The Solar Energy Analysis Project

May 2024

Introduction:

The aim of this Solar Energy Analysis Project is to decide the best solar systems to be used in five different areas with different sizes. These areas are:

- Site 1: $4km \times 4km$
- Site 2: $3km \times 3km$
- Site 3: $6km \times 6km$
- Site 4: $5km \times 5km$
- Site 5:$4km \times 4km$

In order to evaluate the potential and effectiveness of each solar system for each area, different weather, environmental and solar energy factors were measured and calculated across three years from January 2017 to December 2019. The most important of these factors are:

- Temperature (C^o): The background temperature at the time of measurement.
- Clearsky DHI (W/m^2): Diffuse Horizontal Irradiance under clear sky conditions.
- Clearsky DNI (W/m^2): Direct Normal Irradiance under clear sky conditions.
- Clearsky GHI (W/m^2): Global Horizontal Irradiance under clear sky conditions.
- DHI (W/m^2): Diffuse Horizontal Irradiance.
- DNI (W/m^2): Direct Normal Irradiance.
- GHI (W/m^2): Global Horizontal Irradiance.
- Relative Humidity (%): The amount of moisture in the air relative to the maximum amount of moisture the air can hold at that temperature.

- Solar Zenith Angle (degrees): The angle between the sun's rays and the vertical direction.
- Surface Albedo: The fraction of solar radiation that is reflected by the surface.
- Wind Speed (m/s): The speed of the wind.

In order to calculate the power and energy produced by each solar system, we used the equations below:

$$Power = Efficiency \times GHI \times Area\ of\ system$$

$$Energy = Power \times time$$

Based on these equations, GHI is directly proportional to the power produced. As a result, we compared each factor to the GHI, and measured the Pearson correlation coefficient for each relation. Using the correlation coefficients, we were able to deduce the effect of each factor on the effectiveness of the solar system used in each area. This also helped us in our solar system selection process.

Solar Panel Analysis:

Photovoltaic Systems:

Factors	Monocrystalline	Polycrystalline	Thin-Film
Efficiency	Highest (17-24%)(20%)	Balanced (15-18%)(16.5%)	Lower (6-13%)(8.5%)
Applications	Limited space (maximizes power generation)	N/A	Non-traditional surfaces & building integration (BIPV)
Effect of Weather: Cloud Cover	Reduces efficiency (10-25% of clear day)	More affected (lower efficiency in diffuse light)	Performs better in low-light conditions
Effect of Weather: Fog	Reduces efficiency (scatters light, adds moisture)	Similar impact (scatters light, reduces performance)	Slightly better efficiency in diffuse light (proper coatings)
Effect of Temperature	Decreases efficiency (high temperatures)	More affected by high temperatures	Generally more tolerant (better suited for hot climates)
Environmental Impact	Higher albedo surfaces improve performance	Higher albedo surfaces improve performance	Higher albedo surfaces improve performance
Drawbacks	Higher cost	N/A	Lower efficiency
Average Cost	$173.70/m²	$130.20/m²	$75/m²
Maintenance	Minimal (clean, inspect)	Minimal (clean, inspect)	Minimal (clean more often, inspect)

Concentrated Solar Power Systems:

Factors/CSP Technology	Parabolic Troughs	Solar Power Towers	Fresnel Reflectors
Efficiency	20-25% (w/o storage), 40-53% (6h storage)	40-45% (6-7.5h storage), 65-80% (12-15h storage)	Optical Efficiency: 60% to 75% Overall System Efficiency: 10% to 20%
Applications	Large-scale, high direct sunlight	High efficiency, large land area, high DNI	Cost-effective, lower efficiency
Effect of Weather	Limited to hot, dry regions	Limited to hot, dry regions	Limited to hot, dry regions
Effect of Temperature	Requires high temperatures (achieved by concentration)	Requires high temperatures (achieved by concentration)	Requires high temperatures (achieved by concentration)
Environmental Impact	Low, virtually zero fuel cost, no greenhouse gas emissions	Low, virtually zero fuel cost, no greenhouse gas emissions	Low, virtually zero fuel cost, no greenhouse gas emissions
Drawbacks	Lower efficiency (w/o storage)	Large land area, high DNI	Less efficient
Maintenance Cost	High (USD 0.02-0.035/kWh)	High (USD 0.02-0.035/kWh)	High (USD 0.02-0.035/kWh)
Cost per Panel	\$100 - \$300 per m² (\$200)	\$150 - \$500 per m² (\$325)	\$75 - \$250 per m² (\$162.5)

Site 1:

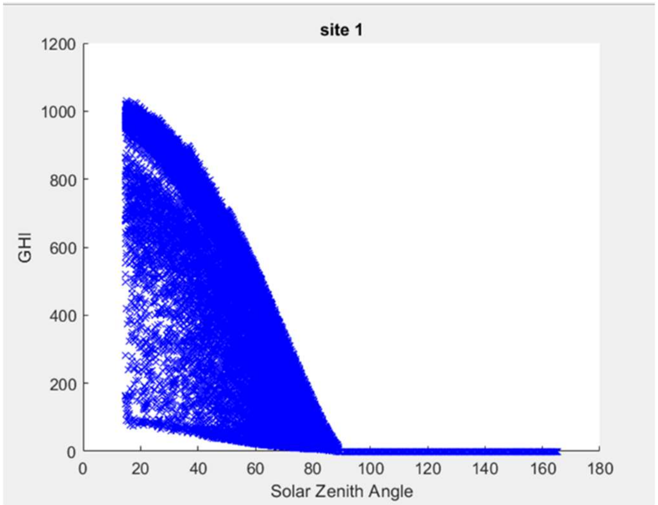
Data Cleaning:

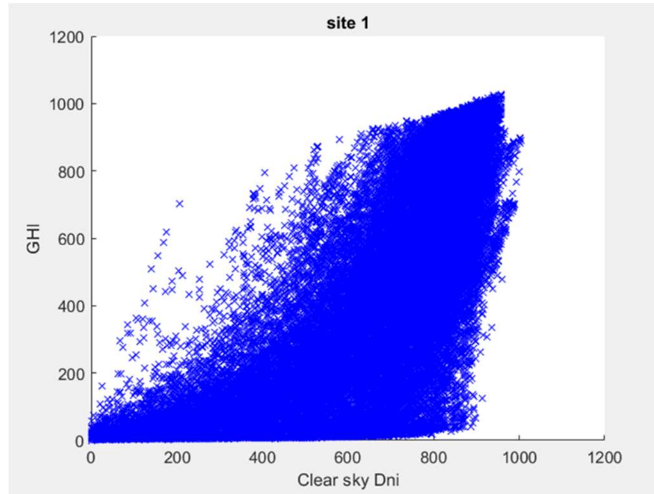
In the datasheet, there were two extra empty columns named “Unnamed: 18”. Before analyzing the data, we made sure to remove them using the removevars() function in Matlab.

We checked if there were any null values present in the columns. We couldn’t use these values in our data analysis, so we replaced them with predicted linear fit values using the fillmissing() function.

Data Analysis:

After comparing the GHI to each and every factor in our data, the two largest correlation coefficients we found were between GHI and Solar Zenith Angle, and GHI and Clearsky DNI. The graphs for these two relations are shown below.





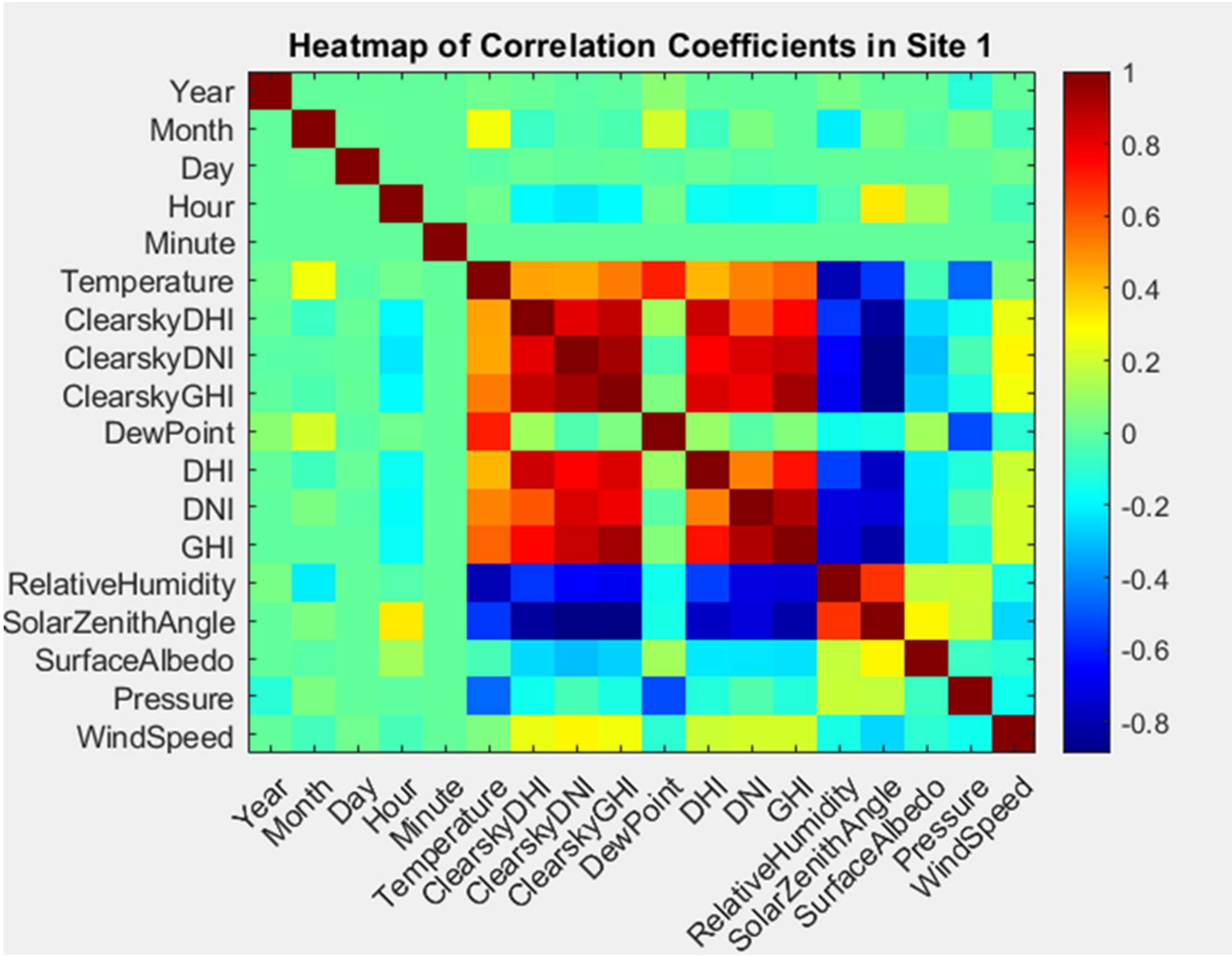
The correlation coefficient between the GHI and the Solar Zenith angle is equal to $r = 0.8163$. This implies an almost exact inverse proportionality between both values. However, looking at the equation for calculating the GHI, we get a different perspective.

$$GHI = DHI + DNI \times \cos(Solar\ Zenith\ Angle)$$

Based on this equation, the GHI should be proportional to the cosine of the angle, which exactly matches the graph we got.

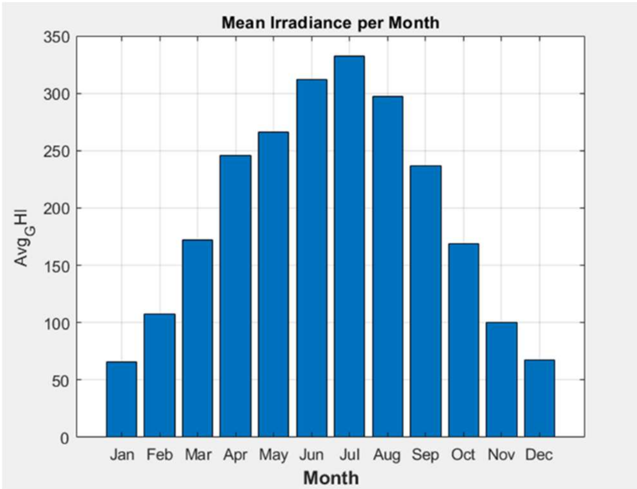
Moving on to the Clearsky DNI, we find a direct proportionality with a correlation coefficient of $r = 0.8629$. This also matches the equation we have for the GHI, as it is directly proportional to the DNI and therefore directly proportional to the Clearsky DNI as well.

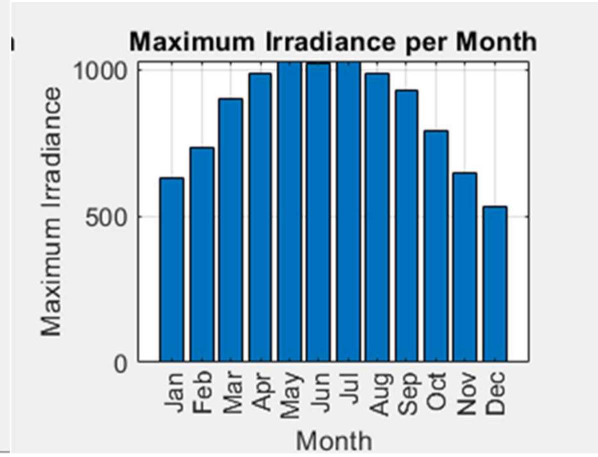
A heatmap was also constructed to represent the correlations between factors clearly.



We see that factors such as Relative Humidity and the Solar Zenith Angle are inversely proportional to the DHI, DNI and GHI values. Moreover, factors such as temperature appear to be directly proportional to the DHI, DNI and GHI.

Since the irradiance varied over time, we calculated mean values for the GHI for every month of the year across the three years. We also calculated the maximum irradiance for every month.





Based on the graphs shown, the GHI varies intensely for each month of the year. It is at its highest from April till September and has a significant decrease from October till March. The values generally vary between 50 and $350W/m^2$. This large range in values shows inconsistency in the solar power in Site 1, which will be important in the solar system selection process.

In terms of weather, the site does not appear to host any rain or snow. In terms of the clear sky, it was apparent that the GHI maintains above 65% of the clearsky GHI's value, which shows that the site has an almost clear to completely clear sky most of the time. When we analyzed other weather factors such as dew point, humidity and pressure, we found a low correlation between them and the GHI, meaning these factors had minimal effect on the power and energy.

In terms of temperature, no correlation was found between it and the GHI, leading to it having practically no effect. The same goes with the Surface Albedo and the Wind Speed, where they had no effect on GHI.

Solar System Selection:

Based on our data analysis, we concluded that Site 1 has low temperatures, little cloud cover since most of the clearsky GHI is preserved, and a high light level. As a result, the two best options for a solar system would be monocrystalline and polycrystalline. Since the weather has very little effect on the Solar Power, we can use **Polycrystalline** for cheaper energy and still maintain effectiveness and high efficiency.

We will use 50% of the area of the site, since we can exploit it for high amounts of energy due to its clear skies. We will also need to leave space for the engineers to work and for the other buildings and utilities. The solar system area accounts for $8km^2$ of the site, which will cost us:

$$130.20 \times \frac{8 \times 10^6}{1} = \$1.042\ billion$$

Energy Production:

Using the power and energy equations, we can calculate the different energy production rates.

	Average GHI (W/m^2)	Power (GW)	Energy (GJ)
January	65.59	8.66E-02	231892.7
February	107.13	1.41E-01	342102.9
March	172.21	2.27E-01	6.09E+05
April	245.35	3.24E-01	839450.3
May	265.95	3.51E-01	940263
June	312.32	4.12E-01	1068584
July	332.01	4.38E-01	1173817
August	296.99	3.92E-01	1050005
September	236.38	3.12E-01	808760
October	168.84	2.23E-01	596931.8
November	99.91	1.32E-01	341836.1
December	67.33	8.89E-02	238044.4

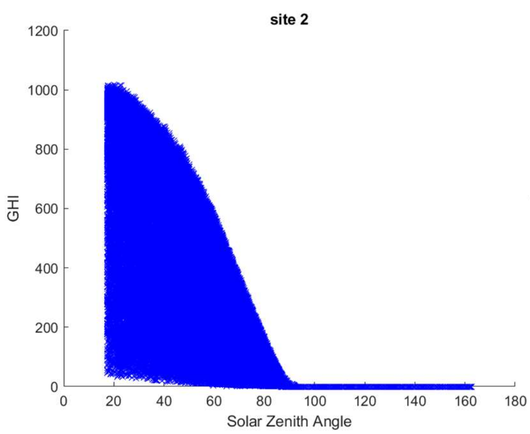
Site 2:

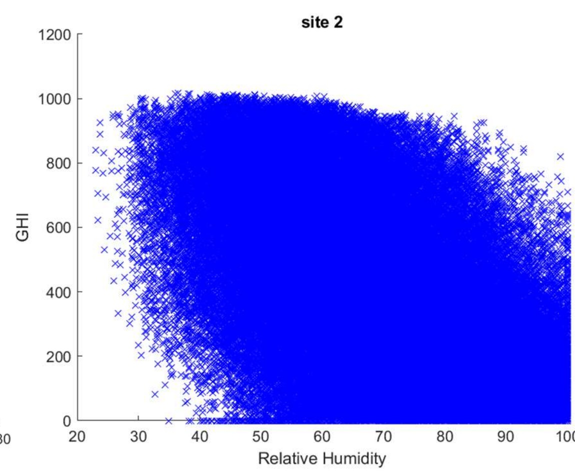
Data Cleaning:

Similarly to Site 1, we replaced null cells with predicted linear fit values.

Data Analysis:

After comparing the GHI to every factor in our data, the two largest correlation coefficients we found were between GHI and Solar Zenith Angle, and GHI and relative humidity. The graphs for these two relations are shown below.

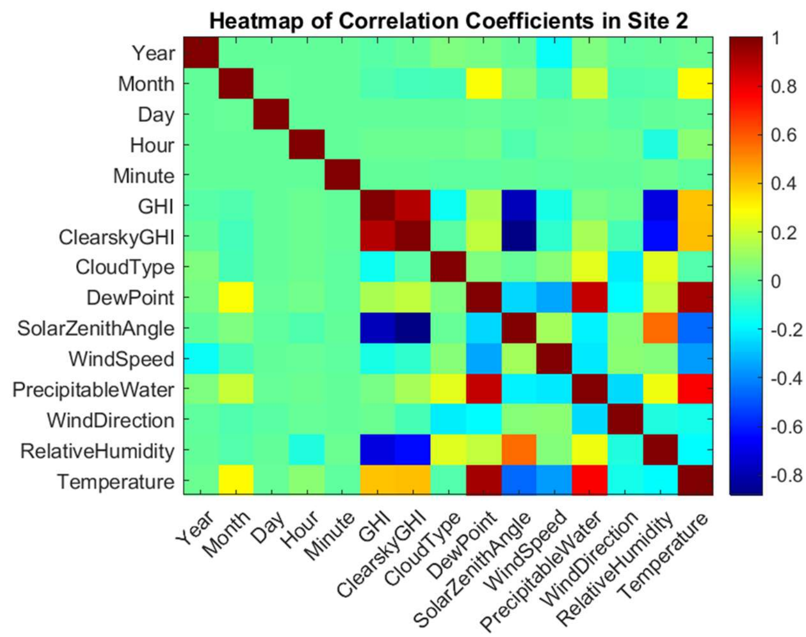




The correlation coefficient between the GHI and the Solar Zenith angle is equal to -0.7831. The relationship between the solar zenith angle and the GHI is the same as Site 1, where the GHI is proportional to the cosine of the solar zenith angle.

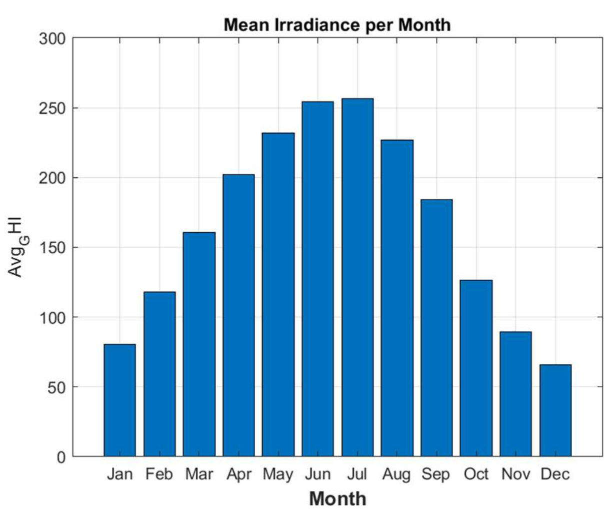
Moving on to the relative humidity, we find an inverse proportionality with a correlation coefficient of -0.7113.

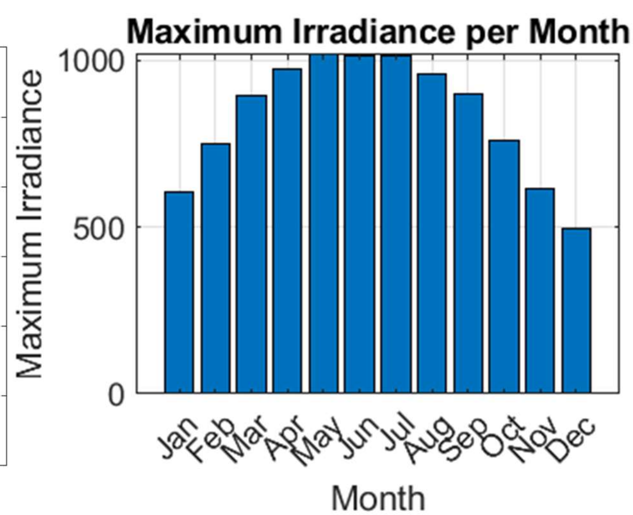
A heatmap was also constructed to represent the correlations between factors clearly.



We see that factors such as Relative Humidity and the Solar Zenith Angle are inversely proportional to GHI.

Since the irradiance varied over time, we calculated mean values for the GHI for every month of the year across the three years. We also calculated the maximum irradiance for every month.





Based on the graphs shown, GHI doesn't vary as intensely as Site 1. It is at its highest from April till September and has a moderate decrease from October till March. The maximum values generally vary between 500 and $1000W/m^2$. However, the mean values only vary from 66 to $256W/m^2$, which highlights consistent irradiance throughout the year.

The temperature was found to vary too little and was found between the range of 20^o to 37^o Celsius. Almost no correlation was found between it and GHI. High Humidity was also found in Site 2. However, it had no effect on GHI.

Solar System Selection:

Due to high humidity and low temperatures, the use of Concentrated Solar Power Systems is not possible since they are limited to hot, dry regions. The available site area is also small compared to the other sites. Therefore, the usage of **Monocrystalline** is advised, since they perform best in clear sky, limited space conditions with moderate temperatures.

We will use 60% of the area of the site, to get as much power as necessary while leaving enough space for other setups and buildings. This accounts for $5.4km^2$ of the site, which will cost us:

$$173.7 \times \frac{5.4 \times 10^6}{1} \approx \$940\ million$$

Energy Production

	Average GHI (W/m^2)	Power (GW)	Energy (GJ)
January	80.5699	8.70E-02	233062.3
February	117.8766	1.27E-01	307980.4
March	160.5193	1.73E-01	4.64E+05
April	202.1386	2.18E-01	565858.7
May	232.1090	2.51E-01	671415.2
June	254.1027	2.74E-01	711324.9
July	256.5697	2.77E-01	742172
August	226.9598	2.45E-01	656520.3
September	184.0684	1.99E-01	515273.7
October	126.6162	1.37E-01	366259.1
November	89.6742	9.68E-02	251030.4
December	66.0844	7.14E-02	191160.5

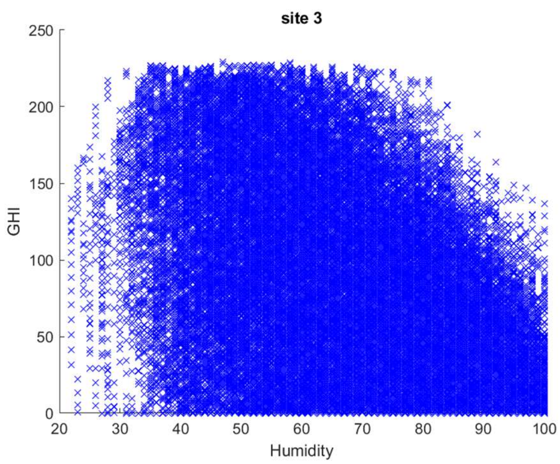
Site 3:

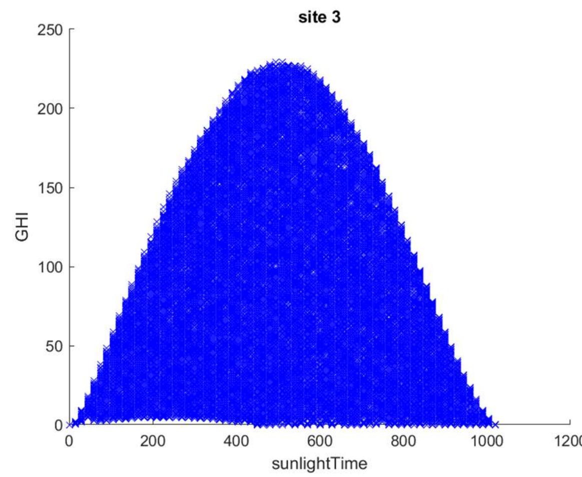
Data Cleaning:

In the datasheet, the time was compressed in one column as a date string, so we separated it to five columns: year, month, day, hour and minute. We checked if there were any null values present in the columns. Just like Sites 1 & 2 and replaced all null values with predicted linear fit values.

Data Analysis:

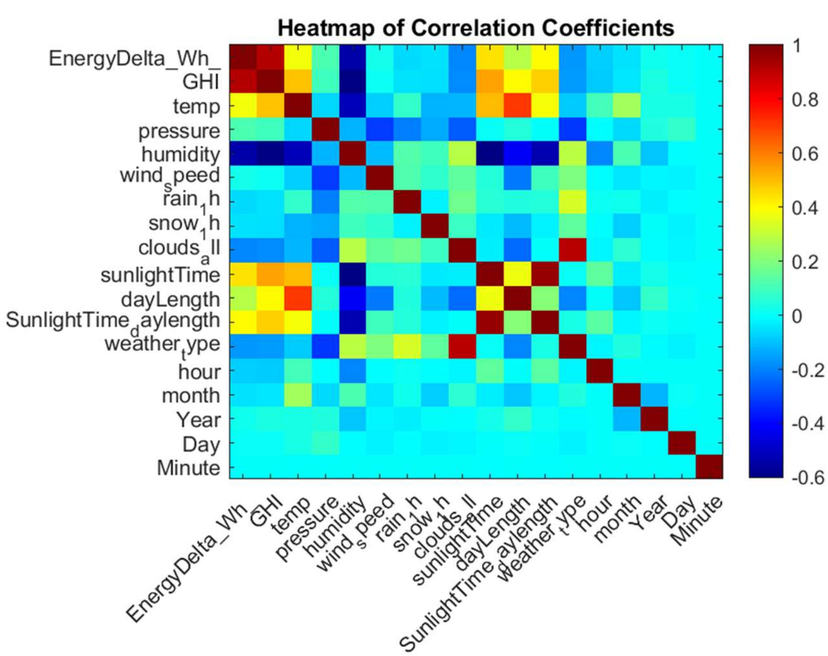
After comparing the GHI to each factor in our data, the two largest correlation coefficients we found were between GHI and Humidity, and GHI and Sunlight time. The graphs for these two relations are shown below.





The coefficient correlation between the GHI and the Humidity is equal to -0.6022. This implies a moderate inverse proportionality between both values. Moving on to sunlighTime, we find a direct proportionality with a correlation coefficient of 0.5402.

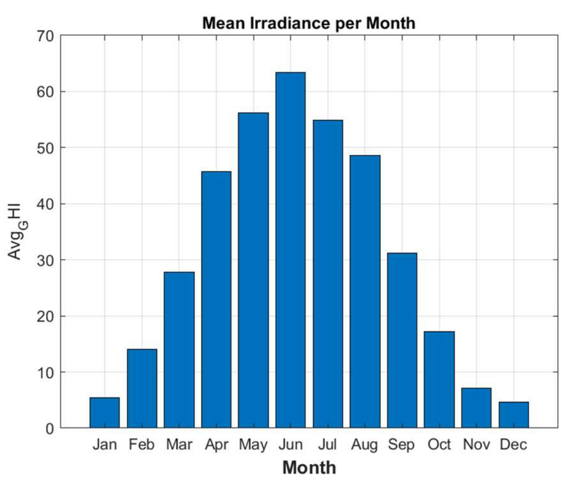
A heatmap was also constructed to represent the correlations between factors clearly.

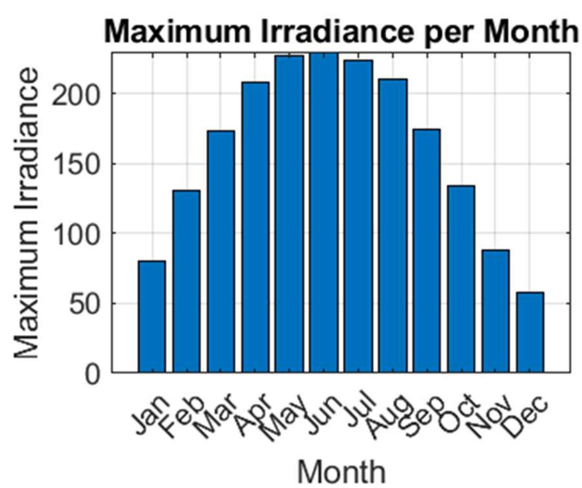


We see that factors such as Pressure and the Sunlight Time are directly proportional to GHI. Moreover, factors such as clouds and Weather type appear to be inversely proportional to the GHI.

Snow & Rain are found on-site, which affects the efficiency of the Panels and will require careful and consistent maintenance of the solar system. Clouds are present densely throughout the day, affecting the efficiency of the panels. On top of that, high humidity is negatively affecting GHI the most out of all-weather factors.

Since the irradiance varied over time, we calculated mean values for the GHI for every month of the year across the three years. We also calculated the maximum irradiance for every month.





As shown on the irradiance graphs, Site 3 has incredibly low GHI, meaning that very little solar power can be used for energy production. The average GHI varies from 4.5834 to

$63.4019 W/m^2$. This is astoundingly low compared to the values of Sites 1 and 2. The highest maximum irradiance is less than $250 W/m^2$.

Solar System Selection:

Due to high humidity and low temperatures, the use of Concentrated Solar Power Systems is not possible since they are limited to hot, dry regions. Therefore, the usage of **Thin-Film** is advised since they perform better in low-light conditions and are generally more tolerant with high temperatures.

We will use 75% of the area of the site, to get the most out of the low efficiency of Thin-Films. This is $27km^2$. To cover this area, the panels will cost us:

$$75 \times \frac{27 \times 10^6}{1} = \$2.025\ billion$$

Energy Production:

	Average GHI (W/m^2)	Power (GW)	Energy (GJ)
January	5.49	1.26E-02	33746.63472
February	14.07	3.23E-02	78117.54048
March	27.74	6.37E-02	1.71E+05
April	45.73	1.05E-01	272031.3072
May	56.16	1.29E-01	345211.4765
June	63.41	1.46E-01	377203.2624
July	54.87	1.26E-01	337281.9394
August	48.58	1.11E-01	298617.7622
September	31.15	7.15E-02	185300.136
October	17.22	3.95E-02	105850.1002
November	7.14	1.64E-02	42473.2896
December	4.58	1.05E-02	28152.93024

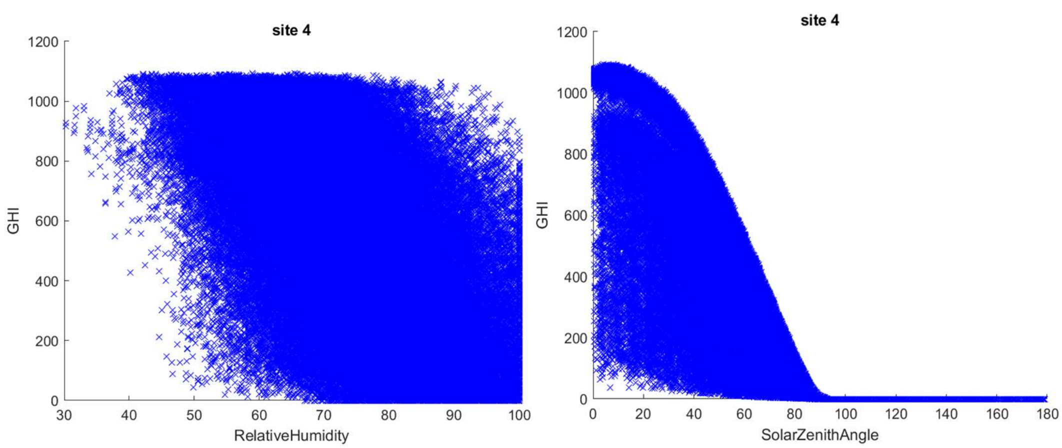
Site 4:

Data Cleaning:

We replaced the null values present with predicted linear fit values. In addition, GHI and Clearsky GHI had the exact same values, so we removed the Clearsky GHI column, assuming it was a duplicate.

Data Analysis:

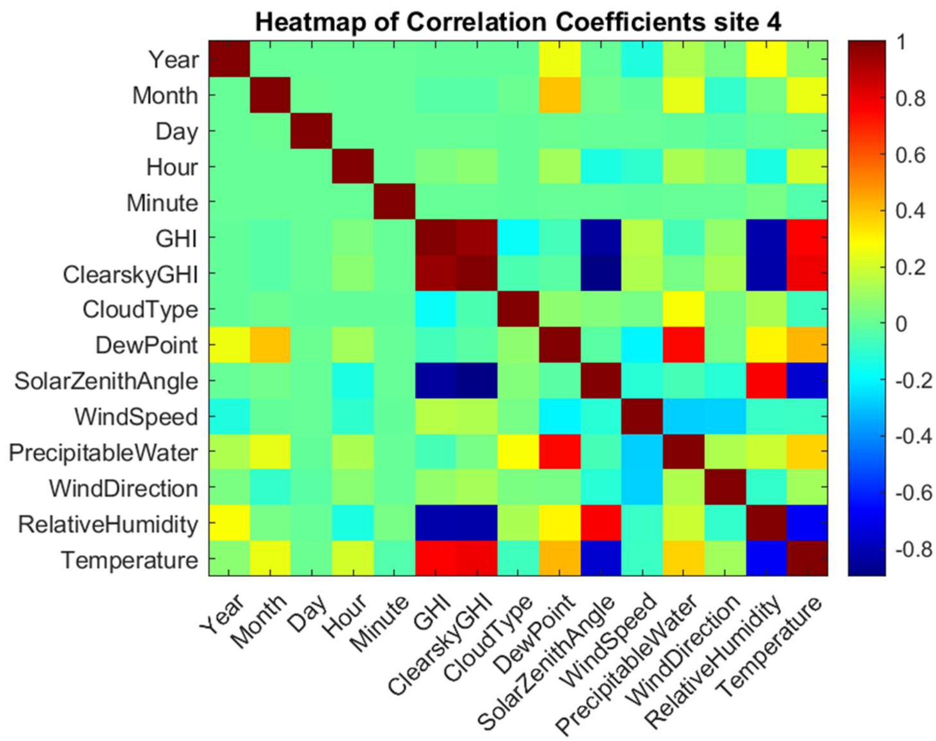
After comparing the GHI to the factors in our data, the two largest correlation coefficients we found were between GHI and Solar Zenith Angle, and GHI and Relative Humidity. The graphs for these two relations are shown below.

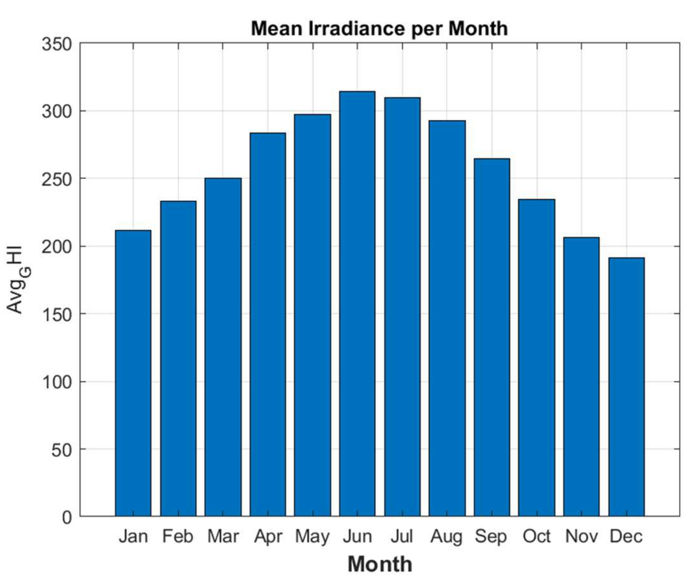


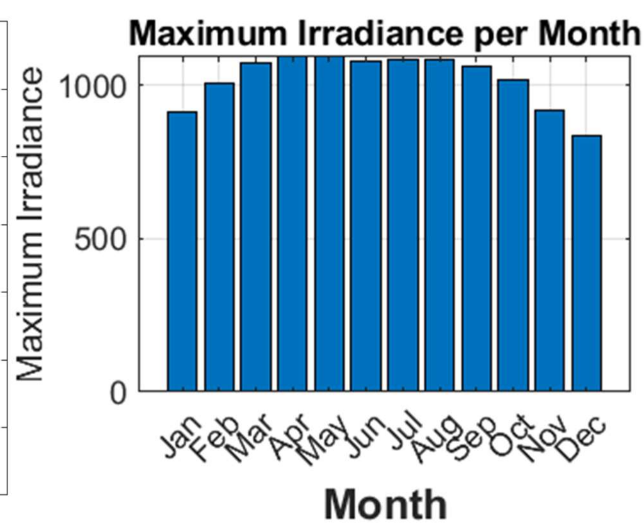
The correlation coefficient between the GHI and the Solar Zenith angle is $r = -0.8488$. We have the same relation between GHI and Solar Zenith Angle as the three previous sites.

Moving on to the Relative Humidity, we find an inverse proportionality with a correlation coefficient of $r = -0.8184$. This implies a strong negative relationship between both values, and a large effect on GHI.

A heatmap was also constructed to represent the correlations between factors clearly.







Based on the graphs shown, the GHI is consistent throughout the year. The values generally vary between 200 and $400W/m^2$. This shows that there is barely any difference in GHI across the year, and that GHI is relatively high for this site.

While the temperature in this site is similar to the other sites' temperatures, it covers a smaller range at higher temperatures between 14^o and 24^o Celsius. This represents consistently high temperatures for the site. A positive correlation of $r = 0.7557$ was found between it and the GHI, meaning it strongly affects GHI in this site.

Solar System Selection:

Due to high humidity, no rain, and high temperatures, the use of Photovoltaic Systems won't be effective. Therefore, the usage of Parabolic Troughs is advised.

Assuming the site is off grid, we will require a large amount of energy from the panels, and energy storage for the site to stay up during the night. Therefore, we will use Parabolic Troughs with 6hr storage. We will need 40% of the area, since parabolic troughs already have high efficiency. This is an area of $10km^2$. This will cost:

$$200 \times \frac{10 \times 10^6}{1} = \$2 \; billion$$

Energy Production:

	Average GHI (W/m^2)	Power (GW)	Energy (GJ)
January	211.7681	9.74E-01	2609118.52
February	233.263	1.07E+00	2595825.31
March	250.1062	1.15E+00	3081468.45
April	283.4972	1.30E+00	3380193.82
May	297.0336	1.37E+00	3659644.05
June	314.2091	1.45E+00	3746377.94
July	310.037	1.43E+00	3819854.26
August	292.576	1.35E+00	3604723.57
September	264.8054	1.22E+00	3157327.75
October	234.702	1.08E+00	2891678.85
November	206.3561	9.49E-01	2460425.05
December	191.5485	8.81E-01	2360000.11

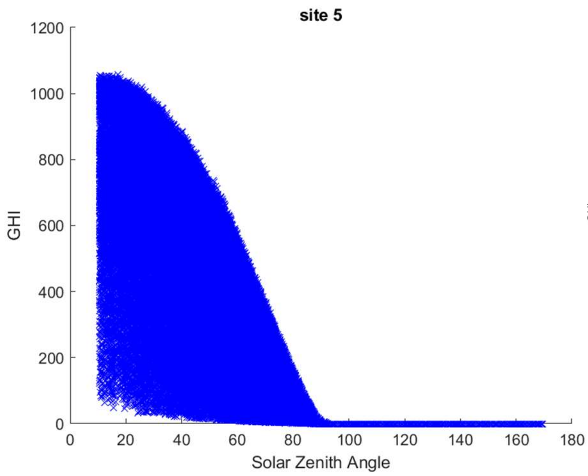
Site 5:

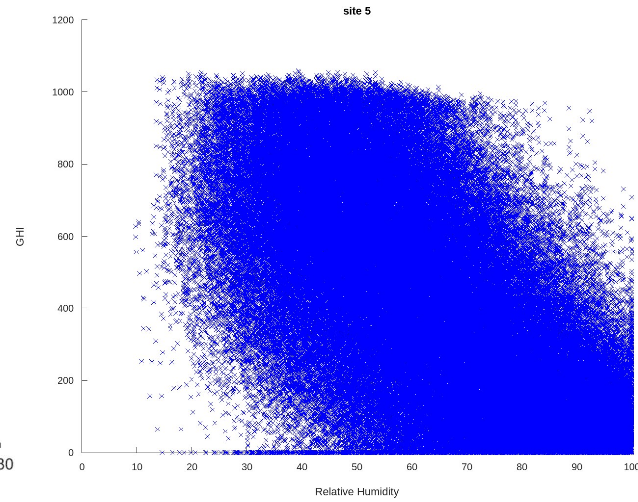
Data Cleaning:

We replaced the null values present with predicted linear fit values. In addition, GHI and Clearsky GHI had the exact same values, so we removed the Clearsky GHI column assuming it was a duplicate.

Data Analysis:

After comparing the GHI to the factors in our data, the two largest correlation coefficients we found were between GHI and Solar Zenith Angle, and GHI and Relative Humidity. The graphs for these two relations are shown below.

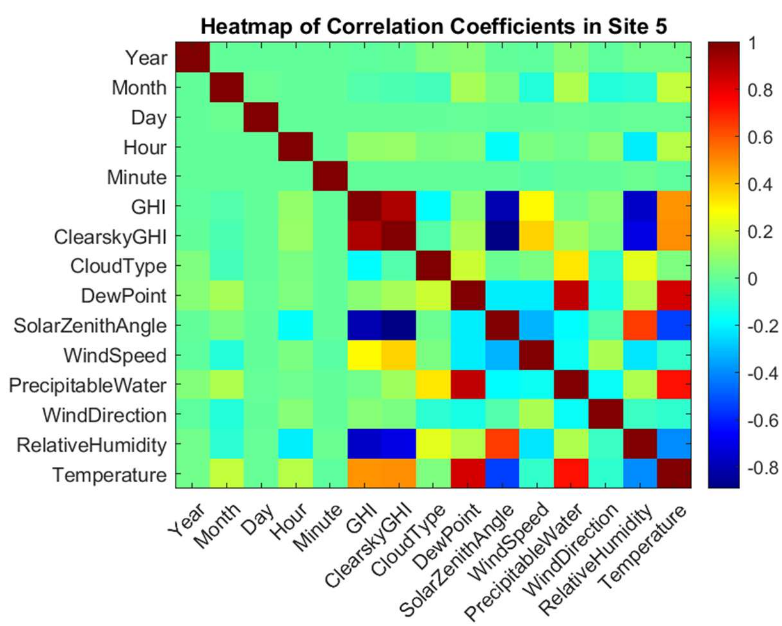




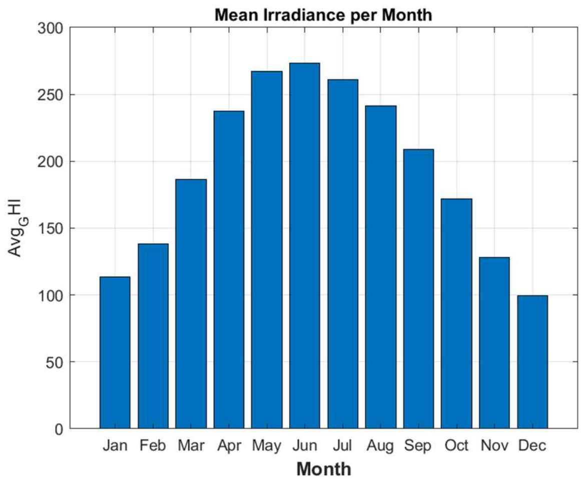
The correlation coefficient between the GHI and the Solar Zenith angle is equal to -0.8054, this implies a high inverse proportionality.

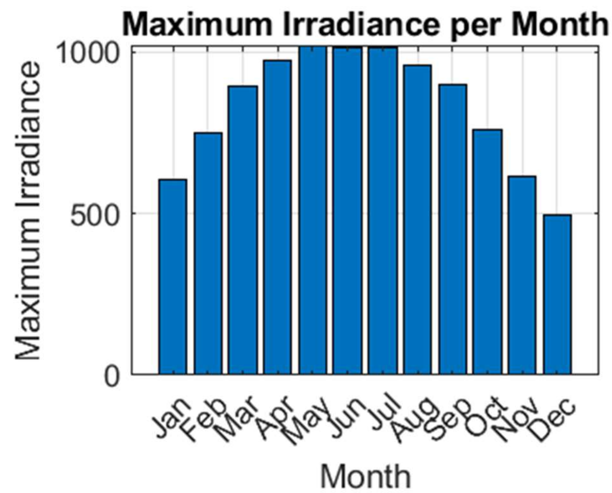
Moving on to the Relative Humidity, we find an inverse proportionality with a correlation coefficient of -0.7634, which means that GHI is largely affected by the high relative humidity in this site.

A heatmap was also constructed to represent the correlations between factors clearly.



We see that the only factors that have a direct proportionality with the GHI are the Temperature and the Wind Speed. Moreover, Cloud Type appears to be inversely proportional to the GHI.





Based on the graphs shown, the GHI is somewhat varying throughout the year while being the highest around the middle of the year. The values generally vary between 100 and $270W/m^2$. This shows that the GHI is relatively high for this site.

In terms of cloud cover, the sky is practically clear most of the time. At cloud types 0-4, GHI is almost the same. However, it begins to significantly decrease in cloud type 6 and higher. These types are rarer than 0-4, meaning that the sky is clear for longer durations.

In terms of temperature, the temperature of the site generally varies from 0^o to 35^o. The correlation between GHI and temperature is 0.4843 , meaning it has somewhat of a positive effect on GHI.

Solar System Selection:

Based on our analysis, CSP systems would be inefficient since the temperature has large variations between cold and hot. GHI is high all-year, meaning that thin films would be ineffective. The two available options are polycrystalline and monocrystalline.

Due to the high GHI and clear skies, it would be best to use 60% of the area available. We will divide this area in two. One half will contain monocrystalline panels, while the other half will have polycrystalline panels. The total area for the solar systems will be $9.6km^2$, and the cost will be:

$$173.7 \times 4.8 \times 10^6 + 130.2 \times 4.8 \times 10^6 = \$1.46\ billion$$

Energy Production:

	Average GHI (W/m^2)	Power (GW)	Energy (GJ)
January	113.507	2.60E-01	697719.36
February	138.381	3.18E-01	768300.17
March	186.483	4.28E-01	1146297.57
April	237.329	5.45E-01	1411784.78
May	267.245	6.13E-01	1642735.77
June	273.368	6.27E-01	1626167.82
July	260.851	5.99E-01	1603432.32
August	241.245	5.54E-01	1482915.65
September	208.868	4.79E-01	1242480.54
October	171.598	3.94E-01	1054800.55
November	127.945	2.94E-01	761098.74
December	99.473	2.28E-01	611453.37

Conclusion:

In conclusion, we were able to analyze the data for five different sites of different sizes. By analyzing the data, we were able to select different Solar Systems and different areas for each site and calculate the monthly energy productions. The chosen solar systems, areas and costs are summarized below.

	Solar System Type	Area of System	Total Cost
Site 1 ($16km^2$)	Polycrystalline	$8km^2$	$\$1.042\ billion$
Site 2 ($9km^2$)	Monocrystalline	$5.4km^2$	$940 $million$
Site 3 ($36km^2$)	Thin film	$27km^2$	$\$2.025\ billion$
Site 4 ($25km^2$)	Parabolic Troughs	$10km^2$	$\$2\ billion$
Site 5 ($16km^2$)	Polycrystalline & Monocrystalline	$9.6km^2$ ($4.8km^2$ each)	$\$1.46\ billion$

Contribution Sheet:

Team Members:

Ziad Ahmed [23-101292]

Yousef Omar [23-101120]

Abdullah Helmy [23-101239]

Abdelrahman Gomaa [23-101279]

Seif Eldin Haythem [23-101282]

Yassin El Khodary [23-101163]

Tasks:

Matlab Code:

Abdelrahman Gomaa/Seif Eldin Haythem/Yassin El Khodary

Data Analysis, Solar System Selection, and Project Report:

Yousef Omar/Abdullah Helmy/Ziad Ahmed

Bibliography:

- “Renewable Energy Cost Analysis: Concentrating Solar ...” IRENA, 2012. https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2012/RE_Technologies_Cost_Analysis-CSP.pdf.
- “Best Research-Cell Efficiency Chart.” NREL. Accessed May 27, 2024. https://www.nrel.gov/pv/cell-efficiency.html.
- “Compare and Save on Clean Home Energy Solutions.” EnergySage. Accessed May 27, 2024. https://www.energysage.com/.
- Solar performance and efficiency | department of energy. Accessed May 27, 2024. https://www.energy.gov/eere/solar/solar-performance-and-efficiency.