

Tanasorn Chindasook - Term Paper

Geographic Information Systems Applications in Colorado Wind Farm Location Selection

Tanasorn Chindasook, M.Sc. Data Engineering

MEGI001-210213 - Geo-Information Systems

Topic description

The purpose of this project is to find optimal locations in Colorado, USA for wind farm construction through the application of GIS methods. The primary focus is to create a multi-layer map with various features such as elevation, cities, highways and wind power and combine them by utilising GIS tools such as ArcGIS Online and QGIS in order to derive essential information through map data visualisation to locate an optimal site for wind farm development.

Initially, this report will explore existing data on the various features in Colorado which provides an overview into the geography of the state. Exploring data is important as knowledge about the terrain is vital in the following decision making on optimal locations. Subsequently, this report will address the methodology in the creation of the project. This section includes the description of each data layer utilised and the combination of the various features into meaningful analyses using ArcGIS Online and QGIS. This process allows for specific conditions to be implemented into the decision making process. Finally, the results of the project will be discussed and a recommendation will be made on the optimal location for wind farm construction.

Data and Methodology

The initial data layers used for this project are listed below by layer name, creator, last updated date and can be found on ArcGIS Online:

1. Highway Traffic, DTDGIS_OpenData, Dec 13, 2018
2. Highways: Functional Class, DTDGIS_OpenData, Dec 28, 2018
3. Colorado City Boundaries, CDPHE_user_community, Feb 19, 2018
4. Continental Divide in Colorado, 2016SP_CDA, Jan 15, 2017
5. Federal Lands, EmanuelVasquez, Oct 11, 2018
6. Elevation Colorado, TI 2017fairfax509, Jan 3, 2018. This is a feature layer generated from interpolated points.
7. Colorado County Boundaries, CDPHE_user_community, Feb 19, 2018
8. Wind Power Class, Learn_ArcGIS, Jun 26, 2018
9. rivers_streams, amelendez_si, May 2, 2017
10. Colorado Major Streams, sskach_PSDSchools, Nov 23, 2016
11. HailWindTornadoDroughtDensity, EDymond@DHSEM, Feb 13, 2018

Colorado Data Exploration

The highest elevations in Colorado is an area of 7,236.58 square miles, with a maximum elevation of 3177.62 metres above sea level. Areas of high elevation occur along the continental divide. The elevation layer used is a rather generalised map layer as it was interpolated from a set of data points.

The major rivers in Colorado are the Arkansas River, North Platte River, Smoky Hill River, Green River and Gunnison River. River flow directions in Colorado are significantly affected by the side of the continental divide that it lies on. Rivers on the West side of the continental divide will flow West, and vice versa.

Cities in Colorado are located almost exclusively on the eastern slope in areas with a maximum elevation of 1,986.40 metres and a minimum elevation of 1,555.57 metres, separated across two different classes (4 and 5) of elevations. A majority of the principle highways are built for ease of access into the cities and therefore congregate near the cluster of cities. The major highways in Colorado appear to follow the path of major rivers. This is an indication of the development of settlement patterns as cities are also located near major rivers. In the lower elevation areas, highways seem to follow a straight route creating a grid pattern, whereas in higher elevation areas, highways seem to be more adapted to the terrain. The most densely populated counties in Colorado are Denver, Jefferson, Douglas and El Paso. The largest five cities in Colorado are Denver (714,516), Colorado Springs (470,885), Aurora (370,859), Fort Collins (164,597) and Lakewood (157,224) according to 2018 ESRI Total Population Data. Wind power in Colorado is stronger in the east than in the west, but the location that experiences the most powerful winds are along or in close proximity to the continental divide with power class 5, 6 and 7.

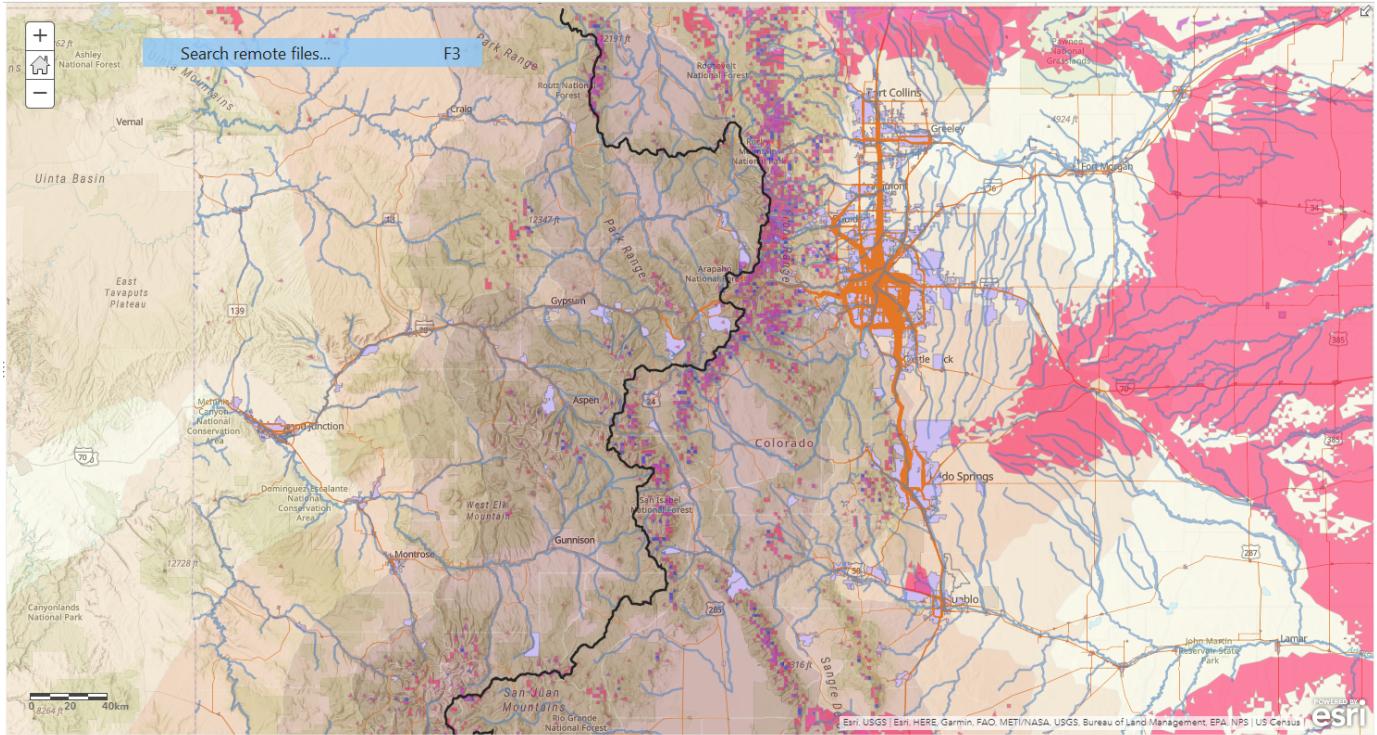


Fig 1. An overlay of the initial layers. Various combinations of these layers were used to provide the data exploration information.

Methodology

The initial step in the project's conception was the research and acquisition of the different required layers in ArcGIS Online. In order to develop a more comprehensive view of the streams layer, two different data layers were merged using the merge layers function of ArcGIS Online. Initial requirements for suitable locations are then fulfilled, followed a set of selection criteria to determine the optimal location out of all suitable locations.

The initial requirements from the government for the project are as follows:

1. In an area where the Wind Power Class is at least 4, which is “good” or better in terms of wind farm resource potential.
2. No further than 5 miles from a highway, so that maintenance crews can more easily access the wind turbines.
3. No further than 50 miles from a city containing at least 25,000 people, to ensure a nearby target market.
4. Not on Federal Land (national parks, forests, grasslands, and so on).
5. On land that meets criteria 1-4 that is also at least 1 square kilometer in size, large enough for a wind farm.

In addition to the initial requirements, one further criterion was added to ensure that the structural integrity of the wind turbines would bear the least risk of being compromised due to natural hazards; more specifically tornadoes. Since Colorado experiences approximately 41 tornadoes per year and tornadoes can cause significant structural damage to wind turbines, it is logical that this condition should be factored into the decision [2]. The additional criterion was later applied to the sites that met all of the initial requirements.

In order to filter out areas that do not satisfy the condition of wind power class 4 or above, the filter field in the layer was selected and the mean value for wind power was set to a mean greater than 16 as the mean for class 4 winds is 16.80. A buffer of 5 miles from highways was added to the Highways: Functional Class layer to create an area that satisfies the second requirement. Subsequently, to find cities with over 25,000 inhabitants, data enrichment was performed. The Colorado City Boundaries layer was enriched with the 2018 Total Population data provided by ESRI. A filter was then applied to the layer to display cities with populations greater than 24,999. Furthermore, buffers of 50 miles were then created around the Colorado City Boundaries layer in order to ensure a target market was in the vicinity of the wind farm. Federal Lands are then excluded from the suitable locations via erasure of areas in the Federal Lands layer from the Wind Power Class, Road Buffer and City Buffer layers.

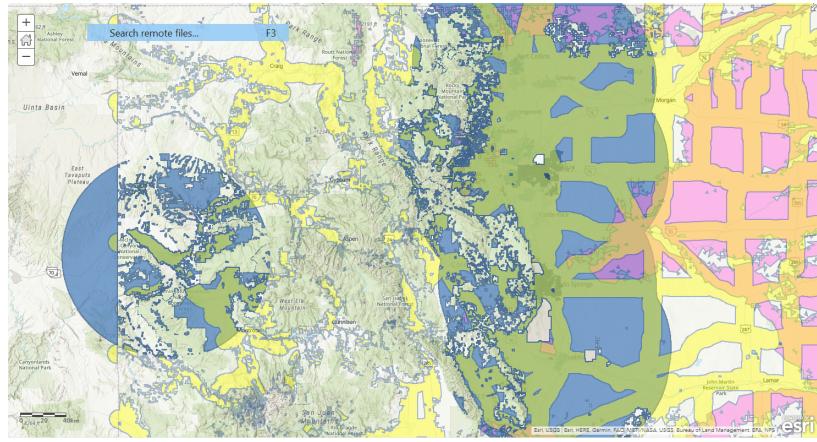


Fig 2. shows the overlay of the layers that have been manipulated to satisfy initial requirements 1 to 4.

With regards to the additional criterion of removing areas exceptionally prone to tornadoes from the existing suitable wind farm locations, the "HailWindTornadoDroughtDensity- Colorado Tornado Density" layer was added to the map and filtered to only show areas with gridcode greater than 1.8. These areas were then excluded from the existing suitable locations via erasure.

Moreover, in order to fulfill the final criterion of displaying land that is more than 1 square kilometre in size, the "intersect hwy wind cities excluding FL" was exported from ArcGIS Online as a Shapefile and opened in QGIS. A grid of 1 square kilometre was then created on the layer extent through the use of Vector Research Tools. The grid layer created in QGIS was then exported as a Shapefile, hosted on ArcGIS Online then added to the map. The areas of 1 square kilometre in size where then generated through the intersection of the grid layer and the existing suitable areas, then filtered on areas equal to 1 square kilometer. The final suitable wind farm locations that are over 1 square kilometre in size and exclude areas prone to tornadoes are displayed in green under the "Intersect of Suitable Areas Excluding Areas at Risk and colorado grid 1km" layer, whilst the suitable locations that are over 1 square kilometre in size and include areas at risk to tornadoes are displayed in yellow under the "Wind Farm Grid 1km Including Areas at Risk" layer.

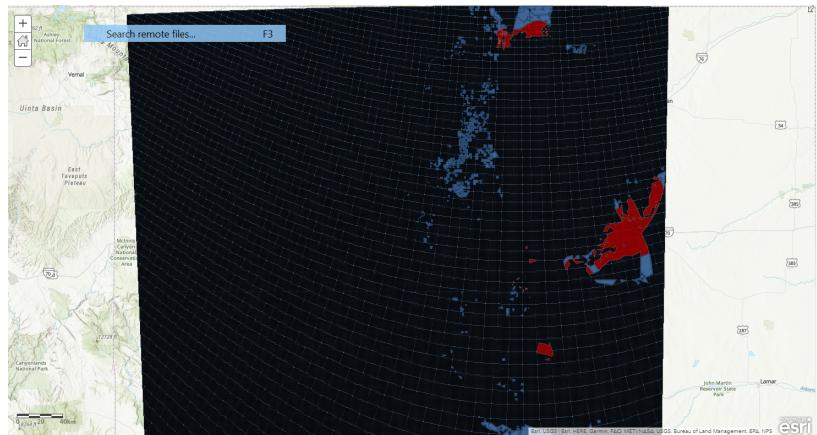


Fig 3 displays the QGIS created grid overlay with the suitable areas, encoded by red to denote areas at risk, and blue to denote safe areas.

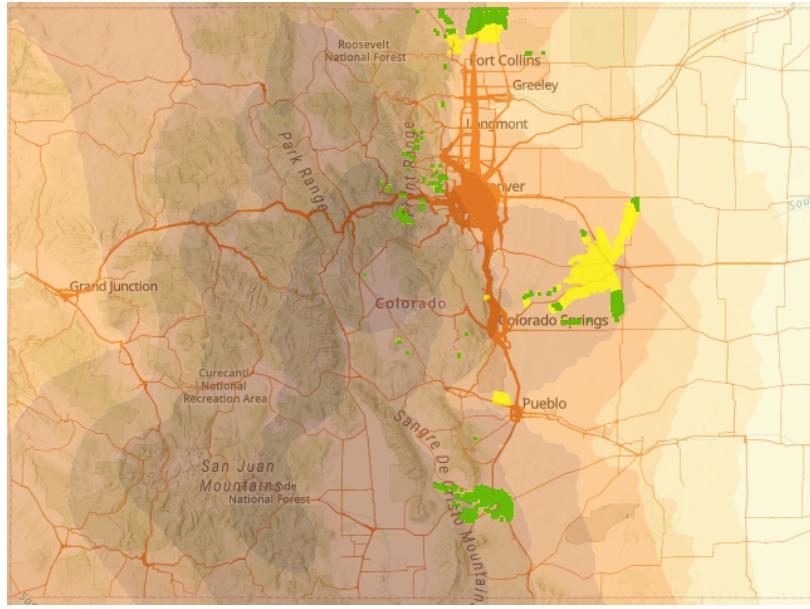


Fig 4. Illustrates the final locations suitable for wind farm construction after grid intersection. The areas at risk of tornadoes are encoded in yellow, whilst the areas with less risk are encoded in green.

Finally, after all of the suitable wind farm locations that satisfy all the requirements above have been located, a decision criteria was formulated in order to determine the optimal site for wind farm construction out of all the suitable areas. The decision criteria is as follows:

1. Areas which are closer to major cities are preferred in order to reduce the commute time for construction workers that live in the city.
2. Areas which are less suspect to traffic are preferred to ensure that minimal disruption to the citizens occurs during the assembly of the power plant. Traffic information for this criteria is displayed using the "Highway Traffic" layer. Roads with high traffic are larger in size than those with less traffic.
3. Areas in higher elevation are preferred to minimise obstruction of winds.
4. Areas in which the wind farm serves the most cities are preferred to ensure the maximum consumption for future expansion purposes.
5. Areas that allow for future expansion are preferred.

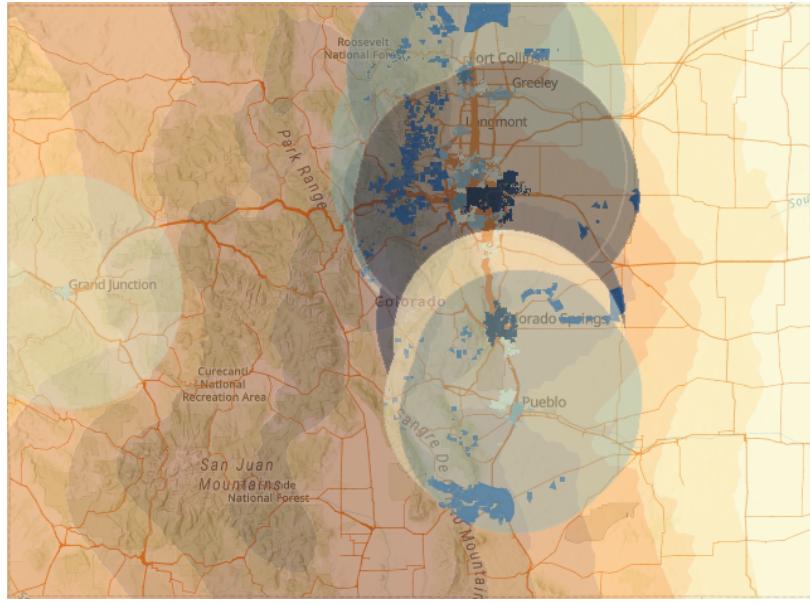


Fig 5. Display of the overlap of the cities' 50 miles buffer, highway traffic and elevation to determine which locations serve the most cities, are less prone to traffic and are located in higher elevations respectively.

Results and discussion

Only one cluster of areas fulfill all of the preferences denoted in the selection criteria. This cluster is located to the left of the map and marked by the green circle. This cluster is situated at an elevation of 3,177.62 – 2292.89 ft., which is the highest elevation out of the four clusters and makes the area less susceptible to structures obstructing the wind. There are many roads that lead into this area, therefore the construction of this power

plant will neither hinder the day-to-day lives of the citizens by obstructing traffic, nor will construction be hindered by a blockage of a road. The cluster also located in the proximity of many major cities in Colorado, ensuring that a target market is in abundance should the government decide to expand in the future.

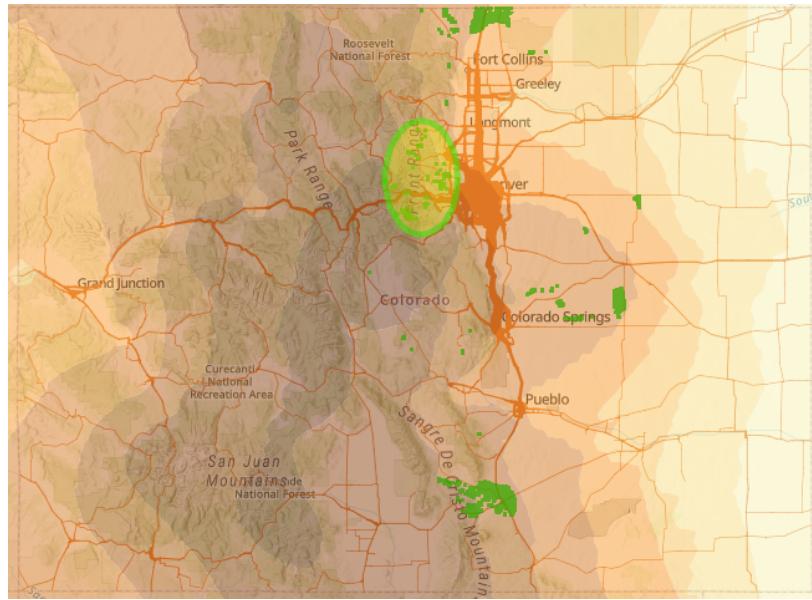


Fig 6. denotes the best cluster of locations by the green circle with the yellow fill.

Within the cluster, there are two main areas that are most suitable for wind farm development. Since the limitation of the one square kilometre grid is that it did not account for areas which are not perfectly square, the original shape of the available areas are reintroduced to provide a better overview of the actual available space. There are trade-offs between the two areas, marked Location 1 and Location 2. The trade-off between the two areas are elevation and accessibility. Location 1 is closer to civilian settlements which will prove more convenient for construction engineers and maintenance to access. The time wasted on commuting will be reduced, thus worker satisfaction could improve and the construction could be finished in a faster period of time, leading to the possibility of lowering labour costs. However, Location 2 is located at a higher elevation, which is beneficial as there are less structures to obstruct the flow of wind. Nonetheless, judging by the terrain visible on the map, Location 1 seems to encompass a flatter terrain, making it easier to build on than Location 2. Ultimately, due to the similarities in all other variables and the advantages in accessibility and terrain, Location 1 seems to be the most optimal location for wind farm construction.

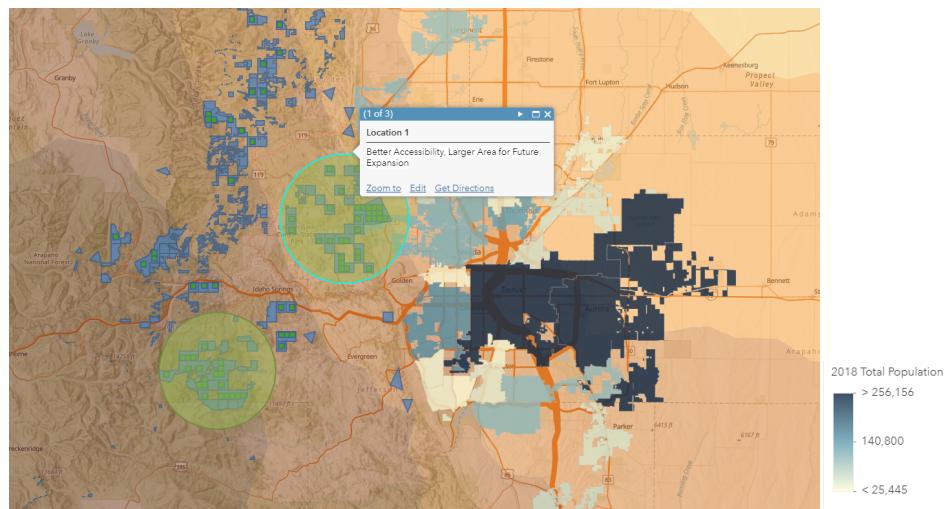


Fig 7. A screen capture displaying one of the two suitable locations for wind farm construction.

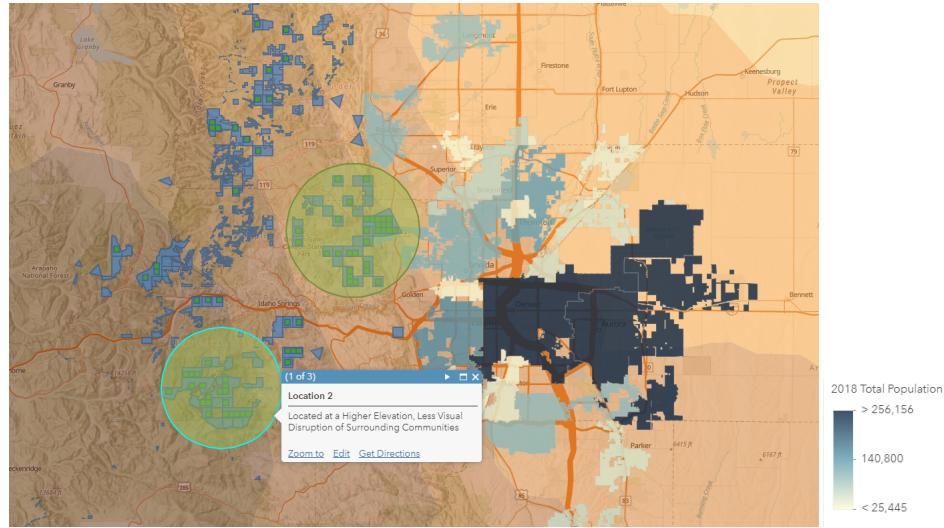


Fig 8. A screen capture displaying the second suitable location for wind farm construction.

The link to the final web application can be found below in the "Links to Files and Repos" section of the paper.

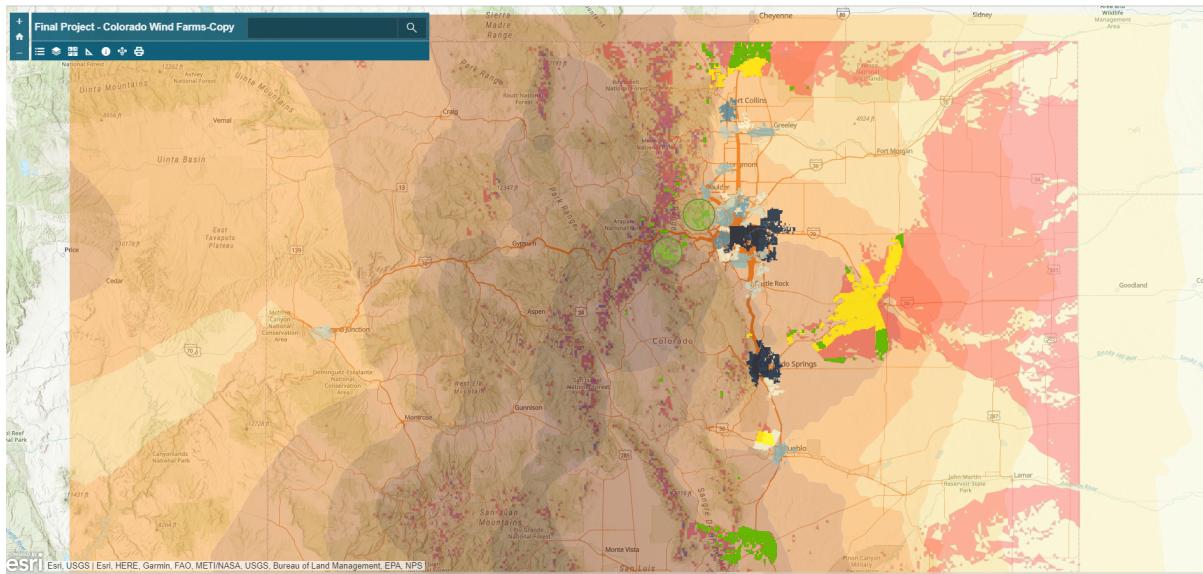


Fig 9. A screen capture of the final Colorado wind farm web application.

Exploring Trends Through Data Visualisation Approaches for Operational Optimisation of Hostels in Taghazout Beach, Morocco

Tanasorn Chindasook, M.Sc. Data Engineering

MEGI001-210103 Introduction to Earth System Data

Topic description

The interest in tourism influx for the coastal town of Taghazout, Morocco has steadily increased in recent years. According to hostel owners in the area, the amount of tourism in Taghazout is highly dependent on the surf conditions in city, as active tourism in Taghazout comprises mainly of individuals looking to participate in surfing related activities. Evidence of this is present through the search results of hostel names around the area on primary booking websites [1,3]; the majority of hostels fall into the category of "surf hostels" as denoted by the inclusion of the word "surf" in the hostels' naming conventions.

The purpose of this project is to utilise web crawling algorithms to retrieve historical data on indicators of optimal surf conditions, such as swell height and swell period, compute its statistics, and apply data visualisation methods on time series data in order to determine trends and seasonality of surf conditions in Taghazout. Determining periods of low-season where the surf conditions are suboptimal or mediocre will allow hostel owners to plan their cash flow to budget for periods of low tourism influx. Some operational strategies that can be implemented once seasonality has been determined include increasing accommodation prices during high-season, reducing accommodation prices during low-season, employing less labour during low season, performing structural maintenance during low season or expansion during low-season to increase profits or reduce incurred costs.

The rest of the paper is structured as followed: the data acquisition section describes the methods used in obtaining data for the project, as well as the source of the data obtained. The data exploration section details the statistics of the data obtained and provides an insight into the properties of the dataset. The trend visualisation section explores various functions that were fit to the data in order to determine seasonality and explains certain chosen parameters for the functions. Finally, the results and discussion section summarises the findings of the report and provides recommendations for hostels that wish to utilise these findings to optimise its operations.

Data and methods used

Data Acquisition

The data primarily used to conduct this research was obtained through the use of an HTML parsing web crawler on the magicseaweed website. The data encompasses daily averages of swell height in metres, swell period in seconds, wind speed in miles per hour, wind direction in degrees from north and swell direction in degrees from north for Taghazout Beach over the period of January 1st 1997 to December 31st 2006 and was retrieved in CSV format. The primary sources of data that magicseaweed draws upon are the NWW3 wave forecast model and the GFS weather forecast model provided by the NOAA [4]. In addition, the magicseaweed organisation also utilises its own software system to operate several nearshore modelling systems that receive inputs of the most recent forecast winds and swells [4]. The algorithm for the crawler predominantly utilises the BeautifulSoup library to parse webpage HTML, and can be found below in the GitHub repository listed in the "Links to Files and Repos" section of the report.

The data for the overall heat map visualisation of swell heights around the world was obtained from the NWW3 server provided by the NOAA. The data encompasses the period from January, 1997 to September, 2006 and provides a monthly measure of swell height data on a global scale.

Data Exploration

Following data acquisition, the CSV dataset was loaded into Jupyter Notebooks in order to compute primary statistics to gain a better insight into the relationships between the variables. The variable means, standard deviations, minimums, maximums and boxplots were computed for swell height, swell period and wind speed. A quadratic function was fit to show the development of each statistic across time.

A notable observation that can be made is that swell height means and wind speed means follow the same pattern, suggesting that wind speed is an contributing factor to the swell height. Another interesting observation that can be made through these results is that the maximum swell height has consistently decreased over time even though maximum wind speed has increased. This could be attributed to the shift in sand

deposit patterns causing the waves to break earlier. It would be interesting to assess the development of sand deposit patterns through the use of satellite imagery in order to understand why maximum swell height is decreasing in the future, however, it is out of scope for this project.

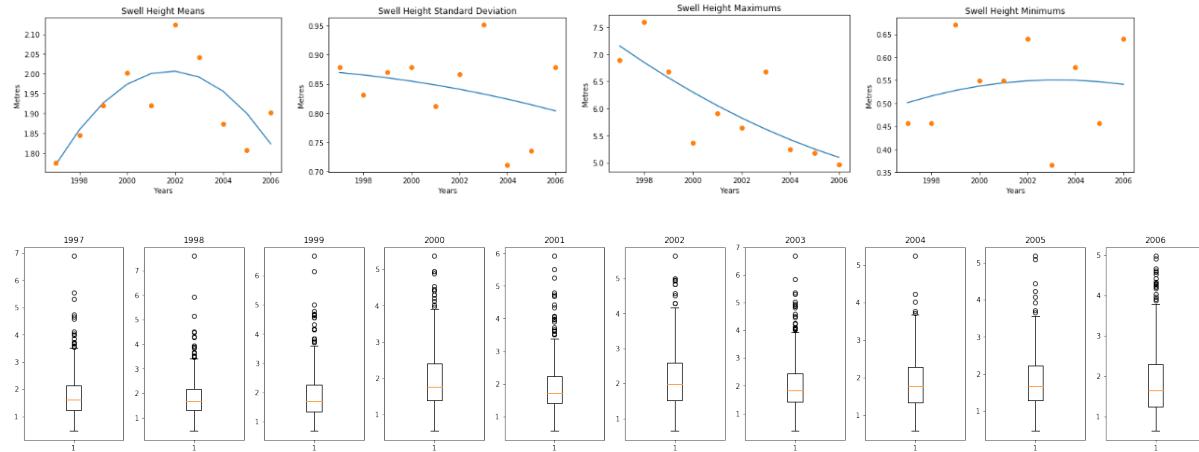


Fig 1. Computed statistics for the swell height variable

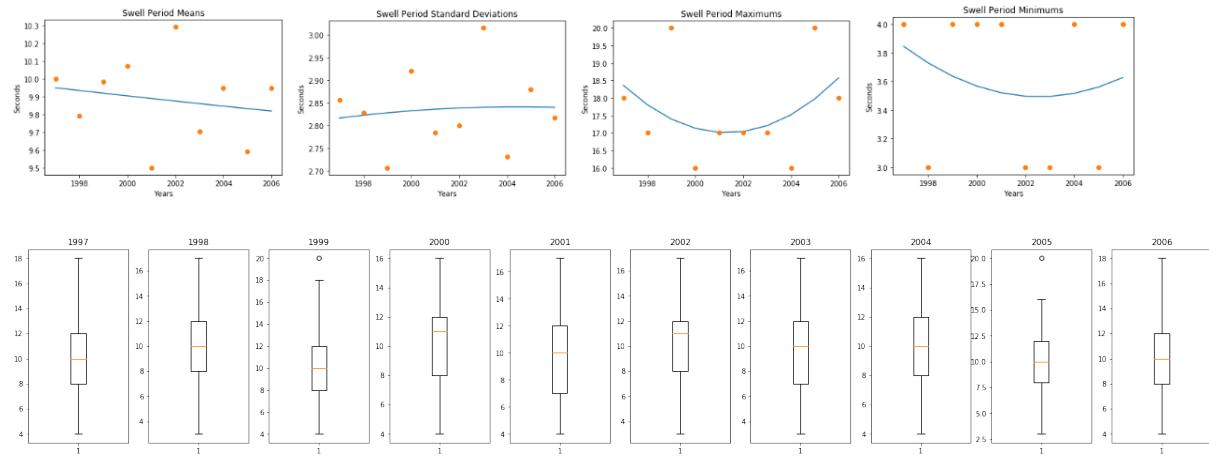


Fig 2. Computed statistics for the swell period variable

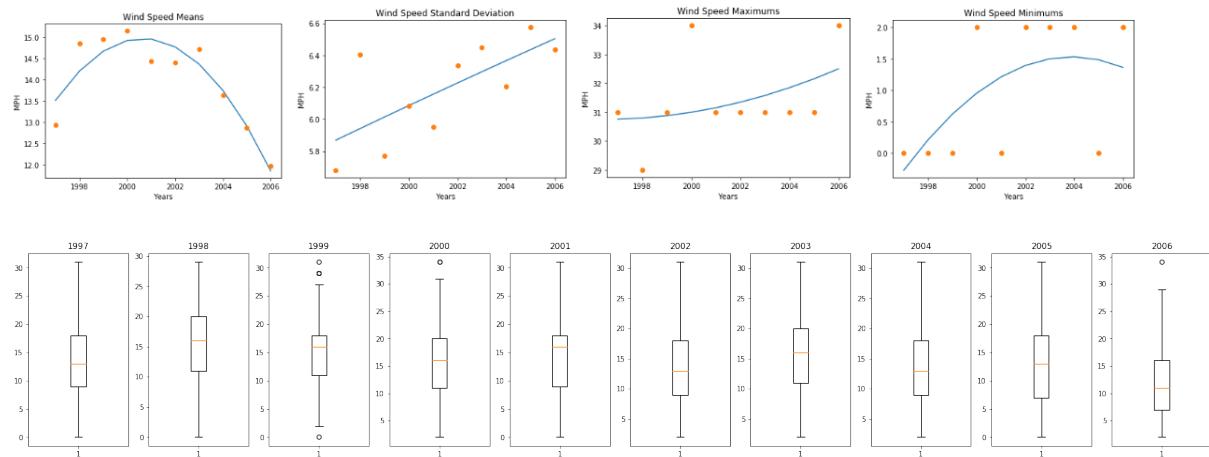


Fig 3. Computed statistics for the wind speed variable.

Wind speed and wind direction are plotted using a Windrose diagram to visualise any possible relationships between the two variables. It is apparent that the majority of winds that Taghazout experiences are north winds, making this variable less important as it remains relatively the same for the majority of the data points. Similarly, swell direction and swell period were visualised as a Windrose diagram. The results show that the majority of swells in Taghazout Beach come from the North West to North direction, thus illustrating that swell direction ranks lower in importance in terms of an identifier due to the low variance within the variable.

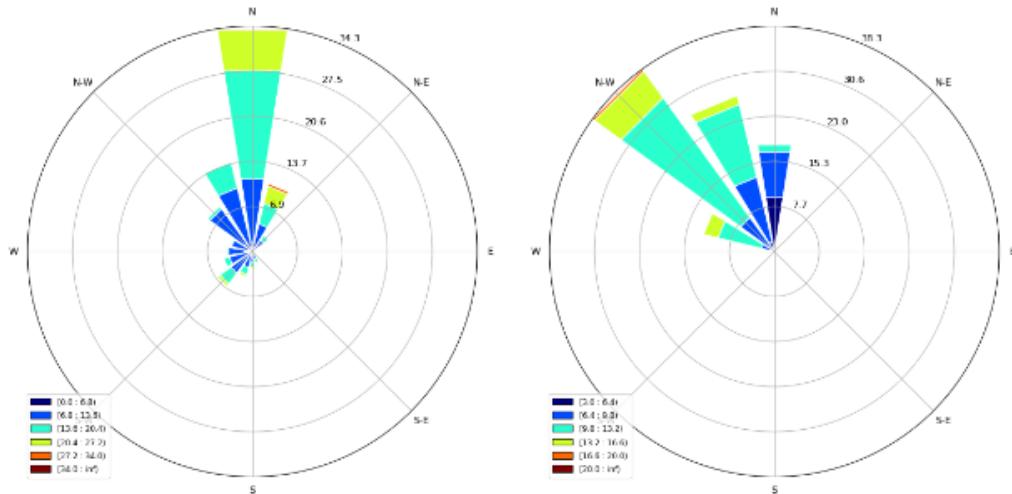


Fig 4.
Windrose
diagram for
wind speed

and wind direction (Left). Windrose visualisation for swell period and swell direction (Right) .

The variables were then plotted against time to determine whether or not they exhibit seasonality. It is apparent that swell height and swell period exhibit seasonal behaviour by the peaks and troughs in the data which the scatter plots convey.



Fig 5. Scatter plots for swell height, swell period and wind speed over time.

Furthermore, the variables are then plotted against each other in a 3D plot and colour categorised by month in order to check if any primary relationships exist between the three main predictors. The variables seem to exhibit a slight correlation with each other, however it is difficult to determine in the 3D plot due to the density.

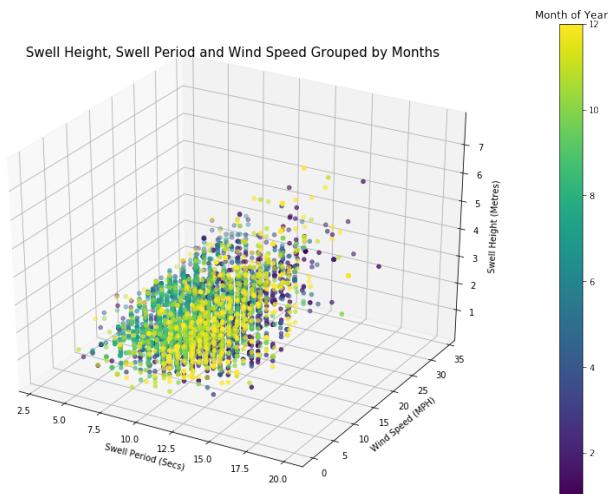


Fig 6. 3D scatter plot for swell height, swell period and wind speed.

3D plots are also utilised to view the relationship between each variable and its development across each month and each year. The points have been colour categorised by month of the year for visual clarification. Through these observations, it is once again apparent that similar monthly patterns exist across all years for swell period and swell height as evident by the curved structure of the 3D plot.

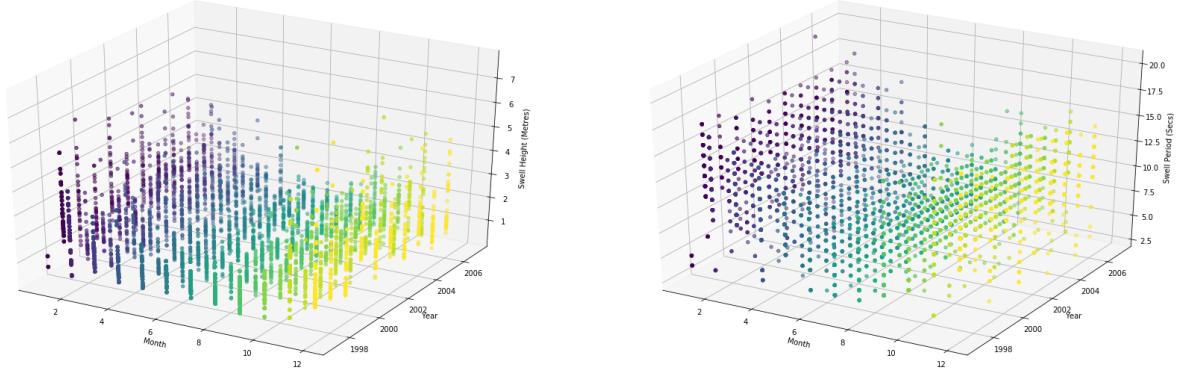


Fig 7. 3D scatter plot for swell height development over time (Left). 3D scatter plot for swell period development over time (Right).

Finally, correlations between the variables are then explored by creating a scatter plot and plotting the three main variables against each other. The variable not on the axis is represented as a colour categorisation in the points to provide a clearer representation, compared to the 3D plot in Fig X, of the relationships between the three main variables. The correlation plot between swell height and swell period shows a strong positive correlation between the two variables whereas wind speed and swell period exhibit a moderate negative correlation. Surprisingly, wind speed and swell height exhibit a very low correlation, suggesting that other factors that were not recorded could be more important in influencing the swell height. For the correlation table, please refer to Appendix 1, 2 and 3.

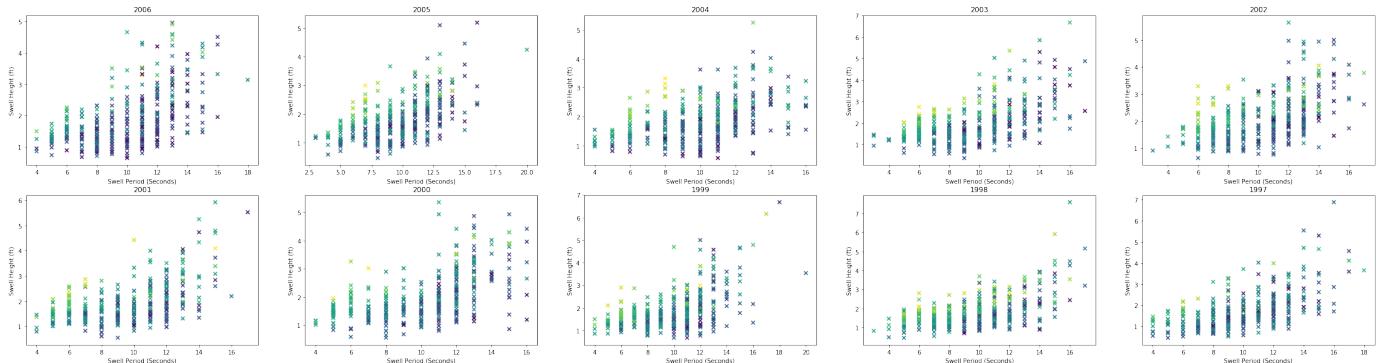


Fig 8. Yearly development of correlation between swell height and swell period

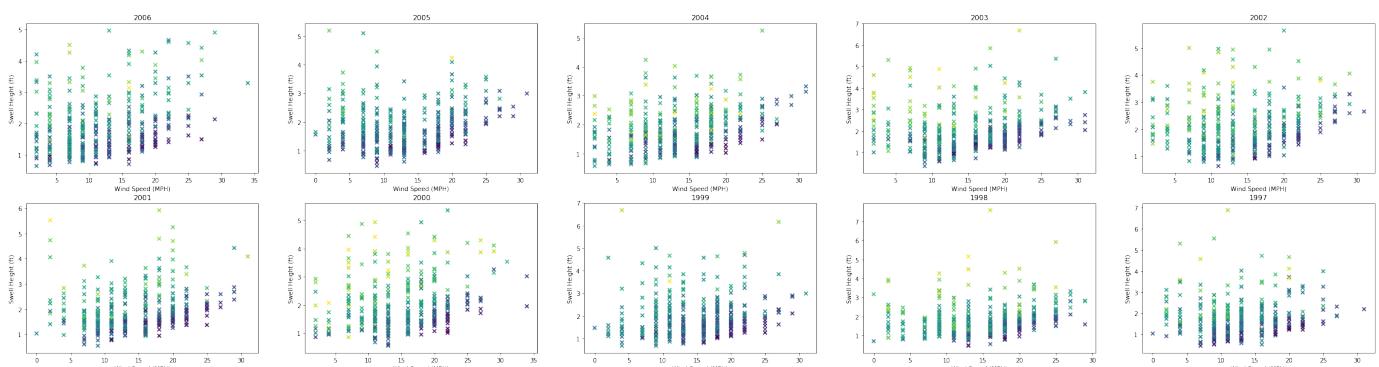


Fig 9. Yearly development of correlation between swell height and wind speed

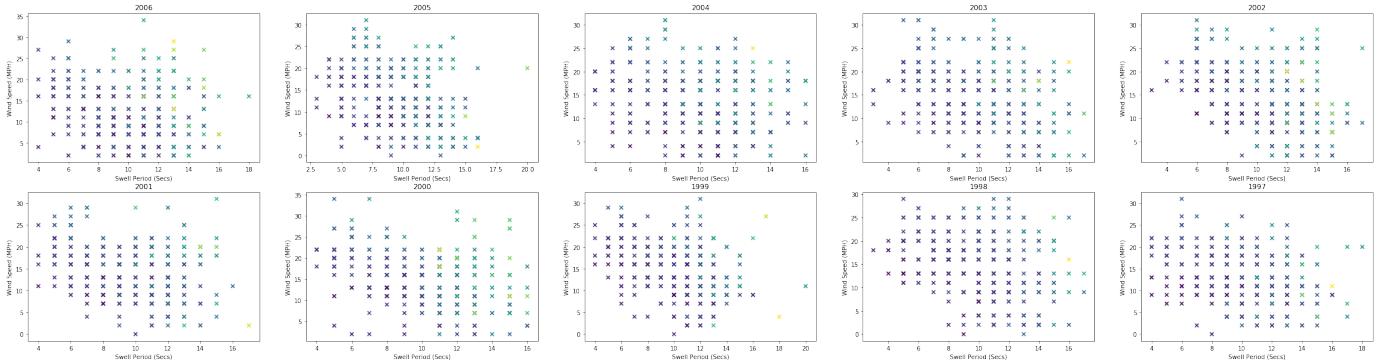


Fig 10. Yearly development of correlation between wind speed and swell period.

Methodology for Trend Visualisation

To determine the periods of low season across the years, three different functions were fit to the swell height data in order to make comparisons between low seasons illustrated by the different types of fit. Since swell height and swell period have a strong correlation, the project has decided to focus on the visualisation of swell height as it is considered a more important factor in determining surf conditions over swell period.

First, a step function with 13 different cut points along the x axis is fit to swell height data points annually as a representation of monthly aggregate. Each step represents a period of 30 days, which is approximately the duration of a month. The plots clearly demonstrate that a trend is present where low season occurs during the beginning of June to the end of September. Conversely, high season is shown to be occurring on approximately beginning November to end of February each year.

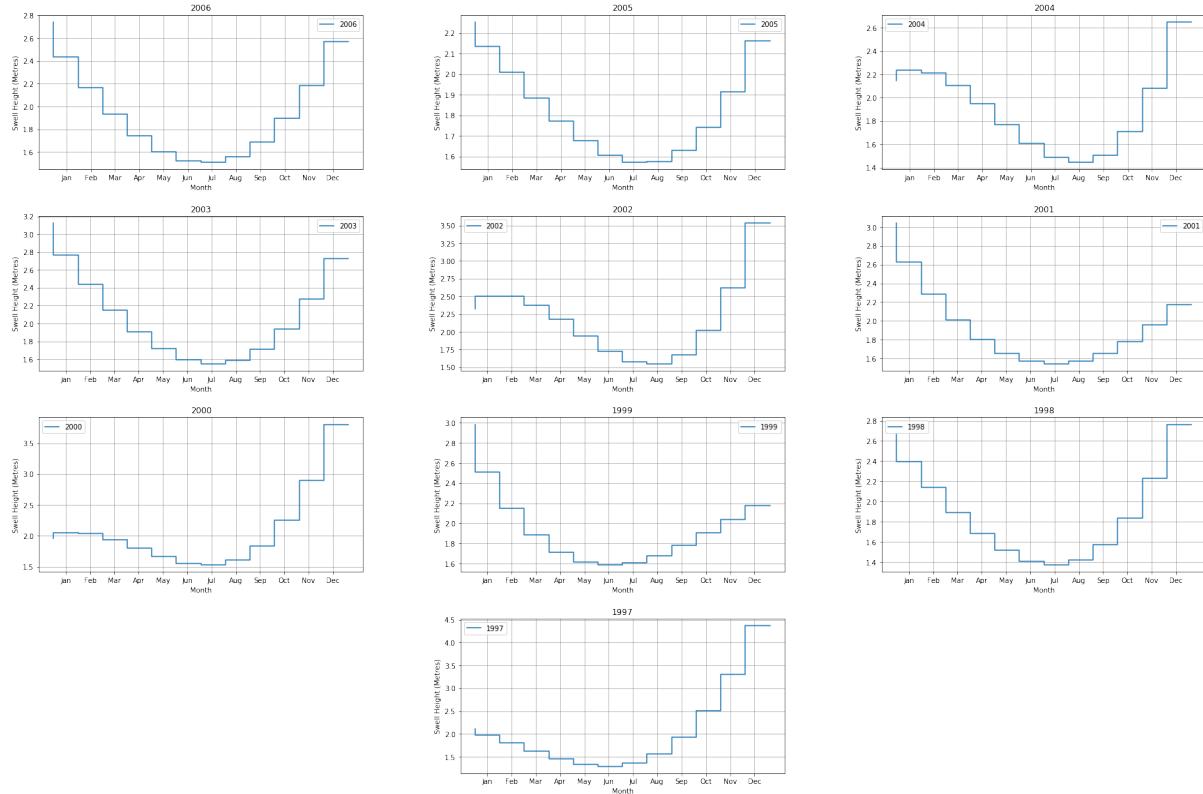


Fig 11. Step function for monthly swell height development by year.

A smoothing spline is then fit to the data to display the development of the shape of the annual trends. Similar to the previous step function, the function fit by the smoothing splines across the years convey a low season period beginning from the early June to mid-September. Furthermore, high season is shown to be approximately around the beginning of November to the end of March. The function exhibits cubic behaviour as the majority of the functions have one or two extrema and one or no inflection points.

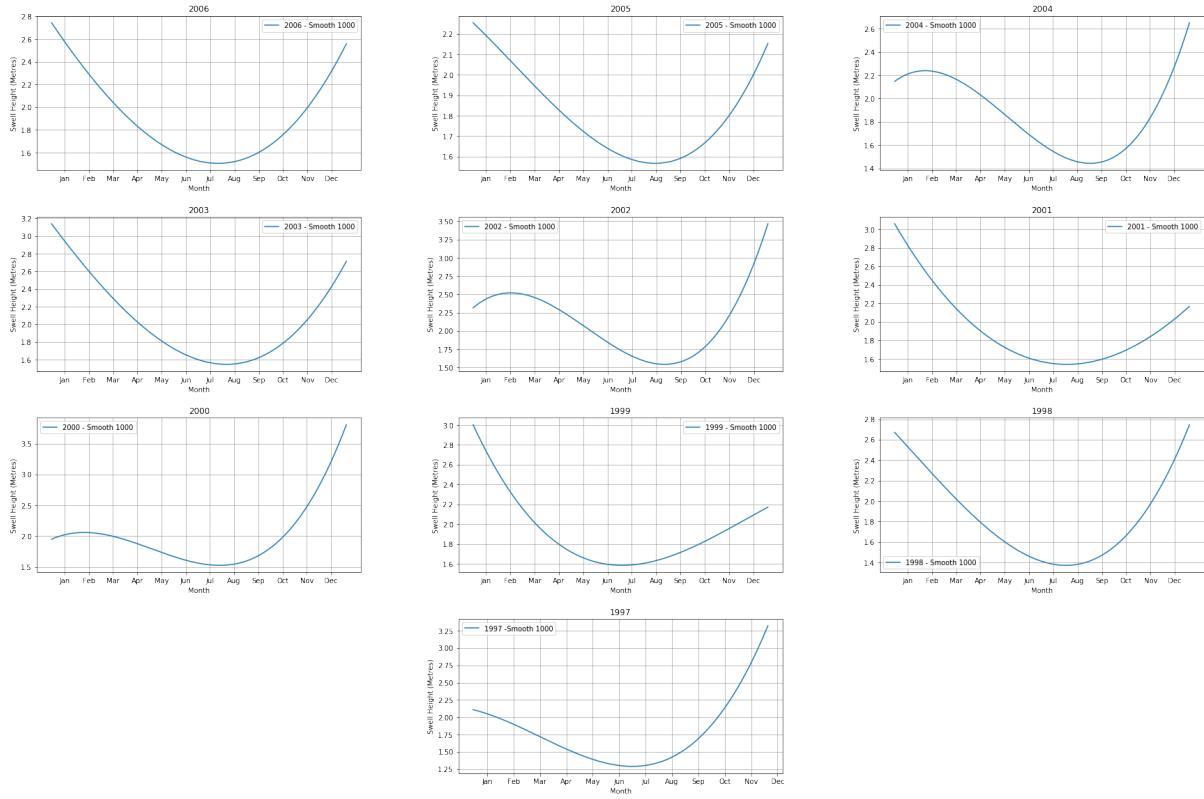


Fig 12. Smoothing spline for monthly swell height development by year.

The last plot is a cubic fit overlay of data from all years. A cubic function is fit for each year as the behaviour of the spline function displayed cubic properties and an overlay of each function was chosen in order to depict a better comparison between the years. Low season in this plot is, once again, shown to be the period from the middle of June to the middle of September, whereas high season can be observed around the middle of November to the end of March.

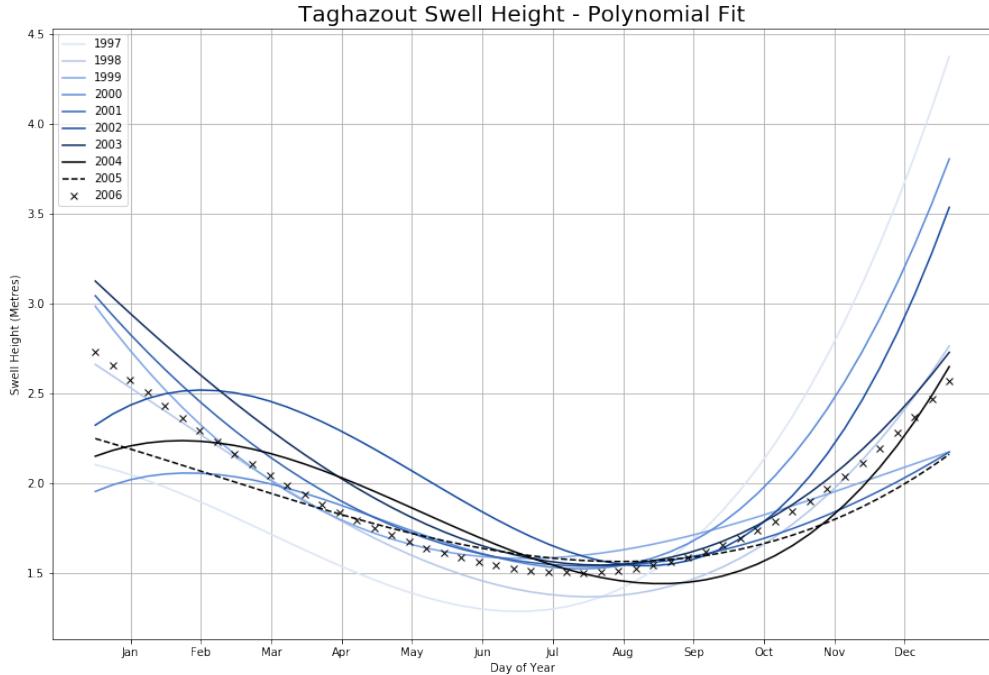


Fig 13. Cubic polynomial for monthly swell height development by year.

The final visualisation is a heat map of swell heights around the world. The GRIB files obtained were downloaded from the NOAA's NWW3 server via FTP on the FileZilla FTP Client and opened in Panoply to create a georeferenced Longitude-Latitude visualisation. The visualisations created in Panoply are then individually exported as a .png file and combined together to create a .gif file using Python's imageio library. The purpose of

this visualisation is to provide a comparison of Taghazout's swell conditions compared to other locations around the world. This method was required for the visualisation of global trends due to limitations of pygrib not being available on Windows Machine without an installation of a virtual machine; therefore, the GRIB files obtained were unable to be parsed for further analysis in Python.

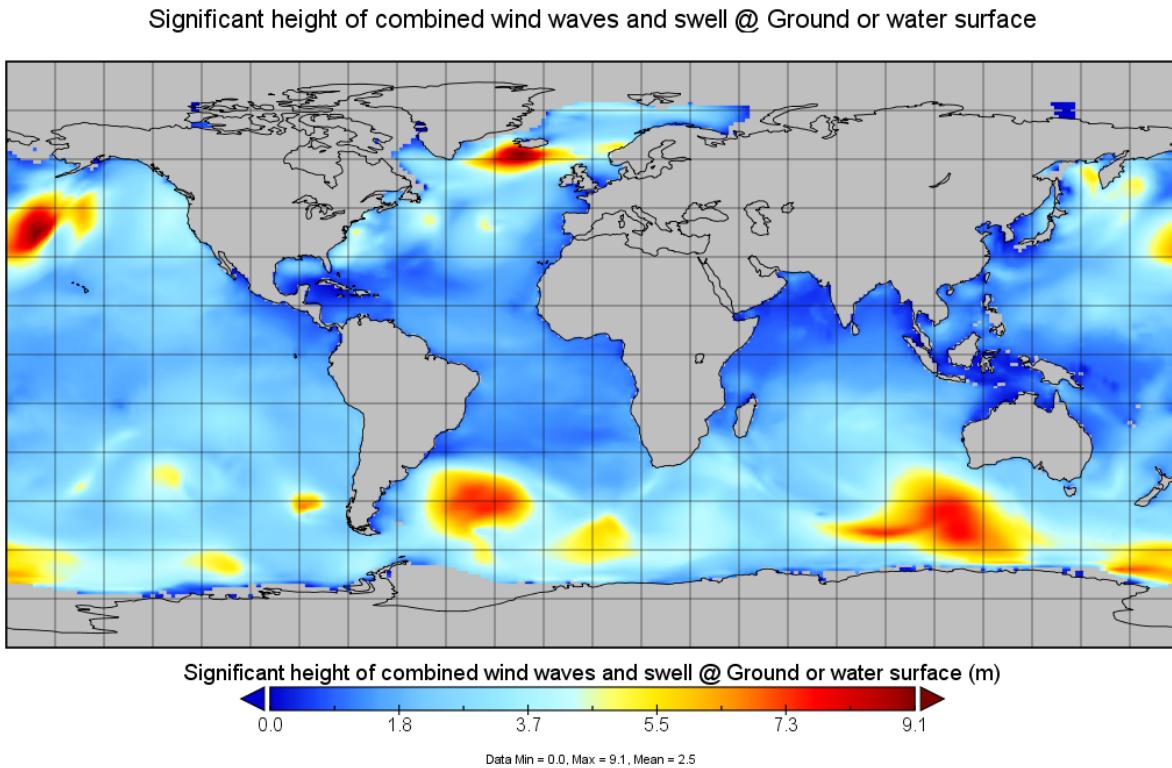


Fig 14. One frame from the gif displaying monthly swell height development world wide.

Results and discussion

Annual trends that depict periods of low and high season for optimal surf conditions in Taghazout have been found through the use of HTML parsing web crawlers and data visualisation. The three different types of functions fit to the dataset yield similar results which convey that the periods of low season in Taghazout occur approximately around beginning of July to end of September. This decision was formulated by observing the points on the curves closest to the minima. An assumption that can be generated from this result is that tourism influx will be at its lowest during the year as surf conditions are generally unfavourable and unstable, thus hostel owners can be advised to decrease the price of accommodation during this time to entice and incentivise tourists to visit, or perform maintenance during these periods to ensure the lowest opportunity costs incurred.

Periods of high season are determined by observing where the swell height is above two metres and occur approximately around mid-November to end of March annually. Here we can assume that the conditions of the surf will attract guests of a more professional nature with regards to surfing abilities. The difficulty that can be expected during this period will be that professional surfers tend to spend less money overall on other services that hostels typically offer, such as residential trips. Since there will be an increased influx in tourism due to optimal surf conditions during this period, it is advised that the best course of operational action during this time is to increase accommodation prices to maximise potential revenue.

Other periods that have not been mentioned are to be considered as transition periods between high and low season, tourists that can be expected here are a mixture between recreational and professional surfers. As the nature of the recreational tourists allows for hostels to monetise on a wider range of services for the clients, it is recommended that hostels organise various trips during this period to create an alternative activity for its clients in order to combat the uncertainty of the surf conditions.

Links to repos/files

Colorado Wind Farm Location Selection Web Application (Part 1 of Report): <http://jub.maps.arcgis.com/apps/View/index.html?appid=daa32c4c1e044c9499746512ea6dc4a0>.

GitHub Repository for the Taghazout Swell Height Analysis Project (Part 2 of Report): https://github.com/mimixtvxq/taghazout_crawler_visualisation_project

Appendix

Appendix 1. Correlation table between swell height and swell period over time.

Year	Correlation
2006	0.53
2005	0.55
2004	0.45
2003	0.54
2002	0.51
2001	0.53
2000	0.54
1999	0.52
1998	0.53
1997	0.56

Appendix 2. Correlation table between swell height and wind speed

Year	Correlation
2006	0.3
2005	0.16
2004	0.3
2003	0.13
2002	0.16
2001	0.21
2000	0.20
1999	0.11
1998	0.18
1997	0.08

Appendix 3. Correlation table between swell period and wind speed

Year	Correlation
2006	-0.21
2005	-0.33
2004	-0.24
2003	-0.39
2002	-0.38

2001	-0.32
2000	-0.37
1999	-0.38
1998	-0.34
1997	-0.30

References

- [1] *Booking.com*, Accessed January 24th 2019, [The Denver Post, The Denver Post, 16 June 2017, Accessed January 20th 2019, \[www.denverpost.com/2017/06/13/colorado-tornado-data/\]\(http://www.denverpost.com/2017/06/13/colorado-tornado-data/\).](https://www.booking.com/searchresults.en-gb.html?aid=375018&label=taghazout-nzhchbQTFQgJ_1lAVksFjwS94039967191%3Ap1%3Ata%3Ap1%3Ap2%3Aac%3Aap1t2%3Aneg%3Af1%3Atiaud-285284111686%3Akwd-21241368998%3Alp9068361%3Ali%3Adec%3Adm&sid=7668ebda71fc304b353f025b7c37e85d&tmpl=searchresults&class_interval=1&dest_id=81435&dest_type=city&from_sf=1&group_adults=2&group_children=0&label_click=undef&nflt=sth%253D4%3Dundefined%3B&no_rooms=1&raw_dest_type=city&room1=A%2CA&sb_price_type=total&shw_aparth=1&slp_r_match=0&src_elem=sb&ss=Taghazout&ssb=empty&ssne=Taghazout&ssne_unouched=Taghazout&rows=15.</p>
<p>[2] Hamm, Kevin.)
- [3] "Hostels in Taghazout - Choose from 20 Taghazout Hostels with Hostelworld." *Hostelworld Blog*, Accessed January 24th 2019, www.hostelworld.com/search?search_keywords=Taghazout%2C%2BMorocco&country=Morocco&city=Taghazout&date_from=2019-01-27&date_to=2019-01-30&number_of_guests=2.
- [4] "Magicseaweed Data Source." *Magicseaweed.com*, Accessed January 24th 2019, magicseaweed.com/Magicseaweed-Data-Sources-Article/321/.