speed v_2 and uses the corresponding power difference. If this happened in a pipe with rigid walls at constant pressure, the wind speed v_2 would change with the initial wind speed v_1 , because the same amount of air that enters the pipe must leave it.

Hence, the mass flow of the air before and after the wind turbine is the same. Wind turbines slow down the wind when converting wind energy into electricity; however, the mass flow before and after the wind turbine remains constant. The wind flows through a larger cross-section after

passing through the wind turbine as shown in Figure 11. For constant pressure and density ρ of the air, the mass flow is:

$$\dot{m} = \rho \cdot \dot{V} = \rho \cdot A_1 \cdot v_1 = \rho \cdot A \cdot v = \rho \cdot A_2 \cdot v_2 = \text{const.}$$

Equation 18

The wind speed:

$$v = \frac{1}{2} \cdot (v_1 + v_2)$$

Equation 19

at the height of the wind turbine is the average of the wind speeds v_1 and v_2 . The power P_T taken from the wind can be calculated from the difference in wind speeds:

$$P_T = \frac{1}{2} \cdot \dot{m} \cdot (v_1^2 - v_2^2).$$

Equation 20

With $\dot{m} = \rho \cdot A \cdot v = \rho \cdot A \cdot \frac{1}{2}(v_1 + v_2)$, the expression becomes:

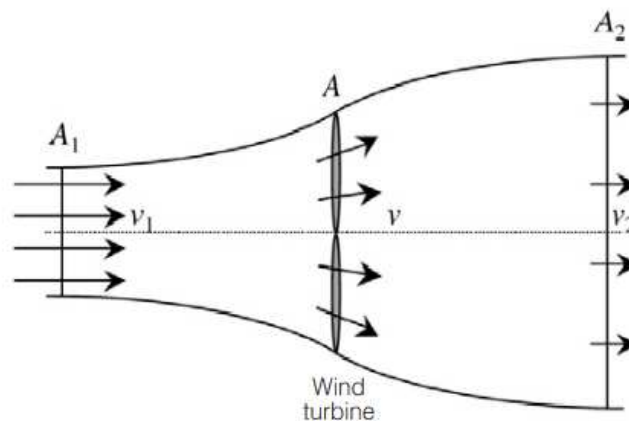


Figure 11 Sketch of changes in wind speed when flowing through wind turbine. As the diameter gets bigger the wind speed slows down up to one third of the original velocity

$$P_T = \frac{1}{4} \cdot \rho \cdot A \cdot (v_1 + v_2)(v_1^2 + v_2^2).$$

Equation 21

The power P_0 of the wind through the area A without the influence of the wind turbine is:

$$P_0 = \frac{1}{2} \cdot \rho \cdot A \cdot v_1^3.$$

Equation 22

The ratio of the power used by the turbine P_T to the power content P_0 of the wind is called the power coefficient c_p and is given by:

$$c_p = \frac{P_T}{P_0} = \frac{(v_1 + v_2) \cdot (v_1^2 + v_2^2)}{2 \cdot v_1^2} = \frac{1}{2} \cdot \left(1 + \frac{v_2}{v_1}\right) \cdot \left(1 - \frac{v_2^2}{v_1^2}\right).$$

Equation 23

Betz has calculated the maximum power coefficient possible, which is called the ideal or Betz power coefficient $c_{p, \text{Betz}}$, Betz:

$$\text{With } \xi = \frac{v_2}{v_1} \text{ and } \frac{dc_p}{d\xi} = \frac{d\left(\frac{1}{2} \cdot (1 + \xi) \cdot (1 - \xi^2)\right)}{d\xi} = -\frac{3}{2} \cdot \xi^2 - \xi + \frac{1}{2} = 0.$$

Equation 24

The ideal wind speed ratio is:

$$\xi_{id} = \frac{v_1}{v_2} = \frac{1}{3}.$$

Equation 25

When substituting the speed ratio in the equation of the power coefficient, the maximum power coefficient becomes:

$$c_{p, \text{Betz}} = \frac{16}{27} = 0.593.$$

Equation 26

If a wind turbine slows down air with an initial wind speed v_1 to one third of v_1 ($v_2 = \frac{1}{3} \cdot v_1$), the theoretical maximum power can be taken, and this maximum is about 60 per cent of the power content of the wind (Quaschnig, 2005).

Chapter 4: Spatial Interpolation

4.1 Data Used in the Project

Wind speed and solar irradiance measurements were obtained for the period 1 Jan 2008 – 31 Dec 2012 at approximately 90 separate locations across the northeast of Mexico. All of the stations were able to measure wind speed and solar irradiance automatically and transmitted the information to the Mexican Weather Ministry via satellite. A great number of the used stations reported infrequent records due to malfunctions, wrong calibrations, and theft of equipment, among others. The distribution of the number of stations in each State is shown here:

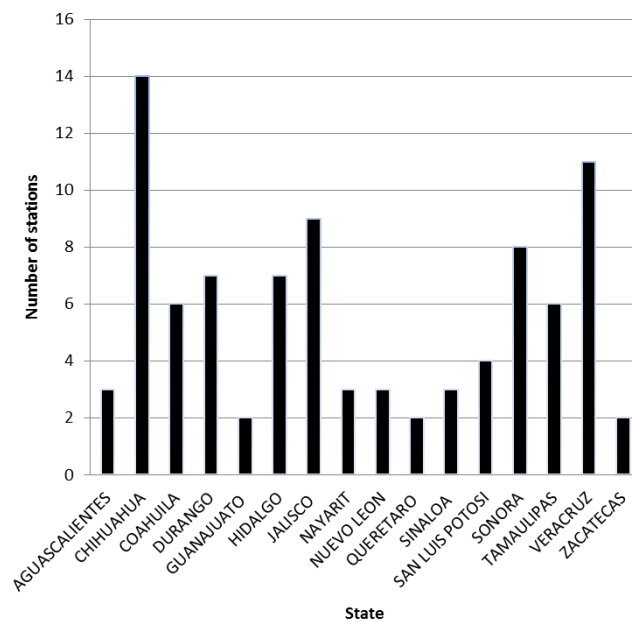


Figure 12 Graphical representation of the number of weather stations in each of the target States of the work

The density of the network was constant in the intervals from 2004 – 2007 and from 2008 – 2011, but kept increasing, because of the installation of new equipment (Figure 13).

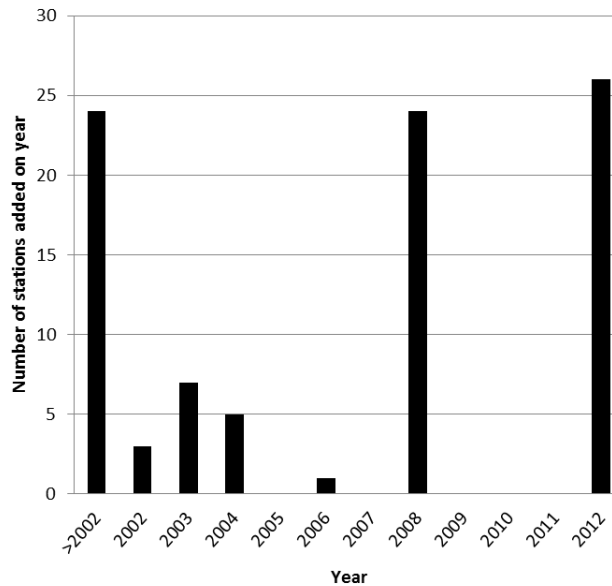


Figure 13 Graphical representation of the number of weather stations that were installed in the period from 2002-2012

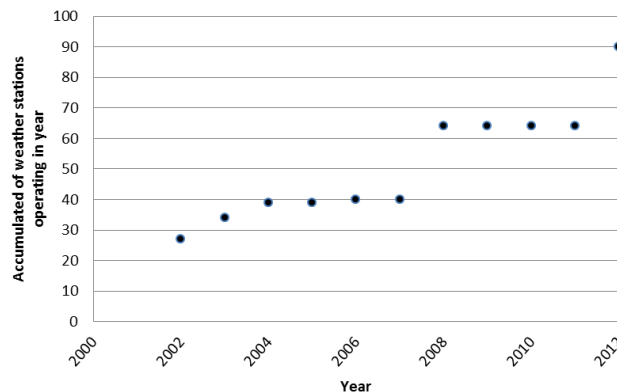


Figure 14 Graphical representation of the accumulated number of weather stations in the period from 2000-2012

A substantial number of stations did not record each hour during a day, so in order to thicken the network used and have more reliable information on the daily wind speed and solar irradiance, all stations were considered, regardless if they measured once a day or once every hour.

Most of the objective regions were not covered well with long-period stations, the study also made use of data from stations with shorter continuous runs and intermittent records where they were available and had sufficient hourly records for dates of interest. As previously mentioned, also stations on neighbor states were considered to carry out the interpolations.

The stations do not follow any pattern for their location, are scattered throughout the 16 states used for the analysis. The elevations of the weather stations ranged from nearly at sea level (Marismas Nacionales, Nayarit, 1 m) to above 3,000 m (Nevado de Colima, Jalisco, 3,461 m), but most were below 1,500 m.

Geographically seen in a map, the stations would be located as shown in Figure 15.

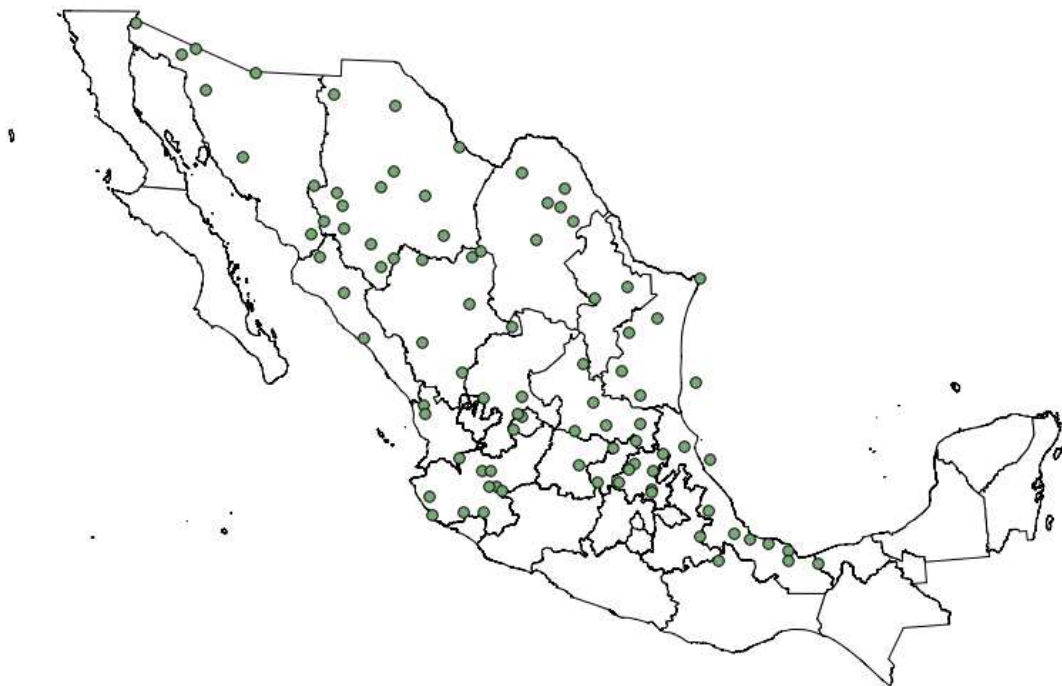


Figure 15 Location of weather stations used. A total of ninety weather stations were used for this work. Some were located on the target States and some on the surrounding States

For the interpolations and further gathering of information, two methods built in with the software QGIS were used (IDW and TIN) as later explained. These methods and the basics of spatial interpolation, spatial sampling and data quality will be described in the next section (Luo, Taylor, & Parker, 2007).

4.2 Automatic Meteorological Stations - AMS

An AMS is a set of electronic and mechanical devices that carry out automatic measurements of different meteorological variables (mostly in a numeric format, Reference OMM 182).

An AMS is made up by a group of sensors that automatically register and transmit weather information of the places where they are strategically located. Their main function is the gathering and monitoring of weather variables to generate 10 minute average files of each variable. That information is sent via satellite in 1 – 3 hour intervals for each station.

The time used for the registry of the data is the UTC (Universal Time Coordinate). The representative area of the stations comprises a 5 km. radius. (Reference OMM # 100, # 168).

Each AMS possesses sensors to measure the following variables:

- Wind direction,
- Wind speed,
- Pressure,
- Temperature and relative humidity,
- Solar irradiation and,
- Precipitation.

There are two types of AMS, which are depicted next (Mexican Weather Ministry, 2013):

Scaffolding Type

Triangular Tower Type

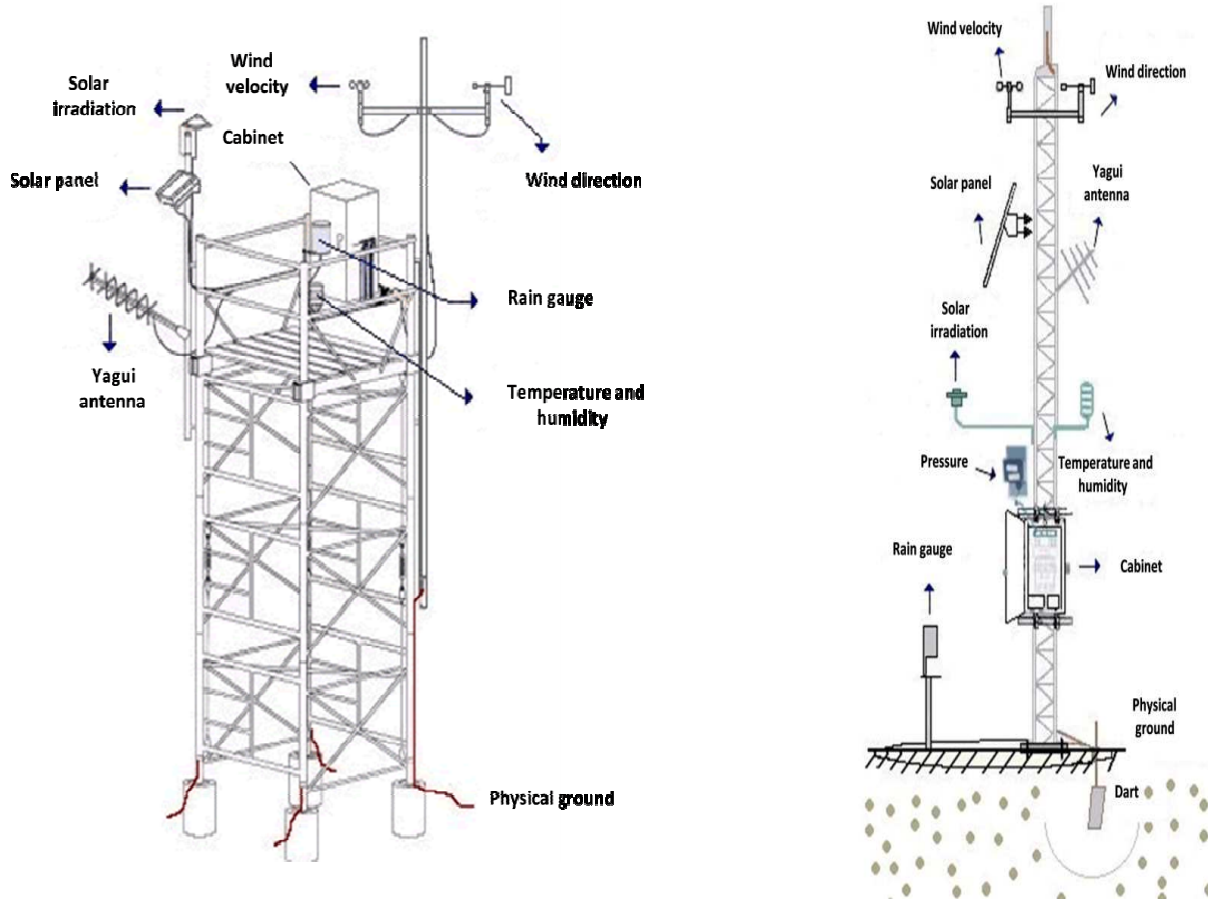


Figure 16 Depiction of the two types of weather stations used with all their parts, structure and sensors

4.3 Spatial Interpolation

Interpolation refers to the process of estimating the unknown data values for specific locations using the known data values for other points.

In many instances we may wish to model a feature as a continuous field (i.e. a 'surface'), yet we only have data values for a finite number of points. It therefore becomes necessary to interpolate (i.e. estimate) the values for the intervening points. For example, we may have measurements of the depth of a particular geological stratum from a number of bore holes, but if we want to model the stratum in 3-dimensions then we need to estimate its depth for places where we do not have bore hole information.

Interpolation may also be required in other situations. For example, as in this project, we have scarce weather information and we would like to know the weather conditions in other places. Interpolation would again be required.

When interpolating from a sample of points we would normally express the estimated values of the intervening points in the same units as were used for the measurements at the sample points, but sometimes we may be more interested in the probability that a certain value is exceeded or that the interpolated value is within a certain range.

Most interpolation methods can be divided into two main types called global and local. Global interpolators use all the available data to provide estimates for the points with unknown values; local interpolators use only the information in the vicinity of the point being estimated. Global interpolators are often used to remove the effects of major trends before using local interpolators to analyze the residuals. Kriging is a particular type of local interpolation using more advanced geostatistical techniques.

Interpolation methods may also be classified as exact or inexact (when inexact is also called “fitting”). Using exact interpolation, the predicted values at the points for which the data values are known will be the known values; inexact interpolation methods remove this constraint (i.e. the observed data values and the interpolated values for a given point are not necessarily the same). Inexact interpolation may produce a smoother (and arguably more plausible) surface.

Interpolation methods may be either deterministic or stochastic. Deterministic methods provide no indication of the extent of possible errors, whereas stochastic methods provide probabilistic estimates.

We will begin by looking at some of the more common forms of spatial sampling; then we will look at global and then local interpolators.

4.4 Techniques

4.4.1 Global interpolators

There are two broad approaches to global interpolation. One uses classification techniques to approximate the values of the desired variable based upon knowledge of the values of other related variables. The other uses regression techniques to infer the value of the variable of interest based upon knowledge of attributes that are easy to measure and approximate with mathematical resources.

4.4.1.1 Classification Techniques

This approach may be used if spatial data are very sparse, although regression techniques (see below) would generally be preferred if there is sufficient data. Classification techniques are global, inexact and deterministic. The basic assumption is that the value of the variable of interest is strongly influenced by another variable which can be used to classify the study area into zones.

4.4.1.2 Trend Surface Analysis

Trend surface analysis is global, inexact and deterministic. A trend surface can be thought of as a high order three dimensional regression surface. To understand this, let us begin with a simple situation in which we can model data values along a transect using a simple regression model (where ε is a random noise factor):

$$z(x) = b_0 + b_1x + \varepsilon.$$

Equation 27

In some cases the data values cannot be adequately summarized as a linear function, in which case a higher order polynomial may provide a better summary. For example, a second order polynomial (or quadratic) equation might provide a better fit:

$$z(x) = b_0 + b_1x + b_2x^2 + \varepsilon.$$

Equation 28

Trend surfaces are similar except, instead of having data values along a transect, the sample points would be in two dimensions (measured by x and y co-ordinates) with the attribute values z modelled as a third dimension. A first order trend surface (analogous to a simple regression line) is an inclined plane with the formula:

$$z(x, y) = b_0 + b_1x + b_2y + \varepsilon.$$

Equation 29

A second order trend surface would be an undulating surface with the formula:

$$z(x, y) = b_0 + b_1x + b_2y + b_3x^2 + b_4xy + b_5y^2 + \varepsilon.$$

Equation 30

Such expressions are called polynomials of second degree in two variables. It should be noted that this model includes a cross product term (i.e. b_4xy). Higher order trend surfaces not only include even higher powers of x and y , but also more cross product terms. A third order trend surface has a total of 10 terms.

Higher order trend surfaces are more convoluted than lower order trend surfaces, and provide closer fits to the observed data values. However, this does not necessarily result in more accurate predictions for the points in between. In fact, trend surfaces higher than third order tend to become counterproductive.

The objective in many instances, consequently, is not to get a 'perfect' fit of the observed data values using a higher order trend, but to identify areas where there are spatially autocorrelated residuals from a low order trend surface as this may indicate the presence of locally important influences upon the variable of interest.

The values of the b terms can be easily determined using the standard regression options available in most statistical packages. The significance of a trend surface can be tested using an analysis of variance test. There is also an analysis of variance test to test whether a trend surface of a given order represents a significant improvement on a trend surface one order lower (Burrough & McDonell, 1998). For more information of this method please refer to the

book "Introduction to Geographical Information Systems" of the authors Burrough & McDonnell, Chapters 5 and 6 and to the book "A Review of Spatial Interpolation Methods for Environmental Scientists" from the authors Jin Li and Andrew Heap, Chapter 2.

4.4.1.3 Regression Techniques

The trend surface techniques discussed in the previous section use only information measured on the variable of interest (i.e. wind velocity or solar irradiance) and the location of the sample points, whereas the classification methods discussed in the previous section make use of 'external' information (i.e. seasonal changes and other limitants). A third strategy is to make use of external information using regression techniques. Regression techniques are global, inexact and stochastic.

There is no restriction on the type of external information that may be used, provided that the regression model is intuitively plausible. However, it obviously makes sense to use information which can be readily obtained.

For example, if it is known that the model that analyzes the variable of interest is the next one:

$$z(x) = b_0 + b_1P_1 + b_2P_2 + \varepsilon.$$

Equation 31

Here P_1 and P_2 are known³, then the regression model determines the relative importance of the two variables using other information available for the sample points. The parameters of the regression model can obviously be estimated using standard regression techniques. Likewise the significance of the regression can be tested using standard techniques. This type of regression which empirically estimates the values of a variable using external information is sometimes referred to as a transfer function (Burrough & McDonnell, 1998). For more information of these methods please refer to the book "Introduction to Geographical Information Systems" of the authors Burrough & McDonnell, Chapters 5 and 6 and to the book

³ The dependent variable could be represented as $z(x,y)$ to maintain consistency with the previous section. However, the handout follows the notation used by Burrough and McDonnell - i.e. the dependent variable is represented as $z(x)$, where x refers to the set of Cartesian coordinates for a particular point. Other texts ignore the locational qualifiers completely, and simply refer to the dependent variable as z .

“A Review of Spatial Interpolation Methods for Environmental Scientists” from the authors Jin Li and Andrew Heap, Chapter 2.

4.4.2 Other Global interpolators

More complex mathematical techniques, such as spectral analysis or Fourier analysis, can be used to model the surface in a manner analogous to trend surface analysis⁴.

They generally require large amounts of data at different scales of resolution.

4.4.3 Local Interpolators

When using global interpolators local variations tend to be dismissed as random noise. However, intuitively this does not make sense as the data values for each point often tend to be very similar to those for neighboring points. Local interpolators therefore attempt to estimate the data values for unknown points using the known data values for points nearby.

The general procedure is to identify a lattice of points for which data values are to be estimated. For each point, the procedure involves the following steps:

1. A search area (neighborhood) is defined around the point,
2. The sample points within the search area are identified;
3. A mathematical function is selected to model the local variation between these points and,
4. The data value for the point is estimated from the function.

Different results will tend to arise depending upon the size of the search area and the type of mathematical function selected. We will discuss a few of the more common functions.

4.4.3.1 Thiessen Polygons

In this approach, Thiessen polygons (also known as Dirichlet or Voronoi diagrams) are constructed around each sample point. All points within a polygon are assumed to have the same data value as the sample point around which the polygon is constructed. This is

⁴ The main difference is that, instead of using polynomials, the surface is modelled as the sum of a number of sinusoidal functions with different wavelengths.

equivalent to saying that each point has the same data value as its nearest sample point. Interpolation using Thiessen polygons would be classed as local, exact and deterministic.

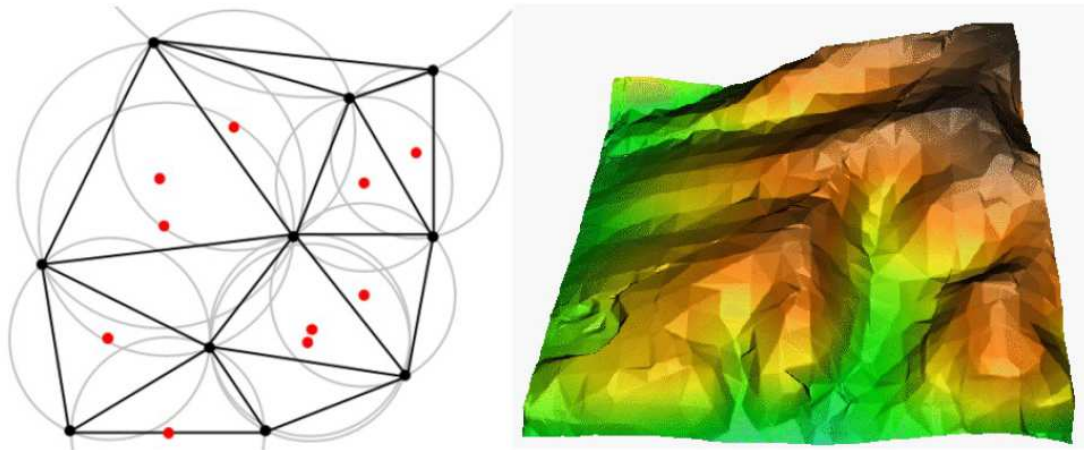


Figure 17 Delaunay triangulation with circumcircles around the red sample data. The resulting interpolated TIN surface created from vector points is shown on the right

Thiessen polygons are constructed by drawing lines between neighboring points - in our case sample points. These lines form the sides of Delaunay triangles (Delaunay is a common Triangulated Irregular Network (TIN) algorithm). The black points represent the vertices of a triangle and the red points represent the centroids of each of the circles made. A Delaunay triangle has the property that a circle drawn through its three corners will never contain any other sample point. If the circle contains a sample point, then the triangles need to be redrawn. To construct the Thiessen polygons, a second set of lines are then constructed to bisect the first set of lines (i.e. the edges of the Delaunay triangles) at right angles at their mid-points (establishment of the circumcircle). The second set of lines form the boundaries of the Thiessen polygons whilst their intersections form the corners of the polygons⁵.

⁵ Figure 5.7 in Burrough and McDonnell, purporting to show Thiessen polygons (and replicated here), is poorly constructed, resulting in some areas being allocated to the wrong polygon.



Figure 18 Formation of Thiessen polygons out of initial sketch

The key property of a Thiessen polygon is that all points within a polygon lie closer to the point around which it was constructed than to any other point.

Using Thiessen polygons to interpolate results in sharp jumps in data values as you move from one polygon to the next, so a technique known as pycnophylactic interpolation is sometimes used to smooth the transition. This technique was developed by Waldo Tobler to 'blur' the boundaries in choropleth maps showing features like population density for administrative regions.

The technique preserves 'volume' (i.e. the total number of people per area), but 'moves' them around within the areas to form a continuous surface. Burrough and McDonnell (pp.116-117) provide details of the mathematics (Burrough & McDonell, 1998).

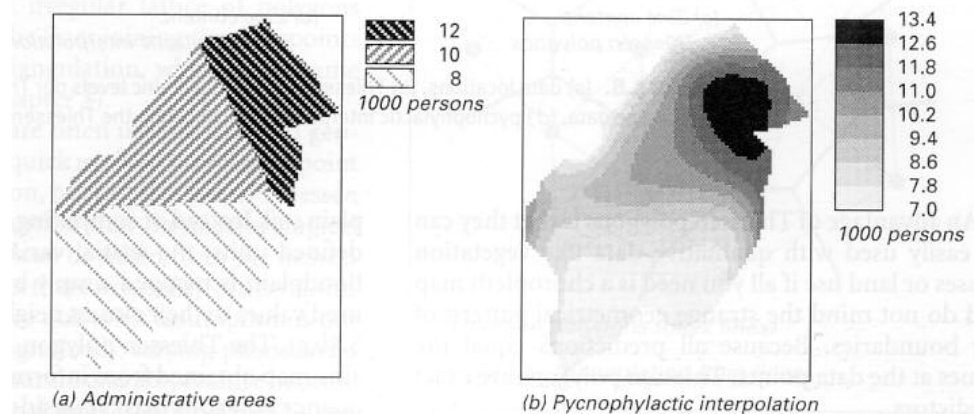


Figure 19 Map used to show the pycnophylactic interpolation technique

For more information of this method please refer to the book "Introduction to Geographical Information Systems" of the authors Burrough & McDonnell, Chapters 5 and 6 and to the book

“A Review of Spatial Interpolation Methods for Environmental Scientists” from the authors Jin Li and Andrew Heap, Chapter 2.

4.4.3.2 Weighted Moving Average Methods

This family of techniques estimates the data value for each point by calculating a distance weighted average of the points within the search radius. These techniques are local, exact and deterministic.

The general formula is:

$$\hat{z}(x_0) = \sum_{i=1}^n \lambda_i z(x_i).$$

Equation 32

In this case $z(x_i)$ are the data values for the n points $(x_1, x_2 \dots x_n)$ within the search radius, and λ_i are weights to be applied to the data values for each point. One constraint is that the weights must add up to 1.0. The weights are usually some function of the distance between the point for which the estimate is being made and the sample points. The most common function is the inverse distance weighting (IDW) predictor. The above formula then becomes:

$$\hat{z}(x_0) = \sum_{i=1}^n \left\{ \frac{d_{ij}^{-r}}{\sum_{i=1}^n d_{ij}^{-r}} z(x_i) \right\}$$

Equation 33

where j represents the point whose value is being interpolated, d_{ij} is the distance from point j to sample point i and r is an arbitrary value which can be selected by the investigator. If r is set equal to 1, then this becomes a simple linear interpolator. The r value is frequently set to 2, in which case the influence of each sample point is in proportion to the square root of its distance from the point to be interpolated. However, r can be set to higher values if required. High values of r give even more weight to the nearer sample points.

If the point to be interpreted corresponds exactly with one of the sample points, d_{ij} would be equal to zero and the sample point would be assigned an infinite weight, producing an infinite value for the estimate. This is obviously undesirable, so in such situations the point to be interpolated is simply assigned the same value as the sample point.

One problem with this particular technique is that isolated data points tend to produce 'duck-egg' patterns⁶. One solution might be to increase the search radius, but this may have the undesirable effect of over-smoothing other parts of the surface.

Another variant is to base the interpolation at each point on a fixed number of sample points. The search radius will then be small in areas where there is a high density of sample points, but larger where there is a lower density of points. If it is believed that there may be a directional bias, the search window could even be divided into, say, 4 quadrants and its radius might be increased until there was a minimum of a specified number of points in each quadrant (Burrough & McDonell, 1998).

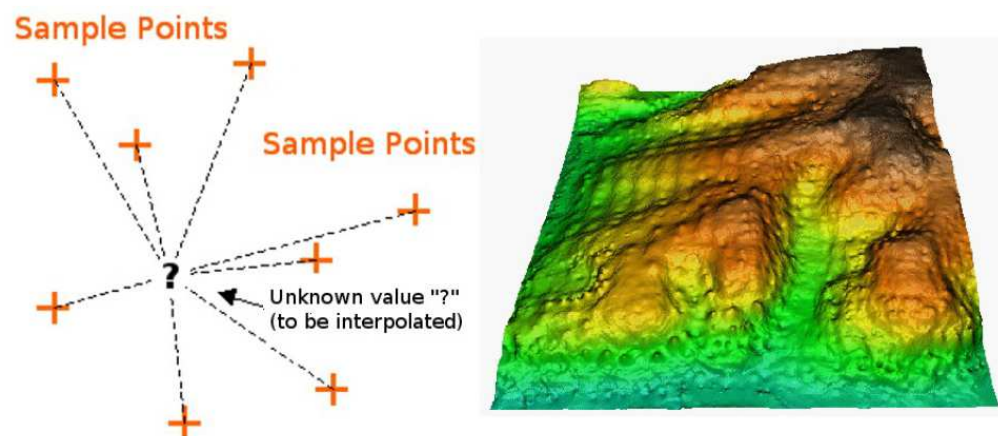


Figure 20 Inverse Distance Weighted interpolation based on weighted sample point distance (left). Interpolated IDW surface from vector points (right)

For more information of this method please refer to the book "Introduction to Geographical Information Systems" of the authors Borrough & McDonnell, Chapters 5 and 6

⁶ 'Duck egg patterns' are ellipses or circles around an isolated point, which are different from the background and therefore very visible. This only occurs in the case when there is only one study point in a large area and the elliptic shape will occur when there is only a small influence from a far away neighbor to the isolated point. For an illustration example of this please refer to: <http://www.chec.pitt.edu/gismapsdetail.html>

and to the book “A Review of Spatial Interpolation Methods for Environmental Scientists” from the authors Jin Li and Andrew Heap, Chapter 2.

4.4.3.3 *Kriging*

Kriging is a form of local interpolation using geostatistical methods developed by a French geostatistician called Georges Matheron and a South African mining engineer called D.G. Krige. Kriging is local, exact and stochastic. Whilst much more complex than the methods discussed above, it provides a number of advantages:

1. Given sufficient data, kriging produces better estimates than the other methods because the method takes explicit account of the effects of random noise,
2. Although the investigator can choose between different types of kriging, the method is less susceptible to arbitrary decisions taken by the investigator (e.g. search distance, number of sample points to use, location of break points, etc.). Kriging identifies the optimal interpolation weights and search radius and,
3. Kriging provides an indication of the reliability of the estimates.

Kriging begins with the recognition that the spatial variation of any continuous attribute is usually too complex to be modelled by a simple, smooth mathematical function. So, instead, it is modelled as a stochastic surface or random field. Regionalized variable theory assumes that the spatial variation of any variable can be expressed as the sum of three components:

1. A structural component having a constant mean or trend,
2. A random, but spatially correlated component, known as the regionalized variable and,
3. Spatially uncorrelated random noise i.e. residual component.

The value of a random variable Z at x is given as:

$$Z(x) = m(x) + \varepsilon'(x) + \varepsilon''$$

Equation 34

where $m(x)$ is a structural function describing the structural component, $\varepsilon'(x)$ is the stochastic but spatially auto correlated residuals from $m(x)$ (i.e. the regionalized variable), and $\varepsilon''(x)$ is random noise having a normal distribution with a mean of 0 and a variance σ^2 .

The first step is to decide on a suitable function for $m(x)$. In the simplest case this can be thought of as a flat surface with no trend. The mean value of $m(x)$ is the mean value within the sample area. This means the expected difference in the values for two points x and $x + h$ (where h is the distance between the points) is zero, i.e

$$E[Z(x) - Z(x + h)] = 0.$$

Equation 35

It is also assumed that the variance of the differences is a function of the distance between the points (i.e. they are spatially autocorrelated, meaning that near points are more likely to have similar values, or smaller differences, than distant points) - i.e.

$$E[\{Z(x) - Z(x + h)\}^2] = E[\{\varepsilon'(x) - \varepsilon'(x + h)\}^2] = 2\gamma(h)$$

Equation 36

where $\gamma(h)$ is known as the semivariance. Under these two assumptions (i.e. Stationarity of difference and stationarity in the variance of differences), the original model can be expressed as:

$$Z(x) = m(x) + \gamma(h) + \varepsilon''.$$

Equation 37

The semivariance can be estimated from the sample data using the formula:

$$\gamma(h) = \frac{1}{2n} \sum_{i=1}^n \{z(x_i) - z(x_i + h)\}^2,$$

Equation 38

where n is the number of pairs of sample points.

The semivariance can be calculated for different values of h . A plot of the calculated semivariance values against h is referred to as an experimental variogram.

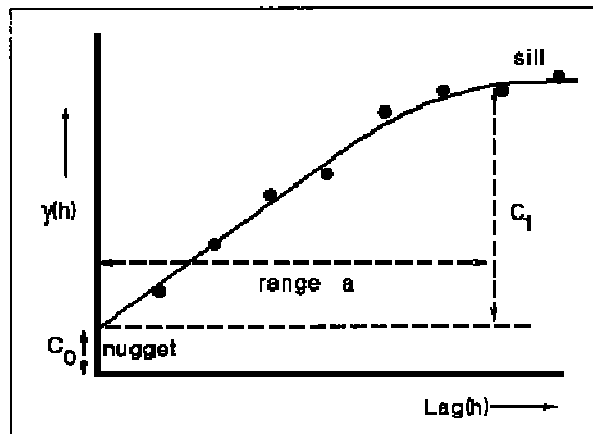


Figure 21 An example of a semivariogram illustrated by an exponential model

Experimental variograms typically have a number of characteristic features, as illustrated in the diagram:

1. The variance is small for low values of h , but increases as the value of h gets larger. However, beyond a certain point the graph levels off to form what is known as the sill,

2. The distance at which the graph levels off is known as the range. At distances less than the range, points closer together are more likely to have similar values than points further apart. At distance larger than the range, points do not exert an influence upon one another. The range therefore provides an indication of how large the search radius needs to be when doing a distance weighted interpolation and,

3. The fitted model does not pass through the origin - i.e. according to the graph the semivariance when h is zero has a positive value referred to as the nugget. However, one would expect the variance to be zero (i.e. one would

expect the difference between points and themselves to be zero). The nugget provides an indication of the amount of non-spatially autocorrelated noise (i.e. ε'').

The semivariance depicted in an experimental variogram must be modelled by a mathematical function taking the form $\gamma(x) = \dots$. Different models may be used, each with a different formula. The choice between these models is determined by the shape of the experimental variogram. The figure 22 shows some of the more commonly used models.

A spherical model (a) is used when the variogram has the 'classic' shape, an exponential model (b) is used when the approach to the sill is more gradual, a Gaussian model (d) may provide a good fit when the nugget is small and the variation is very smooth, and a linear model (c) may be the most appropriate when there is no sill within the study area. Fitting the most appropriate model to the data requires a lot of skill.

Other variogram shapes may indicate a different course of action is required. For example:

1. A variogram that becomes increasingly steep with larger values of h indicates that there is a trend in the data that should be modelled separately,
2. If the nugget variance is large and the variogram shows no tendency to diminish with smaller values of h , then interpolation is not really sensible. The best estimate of $z(x)$ is simply the overall mean of the sample points,
3. A noisy variogram showing no particular pattern may indicate that there are too few sample points,
4. If the range is smaller than the distance between sample points, then the sample points are too far apart to influence one another. The best estimate of $z(x)$ is again the overall mean of the sample points and,
5. If the variogram dips at distances further than the range to create a 'hole effect', then the study area may be too small to capture some long wave-length variation in the data.

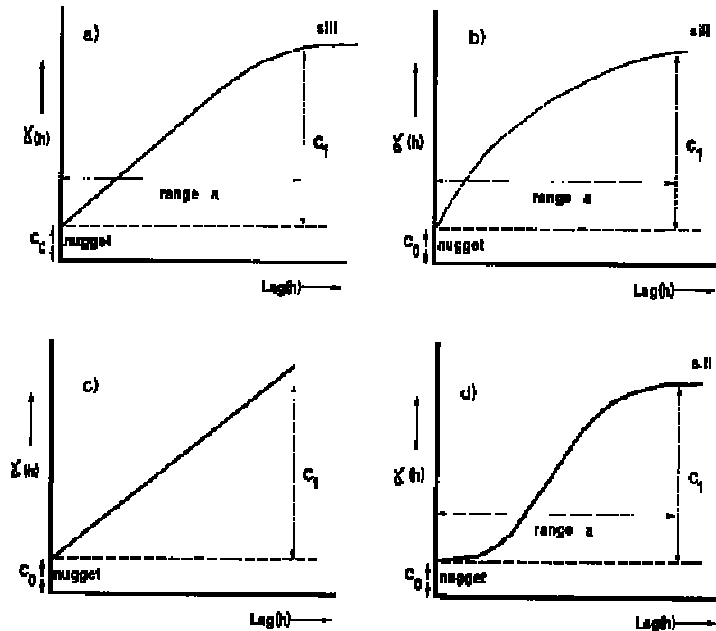


Figure 22 Examples of four commonly used variogram models: (a) spherical; (b) exponential; (c) linear; and (d) Gaussian

More complex variogram models can be developed where it is felt they are required, for example:

1. Anisotropic models may be developed if it is believed that there may be directional effects (i.e. that the relationship between the semivariance and distance is different in different directions). This essentially involves calculating a different experimental variogram for each direction. These can be shown as a contour 'map' rather than a graph,
2. A complex variogram may be required if it is believed that the total variation is the sum of the effects of two or more regionalized variables and,
3. Different variograms may be required for different cover classes (e.g. rock types, land covers, etc.). In such situations a separate variogram should be calculated for each cover class.

Once the variogram has been modelled, the next step is to determine the weights λ_i required for local interpolation. The weights λ_i for each point should sum to 1.0. They are also selected so that the estimated data value for the point is unbiased, and the variance of the estimation is minimized - i.e. the estimate is a best linear unbiased estimate. It can be shown that the estimation variance σ_e^2 is minimized when:

$$\sum_{i=1}^n \lambda_i \gamma(x_i, x_0) + \phi = \gamma(x_j, x_0)$$

Equation 39

where n is the number of samples points used to interpolate the data value for point x_0 and ϕ is a Lagrange multiplier (which ensures that the λ_i values add up to 1.0). This equation can be solved by substituting the estimated values of γ from the variogram model. Burrough and McDonnell provide a worked example on pp. 140-141.

Finally, the data value for each point can be estimated by inputting the weights into the formula:

$$\hat{z}(x_0) = \sum_{i=1}^n \lambda_i z(x_i).$$

Equation 40

The estimation variance σ_e^2 (also known as the kriging variance) for each point can be calculated using the formula:

$$\widehat{\sigma_e^2} = \sum_{i=1}^n \lambda_i \gamma(x_i, x_0) + \phi.$$

Equation 41

The estimation variance σ_e^2 (or more usually the kriging standard deviation σ_e) can be mapped. This provides an indication of the reliability of the estimate at each location.

The validity of the variogram model may also be tested for consistency by comparing the actual values for a data point with the estimated value calculated using a variogram calculated

using all the other data points. This process, known as cross validation, is repeated omitting each point in turn. If the mean difference between the predicted values and actual values is close to zero with a small variance then the variogram can be assumed to unbiased.

The above method, known as ordinary kriging, is the most common form of kriging. However, there are other variants. Block kriging is used to interpolate data values to areas larger than the support, simple kriging can be used under an assumption of second order stationarity and a known mean, but as these conditions rarely apply it is seldom used on its own; non-linear kriging can be used if the data are non-normally distributed (e.g. log-normal data), indicator kriging can be used to interpolate binary data, stratified kriging is used to interpolate data stratified by different cover classes, co-kriging can be used to make use of data different from, but correlated with, the variable to be estimated, whilst universal kriging is used to incorporate information about trends.

Finally conditional simulation techniques (or stochastic imaging) may be used to identify the most likely data value for each place (but at the expense of continuity between adjacent areas). Burrough and McDonnell provide further details on most of these methods (Burrough & McDonnell, 1998). For more information of this method please refer to the book "Introduction to Geographical Information Systems" of the authors Burrough & McDonnell, Chapters 5 and 6 and to the book "A Review of Spatial Interpolation Methods for Environmental Scientists" from the authors Jin Li and Andrew Heap, Chapter 2.

Chapter 5: Factors Affecting the Performance of Spatial Interpolation Methods

In this chapter, several factors that affect the performance of the spatial interpolation methods are discussed. The impacts of sampling density, variation in the data, sampling design and stratification on the estimation of the spatial interpolation methods will be analyzed.

5.1 Sampling Design and Sample Spatial Distribution

Data density plays a significant role in the performance of the spatial interpolation methods. The following sections discuss the effects of data density on the performance of the spatial interpolation methods.

5.1.1 High Density

When data density is high, most methods produce similar results (Burrough & McDonell, 1998)). Using datasets of regularly spaced and high density samples, Gotway et al. (1996) found that the use of wider sample spacings greatly reduced the information in the resultant maps, although the sample density was still relatively high.

5.1.2 Low Density

When data are sparse, the underlying assumptions about the variation among samples may become critical (Burrough & McDonell, 1998). The performance of the spatial interpolation methods is better when the sample density is greater (Englund, Weber, & Leviant , 1992).

However, it is claimed that the accuracy of regression modelling is not really dependent on the sampling density, but rather on how well the data are sampled and how significant the correlation is between the primary variable and secondary variable(s) (Hengl, 2007).

Sample size also affects the predicted error. It was found that with small samples, some methods may dramatically over- or under-predict the predicted estimation error. This suggests that such predicted errors should not be used in an absolute sense, but as a relative measure of spatial estimation accuracy (C.E. & R.L., 1986).

In addition, it is found that the smoothing of the estimations (or map) increased at lower sample densities (Goovaerts, 2000). Issues relating to sample size are further discussed below.

5.1.2.1 *Sample Spatial Distribution*

Sample spatial distribution may affect the performance of the spatial interpolation methods.

Sample clustering affects the accuracy of the estimations and the effects may also depend on the spatial interpolation methods. While spatial scale, relative spatial density and distribution of samples can be determinant factors on the performance of the spatial interpolation methods (Collins & Bolstadt, 1996), other relevant factors may also be important. For example, altitudinal and seasonal changes in data have been shown to play a significant role in predictions (Stahl et al., 2006). Where temporal scales are short, preliminary data analyses are especially important to determine the suitability of a particular spatial.

Sample size and sampling configuration or design affect the reliability of the variogram. Generally, if the sample size is <50, the variograms derived are often erratic with little or no evident spatial structure (Webster & Oliver, 1992). The larger the sample size from which the variogram is computed, the more precisely is it estimated, although the precision is unknown in most instances (Webster & Oliver, 1992). If the sample size is too small, a noisy variogram would be generated (Burrough & McDonell, 1998).

Sample spacing must relate to the scale or scales of variation in a region, otherwise samples might be too sparsely spaced to identify correlation and could result in a pure nugget (Webster & Oliver, 1992). In such cases, the accuracy of the estimation could be reduced, as evidenced by the findings in Gotway et al. (1996) that the use of wider sampling spacings greatly reduced the information in the resultant maps. In addition, the smoothness of the estimations (or map) may increase with the relative nugget effect (Goovaerts, 2000).

The spatial structure of the data may also affect the sample size and variogram. For data with a short range of variograms, intensive sampling with a large proportion of clustered points is required; and conversely for variables with a long range, fewer and more evenly spaced samples are required (Marchant & Lark, 2006). Variogram is also sensitive to sample clustering, particularly when it is combined with a proportional effect that is a form of heteroscedasticity

where the local mean and local variance of data are related (Goovaerts, 2000).

The number of pairs of samples at each lag is an important factor that needs to be considered in modelling the variogram. A rule of thumb, as suggested by Burrough and McDonnell (1998) is that at least 50-100 samples are necessary to achieve a stable variogram.

5.1.3 Sample Size and Spatial Interpolation Methods

Findings imply that the effects of sample size on the estimations depend on the spatial interpolation methods. In practice, it is believed that there is a threshold beyond which any increase in sample size does not improve much the accuracy of the estimations; otherwise sample size is still a critical factor. Other factors like variance inherited in the data also play a significant role.

5.2 Data Quality

Five major factors relevant to the quality of the data are discussed in this section: distribution, isotropism⁷ and anisotropism⁸, variance and range, accuracy, spatial correlation and other factors. The sources of errors in spatial continuous data and factors affecting the reliability of spatial continuous data have been discussed in the previous section.

5.2.1 Distribution

Data normality can influence the estimation of certain spatial interpolation methods that assume that the input data are distributed normally about their mean. Data normality can be tested using such as the Kolgorov-Smirnov test. If this assumption is not met, log transformation is commonly applied, thus resulting in lognormal methods (e.g., lognormal kriging; Cressie, 1993).

The predictions are then transformed back to the original scale by a marginally unbiased back transformation proposed by Cressie (1993). However, back-transforming the estimated values can be problematic because exponentiation tends to exaggerate any interpolation-related error (Goovaerts, 2000). In addition, the prediction error may also be used to determine whether the data should be transformed (Nadler & Wein, 1998).

⁷ *Identical in all directions; invariant with respect to direction.*

⁸ *Having properties that differ according to the direction of measurement, not isotropic.*

Wu et al. (2006) found that data transformation of highly skewed data generally improved the estimations only by some methods.

5.2.2 Isotropism and Anisotropism

Isotropism of data is assumed for kriging methods. Data may display evidence of anisotropism, which should be considered in the modelling; otherwise biased estimation may result. However, in some cases, the anisotropism could be ignored to simplify model fitting and to maintain some consistency between the semivariograms in the multivariate model (Martinez-Cob, 1996). Conditions that allow for this are:

1. Anisotropism is not evident with the specified search distance;
2. The secondary variable and primary variable are collocated, thus the influence of surrounding values would be small, so anisotropy would make little difference and,
3. The directions of maximum and minimum spatial variability for the different variables did not coincide.

A similar result was also observed by Haberlandt (2007). It was found that the impact of the semivariogram on interpolation performance was not great because no significant differences could be found in prediction performance between isotropic and anisotropic variograms, although anisotropy was clearly present in the data. The best estimations were obtained using an automatic fitting procedure with isotropic variograms.

5.2.3 Variance and Range

The variance of the data affects the performance of the spatial interpolators and the resultant predictions. The performance of the spatial interpolation methods decreased rapidly when the coefficient of variation (CV) increased (Martinez-Cob, 1996).

It was also found that the variance and range of the data can influence the performance and choice of a spatial interpolation technique after comparing eight spatial interpolators across two regions for two temperature variables (maximum and minimum) at three temporal scales (Collins & Bolstadt, 1996).

When temperature variances were large, the performance of all spatial interpolation techniques suffered, which means that increasing temperature variance negatively affected the performance of the spatial interpolators. Temporal scale also affected the choice of a spatial interpolator as temperature range, temperature, variance, and temperature correlation with elevation, all changed with temporal scale (Collins & Bolstadt, 1996).

The chosen sampling scheme also affects the performance of the spatial interpolation methods through the variation in the data. Data should be collected at a range of separations to capture changes in the scales of the variation (Laslett, 1994).

5.2.4 Accuracy

Data accuracy is an important factor influencing the estimations of the spatial interpolation methods. Where the data are not representative of the surface being modelled, it may result in interpolation biases (Collins & Bolstadt, 1996). Where sample elevations are not representative of regional elevations, care must be taken in comparing observed and interpolated data.

Data noise can negatively affect the performance of the spatial interpolation methods (i.e., OK, UK and IDS in Zimmerman et al., 1999; NaN in Webster and Oliver, 2001). When data are too noisy, a pure nugget effect is produced in the variogram and the resultant interpolation is not sensible (Burrough & McDonell, 1998). In contrast, sampling precision (i.e., zero error and high-level normally distributed error with a relative standard deviation of 32% of the true value) was found not to be significant in determining the performance of a spatial interpolator, being OK in this case (Englund, Weber, & Leviant, 1992).

Outliers affect the performance of the spatial interpolation methods and interact with sampling schemes. The variogram is sensitive to outliers and to extreme values. Exceptionally large or small values will distort the average as evident from its definition. This effect depends on the location of the data point in the region and also on the spatial pattern of data (Webster & Oliver, 1992).

All outliers must be regarded with suspicion and investigated. Outliers should be removed if they are believed to not belong to the population and strongly skewed distributions need to be transformed to approximate normal before conducting geostatistical analyses (Webster & Oliver,

1992). For example, removing outliers resulted in considerable improvement in the performance of the spatial interpolation methods, particularly when additional samples were included to allow estimation of short-range variation (Laslett, 1994).

5.2.5 Spatial Correlation and Other Factors

Spatial correlation in samples is also essential for reliable estimation. The performance of different spatial interpolation methods changes with the variable estimated. It was found that the best method varied as a function of the region and the spatial scale required for estimation

Accuracy was lower in regions of great topographic complexity and regions with contrasting atmospheric or oceanic influences than in flatter regions or regions with constant atmospheric patterns. Even the performance of the same spatial interpolation method differed considerably with different variables, which resulted from the fact that data quality (in this case variance and range) changed with different variables (Li & Heap, 2008).

Chapter 6: Procedure for Elaborating the Tool

6.1 Description of the Tool

The results of this thesis were programmed into an interactive, user friendly tool using the program Stadt Planet Plus. One file was programmed independently for each State, in which the user is able to choose any municipality for analysis. A total of 354 municipalities were analyzed. Each municipality is divided into 100 km² squares, called features (as explained in this chapter). A total of 12,556 features were analyzed. So when the user runs the mouse over one of those features, the information regarding the wind velocity, the solar irradiance, the distances to the next transmission line and next highway or road and the coordinates of the centroid are displayed.

The program also enables the user to do animations of the parameters (wind speed and solar irradiance) along the 5 year time interval in which the interpolations were carried out, with this the user can notice how the parameters have been changing in the selected time frame. It is also possible for the user to compare different sections of a map and graph parameters. A different number of graphs (bubble chart, time series graph, bar or column chart) can be made to get conclusions or analyze the information.

The values of the parameters are distinguished with different color tones so it is easier to acknowledge where the potential is greater, making the analysis easier. The tool also contains a zoom in case the user would like to enlarge any municipality or section in the maps. A CD containing the different executable files (.exe) are attached to this work. Those executable files are only to be run on Windows as operating system because they are Flash-based, MAC / Apple are not suitable for this task.

Please refer to the Annexes section for instructions on how to operate the tool.

6.2 Building the Data Base

6.2.1 First Step – Standardizing formats

Unfortunately, the format of the weather stations was different in most of the spreadsheets, so in order to have uniformity among them they had to be arranged in the same way. Next is an example of two different formats that were found.

	A	B	C	D	E	F	G	H	I	J	K	L
1	DirViento	DirRafaga	RapViento	RapRafaga	TempAire	HumRelativa	PresBarometric	Precipitacion	RadSolar	TempComb	nombre_estacion	
2	32	21	15.6	38.5	21.4	67	932.8	0	0	20.6	CUMBRES DE MONTEREY I, N.L.	
3	80	64	14.4	40	21.1	68	932.9	0	0	20.3	CUMBRES DE MONTEREY I, N.L.	
4	46	34	14.1	31.3	20.8	71	933	0	0	19.9	CUMBRES DE MONTEREY I, N.L.	
5	26	27	15.5	37.1	20.4	73	933.1	0	0	19.6	CUMBRES DE MONTEREY I, N.L.	
6	16	41	15.7	34.6	20.2	74	933.2	0	0	19.5	CUMBRES DE MONTEREY I, N.L.	
7	25	40	21.2	42.8	20.1	76	933.3	0	0	19.4	CUMBRES DE MONTEREY I, N.L.	

Figure 23 First format found in the information obtained. The different variables and their order can be seen. This format was chosen to work with, meaning that all the obtained spreadsheets and different variables with other formats were arranged this way⁹

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Rh	ATC	SR	GTempC	Rain	AvgWSU	AvgWSV	WSMax	WDMMax	BP	Vis	Batt	SPanel	WVError	WPError	ReplicateHistory		
2	%	C	W/m^2	C	mm	m/s	m/s	m/s	deg	mbar	Km	V	V	code	code			
3	44	25	149	26.2														172.29.71.254 20080901 001407
4	50	24.6	66	25.7														172.29.71.254 20080901 002352
5	54	23.9	56	24.9														172.29.71.254 20080901 003333
6	56	23.4	54	24.3														172.29.71.254 20080901 004404
7	60	23	37	23.7														172.29.71.254 20080901 005348

Figure 24 Second format found in the information obtained. The different variables and their order can be seen. The spreadsheets with this order were arranged to match the format of Figure 24. An approximate of 5,400 files had to be rearranged¹⁰

The first format was chosen to work with and all the 5,400 files had to be arranged in that way, this was neither a short nor an easy task. The objective of this was to have the target weather variables (wind speed and solar irradiance) in the same order to make the information easy to work with. This could not be programmed in Excel VBA because of the many different formats; it had all to be done manually.

6.2.2 Second Step – Calculating monthly averages and arranging information

After having uniformity in the data, it was proceeded to construct for each of the 90 weather stations a yearly file in Microsoft Excel with all the monthly spreadsheets together in different tabs. To those spreadsheets followed a last one with a summary of the monthly averages so in

⁹ DirViento=wind direction, DirRafaga=gust direction, RapViento=wind speed, RapRafaga=gust speed, TempAire=wind temperature, HumRelativa=relative humidity, PresBarometric= barometric pressure, Precipitacion=rain, RadSolar=solar irradiance, TempComb=enviroment temperature, nombre_estacion=station's name

¹⁰ Rh=relative humidity, ATC=environment temperature, SR=solar irradiance, GTempC=wind temperature, Rain=rain, AvgWSU=average wind speed (low), AvgWSV=average wind speed (sustained), WSMMax=average wind speed (max), WDMMax=wind direction, BP=barometric pressure, Vis=visibility, Batt= voltage of battery, SPanel=voltage from panel

the same spirit the yearly averages for each weather station were calculated. Next is shown said summary spreadsheet:

	A	B	C	D	E	F	G	H	I	J
1		DirViento	DirRafaga	RapViento	RapRafaga	TempAire	HumRelativa	PresBarometric	Precipitacion	RadSolar
2	JANUARY	176.73	217.08	5.64	10.72	17.56	66.74	907.92	0.44	336.20
3	FEBRURARY	173.00	212.94	6.12	11.39	18.45	71.81	906.41	0.45	307.02
4	MARCH	159.85	203.78	6.62	13.01	21.74	64.13	904.66	1.14	472.72
5	APRIL	158.12	198.92	6.71	13.51	24.44	58.35	903.46	0.29	501.33
6	MAY	152.23	190.10	6.95	14.30	26.44	56.56	903.20	0.29	502.44
7	JUNE	159.95	202.64	6.64	14.03	26.24	62.04	902.93	0.43	469.15
8	JULY	149.44	197.80	6.59	13.92	25.48	64.33	906.12	0.86	471.27
9	AUGUST	173.59	215.88	5.95	12.27	25.57	66.43	904.78	0.54	448.52
10	SEPTEMBER	170.07	208.80	5.73	11.59	24.65	67.50	906.05	0.88	435.93
11	OCTOBER	168.68	209.63	5.84	11.61	23.28	64.54	905.63	0.78	464.49
12	NOVEMBER	166.87	208.43	5.14	10.02	19.62	70.82	909.19	0.34	321.84
13	DECEMBER	171.85	215.44	5.44	10.43	18.54	60.26	906.00	0.20	343.37
14	TOTAL AVERAGE YEAR	165.03	206.79	6.11	12.23	22.67	64.46	905.53	0.55	422.86

Figure 25 Depiction of the monthly average summary used for each weather station and each variable. An approximate of 500 files were obtained at the end with this format. This format allowed the concentration and easy manipulation of the information¹¹

6.2.3 Third Step – Yearly data bases for each State

After that, every summary spreadsheet of each station was gathered in another file. In this file the weather stations were divided by State as shown next for the weather stations “La Florida” and “Calvillo” in the states of Aguascalientes and Zacatecas respectively.

	A	B	C	D	E	F	G	H	I	J	K	L
1	AGUASCALIENTES											
2												
3												
4	STATION	CALVILLO										
5	YEAR	2008										
6		DirViento	DirRafaga	RapViento	RapRafaga	TempAire	HumRelativ	PresBaromet	Precipitaci	RadSolar		
7	TOTAL AVERAGE YEAR	125.24	126.83	9.41	14.81	19.22	52.68	838.65	0.90	496.22		
8												
9	YEAR	2009										
10		DirViento	DirRafaga	RapViento	RapRafaga	TempAire	HumRelativ	PresBaromet	Precipitaci	RadSolar		
11	TOTAL AVERAGE YEAR	129.88	127.35	8.93	13.95	20.77	54.33	838.39	0.65	477.10		
12												
13	YEAR	2010										
14		DirViento	DirRafaga	RapViento	RapRafaga	TempAire	HumRelativ	PresBaromet	Precipitaci	RadSolar		
15	TOTAL AVERAGE YEAR	130.11	128.63	8.99	13.99	18.88	49.06	609.34	670.94	487.83		
16												
17	YEAR	2011										
18		DirViento	DirRafaga	RapViento	RapRafaga	TempAire	HumRelativ	PresBaromet	Precipitaci	RadSolar		
19	TOTAL AVERAGE YEAR	128.94	126.42	9.61	15.20	20.36	42.60	838.11	0.60	477.49		
20												
21	YEAR	2012										
22		DirViento	DirRafaga	RapViento	RapRafaga	TempAire	HumRelativ	PresBaromet	Precipitaci	RadSolar		
23	TOTAL AVERAGE YEAR	127.83	141.35	8.65	13.65	20.05	48.78	838.61	0.56	472.01		
24												

Figure 26 Screenshot of yearly averages done for each weather station (Aguascalientes)¹²

¹¹ See footnote on page 63 for a translation of each variable

	A	B	C	D	E	F	G	H	I	J	K	L
2058	ZACATECAS											
2059												
2060	STATION	LA FLORIDA										
2061	YEAR	2008										
2062		DirViento	DirRafaga	RapViento	RapRafaga	TempAire	HumRelativ	PresBaromet	Precipitaci	RadSolar		
2063	TOTAL AVERAGE YEAR	136.24	194.05	6.17	11.91	17.37	68.52	820.58	0.91	405.37		
2064												
2065	YEAR	2009										
2066		DirViento	DirRafaga	RapViento	RapRafaga	TempAire	HumRelativ	PresBaromet	Precipitaci	RadSolar		
2067	TOTAL AVERAGE YEAR	143.20	205.11	5.93	12.30	17.36	54.92	820.14	0.67	437.67		
2068												
2069	YEAR	2010										
2070		DirViento	DirRafaga	RapViento	RapRafaga	TempAire	HumRelativ	PresBaromet	Precipitaci	RadSolar		
2071	TOTAL AVERAGE YEAR	141.35	210.84	5.87	12.19	16.76	54.75	683.19	205.30	448.21		
2072												
2073	YEAR	2011										
2074		DirViento	DirRafaga	RapViento	RapRafaga	TempAire	HumRelativ	PresBaromet	Precipitaci	RadSolar		
2075	TOTAL AVERAGE YEAR	134.65	202.69	6.38	13.51	18.09	46.11	819.91	0.66	451.65		
2076												
2077	YEAR	2012										
2078		DirViento	DirRafaga	RapViento	RapRafaga	TempAire	HumRelativ	PresBaromet	Precipitaci	RadSolar		
2079	TOTAL AVERAGE YEAR	138.86	208.86	6.06	12.59	17.68	51.86	820.39	0.69	442.37		
2080												

Figure 27 Screenshot of yearly averages done for each weather station (Zacatecas)¹³

6.2.4 Fourth Step – Concentration of variables of interest

Up to this point, all the information was processed but special attention had to be put to the variables of interest (wind speed and solar irradiance) in the years of interest (2008-2012), so in another spreadsheet the variables of interest for the desired period were concentrated. This is in the following Figure depicted:

	A	B	C	D	E	F	G	H	I	J	K	L
1		2008				2009				2010		
2		<i>Rapidez Viento</i>	<i>Rapidez Rafaga</i>	<i>Radiación Solar</i>		<i>Rapidez Viento</i>	<i>Rapidez Rafaga</i>	<i>Radiación Solar</i>		<i>Rapidez Viento</i>	<i>Rapidez Rafaga</i>	<i>Radiación Solar</i>
3												
4	CALVILLO	9.41	14.81	496.22		8.93	13.95	477.10		8.99	13.99	487.83
5	PRESA 50 ANIVERSARIO	0.00	0.00	0.00		0.00	0.00	0.00		0.00	0.00	0.00
6	SIERRA FRIA	0.00	0.00	0.00		0.00	0.00	0.00		0.00	0.00	0.00
7	BASASEACHI	6.86	12.49	471.45		6.62	12.29	#DIV/0!		7.12	13.42	417.87
8	CD CUAUHTEMOC	9.10	14.37	473.43		11.97	19.45	443.11		13.10	21.40	481.86
9	CD DELICIAS	8.01	16.14	488.94		7.49	15.24	487.67		8.39	17.15	536.25
10	CHINATU	6.59	12.00	2.00		6.22	11.22	#DIV/0!		5.36	9.27	460.15
11	CHINIPAS	6.52	12.66	455.43		7.25	13.45	438.93		7.88	14.09	466.43
12	CUMBRES DE MAJALCA	0.00	0.00	0.00		0.00	0.00	0.00		0.00	0.00	0.00
13	EL VERGEL	14.38	105.61	512.04		8.14	16.06	526.69		8.56	17.06	547.23
14	GUACHOCHI	#DIV/0!	#DIV/0!	451.84		#DIV/0!	#DIV/0!	443.06		1.60	2.23	463.67

Figure 28 Summary of the variables of interest for different years. Information of this file was used to carryout the interpolations, from which the results were later uploaded to Stat Planet Plus¹⁴

The information had to be separated by years to be uploaded to the program where the interpolations were carried out (QGIS), so there was one more file with all the weather stations, the yearly average of the desired variables and the coordinates (they were obtained at the

¹² See footnote on page 63 for a translation of each variable

¹³ See footnote on page 63 for a translation of each variable

¹⁴ Rapidez Viento=wind speed, Rapidez Ráfaga=gust speed, Radiación Solar=solar irradiance

webpage of the Mexican Weather Ministry) of each of the weather stations that we were working with, this file was used to later do the interpolations.

	A	B	C	D	E	F
1					2008	
2	STATE	STATION'S NAME	LATITUD 3	LONGITUDE 3	Wind Speed	Solar Radiation
3	AGUASCALIENTES	CALVILLO	2426746.6500000	1092713.7800000	9.41	496.22
4	AGUASCALIENTES	PRESA 50 ANIVERSARIO	2452323.6400000	1130035.0200000	0.00	0.00
5	AGUASCALIENTES	SIERRA FRIA	2437599.6500000	1138964.0900000	0.00	0.00
6	CHIHUAHUA	BASASEACHI	1891902.4300000	1805841.7300000	6.86	471.45
7	CHIHUAHUA	CD. CUAUHTEMOC	2026731.5900000	1822506.1000000	9.10	473.43
8	CHIHUAHUA	CIUDAD DELICIAS	2157058.8000000	1799636.2800000	8.01	488.94
9	CHIHUAHUA	CHINATU	2025478.4800000	1583041.5200000	6.59	2.00
10	CHIHUAHUA	CHINIPAS	1855788.4700000	1718238.3200000	6.52	455.43
11	CHIHUAHUA	CUMBRES DE MAJALCA	2062682.0700000	1866331.2800000	0.00	0.00
12	CHIHUAHUA	EL VERGEL	2064115.4500000	1608727.2100000	14.38	512.04
13	CHIHUAHUA	GUACHOCHI	1997679.8000000	1648461.1700000	0.00	451.84
14	CHIHUAHUA	JANOS	1883642.8400000	2099126.5700000	0.00	0.00
15	CHIHUAHUA	JIMENEZ	2212787.5600000	1675329.8300000	9.83	490.57
16	CHIHUAHUA	MAGUARICHI	1911318.1400000	1767299.7600000	9.14	470.22
17	CHIHUAHUA	OJINAGA	2260031.8600000	1942550.2100000	10.15	530.85
18	CHIHUAHUA	URIQUE	1915995.7600000	1696031.3600000	7.80	446.91
19	CHIHUAHUA	VILLA AHUMADA	2067126.7800000	2067290.3700000	24.12	431.18

Figure 29 Screenshot of CSV file uploaded to QGIS to carryout interpolations

6.3 Construction of Maps

To carry out the interpolations, maps were needed to delimit the areas that the program was to work with. Also maps were needed in order to be able to project the information in StatPlanet Plus and put up the renewable energy atlas. Therefore, maps of each State with their municipalities were downloaded from the Mexican National Geography and Statistics Institute. For each State 5 files were needed (.shp, .prj, .qpj, .shx, .dbf), without any of them it would have been possible to manipulate or process the information. For this project, a total of 354 municipalities were analyzed. For each of them the following procedure was carried out, this means that 1,770 additional files had to be created manually, this is next explained. The processing and manipulation of the maps was done in QGIS.

When working with QGIS, each region, state, country or municipality (from now on named as feature) has a name and an identification tag (alphanumeric). These have to be different to allow the program to differentiate one feature from another.

6.3.1 First Step – Finding and downloading maps

The maps downloaded were first maps of Mexico divided by States and divided by municipalities, these are shown next:



Figure 30 First maps obtained by INEGI. The one on the left represents the states and the one on the right the municipalities of the country. From the map of the left, the states had to be extracted and then merged with the map of the right to obtain the municipalities in the State. Around 400 maps had to be created

6.3.2 Second Step – Modifying maps to fit program needs

The name and identification tag of each of the 32 States in Mexico (31 States and Federal District) were changed. First because QGIS does not accept non ASCII (American Standard Code for Information Interchange) characters (e.g. ñ, á, ó, etc.) and second because there are municipalities that are named after the State and as mentioned earlier, the identification tags have to be unique for each feature. If this wouldn't have been done, the program wouldn't have distinguished between a municipality and a State with the same name and neither the manipulation nor the results obtained would have been accurate. The manipulation and change of the name and identification tag was carried out by modifying the file with the termination .dbf. This was done using Open Office Calc, next is shown a picture of the before and after of the States map:

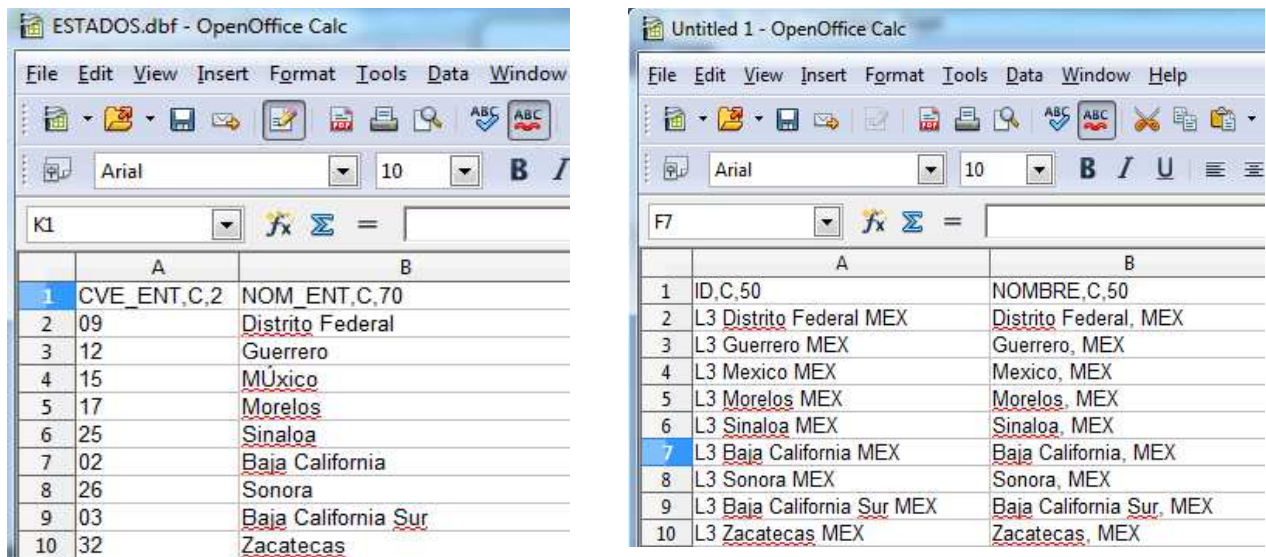


Figure 31 Screenshot of the modifications that had to be made to the original files. A lot of names had to be rewritten to fit the requirements of the program

6.3.3 Third Step – Extraction of State maps

The map of the States that compose this project had to be first extracted from the original files, this was done via QGIS→Layer→Query→Fields: Nombre→All→SQL where clause: “Nombre”=‘Zacatecas MEX’, as an example some of the results obtained are shown next:



Figure 32 Some of the States that were extracted from whole map of Mexico. As shown in Figure 31

6.3.4 Fourth Step – Adding municipalities to State maps

The map of the States had to be then intersected with the map of Mexico with the municipalities (QGIS→Vector→Geoprocessing Tools→Intersect→Input Vector Layer→Output

Vector Layer), these maps were mostly used, modified and processed, some of them are the following:



Figure 33 Some of the States with municipalities that were extracted from whole map of Mexico. As shown in Figure 31

6.3.5 Fifth Step – Division of municipalities

Next each municipality (354 in total) of each State had to be extracted and processed. This was done QGIS→Layer→Query→Fields: Nombre→All→SQL where clause: "Nombre"='Ahualulco, SLP'. Then the municipalities were divided into 100 km² features. This was done with a grid (QGIS→Vector→Research Tools→Vector Grid→Update Extents From Layer→Parameters: 10,000→Output Shapefile→Browse→OK) of 10,000m x 10,000m. Each of these features has a centroid with X and Y coordinates, these will be used later when interpolating, sampling and calculating the distances to the next transmission line or next road. The procedure is next illustrated:

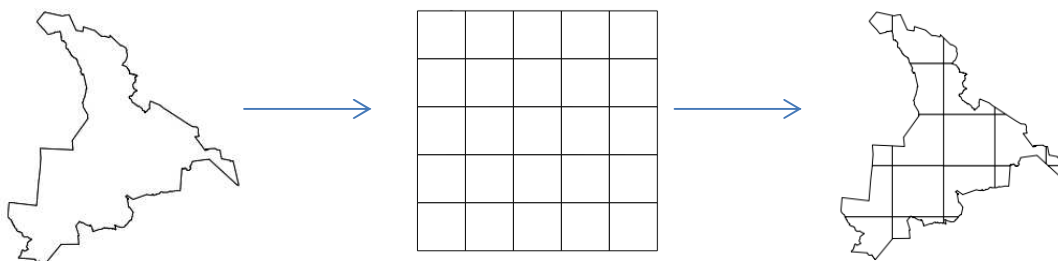


Figure 34 Procedure showing the division of the municipalities

Each of the features of the municipality has also one name and one identification tag. These had to be also modified using Open Office Calc. As already said, they have to be unique, so for this work there are 12,556 features and each one has a different name and a different identification tag.

6.3.6 Sixth Step – Attaching municipalities to State maps

The municipality divided had to be then attached to the map of the State to which that municipality belongs (QGIS→Vector→Geoprocessing Tools→Union→Input Vector Layer→Output Vector Layer), next is an example depicted:

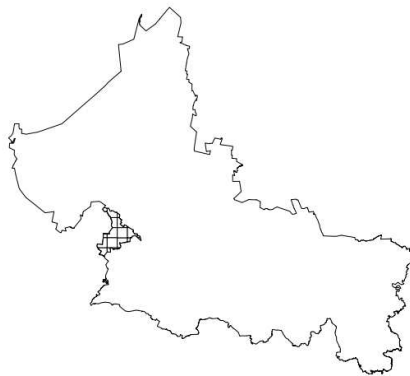


Figure 35 Representation of one divided municipality of the State of San Luis Potosi. The map of the municipality had to be attached to the map of the State This had to be done around 450 times, one time of each municipality

6.3.7 Seventh Step – Merging maps

Last the processed map had to be merged with the map of the State with all the municipalities (Vector→Data Management Tools→Merge Shapefiles To One→Select Layers in Folders→Browse→OK). The merging process is done to have two layers, in this case the one on the top is the map of the State with all the municipalities and the one on the bottom is the one of the map of the sixth step. This was done because when projecting the maps with StatPlanet the layers were needed.

6.4 Interpolations

The next phase of the project was the interpolations. QGIS is a program that has built-in two interpolation methods (Inverse Distance Weighting and Triangulated Irregular Network), both were used to carry out the interpolations for each municipality, for each feature, for each year.

The way the interpolations were carried out is described in the next paragraphs and steps.

6.4.1 First Step – Information input for QGIS

As already mentioned QGIS needs a map when carrying out the interpolations in order to define the boundaries where the mathematical methods should work in. Because of that, the map of Mexico with its states (QGIS→Browser→Load Shapefile) and also the file with all the weather stations (yearly average of the desired variables and the coordinates of the weather stations) were loaded (QGIS→Add Delimited Text Layer→Browse→Select File→Select Delimiters→Comma→X Field→Y Field→OK).



Figure 36 Depiction of the information that had to be uploaded to QGIS to carryout the interpolations. The map on the right contains all the weather stations and weather information used. This was done for each year (5)

6.4.2 Second Step - Interpolations

As already mentioned the interpolations were carried out in QGIS because it's built-in interpolation methods (QGIS→Vector Layer→"Estados" →Interpolation Attribute→Wind Speed (or Solar Irradiance) →Add→Select Interpolation Method→Set to Layer Extent→Browse Output file→OK), the result was a layer with different colors representing the different values. Two

interpolation methods were tried, the Inverse Distance Weighting and the Triangulated Irregular Network methods.

For the IDW method, a coefficient had to be given; therefore 10 different coefficients were tried (from 1 to 10). The greater the weighting coefficient, the less the effect points will have if they are far from the unknown point during the interpolation process. Due that the reach of the measurements of the weather stations is only 5 km, the coefficient # 1 was chosen for all cases, and if this would have been otherwise the results would not be anything alike to the reality. The points in the image are where the values are critical, near the weather stations.

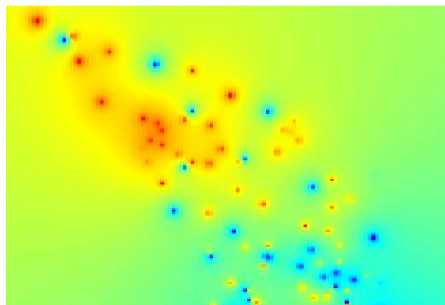


Figure 37 Image of the first obtained results. Blue points are the weather stations used and the other points the interpolated results. From this a sampling was done to each of the centroids of the ten square kilometers squares to obtain a result

6.4.3 Third Step - Sampling

Last the interpolated layer was sampled to obtain the results in each of the centroids of the features that the municipalities were divided into (QGIS→Point Sampling Tool→Layer Containing Sampling Points→Select Layer→Layers with fields/bands to get the values from: →Select Layer→Output Point Vector Layer→Browse→OK). For each interpolation method, each municipality and each year the result was a spreadsheet with all the values. A great amount of spreadsheets were organized and formatted in order to be easier to work with them, as it is shown next.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	CHIHUAHUA				2008										
2	Municipio (division)	X	Y		IDW										TIN
3					1	2	3	4	5	6	7	8	9	10	
4	Ahumada: CHIH 1	2014656.491	2105843.02		324.684	343.931	370.957	395.859	411.552	420.446	425.343	428.013	429.464	430.25	0
5	Ahumada: CHIH 2	2014656.491	2115843.02		323.694	338.438	360.822	385.945	403.834	414.982	421.666	425.616	427.934	429.289	0
6	Ahumada: CHIH 3	2014656.491	2125843.02		322.799	333.224	349.916	374.054	393.681	407.158	415.959	421.594	425.166	427.417	0
7	Ahumada: CHIH 4	2024656.491	2105843.02		329.551	371.469	407.123	422.465	428.039	430.041	430.765	431.028	431.124	431.159	0
8	Ahumada: CHIH 5	2034656.491	2105843.02		329.551	371.469	407.123	422.465	428.039	430.041	430.765	431.028	431.124	431.159	0
9	Ahumada: CHIH 6	2024656.491	2115843.02		327.895	363.6	398.971	417.724	425.672	428.925	430.253	430.798	431.022	431.115	0
10	Ahumada: CHIH 7	2034656.491	2115843.02		327.895	363.6	398.971	417.724	425.672	428.925	430.253	430.798	431.022	431.115	0
11	Ahumada: CHIH 8	2034656.491	2125843.02		326.451	356.106	389.378	411.157	421.931	426.934	429.23	430.283	430.767	430.99	0
12	Ahumada: CHIH 9	2024656.491	2125843.02		326.451	356.106	389.378	411.157	421.931	426.934	429.23	430.283	430.767	430.99	0
13	Ahumada: CHIH 10	2014656.491	2075843.02		327.56	357.405	390.982	412.402	422.729	427.399	429.486	430.419	430.837	431.025	306.795
14	Ahumada: CHIH 11	2004656.491	2075843.02		327.56	357.405	390.982	412.402	422.729	427.399	429.486	430.419	430.837	431.025	255.48

Figure 38 Screenshot of results changing the coefficients of the IDW method, 10 were used in total

6.5 Distances Calculation

One important aspect to determine if a location is suitable for the development of renewable energy projects is the distance to highways and to transmission lines. The distance from the centroid of each of the features into which the municipalities were divided to the next transmission line and next road was also calculated using MatLab.

6.5.1 First Step – Finding and downloading maps

For this purposes, maps of the transmission lines (400 and 230 kV) and highways (including roads) in Mexico were downloaded from the Mexican National Geography and Statistics Institute. Each file had 4 other auxiliary files necessary for the manipulation and processing of the information (.shp, .prj, .qpj, .shx, .dbf). Next images of those two maps are presented:

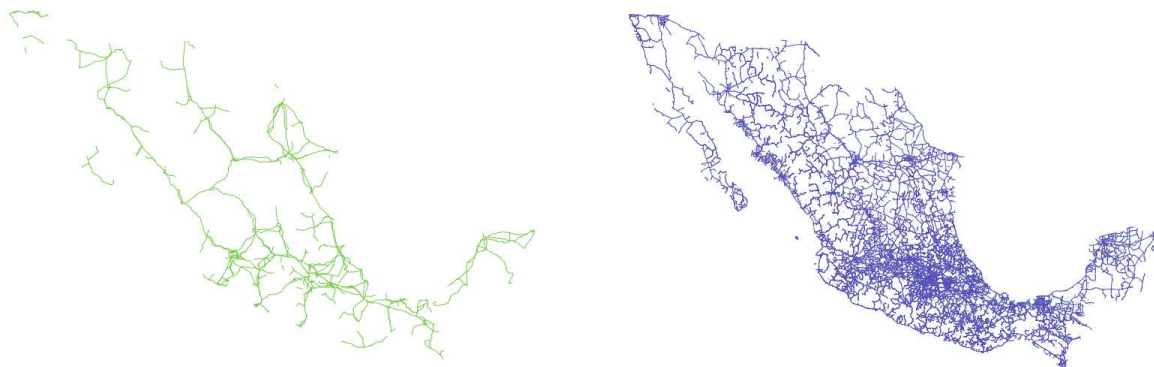


Figure 39 Maps of transmission lines and roads obtained by INEGI. On the left the transmission lines (400 and 230 kV) and on the right the highways and roads are depicted

6.5.2 Second Step – Choosing method for calculation

The distances were calculated using the Pythagoras Theorem. The Pythagorean Theorem is one of the most famous theorems of geometry. It is attributed to Pythagoras of Samos (Greece), who lived in the sixth century b.C. The theorem states that in any right triangle, the square of the hypotenuse of the triangle (the side opposite the right angle) is equal to the sum of the squares of the other two sides (Encyclopedia, 2002):

$$c^2 = a^2 + b^2.$$

Equation 42

So, supposing that the hypotenuse of the triangle was the distance from the centroid of the feature to either the next transmission line or the next road this was possible.

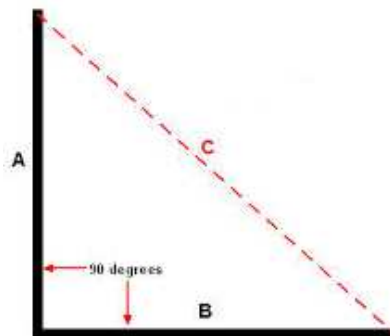


Figure 40 Triangle for representation Pythagoras Theorem used to calculate distances

The distances were calculated using the coordinate system EPSG:6372 a projection for Mexico defined by Mexican National Geography and Statistics Institute which has meters as units. For this, the maps of the transmission lines and the roads were converted from lines to points using QGIS (QGIS →Geometry Tools →Extract Nodes), the coordinates of these points were extracted (QGIS →Open Attribute Table →Copy coordinates →Paste Coordinates in Microsoft Excel) for their later use in the calculations.



Figure 41 Maps of transmission lines and roads converted to points

6.5.3 Third Step – Loading coordinates to MatLab and distances calculation

Those coordinates were then copied and pasted to a Microsoft WordPad file (in this case COORDENADAS-CARRETERAS-12.txt) and saved with the termination .txt to be able to load them into MatLab. There was also a file with the same termination with the coordinates of the centroids of the features into which all the municipalities of a State were divided. Then with the following programming code the minimum distances from the centroids (in this example the State of Coahuila) to the next transmission line or to the next road were calculated.

```
A=load('COORDENADAS-CARRETERAS-12.txt');
B=load('COORDENADAS-DIVISIONES-MUNICIPIOS-COAH.txt');
X1=A(:,1);
Y1=A(:,2);
X2=B(:,1);
Y2=B(:,2);
n=length(A);
m=length(B);

C=zeros(n,m);
D=zeros(n,m);

for W=1:m;
    for E=1:n;
        C(E,W)=(X2(W)-X1(E))^2;
        D(E,W)=(Y2(W)-Y1(E))^2;
    end
end
```

RES12=sqrt(C+D);

6.6 Maps Projection

As already mentioned, this process was carried out to put together a renewable energies atlas for northeast Mexico. This was done through the projection of the created maps / features with their information in StatPlanet. This program works by assigning a Microsoft Excel cell contained in a macro to each of the features that would be projected. Therefore the information of each of the 12,556 features contained in the 7 states had to be arranged in the cell assigned to that feature. This task was not possible to do manually so VBA programming was used. The process is next described.

6.6.1 First Step – Merging the information

A Microsoft Excel file was created where the name of the feature, the coordinates of its centroid, the chosen results from the interpolation methods (IDW with coefficient #1) of the period from 2008 - 2012 and the results of the calculations of the distances were concentrated. Using programming language in Microsoft VBA and the following code a macro was programmed to merge the 5 pieces (name of feature, X-coordinate, Y-coordinate, distance to next transmission line and distance to next road) of information into one Excel cell.

	A	B	C	D	E	F	G
1	MERGE IT						
2	Name	X	Y	Distance to next highway [m]	Distance to next transmission line [m]	Solar irradiance [W/m ²]	
3	Abasolo, NL 1	2659450.01800	1544531.23600	190.36	1270.84	231.54	Name: Abasolo, NL 1 X: 2659450.018 Y: 1544531.236 Distance to next highway [m]: 190.36 Distance to next transmission line [m]: 1270.84
4	Abasolo, NL 2	2659450.01800	1534531.23600	1700.34	2413.66	286.78	Name: Abasolo, NL 2 X: 2659450.018 Y: 1534531.236 Distance to next highway [m]: 1700.34 Distance to next transmission line [m]: 2413.66
5							
6	Agualeguas, NL 1	2748367.42600	1596298.38970	3108.48	16846.51	307.62	Name: Agualeguas, NL 1 X: 2748367.426 Y: 1596298.3897 Distance to next highway [m]: 3108.48 Distance to next transmission line [m]: 16846.51
7	Agualeguas, NL 2	2748367.42600	1576298.38970	1728	1779.15	308.41	Name: Agualeguas, NL 2 X: 2748367.426 Y: 1576298.3897 Distance to next highway [m]: 1728 Distance to next transmission line [m]: 1779.15

Figure 42 Screenshot of Excel VBA file used to arrange information. This information was uploaded to Stat Planet Plus

```

Sub MergeContents()
    Dim i As Integer
    For i = 2 To ActiveSheet.Cells(ActiveSheet.Rows.Count, 1).End(xlUp).Row Step 1
        ActiveSheet.Cells(i, 7).Value = _
            ActiveSheet.Cells(1, 1).Value & ": " & ActiveSheet.Cells(i, 1).Value & Chr(10) & _
            ActiveSheet.Cells(1, 2).Value & ": " & ActiveSheet.Cells(i, 2).Value & Chr(10) & _
            ActiveSheet.Cells(1, 3).Value & ": " & ActiveSheet.Cells(i, 3).Value & Chr(10) & _
            ActiveSheet.Cells(1, 4).Value & ": " & ActiveSheet.Cells(i, 4).Value & Chr(10) & _
            ActiveSheet.Cells(1, 5).Value & ": " & ActiveSheet.Cells(i, 5).Value
    Next i
End Sub

```

6.6.2 Second Step – Assigning merged information to corresponding cell

The information had to be then inserted to the cells assigned to each different feature. For this the function =VLOOKUP(lookup_value, table_array, col_index_num, range_lookup]) of Microsoft Excel was used. So, once the information was assigned, it looked as follows, the next image is an image of the macro that was then loaded to StatPlanet to project the information and conclude the project. In the first row, the name of the feature is to be seen and in the lower rows the information for the analyzed period is attached.

	AG	AH	AI	AJ
1	Ahumada, CHIH 117	Ahumada, CHIH 118	Ahumada, CHIH 119	Ahumada, CHIH 12
2	Name: Ahumada, CHIH 117 X: 2064656.491 Y: 2045843.02 Distance to next highway [m]: 214.87	Name: Ahumada, CHIH 118 X: 2054656.491 Y: 2025843.02 Distance to next highway [m]: 13316.67	Name: Ahumada, CHIH 119 X: 2044656.491 Y: 2025843.02 Distance to next highway [m]: 7789.89	Name: Ahumada, CHIH 12 X: 2014656.491 Y: 2085843.02 Distance to next highway [m]: 34239.53
3	Distance to next transmission line [m]: 1460.44	Distance to next transmission line [m]: 8874.62	Distance to next transmission line [m]: 10149.46	Distance to next transmission line [m]: 50401.94
4		376.92	357.09	357.09
5	Name: Ahumada, CHIH 117 X: 2064656.491 Y: 2045843.02 Distance to next highway [m]: 214.87 Distance to next transmission line [m]: 1460.44	Name: Ahumada, CHIH 118 X: 2054656.491 Y: 2025843.02 Distance to next highway [m]: 13316.67 Distance to next transmission line [m]: 8874.62	Name: Ahumada, CHIH 119 X: 2044656.491 Y: 2025843.02 Distance to next highway [m]: 7789.89 Distance to next transmission line [m]: 10149.46	Name: Ahumada, CHIH 12 X: 2014656.491 Y: 2085843.02 Distance to next highway [m]: 34239.53 Distance to next transmission line [m]: 50401.94

Figure 43 Screenshot of arranged information ready for programming into StatPlanet Plus

Chapter 7: Results and Discussion

Out of the two methods (Triangulated Irregular Network and Inverse Distance Weighting) it was seen that Inverse Distance Weighting represented best the reality of the solar irradiance and wind speed conditions of northeast Mexico because this method does not limit the working area to the area surrounded by the weather stations, whereas the TIN does. This means that there were a lot of features that were left without any result because they are located outside of the scope area of this method. Even though that ten different coefficients were tried to do the interpolations, a coefficient of one was chosen due that the sensors of the weather stations only have a range of 5 kilometers for the measurements (The greater the weighting coefficient, the less the effect points will have if they are far from the unknown point during the interpolation process). If another coefficient would have been taken the results would have been considerably different. It was seen that when the coefficient was greater, the results for the features more far away from the weather stations changes dramatically to even decimal values. Logically there cannot be regions in the northeast of Mexico with irradiance values of less than one.

It is important to notice that the IDW interpolation method has some disadvantages: The quality of the interpolation result can decrease, if the distribution of sample data points is uneven. Furthermore, maximum and minimum values in the interpolated surface can only occur at sample data points. This often results in small peaks and pits around the sample data points.

The main disadvantage of the TIN interpolation is that the surfaces are not smooth and may give a jagged appearance. This is caused by discontinuous slopes at the triangle edges and sample data points. In addition, triangulation is generally not suitable for extrapolation beyond the area with collected sample data points. It must also be taken into consideration that the altitude was not considered for the interpolations and that the analyzed region very mountainous condition presents. It is thought that this has also affected the results considerably.

It is important to remember that there is no single interpolation method that can be applied to all situations. Some are more exact and useful than others but take longer to calculate. They all have advantages and disadvantages. In practice, selection of a particular interpolation method should depend upon the sample data, the type of surfaces to be generated and tolerance of estimation errors. Generally, a three step procedure is recommended:

1. Evaluate the sample data. To get an idea on how data are distributed in the area, as this may provide hints on which interpolation method to use,
2. An interpolation method which is most suitable to both the sample data and the study objectives has to be applied. If in doubt, several methods have to be tried and,
3. Comparison of the results and find the best result and the most suitable method. This may look like a time consuming process at the beginning. However, as experience and knowledge of different interpolation methods are gained, the time required for generating the most suitable surface will be greatly reduced (GIS for educators, 2013).

The distances were calculated using the Theorem of Pythagoras and the curvature of the earth was not considered so there may also be small variations with the real results. It must also be considered that the values were obtained out of maps and numerical methods so an error margin has to be considered. The distances were calculated using the coordinate system EPSG:6372, a projection for Mexico defined by Mexican National Geography and Statistics Institute which has meters as units. Other errors would also have to be taken into consideration as the lack of experience of the author, rounded and truncated figures, processing and procedure errors, etc.

Further research is recommended to test whether other environmental factors might allow variations in the weather variables analyzed. As there are some additional geostatistical methods that were not evaluated in this study, it may be necessary to explore these methods to determine if they could generate a better representation of wind speed and solar irradiance across the study area.

Bibliography

- Burrough, & McDonell. (1998, February 5). Principles of Geographical Information Systems (Spatial Information Systems). USA: Oxford University Press.
- C.E., P., & R.L., B. (1986). Disjunctive kriging, universal kriging, or no kriging: Small sample results with simulated fields. In *Mathematical Geology* (pp. 287-305).
- CIA. (2013, December 5). *Central Intelligence Agency Factbook*. Retrieved December 23, 2013, from <https://www.cia.gov/library/publications/the-world-factbook/geos/mx.html>
- Collins , F., & Bolstadt, P. (1996). A comparison of spatial interpolation techniques in temperature estimation. In *National Center for Geographic Information Analysis* . Santa Monica.
- Encyclopedia. (2002). *Encyclopedia.com*. Retrieved December 10, 2013, from http://www.encyclopedia.com/topic/Pythagorean_theorem.aspx
- Encyclopedia, T. C. (2013). *Nuevo Leon*. Retrieved 12 23, 2013, from Columbia University Press.
- Encyclopedia, T. C. (2013). *Nuevo Leon*. Retrieved 12 23, 2013, from The Columbia University: <http://www.encyclopedia.com/topic/Coahuila.aspx>
- Englund, E., Weber, D., & Leviant , N. (1992). The effects of sampling design parameters on block section. In *Mathematical Geology* (pp. 329-343).
- GIS for educators. (2013). *Geographical Information Systems for educators*. Retrieved December 5, 2013, from <http://www.esri.com/industries/education>
- Goovaerts, P. (2000). Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. In *Journal of Hydrology* (pp. 113–129).
- Hengl, T. (2007). A Practical Guide to Geostatistical Mapping of Environmental Variables. Luxemburg: Office for Official Publication of the European Communities.
- INEGI. (2010, 6 December). *Instituto Nacional de Estadística y Geografía*. Retrieved December 23, 2013, from <http://www.inegi.org.mx/inegi/contenidos/espanol/prensa/comunicados/rpcpyv10.asp>
- Laslett, G. (1994). Kriging and splines: an empirical comparison of their predictive performance in some applications. In *Journal of the American Statistical Association* (pp. 391-400).
- Li, J., & Heap, A. (2008). A review of spatial interpolation methods for environmental scientists. Camberra, Australia: Geoscience Australia.
- Luo, W., Taylor, C., & Parker, R. (2007, August 2). A comparison of spatial interpolation methods to estimate continuous wind speed surfaces using irregularly distributed data from England and Wales. San Hutton, York, UK: John Wiley & Sons.
- Marchant, P., & Lark, R. (2006). Adaptive sampling and reconnaissance surveys for geostatistical mapping of the soil. In *European Journal of Soil Science* (pp. 831-845).
- Martinez-Cob, A. (1996). Multivariate geostatistical analysis of evapotranspiration and precipitation in mountainous terrain. In *Journal of Hydrology* (pp. 19-35).

- Mexican Embassy in the United States. (2012, December 3). *Secretaría de Relaciones Exteriores*. Retrieved December 23, 2013, from <http://embamex.sre.gob.mx/eua/index.php/en/about-mexico>
- Mexican Energy Ministry. (2013, January 1). *Renewable Energy Outlook 2012-2026*. Mexico City, Mexico: Mexican Energy Ministry.
- Mexican Weather Ministry. (2013). *Automatic meteorological station description*. Retrieved December 23, 2013, from <http://smn.cna.gob.mx/emas/>
- Ministry of Foreign Affairs of Japan. (2012, May 7). Ministry of Foreign Affairs of Japan. Retrieved December 23, 2013, from http://en.wikipedia.org/wiki/Mexico#cite_note-16
- Nadler, I., & Wein, R. (1998). Spatial interpolation of climatic Normals: test of a new method in the Canadian boreal forest. In *Agricultural and Forest Meteorology* (pp. 211-225).
- Quaschnig, V. (2005). *Understanding renewable energies*. Carl Hanser Verlag GmbH & Co KG.
- Solar Electricity Handbook. (2013). *Solar Electricity Handbook*. Retrieved February 24, 2014, from Solar Electricity Handbook: <http://solarelectricityhandbook.com/solar-irradiance.html>
- The Columbia Encyclopedia. (2013). Retrieved December 23, 2013, from Columbia University Press: <http://www.encyclopedia.com/>
- Trade & Logistics Innovation Center. (2011). *Mexican National highway and road system*. Georgia Tech and ITESM.
- W. Luo, M. C. Taylor and S. R. Parker. (2007, 8 2). A comparison of spatial interpolation methods to estimate continuous wind speed surfaces using irregularly distributed data from England and Wales. John Wiley & Sons, Ltd.
- Webster, M. (2007). *Merriam-Webster's Geographical Dictionary*. Springfield, Massachusetts: p. 733.
- Webster, R., & Oliver, R. (1992). Sample adequately to estimate variograms of soil properties. In *Journal of Soil Science* (pp. 177-192).
- Willmot, C. (1985). On the evaluation of model performance in physical geography. In W. CJ, *Spatial Statistics and Models* (pp. 443-460). Dordrecht, Holland: Reidel.
- World Bank. (2011, March 5). *World Bank Data*. Retrieved December 23, 2013, from http://data.worldbank.org/about/country-classifications/country-and-lending-groups#Upper_middle_income
- World Tourism Organization. (2011). *UNWTO Tourism Highlights*. Retrieved December 23, 2013, from United Nations World Tourism Organization: <http://wayback.archive.org/web/20120113021355/http://mkt.unwto.org/sites/all/files/docpdf/unwt ohighlights11enlr.pdf>

Annexes

For more information about the municipalities, their location and names please refer to:

Tamaulipas

http://cuentame.inegi.gob.mx/mapas/pdf/entidades/div_municipal/tampsmpio.pdf

Nuevo Leon

http://cuentame.inegi.gob.mx/mapas/pdf/entidades/div_municipal/nvoleonmpios.pdf

Coahuila

http://cuentame.inegi.gob.mx/mapas/pdf/entidades/div_municipal/coahmpios.pdf

Chihuahua

http://cuentame.inegi.gob.mx/mapas/pdf/entidades/div_municipal/chihtmlmpios.pdf

Durango

http://cuentame.inegi.gob.mx/mapas/pdf/entidades/div_municipal/dgompio.pdf

Zacatecas

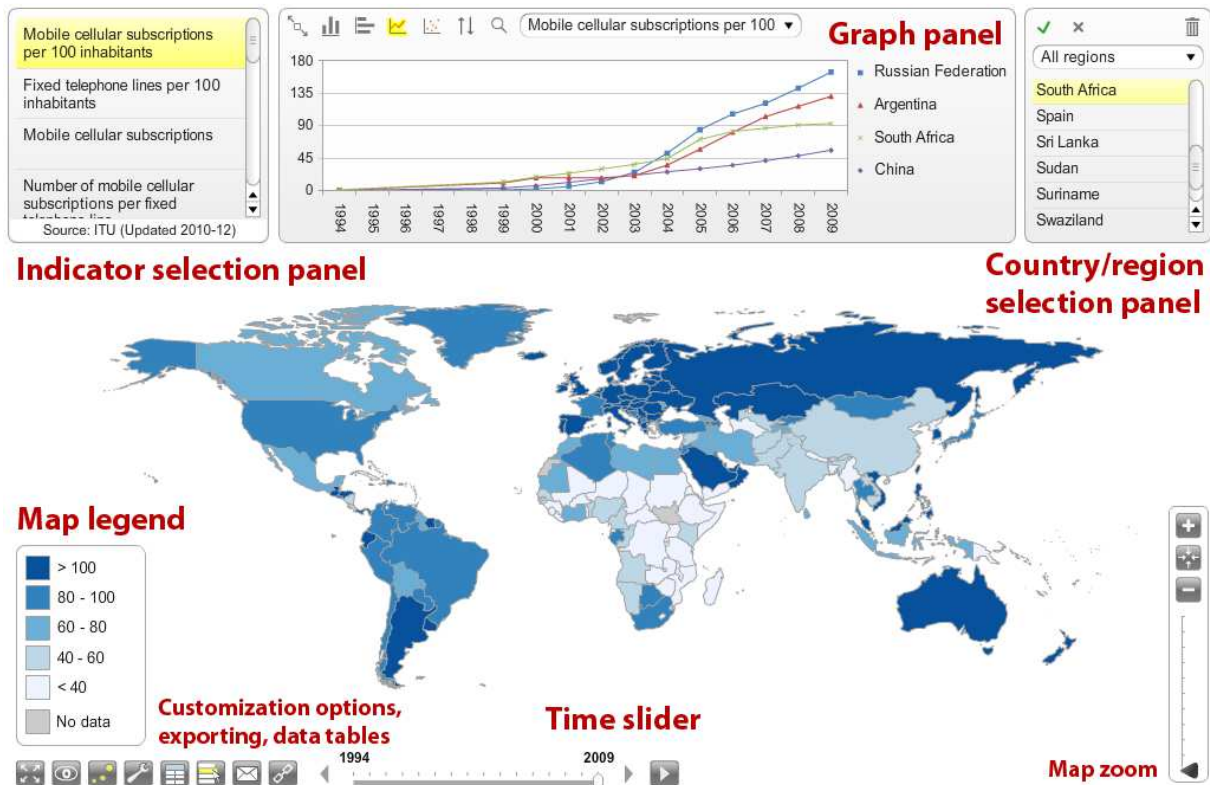
http://cuentame.inegi.gob.mx/mapas/pdf/entidades/div_municipal/zacatecasmpios.pdf

San Luis Potosi

http://cuentame.inegi.gob.mx/mapas/pdf/entidades/div_municipal/slmpios.pdf

Instructions to Use Tool

Features and Options



Thematic map

Choropleth Map



This is the main thematic map type in StatPlanet. The map legend shows which map colors are associated with each data range (for example, higher values may be shaded in increasingly darker colors). The map and map legend colors, ranges and values can be customized.

Proportional Symbol Map



A proportional symbol map scales symbols (usually circles) according to the indicator being mapped. Each symbol can represent a country or other map area. Proportional symbol maps are most suitable when there is a large range of

[D]

values and measurements are continuous like wind speed velocities and solar irradiance values

In StatPlanet the symbol map is overlaid on top of the choropleth map (see above). This means that two data sets can be shown on the same map – one for the choropleth map and one for the symbol map.

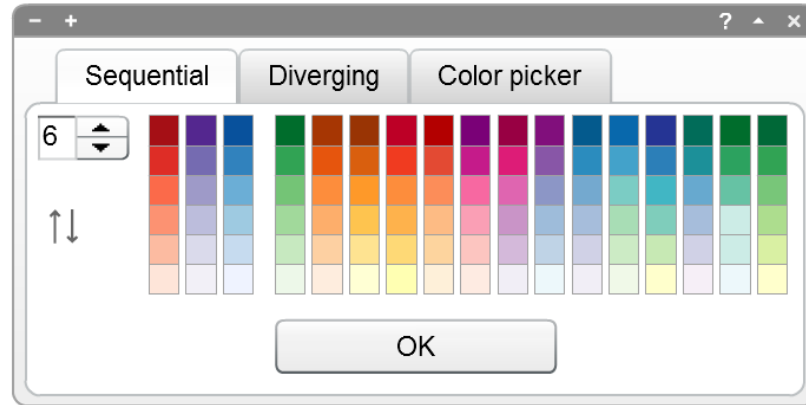
To show the symbol map, click the symbol icon in the bottom-left corner of the screen. If you have bookmarked an indicator, the symbol map represents the data for the bookmarked indicator, whereas the choropleth map represents the data for the selected indicator. (If the bookmarked indicator is currently selected, both the symbol map and the choropleth map represent the bookmarked indicator).

Select Country or Other Map Area Using the Map

- *Mouse over a country or map area:* Moving the mouse over a country or other map area brings up a popup containing information about that particular country or map area for the selected indicator (as well as the bookmarked indicator, if there is one).
- *Click on a country or map area:* You can select a country or other map area by clicking on it in the map. See also the Selection panel for additional selection options.

Map Legend

Map colors: Clicking on any of the colors in the legend will bring up a color selection panel. In this panel you can change both the colors (either Sequential or Diverging color schemes), as well as the number of color classes (between 3 and 9). The color schemes are from the Color Brewer website (<http://colorbrewer2.org/>**Fehler! Hyperlink-Referenz ungültig.**



Data range: To adjust the data range of the map legend, click on the top or bottom value. Use the popup to increment or decrement the value, or enter a whole new value. The intermediate values will be adjusted automatically.

Map Zoom



The map zoom controls are normally hidden from view. Move the mouse towards the bottom-right of the screen to make them appear.

Zoom: You can zoom in and out of the map using the 'zoom in' and 'zoom out' buttons, or by dragging the zoom slider up or down. If your mouse has a scroll wheel, you can also use this to zoom in and out.

Moving the map: click and drag the map with the mouse to move it to a new position.

Restore map position: the button shown on the left restores the map to the original coordinates for the selected region.

Indicators Panel

Select Category



Use the drop-down menu in the top of the Indicators panel to select a new category. When switching categories, StatPlanet will remember which year was selected and check if data exists for this year in the new category. If there is no data for this year, it will select the year closest to the previously selected year for which data is available in the new category.

Indicator Bar



The bars are scaled in proportion to the maximum value of that particular indicator for all countries in the selected region. For example, if the value for country X is 20 and the maximum value for all countries is 100, the bar will be scaled at 20%.

Graph panel

In StatPlanet, click the 'graph' button in the bottom-left corner of the screen to open or close the Graph panel.

Bar Chart / Column Chart



The bar chart and column chart buttons are located in the top left corner of the Graph panel.



Use the "sort" button to sort the graph from lowest to highest, highest to lowest, highest to lowest starting in the middle, or alphabetically.

Time Series



The time series button is located in the top left corner of the Graph panel. Use the Selection panel to add more features (depending on your data set) to the time series graph. Click on it again if you wish to remove it. Features can also be selected directly from the map. Use the "sort" button to sort the time series labels.

Vertical Bubble Chart



This button enables the display of country data as with the column chart, but with the values on the y-axis marked as bubbles rather than the top of a column. This display type has the advantage that it allows a second indicator to be visualized in the form of the bubble size. The bubble size follows the formula: $\text{value} / \text{maximum value}$.

Scatter Plot (Bubble Chart)




The scatter plot button is located in the top left corner of the Graph panel. Clicking on the button will automatically use the selected indicator as the x-axis

variable. You need to select a second indicator as the y-axis variable. The x-axis and y-axis variables can be selected in the Graph panel (see below), or in the Indicators panel (see Bookmark indicator above).

Press the 'play' button to see an animation of changes over time, with each bubble (point) moving to the corresponding x and y positions (depending on whether data is available for each time interval). If 'Show trails' is selected (next to the play button), each bubble will leave a trail to mark previous positions over time.

Clicking on a scatter plot bubble will display the associated label. They can also be repositioned through 'drag and drop', by right-clicking and selecting 'move text labels or map points'.

A third indicator can be visualized through the bubble size parameter. This indicator can be selected through the drop-down above the scatter plot. The bubble size follows the formula: $\text{value} / \text{maximum value}$.

The scatter plot has an 'options' icon  which can be used to show or hide the trend line. Move the mouse over the trend line to see the slope and trend line equation.

Select X-Axis / Y-Axis Indicator

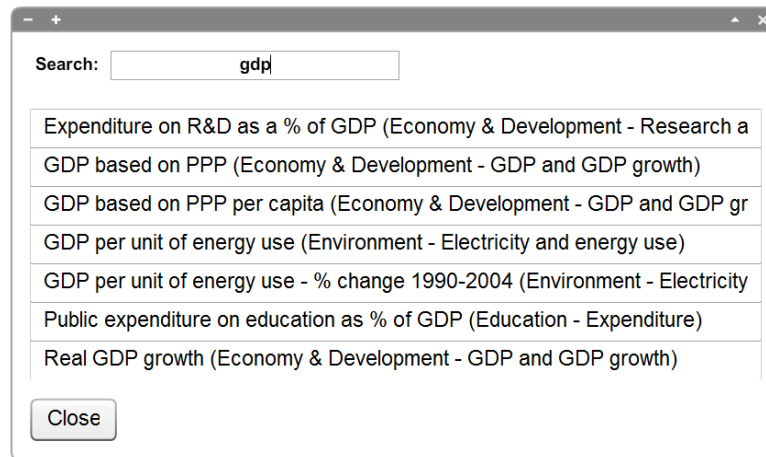


Click on the x-axis or y-axis label, and then select an indicator from the drop-down menu. Use the drop-down menu in the top of the Graph panel to change the scale of the 'bubbles' according to a selected indicator.

Search



The search function brings up a popup which enables you to filter the list of indicators to find the one you need.



Adjust Graph Size



To change the graph size, move the mouse to the sides or corners of the Graph panel until you see the cursor change to look like the one shown on the left. Click and hold down the left mouse button, then drag the panel to the size you want. Release the mouse button once you have the right size.

Adjust Graph Scale

The graph scale can be adjusted by clicking on the top or bottom graph values. The value can then be edited in the popup window.

Options Panel



Click the 'options' button in the bottom-left corner of the screen to open the options panel.

Map:

- Map colors: map background, map borders, map text color, map text outline color, etc.,
- Map text size,
- Map symbol size (proportional symbol map symbols),
- Map legend - estimate best value distribution: adjust the values so that there is a more equal distribution of countries for each color class. This will usually result in a map with a better distribution of colors. If this is switched off, the value range for each

color class will be set at equal intervals based on the highest and lowest value in the data range,

- Map legend - show maximum & minimum values and,
- Map legend colors.

Graph/Chart:

- Graph colors: background, bar & scatter points, scatter point borders,
- Graph text size,
- Transparency level of graph area (bars / bubbles),
- Size of bubbles (scatter plot graph) and,
- Bullet graph.

General options:

- Animation duration;
- Decimal places shown.

Adjust the map/graph scale: StatPlanet automatically adjusts the map and graph scale to suit the data set. However, in some cases you may wish to keep it fixed, for example if you made some changes to it yourself. You can set (or prevent) the automatic updating of the map/graph scale on (i) changing indicator, (ii) changing region, (iii) changing year. The default setting is for StatPlanet to adjust the scale when changing indicator or region, but not year.

Data-Table Panel



Click the 'table' button in the bottom-left corner of the screen to get a data table of the selected indicator. If an indicator has been bookmarked, the data for both the bookmarked and selected indicator will be displayed

Selection Panel

Countries, map areas can be selected in various ways. An efficient way of finding and selecting a country or map area is through the selection panel, as explained below. However, a country or map area can also be selected by clicking on it in the map, in the data table panel, or in the graph panel. In each case it will be highlighted in all of these StatPlanet components.

Click on a country/map area/variable in the list to select it. Clicking on a selected item will deselect it. You can press the first one or two letter of an item in the list to quickly jump to that item. For example, if it contains a list of countries, press 'b' to quickly jump to countries starting with the letter 'b'.



If the Selection panel is hidden, click the 'select' button in the bottom-left corner of the screen to make it appear.

Select Regions



Use the drop-down menu in the top of the indicators panel to select and zoom into a different region, such as 'Africa' or 'Europe'. It is also possible to select countries to define a custom region (see Selection panel).

Select Button



Press the Select button to reduce the list to your selection. Any items which are not selected will be removed from view.

Deselect All Button



Press the Deselect All button to remove all selections. Any items which are not selected will be removed from view.

Refresh Button



The Refresh button appears once you have created a custom region or group of items. It can be used to return to the original list.

Remove Button



To remove items from the list, select those you wish to remove and press the Remove button

Story Panel

The Story Panel can optionally be used to display 'stories' or descriptive text for each indicator. The text can be formatted using HTML, with support for image embedding and links to web-pages or documents.

Time Slider



Use the slider or click on the arrow buttons to change the year. Click on the play button to show changes over time as an animation, starting from the beginning. The animation speed can be set in the Options panel.

Interface Options

View Panel



Move the mouse over the 'View' button in the bottom-left of the screen to see various Options for showing or hiding map & graph elements, and other components.

- Show or hide country or map area names on the map or graph (country names can be shown in full, in abbreviated form or as [ISO3 codes](#)**Fehler! Hyperlink-Referenz ungültig.**)
- Show or hide country or map area statistics on the map,
- Show or hide the map popup and its components - bar chart, indicator and statistic and,
- Show or hide various panels.

Shrink / Enlarge



The buttons in the top-left corner of the panel can be used to shrink or enlarge them. This may be useful for space management when you have several panels open at the same time, or to focus in on certain areas.

Help / Minimize / Close



The buttons in the top-right corner of the screen are the help, minimize and close buttons.

Adjust Graph Panel Size



Drag and drop the graph panel borders to adjust the size.

Full Screen



Click on this button in the bottom-left of the screen to either switch to full screen mode, or go back to normal panel mode.

Changing Category

When switching to another category the program will remember which year was selected and check if data exists for this year in the new category. If there is no data for this year, it will select the year closest to the previously selected year for which data is available in the new category. To change the position and size of the map, move the mouse to the bottom-right of the screen to see the map zoom controls: